

# **The Molecular Genetics of Familial Cardiomyopathy**

Stephen Nzeki Kamuli

Thesis Presented for the Degree of

**DOCTOR OF PHILOSOPHY**

In the Department of Medicine

UNIVERSITY OF CAPE TOWN

Supervisor: A/Prof Gasnat Shaboodien

Co-supervisor: Professor Ntobeko Ntusi

August 2022

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.

# TABLE OF CONTENTS

|                                               |            |
|-----------------------------------------------|------------|
| <b>TABLE OF CONTENTS .....</b>                | <b>I</b>   |
| <b>DECLARATION .....</b>                      | <b>I</b>   |
| <b>ABSTRACT .....</b>                         | <b>II</b>  |
| <b>ACKNOWLEDGMENTS.....</b>                   | <b>III</b> |
| <b>LIST OF FIGURES .....</b>                  | <b>IV</b>  |
| <b>LIST OF TABLES.....</b>                    | <b>V</b>   |
| <b>LIST OF ABBREVIATIONS AND SYMBOLS.....</b> | <b>VII</b> |
| <b>LIST OF AMINO ACID CODES .....</b>         | <b>XI</b>  |
| <b>CHAPTER 1 - INTRODUCTION .....</b>         | <b>1</b>   |
| <b>1.1 Cardiomyopathy .....</b>               | <b>3</b>   |
| 1.1.1 Clinical presentation .....             | 5          |
| 1.1.2 Classification .....                    | 6          |
| 1.1.3 Imaging modalities.....                 | 8          |
| 1.1.4 Therapies .....                         | 10         |
| 1.1.5 Genetics .....                          | 11         |
| 1.1.5.1 Single genes.....                     | 11         |
| 1.1.5.2 Multiple genes.....                   | 12         |
| 1.1.5.3 Modifiers .....                       | 12         |
| 1.1.5.4 Polygenetic Risk Scores .....         | 14         |
| 1.1.5.5 Genetics in sub-Saharan Africa.....   | 15         |
| <b>1.2 Dilated cardiomyopathy .....</b>       | <b>19</b>  |
| 1.2.1 Clinical presentation .....             | 19         |
| 1.2.2 Aetiology .....                         | 20         |

|                                                          |           |
|----------------------------------------------------------|-----------|
| 1.2.3 Disease burden .....                               | 20        |
| 1.2.4 Molecular genetics of DCM .....                    | 20        |
| <b>1.3 Left ventricular noncompaction .....</b>          | <b>22</b> |
| 1.3.1 Clinical presentation .....                        | 22        |
| 1.3.2 Aetiology .....                                    | 23        |
| 1.3.3 Disease burden .....                               | 23        |
| 1.3.4 Molecular genetics.....                            | 23        |
| <b>1.4 Arrhythmogenic cardiomyopathy .....</b>           | <b>25</b> |
| 1.4.1 Clinical presentation .....                        | 25        |
| 1.4.2 Aetiology .....                                    | 26        |
| 1.4.3 Disease burden .....                               | 26        |
| 1.4.4 Molecular genetics.....                            | 26        |
| <b>1.5 Next-generation sequencing.....</b>               | <b>28</b> |
| 1.5.1 Short-read sequencing:.....                        | 28        |
| 1.5.1.1 Sequencing-by-synthesis (SBS) .....              | 29        |
| 1.5.1.2 Sequencing-by-ligation (SBL) .....               | 29        |
| 1.5.2 Long-read sequencing:.....                         | 30        |
| 1.5.2.1 Single-molecule real-time sequencing (SMRT)..... | 31        |
| 1.5.2.2 Synthetic long-reads .....                       | 31        |
| 1.5.3 NGS applications .....                             | 33        |
| <b>1.6 Functional models for cardiomyopathy .....</b>    | <b>34</b> |
| <b>1.7 Rationale for the study.....</b>                  | <b>35</b> |
| <b>1.8 Hypothesis.....</b>                               | <b>36</b> |
| <b>1.9 Aims and objectives .....</b>                     | <b>36</b> |
| <b>CHAPTER 2 - METHODS.....</b>                          | <b>37</b> |
| <b>2.1 Study design .....</b>                            | <b>37</b> |

|                                                                                                                                         |           |
|-----------------------------------------------------------------------------------------------------------------------------------------|-----------|
| <b>2.2 Study participants and informed consent.....</b>                                                                                 | <b>38</b> |
| 2.2.1 Collection of demographic and clinical information .....                                                                          | 39        |
| 2.2.2 Diagnostic evaluation.....                                                                                                        | 40        |
| 2.2.2.1 DCM .....                                                                                                                       | 40        |
| 2.2.2.2 HCM .....                                                                                                                       | 41        |
| 2.2.2.3 ACM .....                                                                                                                       | 42        |
| 2.2.2.4 Exclusion criteria .....                                                                                                        | 44        |
| <b>2.3 DNA preparation .....</b>                                                                                                        | <b>44</b> |
| 2.3.1 DNA extraction.....                                                                                                               | 44        |
| 2.3.2 DNA quality control .....                                                                                                         | 44        |
| 2.3.2.1 Nanodrop quantification .....                                                                                                   | 44        |
| 2.3.2.2 Agarose Gels .....                                                                                                              | 44        |
| <b>2.4 Whole exome sequencing.....</b>                                                                                                  | <b>45</b> |
| 2.4.1 Principle of next generation sequencing (NGS) .....                                                                               | 45        |
| 2.4.2 Library preparation for whole exome sequencing.....                                                                               | 45        |
| 2.4.3 Library quality control.....                                                                                                      | 46        |
| 2.4.4 Data quality control (QC) .....                                                                                                   | 46        |
| 2.4.5 Whole exome sequencing data analysis .....                                                                                        | 47        |
| 2.4.5.1 American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) classification ..... | 47        |
| 2.4.5.2 Variant effect predictor (VEP) .....                                                                                            | 50        |
| 2.4.5.3 Variant filtering and assessment of pathogenicity .....                                                                         | 51        |
| 2.4.6 Bioinformatics tools.....                                                                                                         | 52        |
| 2.4.6.1 Variant prediction.....                                                                                                         | 52        |
| 2.4.6.2 Tissue expression.....                                                                                                          | 53        |
| 2.4.6.3 Protein folding.....                                                                                                            | 54        |
| 2.4.6.4 Conservation .....                                                                                                              | 55        |
| 2.4.6.5 Variant prioritisation .....                                                                                                    | 55        |
| 2.4.7 Whole exome sequencing variant validation.....                                                                                    | 55        |

|                                                      |           |
|------------------------------------------------------|-----------|
| 2.4.7.1 Gene annotation and primer design .....      | 56        |
| 2.4.7.2 Primers and parameters for DCM334 .....      | 56        |
| 2.4.7.3 Primers and parameters for DCM343 .....      | 57        |
| 2.4.7.4 Primers and parameters for ACM145.....       | 57        |
| 2.4.8 High resolution melt analysis.....             | 57        |
| 2.4.8.1 HRM protocol .....                           | 59        |
| 2.4.8.2 Purification of HRM products .....           | 59        |
| 2.4.9 Sanger sequencing .....                        | 60        |
| 2.4.9.1 Sequencing parameters.....                   | 60        |
| 2.4.9.2 Capillary electrophoresis .....              | 61        |
| 2.4.9.3 Analysis of Sanger sequencing products ..... | 61        |
| <b>2.5 Targeted resequencing .....</b>               | <b>62</b> |
| 2.5.1 Principle of targeted enrichment.....          | 62        |
| 2.5.2 HaloPlex Target Enrichment Workflow.....       | 63        |
| 2.5.3 Designing Haloplex probes .....                | 63        |
| 2.5.4 Sample Preparation .....                       | 63        |
| 2.5.5 Integrative Genomics Viewer .....              | 64        |

|                                                                                                                       |           |
|-----------------------------------------------------------------------------------------------------------------------|-----------|
| <b>CHAPTER 3 – WHOLE EXOME SEQUENCING IDENTIFIES A RARE PATHOGENIC DMPK MUTATION IN A FAMILY WITH SEVERE DCM.....</b> | <b>67</b> |
| <b>3.1 Introduction.....</b>                                                                                          | <b>67</b> |
| <b>3.2 Aim.....</b>                                                                                                   | <b>68</b> |
| <b>3.3 Objectives.....</b>                                                                                            | <b>68</b> |
| <b>3.4 Methods .....</b>                                                                                              | <b>68</b> |
| 3.4.1 Whole exome sequencing.....                                                                                     | 68        |
| 3.4.2 Cohort screen of <i>DMPK</i> through HRM and Sanger sequencing.....                                             | 68        |
| <b>3.5 Results.....</b>                                                                                               | <b>73</b> |
| 3.5.1 Clinical history of the DCM334 family.....                                                                      | 73        |

|                                                                                                                                     |               |
|-------------------------------------------------------------------------------------------------------------------------------------|---------------|
| 3.5.2 Whole exome sequencing.....                                                                                                   | 74            |
| 3.5.3 Variant validation.....                                                                                                       | 77            |
| 3.5.4 <i>DMPK</i> c.1067C>T(p.P356L) protein structural prediction.....                                                             | 78            |
| 3.5.5 <i>DMPK</i> IGV visualization.....                                                                                            | 79            |
| 3.5.6 <i>DMPK</i> Familial segregation .....                                                                                        | 79            |
| 3.5.7 <i>DMPK</i> cardiomyopathy cohort screen through HRM and Sanger sequencing .....                                              | 80            |
| 3.5.8 Modifiers.....                                                                                                                | 81            |
| <b>3.6 Discussion .....</b>                                                                                                         | <b>83</b>     |
| 3.6.1 <i>DMPK</i> variants .....                                                                                                    | 83            |
| 3.6.2 Modifiers.....                                                                                                                | 85            |
| 3.6.2.1 <i>OBSCN</i> c.2997-2998TG/CA (p.A1000T).....                                                                               | 86            |
| 3.6.2.2 <i>IQGAP2</i> c.932A>G (p.N311S) .....                                                                                      | 87            |
| <b>3.7 Conclusion .....</b>                                                                                                         | <b>88</b>     |
| <br><b>CHAPTER 4 – WHOLE EXOME SEQUENCING IDENTIFIES A <i>BAG3</i> MUTATION AS THE CAUSE OF DCM IN A SOUTH AFRICAN FAMILY .....</b> | <br><b>90</b> |
| <b>4.1 Introduction.....</b>                                                                                                        | <b>90</b>     |
| <b>4.2 Aims .....</b>                                                                                                               | <b>91</b>     |
| <b>4.3 Objectives.....</b>                                                                                                          | <b>91</b>     |
| <b>4.4 Methods .....</b>                                                                                                            | <b>91</b>     |
| 4.4.1 Whole exome sequencing.....                                                                                                   | 91            |
| 4.4.2 Targeted resequencing .....                                                                                                   | 91            |
| <b>4.5 Results.....</b>                                                                                                             | <b>92</b>     |
| 4.5.1 Clinical history of the DCM 343 family.....                                                                                   | 92            |
| 4.5.2 Whole exome sequencing.....                                                                                                   | 94            |
| 4.5.3 Variant validation.....                                                                                                       | 97            |
| 4.5.4 <i>BAG3</i> c.925C>T (p.R309Ter) protein structural prediction .....                                                          | 98            |
| 4.5.5 <i>BAG3</i> IGV visualization.....                                                                                            | 98            |

|                                                                                                                        |                |
|------------------------------------------------------------------------------------------------------------------------|----------------|
| 4.5.6 BAG3 c.925C>T familial segregation .....                                                                         | 99             |
| 4.5.7 BAG3 targeted resequencing in an extended cardiomyopathy cohort .....                                            | 100            |
| 4.5.8 Modifiers.....                                                                                                   | 101            |
| <b>4.6 Discussion .....</b>                                                                                            | <b>104</b>     |
| 4.6.1 BAG3 c.925C>T mutation .....                                                                                     | 104            |
| 4.6.2 BAG3 extended cardiomyopathy cohort screen .....                                                                 | 109            |
| 4.6.3 Modifiers.....                                                                                                   | 110            |
| 4.6.3.1 RYR1 c.10616G>A(p.R3539H) .....                                                                                | 110            |
| 4.6.3.2 CRK c.48_49delinsCG(p.R17G).....                                                                               | 111            |
| 4.6.3.3 NEB c.16582C>G p.D5528H .....                                                                                  | 112            |
| 4.6.3.4 NEB c.10531C>A p.A3511S .....                                                                                  | 113            |
| <b>4.7 Conclusion .....</b>                                                                                            | <b>113</b>     |
| <br><b>CHAPTER 5 – INVESTIGATION OF PKP2 AS THE DISEASE-CAUSING VARIANT IN THIS SOUTH AFRICAN<br/>ACM FAMILY .....</b> | <br><b>114</b> |
| <b>5.1 Introduction.....</b>                                                                                           | <b>114</b>     |
| <b>5.2 Aim.....</b>                                                                                                    | <b>115</b>     |
| <b>5.3 Objectives.....</b>                                                                                             | <b>116</b>     |
| <b>5.4 Methods .....</b>                                                                                               | <b>116</b>     |
| 5.4.1 Whole exome sequencing.....                                                                                      | 116            |
| 5.4.2 Targeted resequencing .....                                                                                      | 116            |
| <b>5.5 Results.....</b>                                                                                                | <b>117</b>     |
| 5.5.1 Clinical history of the ACM145 family.....                                                                       | 117            |
| 5.5.2 Whole exome sequencing of combined affected individuals in ACM145 .....                                          | 118            |
| 5.5.2.1 Known cardiac panel screen.....                                                                                | 119            |
| 5.5.2.2 Novel candidate gene screen .....                                                                              | 121            |
| 5.5.2.3 Variant validation .....                                                                                       | 124            |
| 5.5.3 Whole exome sequencing of individual affected members of ACM145 .....                                            | 132            |

|                                                                                                      |            |
|------------------------------------------------------------------------------------------------------|------------|
| 5.5.3.1 Proband .....                                                                                | 133        |
| 5.5.3.2 Proband's affected brother .....                                                             | 136        |
| 5.5.4 Targeted resequencing .....                                                                    | 140        |
| 5.5.4.1 PKP2 c.2197_2202delInsG .....                                                                | 140        |
| 5.5.4.2 ALPK3 c.2701C>T(p.Q901Ter).....                                                              | 141        |
| 5.5.5 Integrative Genomics Viewer .....                                                              | 141        |
| 5.5.5.1 PKP2 c.2197_2202delInsG .....                                                                | 141        |
| 5.5.5.2 ALPK3 c.2701C>T(p.Q901Ter).....                                                              | 143        |
| 5.5.6 Variant summary for ACM145 .....                                                               | 144        |
| <b>5.6 Discussion .....</b>                                                                          | <b>145</b> |
| 5.6.1 Known cardiac gene panel.....                                                                  | 145        |
| 5.6.1.1 PKP2 c.2197_2202del(p.H733Afs*8) .....                                                       | 146        |
| 5.6.1.2 ALPK3 c.2701C>T(p.Q901Ter).....                                                              | 149        |
| 5.6.2 Possible novel candidate genes .....                                                           | 152        |
| 5.6.2.1 CFTR c.1727G>C(p.G576A) .....                                                                | 153        |
| 5.6.2.2 CFTR c.2002C>T(p.R668C) .....                                                                | 153        |
| <b>5.7 Conclusion .....</b>                                                                          | <b>154</b> |
| <b>CHAPTER 6 - CONCLUSION AND FUTURE DIRECTIONS.....</b>                                             | <b>156</b> |
| <b>6.1 Conclusion .....</b>                                                                          | <b>156</b> |
| <b>6.2 Future directions.....</b>                                                                    | <b>157</b> |
| <b>REFERENCES .....</b>                                                                              | <b>160</b> |
| <b>APPENDICES .....</b>                                                                              | <b>184</b> |
| <b>Appendix A: List of genes implicated in dilated cardiomyopathy<sup>346-348</sup> .....</b>        | <b>184</b> |
| <b>Appendix B: List of genes implicated in left ventricular non-compaction<sup>349</sup> .....</b>   | <b>184</b> |
| <b>Appendix C: List of genes implicated in arrhythmogenic cardiomyopathy<sup>350-353</sup> .....</b> | <b>184</b> |
| <b>Appendix D: Ethics Annual Renewal for study HREC REF 198/2020 .....</b>                           | <b>185</b> |

|                                                                     |     |
|---------------------------------------------------------------------|-----|
| Appendix E: Ethics Approval for study HREC REF 198/2020 .....       | 186 |
| Appendix F: Informed DNA consent for genetic analysis.....          | 187 |
| Appendix G: Storage and governance SOP for IMHOTEP .....            | 188 |
| Appendix H: Buffers and solutions.....                              | 190 |
| Appendix I: New England Biolabs <sup>R</sup> 100bp DNA ladder ..... | 192 |
| APPENDIX J: CMO Genes.....                                          | 193 |
| Appendix K: 38 gene panel used in Haloplex targeted analysis .....  | 194 |
| Appendix L: Puregene DNA extraction protocol .....                  | 195 |

## DECLARATION

I, ...*Stephen Nzeki Kamuli*....., hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature:

Signed by candidate

Date: 1 August 2022

# ABSTRACT

## Introduction

The cardiomyopathies are responsible for approximately 5.9 of 100,000 deaths in the general global population and in sub-Saharan Africa (SSA), these myocardial diseases are observed in 21.4% of patients with heart failure. The precise etiology of the cardiomyopathies is currently not well known and through our research we aim to contribute to the genetic landscape and bridge the gaps in knowledge for the different cardiomyopathies as SSA could provide some very important insights into the cardiomyopathies and identify other possible disease mechanisms.

## Methods

Through next generation sequencing techniques such as whole exome sequencing and targeted resequencing we studied three South African families with severe cardiomyopathy. Clinical diagnosis and recruitment of cardiomyopathy patients into the study was done at Groote Schuur Hospital, Cape Town by a panel of experts. Next generation sequencing data was analysed and filtered through various stringent criteria and the final list of variants were validated through Sanger sequencing.

## Results

In the first multi-generational family with severe dilated cardiomyopathy (DCM) (DCM 334), we identified a pathogenic *DMPK* c.1067C>T(p.P356L) variant in the proband and her affected father. We also screened a cohort of 542 cardiomyopathy probands through Sanger sequencing of the *DMPK* gene and identified the *DMPK* c.1477C>T(p.R493C) variant as a variant of unknown significance. We then investigated a three-generation family with four affected family members who were also affected with severe DCM (DCM343). We used whole exome sequencing and identified the pathogenic *BAG3* c.925C>T (p.R309Ter) variant as the cause of disease within this family. Viral infection, anti-hypertensive medication and genetic modifiers in *RYR1* and *NEB* contributed to the variable phenotype among the individuals with the *BAG3* variant. Through targeted resequencing we also identified the same pathogenic *BAG3* variant in 2 of the 634 cardiomyopathy probands screened. In the third family, we investigated a South African family affected with severe arrhythmogenic cardiomyopathy (ACM). We used whole exome sequencing and targeted resequencing in combination and identified the pathogenic *PKP2* c.2197\_2202InsGdelCACACC (p.H733Afs\*8) as the cause of disease in the proband and his father. We also present evidence of the *ALPK3* c.2701C>T(p.Q901Ter) variant modifying the phenotypic manifestation which correlates with the variable penetrance that is seen among ACM families.

## Conclusion

Through this project, we have identified many firsts. To the best of our knowledge, we are the first to show that *DMPK* is associated with primary DCM in severely affected young patients. As a first for South Africa, we not only identified the pathogenic *BAG3* variant in a family with severe DCM, but we also identified the same variant in two additional probands, raising the possibility of a founder effect. In the third and final family with ACM, we identified the pathogenic *PKP2* variant as the cause of disease within this family with the novel *ALPK3* variant acting as a possible modifier. Our research has added to what is currently known about the cardiomyopathies in Africa but there is still much work to be done as we believe we have just scratched the tip of the iceberg.

## ACKNOWLEDGMENTS

- To my supervisor, A/Prof. Gasnat Shaboodien, for her advice, assistance and guidance throughout this project
- To my co-supervisor, Professor Ntobeko Ntusi, for his support and insight, without whom this research would not have been possible
- To the late Professor Bongani Mayosi, for his unequivocal generosity and mentorship that ensured I grew professionally and personally while doing my project
- To members of the Cardiovascular Genetics Team at the Cape Heart Mrs. Lameez Pearce, Polycarp Ndibangwi and Kim Brooks for all their valuable time during the research and help in proofreading my thesis
- To Dr. Sarah Kraus for her expertise in the clinical aspects of the research
- To Professor Karen Sliwa, director of the Cape Heart Institute for her availability and generosity to ensure quality resources are available at the institute and the other members of the Cape Heart Institute for making it a conducive place to work
- To the IMHOTEP project and University of Cape Town for funding my studies
- To Professor Bernard Keavney, Dr. Simon Williams and other members at the University of Manchester, UK for the indispensable training they gave me to be able to carry out my research
- To my mother and brother for always being there for me and their prayers that carried me throughout the course of this project
- To my wife, Mrs. Janice Kamuli for her constant encouragement and understanding at all times from the advent of my research
- Above all I would like to thank The Almighty GOD for the gift of sound health and strength to completion of this dissertation

# LIST OF FIGURES

## CHAPTER 1

|                                                                                            |    |
|--------------------------------------------------------------------------------------------|----|
| Figure 1.1: Global distribution of the causes of death.....                                | 1  |
| Figure 1.2: Distribution of CVDs globally.....                                             | 2  |
| Figure 1.3: Structure of the heart showing the myocardium.....                             | 3  |
| Figure 1.4: Images of a normal heart in comparison to the cardiomyopathy subtypes.....     | 6  |
| Figure 1.5: A timeline of the evolution of the classification of the cardiomyopathies..... | 7  |
| Figure 1.6: ESC classification of the cardiomyopathies.....                                | 7  |
| Figure 1.7: The MOGES classification for cardiomyopathies.....                             | 8  |
| Figure 1.8: Cardiovascular imaging modalities used for the assessment of CVD.....          | 9  |
| Figure 1.9: Comparison between a normal heart and a DCM heart.....                         | 16 |
| Figure 1.10: Genetic variants that yield a DCM phenotype.....                              | 18 |
| Figure 1.11: Comparison between a normal heart and an LVNC heart.....                      | 19 |
| Figure 1.12: The genetic landscape of genes underlying the LVNC phenotype.....             | 21 |
| Figure 1.13: Comparison between a normal heart and an ACM heart.....                       | 22 |
| Figure 1.14: Desmosomal and non-desmosomal genes in ACM.....                               | 24 |
| Figure 1.15: Short read sequencing approaches.....                                         | 27 |
| Figure 1.16: Long read sequencing approaches.....                                          | 28 |
| Figure 1.17: Sequencing output of the different NGS approaches.....                        | 30 |

## CHAPTER 2

|                                                                   |    |
|-------------------------------------------------------------------|----|
| Figure 2.1: Flow diagram showing work done.....                   | 35 |
| Figure 2.2: Criteria organization for variant classification..... | 44 |
| Figure 2.3: Stepwise use of VEP to analyze NGS data.....          | 48 |
| Figure 2.4: Melt curve profiles of different PCR amplicons.....   | 55 |
| Figure 2.5: Haloplex sample preparation overflow.....             | 60 |
| Figure 2.6: IGV coverage track.....                               | 62 |
| Figure 2.7: IGV alignment track.....                              | 62 |
| Figure 2.8: IGV structural variations.....                        | 63 |

## CHAPTER 3

|                                                                                               |    |
|-----------------------------------------------------------------------------------------------|----|
| Figure 3.1: Pedigree of family DCM334.....                                                    | 71 |
| Figure 3.2: Filtering analysis of WES data from family DCM334.....                            | 72 |
| Figure 3.3: Validation of the candidate variants identified in the proband of DCM334.....     | 74 |
| Figure 3.4: Conservation of the DMPK p.P356L residue.....                                     | 74 |
| Figure 3.5: Alignment coverage of DMPK using Chimera.....                                     | 75 |
| Figure 3.6: Segregation pattern of the <i>DMPK</i> , <i>TTN</i> and <i>SYNE</i> variants..... | 76 |
| Figure 3.7: Conservation of the DMPK p.R493C residue.....                                     | 77 |
| Figure 3.8: HRM and Sanger sequencing results of <i>DMPK</i> c.1477C>T variant.....           | 77 |
| Figure 3.9: Structure of the <i>DMPK</i> gene.....                                            | 80 |
| Figure 3.10: <i>OBSCN</i> mutations identified to cause cardiomyopathies.....                 | 83 |

## CHAPTER 4

|                                                                                       |    |
|---------------------------------------------------------------------------------------|----|
| Figure 4.1: Pedigree of DCM family 343.....                                           | 88 |
| Figure 4.2: Electron and light microscopy images of the heart of individual VI:4..... | 89 |
| Figure 4.3: VEP data analysis of variants for family DCM343.....                      | 90 |
| Figure 4.4: Validation of the <i>BAG3</i> c.925C>T variant.....                       | 93 |
| Figure 4.5: Conservation of the <i>BAG3</i> p.R309 residue.....                       | 93 |

|                                                                                        |     |
|----------------------------------------------------------------------------------------|-----|
| Figure 4.6: Segregation of <i>BAG3</i> c.925C>T variant.....                           | 95  |
| Figure 4.7: Segregation of variants identified in <i>RYR1</i> .....                    | 97  |
| Figure 4.8: Segregation of variants identified in <i>CRK</i> and <i>NEB</i> genes..... | 99  |
| Figure 4.9: Structure of <i>BAG3</i> protein.....                                      | 100 |

## CHAPTER 5

|                                                                                    |     |
|------------------------------------------------------------------------------------|-----|
| Figure 5.1: Pedigree of family ACM145.....                                         | 113 |
| Figure 5.2: Filtering for WES data from the cardiac panel .....                    | 114 |
| Figure 5.3: Filtering for WES data from the non-cardiac gene panel.....            | 116 |
| Figure 5.4: Validation analysis of <i>CFTR</i> c.1727G>C.....                      | 119 |
| Figure 5.5: Amino acid conservation of the <i>CFTR</i> p.G576A residue.....        | 120 |
| Figure 5.6: Chimera visualization of the <i>CFTR</i> p.G576A residue.....          | 120 |
| Figure 5.7: Validation analysis of <i>CFTR</i> c.2002C>T.....                      | 121 |
| Figure 5.8: Amino acid conservation of the <i>CFTR</i> p.R668C residue.....        | 122 |
| Figure 5.9: Chimera visualization of <i>CFTR</i> .....                             | 122 |
| Figure 5.10: Validation analysis of <i>PKD1</i> c.6749C>T .....                    | 123 |
| Figure 5.11: Amino acid conservation of the <i>PKD1</i> p.T2250M residue.....      | 123 |
| Figure 5.12: Chimera visualization of the <i>PKD1</i> p.T2250M residue.....        | 124 |
| Figure 5.13: Residue coverage of <i>PKD1</i> by UNIPROT.....                       | 124 |
| Figure 5.14: Validation analysis of <i>LRP1</i> c.9927_9928delCG.....              | 125 |
| Figure 5.15: Amino acid conservation of the <i>LRP1</i> p.G3311R fs residue.....   | 125 |
| Figure 5.16: Chimera visualization of <i>LRP1</i> .....                            | 126 |
| Figure 5.17: Pedigree of variants identified in family ACM145.....                 | 127 |
| Figure 5.18: Filtering for WES data from the proband.....                          | 128 |
| Figure 5.19: Validation HRM analysis of <i>ALPK3</i> c.2701C>T variant.....        | 128 |
| Figure 5.20: Conservation of the <i>ALPK3</i> p.Q901* residue.....                 | 130 |
| Figure 5.21: Filtering for WES data from the proband's brother.....                | 130 |
| Figure 5.22: Validation of <i>PKP2</i> c.2197_2202del.....                         | 134 |
| Figure 5.23: IGV visualisation of <i>PKP2</i> c.2197_2022delInsG variant.....      | 136 |
| Figure 5.24: Conservation of the <i>PKP2</i> p.H733fs residue.....                 | 137 |
| Figure 5.25: IGV visualisation of <i>ALPK3</i> c.2701C>T variant.....              | 138 |
| Figure 5.26: Segregation of variants identified after individual WES analysis..... | 139 |
| Figure 5.27: Structure of <i>PKP2</i> .....                                        | 140 |
| Figure 5.28: <i>PKP2</i> c.2197-2202del variant in <i>PKP2</i> ARM 6 motif.....    | 142 |
| Figure 5.29: Location of the <i>ALPK3</i> c.2701C>T variant in exon 6.....         | 144 |

## LIST OF TABLES

### CHAPTER 1

|                                                                              |    |
|------------------------------------------------------------------------------|----|
| Table 1.1: Causes of the different cardiomyopathies.....                     | 5  |
| Table 1.2: Contribution of genes implicated in DCM.....                      | 18 |
| Table 1.3: Contribution of genes reported in LVNC.....                       | 21 |
| Table 1.4: Contribution of genes implicated in ACM.....                      | 24 |
| Table 1.5: Application parameters of the sequencing platforms available..... | 29 |
| Table 1.6: Cardiomyopathy genes identified through NGS.....                  | 31 |

### CHAPTER 2

|                                             |    |
|---------------------------------------------|----|
| Table 2.1: Diagnostic criteria for DCM..... | 37 |
| Table 2.2: Diagnostic criteria for HCM..... | 38 |

|                                                                                           |    |
|-------------------------------------------------------------------------------------------|----|
| Table 2.3: International Task Force criteria for ACM.....                                 | 39 |
| Table 2.4: Criteria for classifying Pathogenic variants.....                              | 46 |
| Table 2.5: Criteria for classifying Benign variants.....                                  | 47 |
| Table 2.6: Primer pairs used for HRM analysis of candidate variants in family DCM334..... | 53 |
| Table 2.7: Standard PCR cycling parameters for DCM343.....                                | 54 |
| Table 2.8: Standard PCR cycling parameters for ACM145.....                                | 54 |
| Table 2.9: Protocol for PCR and HRM of the exons.....                                     | 55 |
| Table 2.10: Cycling conditions for the PCR and HRM of variant exons.....                  | 56 |
| Table 2.11: Cleanup protocol for variant HRM products.....                                | 56 |
| Table 2.12: PCR protocol for the sequencing reactions.....                                | 57 |
| Table 2.13: Cycling conditions for the sequencing .....                                   | 57 |
| Table 2.14: Database numbers of genes investigated.....                                   | 58 |

### CHAPTER 3

|                                                                                              |    |
|----------------------------------------------------------------------------------------------|----|
| Table 3.1: Primer pairs used for HRM analysis of <i>DMPK</i> .....                           | 66 |
| Table 3.2: Standard PCR cycling parameters.....                                              | 66 |
| Table 3.3: Protocol for PCR and HRM of exon 1.1 and 6.1 (2mM MgCl <sub>2</sub> ).....        | 67 |
| Table 3.4: Temperature cycling conditions for exons 1.1 and 6.1.....                         | 67 |
| Table 3.5: Protocol for PCR and HRM of exon 3 (1mM MgCl <sub>2</sub> ) .....                 | 67 |
| Table 3.6: Temperature cycling conditions for exon 3.....                                    | 67 |
| Table 3.7: Protocol for PCR and HRM of exon 9 (2mM MgCl <sub>2</sub> ) .....                 | 68 |
| Table 3.8: Temperature cycling conditions for exon 9.....                                    | 68 |
| Table 3.9: Primer redesign, magnesium chloride and temperature optimisation.....             | 68 |
| Table 3.10: Protocol for PCR and HRM of exon 13 (2mM MgCl <sub>2</sub> ) .....               | 69 |
| Table 3.11: Temperature cycling conditions for exon 13.....                                  | 69 |
| Table 3.12: Protocol for PCR and HRM of exon 1.2 (1.5 mM MgCl <sub>2</sub> ).....            | 69 |
| Table 3.13: Temperature cycling conditions for exon 1.2.....                                 | 70 |
| Table 3.14: Protocol for PCR and HRM of exons 2 and 12 (1 mM MgCl <sub>2</sub> ) .....       | 70 |
| Table 3.15: Temperature cycling conditions for exons 2 and 12.....                           | 70 |
| Table 3.16: Variants identified after filtering of WES data for family DCM334.....           | 73 |
| Table 3.17: SDM prediction outcome of <i>DMPK</i> , <i>PLN</i> and <i>PKP2</i> variants..... | 75 |
| Table 3.18: Modifier variants in known cardiac genes.....                                    | 78 |
| Table 3.19: Modifier variants in novel genes.....                                            | 78 |

### CHAPTER 4

|                                                                                   |    |
|-----------------------------------------------------------------------------------|----|
| Table 4.1: Variants identified after filtering of WES data for family DCM343..... | 92 |
| Table 4.2: Modifier variants in family DCM 343.....                               | 98 |

### CHAPTER 5

|                                                                                           |     |
|-------------------------------------------------------------------------------------------|-----|
| Table 5.1: List of rare variants identified in ACM 145 family from the cardiac panel..... | 115 |
| Table 5.2: Filtering for WES data from the novel candidate genes.....                     | 118 |
| Table 5.3: SDM prediction outcome of <i>CFTR</i> c.1727G>C variant.....                   | 121 |
| Table 5.4: Residues coverage of PKD1 with UNIPROT.....                                    | 126 |
| Table 5.5: Filtering for WES data in the proband.....                                     | 129 |
| Table 5.6: Filtering for WES data in the proband's brother.....                           | 133 |

## LIST OF ABBREVIATIONS AND SYMBOLS

|                  |                                                               |
|------------------|---------------------------------------------------------------|
| 1D               | One-dimensional                                               |
| 2D               | Two-dimensional                                               |
| 3D               | Three-dimensional                                             |
| 1000G            | 1000Genomes                                                   |
| ACA              | American College of Cardiology                                |
| ACE              | Angiotensin I Converting Enzyme                               |
| ACE              | Angiotensin-converting enzyme                                 |
| ACEis            | ACE inhibitors.                                               |
| ACMG             | American College of Medical Genetics and Genomics             |
| ACTC1            | Cardiac actin                                                 |
| ADRA1A           | Adrenoceptor Alpha 1A                                         |
| AHA              | American Heart Association                                    |
| ACM              | Arrhythmogenic cardiomyopathy                                 |
| ALFA             | Allele frequency aggregator                                   |
| AMP              | Association for Molecular Pathology                           |
| API              | Application program interface                                 |
| APOBEC3A         | Apolipoprotein B mRNA Editing Enzyme Catalytic Subunit 3A     |
| APOBEC3B         | Apolipoprotein B mRNA Editing Enzyme Catalytic Subunit 3B     |
| ARBs             | Angiotensin receptor blockers                                 |
| ARVC             | Arrhythmogenic right ventricular cardiomyopathy               |
| ATP              | Adenosine triphosphate                                        |
| AVF              | Arteriovenous                                                 |
| BAG3             | Bcl-2-associated athanogene 3                                 |
| BGI              | Beijing Genomics Institute                                    |
| $\beta$ -MHC     | Beta-myosin heavy chain gene                                  |
| BPs              | Base Pairs                                                    |
| BP               | Base Pair                                                     |
| BP               | Benign supporting                                             |
| BS               | Benign strong                                                 |
| BSA              | Bovine serum albumin                                          |
| Ca <sup>2+</sup> | Calcium ion                                                   |
| CACNA2D4         | Calcium Voltage-Gated Channel Auxiliary Subunit Alpha2delta 4 |
| CAP              | College of American Pathologists                              |
| CASA             | Chaperone-assisted selective autophagy                        |
| CASQ2            | Calsequestrin 2                                               |
| CDH2             | Cadherin 2                                                    |
| CCS              | Circular consensus sequence                                   |
| cDNA             | Complementary DNA                                             |
| CFTR             | CF Transmembrane Conductance Regulator                        |
| CHD              | Congenital heart disease                                      |
| CM               | Cardiomyopathy                                                |
| cm               | Centimetre                                                    |
| CMOs             | Cardiomyopathies                                              |
| CMOS             | Complementary metal-oxide semiconductor                       |
| CMR              | Cardiac magnetic resonance imaging                            |
| CNVs             | Copy number variants                                          |
| CP               | Cardiac panel                                                 |
| CPGR             | Centre for Proteomic and Genomic Research                     |
| CREBBP           | CREB Binding Protein                                          |
| CRK              | Proto-Oncogene, Adaptor Protein                               |
| CSRP3            | Cysteine And Glycine Rich Protein 3                           |

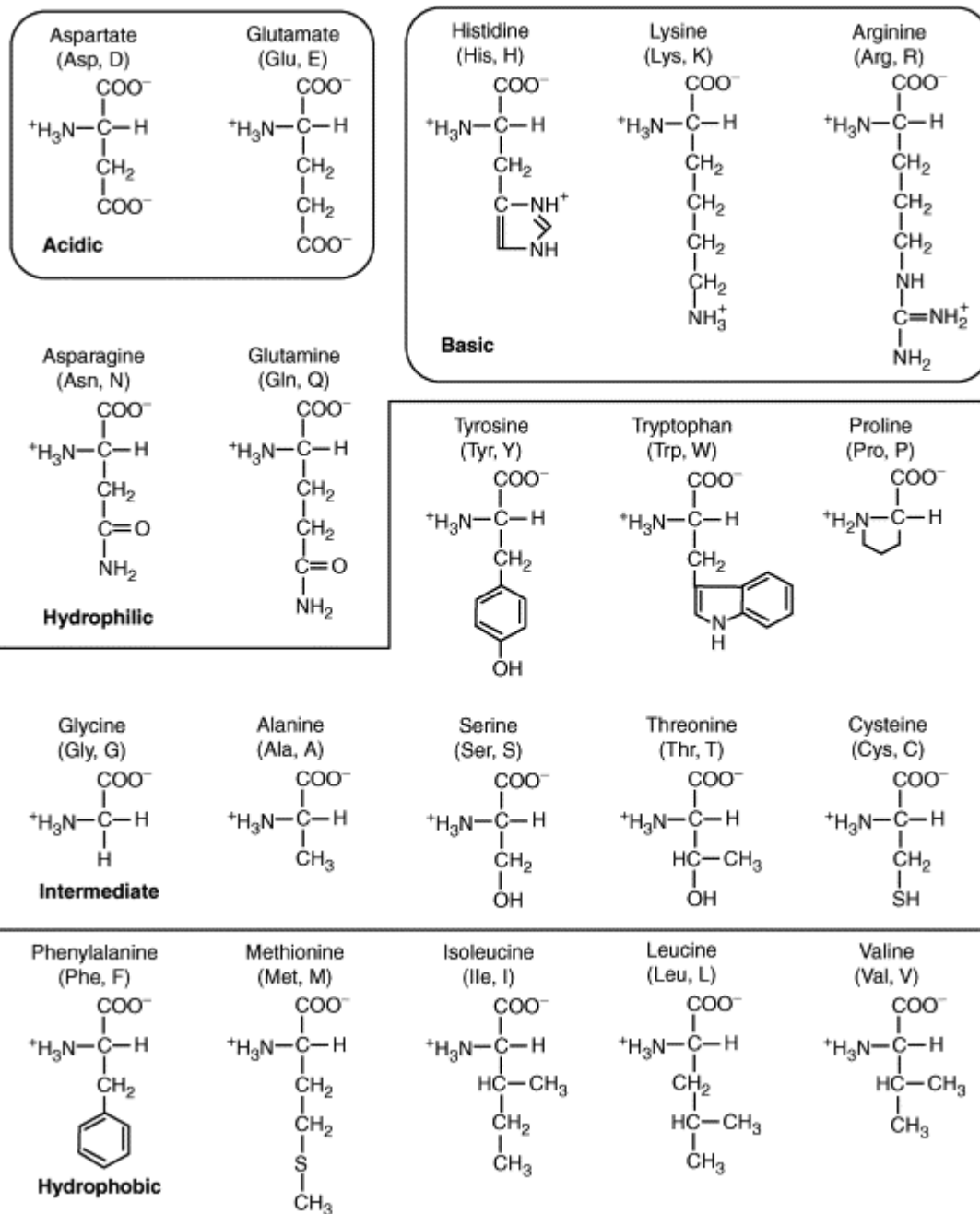
|                 |                                                               |
|-----------------|---------------------------------------------------------------|
| CVD             | Cardiovascular disease                                        |
| CPVT            | Catecholaminergic polymorphic ventricular tachycardia         |
| CAD             | Coronary artery disease                                       |
| <i>DES</i>      | Desmin                                                        |
| <i>DLC1</i>     | DLC1 Rho GTPase Activating Protein                            |
| DM              | Muscular dystrophy                                            |
| <i>DMPK</i>     | Dystrophin Myotonia Protein Kinase                            |
| DNA             | Deoxyribonucleic acid                                         |
| dNTP            | Deoxynucleotide triphosphate                                  |
| DdNTP           | Dideoxyribonucleotides                                        |
| <i>DSG2</i>     | Desmoglein-2                                                  |
| <i>DSC2</i>     | Desmocollin-2                                                 |
| <i>DSP</i>      | Desmoplakin                                                   |
| DCM             | Dilated cardiomyopathy                                        |
| <i>DMD</i>      | Dystrophin                                                    |
| EDTA            | Ethylenediaminetetraacetic acid                               |
| ECG             | Electrocardiogram                                             |
| EMF             | Endomyocardial fibrosis                                       |
| ESC             | European Society of Cardiology                                |
| ExAC            | Exome Aggregation Consortium                                  |
| F-ACTIN         | Filamentous actin                                             |
| FH              | Family history                                                |
| <i>FHL1</i>     | Four And A Half LIM Domains 1                                 |
| <i>FLNC</i>     | Filamin C                                                     |
| <i>FXD1</i>     | FXD Domain Containing Ion Transport Regulator 1               |
| gDNA            | genomic DNA                                                   |
| <i>GFAP</i>     | Glial Fibrillary Acidic Protein                               |
| <i>GLA</i>      | Galactosidase Alpha                                           |
| GNOMAD          | Genome Aggregation Database                                   |
| GSH             | Groote Schuur Hospital                                        |
| <i>GSTM1</i>    | Glutathione S-Transferase Mu 1                                |
| GTex            | Genotype-Tissue Expression                                    |
| GVUS            | Genetic variant of unknown significance                       |
| HF              | Heart failure                                                 |
| HCM             | Hypertrophic cardiomyopathy                                   |
| HG19            | Human genome assembly 19                                      |
| <i>HIP1R</i>    | Huntingtin Interacting Protein 1 Related                      |
| HIV             | Human immunodeficiency virus                                  |
| <i>HLA-DRB1</i> | Major Histocompatibility Complex, Class II, DR Beta 1         |
| <i>HLA-DRB5</i> | Major Histocompatibility Complex, Class II, DR Beta 5         |
| HREC            | Human Research Ethics Committee                               |
| HRM             | High resolution melting                                       |
| HSC70           | HEAT SHOCK COGNATE 71 KDA PROTEIN                             |
| HSPB6           | Heat Shock Protein Family B (Small) Member 6                  |
| HSPB8           | Heat Shock Protein Family B (Small) Member 6                  |
| ICDs            | Implantable cardioverter defibrillators                       |
| IMHOTEP         | The African Cardiomyopathy and Myocarditis Registry Programme |
| IPSCs           | Induced pluripotent stem cells                                |
| IQGAP           | IQ-motif containing Ras GTPase-activating-like protein        |
| <i>IQGAP2</i>   | IQ Motif Containing GTPase Activating Protein 2               |
| ISFET           | Ion-sensitive field-effect transistor                         |
| ITF             | International Task Force                                      |
| <i>JUP</i>      | Plakoglobin                                                   |
| Kb              | Kilobyte                                                      |

|                  |                                                                |
|------------------|----------------------------------------------------------------|
| <i>KCNJ12</i>    | Potassium Inwardly Rectifying Channel Subfamily J Member 12    |
| KDa              | Kilodalton                                                     |
| <i>KRT14</i>     | Keratin 14                                                     |
| <i>LAMA4</i>     | Laminin Subunit Alpha 4                                        |
| LBBB             | Left bundle branch block                                       |
| <i>LMNA</i>      | Lamin A/C                                                      |
| LOF              | Loss-of-function                                               |
| LQTS             | Long QT syndrome                                               |
| <i>LRP1</i>      | LDL Receptor Related Protein 1                                 |
| LV               | Left ventricular                                               |
| LVEDD            | Left Ventricular End-Diastolic Diameter                        |
| LVEDV            | Left ventricular end diastolic volume                          |
| LVEF             | Left ventricular ejection fraction                             |
| LVESV            | Left ventricular end systolic volume                           |
| LVNC             | Left ventricular noncompaction                                 |
| MAF              | Minor allele frequency                                         |
| MG               | Miligram                                                       |
| Mg <sup>2+</sup> | Magnesium ion                                                  |
| MGI              | Mouse Genome Informatics                                       |
| MH               | Malignant hyperthermia                                         |
| MRI              | Magnetic resonance imaging                                     |
| MSEC             | Millisecond                                                    |
| <i>MSH6</i>      | MutS Homolog 6                                                 |
| MT               | Mutationtaster                                                 |
| <i>MYH7</i>      | Beta-myosin heavy chain                                        |
| <i>MYH6</i>      | Myosin-6 ( $\alpha$ -myosin heavy chain)                       |
| <i>MYBPC3</i>    | Cardiac myosin-binding protein C                               |
| <i>MYL2</i>      | Regulatory light chain of myosin                               |
| <i>MYL3</i>      | Myosin light chain 3                                           |
| <i>MYPN</i>      | Myopalladin                                                    |
| <i>MYBPH</i>     | Myosin Binding Protein H                                       |
| NCBI             | National Center for Biotechnology Information                  |
| <i>NEB</i>       | Nebulin                                                        |
| NGS              | Next generation sequencing                                     |
| NM               | Nemaline myopathy                                              |
| Nm               | Nanometre                                                      |
| Ng               | Nanogram                                                       |
| <i>NOS2</i>      | Nitric Oxide Synthase 2                                        |
| NYHA             | New York Heart Association                                     |
| <i>OBSCN</i>     | Obscurin, Cytoskeletal Calmodulin And Titin-Interacting RhoGEF |
| OR               | Odds ratio                                                     |
| PCR              | Polymerase chain reaction                                      |
| qPCR             | quantitative PCR                                               |
| Path             | Pathogenic                                                     |
| PDB              | Protein data bank                                              |
| <i>PDE4DIP</i>   | Phosphodiesterase 4D Interacting Protein                       |
| PERL             | practical extraction and report language                       |
| <i>PKD1</i>      | Polycystin 1, Transient Receptor Potential Channel Interacting |
| <i>PKP2</i>      | Plakophilin-2                                                  |
| <i>PLEC</i>      | Plectin                                                        |
| <i>PLM</i>       | Phospholemman                                                  |
| <i>PLN</i>       | Phospholamban                                                  |
| PM               | Pathogenic moderate                                            |
| PND              | Paroxysmal nocturnal dyspnea                                   |

|                |                                                                           |
|----------------|---------------------------------------------------------------------------|
| PP-2           | Polyphen-2                                                                |
| <i>PRKAG2</i>  | Protein Kinase AMP-Activated Non-Catalytic Subunit Gamma 2                |
| PS             | Pathogenic supporting                                                     |
| PVS            | Pathogenic very strong                                                    |
| <i>PSEN1</i>   | Presenilin-1                                                              |
| RBBB           | Left bundle branch block                                                  |
| <i>RBM20</i>   | RNA-binding protein 20                                                    |
| RCM            | Restrictive cardiomyopathy                                                |
| RR             | Relative risk                                                             |
| RV             | Right ventricular                                                         |
| RVEF           | Right ventricular ejection fraction                                       |
| RVEDV          | Right ventricular end-diastolic volume                                    |
| <i>RYR1</i>    | Ryanodine Receptor 1                                                      |
| <i>RYR2</i>    | Ryanodine receptor 2                                                      |
| SBL            | Sequencing by ligation                                                    |
| SBS            | Sequencing by synthesis                                                   |
| SDM            | Site directed mutator                                                     |
| <i>SERCA2a</i> | ATPase Sarcoplasmic/Endoplasmic Reticulum Ca <sup>2+</sup> Transporting 2 |
| SIFT           | Sorting Intolerant From Tolerant                                          |
| SNPs           | Single-nucleotide polymorphisms                                           |
| SCD            | Sudden cardiac death                                                      |
| <i>SCN5A</i>   | Sodium channel protein type 5 subunit $\alpha$                            |
| SPECT          | Single-photon emission computerized tomography                            |
| SSA            | Sub-Saharan Africa                                                        |
| SUNDS          | Sudden unexplained nocturnal death syndrome                               |
| <i>SYNE1</i>   | Spectrin Repeat Containing Nuclear Envelope Protein 1                     |
| TBE            | Tris/Borate/EDTA                                                          |
| <i>TGFB3</i>   | Transforming Growth Factor Beta 3                                         |
| <i>TNNI3</i>   | Troponin I3, Cardiac Type                                                 |
| <i>TNNT2</i>   | Cardiac muscle troponin T                                                 |
| <i>TPM1</i>    | Alpha-tropomyosin                                                         |
| TS             | Targeted sequencing                                                       |
| <i>TTN</i>     | Titin                                                                     |
| UCP            | Unknown cardiac panel                                                     |
| UCSF           | University of California, San Francisco                                   |
| UK             | United Kingdom                                                            |
| $\mu$ l        | Microlitre                                                                |
| $\mu$ M        | Micromolar                                                                |
| UCSC           | University of California, Santa Cruz                                      |
| UCT            | University of Cape Town                                                   |
| USA            | United States of America                                                  |
| VCF            | Variant call format                                                       |
| VEP            | Variant Effect Predictor                                                  |
| VT             | Ventricular tachycardia                                                   |
| VUS            | Variants of unknown significance                                          |
| <i>VWF</i>     | Von Willebrand Factor                                                     |
| WES            | Whole exome sequencing                                                    |
| WGS            | Whole genome sequencing                                                   |
| WHO            | World Health Organization                                                 |
| <i>ZNF469</i>  | Zinc Finger Protein 469                                                   |

## LIST OF AMINO ACID CODES

Amino acid nomenclature and structure obtained from (<https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/amino-acids>).



## CHAPTER 1 - INTRODUCTION

The major threat to human health and development is non-communicable diseases (NCDs).

Globally, the morbidity and mortality rates from NCDs, specifically cardiovascular diseases (CVD), respiratory diseases, cancer and diabetes, are higher than those from all other causes combined.

While the burden of infectious diseases attributed to HIV and TB is still a significant concern in sub-Saharan Africa (SSA), the disease pattern according to current observations indicates that the epidemiological landscape in SSA is gradually evolving into communicable and NCDs.<sup>1</sup> It is estimated that by 2030, NCDs are to surpass the combined death toll due to neonatal, communicable diseases, nutritional and maternal conditions(WHO).<sup>2</sup> This contrasts with wealthy Western countries, where death and disability are primarily due to unhealthy diet and behaviours (Figure 1.1).<sup>3</sup>

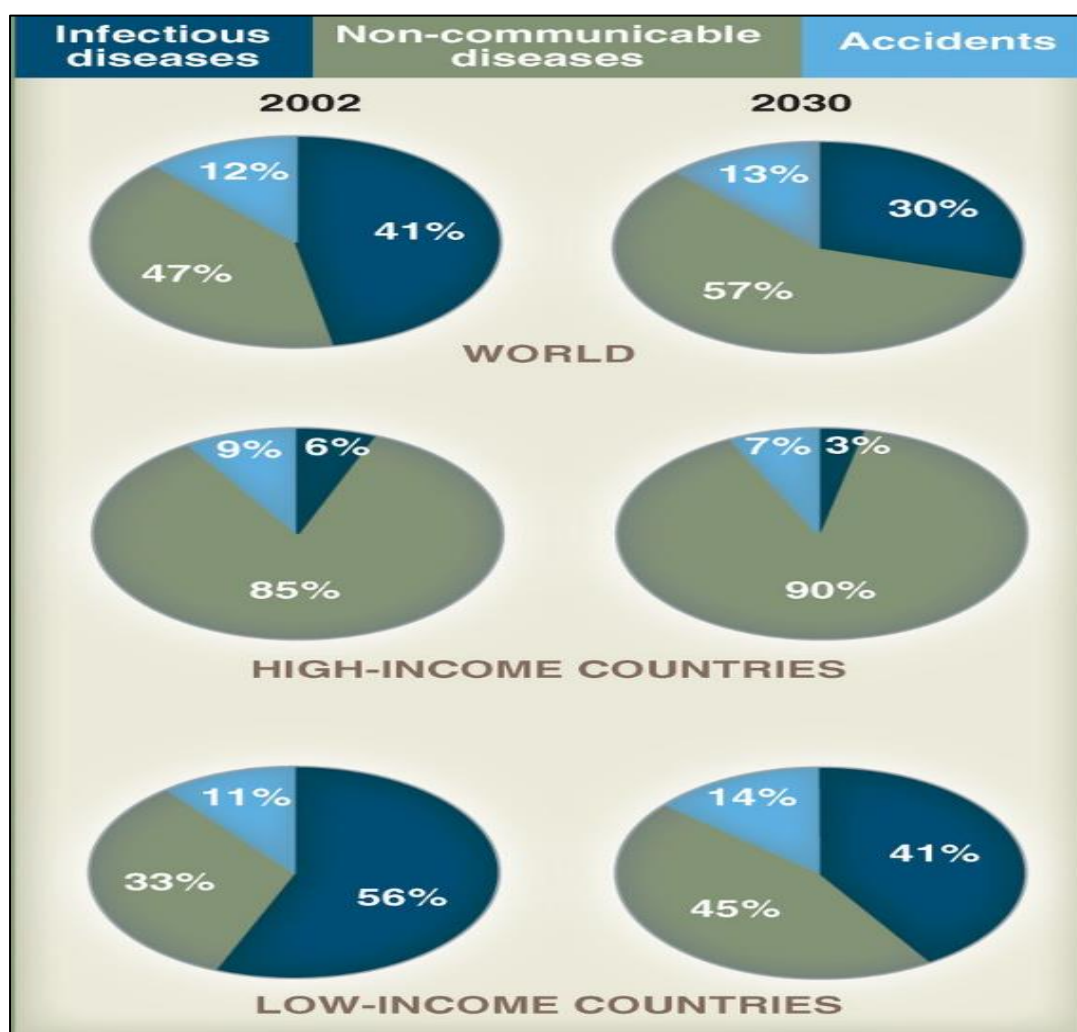


Figure 1.1: The global distribution of the causes of death from 2002 projecting that by 2030 non-communicable diseases will surpass the proportion of global deaths due to infectious deaths and accidents.<sup>3</sup>

According to investigations carried out in 2017 on the burden of disease worldwide, CVDs account for 73.4% of all deaths worldwide and within three decades, the total number of deaths due to CVDs in SSA has risen by over 50%. In SSA, the fatalities tend to occur at significantly younger ages, compared to western countries their western world counterparts which compounds the health sector and socioeconomic challenges the region faces.<sup>4-6</sup> Figure 1.2 shows the global distribution and stratification of CVD deaths due to the various cardiovascular diseases in males and females.

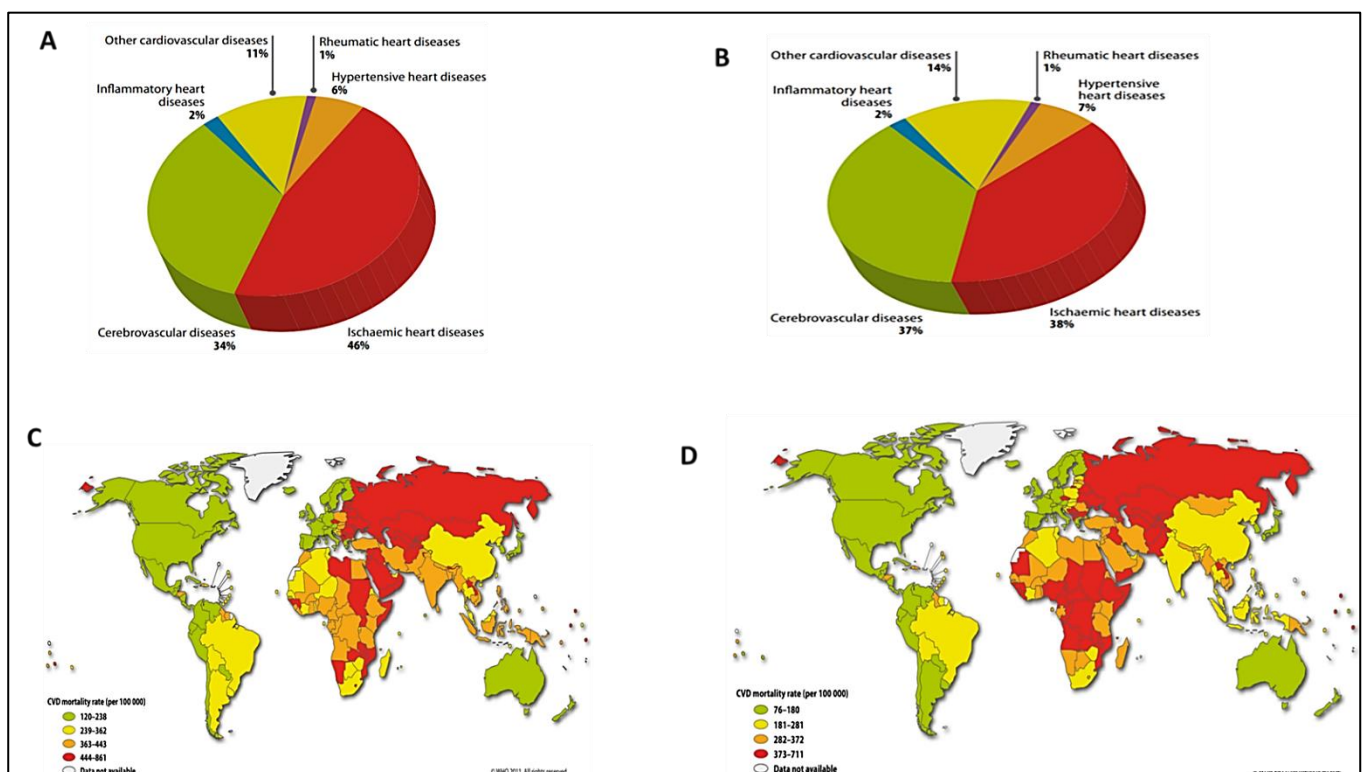


Figure 1.2: Distribution of the CVDs: A) Distribution of CVD deaths due to various cardiovascular diseases in males. B) Distribution of CVD deaths due to various cardiovascular diseases in females. C) Global stratification of CVD mortality rates in males D.) Global stratification of CVD mortality rates in females.<sup>7</sup>

Within CVD, heart failure (HF) is a common condition. The adult population in the western world has a HF prevalence of up to 2% which increases to over 10% among people 70 years and older.<sup>9-11</sup> Survival rates have improved although HF leading to re-hospitalization rates within 30 days remain at approximately 25% and death due to HF are still high at 50% within five years of diagnosis in these developed countries.<sup>11</sup> In SSA, these HF numbers are unknown due to the lack of investigations that highlight their mortality and prevalence, however, an estimated six-month mortality rate of 18% has been recorded.<sup>8</sup> It has been determined that IHD is responsible for HF among

the adult population in the develop countries while in SSA HF is mainly due to cardiomyopathy, rheumatic heart disease and hypertensive heart disease.<sup>6,8,10,12-14</sup> If we consider CVD in Africa, the cardiomyopathies are a subgroup that pose the greatest challenge because of their greater prevalence in areas still affected by diseases of famine and pestilence<sup>15</sup> together with lack of resources for proper cardiological investigations needed for accurate diagnosis. In addition, the high mortality that is associated with cardiomyopathies, together with limited availability of necessary remedied including heart transplantation, highlights the significant challenges posed by cardiomyopathy in SSA.<sup>15,16</sup>

## 1.1 Cardiomyopathy

Cardiomyopathies are abnormalities of the heart muscle which renders the heart muscle compromised in function and composition without the presence of anomalous conditions that can elucidate the myocardial anomaly seen.<sup>17,18</sup> The cardiomyopathies represent aberrations of the myocardium, which is the thick muscle layer of the heart wall. (Figure 1.3).<sup>19,20,21</sup>

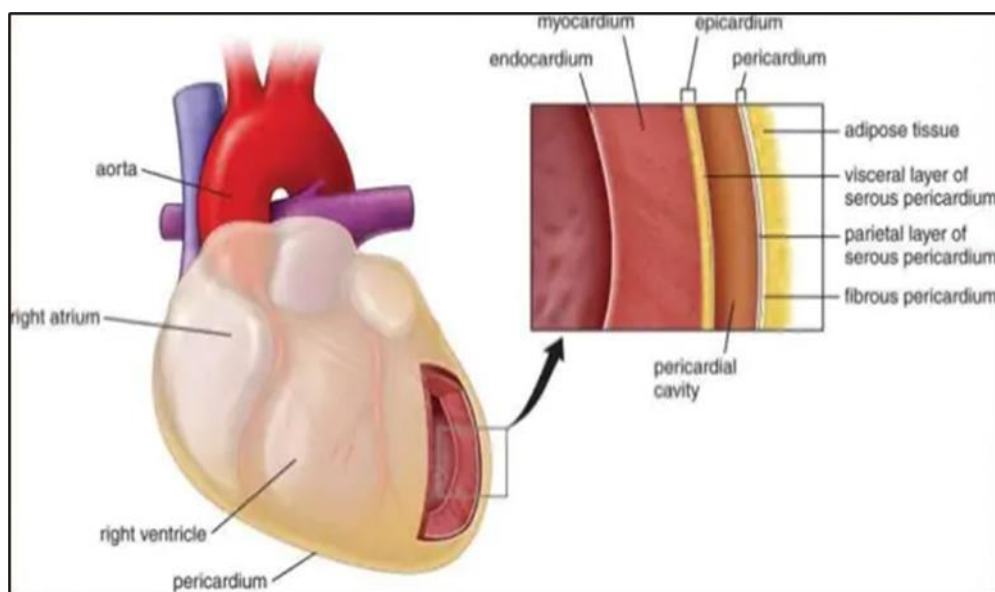


Figure 1.3: Structure of the heart showing the myocardium.<sup>19</sup>

Primary or genetic cardiomyopathies are classified morphofunctionally into dilated (DCM), hypertrophic (HCM), arrhythmogenic (ACM) and restrictive cardiomyopathy (RCM). While clearly a disorder of heart muscle, left ventricular noncompaction (LVNC) is not considered an independent primary cardiomyopathy. The acquired aetiologies of secondary cardiomyopathies include tachycardia, infections, endocrinopathies, peripartum cardiomyopathy (PPCM), immune-related disorders, and nutritional deficiencies while other important secondary/infiltrative aetiologies include amyloidosis, glycogen storage disorders, mitochondrial myopathies and Fabry disease (Table 1.1).<sup>18,19,22-24</sup>

**Table 1.1: Causes of the different cardiomyopathies<sup>19</sup>**

| Primary Cardiomyopathies                                                                                                                                                     |                                                                                                                   |                                         |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|-----------------------------------------|
| Genetic                                                                                                                                                                      | Mixed (Genetic/acquired)                                                                                          | Acquired                                |
| Hypertrophic cardiomyopathy                                                                                                                                                  | Dilated cardiomyopathy                                                                                            | Inflammatory (myocarditis)              |
|                                                                                                                                                                              |                                                                                                                   | Stress produced (Tako-tsubo)            |
| Arrhythmogenic cardiomyopathy                                                                                                                                                | Restrictive (non-hypertrophic and non-dilated)                                                                    | Peripartum                              |
|                                                                                                                                                                              |                                                                                                                   | Tachycardia induced                     |
|                                                                                                                                                                              |                                                                                                                   | Infants of mothers with Type 1 diabetes |
| Left ventricular non-compaction                                                                                                                                              |                                                                                                                   |                                         |
| Conduction defects                                                                                                                                                           |                                                                                                                   |                                         |
| Mitochondrial myopathies                                                                                                                                                     |                                                                                                                   |                                         |
| *Ion channel disorders                                                                                                                                                       |                                                                                                                   |                                         |
| LQTS                                                                                                                                                                         |                                                                                                                   |                                         |
| Brugada                                                                                                                                                                      |                                                                                                                   |                                         |
| SQTS                                                                                                                                                                         |                                                                                                                   |                                         |
| CPVT                                                                                                                                                                         |                                                                                                                   |                                         |
| Asian SUNDs                                                                                                                                                                  |                                                                                                                   |                                         |
| Secondary cardiomyopathies                                                                                                                                                   |                                                                                                                   |                                         |
| Infiltrative ✕                                                                                                                                                               | e.g amyloidosis, Gaucher's disease, Hurler's disease, Hunter's disease                                            |                                         |
| Storage ‡                                                                                                                                                                    | e.g hereditary haemachromatosis, Fabry's disease, glycogen storage disease (type II, Pompe), Niemann-Pick disease |                                         |
| Toxicity                                                                                                                                                                     | e.g drugs (e.g cocaine), alcohol, heavy metals (e.g cobalt), chemical agents                                      |                                         |
| Endomyocardial                                                                                                                                                               | e.g endomyocardial fibrosis, Loeffler's endocarditis                                                              |                                         |
| Inflammatory (granulomatous)                                                                                                                                                 | e.g. sarcoidosis, post-infective (e.g. Chagas' disease)                                                           |                                         |
| Endocrine                                                                                                                                                                    | e.g. diabetes mellitust, hyper- or hypo-thyroidism,hyperparathyroidism, phaeochromocytoma,acromegaly              |                                         |
| Cardiofacial                                                                                                                                                                 | e.g. Noonan syndrome1, lentiginosist                                                                              |                                         |
| Neuromuscular/neurological                                                                                                                                                   | e.g. Friedrich's ataxiat, Duchenne-Becker muscular dystrophyt, myotonic dystrophy1 neurofibromatosist             |                                         |
| Nutritional deficiencies                                                                                                                                                     | e.g. beriberi {thiamin), pellagra, scurvy, selenium, carnitine, kwashiorkor                                       |                                         |
| Autoimmune/collagen                                                                                                                                                          | e.g. systemic lupus erythematosus, dermatomyositis, systemic sclerosis                                            |                                         |
| Consequences of cancer therapy                                                                                                                                               | e.g. anthracyclines {e.g. doxorubicin), daunorubicin, cyclophosphamide, radiation                                 |                                         |
| LQTS, long QT syndrome; SQTS, short QT syndrome; CPVT, catecholaminergic polymorphic ventricular tachycardia; Asian SUNDs, Asian sudden unexpected nocturnal death syndrome. |                                                                                                                   |                                         |
| *Excluded from the European Society of Cardiology classification.                                                                                                            |                                                                                                                   |                                         |
| ✕ Accumulation of abnormal substances between myocytes [I e. extracellular).                                                                                                 |                                                                                                                   |                                         |
| § Genetic (familial) origin.                                                                                                                                                 |                                                                                                                   |                                         |
| ‡ Accumulation of abnormal substances within myocytes [I.e. intracellular).                                                                                                  |                                                                                                                   |                                         |

### 1.1.1 Clinical presentation

Early clinical stages of cardiomyopathies may be asymptomatic but typical symptoms include reduced effort tolerance, fatigue, oedema, paroxysmal nocturnal dyspnoea, orthopnoea, cough and shortness of breath.<sup>25</sup> Each of the cardiomyopathies affect different aspects of the

cardiovascular axis and the myocardium through different mechanisms with consequent varying phenotypes associated with individual cardiomyopathies. (Figure 1.4).<sup>26</sup>

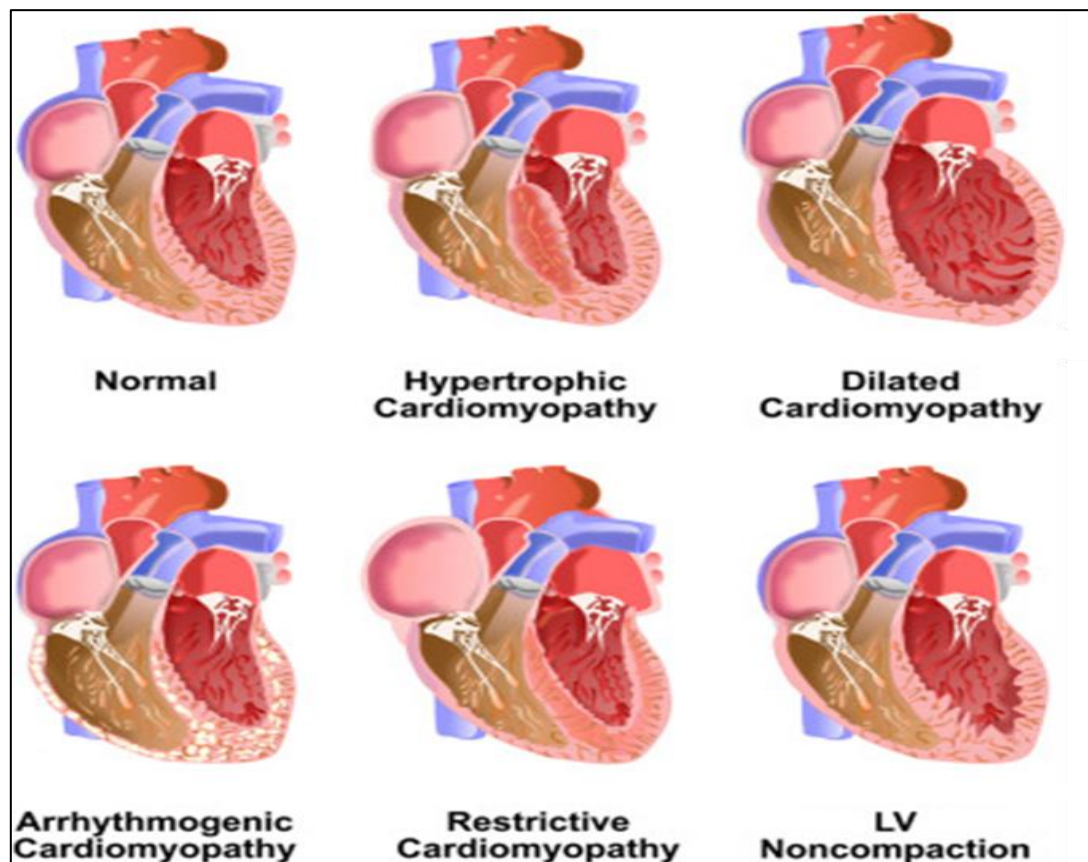


Figure 1.4: Images of a normal heart and the different morphofunctional subtypes of primary cardiomyopathies.<sup>26</sup>

### 1.1.2 Classification

The nomenclature and classification of cardiomyopathies have evolved over time due to the significant technological shifts in diagnoses, the discovery of new disease entities such as channelopathies in the course of a decade and well-defined understanding of the basis of these myocardial disorders (Figure 1.5). The improved classification scheme has enabled a better opportunity to develop robust frameworks to diagnose and gain more knowledge on the cardiomyopathies alongside facilitating interaction among clinicians thus providing more clarity in the genetic analyses of identifying their molecular mechanisms.<sup>27,28</sup>

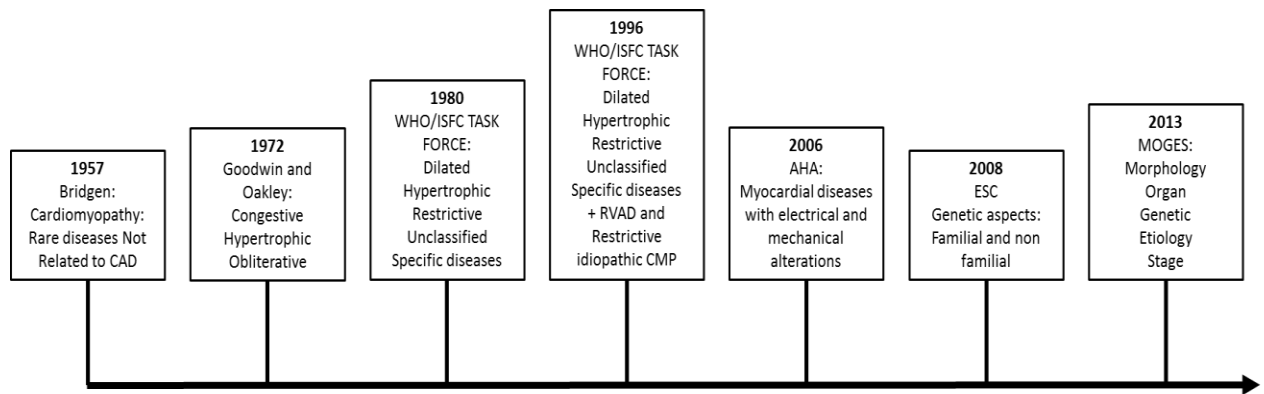


Figure 1.5: A timeline of the evolution of the classification of the cardiomyopathies.<sup>28</sup>

The first classification of cardiomyopathies was based on the predominant phenotypes observed at the time, including DCM, HCM, and RCM.<sup>30</sup> Further classification systems were proposed in the 1970s by Goodwin and colleagues,<sup>31</sup> and in 1996 by the World Health Organization (WHO) and the International Society and Federation of Cardiology.<sup>32</sup> The classification of cardiomyopathies was reviewed again by the American Heart Association (AHA) in 2006 and the European Society of Cardiology (ESC) in 2008.<sup>19,20</sup> The ESC position emphasised the important distinction between genetic and non-genetic forms of disease to highlight that cardiomyopathy results from a spectrum of disorders (Figure 1.6).<sup>20</sup>

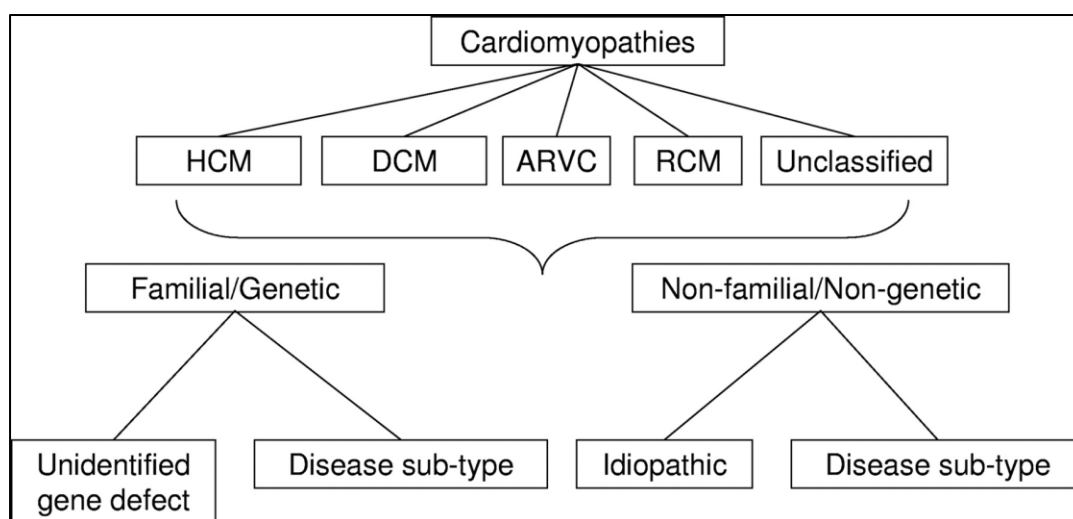


Figure 1.6: ESC classification of the cardiomyopathies into familial and non-familial sub-types.<sup>20</sup>

To date, four classic morphologic subtypes of dilated (DCM), arrhythmogenic (ACM), hypertrophic (HCM), and restrictive (RCM) cardiomyopathies are described and comprise primary cardiomyopathies.<sup>20</sup> In 2013, the World Heart Federation (WHF) proposed the MOGE(S) classification, the most recent classification system (Figure 1.7), and unlike the ESC classification which defines cardiomyopathies based on morphologic phenotypes, the MOGE(S) scheme includes both morphologic phenotype, genotype, systemic involvement and functional status.<sup>33</sup>

| NOTATION        | M<br>MORPHO-FUNCTIONAL<br>PHENOTYPE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | O<br>ORGAN/SYSTEM<br>INVOLVEMENT                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | G<br>GENETIC INHERITANCE<br>PATTERN                                                                                                                                                                                                                                                                                                                                                                    | E<br>ETIOLOGY                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | S<br>STAGE                                                                                                                                                                           |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| CHARACTERISTICS | Proband's cardiomyopathy (CM) diagnosis (DCM, HCM, RCM, ARVC/D, LVNC)                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Clinical history and evaluation<br>Organ involvement: Extracardiac organs/tissues<br>Multidisciplinary evaluation according per clinical needs or diagnostic hypothesis                                                                                                                                                                                                                                                                                                                               | Genetic counseling with pedigree<br>Clinical family screening<br>Affected, asymptomatic relative unaware of the disease<br>Relatives with ECG and/or Echo abnormalities<br>Healthy family members with normal ECG and ECHO                                                                                                                                                                             | Genetic testing in the proband<br>Positive: Cascade genetic testing in relatives<br>Negative: New tests novel genes<br>Regular monitoring in relatives                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | Functional status ACC/AHA, NYHA                                                                                                                                                      |
| SUBSCRIPT       | <b>D</b> Dilated<br><b>H</b> Hypertrophic<br><b>R</b> Restrictive<br><b>REMF</b> Endomyocardial fibrosis<br><b>LV</b> =left ventricle<br><b>RV</b> =right ventricle<br><b>RLV</b> =biventricular<br><b>A</b> ARVC<br><b>M</b> =major<br><b>m</b> =minor<br><b>c</b> =category<br><b>LV</b> = left ventricle<br><b>RV</b> =right ventricle<br><b>RLV</b> =biventricular<br><b>NC</b> LVNC<br><b>E</b> Early, with type in parentheses<br><b>NS</b> Nonspecific phenotype<br><b>NA</b> Information non available<br><b>O</b> Unaffected* | <b>H</b> Heart<br><b>LV</b> =left ventricle<br><b>RV</b> =right ventricle<br><b>RLV</b> =biventricular<br><b>M</b> Muscle (skeletal)<br><b>N</b> Nervous<br><b>C</b> Cutaneous<br><b>E</b> Eye, Ocular<br><b>A</b> Auditory<br><b>K</b> Kidney<br><b>G</b> Gastrointestinal<br><b>Li</b> Liver<br><b>Lu</b> Lung<br><b>S</b> Skeletal<br><b>O</b> Absence of organ/system involvement*, e.g. in family members who are healthy mutation carriers; the mutation is specified in E and inheritance in G | <b>N</b> Family history negative<br><b>U</b> Family history unknown<br><b>AD</b> Autosomal dominant<br><b>AR</b> Autosomal recessive<br><b>XLD</b> X-linked dominant<br><b>XLR</b> X-linked recessive<br><b>XL</b> X-linked<br><b>M</b> Matrilineal<br><b>O</b> Family history not investigated*<br><b>Undet</b> Inheritance still undetermined<br><b>S</b> Phenotypically Sporadic (apparent or real) | <b>G</b> Genetic cause<br><b>OC</b> Obligate carrier<br><b>ONC</b> Obligate non-carrier<br><b>DN</b> De novo<br><b>Neg</b> Genetic test negative for the known familial mutation<br><b>N</b> Genetic defect not identified<br><b>O</b> No genetic test, any reason*<br><b>G-A-TTR</b> Genetic amyloidosis<br><b>G-HFE</b> Hemochromatosis<br><b>Non-genetic etiologies:</b><br><b>M</b> Myocarditis<br><b>V</b> Viral infection (add the virus identified in affected heart)<br><b>AI</b> Autoimmune/immune-mediated; suspected (AI-S), proven (AI-P)<br><b>A</b> Amyloidosis (add type: A-K, A-L, A-SAA)<br><b>I</b> Infectious, non viral (add the infectious agent)<br><b>T</b> Toxicity (add cause/drug)<br><b>EO</b> Hypereosinophilic heart disease<br><b>O</b> Other | <b>ACC-AHA</b> stage represented as letter A, B, C, D<br><b>NA</b> not applicable<br><b>NU</b> not used<br><b>followed by</b> NYHA class represented as Roman numeral I, II, III, IV |

Figure 1.7: The MOGES classification for cardiomyopathies.<sup>33</sup>

### 1.1.3 Imaging modalities

The importance of these imaging modalities cannot be overstated, as they play a critical role in the ability to diagnose, guide therapy and predict outcomes. The function and architecture of the heart has been determined more precisely with current imaging techniques and this has translated into an improved diagnosis of cardiomyopathies in the patients.<sup>150</sup> Modern cardiology has been revolutionised by imaging modalities such as echocardiography, positron emission tomography (PET) (Figure 1.8), cardiovascular magnetic resonance (CMR) and computed tomography

(CT)<sup>19,20,34,35</sup> and as they continue to evolve, they promise to lead directly to improved patient outcome by allowing for earlier and more accurate diagnosis, and in some cases, to a better and more successful treatment.

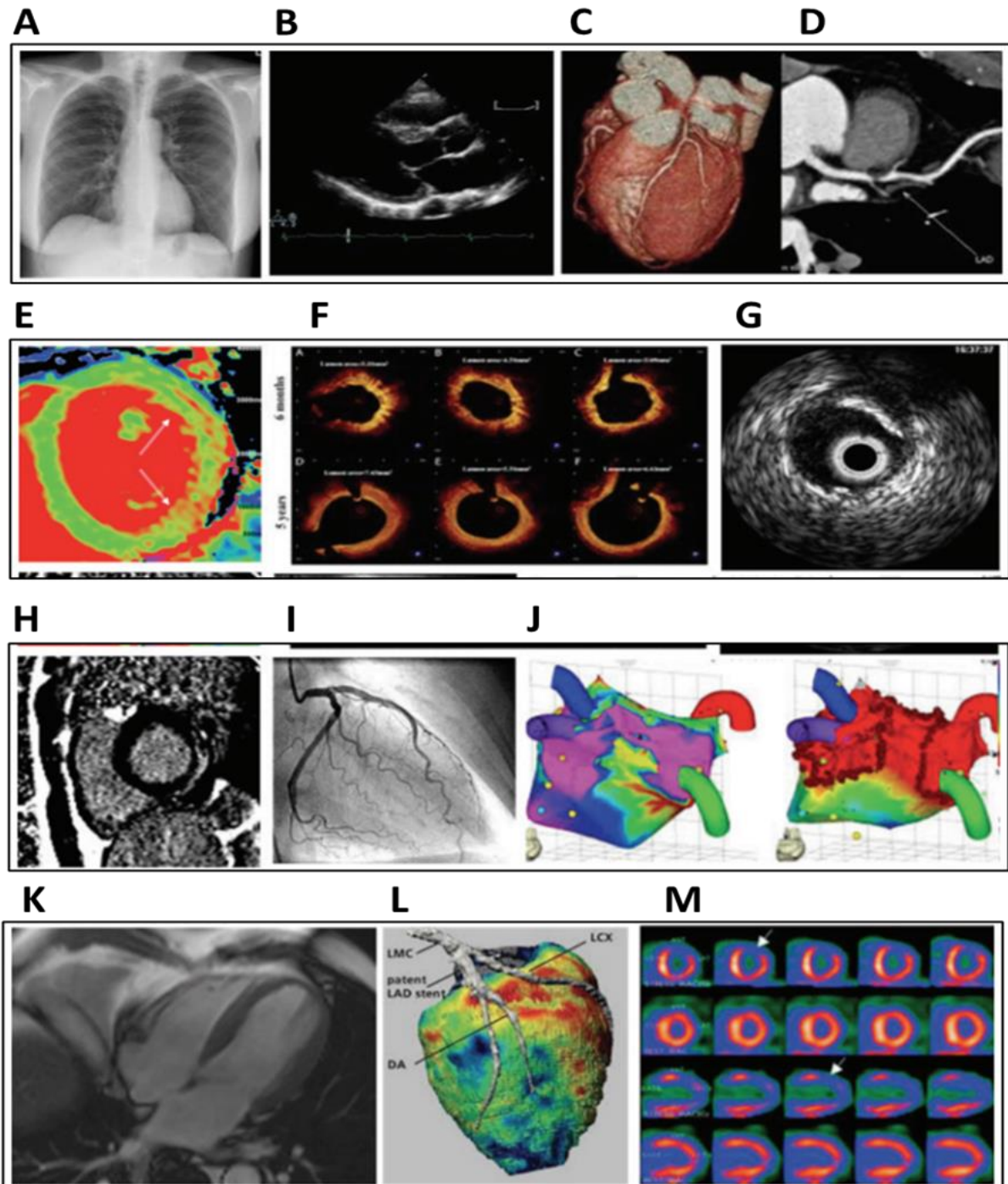


Figure 1.8: Cardiovascular imaging modalities used for the assessment of CVD A.) Chest x-ray showing the position of the heart. B.) Normal echocardiogram of the heart. C.) Computed tomography showing the heart. D.) Computed tomography showing the left anterior descending artery. E.) CMR tissue characterization of the left ventricle from base to apex using T1 mapping. F.) Cardiac Intracoronary Optical Coherence Tomography (OCT). G.) Intravascular ultrasound imaging for percutaneous coronary intervention. H.) Computed tomography angiography of the heart showing calcification of the heart blood vessels. I.) Coronary artery angiography (CAG) of the left coronary artery. J.) Color doppler imaging of the heart vessels showing flow. K.) CMR of the heart. L.) Cardiac PET showing 3-D functional processes in the heart. M.) SPECT myocardial perfusion for assessment of LV function.<sup>19,20,34,35,283-287</sup>

The lack of proper healthcare resources, few cardiac specialists and a significant budget neglect towards NCDs makes it challenging to properly diagnose the myocardial disorders and while these issues are seen largely in SSA, South Africa shows greater promise due to its development of electrophysiological approaches geared towards cardiological interventions that match international standards.<sup>37-40</sup>

#### 1.1.4 Therapies

Therapeutic interventions in cardiomyopathies are usually aimed at symptomatic relief as the management plan, alongside delaying heart failure, preventing sudden death and prolonging survival.<sup>60</sup> The treatment options are in form of pharmacological considerations as seen in patients with systolic dysfunction being administered inhibitors of adrenergic or neurohormonal pathways. Lifestyle changes and implantable cardioverter defibrillators (ICD) depending on risk factors such as electrophysiological evaluation, proband status and gender have also been used as approaches to therapeutically deal with the cardiomyopathies.<sup>61-63</sup> Targeted therapies are considered when conventional therapeutic interventions do not significantly improve the prognosis of some cardiomyopathy patients and these will have a larger role to play as the molecular mechanisms and pathogenesis of the cardiomyopathies are increasingly unravelled.<sup>11,64</sup>

New therapies termed as emerging therapies are gradually being introduced and these aim at timing the disease before manifestation due to genetic or environmental drivers of disease pathophysiology.<sup>55</sup> The emerging therapies include somatic gene therapies through directly targeting the cells of interest with nucleotide sequences delivered through mechanisms that include viral vectors and lipid nanoparticles.<sup>65,282</sup> Regulators of primary and secondary disease pathways are other emerging techniques that are being explored. Primary disease regulation involves the primary disease trajectory being affected by the protein encoded due to a genetic anomaly and this includes approaches such as protein stabilizers, protein quality control and direct contractile modulators.

The regulators of secondary disease pathways usually consider the downstream effects of a primary genetic aberration and they are possible through the use of targets for metabolism regulators and pathways for downstream signaling.<sup>65</sup>

However, heart failure with reduced ejection fraction (HFrEF) is still associated with undesirable morbidity and mortality rates, with as high as 50% of those who are affected passing away within 5 years of diagnosis. This has prompted researchers to consider how gene therapy may be used to treat or prevent HFrEF by raising the level of a particular cardiac protein.<sup>64,65,66</sup> The effective transition of gene therapy for HF patients from the laboratory to the clinical setting is an achievable venture, with many patients having a significant opportunity to benefit from the numerous genomic investigations that lead to the discovery of mutations responsible for HFrEF.

### 1.1.5 Genetics

The primary cardiomyopathies have significant heritable components which can be caused by either single or multiple genetic anomalies. Each primary cardiomyopathy can be caused by missense, stop-gain, frameshift, intronic or splice site mutations in the DNA.<sup>41</sup>

#### 1.1.5.1 Single genes

A rare disease-causing mutation in a particular gene can be causative for cardiomyopathy, resulting in a monogenic disorder and even though such monogenic disorders are individually rare, collectively they are common and are associated with adverse clinical presentations with high morbidity and mortality.<sup>42,43</sup> The monogenic forms of cardiomyopathies have been attributed to at least 15 genes located at various parts of the cell.<sup>44-47</sup> *TTN* and *BAG3* genetic variations have been reported to cause monogenic DCM while monogenic forms of ARVC have been associated with *CDH2* and *DSG2* genetic changes.<sup>42,48</sup> The frequent misdiagnosis of monogenic myocardial disorders

highlights an area that can be resolved by using approaches such as precision medicine for genetic testing of patients purported to have heritable cardiomyopathy.<sup>49</sup>

#### *1.1.5.2 Multiple genes*

Polygenic variants, which represent more than one genetic variant, may frequently be the cause of cardiomyopathy. Allelic variability of cardiomyopathies is highlighted by several gene correlations in various cardiomyopathies and associated harmful mutations in numerous genes.<sup>50</sup> Multiple variants are usually present in DCM, HCM and ACM patients, which typically results in an earlier age of disease onset and a more severe disease.<sup>51</sup> Cardiomyopathies exhibit heterogeneity in both their clinical and genetic manifestations and this can even be seen in family members sharing a disease-causing gene anomaly having a variable expressivity that is reflected in distinct disease manifestations among them.<sup>50,52</sup> We also address the effect of multiple genes under the section for Polygenic Risk Scores (1.5.5.4).

#### *1.1.5.3 Modifiers*

Cardiomyopathy disease severity and prognosis can be affected by additional factors collectively known as modifiers which influence the expressivity and penetrance of mendelian inheritance patterns even in the cardiomyopathies caused by rare monogenic variants.<sup>50</sup> The modifiers are genetic and non-genetic factors and their effects are evidenced by a spectrum of phenotypic observations ranging from mild with normal longevity to early arrhythmic death in ACM, differential LVESV in DCM and differential cardiac contractility in HCM.<sup>53,54</sup>

The genetic modifiers could be single rare variants and/or polygenic common variants in the human genome while the non-genetic modifiers could include sex and the environment and together, they

interact to bring about the variable expressivity and incomplete penetrance of disease states in the cardiomyopathies. It is worthwhile to note that the diverse phenotypes seen within the same family as in the mendelian cardiovascular diseases is particularly prominent in ACM with 67% of families having a subtype discordance which describes the coexistence of classic or left-dominant with biventricular disease patterns.<sup>294</sup>

Identification of the genetic modifiers is through the investigation of gene networks that drive the biological processes of the primary disease-causing variants, as will be described in this thesis and this is mostly a hypothesis-based approach formulated on evidence from current experimental and in silico knowledge. It has been shown that while both monogenic and polygenic risk variants influence the risk of developing disease, polygenic risk scores account for significant variability in Mendelian disease inheritance because genetic testing identifies a causal rare variant in less than half of the cases yet data from both families and population cohorts show a complex genetic interplay due to the reduced penetrance and variable expressivity.<sup>291</sup> These observations are seen in cardiomyopathies such as DCM where common variants in genes that encode for regulatory proteins, ion-binding and peptide receptors modify the risk of progressive heart failure and arrhythmic events thus predicting the survival rate and also in HCM where common variants in the renin-angiotensin-aldosterone system (*AGT*, *CYP11B2*, *ACE*, *ACE2* and *AGTR1*) increase phenotypic severity as seen through earlier disease onset and increased LV hypertrophy.<sup>291</sup> In ACM, SCD is the most conspicuous first clinical manifestation in 20% of the probands whereas most of the affected relatives show favorable prognosis with an annual mortality rate of 0.1%.<sup>294</sup> Taken together, the above strongly support the gene-dose effects and environmental factors as growing factors in modifying disease risk.

Rare monogenic variants in themselves are not able to sufficiently explain the widespread variable expressivity and penetrance observed in the cardiomyopathies and through improved systematic investigation of both modifier genes and environmental influences, the penetrance and expressivity

of genetic variants will increasingly become pivotal in identifying causative mechanisms and providing refined disease risks among the population with and at risk for developing cardiomyopathy.<sup>290</sup>

#### *1.1.5.4 Polygenetic Risk Scores*

Polygenic risk scores (PRS) can also play a role in the heterogeneity observed within families, which are the weighted summations of single nucleotide variants (SNVs) which individually confer trivial risk to cardiomyopathy but have significant risk to disease collectively.<sup>288</sup>

The PRSs are determined from large external data, typically GWAS which represents the association of the SNV with the particular disease and thus, each PRS is autonomously associated with its respective disease.<sup>289</sup> Both PRSs and discovery of rare monogenic variants are used to identify the risk factors in cardiovascular disease but the two distinctly differ because polygenic scores are used for the simultaneous calculation of the risk scores for disease while rare monogenic variant discovery necessitates the sequencing of particular genes and thorough interrogation of any functional anomalies the monogenic variants have.<sup>289</sup>

PRSs play a role in the genetic landscape of the cardiomyopathies and studies have candidate genes have been identified at genomic loci implicated in the cardiomyopathies in genes such as *TTN*, *FLNC*, *BAG3*, *ALPK3* and *PLEKHM2*.<sup>290</sup> The studies have shown that while common variants identified in the candidate genes have little to no effect in disease causality in isolation, the effects of the variants confer additive effects involving crucial cardiomyocyte architecture such as calcium channeling, sarcomere regulation and protein homeostasis.<sup>290,291</sup> It is worthwhile to note that while the studies have shown a link between the common variants in genes and the cardiomyopathies such as DCM and HCM, some of them have also shown that there is no evidence to suggest an interaction in the molecular pathway between the variants to bring about an additive effect of their PRSs despite the

increased susceptibility to disease in genes such as *HSPB7* and *BAG3* (rs945417 and rs2234962 respectively).<sup>292</sup>

Population studies to profile cardiomyopathy risks through PRSs have improved the study of these myocardial disorders, although the Eurocentric biases in genetics studies have led to major scientific opportunities being missed such as the study of disease-causing variants in African and South Asian populations which have the most genetic diverse populations and most loss-of-function variants, which can significantly improve the global interpretation of genomic variation thus helping to bridge the gap in health disparities.<sup>293</sup>

#### *1.1.5.5 Genetics in sub-Saharan Africa*

SSA, the cradle of mankind, is a region with greater genetic variation than the rest of the world put together and studying the genetic cause of disease in our South African families will yield significant opportunities for scientific discovery.<sup>55</sup>

However, there is a scarcity of publications on the genetics of cardiomyopathies. Through improved technologies such as next generation sequencing (NGS), this is slowly changing and we are now able to study the genetics of cardiomyopathy in multiple African countries including South Africa, Tunisia, Egypt, Uganda and Morocco.<sup>56</sup> Through research done in Africa by the Cardiovascular Genetics (CVG group) at the University of Cape Town (UCT) and other African researchers, small strides have been taken in bridging the gaps in knowledge (Table 1.2).<sup>56</sup> Collectively, African researchers have detected new genes for ARVC, reported on genes responsible for severe forms of DCM and identified founder mutations in both HCM and ARVC. However, larger studies are needed to fill this cardiomyopathy void in SSA.<sup>48,54,57,58</sup>

**Table 1.2: African studies that have studied the genetics of inherited cardiomyopathies<sup>56</sup>**

| YEAR                        | REF | COHORT SIZE       | TECHNIQUES     | STUDY TYPE                     | GENE(S)                                                                                                                             | VARIANT/S                                                                                                                                                                                                                                                                                                     | COUNTRY      |
|-----------------------------|-----|-------------------|----------------|--------------------------------|-------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| Dilated cardiomyopathy      |     |                   |                |                                |                                                                                                                                     |                                                                                                                                                                                                                                                                                                               |              |
| 1999                        | 318 | 171               | Genotyping     | SNP assoc. (LVH) (+ve)         | <i>ACE</i>                                                                                                                          | 287bp alu repeat (intron16)                                                                                                                                                                                                                                                                                   | South Africa |
| 2002                        | 319 | 107               | Candidate gene | SNP assoc. (impr. LVEF) (+ve)  | <i>ACE</i> †, <i>AGT</i> ‡, <i>CYP11B2</i> §                                                                                        | Insertion/deletion †, c.704T > C‡ and c.344C > T§                                                                                                                                                                                                                                                             | South Africa |
| 2007                        | 320 | 394               | Genotyping     | SNP assoc. (LVEF&LVEDD)        | <i>β2-AR</i>                                                                                                                        | p.R16G and p.Q27E                                                                                                                                                                                                                                                                                             | South Africa |
| 2008                        | 321 | 37                | Genotyping     | SNP assoc. (risk HF)           | <i>α2C-AR</i>                                                                                                                       | Del322-325                                                                                                                                                                                                                                                                                                    | South Africa |
| 2008                        | 322 | 132               | Genotyping     | SNP assoc. (HF risk)           | <i>α2C-AR</i> ¶ and <i>β1-AR</i> &                                                                                                  | p.G389R¶ and Del322-325&                                                                                                                                                                                                                                                                                      | South Africa |
| 2010                        | 323 | 76                | Genotyping     | SNP assoc. (risk of DCM) (+ve) | <i>ACE</i>                                                                                                                          | DD vs. ID genotype                                                                                                                                                                                                                                                                                            | Tunisia      |
| 2015                        | 324 | 60                | Genotyping     | PCR                            | <i>HO-1</i>                                                                                                                         | GT (n) repeats                                                                                                                                                                                                                                                                                                | Egypt        |
| 2016                        | 57  | 315               | Candidate gene | Gene screen                    | <i>PLN</i>                                                                                                                          | p.R9C                                                                                                                                                                                                                                                                                                         | South Africa |
| 2018                        | 325 | One family        | NGS            | Targeted resequencing          | <i>LMNA</i> (50 genes)                                                                                                              | p.R541C                                                                                                                                                                                                                                                                                                       | Morocco      |
| 2022                        | 355 | 68                | NGS            | Targeted resequencing          | <i>TNNI3</i><br><i>TTN</i><br><i>DSG2</i><br><i>MYH7</i><br><i>VCL</i><br><i>LAMA</i><br><i>RBM20</i><br><i>TPM1</i><br><i>BAG3</i> | <i>TNNI3</i> (p.R204C); <i>TTN</i> (p.Q34757T); <i>TTN</i> (p.Y2336H); <i>TTN</i> (p.G24976A); <i>DSG</i> (p.K294G); <i>MYH7</i> (p.R941C); <i>MYH7</i> (p.R1359H); <i>MYH7</i> (p.V1670L); <i>VCL</i> (p.P347Q); <i>LAMA4</i> (p.P1359S); <i>RBM20</i> (p.R7166E); <i>TPM1</i> (p.T72A); <i>BAG3</i> (p.H24) | Egypt        |
| 2018                        | 356 | Pediatric patient | NGS            | WES                            | <i>ALPK3</i>                                                                                                                        | (p.K511Rfs*12)                                                                                                                                                                                                                                                                                                | Tunisia      |
| Hypertrophic cardiomyopathy |     |                   |                |                                |                                                                                                                                     |                                                                                                                                                                                                                                                                                                               |              |
| 2009                        | 326 | 143               | Genotyping     | SNP assoc. (BP) (+ve)          | <i>TNNT2</i> , <i>MYH7</i>                                                                                                          | p.R92W                                                                                                                                                                                                                                                                                                        | South Africa |

|      |     |                   |                |                              |                                                                         |                                                                                                    |              |
|------|-----|-------------------|----------------|------------------------------|-------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------|--------------|
| 1993 | 327 | 5                 | Candidate gene | Gene screen                  | <i>βMHC</i>                                                             | p.R403                                                                                             | South Africa |
| 1995 | 328 | 32                | Candidate gene | Gene screen                  | <i>βMHC</i>                                                             | p.A797T                                                                                            | South Africa |
| 1995 | 329 | 2 large pedigrees | Candidate gene | Gene screen                  | <i>βMHC</i>                                                             | p.R403Trp and p.R249Q                                                                              | South Africa |
| 1997 | 330 | 114               | Genotyping     | SNP assoc. (LVH) (+ve)       | <i>ACE</i>                                                              | p.R403                                                                                             | South Africa |
| 1997 | 331 | 48                | Candidate gene | Gene screen                  | <i>TNNT2</i>                                                            | p.R92W                                                                                             | South Africa |
| 1998 | 332 | 4                 | Candidate gene | Gene screen                  | <i>MYBPC3</i>                                                           | p.R654H                                                                                            | South Africa |
| 1999 | 333 | 40                | Candidate gene | Gene screen                  | <i>βMHC, MYBPC</i>                                                      | p.Q499K in <i>βMHC</i> and p.V896M and 756 in <i>MYBPC</i>                                         | South Africa |
| 2001 | 334 | 68                | Candidate gene | Gene screen                  | <i>MYL2</i> and <i>MYL3</i>                                             | None found                                                                                         | South Africa |
| 2002 | 335 | 82                | Candidate gene | Gene screen                  | <i>MYH7</i>                                                             | p.A1379R and p.S1776G                                                                              | South Africa |
| 2000 | 336 | 66                | Genotyping     | SNP assoc. (prognosis) (+ve) | <i>MYH7</i>                                                             | c.797A > T                                                                                         | South Africa |
| 2008 | 337 | 227               | Genotyping     | SNP assoc. (LVH) (+ve)       | <i>ACE2</i>                                                             | rs1978124, rs2285666 or rs4646179, and rs879922                                                    | South Africa |
| 2011 | 338 | 353               | Genotyping     | SNP assoc. (LVH) (+ve)       | <i>AGTR2</i>                                                            | +1645G > A                                                                                         | South Africa |
| 2012 | 339 | 192               | Candidate gene | Gene screen                  | <i>MYBPC3, MYH7, and TNNT2</i>                                          | 40% mutation rate in cohort                                                                        | Egypt        |
| 2016 | 340 | 45                | Targeted       | NGS                          | <i>MYH7, MYBPC3, TNNT2, TNNI3, ACTC1, TNNC1, MYL2, MYL3, TPM1, FLNC</i> | 27% cases mutation positive, predominantly in <i>MYBPC3</i> and <i>MYH7</i> as well as <i>MYL3</i> | Tunisia      |

|                                |     |                   |                |                           |                                                                                                                                                                                                                        |                                                                                                                                                                            |              |
|--------------------------------|-----|-------------------|----------------|---------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| 2016                           | 341 | 388               | Genotyping     | 7 SNPs as modifiers (+ve) | <i>MYBPH</i>                                                                                                                                                                                                           | c.1093+449T > C (rs2250509)                                                                                                                                                | South Africa |
| 2016                           | 128 | 43                | NGS            | Targeted resequencing     | <i>MYBPC3</i> , <i>MYH7</i> , <i>TNNT2</i> , <i>TNNI3</i> , <i>TPM1</i> , <i>MYL2</i> , <i>MYL3</i> , <i>ACTC1</i> , <i>PLN</i> , <i>CSRP3</i> , <i>FHL1</i> , <i>PRKAG2</i> , <i>GLA</i> , <i>LMNA</i> , <i>LAMP2</i> | MYH7 (p.R204H); MYH7 (p.T761N); MYH7 31 (p.R1420W); MYBPC3 (p.E334K); MYH7 (p.R723C); MYH7 (p.R249Q); MYBPC3 (p.E258K); MYBPC3 (p.R177H); MYH7 (p.R453C); MYBPC3 (p.G416S) | South Africa |
| 2017                           | 342 | One family        | NGS            | WES                       | <i>SLC22A5</i>                                                                                                                                                                                                         | p.Y396*                                                                                                                                                                    | Morocco      |
| 2020                           | 354 | 24                | NGS            | Targeted resequencing     | <i>MYBPC3</i> , <i>TTN</i> , <i>MYH7</i> , <i>MYL2</i> , <i>VCL</i> , <i>CSRP3</i> , <i>RBM20</i>                                                                                                                      | MYBPC3 (p.R495G); TTN (p.A16018V); MYH7 (p.R403Q); MYL2 (p.Q38R); TTN (p.L25430P); MYBPC3 (p.D75E); VCL (p.L682F); CSRP3 (p.R122Q); RBM20 (p.E1012G)                       | Egypt        |
| 2018                           | 356 | Pediatric patient | NGS            | WES                       | <i>ALPK3</i>                                                                                                                                                                                                           | (p.K511Rfs*12)                                                                                                                                                             | Tunisia      |
| Arrhythmogenic cardiomyopathy  |     |                   |                |                           |                                                                                                                                                                                                                        |                                                                                                                                                                            |              |
| 2006                           | 343 | 13                | Linkage        | Gene screen               | <i>ITG8</i> , <i>FRMD4A</i> , <i>LAMRIP6</i>                                                                                                                                                                           | None found                                                                                                                                                                 | South Africa |
| 2009                           | 59  | 50                | Candidate gene | Gene screen               | <i>PKP2</i>                                                                                                                                                                                                            | p.IVS2146-1G-C, p.A733fsX740, p.Q378X, p.R388W, p.G489R, p.S837fs, p.L847P                                                                                                 | South Africa |
| 2017                           | 344 | 1                 | NGS            | Targeted resequencing     | <i>MYBPC3</i>                                                                                                                                                                                                          | p.E1179K                                                                                                                                                                   | East Africa  |
| 2017                           | 48  | 2                 | NGS            | WES                       | <i>CDH2</i>                                                                                                                                                                                                            | p.Q229P and p.D407N                                                                                                                                                        | South Africa |
| Left ventricular noncompaction |     |                   |                |                           |                                                                                                                                                                                                                        |                                                                                                                                                                            |              |
| 2022                           | 355 | 68                | NGS            | Targeted resequencing     | <i>TTN</i> , <i>MYPN</i> , <i>DTNA</i>                                                                                                                                                                                 | TTN (p.S29003F); TTN (p.R13835Y); MYPN (p.V1183A); DTNA (p.P121L)                                                                                                          | Egypt        |
| Restrictive cardiomyopathy     |     |                   |                |                           |                                                                                                                                                                                                                        |                                                                                                                                                                            |              |
| 2015                           | 345 | 113               | Candidate gene | Gene screen               | <i>TNNI3</i>                                                                                                                                                                                                           | p.L144H, de novo p.R170Q                                                                                                                                                   | South Africa |

SNP- Single nucleotide polymorphism; *ACE*- Angiotensin I converting enzyme; *AGT*- Angiotensinogen; *CYP11B2*- Cytochrome P450 family 11 subfamily B member 2; *β2-AR*- Adrenoceptor beta 2; *α2C-AR*- Adrenoceptor alpha 2C; *β1-AR*- Adrenoceptor beta 1; *HO-1*- Haem oxygenase 1; *PLN*- Phospholamban; *LMNA*- Lamin A/C; *TNNI3*- Troponin I3, cardiac type; *TTN*- Titin; *DSG2*- Desmoglein-2; *MYH7*- b-myosin heavy chain; *LAMA*- ; *RBM20*- ; *TPM1*- Tropomyosin 1; *BAG3*-BCL2 associated athanogene 3; *ALPK3*- ; *TNNT2*- Troponin T2; *MYH7*-b-myosin heavy chain; *βMHC*- Beta myosin heavy chain; *MYBPC3*- myosin-binding protein C; *MYL2*-Myosin light chain 2; *MYL3*-Myosin light chain 3; *AGTR2*- Angiotensin II receptor type 2; *MYBPH*- ; *ACTC1*- ; *CSRP3*-Cysteine and glycine rich protein 3; *FHL1*- Four and a half LIM domains 1; *PRKAG2*- Protein kinase AMP-activated non-catalytic subunit gamma 2; *GLA*-

## 1.2 Dilated cardiomyopathy

### 1.2.1 Clinical presentation

DCM is a progressive disease that leads to left ventricular (LV) dilation and depressed myocardial performance which in turn results in a reduced systolic function, in the absence of significant overloading (as seen in hypertension), valvular, ischaemic, pericardial or congenital heart disease that may account for the abnormalities (Figure 1.9).<sup>67</sup>

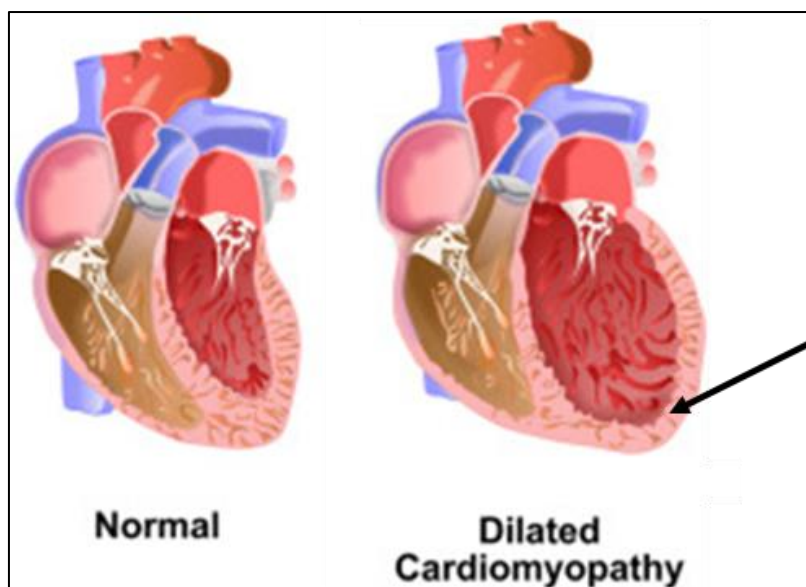


Figure 1.9: Comparison between a normal heart and a DCM heart. The black arrow points to the thinning wall of a dilated ventricle.<sup>67</sup>

DCM is a frequent indication for orthotopic heart transplantation in children and adults because of its deteriorating effects leading to HF which progresses to multiple organ failure, with a 20% mortality rate at 1 year and a 56% mortality rate at 4 years.<sup>68,69</sup>

Initial clinical presentation of DCM may be non-specific with symptoms ranging from dyspnoea on exertion, fatigue, oedema and some patients even presenting for the first time with sudden cardiac death (SCD), arrhythmia and/or systemic thromboembolism.

### 1.2.2 Aetiology

DCM has a varied aetiology, though this may remain idiopathic in up to 70% of the cases. In SSA, important causes include genetic, human immunodeficiency virus (HIV) infection, other cardiac infections, myocarditis, infiltrative disease, iron overload, endocrinopathies, nutritional deficiencies, PPCM, toxins, and alcohol-induced cardiomyopathy.<sup>70</sup>

### 1.2.3 Disease burden

DCM is the commonest cardiomyopathy observed in SSA; with an estimated prevalence of 1:250-1:400 in the population and a significant cause of HF among the young.<sup>68,69,71</sup> It is generally acknowledged that DCM affects many black Africans although the true burden of DCM in SSA is unknown because no population studies have been conducted to assess its prevalence; despite this, it is widely accepted that DCM is widespread among black Africans, based on data from hospitalised and ambulant patients.<sup>72</sup> DCM was responsible for 11.6% to 37.5% of cardiovascular diagnoses in clinical series conducted in South African populations and 15.4% to 48% of HF admissions.<sup>73,74</sup> In 14.1% to 17% of CVD in necropsy series, DCM was determined to result in death in up to 17% of autopsies studies and HF in 12.7% of HF deaths.<sup>75,76</sup> DCM mortality is relatively high despite modern medical and surgical interventions, and larger trials are needed to evaluate the putative therapies and health systems strengthening in SSA.<sup>77,78</sup>

### 1.2.4 Molecular genetics of DCM

Historically, most DCM cases have been reported to be idiopathic and many of these DCM cases have a genetic variant and are termed familial DCM.<sup>79,80</sup> Pathogenic mutations have been found in 20-30% of familial DCM patients.<sup>81,82</sup> Autosomal dominant (AD) inheritance pattern is the most common mode of familial DCM inheritance, reported in 20%-70% of cases; while less common modes of inheritance include X-linked and autosomal recessive (AR).<sup>56,79</sup> Patients with the familial

DCM often have incomplete penetrance, accompanied by heterogeneity of the clinical phenotype and variable expressivity.<sup>83,84</sup>

Molecular genetic studies have led to the discovery of DCM causal mutations in more than 60 genes which are involved in the function of various cellular components, most notably the sarcomeric proteins (myosin and troponin) which help to form the basic structural units of cardiac muscle and are essential for contraction and relaxation of the heart. Any mutations in the sarcomeric components leads to dysfunction in cardiac contraction which transitions to eventual heart failure.<sup>45</sup> Other vital cellular components that are affected by mutations in DCM genes include the cytoskeleton, desmosomes, ion channels and nuclear envelope (Figure 1.10). Genes implicated in DCM encode a set of proteins whose functions are closely related to the structural and functional integrity of cardiomyocytes.<sup>79,83,84</sup> Table 1.3 provides a summary of the more predominant DCM genes but a comprehensive list of genes implicated in DCM is highlighted in appendix A.<sup>346-348</sup>

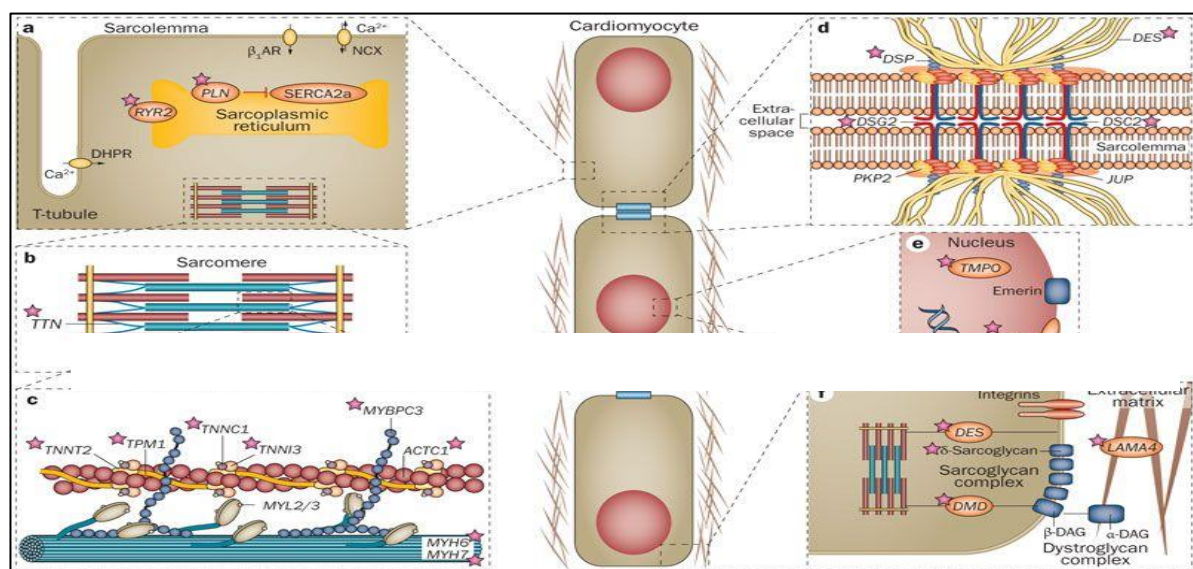


Figure 1.10: An illustration of the heterogeneous group of genetic variants that yield a DCM phenotype.<sup>85</sup>

**Table 1.3: Contribution of genes implicated in DCM**

| Gene         | Protein                                        | Estimated frequency among patients | References   |
|--------------|------------------------------------------------|------------------------------------|--------------|
| <i>TTN</i>   | Titin                                          | 25%                                | 18, 44,71,86 |
| <i>MYH6</i>  | Myosin-6 ( $\alpha$ -myosin heavy chain)       | 4%                                 | 18           |
| <i>MYPN</i>  | Myopalladin                                    | 3–4%                               | 71           |
| <i>TNNT2</i> | Cardiac muscle troponin T                      | 3%                                 | 18,44        |
| <i>SCN5A</i> | Sodium channel protein type 5 subunit $\alpha$ | 2–3%                               | 44           |
| <i>RBM20</i> | RNA-binding protein 20                         | 2%                                 | 18           |
| <i>PSEN1</i> | Presenilin-1                                   | <1%                                | 44,86        |
| <i>PLN</i>   | Phospholamban                                  | <1%                                | 18,86        |
| <i>DES</i>   | Desmin                                         | <1%                                | 44           |

## 1.3 Left ventricular noncompaction

### 1.3.1 Clinical presentation

LVNC is a cardiac disorder in which the myocardial structure is two-layered with a thin compacted epicardium surrounding a spongy noncompacted endocardium (Figure 1.11).<sup>63,64</sup>

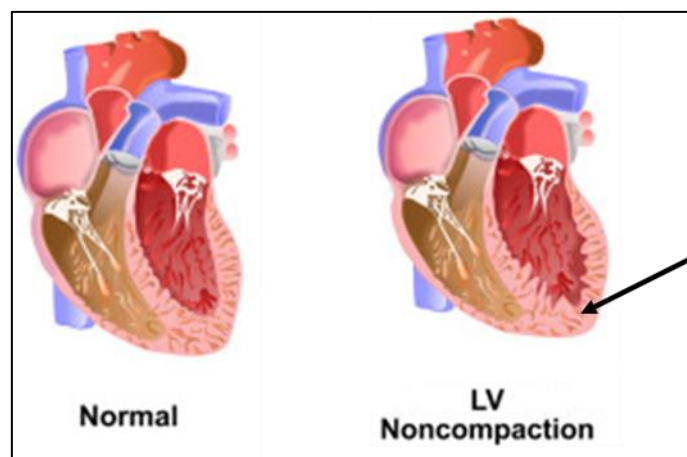


Figure 1.11: Comparison between a normal heart and an LVNC heart with the spongy myocardium shown with the black arrow.<sup>26</sup>

LVNC has wide phenotypic overlap with DCM, and usually presents with the clinical triad of HF, arrhythmia and embolic phenomena. LVNC may be seen together with other cardiac conditions including HCM, atrial/ventricular septal defects, pulmonary atresia, patent ductus arteriosus and Ebstein anomaly.<sup>50</sup>

### 1.3.2 Aetiology

The aetiology of LVNC is suggested to be the arrest of morphogenesis with the morphological features of the developing heart and noncompaction very evident.<sup>87,88,89</sup> The myocardium is a mesh of trabecular sponge-like muscle during mid-late embryonic life when myocardial blood supply is through the diffusion from the spaces in the intertrabeculae that are in communication with the chambers of the heart.<sup>90</sup> During the second embryonic month coronary vessels are formed after completion of cardiac looping and following valve development and ventricular and atrial septation. Disappearance of the “sinusoids” characterises maturation of the epicardial coronary with compact musculature through transformation of the spongy myocardium.<sup>88,89</sup> Ventricular noncompaction is stated to be attributed to aberrant remodelling of trabeculae during formation of the ventricular free wall while trabeculation reduction is involved in deficiencies in the ventricular compacted layer leading to embryonic heart failure and death.<sup>90</sup>

### 1.3.3 Disease burden

The prevalence of LVNC is detected in the general population at a rate of 1:5000 as more awareness is raised with the aid of better diagnostic tools.<sup>91</sup>

In Africa, LVNC has been found at a prevalence range of 0.5-6.9% based on five studies done in Northern Africa, Eastern Africa, and Southern Africa.<sup>56</sup>

### 1.3.4 Molecular genetics

Similar to the other cardiomyopathies, LVNC is genetically heterogenous and it can be inherited through AD, AR or x-linked recessive manner and is familial in 30-50% of cases.<sup>56</sup> The first genetic anomaly responsible for LVNC was reported in *TAZ* in affected male and female carriers in an X-linked family with LVNC and since then, mutations in >40 genes have been linked to LVNC.<sup>56,90</sup>

Mutations in sarcomeric genes are notably the significant identifiable cause in up to 30% of LVNC cases, with mutations in nonsarcomeric genes and the NOTCH signalling pathway being implicated in the genetic landscape of LVNC (Figure 1.12).<sup>56,91</sup> The estimated contribution of these genes to LVNC is shown in table 1.4 and a more comprehensive list of genes known to cause LVNC is provided in appendix B.<sup>349</sup>

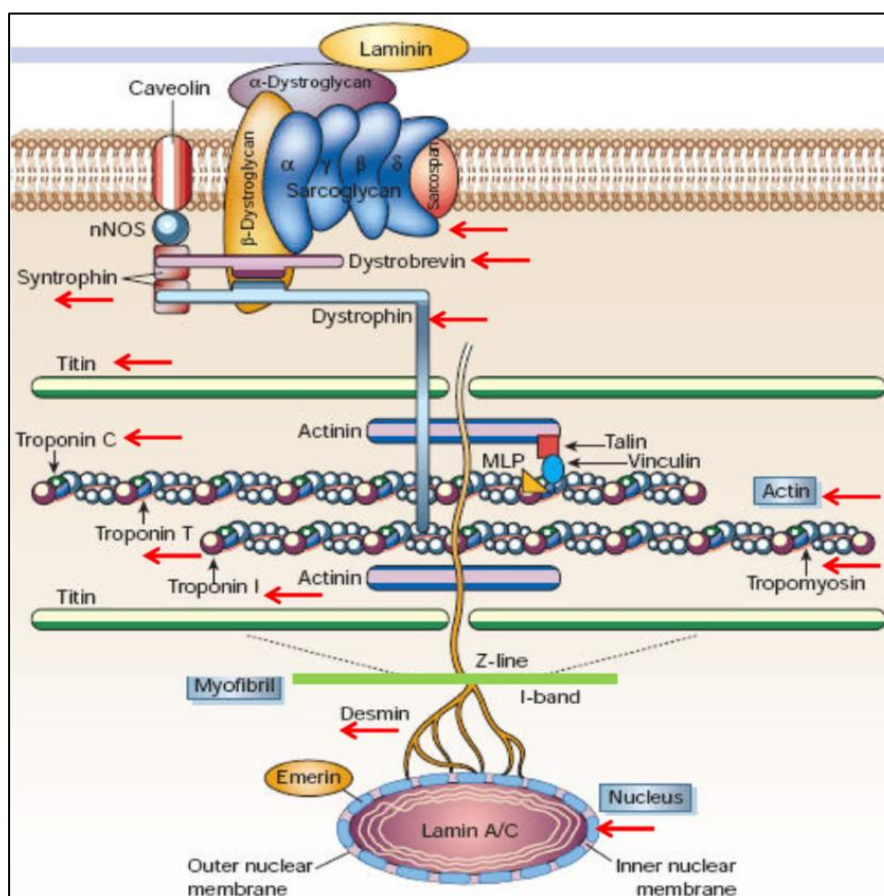


Figure 1.12: The genetic landscape of genes underlying the LVNC phenotype.<sup>56,91</sup>

**Table 1.4: Contribution of genes reported in LVNC**

| Gene         | Protein                          | Estimated frequency among patients (%) | References |
|--------------|----------------------------------|----------------------------------------|------------|
| <i>TTN</i>   | Titin                            | 19                                     | 92         |
| <i>LMNA</i>  | Lamin A/C                        | 5                                      | 92         |
| <i>MYBC3</i> | Cardiac myosin-binding protein C | 4                                      | 92         |
| <i>MYH7</i>  | β-Myosin heavy chain             | 3                                      | 92         |
| <i>CASQ2</i> | Calsequestrin 2                  | 2                                      | 92         |
| <i>TPM1</i>  | α-Tropomyosin                    | <1                                     | 92         |
| <i>TNNT2</i> | Cardiac troponin T               | <1                                     | 92         |
| <i>RYR1</i>  | Ryanodine receptor 1             | <1                                     | 93         |
| <i>RBM20</i> | RNA-binding motif protein 20     | <1                                     | 93         |
| <i>ACTC1</i> | Actin Alpha Cardiac Muscle 1     | <1                                     | 93         |

## 1.4 Arrhythmogenic cardiomyopathy

### 1.4.1 Clinical presentation

ACM is an arrhythmogenic disorder of the LV or right ventricle (RV) characterised by fibrofatty replacement of the myocardium with subsequent scar that serves as an arrhythmogenic substrate leading to HF, arrhythmia and SCD (Figure 1.13).<sup>94,95</sup>

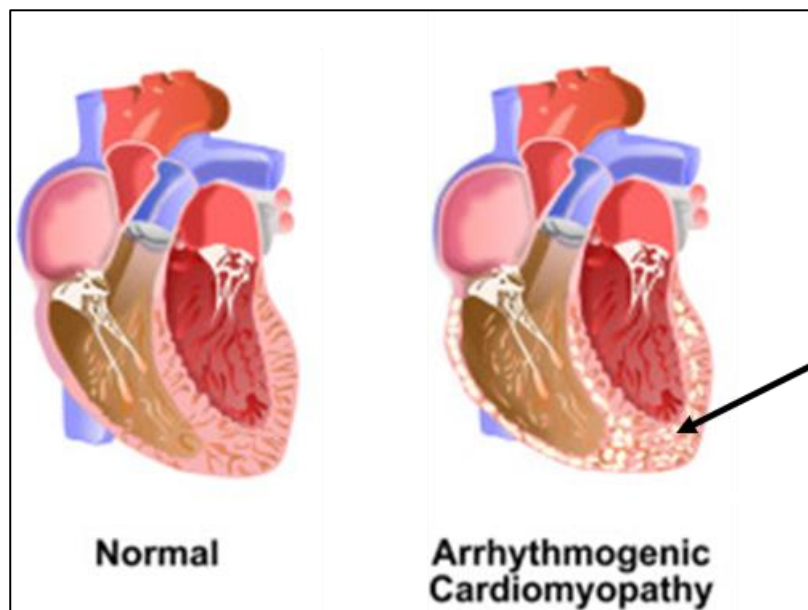


Figure 1.13: Comparison between a normal heart and an ACM heart with the fibrofatty infiltrations indicated by the black arrow.<sup>18,94</sup>

The ACM phenotype overlaps with DCM, RCM or HCM with arrhythmia presentation that may be associated with ventricular dilatation and/or impaired systolic function.

Clinically ACM manifests in the third to fourth decades of life with structural features, with a variable expression with a penetrance that is age-related.<sup>96,97</sup> Sustained ventricular arrhythmias, frequent palpitations, cardiac arrest and arrhythmic syncope, SCD are all common presentations.<sup>98</sup> Later in the disease, HF may occur. In most cases of desmoplakin-mediated LV disease may present as acute myocarditis, with elevated serum troponins and chest pain with ST-segment changes on the electrocardiogram (ECG).<sup>99</sup> Depending on the predominantly affected ventricle, clinical phenotype is varied.<sup>101</sup>

### 1.4.2 Aetiology

Sequential modifications in the expression of many genes characterises the adipocyte phenotype transition,<sup>102</sup> with some of these genes are involved in the Hippo and *Wnt* pathways which regulate the adipogenesis process in normal heart to ensure cardiomyocyte survival and proliferation.<sup>103</sup> Dysregulation of these pathways leads to the transmural fibrofatty replacement of contractile myocardium as the disease progresses.<sup>104</sup> Cardiomyocyte defects, including dysregulation of sodium channels, intracellular calcium regulating channels, gap junctions and eventually death of the cardiomyocytes, lead to the electrical instability in ACM.<sup>105</sup>

### 1.4.3 Disease burden

Prevalence of ACM has not been systematically studied in most countries but is estimated at 1:2000 to 1:5000 globally.<sup>97</sup> Disease expression is usually seen in adolescence and early adulthood, rarely in younger children, but may develop at any age.<sup>106</sup> In those younger than 35 years, ACM accounts for 11% of SCD cases and for 22% of SCD in athletes.<sup>107,108</sup> In South Africa, ACM is present in all racial and ethnic groups.<sup>59</sup>

### 1.4.4 Molecular genetics

ACM often has an AD pattern of inheritance, with mutations in elements of the desmosome and gap junctions resulting in an age-dependent and variable expressivity.<sup>109</sup> The desmosome is a membrane protein complex that has an integral role in stabilising cellular adhesion between cells and in maintaining structural integrity of tissues that undergo mechanical stress such as those in the heart and skin. Desmosomes are composed of 5 proteins namely plakoglobin (*JUP*), plakophilin-2 (*PKP2*), desmocollin-2 (*DSC2*) desmoglein-2 (*DSG2*) and desmoplakin (*DSP*). (Figure 1.14).<sup>103,110</sup>

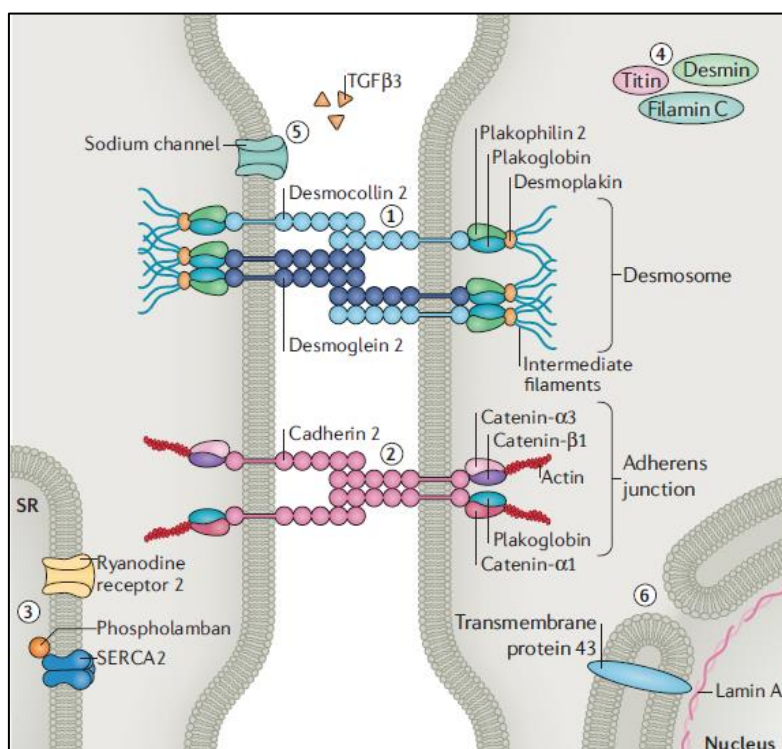


Figure 1.14: An illustration of the desmosomal and non-desmosomal genes in ACM.<sup>103</sup>

Desmosomal mutations have been reported in 33% to 63% of ACM probands, with *PKP2* being the major disease-causing gene accounting for 36% to 92% of desmosomal mutations identified.<sup>111,112</sup>

Other non-desmosomal gene variations have also been established to be responsible for ARVC as shown in Table 1.5 below with a complete list of ACM genes given in appendix C.<sup>350-353</sup>

**Table 1.5: Contribution of genes implicated in ACM**

| Gene        | Protein              | Estimated frequency among patients (%) | References |
|-------------|----------------------|----------------------------------------|------------|
| <i>PKP2</i> | Plakophilin 2        | 19–46                                  | 111,112    |
| <i>DSP</i>  | Desmoplakin          | 1–16                                   | 111,112    |
| <i>DSG2</i> | Desmoglein 2         | 2.5–10.0                               | 111        |
| <i>DSC2</i> | Desmocollin 2        | 1–8                                    | 111        |
| <i>JUP</i>  | Junction plakoglobin | 1                                      | 111        |
| <i>CDH2</i> | Cadherin 2           | <1–2                                   | 111        |
| <i>RYR2</i> | Ryanodine receptor 2 | <1                                     | 111        |
| <i>PLN</i>  | Phospholamban        | <1                                     | 111        |
| <i>FLNC</i> | Filamin C            | <1                                     | 111        |

## 1.5 Next-generation sequencing

Sanger sequencing has been the benchmark for DNA sequencing for a while and it produces relatively long DNA sequences of unmatched quality, though it is laborious and costly even when used in small genomes.<sup>113</sup> The release of pyrosequencing in 2005 was a milestone enabled one machine cycle to produce up to millions of short sequence reads as a high throughput technological advancement termed as next generation sequencing (NGS).<sup>114,115</sup>

NGS and Sanger share analogous principles involving the addition of fluorescent nucleotides by DNA polymerase onto an elongating DNA strand and a fluorescent signal being linked to each added nucleotide. The significant difference between NGS and Sanger sequencing is seen in their sequencing volume whereby NGS can sequence millions of fragments at once in a cycle while Sanger sequencing can only sequence one DNA fragment per cycle thus offering NGS a potent capability of identifying rare or novel variants.<sup>116,117,118</sup>

In NGS technology, short-read sequencing and long-read sequencing are the two major paradigms with their own benefits and shortcomings.

### 1.5.1 Short-read sequencing:

This approach offers more precise data at a lower cost thus being very vital in identifying variants of clinical significance and research at the population level. Short read sequencing techniques have an output ranging from at least 0.144 gigabases (Gb) to 3000 Gb with read lengths that are between 36 base pairs (bp) to 600bp thus resulting in turnaround times of between 9 hours to 213 hours. Sequencing by synthesis (SBS) and sequencing by ligation (SBL) are the two arms of short-read sequencing and are differentiated by the underlying detection chemistries (Figure 1.15).<sup>117</sup>

#### *1.5.1.1 Sequencing-by-synthesis (SBS)*

SBS approaches can fall into the category of either cyclic reversible termination (CRT) or single-nucleotide addition (SNA).

##### *Cyclic reversible termination (CRT)*

CRT is the most widely used SBS technique and it uses direct imaging to distinguish between fluorescent nucleotides which have been “bridge-amplified”.<sup>117</sup> NGS platforms include Illumina and GeneReader (Qiagen).

##### *Single-nucleotide addition (SNA)*

SNA approaches such as proton detection and pyrosequencing mark the inclusion of a dNTP onto a growing strand through solitary feedback. The SNA approaches are seen in Proton detection techniques including Ion Torrent (Thermo Fisher) or in Pyrosequencing through pyrophosphate generation used by 454 pyrosequencing (Roche).<sup>119-121</sup>

#### *1.5.1.2 Sequencing-by-ligation (SBL)*

Fundamentally, SBL approaches attach a hybridized probe onto a DNA sequence with imaging happening due to the hybridizing fluorophore indicating the particular position of a particular base. SBL options such as DNA nanoball sequencing anchor DNA sequences in form of DNA nanoballs onto slots for the sequencing reactions.<sup>122,123</sup> Examples of platforms include SOLiD and Complete Genomics (Beijing Genomics Institute (BGI) (Figure 1.15).

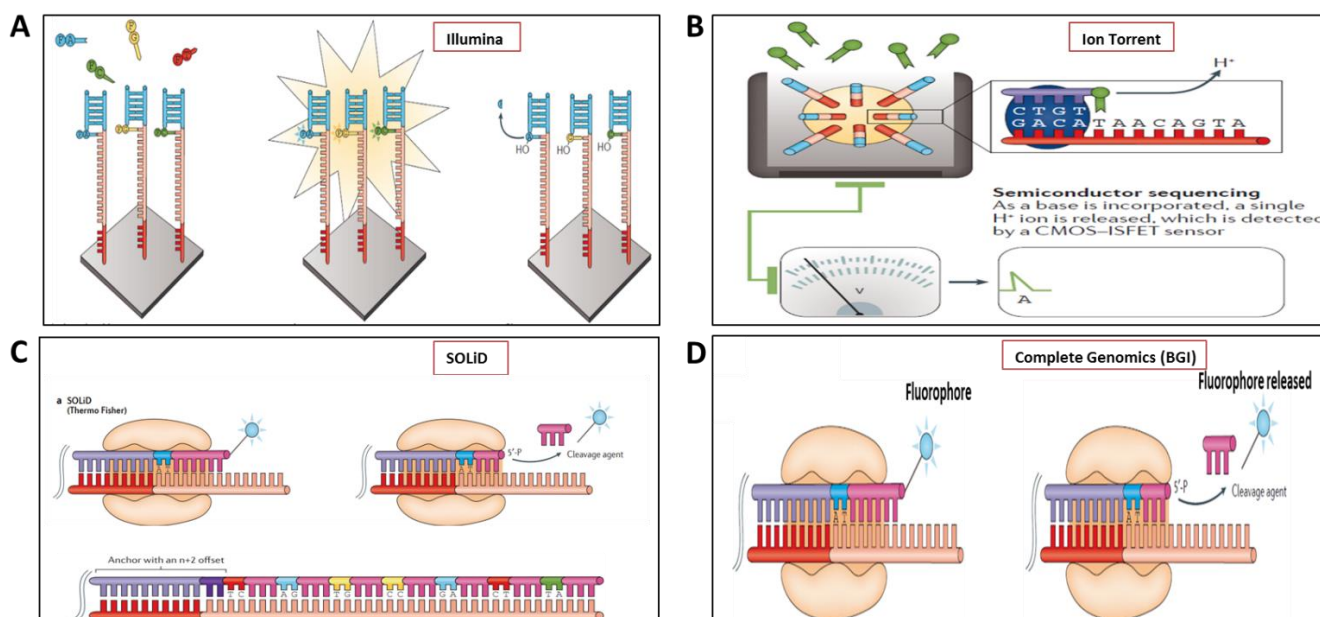


Figure 1.15: **Short read sequencing approaches** A.) ILLUMINA (sequencing-by-synthesis). Illumina uses fluorophore-labelled, terminally blocked nucleotides hybridise to complementary base. Each cluster on a slide can incorporate a different base. B.) Slides are imaged with either two or four laser channels. Each cluster emits a colour corresponding to the base incorporated during this cycle. C.) Fluorophores are cleaved and washed from flow cells and the 3'-OH group is regenerated. A new cycle begins with the addition of new nucleotides. B. ION TORRENT (sequencing-by-synthesis) Ion Torrent sequencing flow: The signal generated from the release of a  $H^+$  ion as a pH change is captured as a signal and recorder corresponding to a particular nucleotide<sup>116</sup> C. SOLiD (sequencing-by-ligation) The SOLiD sequencing procedure is composed of a series of probe-anchor binding, ligation, imaging and cleavage cycles to elongate the complementary strand. Over the course of the cycles, single-nucleotide offsets are introduced to ensure every base in the template strand is sequenced. D.) Complete Genomics (Beijing Genomics Institute) sequencing ((sequencing-by-ligation), after each sequencing round of fluorophore hybridization and release, the fluorophore release is captured as a signal which corresponds to a particular nucleotide base It combines probe-anchor ligation chemistry with low reagent consumption and quality imaging, and is the only current sequencing technique that employs SBL<sup>117,122,123</sup>

### 1.5.2 Long-read sequencing:

In contrast to short-read sequences, long-read sequencers generate sequencing data more than several kilobases to one megabase (Mb) of read lengths with time taken to generate 20 Gb to 15,000 Gb of output data taking between 10 to 72 hours from single molecules and thus offer sequencing output that is best used in characterizing massive structural assemblies. Single-molecule real-time sequencing (SMRT) and synthetic approaches are the two major long-read platforms in use (Figure 1.16).<sup>124</sup>

### 1.5.2.1 Single-molecule real-time sequencing (SMRT)

SMRT provided by platforms such as Pacific Biosciences (PacBio) is based on optical detection in conjunction with SBS for read lengths >1 Kb.<sup>124</sup>

### 1.5.2.2 Synthetic long-reads

Synthetic long-read sequencing identifies reads from through barcodes, unlike true sequencing platforms. DNA segments are divided into either an emulsion or wells so that few nucleic acid molecules are present in each partitioned well (Figure 1.16).<sup>117</sup>

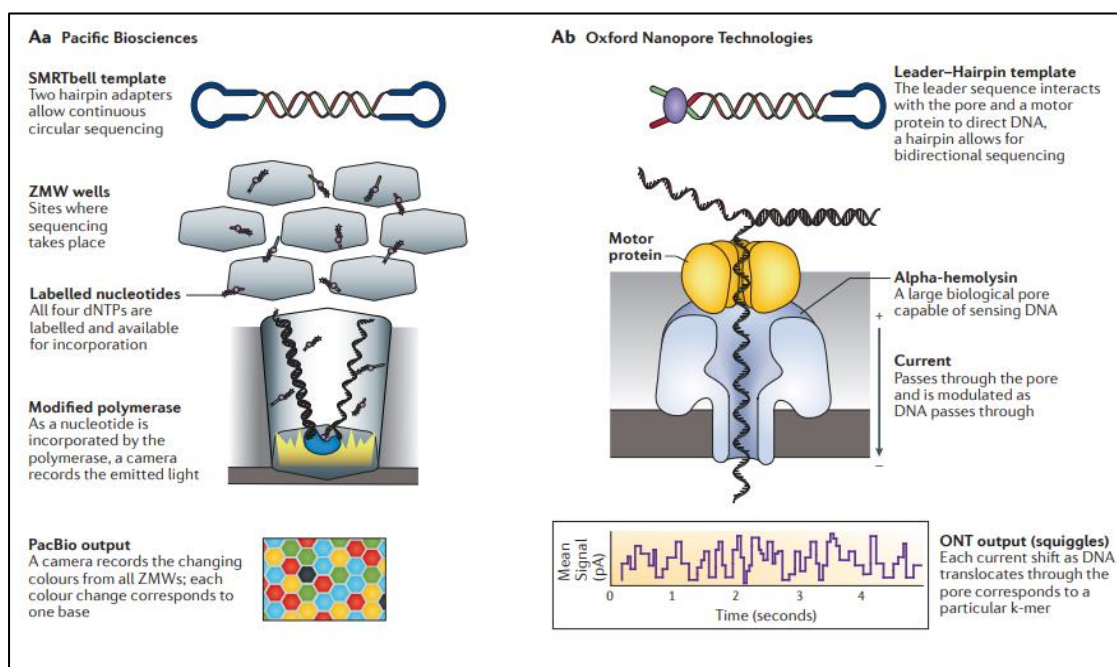


Figure 1.16: **Long-read sequencing approaches** A.) PacBio that uses ZMW wells for multiple parallel sequencing. B.) Nanopore that uses a motor protein to pass DNA through a biological protein pore and the resultant energy change is interpreted per base.<sup>117</sup>

The various NGS platforms and their application and parameters are summarised as second and third generation technologies in Table 1.6 below.

**Table 1.6 Application parameters of the sequencing platforms available**

| PLATFORM                         | MAXIMUM<br>READS<br>RANGE PER<br>RUN | MAXIMUM<br>OUTPUT<br>RANGE<br>(GIGABASES) | READ<br>LENGTH<br>RANGE<br>(BP) | SEQUENCING<br>RUN TIME<br>(HOURS) | ADVANTAGES                                                                                                                                                                                         | DISADVANTAGES                                                                                                                                          | MAIN APPLICATIONS                                                                                                        |
|----------------------------------|--------------------------------------|-------------------------------------------|---------------------------------|-----------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|
| <b>SHORT READ SEQUENCING</b>     |                                      |                                           |                                 |                                   |                                                                                                                                                                                                    |                                                                                                                                                        |                                                                                                                          |
| <b>ILLUMINA</b>                  | 4M-20B                               | 0.144-3000                                | 36-150                          | 9-44                              | <ul style="list-style-type: none"> <li>- Has an overall accuracy&gt;99.5%</li> <li>- Less prone to homopolymers errors</li> <li>- High similarity with SNPs from genotyping microarrays</li> </ul> | <ul style="list-style-type: none"> <li>- Shows some under-representation of GC and AT regions</li> <li>- Possibility of substitution errors</li> </ul> | 16S metagenomic sequencing, targeted gene expression profiling sequencing, amplicon sequencing, transcriptome sequencing |
| <b>ION TORRENT</b>               | 2M-1.8B                              | 0.3-6000                                  | 200-600                         | 2.5-213                           | <ul style="list-style-type: none"> <li>- Reduced operational and upfront costs because it doesn't use optical tools and modified bases</li> </ul>                                                  | <ul style="list-style-type: none"> <li>- Short read sequences compared to other technologies such as pyrosequencing</li> </ul>                         | Transcriptome sequencing, RNA sequencing, whole exome sequencing                                                         |
| <b>NANOBALL</b>                  | 300M-1.8B                            | 60-6000                                   | 50-150                          | 24-213                            | <ul style="list-style-type: none"> <li>- Reduces running costs</li> </ul>                                                                                                                          | <ul style="list-style-type: none"> <li>- Short sequences generated might be a challenge for read mapping</li> </ul>                                    | Deep exome sequencing, whole-genome sequencing, targeted panel sequencing                                                |
| <b>LONG READ SEQUENCING</b>      |                                      |                                           |                                 |                                   |                                                                                                                                                                                                    |                                                                                                                                                        |                                                                                                                          |
| <b>SINGLE MOLECULE REAL TIME</b> | 0.5M-375M                            | 20-15,000                                 | 10 Kb-1Mb                       | 10-72                             | <ul style="list-style-type: none"> <li>- Each read is sequenced multiple times</li> <li>- Short DNA serve as templates for circular consensus sequences</li> </ul>                                 | <ul style="list-style-type: none"> <li>- DNA templates&gt;3 Kb are a challenge to sequence multiple times</li> </ul>                                   | Epigenetic interrogation, gene expression, compound population sequencing,                                               |
| <b>SYNTHETIC LONG READS</b>      | 2B                                   | 47-72                                     | 100 Kb                          | 29                                | <ul style="list-style-type: none"> <li>-Generate reads &gt;10 Kb</li> <li>- Reduced error rate</li> </ul>                                                                                          | <ul style="list-style-type: none"> <li>- Challenges in scalability due to the large volumes of data generated</li> </ul>                               | Amplicon sequencing, methylation sequencing, long range PCR                                                              |

### 1.5.3 NGS applications

Technical and analytical improvements responsible for the advancement of NGS have improved their clinical utility and increased use as diagnostic tools,<sup>125</sup> enabling accurate, timely and cost-effective association of gene anomalies with diseases.<sup>125</sup> Currently, there are three main types of NGS sequencing of DNA used for the investigation of genomic variability namely, whole genome sequencing (WGS), targeted sequencing (TS) and whole exome sequencing (WES) with all of them having distinct cost and output (Figure 1.17).<sup>125,126</sup>

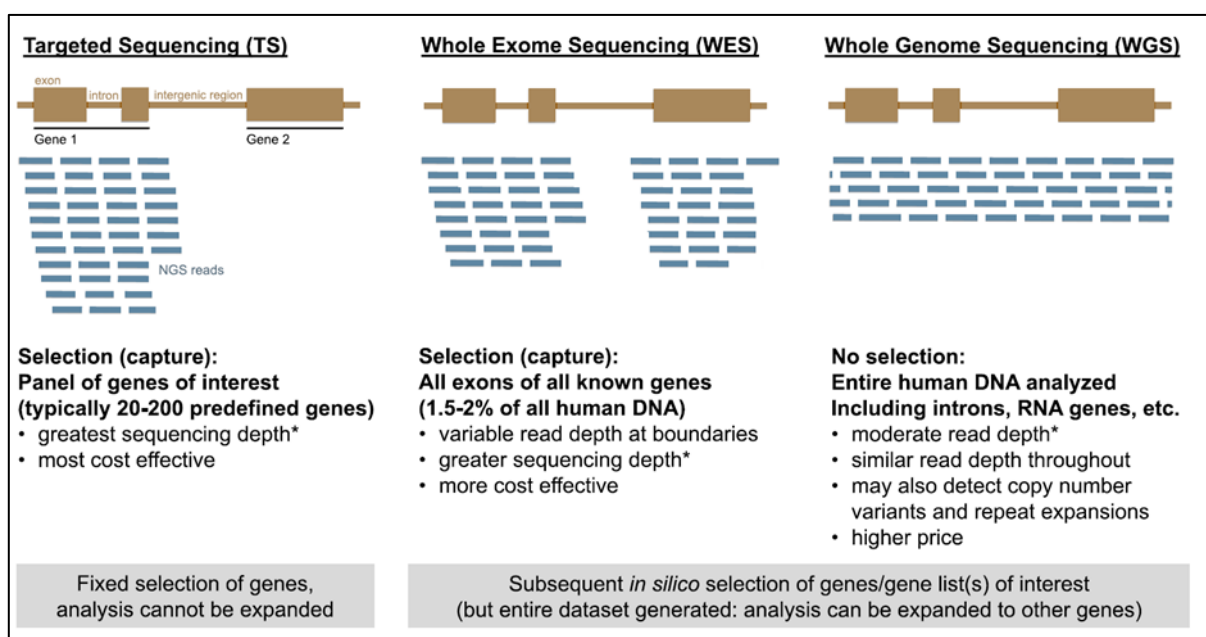


Figure 1.17: Sequencing output of the different NGS approaches for mutation identification.<sup>126</sup>

NGS techniques have been used to study the genes associated with the cardiomyopathies and thus contributing to the knowledge on the genetic architecture of these myocardial disorders. In the landscape of cardiovascular anomalies, the complex forms of the disorders are characterised by genetic variants commonly observed in the population with >100,000 individuals harbouring these variants thus conferring a differential risk of disease development based on the gene-gene and gene-environment association on the individuals with these variants.<sup>127,317</sup> The monogenic forms of cardiovascular disease are attributed to a specific gene variant and are relatively rare, being observed in less than 10 individuals.<sup>127,317</sup> NGS enables both the complex and monogenic forms of

the disease and has thus become increasingly vital in identifying and characterising genes implicated in these pathologies for improved patient care and their family members.<sup>127</sup> Table 1.7 provides an overview of the cardiomyopathy causing genes identified through the different NGS techniques.

**Table 1.7: Cardiomyopathy genes identified through NGS**

| NGS | Genes investigated                                                              | References |
|-----|---------------------------------------------------------------------------------|------------|
| TS  | <i>ACTC1, LAMP2, MYH7, GLA, PLN, TPM1, FHL1, LMNA, CSRP3, PRKAG2</i>            | 53,128     |
| WES | <i>TNNT2, TTN, DSP, RYR2, LMNA, BAG3, ACTC1, CDH2, SLC22A5</i>                  | 129        |
| WGS | <i>PLEC, BAG3, DMD, NEB, ADRA1A, SYNE1, TTN, MYH7, DSC2, LAMP2, LAMA2, RYR2</i> | 130        |

## 1.6 Functional models for cardiomyopathy

With the discovery of new genes, there is a promise of more accurate diagnostics, new therapies, and tailored regimens although this is fundamentally dependent on the comprehensive elucidation of pathogenesis at different cellular levels.<sup>131</sup>

A common approach to study the cardiomyopathies has been the use of models to recapitulate the human disease; such models have helped to characterise disease mechanisms with an end goal of developing useful therapeutics. The models used have been cellular and animal models each with their own specific utility.<sup>131</sup>

Different cell types in the heart have been studied through different experimental methods which have proved to be useful study models because of their articulation of signalling pathways with significant involvement in cardiac function and structure, abundant distribution in the heart, reproducibility of *in vivo* observations, unlimited lifespan and preservation of both protein biomarkers and an individual's genetic background.<sup>132-135</sup> Cell types used include fibroblasts, cardiomyocytes, induced pluripotent stem cells (iPCS), primary cells and immortalised cells lines.<sup>131-</sup>

<sup>134</sup> Genetic anomalies in *MYH7* and *TNNT2* have been studied in cellular models and these models

exhibited phenotypic features seen in cardiomyopathies including LVH, arrhythmia and a reduced force generation.<sup>136</sup> The use of cellular models has enabled to a clearer comprehension of the cardiomyopathies and their molecular underpinnings.

Use of animal models has increased over the years due to the evolution of new technologies which enabled researchers to gain more insight into the pathogenesis of human cardiomyopathies. The complexity of pathological modulators and disease-causing pathways has been a challenge while studying cardiomyopathy in animal models however, despite the challenges, the models have been successfully used in genotype-phenotype studies, assays that involve the development and discovery of drugs and assessing cardiac phenotypes.<sup>137</sup> Naturally occurring animal models include dogs and cats, and have been used to recapitulate human HCM phenotype caused by genetic variants in *MYBPC3*. The genetically engineered animal models include fruit flies, zebrafish, rats, mice, pigs and rabbits and have been utilised to study DCM, HCM and ACM due to the similarity in gene expression or similar phenotypic expression seen in humans.<sup>137-139</sup> These animal models have been a great platform to recapitulate patient disease and study specific therapeutic targets with the focus on personalised medicine.

## 1.7 Rationale for the study

NGS has provided a platform to have rapid analyses of genomic areas of interest thus contributing to the body of knowledge in inherited diseases. Cardiomyopathies have a genetic component to their heritability. In up to 50% of cardiomyopathies, a genetic variant is identified responsible for causing disease.<sup>50</sup> With NGS, the genetic variability observed in cardiomyopathy has been established, and the data emanating from the sequencing techniques has provided greater context into these variants. Together with validation experiments of the identified variants, diagnostic and intervention strategies on the affected individuals and cascade screening in family members can be done for better management of the disease.<sup>140</sup> The paucity of data on molecular genetic causes of cardiomyopathies in Africa presents a large gap in knowledge that requires further investigation.

## 1.8 Hypothesis

We hypothesise that NGS will identify novel genetic causes of disease in South African families with cardiomyopathy.

## 1.9 Aims and objectives

The aim of this thesis was to investigate multigenerational South African families with ACM and DCM to identify the genetic mutations responsible for their disease.<sup>56</sup>

The objectives for this project are as follows:

1. Performing WES on probands and family members diagnosed with cardiomyopathy.
2. Identifying disease-causing variants from the sequencing data of the families.
3. Validating plausible disease-causing variants through high resolution melting and Sanger sequencing.

## CHAPTER 2 - METHODS

### 2.1 Study design

This project is a sub-study (HREC REF 198/2020) that forms part of the larger principal study called, The African Cardiomyopathy and Myocarditis Registry Program (IMHOTEP) (HREC 766/2014) (Appendix E). IMHOTEP is a multi-centre, multi-national prospective cohort study that has been designed as a platform to study the genetic and clinical characteristic of cardiomyopathy and myocarditis in Africa. IMHOTEP aims to disseminate established (and gained) expertise in the fields of cardiology and genetics to other African countries lacking the capacity to support a registry of this nature, and to provide the opportunity for continued research across the continent.

The primary objectives of the principal study, IMHOTEP, is to: (1) identify the molecular genetic causes in all probands and their families, (2) to define the clinical, etiologic, electrocardiographic, autonomic, imaging and histological characteristics of cardiomyopathy and myocarditis in children and adults in Africa according to the European Society of Cardiology (ESC) criteria and (3) to describe the management of cardiomyopathy and myocarditis in children and adults in Africa.

This PhD forms part of aim 1, where the aim was to identify the disease-causing variants in South African families affected with cardiomyopathy.

Participants (probands and affected (or at-risk) family members) were recruited into the two arms of IMHOTEP: the retrospective cohort and the prospective hospital-based cohort, respectively.

The retrospective cohort included probands who have been studied at the Groote Schuur Hospital (GSH) Cardiomyopathy Clinic, the Tygerberg Hospital Paediatric Cardiology Clinic and Red Cross War Memorial Children's Hospital Paediatric Cardiology Clinic over the last 18 years (from 1996-2014). The prospective hospital-based cohort included probands who were newly diagnosed with cardiomyopathy and who were investigated and initiated on therapy at the major referral center for cardiomyopathy in the Western Cape Province and South Africa as a whole, the GSH

Cardiomyopathy Clinic. Figure 2.1 below gives overview of the comprehensive work done in this thesis starting with recruitment of participants as described above.

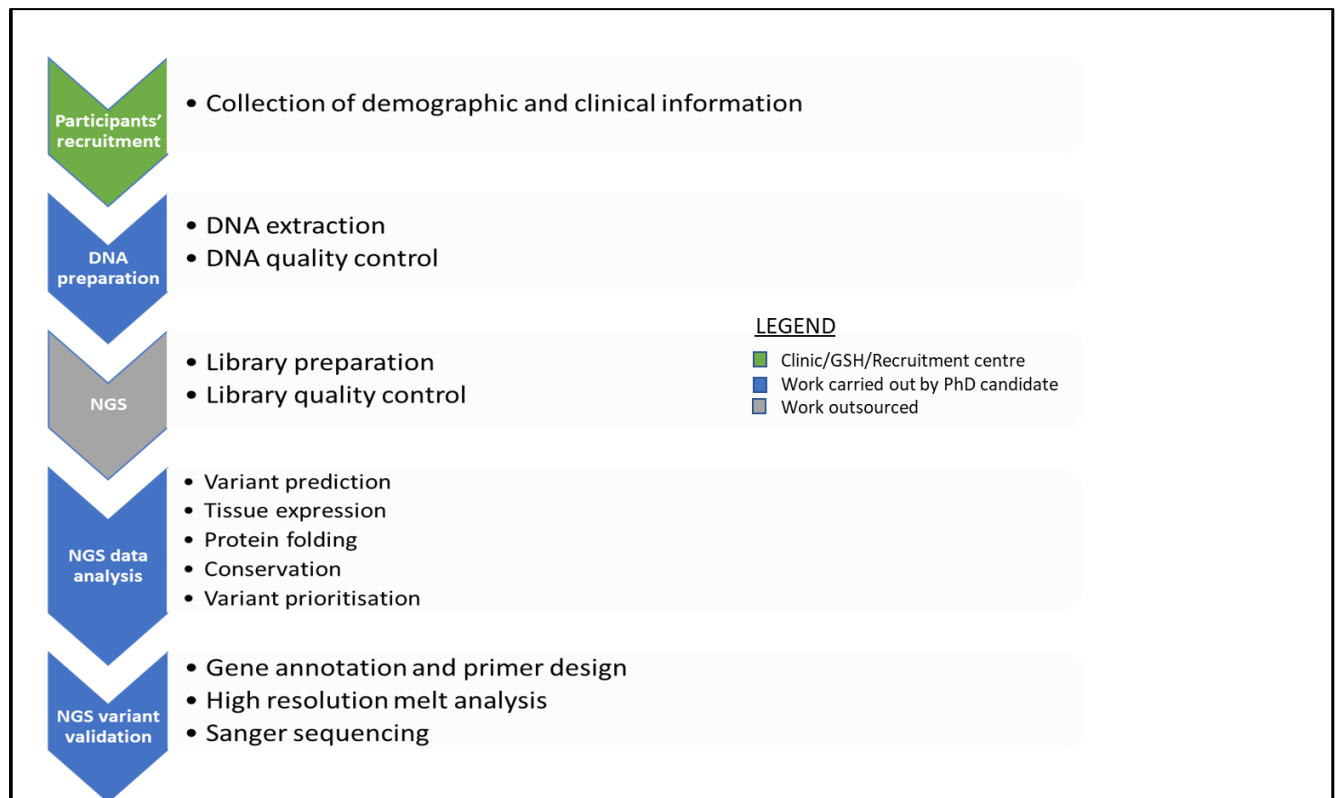


Figure 2.1: A workflow diagram illustrating the various stages of the work and where it was carried out.

## 2.2 Study participants and informed consent

Ethics approval for the principal study, IMHOTEP, was obtained from the University of Cape Town Faculty of Health Sciences Human Research Ethics Committee (HREC), UCT (HREC reference no. 766/2014) in 2014. For the three families discussed in this PhD thesis, independent HREC approval was obtained from the University of Cape Town Faculty of Health Sciences HREC REF 198/2020 (Appendix E).

Informed DNA consent (Appendix F) was obtained for all participants in the IMHOTEP study through a genetic counsellor, registered nurse or doctor. IMHOTEP does not have any age restriction as cardiomyopathy is seen in all age groups. This inclusion would enable researchers to determine how

specific genetic variants and additional environmental factors affect the disease onset, natural history and long-term outcomes of patients at a very young age. Annual follow-up of participants was done to identify hospitalisations, recurrent symptoms, use of medications and implanted devices, and therapeutic interventions.

Research participant details such as name, telephone number, address and hospital folder numbers are protected through two levels of security to ensure anonymity. Links that identify a participant are coded by letters, numbers or symbols in addition to a data matrix barcode that is used to decrypt the code for secure storage. Patient confidentiality is ensured required through HREC rules and regulations as well as the Good Clinical Practice (GCP) training received by staff and students handling the data and biological material (Appendix F)

### 2.2.1 Collection of demographic and clinical information

Information recorded for each participant included symptoms, physical examination findings, electrocardiographic findings, cardiovascular imaging findings, histology, family history, ethnicity and gender. The family histories, together with clinical information, were used to ascertain possible disease inheritance patterns and link observed disease with genetic variants, respectively. Clinical information included diagnosis, disease severity, disease status (classified as affected, unaffected, at-risk or unknown), disease onset and outcomes, including complications. The inheritance patterns were determined through analysing the pedigree, which contains a detailed, multi-generational, clinical family history of the affected family. We chose the best-fit model/s for the family, with some families having more than one possible pattern. These patterns were then carefully interrogated.

The information collected is stored in OpenClinica, which is a web-based Clinical Data Management Systems(CDMS). This open-source software is used to store and annotate patient information electronically, which is preferable to paper-based data management systems in terms of traceability and consistency. It is a cost-effective solution unlike other commercial options that incur

purchasing, deployment and regular management which can exceed the resources available to a project.<sup>141</sup> Furthermore, it is very useful in saving resources through its ability to generate and reuse case report forms that are increasingly used in the expanding clinical research field.<sup>142</sup>

## 2.2.2 Diagnostic evaluation

All probands with overt familial forms of cardiomyopathy and their available relatives were recruited into the IMHOTEP study. The participants were reviewed by a diagnostic panel of experts specialising in electrophysiology, cardiovascular imaging, cardiomyopathy and histopathology to determine their diagnostic status as per the European Society of Cardiology (ESC) guidelines on cardiomyopathy.<sup>63</sup> Echocardiography, chest X-ray, electrocardiography, cardiovascular magnetic resonance (CMR), Holter monitoring, renal tests, liver tests and full blood counts were assessed in study participants.

### 2.2.2.1 DCM

DCM was diagnosed in accordance with ESC guidelines in which DCM is defined by presence of LV dilatation and LV systolic dysfunction in absence of abnormal loading conditions (hypertension, valve disease) or congenital, pericardial, and coronary artery disease sufficient to cause global systolic impairment. RV dysfunction and dilation are not essential for the diagnosis but may be present (Table 2.1).<sup>315</sup>

**Table 2.1: Diagnostic criteria for DCM<sup>315</sup>**

| Echocardiographic parameters | Cutoff/features                                                                                                                                               | Comments                                                          |
|------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------|
| LV dilatation                | LV end-diastolic diameter >2.7 cm/m <sup>2</sup> or >117% predicted value corrected for age and body surface area<br><br>Increased sphericity index is common | Not necessary for diagnosis (e.g., mildly dilated cardiomyopathy) |
| LV systolic dysfunction      | EF < 45%                                                                                                                                                      | Impaired global contractility                                     |

|                                                      |                                                                                                         |                                                                                                                       |
|------------------------------------------------------|---------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------|
| LV wall motion abnormalities                         | Diffuse hypokinesis                                                                                     | Possible regional wall motion abnormalities mostly in LV septum and apex                                              |
| LV wall thickness                                    | Normal or only mildly increased                                                                         | Common presence of LV eccentric hypertrophy                                                                           |
| LV diastolic dysfunction                             | “Restrictive pattern” ( $E < 150$ ms and $E/A$ ratio $> 2$ ) is related to increased LV stiffness       | Useful hallmark of advanced diastolic dysfunction and elevated LV filling pressure. Can vary during follow-up         |
| LV dyssynchrony                                      | Qualitative + quantitative polyparametric evaluation                                                    | Frequent if severe LV dysfunction and LBBB; not a selection criteria for CRT                                          |
| RV dilation and dysfunction                          | TAPSE $< 14$ mm, RV FAC $< 35\%$                                                                        | Secondary of biventricular involvement and/or pulmonary hypertension                                                  |
| LA dilation                                          | End-systolic LA volume index $> 34$ ml/m <sup>2</sup>                                                   | Associated with diastolic dysfunction, MR, atrial fibrillation                                                        |
| Functional MR                                        | EROA $> 0.20$ cm <sup>2</sup> identifies a significant functional MR                                    | Contributes to increase of LV filling pressure and decrease of forward stroke volume; increases LV adverse remodeling |
| Functional TR                                        |                                                                                                         | Common in presence of RV dilation and dysfunction and pulmonary hypertension                                          |
| Dobutamine or exercise stress echocardiographic test | Assessment of presence or absence of LV inotropic response; sustained improvement vs. biphasic response |                                                                                                                       |

CRT-cardiac resynchronization therapy; EF-ejection fraction; EROA-effective regurgitant orifice area; FAC-fractional area change; LA-left atrial; MR-mitral regurgitation; TAPSE-tricuspid annular peak systolic excursion; TR-tricuspid regurgitation

#### 2.2.2.2 HCM

HCM was diagnosed on presence of increased LV wall thickness or mass in the absence of loading conditions (hypertension, valve disease) sufficient to cause the observed abnormality, according to ESC guidelines (Table 2.2).<sup>63,316</sup>

**Table 2.2: Diagnostic criteria for HCM**<sup>316</sup>

|                        | Echocardiographic parameter                                                                                                                  | Cutoff values suggesting HCM                                                                                                                                                                        |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Hypertrophy            | Wall thickness / IVS to PW ratio<br>Distribution of hypertrophy                                                                              | $> 15$ mm <sup>a</sup> , $> 1.3$ <sup>b</sup><br>Asymmetric hypertrophy<br>RV free wall hypertrophy $\geq 7$ mm <sup>c</sup><br>Reverse hypertrophic IVS                                            |
| Mitral valve apparatus | Anterior leaflet elongation<br>Posterior leaflet elongation<br>Papillary muscle abnormalities<br>Aorto-mitral angle<br>Mitral chordae<br>SAM | AML $> 30$ mm (17 mm/m <sup>2</sup> )<br>Absolute height of PL $> 15$ mm<br>Anterior displacement of AL PM<br>$< 120^\circ$<br>Elongation/thickening/buckling<br>$> 30\%$ systolic contact with IVS |

|                                 |                                                                           |                                                                                                                                                                                           |
|---------------------------------|---------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Systolic function               | Systolic longitudinal dysfunction<br><br>Normal/supranormal radial strain | Lateral S (TDI) < 4 cm/s<br>Worse GLS (> -10.6%) <sup>d</sup><br>Paradoxical apical strain (apical HCM)                                                                                   |
| Diastolic function <sup>e</sup> | Impaired relaxation<br>Elevated filling pressures                         | Lateral e' < 4 cm/s<br>Increase of A wave velocity during Valsalva manœuvre<br>LAVI > 34 mL/m <sup>2</sup> <sup>f</sup><br>Ar-A ≥ 30 ms<br>E/e' ratio > 10 <sup>g</sup><br>PAPs > 35 mmHg |
| Intraventricular obstruction    | LVOT gradient /Midventricular obstruction                                 | > 30 mmHg<br>“Dagger shaped”/“Lobster claw” Doppler envelope                                                                                                                              |

IVS-interventricular septum; PW-posterior wall; AML-anterior mitral leaflet length; PL-posterior leaflet; ALPM- anterolateral papillary muscle; SAM-systolic anterior motion; TDI-tissue Doppler imaging; Ar-duration of atrial reverse wave of the pulmonary venous flow; PAPs-systolic pulmonary artery pressure; <sup>a</sup>-Absence of abnormal loading conditions. 13 mm cutoff for HCM relatives; <sup>b</sup>-1.5 for hypertensive patients; <sup>c</sup>-Absence of abnormal loading conditions for the RV; <sup>d</sup>-Reduction in longitudinal strain is greater for hypertrophied segments; <sup>e</sup>-Diastolic dysfunction is the hallmark of the disease; filling pressures are elevated, even in the presence of an impaired relaxation pattern of the transmitral flow; <sup>f</sup>-Absence of atrial fibrillation/significant mitral regurgitation; <sup>g</sup>-Less specific in HCM as a surrogate for elevated filling pressures

### 2.2.2.3 ACM

ACM was diagnosed in presence of RV dysfunction which may be regional or global, with or without LV disease, in the presence of histological evidence for the disease and/or

electrocardiographic abnormalities in accordance with International Task Force (ITF) Criteria.<sup>151</sup>

Participants were classified as having definite ARVC (if Task Force criteria were met), probable or possible ARVC (if some criteria were met and no alternative diagnosis was found), or not having ARVC (if there was no evidence of ARVC and/or an alternative diagnosis was present) (Table 2.3).

The modified standards of Hamid and associates were used to diagnose ACM in the first-degree relatives of those who were affected.<sup>152</sup>

**Table 2.3: International Task Force criteria for ACM<sup>151</sup>**

| Major criteria                                                                                                | Minor criteria                                                                                      |
|---------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|
| <b>1. Global and/or regional dysfunction and structural alterations</b>                                       |                                                                                                     |
| By 2D echocardiogram<br>Regional RV akinesia, dyskinesia or aneurysm and one of the following (end diastole): | By 2D echocardiogram<br>Regional RV akinesia or dyskinesia and one of the following (end diastole): |
| 1. Parasternal long axis view ≥ 32 mm or ≥ 19 mm/m <sup>2</sup>                                               | 3. Parasternal long axis view 29-32 mm or 16-19 mm/m <sup>2</sup>                                   |
| 2. Parasternal short axis view ≥ 36 mm or ≥ 21 mm/m <sup>2</sup>                                              | 4. Parasternal short axis view 32-36 mm or 18-21 mm/m <sup>2</sup>                                  |
| 3. Fractional area change ≤ 33%                                                                               | 5. Fractional area change 33-40%                                                                    |

|                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>By CMR</p> <p>Regional RV akinesia or dyskinesia or dyssynchronous RV contraction and one of the following:</p> <ol style="list-style-type: none"> <li>RV end diastolic volume <math>\geq 110</math> ml/m<sup>2</sup> male or <math>\geq 100</math> ml/m<sup>2</sup> female</li> <li>RV ejection fraction <math>\leq 40\%</math></li> </ol> <p>By RV angiography</p> <p>Regional RV akinesia, dyskinesia or aneurysm</p> | <p>By CMR</p> <p>Regional RV akinesia or dyskinesia or dyssynchronous RV contraction and one of the following:</p> <ol style="list-style-type: none"> <li>RV end diastolic volume 100-110 ml/m<sup>2</sup> male or 90-100 ml/m<sup>2</sup> female</li> <li>RV ejection fraction 40-45%</li> </ol>                                                                                                                                                                                                                                                                                                                                                                                                              |
| <b>2. Tissue characterisation of wall</b>                                                                                                                                                                                                                                                                                                                                                                                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| Residual myocytes $< 60\%$ by morphometric analysis, (or $< 50\%$ if estimated), with fibrous replacement of the RV free wall myocardium in at least one sample, with or without fatty replacement of tissue on endomyocardial biopsy                                                                                                                                                                                       | Residual myocytes 60-75% by morphometric analysis, (or 50-65% if estimated), with fibrous replacement of the RV free wall myocardium in at least one sample, with or without fatty replacement of tissue on endomyocardial biopsy                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| <b>3. Repolarisation abnormalities</b>                                                                                                                                                                                                                                                                                                                                                                                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| Inverted T waves in right precordial leads (V <sub>1</sub> , V <sub>2</sub> and V <sub>3</sub> ) or beyond in individuals $> 14$ years of age (in the absence of complete RBBB QRS $\geq 120$ msec)                                                                                                                                                                                                                         | Inverted T waves in leads V <sub>1</sub> and V <sub>2</sub> in individuals $> 14$ years of age (in the absence of complete RBBB), or in V <sub>4</sub> , V <sub>5</sub> , or V <sub>6</sub> .<br>Inverted T waves in leads V <sub>1</sub> , V <sub>2</sub> , V <sub>3</sub> and V <sub>4</sub> in individuals $> 14$ years of age in the presence of complete RBBB                                                                                                                                                                                                                                                                                                                                             |
| <b>4. Depolarisation/conduction abnormalities</b>                                                                                                                                                                                                                                                                                                                                                                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| Epsilon wave (reproducible low amplitude signals between end of QRS complex to onset of the T wave) in the right precordial leads (V <sub>1</sub> to V <sub>3</sub> )                                                                                                                                                                                                                                                       | <ol style="list-style-type: none"> <li>Late potentials by signal averaged ECG in at least one of three parameters in the absence of a QRS duration of <math>\geq 110</math> msec on the standard ECG</li> <li>Filtered QRS duration: <math>\geq 114</math> ms</li> <li>Duration of terminal QRS <math>&lt; 40</math> <math>\mu</math>V (LAS): <math>\geq 38</math> ms</li> <li>Root mean square voltage of terminal 40 ms: <math>\geq 20</math> <math>\mu</math>V</li> <li>Terminal activation duration of QRS <math>\geq 55</math> ms measured from the nadir of the S wave to the end of the QRS, including R', in V<sub>1</sub>, V<sub>2</sub> or V<sub>3</sub>, in the absence of complete RBBB</li> </ol> |
| <b>5. Arrhythmias</b>                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| Non-sustained or sustained VT of left bundle branch morphology with superior axis (negative or indeterminate QRS in II, III, AVF and positive in AVL)                                                                                                                                                                                                                                                                       | <ol style="list-style-type: none"> <li>Non sustained or sustained VT of RV outflow configuration, LBBB morphology with inferior axis (positive QRS in II, III, AVF and negative in AVL) or of unknown axis</li> <li>Greater than 500 ventricular extrasystoles/24 hours by Holter</li> </ol>                                                                                                                                                                                                                                                                                                                                                                                                                   |
| <b>6. Family history</b>                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| <ol style="list-style-type: none"> <li>ACM confirmed in a first-degree relative who meets current task force criteria</li> <li>ACM confirmed pathologically at autopsy or surgery in a first-degree relative</li> <li>Identification of a pathogenic mutation categorised as associated or probably associated with ACM in the patient under evaluation</li> </ol>                                                          | <ol style="list-style-type: none"> <li>History of ACM in a first-degree relative in whom it is not possible or practical to determine if the family member meets current task force criteria</li> <li>Premature sudden death (<math>&lt; 35</math> years) due to suspected ACM in a first-degree relative</li> <li>ACM confirmed pathologically or by current Task Force Criteria in second-degree relative</li> </ol>                                                                                                                                                                                                                                                                                         |

#### *2.2.2.4 Exclusion criteria*

Probands with secondary causes of cardiomyopathy or LV dysfunction (such as hypertension, valvular heart disease, diabetes, coronary artery disease, or sepsis) were excluded from this study.

### **2.3 DNA preparation**

#### **2.3.1 DNA extraction**

Peripheral blood samples were collected from participants in 10ml ethylenediaminetetraacetic acid (EDTA) tubes. Genomic DNA (gDNA) was extracted from peripheral blood using the PureGene™ DNA isolation kit (Gentra system, USA) and was extracted according to the manufacturer's instructions (Appendix L). The gDNA samples were given disease code numbers to anonymise them and were filed in the database of the Cardiovascular Genetics Laboratory at the Cape Heart Institute, Department of Medicine, Faculty of Health Sciences, University of Cape Town. Extracted DNA samples were stored at -80°C for long term storage.

#### **2.3.2 DNA quality control**

##### *2.3.2.1 Nanodrop quantification*

All DNA samples were mixed prior to measurement and were quantified at an absorbance of 260 nm using a NanoDrop ND-2000 spectrophotometer (ThermoScientific, UK). Each sample was measured in duplicate, and the average of the two measurements was calculated. All DNA extractions were also validated using the Qubit™ Fluorometer (Invitrogen, USA), which is considered an accurate determinant of nucleic acid concentrations.

##### *2.3.2.2 Agarose Gels*

DNA samples were electrophoresed on an agarose gel to determine the quality of the DNA (Appendix H). DNA in the gel was stained with GelRed™ (Biotium, California, US), a fluorescent nucleic acid stain, for visualisation of the DNA. GelRed™ intercalates between the base pairs of

the DNA strands and is visualised under ultraviolet (UV) light. Loading of the DNA and the monitoring of electrophoretic progression were assisted with the electrophoresis buffer, 3µl of 1X loading dye with a DNA ladder (New England Biolabs<sup>R</sup> 100bp DNA ladder) (Appendix I). The DNA ladder was included to verify the sizes of the samples electrophoresed through a 1.5% agarose gel in 1X TBE buffer.

Genomic DNA was then stored at a -80°C freezer. Working stocks of 100ng/ul of DNA were stored at 4°C. DNA of affected individuals that passed the quality control step was then sent for whole exome sequencing.

## 2.4 Whole exome sequencing

Following DNA quality control, the DNA of the proband and at least one affected family member was sent for whole exome sequencing.

### 2.4.1 Principle of next generation sequencing (NGS)

The Ion Torrent sequencing technology we used for WES in this thesis relies on the principle of pH change to identify a particular nucleotide. Each time a nucleotide is added by DNA polymerase onto a strand, pH change is noted through proton emission. Double voltage is registered when two nucleotides are inserted and if a wrong nucleotide is encountered then no voltage is registered. IonTorrent does not need any camera utility or fluorescence as its utility is defined by pH shifts.<sup>143</sup>

### 2.4.2 Library preparation for whole exome sequencing

Human gDNA samples were submitted to the Centre for Proteomic and Genomic Research (CPGR) for whole exome sequencing on the ion proton sequencer. A total of 100 ng of each sample was used as template in the exome amplification PCR protocol using the Ion AmpliSeq<sup>TM</sup> Exome RDY Kit

(Life Technologies). The amplicons were treated with the FuPa Reagent (Life Technologies) to partially digest the primers and phosphorylate the 3'-ends thus ensuring efficient ligation reactions downstream. Amplification products were ligated to Ion Xpress™ Barcode Adapters (Life Technologies) and purified using Agencourt Ampure XP beads (Beckmann Coulter). Amplification products with ligated adaptors were then amplified through a limited cycle PCR and purified with Agencourt Ampure XP beads (Beckmann Coulter). The barcoded amplification products (sequencing libraries) were quantified by quantitative PCR (qPCR) using the KAPA Library Quantification kit for Ion Torrent (KAPA BioSystems) and fragment size distribution was determined on the BioAnalyzer instrument using the BioAnalyzer High Sensitivity DNA Assay Kit (Agilent).

#### 2.4.3 Library quality control

Libraries were diluted to 40 pM for templating on the Ion Chef™ instrument (Life Technologies) using the Ion PI™ Hi-Q™ Chef Kit and Ion PI™ Chip Kit v3 (Life Technologies). Sequencing was carried on the Ion Proton™ System using the Ion PI™ Hi-Q™ Sequencing 200 Kit (Life Technologies).

#### 2.4.4 Data quality control (QC)

Data quality was achieved in two steps. Firstly, the quality reports from the service provider in Oxford (United Kingdom) and Central analytical facilities (CAF) (Stellenbosch) showed high quality reads based on the AQ20 phred score quality recalibration used by the vendor equating to a 99% confidence of the base calls with a depth of at least 120 for the called bases. Secondly the validation done by HRM, which is followed by Sanger sequencing identified which variants were false positives.

#### 2.4.5 Whole exome sequencing data analysis

Sequencing reads were aligned to the *H. sapiens* reference genome sequence (UCSC version hg19) using the Torrent Suite Version 5.4.0 Software Ion Reporter™ platform (Thermo Fisher Scientific).

Targeted sequencing was carried out at the University of Oxford where they used hg19 as the reference genome. We also used hg19 for continuity when designing primers for Sanger sequencing.

Exome variant data was downloaded for annotation and filtering through variant effect predictor. Two standard variant effect prediction tools were used for variant analysis (1) American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) classification and (2) Variant Effect Predictor (VEP).<sup>144,145</sup>

##### *2.4.5.1 American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) classification*

###### *2.4.5.1.1 Assessment of variant pathogenicity*

This gene prediction tool serve to classify the variants into 5 classes and this classification was revised by expert input from the American College of Medical Genetics and Genomics (ACMG), the Association for Molecular Pathology (AMP) and the College of American Pathologists (CAP), all consisting of clinical laboratory directors and clinicians.<sup>144</sup> The 5 classes of variants in Mendelian disorders are thus assigned specific standard terminology which are:

- Class one – benign
- Class two – likely benign
- Class three -variants of unknown significance (VUS)
- Class four – likely pathogenic
- Class five - pathogenic

Functional data, population data, computational data and segregation data are incorporated as criteria to classify any particular variant into its respective class. Within the classes, numbered

weights exist which correspond to different criteria used to assign a particular weight number (Figure 2.2)

|                                   | Benign                                                                                                                 |                                                                                                                                                                                                                              | Pathogenic                                                                                                 |                                                                                                                                                                           |                                                                          |                                                                                        |
|-----------------------------------|------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------|----------------------------------------------------------------------------------------|
|                                   | Strong                                                                                                                 | Supporting                                                                                                                                                                                                                   | Supporting                                                                                                 | Moderate                                                                                                                                                                  | Strong                                                                   | Very Strong                                                                            |
| Population Data                   | MAF is too high for disorder <i>BA1/BS1</i> OR observation in controls inconsistent with disease penetrance <i>BS2</i> |                                                                                                                                                                                                                              |                                                                                                            | Absent in population databases <i>PM2</i>                                                                                                                                 | Prevalence in affecteds statistically increased over controls <i>PS4</i> |                                                                                        |
| Computational And Predictive Data |                                                                                                                        | Multiple lines of computational evidence suggest no impact on gene /gene product <i>BP4</i><br>Missense in gene where only truncating cause disease <i>BP1</i><br>Silent variant with non predicted splice impact <i>BP7</i> | Multiple lines of computational evidence support a deleterious effect on the gene /gene product <i>PP3</i> | Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before <i>PM5</i><br>Protein length changing variant <i>PM4</i> | Same amino acid change as an established pathogenic variant <i>PS1</i>   | Predicted null variant in a gene where LOF is a known mechanism of disease <i>PVS1</i> |
| Functional Data                   | Well-established functional studies show no deleterious effect <i>BS3</i>                                              |                                                                                                                                                                                                                              | Missense in gene with low rate of benign missense variants and path. missenses common <i>PP2</i>           | Mutational hot spot or well-studied functional domain without benign variation <i>PM1</i>                                                                                 | Well-established functional studies show a deleterious effect <i>PS3</i> |                                                                                        |
| Segregation Data                  | Non-segregation with disease <i>BS4</i>                                                                                |                                                                                                                                                                                                                              | Co-segregation with disease in multiple affected family members <i>PP1</i>                                 | Increased segregation data →                                                                                                                                              |                                                                          |                                                                                        |
| De novo Data                      |                                                                                                                        |                                                                                                                                                                                                                              |                                                                                                            | <i>De novo</i> (without paternity & maternity confirmed) <i>PM6</i>                                                                                                       | <i>De novo</i> (paternity & maternity confirmed) <i>PS2</i>              |                                                                                        |
| Allelic Data                      |                                                                                                                        | Observed in <i>trans</i> with a dominant variant <i>BP2</i><br>Observed in <i>cis</i> with a pathogenic variant <i>BP2</i>                                                                                                   |                                                                                                            | For recessive disorders, detected in <i>trans</i> with a pathogenic variant <i>PM3</i>                                                                                    |                                                                          |                                                                                        |
| Other Database                    |                                                                                                                        | Reputable source w/out shared data = benign <i>BP6</i>                                                                                                                                                                       | Reputable source = pathogenic <i>PP5</i>                                                                   |                                                                                                                                                                           |                                                                          |                                                                                        |
| Other Data                        |                                                                                                                        | Found in case with an alternate cause <i>BP5</i>                                                                                                                                                                             | Patient's phenotype or FH highly specific for gene <i>PP4</i>                                              |                                                                                                                                                                           |                                                                          |                                                                                        |

Figure 2.2: Criteria organization for variant classification based on type of evidence and strength of the criteria. Abbreviations: BS, benign strong; BP, benign supporting; FH, family history; LOF, loss-of-function; MAF, minor allele frequency; path., pathogenic; PM, pathogenic moderate; PP, pathogenic supporting; PS, pathogenic strong; PVS, pathogenic very strong.<sup>144</sup>

Criteria were combined according to a 5-tier system that offers scoring rules which apply to all data available on a particular variant, whether it has been previously published before or currently under investigation. An in-house laboratory database and public resources (i.e., ClinVar) was used to provide data on unpublished variants and critical flexibility was provided to the classification of variants as some criteria listed in one weight can be shifted to another weight, subject to evidence collected while using professional judgement, accounting for differences in the significance of variant evidence as highlighted in Tables 2.4-2.5.<sup>144</sup> Expert judgment was applied when evaluating the full body of information to account for differences in the strength of variant evidence.

**Table 2.4: The Criteria for classifying Pathogenic variants according to ACMG**

| Pathogenic Criteria |                |                   |                                                                                                                                                                                        | Combination of Criteria Needed to Meet Classification |   |   |   |                   |   |   |   |
|---------------------|----------------|-------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------|---|---|---|-------------------|---|---|---|
| Rule                |                | Modification Type | Rule Description                                                                                                                                                                       | Pathogenic                                            |   |   |   | Likely Pathogenic |   |   |   |
| VS                  | PVS1           | RE                | Null variant in gene with established LOF as disease mechanism                                                                                                                         |                                                       |   |   |   |                   |   |   |   |
| STRONG              | PS1            | NC                | Different nucleotide change (same amino acid) as a previously established pathogenic variant                                                                                           |                                                       |   |   |   |                   |   |   |   |
|                     | PS2            | DG                | <i>De novo</i> (paternity confirmed) in a patient with disease and no family history                                                                                                   |                                                       |   |   |   |                   |   |   |   |
|                     | PS3            | DG                | Functional studies of mammalian knock-in models supportive of a damaging effect on the gene or gene product                                                                            | 2                                                     | 1 | 1 | 1 | 1                 | 1 |   |   |
|                     | PS4            | DG                | Prevalence of the variant in affected individuals is significantly increased compared to the prevalence in controls -OR- Variant identified in ≥15 probands with consistent phenotypes |                                                       |   |   |   |                   |   |   |   |
|                     | PP1_Strong     | MS                | Variant segregates with ≥7 meioses                                                                                                                                                     |                                                       |   |   |   |                   |   |   |   |
| MODERATE            | PM1            | DG                | Hotspot/est. functional domain (amino acids 181-937) without benign variation                                                                                                          |                                                       |   |   |   |                   |   |   |   |
|                     | PM2            | DG                | Absent/extremely rare (<0.004%) from large population studies                                                                                                                          |                                                       |   |   |   |                   |   |   |   |
|                     | PM3            | RE                | Detected in trans with a pathogenic variant (recessive)                                                                                                                                |                                                       |   |   |   |                   |   |   |   |
|                     | PM4            | DG                | Protein length changes due to in-frame deletions/insertions of any size in a non-repeat region or stop-loss variants                                                                   |                                                       |   |   |   |                   |   |   |   |
|                     | PM5            | NC                | Missense change at an amino acid residue where a different missense change previously established as pathogenic                                                                        | 3                                                     | 2 | 1 | 1 |                   | 3 | 2 | 1 |
|                     | PM6            | DG                | Confirmed <i>de novo</i> without confirmation of paternity                                                                                                                             |                                                       |   |   |   |                   |   |   |   |
|                     | PVS1_Moderate  | MS                | Null variant in gene with evidence supporting LOF as disease mechanism                                                                                                                 |                                                       |   |   |   |                   |   |   |   |
|                     | PS4_Moderate   | MS                | Variant identified in ≥6 probands with consistent phenotypes                                                                                                                           |                                                       |   |   |   |                   |   |   |   |
|                     | PP1_Moderate   | MS                | Variant segregates in ≥5 meioses                                                                                                                                                       |                                                       |   |   |   |                   |   |   |   |
| SUPPORTING          | PP1            | DG                | Variant segregates in ≥3 meioses                                                                                                                                                       |                                                       |   |   |   |                   |   |   |   |
|                     | PP2            | RE                | Missense variant in a gene that has a low rate of benign missense variation and where missense variants are a common mechanism of disease                                              |                                                       |   |   |   |                   |   |   |   |
|                     | PP3            | NC                | Multiple lines of computational evidence support a deleterious effect on the gene or gene product                                                                                      |                                                       | 2 | 4 |   | 2                 |   | 2 | 4 |
|                     | PP4            | RE                | Phenotype specific for disease with single genetic etiology                                                                                                                            |                                                       |   |   |   |                   |   |   |   |
|                     | PP5            | RE                | Reputable source reports as pathogenic                                                                                                                                                 |                                                       |   |   |   |                   |   |   |   |
|                     | PS4_Supporting | MS                | Variant identified in ≥2 probands with consistent phenotypes                                                                                                                           |                                                       |   |   |   |                   |   |   |   |

Table 2.4: Summary of the adapted ACMG/AMP pathogenic criteria. VS = Very Strong. DG: Disease/gene modification; RE: Removed - not applicable to MYH7- associated disease; MS: Modified rule strength; NC: No change. Numbers under each classification refer to the number of rules with that strength required to classify the variant as its header category. Example: A likely pathogenic classification may be made with 1 piece of strong and 2 pieces of supporting evidence.<sup>144</sup>

**Table 2.5: The Criteria for classifying Benign variants according to ACMG**

| Benign Criteria |     |                   |                                                                                                                                                                                                                      | Combination of Criteria Needed to Meet Classification |   |               |   |
|-----------------|-----|-------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------|---|---------------|---|
| Rule            |     | Modification Type | Rule Description                                                                                                                                                                                                     | Benign                                                |   | Likely Benign |   |
| SA              | BA1 | DG                | Allele frequency is $\geq 0.1\%$ based on the filtering allele frequency (FAF) in ExAC                                                                                                                               | 1                                                     |   |               |   |
| STRONG          | BS1 | DG                | Allele frequency is $\geq 0.02\%$ based on the filtering allele frequency (FAF) in ExAC provided there is no conflicting information                                                                                 |                                                       | 2 | 1             |   |
|                 | BS2 | RE                | Observed in healthy adult with full penetrance expected at an early age                                                                                                                                              |                                                       |   |               |   |
|                 | BS3 | NC                | Functional studies of mammalian knock-in models supportive of no damaging effect on protein function or splicing                                                                                                     |                                                       |   |               |   |
|                 | BS4 | DG                | Non-segregation in affected members of a family                                                                                                                                                                      |                                                       |   |               |   |
| SUPPORTING      | BP1 | RE                | Missense variant in gene where only LOF causes disease                                                                                                                                                               |                                                       |   |               | 2 |
|                 | BP2 | DG                | Observed as comp het (in trans) or double het in genes with overlapping function (e.g. sarcomere genes) without increased disease severity -OR- Observed in cis with a pathogenic variant in any inheritance pattern |                                                       |   |               |   |
|                 | BP3 | RE                | In-frame deletions/insertions in a repetitive region without a known function                                                                                                                                        |                                                       |   |               |   |
|                 | BP4 | NC                | Multiple lines of computational evidence suggest no impact on gene or gene product                                                                                                                                   |                                                       |   |               |   |
|                 | BP5 | DG                | Variant found in a case with an alternate molecular basis for disease                                                                                                                                                |                                                       |   |               |   |
|                 | BP6 | RE                | Reputable source reports as benign                                                                                                                                                                                   |                                                       |   |               |   |
|                 | BP7 | NC                | A silent variant for which splicing prediction algorithms predict no impact to the splice consensus sequence nor the creation of a new splice site -AND- the nucleotide is not highly conserved                      |                                                       |   |               |   |

Table 2.5: Summary of the adapted ACMG/AMP benign criteria. SA = Stand alone. DG: Disease/gene modification; RE: Removed - not applicable to MYH7-associated disease; MS: Modified rule strength; NC: No change. Numbers under each classification refer to the number of rules with that strength required to classify the variant as its header category. Example: A likely benign classification may be made with 1 piece of strong or 2 pieces of supporting evidence.<sup>144</sup>

#### 2.4.5.2 Variant effect predictor (VEP)

Data analysis skills training for the ENSEMBL variant effect predictor (VEP) occurred at the University of Manchester in the United Kingdom under the guidance of Dr Simon Williams. The PhD candidate was trained in writing codes, using R package as well as VEP.

All VEP training was done in MobaXterm (<https://mobaxterm.mobatek.net>), an enhanced utility in Windows which brings all the essential Unix commands to a Windows environment, enabling access to a remote machine's command line, has a tabbed interface for easy access of all active sessions, and offers support to many communication protocols that enable two computers to communicate.

The various stages of VEP are summarised in Figure 2.3 VEP's versatile utility enables it to be used to interrogate genomic data from single samples, trio analysis as well as thousands of samples simultaneously.

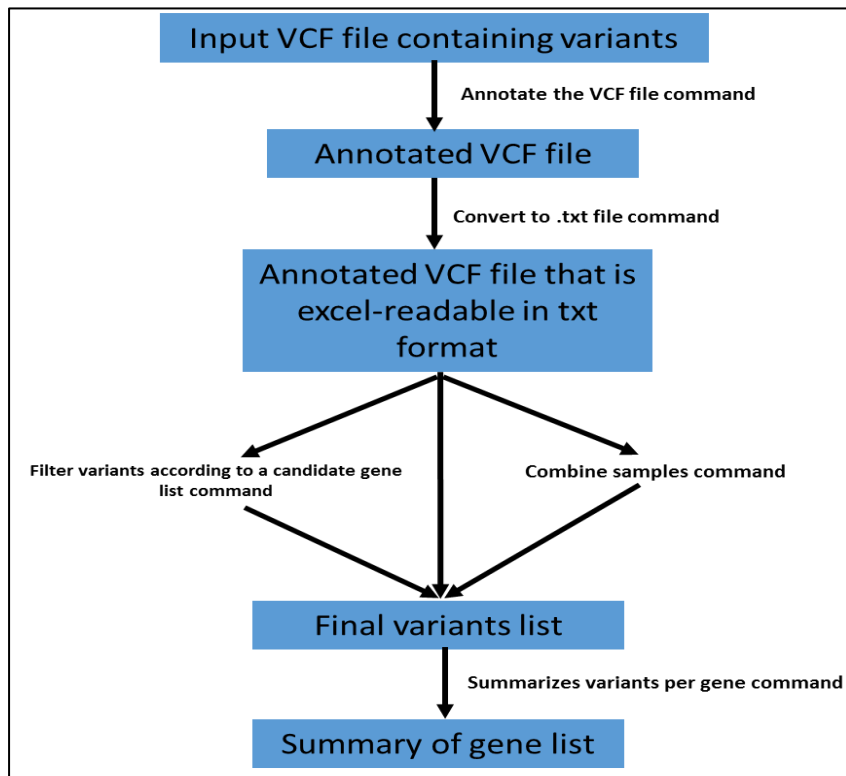


Figure 2.3: Stepwise use of VEP to analyze NGS data. The commands written in the script for the pipeline are highlighted in grey while the input and output after the commands are highlighted in green

Through the supercomputer at the High Performance Computing Facility at the University of Cape Town, we applied the downloadable VEP script for variant analysis. VCF files from the sequencing facility were loaded onto VEP where annotation of the variants from the WES was done. Annotation was proceeded by a combination of all the variants amongst all the individual files into one file in text (txt) format for ease of analysis.

#### 2.4.5.3 Variant filtering and assessment of pathogenicity

A list of 189 known cardiomyopathy-associated genes (Appendix J), generated from the NGS panel (n=48 genes (GENETests™ ©2016)), PAN cardiomyopathy panel (n=103 genes (Blueprint Genetics © 2016)), BGL cardiomyopathy panel (n=120 genes) and elife (n=159 genes), was used for preliminary filtration of WES data.

Variants were filtered against a cardiac gene list (Appendix J). From the list of the 189 genes, we sought those variants shared amongst multiple affected individuals. Additional filters applied were: (1) variants that had no minor allele frequency (MAF) or  $MAF \leq 1\%$  as determined by gnomAD (2)

variants found in the exons or within the margins of exons and introns (3) variants that are nonsense, stop/loss, missense and insertions and deletions (4) variants that were heterozygous in nature because of the AD inheritance (5) the pathogenicity prediction of the variants which was determined using the gene prediction tools: Pathogenicity of variants was determined using the gene prediction tools Varsome<sup>295</sup> (<https://varsome.com>), MutationTaster<sup>296</sup>, SIFT<sup>297</sup> and PolyPhen-2<sup>298</sup> and M-CAP<sup>299</sup>. All synonymous variants were excluded from the primary analysis. All nonsynonymous variants were screened with dbSNP<sup>300</sup>, the 1000 Genomes Project,<sup>301</sup> The Exome Aggregation Consortium (ExAC) database<sup>302</sup> and Ensembl<sup>303</sup>. All variants were validated using high resolution melt analysis (HRM) and Sanger sequencing. Novel variants were defined as those not present in any of the databases mentioned above.

The pathogenicity of a variant was determined by various factors which included the ACMG classification (in which class 5 represents pathogenic variants and 4 represents likely pathogenic variants), frequency in the population, amino acid consequence on protein structure, residue conservation, segregation with disease in affected family members and if there were any supporting literature (functional studies) strengthening the case for this variant.

Variants were determined to be modifiers if they were in ACMG class 3 variant classification and were expressed in the heart. Furthermore, the modifier variants had to show some association in causing a myocardial phenotype based on thorough literature search and affects disease severity caused by a primary disease-causing variant.

## 2.4.6 Bioinformatics tools

### 2.4.6.1 Variant prediction

All the *in-silico* tools, both open-source and commercially available, provide interpretation of variants from sequence data and include determination of the effect of the sequence variant at the nucleotide and amino acid level, determination of the effect of the variant on the primary and

alternative gene transcripts, and probable impact of the variant on resultant protein and other genomic elements. Similarity in core function thus makes the combination of predictions from different *in-silico* tools vital as a comprehensive and single piece of evidence for interpretation of sequence variants, contrary to what independent pieces of evidence offer. In addition, each of these tools have their individual strengths and weaknesses. Despite the comprehensive information provided by these prediction tools, they still remain to be predictions therefore their use in interpreting sequence variants should not be used solely to make a clinical decision.

#### *2.4.6.2 Tissue expression*

Genes were searched in the PubMed (<https://www.ncbi.nlm.nih.gov/pubmed/>),<sup>304</sup> Genotype-Tissue Expression (GTEx)<sup>305</sup> and Mouse Genome Informatics (MGI)<sup>306</sup> databases to evaluate any links to the heart and its function. PubMed was systematically used to find literature linking the gene of interest in either the development of normal heart function or cardiomyopathy by using the Medical Subject Heading (MeSH) terms (((((((("BAG3 protein, human" [Supplementary Concept]) OR "BAG3 protein, animal" [Supplementary Concept]) OR "DMPK protein, human" [Supplementary Concept]) OR "DMPK protein, animal" [Supplementary Concept]) OR "PKP2 protein, human" [Supplementary Concept]) OR "PKP2 protein, animal" [Supplementary Concept]) AND "Gene Expression"[Mesh]) AND "Cells"[Mesh]. GTEx was used to determine the gene of interest's expression in the heart and any genes with no recorded expression in the heart (LV) were excluded. Genes that were confirmed to have been functionally investigated in the MGI database of mouse models and showed cardiovascular defects were prioritised. Genes of interest were identified based on their function, expression in the heart and the results of prior animal studies.

#### *2.4.6.3 Protein folding*

Image visualisation of the protein structural effect caused by any identified variants was done using Phyre2 and Chimera programs. Site directed mutator (SDM)<sup>149</sup> was used to predict the effect of variants of interest on protein function and stability. Phyre2 is an open-source suite of tools online whose focus is to provide users with an intuitive and simple interface to state-of-the-art protein bioinformatic tools.<sup>147</sup> Amongst some of the widely used web servers for protein modelling including i-TASSER8, Robetta12 and HHpred10, Phyre2 differentiates itself through its ease of use; it is reported that there is no significant difference in the accuracy of the methods.<sup>147</sup>

UCSF Chimera is an open-source program for the interactive analysis and visualisation of molecular structures alongside sequence alignments and density maps. It is divided into extensions and core with the extensions being responsible for higher level functionalities that includes fulfilling the demands of outside developers who choose to incorporate new features.<sup>307</sup> The core is used for molecular graphics capabilities and basic services.<sup>148</sup> Chimera can read and write sequence alignments in several popular formats including Clustal ALN and “aligned” FASTA. The different Chimera extensions enable the visualisation of ligand orientations, trajectories of molecular dynamics and display of 3D grid data reading from several file formats.<sup>148</sup>

Site directed mutator (SDM) is used to inform on the protein instability and function caused by mutations that affect the stabilisation interaction of a folded protein with effects of protein destabilisation being responsible for up to 80% of Mendelian diseases due to single mutations in protein coding regions.<sup>149</sup> SDM predicts the effects of mutations by using environment specific frequencies of amino acid substitution within homologous protein families to create free energy difference between wild type sequence and mutant protein and given as a calculated stability score and disease association.<sup>149</sup>

#### 2.4.6.4 Conservation

For conservation analysis, mutated amino acids were checked in chimpanzee (*P. troglodites*), macaque (*M. mulatta*), mouse (*M. musculus*), cat (*F. catus*), chicken (*G. gallus*), zebrafish (*D. rerio*), and frog (*X. tropicalis*) by using the ClustalW multiple sequence alignment tool in BioEdit v 7.2<sup>308</sup> to compare the protein sequence of each gene to the *H. sapiens* reference sequence. Sequence data for other species is available for conservation comparison although inclusion of additional species does not have significant accuracy on prediction. Evolutionary conservation status is classified as either “*all identical*” (i.e., the same amino acid in the homologue and human amino acid sequence), “*partly conserved*” (i.e., similar amino acids in the homologue and human amino acid sequence) or “*not conserved*” (i.e., different amino acids in the homologue and human amino acid sequence). When no alignment could be made or no homologue gene identified, this is also stated.<sup>296</sup>

#### 2.4.6.5 Variant prioritisation

Variants not occurring in the cardiomyopathy panel were also considered for investigation amongst the affected individuals to ascertain their impact on the observed phenotype. The significance of this is seen when interrogating for modifiers in NGS data or when no pathogenic mutation is found in the cardiomyopathy panel. Despite their absence in the cardiomyopathy panel, the variants for prioritisation must occur in genes that cause a similar cardiomyopathy phenotype, must be expressed in the heart, be rare, track with the disease in familial studies and fall within class 3-5 variants according to the ACMG classification. The bioinformatic tools used for gene prioritisation were ToppGene,<sup>309</sup> Varelect<sup>310</sup> and Endeavour (<https://endeavour.esat.kuleuven.be/>).<sup>311</sup>

#### 2.4.7 Whole exome sequencing variant validation

The validation process involves designing the correct primers for the variants of interest, performing PCR-based HRM which is then followed by Sanger sequencing. HRM is used for mutation detection

while Sanger sequencing is used as a confirmatory method to validate the variants identified after data analysis. Sanger sequencing thus confirms that any variants identified are true variants and thus eliminate any false positive variants.

#### *2.4.7.1 Gene annotation and primer design*

Primers were designed to amplify and validate all variants of interest identified through WES. Genes were annotated using the AnnotV9 annotation programme. Primer sequences were chosen manually by referring to the annotated variants gene sequence and analysing the chosen sequences using IDT OligoAnalyze<sup>312</sup>(<http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/Default.aspx>) and NCBI BLAST<sup>314</sup> web-based tools for the determination of the optimal primer sequences for the gene amplification.

The primers were required to have:-

- (1) A length of between 18 and 25 base pairs
- (2) Good sequence diversity with minimal sequence repeats and secondary structure
- (3) A melting temperature of between 50°C and 60°C
- (4) A GC content of between 45% and 65%, and
- (5) A G or C clamp at the 3' end.

An increase in amplicon size in HRM leads to a subsequent decrease in resolution of both the amplicon being screened and the mutation to be detected. Because of this, the amplicon sequence lengths were restricted to 250-300 bp in length. Using these criteria, primer sets were designed for the combined PCR and HRM amplification and analysis of the gene exons and synthesised by Integrated DNA Technologies (IDT) (IDT, USA) (Table 2.6).

#### *2.4.7.2 Primers and parameters for DCM334*

The primers that were used to carry out HRM analysis of the candidate variants in DCM334 from preliminary analysis following DCM334 WES data analysis are as shown in Table 2.6 below.

**Table 2.6: Primer pairs used for HRM analysis of candidate variants in family DCM334**

| PRIMER       | SEQUENCE (5' – 3')             | SIZE (BP) | TM (°C)°C | PRODUCT SIZE (BP) | VARIANT             |
|--------------|--------------------------------|-----------|-----------|-------------------|---------------------|
| DMPK Ex 8_F  | GTC ACT TTC ATA TCA GTC ACT G  | 22        | 50.9      | 229               | c.1067C>T           |
| DMPK R       | CTT CAG GTC ACG GGA AAC CAT C  | 22        | 57.9      |                   |                     |
| TTN Ex 274_F | CAA CCC CAA CAT CTT CCA TGC    | 21        | 56.7      | 271               | c.64207C>T          |
| TTN R        | GTC CAC AAA AGT CGG CTG GTT TC | 23        | 58.8      |                   |                     |
| SYNE Ex 19_F | GTC ACT TTC ATA TCA GTC ACT G  | 22        | 50.9      | 176               | c.1964A>G           |
| SYNE R       | CTT CAG GTC ACG GGA AAC CAT C  | 22        | 57.9      |                   |                     |
| DMPK F       | GTC ACT TTC ATA TCA GTC ACT G  | 22        | 50.9      | 176               | c.1477C>T           |
| DMPK R       | CTT CAG GTC ACG GGA AAC CAT C  | 22        | 57.9      |                   |                     |
| BAG3 Ex 4_F  | GGG GTG ATC AAT GGA AGC CTG AC | 23        | 60        | 298               | c.925C>T            |
| BAG3 Ex 4_R  | CAA CTG GGC CTG GCT TAC TTT C  | 22        | 58.4      |                   |                     |
| CFTR Ex 13_F | AAA AGT GCT ACT TCT GCA CCA C  | 22        | 55.7      | 319               | c.1727G>C           |
| CFTR Ex 13_R | ATG ATG GGA CAG TCT GTC TTT C  | 22        | 54.4      |                   |                     |
| CFTR Ex 14_F | CTGTGT CTGTAACTG ATGGC         | 21        | 53.6      | 319               | c.2002C>T           |
| CFTR Ex 14_R | TT CCC CAA ACT CTC CAG TCT G   | 21        | 56.5      |                   |                     |
| PKD1 Ex 15_F | TGCGTCA CCTACCAGAC TGAG        | 21        | 58.5      | 357               | c.6749C>T           |
| PKD1 Ex 15_R | CAG GTT GGG GTC GTA GGA CTC    | 21        | 59.2      |                   |                     |
| LRP1 Ex 62_F | TCTC CTATCTCC AG CTTCAGCTCC    | 24        | 59.6      | 409               | c.9927_9928delCG    |
| LRP1 Ex 62_R | TGG GAA AAC CAG ATG CAT GAG G  | 22        | 57.6      |                   |                     |
| PKP2 Ex 11_F | CTC AGA CAG TTG TCC AGA AG     | 20        | 52.6      | 298               | c.9927_9928delGGTGT |
| PKP2 Ex 11_R | CGG GAG GTG ATA CAG ACA AC     | 20        | 50        |                   |                     |

### 2.4.7.3 Primers and parameters for DCM343

**Table 2.7 Standard PCR cycling parameters**

| CONDITION                     | TEMPERATURE (TIME)            |
|-------------------------------|-------------------------------|
| <b>Initial denaturation</b>   | 95°C - 10 minutes             |
| <b>Denaturation</b>           | 95°C – 10 seconds             |
| <b>Primer Annealing</b>       | 55°C - 10 seconds   50 cycles |
| <b>Temperature Elongation</b> | 72°C - 10 seconds             |

### 2.4.7.4 Primers and parameters for ACM145

**Table 2.8: Standard PCR cycling parameters**

| CONDITION                     | TEMPERATURE (TIME)            |
|-------------------------------|-------------------------------|
| <b>Initial denaturation</b>   | 95°C - 10 minutes             |
| <b>Denaturation</b>           | 95°C – 10 seconds             |
| <b>Primer Annealing</b>       | 55°C - 10 seconds   50 cycles |
| <b>Temperature Elongation</b> | 72°C - 10 seconds             |

## 2.4.8 High resolution melt analysis

Detection of mutation by HRM analysis is a two-step process involving amplification of the target DNA sequence using PCR and analysis of the produced amplicon by HRM. With the Corbett RotorGene 6000 (Qiagen, Germany), these two steps can be combined into a single procedure

where DNA samples are amplified and then immediately subjected to HRM. DNA fragments of interest are amplified by means of the PCR prior to HRM in the presence of a third generation intercalating dye called Evagreen™ (Biotium, USA).

During HRM, the PCR amplicon is exposed to increasing temperature. The fluorescence of third-generation intercalating dyes (e.g., EvaGreen™) is measured continuously and is plotted against increasing temperature. At lower temperatures, all DNA is double-stranded thus fluorescence is high. The DNA will start to disassociate into two single strands as the temperature increases, resulting in DNA melting. At this point, the dye is released and the fluorescence will decrease. The melting temperature of the DNA sample under analysis is the point of the melt phase where the rate of change in fluorescence is greatest. When the DNA is completely melted, only some background fluorescence will be detected.

DNA strands of a PCR product are bound together by hydrogen bonds and additional interactions such as base stacking forces. Depending on the strength of these interactions and different parameters surrounding a particular PCR amplicon such as the length of the amplicon, the overall base composition, and the local GC content within the PCR amplicon, each amplicon will have a differential melting behaviour thus deliver a characteristic melting profile (Figure 2.4).

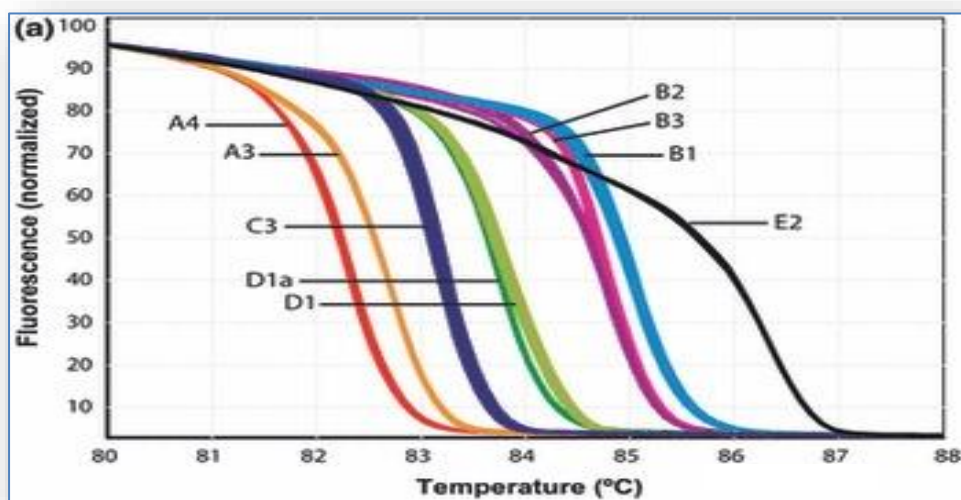


Figure 2.4: Melt curve profiles of different PCR amplicons each with differential melt curves.

#### 2.4.8.1 HRM protocol

Tables 2.9 and 2.10 show the reagents and reaction conditions respectively applied for the RT-PCR and HRM of the variant amplicons. Scientific Specialities Inc (USA) kits were used for HRM analysis.

**Table 2.9: Protocol for PCR and HRM of the exons**

| REAGENT (STOCK CONCENTRATION) | VOLUME (μl) |
|-------------------------------|-------------|
| Forward primer (20μM)         | 0.5         |
| Reverse primer (20μM)         | 0.5         |
| dNTPS (20uM)                  | 1           |
| GoTaq Flexi buffer (5X)       | 5           |
| GoTaq Polymerase (5U/μl)      | 0.1         |
| MgCl <sub>2</sub> (25mM)      | 3           |
| Evagreen (20x)                | 0.5         |
| DNA (100ng/ul)                | 1           |
| dH <sub>2</sub> O             | 13.4        |
| Total                         | 25          |

**Table 2.10: Optimised temperature cycling conditions for the PCR and HRM of the variant exons**

| CONDITION              | TEMPERATURE (TIME)            |
|------------------------|-------------------------------|
| Initial denaturation   | 95°C - 10 minutes             |
| Denaturation           | 95°C – 5 seconds              |
| Primer Annealing       | 55°C - 10 seconds   50 cycles |
| Temperature Elongation | 72°C - 10 seconds             |
| High Resolution Melt   | 72-95°C<br>(0.1°C increments) |

#### 2.4.8.2 Purification of HRM products

After HRM analysis, the samples identified for sequence analysis based on a differing melt profile on the HRM were purified using the Thermosensitive Alkaline Phosphatase (Applied Biosystems, USA) and Exonuclease I (Applied Biosystems, USA) enzymes. The reagents used are shown in Table 2.24. Reactions were incubated at 37°C for 1 hour, after which a 75°C deactivation step was conducted for 15 mins in a MultiGene Gradient thermal cycler (Labnet International, USA).

**Table 2.11: Cleanup protocol for variant HRM products**

| REAGENT                                             | VOLUME (μl) |
|-----------------------------------------------------|-------------|
| <i>Exonuclease I</i> (20U/μL)                       | 0.1         |
| <i>Thermosensitive Alkaline Phosphatase</i> (1U/uL) | 1           |
| Distilled/deionized water                           | 13.9        |
| HRM product                                         | 5           |
| Final reaction volume                               | 20          |

## 2.4.9 Sanger sequencing

Sanger sequencing was performed to validate all WES variants of interest. Sanger sequencing was performed on all purified HRM sample products.

### 2.4.9.1 Sequencing parameters

Sequencing of the HRM products was performed using the BigDye® Terminator v3.1 Sequencing Kit (Applied Biosystems, USA). The reagents used for the direct sequencing reaction of the variants are described in Table 2.12. The temperature cycling conditions were conducted in a MultiGene Gradient thermal cycler (Labnet International, USA) and the sequencing reactions are as described in Table 2.13.

**Table 2.12: PCR protocol for the sequencing reactions of variant amplicons**

| REAGENT                                     | VOLUME (μl) |
|---------------------------------------------|-------------|
| Forward Primer (1 μM)                       | 2           |
| BigDye® Terminator v3.1. Ready Reaction Mix | 2           |
| 5 X Sequencing Buffer                       | 4           |
| Distilled/deionized water                   | 9 μl        |
| HRM product                                 | 3 μl        |
| FINAL REACTION VOLUME                       | 20 μl       |

**Table 2.13: Optimised cycling conditions for the sequencing reactions of variant amplicons**

| CONDITION                                               | TEMPERATURE (TIME)                                                                    |
|---------------------------------------------------------|---------------------------------------------------------------------------------------|
| Initial denaturation                                    | 96°C - 5 minutes                                                                      |
| Denaturation<br>Primer Annealing<br>Template Elongation | <div> 96°C - 30 seconds<br/> 50°C - 15 seconds<br/> 60°C - 4 minutes </div> 25 cycles |

#### 2.4.9.2 Capillary electrophoresis

Analysis of the sequencing products was done using capillary electrophoresis at the DNA sequencing unit (Department of Genetics, Stellenbosch University), by means of an ABI PRISM® 3130xl Genetic Analyser (Applied Biosystems, USA) which is a fully automated system with a 16-capillary throughput that offers reduced sample usage and higher resolution of sequencing data.

#### 2.4.9.3 Analysis of Sanger sequencing products

BioEdit software v 7.2 was used to visualise the electropherograms generated from sequencing. Reference sequence downloaded from the Ensembl database (Table 2.14) was aligned to the visualised sequences using the ClustalW multiple sequence alignment tool which accompanies the BioEdit software and is used to align DNA or protein multiple sequences. To confirm the observed variant from exome sequence analysis, the Sanger sequencing alignment was investigated at the position of interest.

**Table 2.14 Database numbers of genes investigated**

| Gene          | Ensembl gene ID   | Ensembl transcript ID | NCBI RefSeq ID |
|---------------|-------------------|-----------------------|----------------|
| <b>DMPK</b>   | ENSG00000104936   | ENST00000343373       | NM_004409      |
| <b>TTN</b>    | ENSG00000155657   | ENST00000589042       | NM_001267550   |
| <b>SYNE1</b>  | ENSG00000131018   | ENST00000367255       | NM_182961      |
| <b>OBSCN</b>  | ENSP00000455507   | ENST00000570156       | NM_001386125   |
| <b>IQGAP2</b> | ENSP00000274364   | ENST00000274364       | NM_006633      |
| <b>BAG3</b>   | ENSG00000151929   | ENST00000369085       | NM_004281      |
| <b>RYS1</b>   | ENSG00000196218   | ENST00000359596       | NM_000540      |
| <b>CRK</b>    | ENSG00000167193   | ENST00000300574       | NM_016823      |
| <b>NEB</b>    | ENSG00000183091   | ENST00000397345       | NM_001164508   |
| <b>CFTR</b>   | ENSG00000001626   | ENST00000003084       | NM_000492      |
| <b>PKD1</b>   | ENSG00000008710   | ENST00000262304       | NM_001009944   |
| <b>LRP1</b>   | ENSG00000123384   | ENST00000243077       | NM_002332      |
| <b>PKP2</b>   | ENSG00000057294.9 | ENST00000070846       | NM_004572      |
| <b>ALPK3</b>  | ENSG00000136383   | ENST00000258888       | NM_020778      |

## 2.5 Targeted resequencing

WES was our primary NGS method for the work in this thesis while targeted resequencing was used as the NGS technique for IMHOTEP, which is the parent study. Through IMHOTEP a cohort of 634 cardiomyopathy probands (DCM=479, HCM=69, ACM=78, RCM=39)(unpublished data) were screened for 38 genes (Appendix K). We thus had WES and targeted sequencing data for the proband and the nuclear family. We only extracted and analysed the sequencing data for this proband as the other probands did not form part of this PhD study.

We used the customised Haloplex Target Enrichment System Reagent Kit for targeted analysis of a panel of 38 genes associated with cardiomyopathy and arrhythmias (Appendix K). The HaloPlex Target Enrichment System amplifies thousands of targets in the same reaction, incorporating standard Illumina paired- end sequencing motifs in the process. We achieved median coverage of 400× coverage of the targeted regions with the HaloPlex gene panel across this sample, compared to median coverage of 70× for these same target regions by WES. Variants detected in samples that passed quality control were filtered for likely pathogenicity as described in section 2.5 above. The variants were called using software Platypus v 0.8.1 and GATK v 3.8. The thresholds were set at 400 (as determined by the University of Oxford diagnostic laboratory) and scores above this threshold were considered true positives. Prioritised variants validated in this study are listed in section 5.5.4.

### 2.5.1 Principle of targeted enrichment

Targeted enrichment relies on the principle of single-stranded and sequence-specific oligonucleotides known as probes hybridized complementarily to the regions of interest which are then isolated to capture the sequence at the targeted region. The hybridization is highly stable and specific when bound to their genomic region of interest thus ensuring that the targeted sequences are robustly represented in the sequenced DNA. This serves to increase the efficiency of their selective targeting and minimizing off-targets from the expansive background of the entire genome.

## 2.5.2 HaloPlex Target Enrichment Workflow

Figure 2.5 below gives an overview of the HaloPlex procedure workflow with indexed libraries made for each sample identified for sequencing. Depending on the specific sequencing platform used, up to 96 samples can be pooled and sequenced in a single lane.

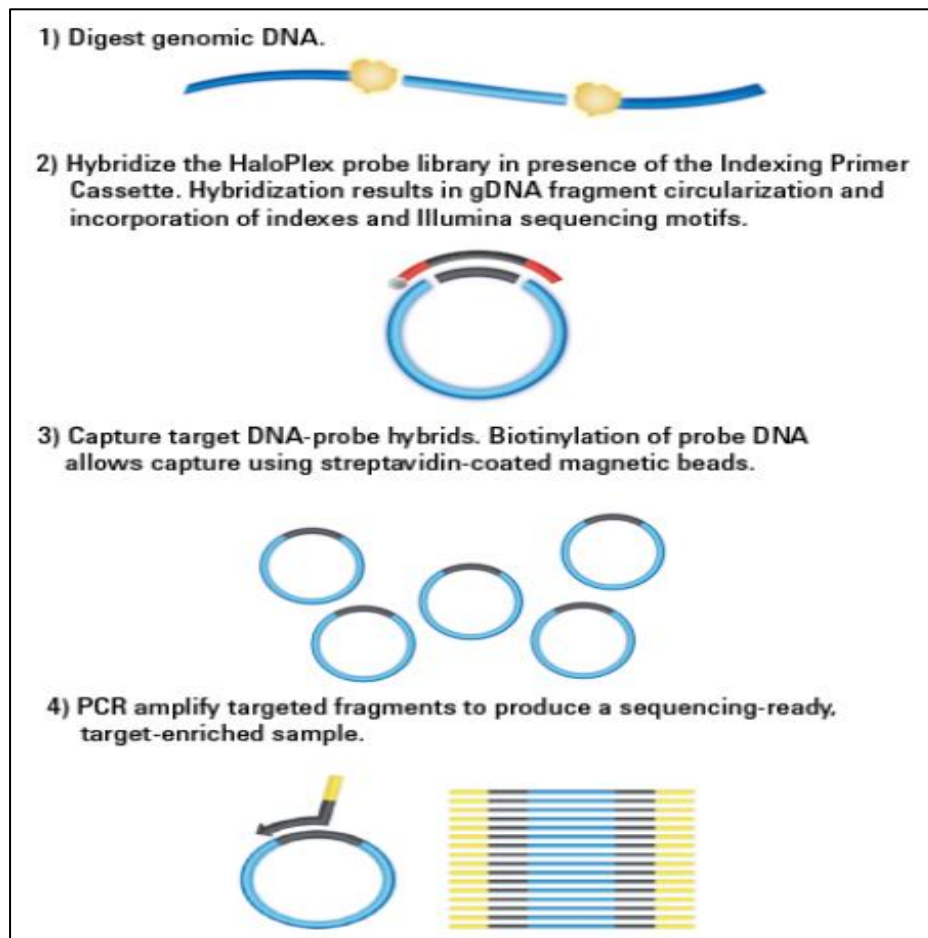


Figure 2.5: A stepwise illustration of the HaloPlex enrichment procedure which uses hybridized probes for DNA sequencing.

## 2.5.3 Designing Haloplex probes

Specific probes which can also be customized to a particular disease were designed from Agilent's SureDesign tool at [www.agilent.com/genomics/suredesign](http://www.agilent.com/genomics/suredesign).

## 2.5.4 Sample Preparation

Restriction enzymes were used to digest DNA which was then hybridized to HaloPlex probes for enriching. The enriched DNA was purified and pooled in preparation for multiplexed sequencing

proceeding qualitative DNA analysis using gel electrophoresis, though the qualitative check is optional. Sample pooling is carried out using a TapeStation or Bioanalyzer with a concentration ranging from 175 to 625 of equal bp products for maximum sequencing. The resultant enriched samples are directly sequenced using 100 + 100 bp or 150 + 150 bp paired-end sequencing with the runs organized for an 8- nt index read before aligning the sequencing reads to a reference genome.

### 2.5.5 Integrative Genomics Viewer

For viewing alignment data we used the Integrative Genomics Viewer (IGV) where we loaded the BAM files of each family members. IGV is a high-performance, open-source, easy-to-use, interactive tool for the visual exploration of genomic data. It supports flexible integration of all the common types of genomic data and metadata, investigator-generated or publicly available, loaded from local or cloud sources. The IGV application was downloaded .<sup>313</sup>

In IGV, loading an alignment file creates up to 3 associated tracks: (1) Alignment Track to view individual aligned reads (2) Coverage Track to view depth of coverage (3) Splice Junction Track which provides an alternative view of reads spanning splice junctions. The coverage track: by default IGV dynamically calculates and displays the default coverage track for an alignment file (Figure 2.5 below). The visibility range threshold: IGV reduces memory usage in the following two ways to improve performance of viewing alignments (1) a visibility range threshold which the size of the region in view at which alignments are loaded. If the region in view is larger than this threshold, no alignments are visible (default: 30 Kb) and (2) in areas of deep read coverage, by default the reads are downsampled, i.e. not all the reads are displayed. The level of downsampling is controlled with the parameters the max read count (default 100) and per window (default 50 bases). The alignment track: When zoomed in to the alignment read visibility threshold, by default 30 Kb, IGV shows the reads (Figures 2.6 and 2.7).

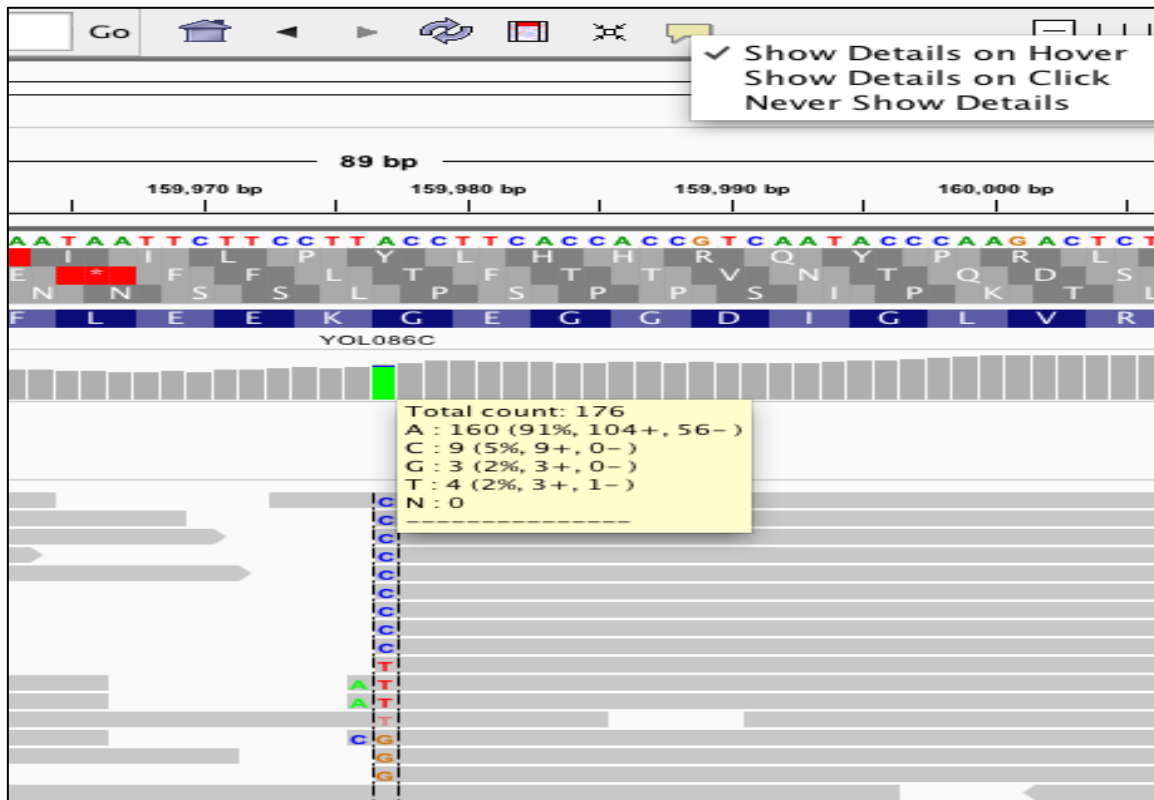


Figure 2.6: IGV coverage track. By default IGV dynamically calculates and displays the default coverage track for an alignment file. When IGV is zoomed to the alignment read visibility threshold (by default, 30 Kb), the coverage track displays the depth of the reads displayed at each locus as a gray bar chart. If a nucleotide differs from the reference sequence in greater than 20% of quality weighted reads, IGV colors the bar in proportion to the read count of each base (A, C, G, T).

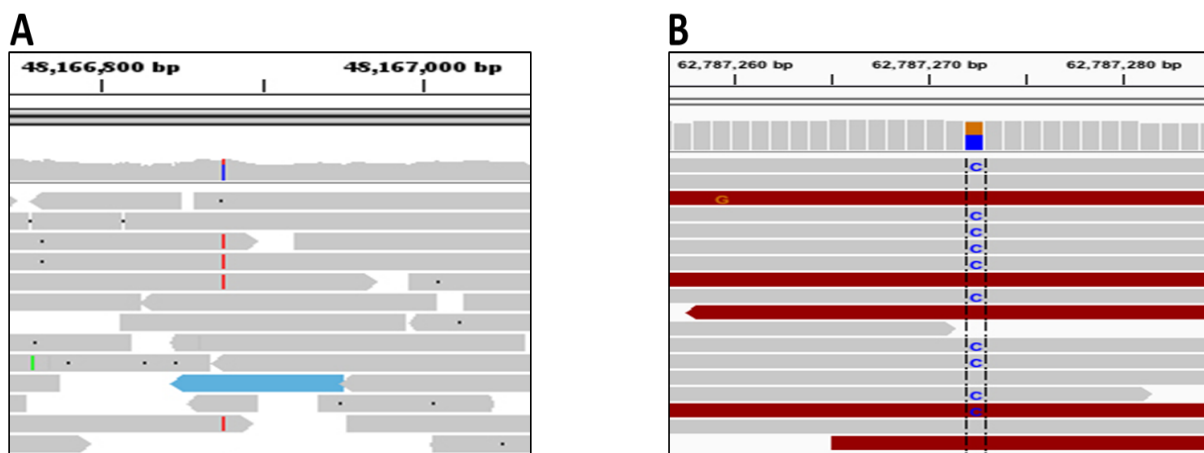


Fig 2.7: **The alignment tracks** A.) IGV shows the reads when zoomed in to the alignment read visibility threshold (default 30 Kb). IGV can display a line at the center of the display. B.) At higher resolutions, the center line becomes two lines that frame the bases centered in the display, as shown in the figure above.

When zoomed in sufficiently, IGV uses color and other visual markers to highlight potential genetic alterations in reads against a reference sequence. Genetic alternations include single nucleotide variations, structural variations, and aneuploidy. Structural variations include insertions, deletions,

inversions, tandem duplications, translocations, and other more complex rearrangements with some examples shown in figure 2.8.

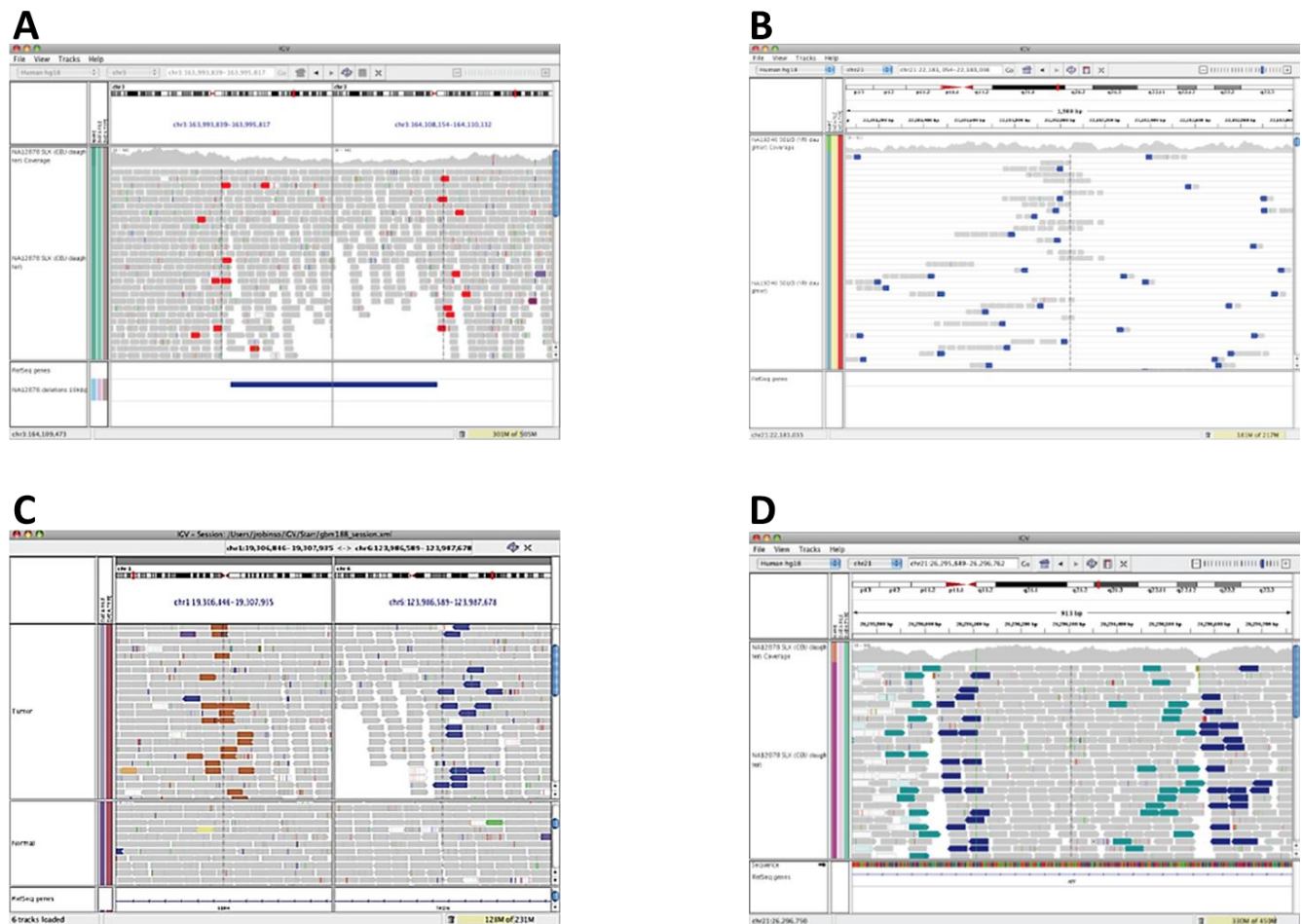


Figure 2.8: **IGV structural variations** A.) Deletions: Reads that are colored red have larger than expected inferred sizes, and therefore indicate possible deletions. B.) Insertions: in the example above reads that are colored blue have smaller than expected inferred sizes, and therefore indicate insertions. C.) Inter-chromosomal Rearrangement: IGV codes inserts for inter-chromosomal rearrangements. For instance, in this case, one end is on chromosome 1 and the other is on chromosome 6. D) Inversions An inversion is a large section of DNA that is reversed in the subject genome compared to the reference genome. Here an inversion shows up in paired-end reads, the reads are distinctively variant from the reference genome.

## CHAPTER 3 – WHOLE EXOME SEQUENCING IDENTIFIES A RARE PATHOGENIC *DMPK* MUTATION IN A FAMILY WITH SEVERE DCM

### 3.1 Introduction

Of the main inherited cardiomyopathies, DCM in particular, has a varied aetiology with an unknown cause reported in up to 70% of cases, while in SSA, important secondary causes include HIV, myocarditis, infiltrative disease, PPCM and alcohol-induced cardiomyopathy.<sup>70</sup> Population prevalence of DCM in Africa is unknown; but it is widely accepted that cardiomyopathy is common among Africans.<sup>15,70,75,154</sup>

A study by Ntusi, *et al.* reports DCM mortality as still relatively high in SSA, despite modern medical and surgical interventions.

Up to 50% of DCM cases have a genetic cause, with most families showing an AD mode of inheritance.<sup>79,155</sup> DCM causal mutations have been discovered in >60 genes which are involved in the function of various cellular components and encode a set of proteins whose functions are closely intertwined with the structural and functional aspects of cardiomyocyte biology. Consequently, pathogenic variants in genes directly disrupt cardiomyocyte homeostasis, resulting in DCM.<sup>155,83,84</sup> Also noted has been the clinical variation even in members of the same DCM family due to incomplete penetrance or genetic modifiers.<sup>156</sup> Genetic modifiers could thus provide an important contribution to the variation seen in the genotype-phenotype relationship observed in some families.<sup>157</sup>

In South Africa, the genetic landscape of cardiomyopathy necessitates more exploration, and, in this chapter, we investigate the genetic variants identified in this family with severe DCM.

## 3.2 Aim

The aim of this sub-study was to determine the genetic cause of disease in a family severely affected with DCM and to identify variants that might modify the observed phenotype.

## 3.3 Objectives

1. To use WES to identify the disease-causing mutation in a South African family with severe DCM.
2. To determine the pathogenicity of novel variants identified through screening of the extended cardiomyopathy cohort and Sanger sequencing.
3. To use WES analysis to identify genetic modifiers that might alter the expression of the phenotype.

## 3.4 Methods

### 3.4.1 Whole exome sequencing

A detailed methodology for whole exome sequencing is provided in Chapter 2, sections 2.4.

### 3.4.2 Cohort screen of *DMPK* through HRM and Sanger sequencing

To determine if additional disease-causing *DMPK* variants were present in any of the probands in our extended cardiomyopathy cohort, we screened 542 probands which included DCM (n=294), HCM (n=41), ACM (n=183), RCM (n=14), and LV non-compaction (n=10). The probands were screened for each of the exons of *DMPK* using HRM (Rotor gene 6000™) and validated using Sanger sequencing. The primer pairs used for the HRM analysis of *DMPK* exons for candidate gene screening of probands in our cardiomyopathy cohort following WES data analysis are shown in Table 3.1 below.

**Table 3.1: Primer pairs used for HRM analysis of *DMPK* exons in the cardiomyopathy cohort**

| Primer        | Sequence (5' – 3')              | Size (Bp) | T <sub>m</sub> (°C) | Product size (Bp) |
|---------------|---------------------------------|-----------|---------------------|-------------------|
| DMPK Ex 1.1 F | GGG GTG ATC AAT GGA AGC CTG AC  | 23        | 60                  | 229               |
| DMPK Ex 1.1R  | CAA CTG GGC CTG GCT TAC TTT C   | 22        | 58.4                |                   |
| DMPK Ex 1.2F  | CAC GGT GTT TAT GTG GGG AAG     | 21        | 56                  | 271               |
| DMPK Ex 1.2R  | GGT TCT AAG GCT CGG TCA TTC ATC | 23        | 59.9                |                   |
| DMPK Ex 2F    | GAT GAA TGA CCG AGC CTT AGA ACC | 20        | 57.4                | 365               |
| DMPK Ex 2R    | CCC TCT CCT CAC GGA AGC AC      | 21        | 59.9                |                   |
| DMPK Ex 3F    | CTC AGG TAG CGG TAG TGA AGA TG  | 23        | 56.6                | 327               |
| DMPK Ex 3R    | CCT CCT CGT CCA GTG ACA GC      | 20        | 59.8                |                   |
| DMPK Ex 4F    | GTC TCT GGG AGG GGT TGT CC      | 22        | 59.6                | 306               |
| DMPK Ex 4R    | CTC TGG GCT CTG ACC TTC CAA G   | 22        | 59.4                |                   |
| DMPK Ex_5F    | GGT CAG AAC TGT AGA AGC CC      | 20        | 55                  | 331               |
| DMPK Ex_5R    | CAC CTC TTT TCC CCT CCA AAT C   | 22        | 55.7                |                   |
| DMPK Ex_6F    | GAG CGG GAC TGG ATT TGG AG      | 20        | 57.9                | 257               |
| DMPK Ex_6R    | GCG TCT GCC CAT AGA ACA TTT C   | 22        | 56.3                |                   |
| DMPK Ex_7.1F  | AGA CCT ATG GCA AGA TCG TCC AC  | 23        | 58.5                | 244               |
| DMPK Ex_7.1R  | CCA AAG AAG AAG GGA TGT GTC CG  | 23        | 57.8                |                   |
| DMPK Ex 7.2F  | CGA GAC TTC ATT CAG CGG TTG     | 21        | 56.1                | 318               |
| DMPK Ex 7.2R  | TGA GTG ATT CAG GAC CCC AG      | 20        | 56.4                |                   |
| DMPK Ex_8F    | GTC ACT TTC ATA TCA GTC ACT G   | 22        | 50.9                | 176               |
| DMPK Ex_8R    | CTT CAG GTC ACG GGA AAC CAT C   | 22        | 57.9                |                   |
| DMPK Ex_9F    | ATT CTG TCT CCC TCC CCA CTC     | 21        | 58.2                | 271               |
| DMPK Ex_9R    | GGT TTC TGT CTG CTT CTG TTC AGG | 24        | 57.9                |                   |
| DMPK Ex_10F   | CAG AAA TGT TGA GGG GAC GGA GG  | 23        | 59.6                | 342               |
| DMPK Ex_10R   | AGA GGG GAT GCC AAG GGT TG      | 20        | 59.7                |                   |
| DMPK Ex_11F   | AAC CAG AAC TTC GCC AGG TCG     | 21        | 59.8                | 310               |
| DMPK Ex_11R   | TTC CCT CCT CCA GGT GTC TAT AC  | 23        | 57.2                |                   |
| DMPK Ex_12F   | CTT GGC ATC CCC TCT CGT CC      | 20        | 59.7                | 263               |
| DMPK Ex_12R   | TCC CTC CTC CAG GTG TCT ATA C   | 22        | 59.1                |                   |
| DMPK Ex_13F   | ATT CCC CTC GCT GTC TCT CG      | 19        | 58.9                | 334               |
| DMPK Ex_13R   | CGA ACT CGT CAT TGG CTG CTT C   | 22        | 58.6                |                   |
| DMPK Ex_14F   | TAG ATA ACC TCC CCA ACC TCG     | 21        | 59.8                | 259               |
| DMPK Ex_14R   | ATT GGC TGC CCG AAG ATC C       | 19        | 57.8                |                   |

Under the PCR conditions specified in Tables 3.2, the primers for exons 4,5,6,2,7,8,10,11 and 14 of *DMPK* worked efficiently.

**Table 3.2 : Standard PCR cycling parameters**

| CONDITION              | TEMPERATURE (TIME)            |
|------------------------|-------------------------------|
| Initial denaturation   | 95°C - 10 minutes             |
| Denaturation           | 95°C – 10 seconds             |
| Primer Annealing       | 55°C - 10 seconds   50 cycles |
| Temperature Elongation | 72°C - 10 seconds             |

**Table 3.3: Protocol for PCR and HRM of exon 1.1 and 6.1 (2mM MgCl<sub>2</sub>)**

| REAGENT (STOCK CONCENTRATION) | VOLUME (μl) |
|-------------------------------|-------------|
| Forward primer (20μM)         | 0.5         |
| Reverse primer (20μM)         | 0.5         |
| dNTPS (20uM)                  | 1           |
| GoTaq Flexi buffer (5X)       | 5           |
| GoTaq Polymerase (5U/μl)      | 0.1         |
| MgCl <sub>2</sub> (25mM)      | 2           |
| Evagreen (20x)                | 0.5         |
| DNA (100ng/ul)                | 1           |
| dH <sub>2</sub> O             | 14.4        |
| Total                         | 25          |

**Table 3.4 : Optimised temperature cycling conditions for the PCR and HRM of exons 1.1 and 6.1**

| CONDITION              | TEMPERATURE (TIME)            |
|------------------------|-------------------------------|
| Initial denaturation   | 95°C - 10 minutes             |
| Denaturation           | 95°C – 10 seconds             |
| Primer Annealing       | 55°C - 10 seconds   50 cycles |
| Temperature Elongation | 72°C - 10 seconds             |
| High Resolution Melt   | 72-95°C<br>(0.1°C increments) |

**Table 3.5: Protocol for PCR and HRM of exon 3 (1mM MgCl<sub>2</sub>)**

| REAGENT (STOCK CONCENTRATION) | VOLUME (μl) |
|-------------------------------|-------------|
| Forward primer (20μM)         | 0.5         |
| Reverse primer (20μM)         | 0.5         |
| dNTPS (20uM)                  | 1           |
| GoTaq Flexi buffer (5X)       | 5           |
| GoTaq Polymerase (5U/μl)      | 0.1         |
| MgCl <sub>2</sub> (25mM)      | 1           |
| Evagreen (20x)                | 0.5         |
| DNA (100ng/ul)                | 1           |
| dH <sub>2</sub> O             | 15.4        |
| Total                         | 25          |

**Table 3.6: Optimised temperature cycling conditions for the PCR and HRM of exon 3**

| CONDITION              | TEMPERATURE (TIME)              |
|------------------------|---------------------------------|
| Initial denaturation   | 95°C - 10 minutes               |
| Denaturation           | 95°C – 10 seconds               |
| Primer Annealing       | 64.6°C - 30 seconds   50 cycles |
| Temperature Elongation | 72°C - 10 seconds               |
| High Resolution Melt   | 72-95°C<br>(0.1°C increments)   |

**Table 3.7: Protocol for PCR and HRM of exon 9 (2mM MgCl<sub>2</sub>)**

| REAGENT (STOCK CONCENTRATION) | VOLUME (μl) |
|-------------------------------|-------------|
| Forward primer (20μM)         | 0.5         |
| Reverse primer (20μM)         | 0.5         |
| dNTPS (20uM)                  | 1           |
| GoTaq Flexi buffer (5X)       | 5           |
| GoTaq Polymerase (5U/μl)      | 0.1         |
| MgCl <sub>2</sub> (25mM)      | 2           |
| Evagreen (20x)                | 0.5         |
| DNA (100ng/ul)                | 1           |
| dH <sub>2</sub> O             | 14.4        |
| Total                         | 25          |

**Table 3.8: Optimised temperature cycling conditions for the PCR and HRM of exon 9**

| CONDITION              | TEMPERATURE (TIME)              |
|------------------------|---------------------------------|
| Initial denaturation   | 95°C - 10 minutes               |
| Denaturation           | 95°C – 10 seconds               |
| Primer Annealing       | 61.6°C - 10 seconds   50 cycles |
| Temperature Elongation | 72°C - 10 seconds               |
| High Resolution Melt   | 72-95°C<br>(0.1°C increments)   |

For the problematic primer pairs that required troubleshooting after a magnesium temperature gradient, we had to redesign the primers and do a magnesium temperature gradient on the new primer sets to determine their optimum working conditions. New primers were redesigned for exons, 1.2, 2, 12, 13 and the new primers used were as below (Table 3.9).

**Table 3.9: Primer redesign, magnesium chloride and temperature optimisation**

| Primer       | Sequence (5' – 3')              | Size (Bp) | Tm (°C) | Product size (Bp) |
|--------------|---------------------------------|-----------|---------|-------------------|
| DMPK Ex 1.2F | CAG GGA CAG GCA GAC ATG CAG C   | 20        | 61      | 336               |
| DMPK Ex 1.2R | CAT GCC CAT CCT GCC ATC CTG C   | 18        | 57.2    |                   |
| DMPK Ex 2F   | GCT TTC CCC CTT GTC TCC AGT G   | 20        | 58.2    | 264               |
| DMPK Ex 2R   | CAA TTT CTA AGG CCC CGC CCC AAC | 21        | 56.5    |                   |
| DMPK Ex_12F  | CTT GGC ATC CCC TCT CGT CC      | 20        | 60.1    | 244               |
| DMPK Ex_12R  | TCC CTC CTC CAG GTG TCT ATA C   | 22        | 57.1    |                   |
| DMPK Ex_13F  | CGG GGC GTG TAT AGA CAC C       | 19        | 57.7    | 282               |

|             |                               |    |      |  |
|-------------|-------------------------------|----|------|--|
| DMPK Ex_13R | TCT GCC CTC TAA AGT CGC AAA G | 22 | 57.1 |  |
|-------------|-------------------------------|----|------|--|

**Table 3.10: Protocol for PCR and HRM of exon 13 (2mM MgCl)**

| REAGENT (STOCK CONCENTRATION) | VOLUME (μl) |
|-------------------------------|-------------|
| Forward primer (20μM)         | 0.5         |
| Reverse primer (20μM)         | 0.5         |
| dNTPS (20uM)                  | 1           |
| GoTaq Flexi buffer (5X)       | 5           |
| GoTaq Polymerase (5U/μl)      | 0.1         |
| MgCl <sub>2</sub> (25mM)      | 2           |
| Evagreen (20x)                | 0.5         |
| DNA (100ng/ul)                | 1           |
| dH <sub>2</sub> O             | 14.4        |
| Total                         | 25          |

**Table 3.11: Optimised temperature cycling conditions for the PCR and HRM of exon 13**

| CONDITION              | TEMPERATURE (TIME)              |
|------------------------|---------------------------------|
| Initial denaturation   | 95°C - 10 minutes               |
| Denaturation           | 95°C – 10 seconds               |
| Primer Annealing       | 64.6°C - 30 seconds   50 cycles |
| Temperature Elongation | 72°C - 10 seconds               |
| High Resolution Melt   | 72-95°C<br>(0.1°C increments)   |

**Table 3.12: Protocol for PCR and HRM of exon 1.2 (1.5mM MgCl<sub>2</sub>)**

| REAGENT (STOCK CONCENTRATION) | VOLUME (μl) |
|-------------------------------|-------------|
| Forward primer (20μM)         | 0.5         |
| Reverse primer (20μM)         | 0.5         |
| dNTPS (20uM)                  | 1           |
| GoTaq Polymerase (5U/μl)      | 0.1         |
| MgCl <sub>2</sub> (25mM)      | 1.5         |
| Evagreen (20x)                | 0.5         |
| DNA (100ng/ul)                | 1           |
| dH <sub>2</sub> O             | 14.9        |
| Total                         | 25          |

**Table 3.13: Optimised temperature cycling conditions for the PCR and HRM of exon 1.2**

| CONDITION              | TEMPERATURE (TIME)              |
|------------------------|---------------------------------|
| Initial denaturation   | 95°C - 10 minutes               |
| Denaturation           | 95°C – 10 seconds               |
| Primer Annealing       | 59.5°C - 30 seconds   50 cycles |
| Temperature Elongation | 72°C - 10 seconds               |
| High Resolution Melt   | 72-95°C<br>(0.1°C increments)   |

**Table 3.14: Protocol for PCR and HRM of exons 2 and 12 (1mM MgCl<sub>2</sub>)**

| REAGENT (STOCK CONCENTRATION) | VOLUME (μl) |
|-------------------------------|-------------|
| Forward primer (20μM)         | 0.5         |
| Reverse primer (20μM)         | 0.5         |
| dNTPS (20uM)                  | 1           |
| GoTaq Polymerase (5U/μl)      | 0.1         |
| MgCl <sub>2</sub> (25mM)      | 1           |
| Evagreen (20x)                | 0.5         |
| DNA (100ng/ul)                | 1           |
| dH <sub>2</sub> O             | 15.4        |
| Total                         | 25          |

**Table 3.15: Optimized temperature cycling conditions for the PCR and HRM of exons 2 and 12**

| CONDITION              | TEMPERATURE (TIME)              |
|------------------------|---------------------------------|
| Initial denaturation   | 95°C - 10 minutes               |
| Denaturation           | 95°C – 10 seconds               |
| Primer Annealing       | 50.2°C - 30 seconds   50 cycles |
| Temperature Elongation | 72°C - 10 seconds               |
| High Resolution Melt   | 72-95°C<br>(0.1°C increments)   |

## 3.5 Results

### 3.5.1 Clinical history of the DCM334 family

We present a three-generation, Caucasian South African family that was diagnosed with severe familial DCM at the Cardiac Clinic, Groote Schuur Hospital, Cape Town.

The proband (III:2) and her father (II:3) were diagnosed with early onset DCM at ages 12 and 17 respectively; both required heart transplants at ages 13 and 21, respectively. At the time of transplant, the father's heart was reported to be dilated with poor contraction and high venous

pressure – features consistent with a diagnosis of DCM. Upon pathological examination, both the LV and RV showed diffuse dilatation and hypertrophy with focal endocardial scarring; no signs of a myocardial infarct was present. The clinical screening and electrocardiograms for the proband's mother and brother, II:4 and III:1 respectively, were both normal. No neuromuscular phenotypes were observed in the family during examination. Due to the SARS-CoV-2 pandemic, no additional family members could be recruited for clinical screening.

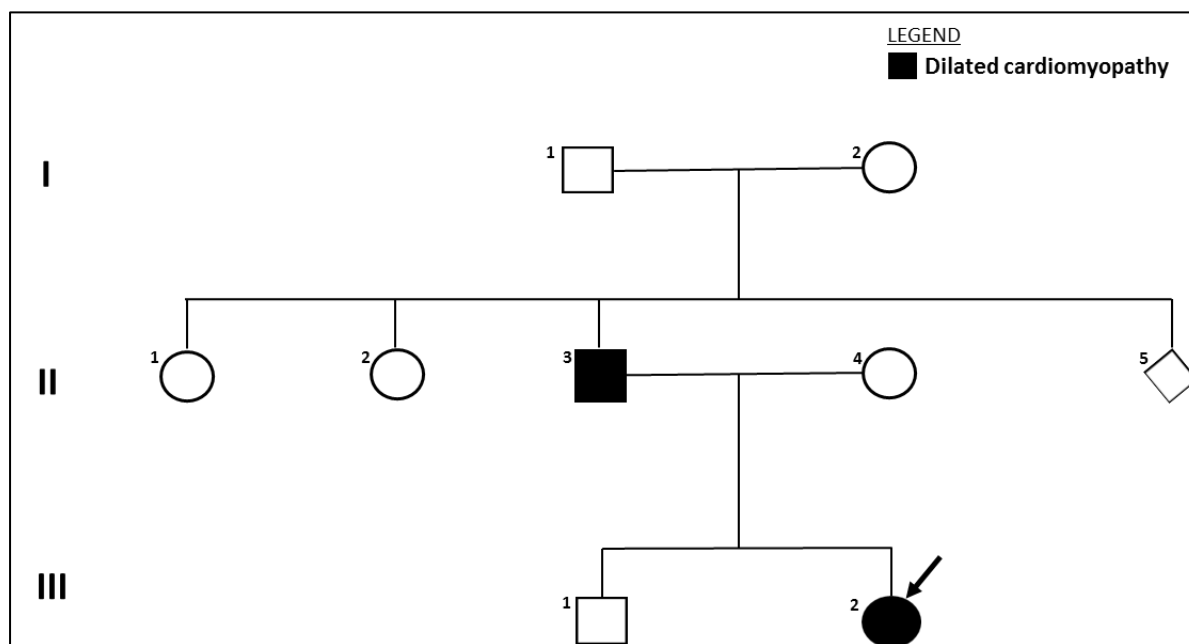


Figure 3.1: Pedigree of family DCM334. Squares represent males, circles represent females. Shaded squares and circles represent affected individuals. Arrow points towards the proband.

### 3.5.2 Whole exome sequencing

The proband (III:2) and her father (II:3) were investigated using whole exome sequencing. We focussed on variants that were common to both affected individuals. The WES data analysis generated 29,889 variants in total, at a mean sequence coverage of 136x (Figure 3.2). Of the variants, only 3 variants met the filtering criteria: *TTN* c.64207C>T, *SYNE* c.1964A>G and the rare *DMPK* c.1067C>T variant (Table 3.16). *SYNE* and *TTN* had a MAF of <0.01, whereas the *DMPK* variant was rare and has a reported frequency of  $3.1 \times 10^{-7}$ . All three variants occurred in a conserved amino acid and according to Varsome, were predicted to be damaging by 15/18, 10/17 and at 10/15 tools for the *DMPK*, *SYNE* and *TTN* variants respectively.

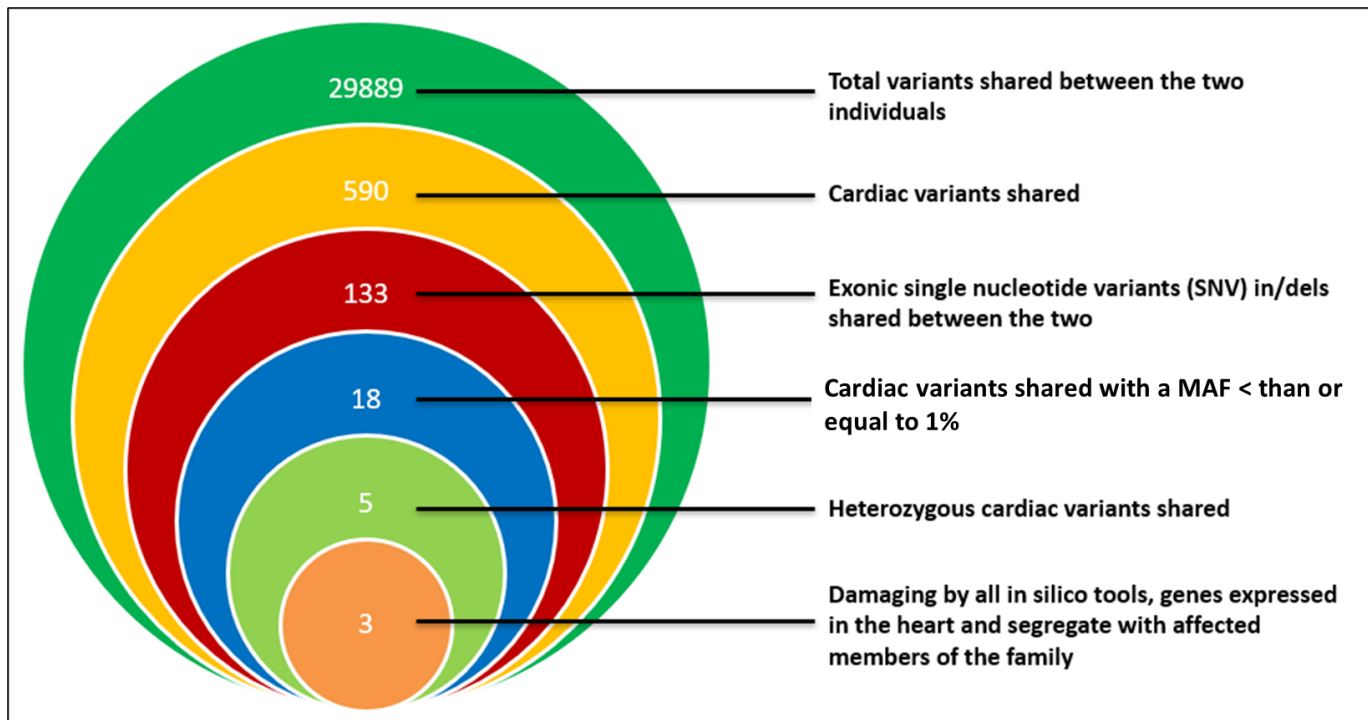


Figure 3.2: Filtering analysis of WES data from family DCM334.

**Table 3.16: Variants identified after filtering of WES data for family DCM334**

| VARIANT IDENTIFIERS (CP) |                       |           |              | POPULATION FREQUENCIES |       |                        |       | ACMG Classification | PATHOGENICITY PREDICTION |          |               |          | PROTEIN EXPRESSION |      |
|--------------------------|-----------------------|-----------|--------------|------------------------|-------|------------------------|-------|---------------------|--------------------------|----------|---------------|----------|--------------------|------|
| Gene                     | Variant               | Protein   | rsNumber     | ALFA                   | 1000G | gnomAD                 | ExAC  |                     | Varsome                  | MT       | SIFT          | PP-2     | Genecards          | Gtex |
| <i>PDE4DIP</i>           | c.6889A>G             | p.T2297A  | rs78371650   | 0.27                   |       | 0.04                   | 0.47  | -                   | -                        | PL 0.999 | T 0.74        | B 0.000  | Yes                | 2    |
| <i>PDE4DIP</i>           | c.6872_6873delGAinsAG | p.R2291Q  | -            | -                      | -     | -                      | -     | -                   | -                        | -        | -             | B 0.002  | Yes                | 2    |
| <i>PDE4DIP</i>           | c.5180T>C             | p.L1727P  | rs1778158    | 0.22                   | -     | -                      | 0.5   |                     |                          | DC 0.831 | T 1           | B 0.000  | Yes                | 2    |
| <i>PDE4DIP</i>           | c.4793A>G             | p.H1598R  | rs1778155    | 0.4                    | -     | -                      | 0.4   | -                   | -                        | PL 0.999 | T 0.15        | B 0.001  | Yes                | 2    |
| <i>OBSCN</i>             | c.4799T>A             | p.V1600D  | rs7532342    | 1                      | 1     | 1                      | 1     | 1                   | B                        | PL 0.81  | T 1           | -        | Yes                | 20   |
| <i>TTN</i>               | c.64207C>T            | p.P21403S | rs55980498   | 0.005                  | 0.001 | 0.004                  | 0.003 | 1                   | Benign                   | DC 0.999 | Not predicted | PD 0.957 | Yes                | 1056 |
| <i>TTN</i>               | c.62323G>C            | p.A20775P | rs4145333    | 1                      | 1     | 1                      | 1     | 1                   | B                        | PL 0.09  | T 0.37        | -        | Yes                | 1056 |
| <i>SCN10A</i>            | c.5137A>G             | p.M1713V  | rs6599241    | 1                      | 1     | 1                      | 1     | 1                   | B                        | PL 0.81  | T 1           | -        | No                 | 0    |
| <i>SYNE1</i>             | c.21904T>G            | p.F7302V  | rs2147377    | 1                      | 1     | 1                      | 1     | 2                   | LB                       | PL 1     | T 1           | -        | Yes                | 0.7  |
| <i>SYNE1</i>             | c.1964A>G             | p.Q655R   | rs9397509    | 0.008                  | 0.004 | 0.006                  | 0.006 |                     |                          | DC 0.999 | T 0.29        | PD 0.996 | Yes                | 0.7  |
| <i>AKAP9</i>             | c.510G>C              | p.E170D   | rs144888041  | 0.002                  | 0.001 | 0.002                  | 0.002 |                     | -                        | PL 0.9   | T 0.01        | -        | Yes                | 0.09 |
| <i>AKAP9</i>             | c.8935C>T             | p.P2979S  | rs1063242    | 1                      | 1     | 1                      | 1     | 2                   | LB                       | PL 1     | T 0.1         | -        | Yes                | 0.09 |
| <i>SVIL</i>              | c.1264G>A             | p.V422I   | rs1247696    | 1                      | 1     | 1                      | 1     | 1                   | B                        | PL 1     | D 0.02        | -        | Yes                | 3    |
| <i>RBM20</i>             | c.2303G>C             | p.W768S   | rs1417635    | 0.006                  | 0.005 | -                      | -     | 1                   | B                        |          | -             | -        | Yes                | 0.6  |
| <i>CACNA1C</i>           | c.5678A>G             | p.K1893R  | rs10774054   | 1                      | 1     | -                      | -     | 1                   | B                        |          | -             | -        | Yes                | 0.54 |
| <i>SYNM</i>              | c.702_703insG         | p.R235fs  | -            | -                      | -     | -                      | -     | -                   | -                        | DC 0.999 | -             | -        | Yes                | 34   |
| <i>DMPK</i>              | c.1067C>T             | p.P356L   | rs1026834789 | -                      | -     | 3.1 x 10 <sup>-7</sup> | -     | 3                   | VUS/LP                   | DC 0.999 | D 0           | PD 1.000 | Yes                | 1    |
| <i>TXNRD2</i>            | c.110A>T              | p.E37L    | rs148092370  | 0.001                  | 0.001 | 0.004                  | 0.003 | 1                   | Benign                   | PL 0.999 | T 0.09        | B 0.309  | Yes                | 31   |

B; Benign; DC- disease causing; D-Deleterious/damaging; MT-MutationTaster PD-probably damaging; PL-Polymorphism; PS-possibly damaging; T-Tolerated; VUS-variant of unknown significance

### 3.5.3 Variant validation

The three variants were screened in the proband by HRM analysis and were subsequently validated through Sanger sequencing (Figure 3.3).

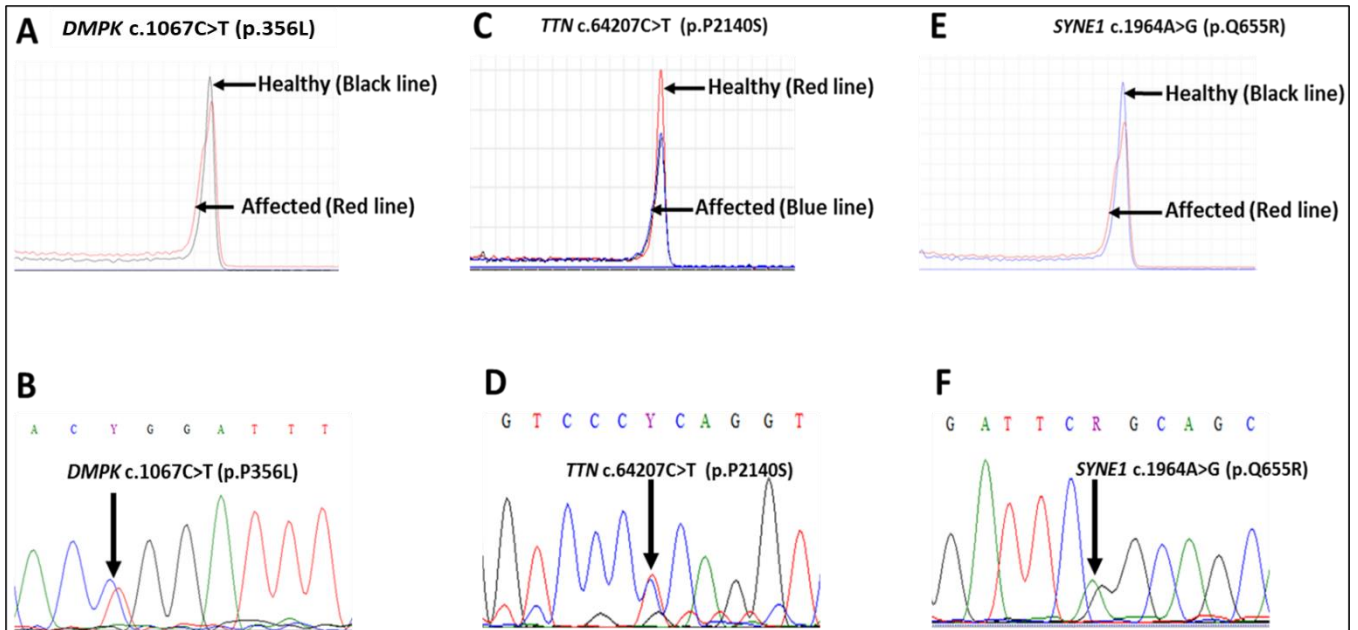


Figure 3.3: **Validation of the candidate variants identified in the proband of DCM334.** A, C and E: HRM melt profiles of control and affected individuals. B.) Sanger sequencing results reveal a heterozygous C>T change in *DMPK* at cDNA position 1067. The presence of the two peaks indicate a C (blue) and a T (red) nucleotide at this locus. D.) Sanger sequencing results reveal a heterozygous C>T change in *TTN* protein found in the proband of DCM334. The two peaks indicate a C (blue) and a T (red) nucleotide at this locus. F.) Sanger sequencing showed a heterozygous A>G genetic change in *SYNE1* as seen in the proband of DCM334 of family DCM334. The two peaks indicate an A (green) and a G (black) nucleotide at this locus.

When we looked at the conservation of the p.P356L variant across multiple species, we found that the amino acid is highly conserved across various species (Figure 3.4)

|                 |     |            |         |         |     |
|-----------------|-----|------------|---------|---------|-----|
| H. sapiens      | 347 | LRDSVPPFTP | PDFEGAT | 362     |     |
| M. mulatta      | 400 |            | FTP     | PDFEGAT | 406 |
| F. catus        | 373 | LRDSVPPFTP | PDFEGAT | 388     |     |
| M. musculus     | 347 | LRDSVPPFTP | PDFEGAT | 362     |     |
| D. Melanogaster | 377 | IRQGPAPYV  | PEVSSPT | 392     |     |
| C. elegans      | 367 |            | V       | PEVSSPE | 374 |

Figure 3.4: Conservation of the *DMPK* protein residue across multiple species.

### 3.5.4 *DMPK* c.1067C>T(p.P356L) protein structural prediction

Investigation of the secondary structural changes caused by the *DMPK* c.1067C>T(p.P356L) variant, through PHYRE2, showed that residues 1 to 406 of the mutant protein were correctly aligned and also included our region of interest (residue 356). Visualisation by Chimera identified the *DMPK* c.1067C>T(p.P356L) in the variant sequence(3.5A) when compared to the wild type (Figure 3.5B)

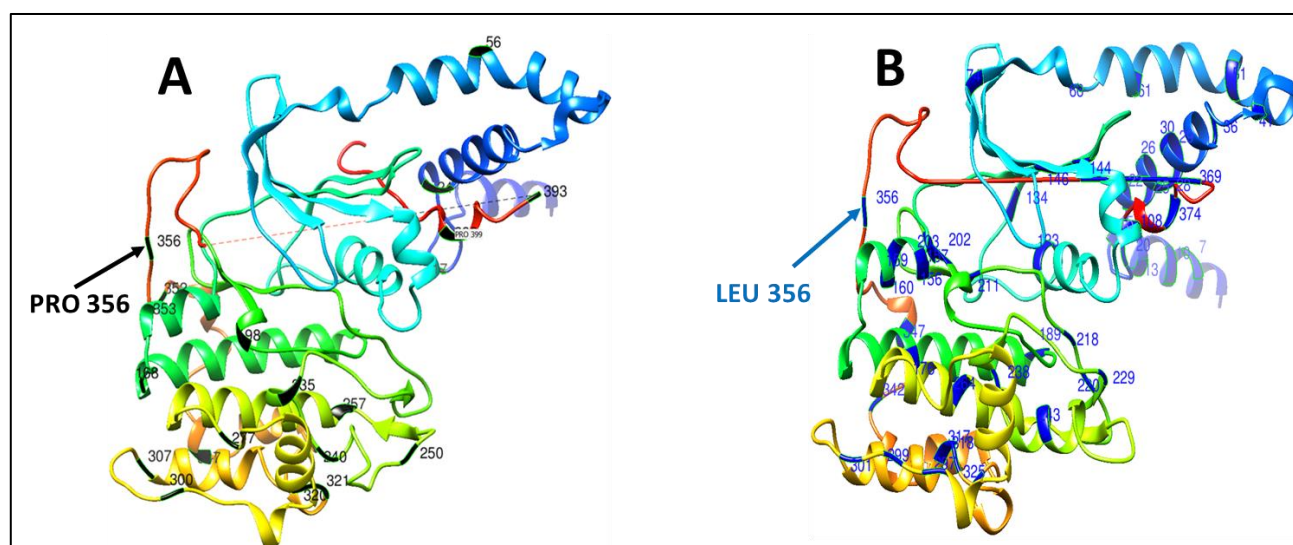


Figure 3.5: Alignment coverage of *DMPK* with the *DMPK* c.1067C>T(p.P356L) residue using Chimera to predict the effect on protein secondary structure of nonsynonymous mutations.

Site directed mutator (SDM) database predicted the outcome of the *DMPK* c.1067C>T(p.P356L) mutation to cause a reduction in the stability of the resultant protein, similar to outcomes of reduced protein stability in known DCM causing *PLN* c.25C>T9(p.R9C) mutation and known ACM causing *PKP2* c.1162C>T (R388W) mutation (Table 3.17).

**Table 3.17: SDM prediction outcome of *DMPK*, *PLN* and *PKP2* variants**

| PDB file  | Chain ID | Mutation                        | Predicted ddg | Outcome           |
|-----------|----------|---------------------------------|---------------|-------------------|
| input.pdb | A        | <i>DMPK</i> c.1067C>T (p.P356L) | -1.09         | Reduced stability |
| 1PLP.pdb  | A        | <i>PLN</i> c.25C>T(p.R9C)       | -0.52         | Reduced stability |
| 3TT9.pdb  | A        | <i>PKP2</i> c.1162C>T (p.R388W) | -0.14         | Reduced stability |

### 3.5.5 *DMPK* IGV visualization

Through IGV, we manually investigated the *DMPK* IGV tracks for the proband and her father which confirmed the presence of the *DMPK* c.1067C>T variants in the affected proband (II:3) and his father (III:2) with 99% confidence of base calls due to the high threshold IGV uses in its base calls (Figure 3.6). The *DMPK* gene has a canonical reverse orientation and with IGV having no option of changing the orientation of genes that are biologically in the reverse orientation, the observation of *DMPK* using IGV shows the bases reversed in a complementary order.

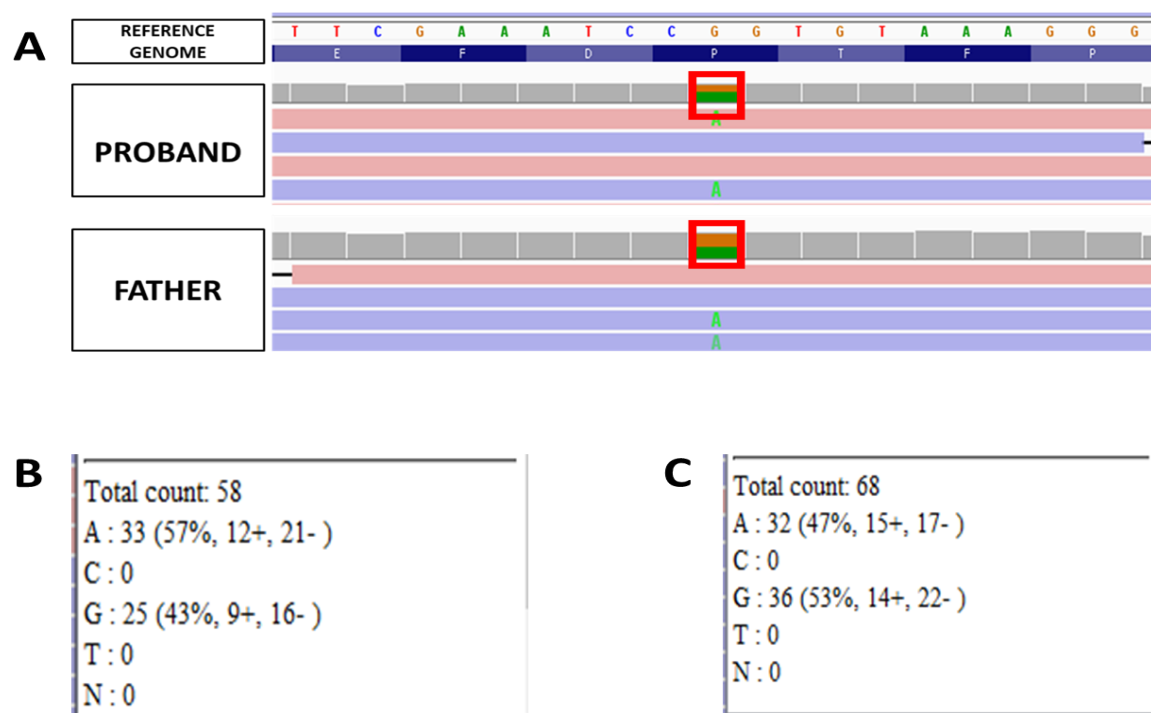


Figure 3.6: Visualization of *DMPK* c.1067C>T variant using IGV with the reference genome hg19. A.) A default threshold of 20 in IGV identifies allelic heterogeneity highlighted in red at this locus in both the proband and father with the C>T change in the canonical reverse orientation for *DMPK* thus G>A. B.) The allelic heterogeneity identified in the proband indicates a total of 58 bases called at this locus with 25 being G(brown color) and 33 being A (green color). C.) Allelic heterogeneity in the father shows 68 base calls at this locus with 36 being G (brown color) and 32 being A (green colour).

### 3.5.6 *DMPK* Familial segregation

HRM analysis for the *DMPK* c.1067C>T variant in this family showed that the proband's father had a different melting curve when compared to the wildtype. This was confirmed via

Sanger sequencing. Familial segregation of the *DMPK* variant also revealed that the two unaffected family members were negative for the variant. The results revealed that the novel *DMPK*c.1067C>T variant segregated with disease in this DCM family (Figure 3.7).

For the *TTN*c.64207C>T and *SYNE1*c.1964C>T variants, HRM analysis showed that both the affected individuals and wildtype harboured the two variants. Familial segregation analysis excluded the *TTN* and *SYNE* variants as they did not segregate with disease. (Figure 3.6).

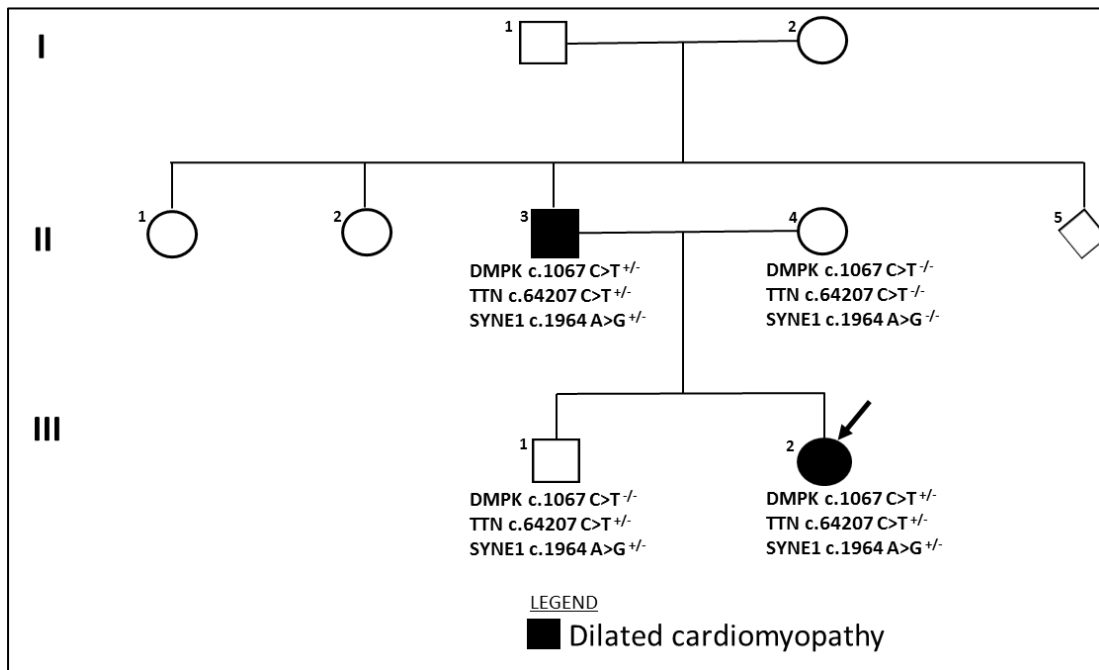


Figure 3.7: Pedigree showing segregation pattern of the *DMPK*, *TTN* and *SYNE* variants identified after WES data analysis of family DCM334. Squares represent males, circles represent females. Shaded squares and circles represent affected individuals.

### 3.5.7 *DMPK* cardiomyopathy cohort screen through HRM and Sanger sequencing

Through HRM and Sanger sequencing of the *DMPK* gene, we screened our extended cardiomyopathy cohort of 542 probands and identified the *DMPK* c.1477C>T(p.R493C) variant in DCM422.1 This variant was reported to be rare (MAF of  $7.49 \times 10^{-4}$ ), was predicted to be pathogenic by 11/18 prediction tools in Varsome (including MutationTaster, SIFT and Polyphen); was also shown to be conserved across multiple species (Figure 3.8) and was validated through HRM and sanger sequencing (Figure 3.9). However, the ACMG criteria classified the variant as benign as there have been three conflicting submissions in ClinVar, two of which report this variant as uncertain significance and another as likely benign. The patient with this variant was diagnosed with idiopathic DCM with LVEF of 20%

and passed away due to advanced HF at 17 years of age. The patient had presented with fatigue and mild, flu-like illness preceding a syncopal event, chest pain and palpitations. DCM was confirmed on echocardiography and CMR. After starting treatment, he showed symptomatic improvement with no orthopnoea, paroxysmal nocturnal dyspnoea or leg swelling.

| species     | match         | gene                               | aa     | alignment              |
|-------------|---------------|------------------------------------|--------|------------------------|
| Human       |               |                                    | 493 RQ | SLSREMEAIRTDNQNFASQLRE |
| mutated     | not conserved |                                    | 493 RQ | SLSREMEAICTDNQNFASQLR  |
| Ptroglyotes | no homologue  |                                    |        |                        |
| Mmulatta    | all identical | <a href="#">ENSMUG00000016457</a>  | 532 RQ | SLSREMEAIRMANQNFASQLR  |
| Fcatus      | all identical | <a href="#">ENSFCAG00000007334</a> | 518 RQ | SLSRELEAIRTANQSFASQLH  |
| Mmusculus   | all identical | <a href="#">ENSMUSG00000030409</a> | 495 RQ | SLSRELEAIRTANQNFSSQLQ  |

Figure 3.8 Conservation of the DMPK p.R493C residue across multiple species

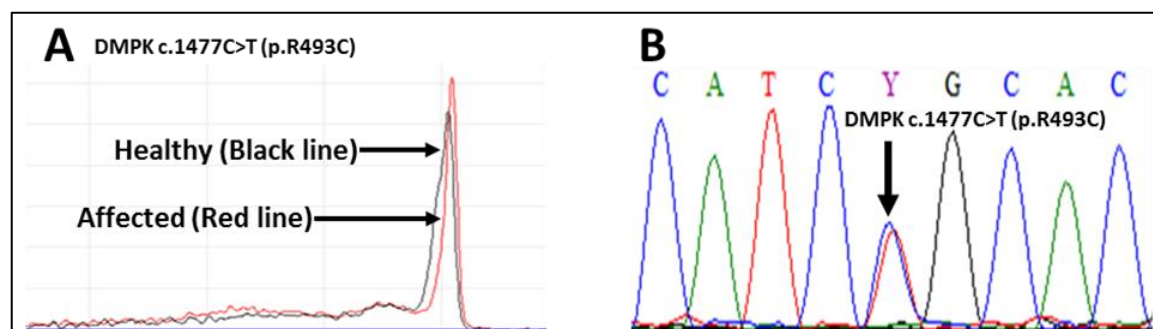


Figure 3.9: HRM and Sanger sequencing results of *DMPK* c.1477C>T variant

### 3.5.8 Modifiers

When we explored the data for possible modifiers, we identified one class three variant amongst genes on the cardiac panel and one class three variant amongst the genes not on the cardiac panel, but novel genes that are expressed in the heart (Table 3.18 and 3.19, respectively).

**Table 3.18 Possible modifier variants occurring in known cardiac genes in family DCM334**

| VARIANT IDENTIFIERS (CP) |                |         |             | POPULATION FREQUENCIES |       |        |      | ACMG Classification | PATHOGENICITY PREDICTION |           |      |      | PROTEIN EXPRESSION |      |
|--------------------------|----------------|---------|-------------|------------------------|-------|--------|------|---------------------|--------------------------|-----------|------|------|--------------------|------|
| Gene                     | Variant        | Protein | rsNumber    | ALFA                   | 1000G | gnomAD | ExAC | 3                   | Varsome                  | MT        | SIFT | PP-2 | Genecards          | Gtex |
| <i>OBSCN</i>             | 2997-2998TG/CA | A1000T  | rs386640006 | -                      | -     | -      | -    |                     | VUS                      | DC (0.97) | -    | -    | YES                | 20   |

CP, Cardiac Panel; MT, MutationTaster; PP-2, Polyphen-2; VUS-variant of unknown significance DC- disease causing

**Table 3.19: Possible modifier variants occurring in novel genes in family DCM334**

| VARIANT IDENTIFIERS (UCP) |          |         |             | POPULATION FREQUENCIES |       |        |      | ACMG classification | PATHOGENICITY PREDICTION |            |          |        | PROTEIN EXPRESSION |      |
|---------------------------|----------|---------|-------------|------------------------|-------|--------|------|---------------------|--------------------------|------------|----------|--------|--------------------|------|
| Gene                      | Variant  | Protein | rsNumber    | ALFA                   | 1000G | gnomAD | ExAC | 3                   | Varsome                  | MT         | SIFT     | PP-2   | Genecards          | Gtex |
| <i>IQGAP2</i>             | c.932A>G | p.N311S | rs371271172 | -                      | -     | -      | -    |                     | VUS                      | DC (0.999) | D (0.04) | PD (1) | YES                | 2    |

UCP, Unknown cardiac panel; DC- disease causing; D-Deleterious; MT, MutationTaster; PD-probably damaging

## 3.6 Discussion

We investigated a family that had presented with early-onset severe DCM to determine the genetic cause of the heart disease in both the proband and her father.

### 3.6.1 *DMPK* variants

WES was performed on the two severely affected DCM family members of DCM334, (DCM334) and through the cardiomyopathy cohort screen we identified the *DMPK* c.1067C>T variant as a possible candidate of disease in this family. Through the extended cardiomyopathy cohort screen of 542 probands, we also identified the *DMPK* c.1477C>T variant in DCM 361; the variant was classified as benign. However, further interrogation of this variant in ClinVar found three conflicting reports: two diagnostic laboratories reported this variant to be of unknown significance as it was detected in affected patients (no information was provided on the condition); the one publication which reported the variant to be likely benign, detected it in a patient having WES for a myotonic dystrophy study, however, it was noted that she had no known history of myotonic dystrophy. We propose that this variant does not lead to myotonic dystrophy but could lead to DCM in an individual with additional risk factors; we have classified this variant as a variant of unknown significance.

The protein, myotonic dystrophy protein kinase (DMPK) is involved in regulating the production and function of the cytoskeleton by interacting with regulatory proteins in muscle cells and is highly expressed in the heart (Figure 3.10).<sup>158</sup> DMPK is a membrane-bound protein that has been localised to the cardiomyocyte sarcoplasmic reticulum, intercalated discs and Purkinje fibres.<sup>159</sup> Studies have shown that DMPK modulates cardiac contractility and conduction activity, probably through the regulation of cellular calcium homeostasis.<sup>159</sup>

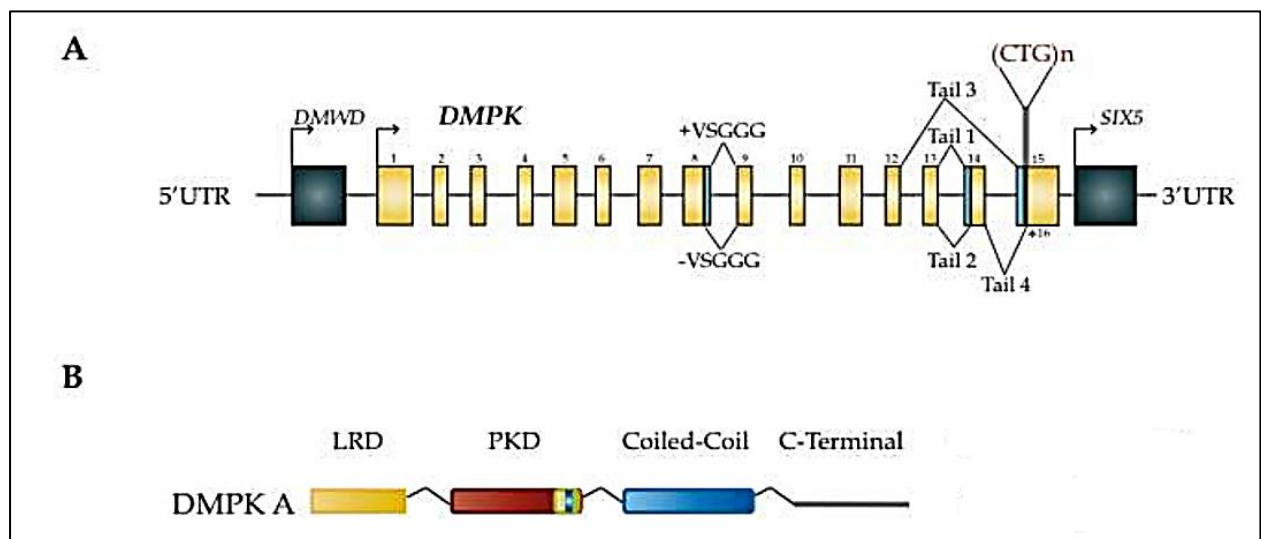


Figure 3.10: **The *DMPK* gene** A.) is composed of 15 exons, represented by rectangles. The (CTG)<sub>n</sub> expansion mutation over exon 15 is shown. Exons that are subject to alternative splicing are also indicated. B.) All isoforms have a leucine-rich N-terminus domain (LRD), a serine/threonine protein kinase domain (PDK), a VSGGG motif, and an  $\alpha$ -helical coiled-coil domain. C-terminal tail is specific to each isoform and determines their subcellular localization.<sup>158</sup>

Contraction-relaxation cycles in the heart are modulated by the continual cycling of calcium via the sarcoplasmic reticulum (SR) and the activity of the SR  $\text{Ca}^{2+}$ -ATPase (SERCA2a) which is closely regulated by phospholamban, a protein localized in the SR membrane.<sup>160,161</sup> Phosphorylation of PLN at serine 16 by PKA under physiological conditions escalates ventricular relaxation and accelerates  $\text{Ca}^{2+}$  cycling into the SR.<sup>162</sup>

The central feature of HF in humans and animals is caused by a decrease in  $\text{Ca}^{2+}$  uptake and a rise in PLN levels bound with SERCA2a, which is highly significant in determining any cardiac SR anomaly.<sup>159,160</sup> In most acquired forms of heart failure, PLN is evidently hypophosphorylated.<sup>160</sup> DCM patients who have different PLN expression have exhibited severe cardiomyopathy and even required heart transplantation at a young age of 24, which is a similar observation to what we have seen in the proband and her father where both required heart transplants at young ages.

In vitro studies done have shown that DMPK is responsible for the phosphorylation of PLN and modulation of calcium uptake in cardiomyocyte SR.<sup>163-165</sup> In cardiac homogenates obtained from *dmpk*<sup>-/-</sup> mice, PLN was shown to be hyperphosphorylated in addition to a hypophosphorylation of.<sup>166</sup> PLN serine 16 that corresponds to a significant reduction in SR  $\text{Ca}^{2+}$  uptake in *dmpk*<sup>-/-</sup> ventricular homogenates.<sup>159</sup> This loss of PLN phosphorylation by Dmpk highlights the significance of LV dysfunction in neuromuscular diseases like muscular dystrophy.<sup>158</sup>

DMPK is also renowned for causing Myotonic dystrophy type 1 (DM1), which is responsible for up to 43% of SCD in patients.<sup>167</sup> Ventricular and atrial arrhythmias, compromised cardiac contractility, atrioventricular conduction delay and mesocardial fibrosis highlight some of the manifestations unique to the heart in DM1. The aberrant expansion of DMPK CTG triplet at the 3' untranslated region resulting in mutant transcripts sequestering in cell nuclei is responsible for DM1 due to the abnormal transcription and splicing.<sup>167,168</sup> Depletion of the DMPK protein may additionally be responsible for the cardiac conduction irregularities observed in DM1 patients.<sup>166,169</sup>

In our study, the *DMPK* c.1067C>T (p.P356L) variant occurs at the kinase domain of DMPK protein (Figure 3.10B). This kinase domain is responsible for the catalytic properties of DMPK and it displays the characteristics of a serine/threonine kinase domain.<sup>170,171</sup> The kinase function is carried out by the kinase and protein kinase C-terminal domain between residues 70 and 405 leading to a conformational change in the target protein by shuttling a gamma phosphate from ATP to the serine or threonine on the protein of interest.<sup>158</sup> Residues 463 to 530 have the coiled-coil domain which is a prevailing structure exhibited by up to 5% of all the amino acids in a specific protein.<sup>158</sup> This coiled coil feature is responsible for the substrate binding, multimerisation, localisation and kinase properties conferred by *DMPK* on its targets which take part in significant cellular activities including ion homeostasis and cytoskeleton remodeling thus underpinning the wide functional spectrum of *DMPK*.<sup>172,173</sup>

We postulate that the pathogenic *DMPK* variant could result in an abnormal phosphorylation of PLN which in turn, would lead to under-phosphorylated PLN in SR vesicles and that this correlates with highly reduced SR Ca<sup>2+</sup> uptake. This loss of PLN phosphorylation due to the mutations in DMPK therefore constitute a significant factor in severe DCM as ventricular function is shown to be regulated through DMPK's kinase domain phosphorylating the Ser-16 of PLN. This study highlights the potential role of *DMPK* as a new candidate gene for DCM due to its association and phosphorylating effect on PLN, thus providing a potential molecular mechanism for the ventricular aberrations observed in DCM patients.<sup>158,163-165</sup>

### 3.6.2 Modifiers

We also explored the possibility of genetic modifiers to determine if the *DMPK* c.1067C>T (p.P356L) variant could possibly interact with other plausible candidates to exacerbate or alter the phenotype seen within this family.

Our search for modifier variants yielded one class 3 variant, the *OBSCN* c.2997-2998TG/CA (p.A1000T) mutation from the cardiac panel and the *IQGAP2* c.932A>G (p.N311S) variant from the unknown cardiac panel.

#### 3.6.2.1 *OBSCN* c.2997-2998TG/CA (p.A1000T)

WES data for the known cardiomyopathy panel identified the *OBSCN* c.2997-2998TG/CA (p.A1000T) mutation in both affected individuals (Individual III:2 the proband and her father individual II:3). This variant is novel as it does not appear in gnomAD or 1000 genomes and was predicted to be disease causing by MutationTaster variant prediction tool, while SIFT and Polyphen prediction tools did not have data available on this variant (Table 3.4).

*OBSCN* which encodes obscurins, which are located at different subcellular locations with a host of binding partners involved in structural linkage between the cytoskeleton, sarcolemma, superficial myofibrils with the internal membrane.<sup>174</sup> They also serve to ensure the stabilisation and proper assembly of sarcomeric A-, M-bands and the sarcoplasmic reticulum membranes, homeostatic regulation of Ca<sup>2+</sup>, cell adhesion and mechanotransduction.<sup>175-179</sup>

Previous reports of *OBSCN* mutations causing cardiomyopathies highlight the potential role of our identified *OBSCN* c.2997-2998TG/CA(p.A1000T) variant in modifying the DMPK c.1067C>T (p.P356L) variant to cause the observed phenotype.<sup>180</sup> In addition, the *OBSCN* variant we have identified occurs four residues downstream of another *OBSCN* residue (A996fs) that is known to cause a cardiomyopathy phenotype (Figure 3.11).<sup>181</sup> Four mutations in *OBSCN* have been linked to DCM (Figure 3.10, red), seven linked to HCM (blue), three with LVNC (green) and one mutation each identified in a case of HF and distal myopathy (orange and purple, respectively)

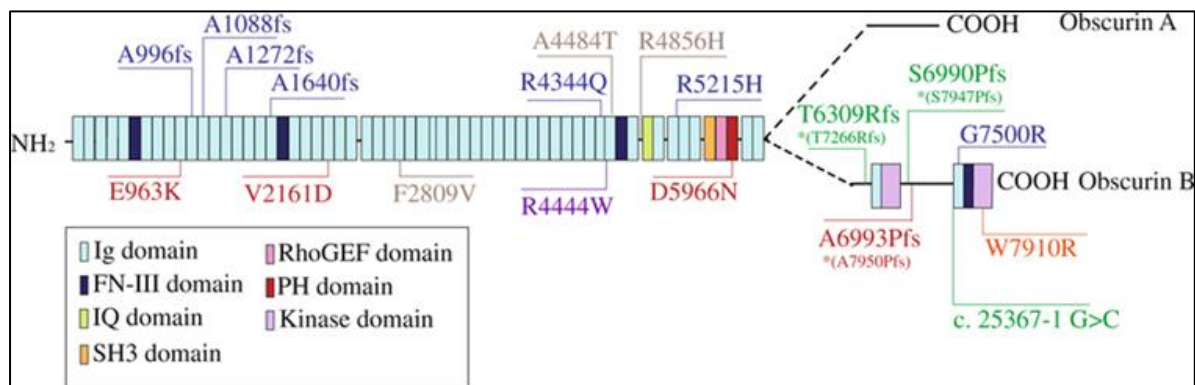


Figure 3.11: *OBSCN* mutations identified to cause cardiomyopathies.<sup>171</sup>

Animal studies have strengthened the association of *OBSCN* variants with an abnormal phenotype. In a zebrafish knockout model, abnormal heart contraction and abnormal heart structure were observed.<sup>182</sup>

The *OBSCN* c.2997-2998TG/CA(p.A1000T) variant we identified is supported in its role as a potential genetic modifier due to its pathogenicity prediction and its novelty was attributed to the absence of any frequency data linked to it, though caution is to be taken in interpreting any novel genetic variant with no frequency data as automatically disease causing.

### 3.6.2.2 *IQGAP2* c.932A>G (p.N311S)

Using the WES dataset, we explored the possibility of novel candidate genes with expression in the heart; we identified the *IQGAP2* c.932A>G(p.N311S) variant in both the proband (Individual III:2) and her father (Individual II:3). The variant was predicted to be pathogenic by at least 3 *in silico* tools and did not have any reported population frequencies in the public databases. *IQGAP2* belongs to a family of evolutionary conserved proteins which are involved in cytoskeletal dynamics, intracellular signalling, cell proliferation and cytokinesis alongside mediating protein-protein interactions through their domains.<sup>183-186</sup>

In vitro studies, *IQGAP1*-null mice were found to have unfavourable ventricular wall thinning, increased apoptosis and decreased contractility as part of cardiac remodelling. *IQGAP1* modulates Akt activation and ERK that manage cardiac remodelling.<sup>187</sup> *IQGAPs* share some binding partners including F-actin and calmodulin and it is possible that the binding partners *IQGAP1* interacts with to cause the cardiac remodelling might be the same binding partners *IQGAP2* interacts with to induce a cardiomyopathy phenotype due to cardiac remodelling manifested through impaired ejection fraction, LV wall thinning and reduced cardiac contractility.<sup>188</sup>

String pathway analysis of *IQGAP2* shows it interacts with beta-myosin heavy (*MYH7*) protein, which is a product of *MYH7*, which is associated with SCD caused by HCM and has been shown to be responsible of up to 3% of DCM amongst DCM patients in studies done.<sup>189,190</sup> Ventricular remodelling happens in *MYH7* gene anomalies and thus a variation in the wild-type form of *IQGAP2* could have a possible effect in *MYH7*-related interactions to cause DCM.<sup>46,191</sup>

The *OBSCN* c.2997-2998TG/CA(p.A1000T) and the *IQGAP2* c.932A>G(p.N311S) variants have thus been identified as potential genetic modifiers along with the primary *DMPK* c.1067C>T(p.P356L) in the DCM family studied.

### 3.7 Conclusion

WES analysis was done on a South African family with severe early onset DCM and a novel *DMPK* variant was identified in all affected family members. Through additional screening of our extended cardiomyopathy cohort, we identified another rare *DMPK* variant of unknown significance in a young, severely affected DCM proband. We propose that this variant could lead to DCM in an individual with additional risk factors and have classified this variant as a variant of unknown significance. To the best of our knowledge, we are the first to report *DMPK* as a primary cause of DCM. Analysis for any potential genetic modifiers also identified potential modifying variants in *OBSCN* and *IQGAP2* in both the affected individuals. During SARS-Cov2, patients had direct contact with the doctor on the study who then counselled the patients if they had any concerns. The doctor sent the patients to the hospital if there was any major concern.

Improved understanding of the aetiology of DCM and advancements in NGS technologies can help improve current methods of diagnosis, intervention and personalised treatment of DCM. By identifying and associating specific variants or genes with the pathogenesis of the disease, family members of affected individuals can be provided with the option of being screened for that variant. Prevention and treatment options can then be offered should there be a risk of developing the disease. Studies, such as these, where next generation sequencing

technologies are used, pose a benefit in that the gap between the known and unknown genetic causes of DCM can be filled. By associating the two *DMPK* variants identified above with cardiomyopathy, pharmacological interventions that include beta-adrenergic blockers or angiotensin-converting enzyme inhibitors that have been shown to be well tolerated in DM patients with cardiac involvement including left ventricular dysfunction are a possible treatment regimen. However, additional studies are needed to cement the role of the modifiers in this family.

# CHAPTER 4 – WHOLE EXOME SEQUENCING IDENTIFIES A *BAG3* MUTATION AS THE CAUSE OF DCM IN A SOUTH AFRICAN FAMILY

## 4.1 Introduction

The most prevalent cardiomyopathy is DCM although its precise prevalence remains to be determined. It represents the most common cause for heart transplants and leads to the highest proportions of heart failure amongst the young who are affected.<sup>192</sup> Researchers have been attempting to define the underlying causes of DCM but it is still a challenge, however, disease-causing genetic variations have increasingly been found to be associated with an increased risk for arrhythmias in recent years.

The advent of NGS, thus changed the approach to DCM by enabling the identification of multiple mutations in genes related to DCM, with a pathogenic variant being discovered in up to 40% of DCM cases.<sup>193</sup> These genes have been known to exist in concert with other cardiomyopathies presenting difficulty in their characterization and evaluation.

Most of the mutations implicated in DCM are rare and often mutations found in DCM families have not been reported, as seen from a study conducted by Norton and colleagues where they showed that from an examination of 197 specific mutations in 235 individuals, 88% of them were private and 89% being single-base variants.<sup>85,194</sup> In sub-Saharan Africa DCM has the highest prevalence and thus far, no widespread evaluation of the disease among African families has been done and this underpins the urgent importance of African patients to undergo extensive genetic testing to gain more insight on the disorder and offer effective treatments. The few studies emanating from Africa pertaining to DCM genetics are insufficient thus pointing towards the critical urgency of a full-scale genetic evaluation of

individuals affected with DCM in Africa to get more insight into the genetic architecture and develop better management strategies for the patients.<sup>8,13,15</sup>

In this chapter, we investigate a six-generation South African family with a generational history of heart failure which was characterised by a variable phenotype expression, that included DCM and LVNC.<sup>56</sup>

## 4.2 Aims

To identify the disease-causing variant and any associated modifying gene variants in a multi-generational family with DCM and HF.

## 4.3 Objectives

1. Utilising WES to establish a genotype-phenotype association in a DCM multi-generation family
2. To use WES and validation techniques to delineate any variants that modify the DCM phenotype in the family

## 4.4 Methods

### 4.4.1 Whole exome sequencing

Whole exome sequencing protocols and parameters are described in Chapter 2, section 2.4.

### 4.4.2 Targeted resequencing

Targeted resequencing was done using the Haloplex Target Enrichment System Reagent Kit at the University of Manchester, Oxford Medical Laboratory (Chapter 2, section 2.5).

Targeted resequencing for 38 genes (Appendix K) was performed in the principal study

called IMHOTEP where we sequenced a cohort of 634 cardiomyopathy probands (DCM=459, HCM=64, ACM=75, RCM=36)(unpublished data). We only extracted and analysed the sequencing data for *BAG3*, as the other 37 genes did not form part of this PhD study.

## 4.5 Results

### 4.5.1 Clinical history of the DCM 343 family

We present a six-generation Caucasian family with history of HF and DCM phenotype with an AD mode of inheritance (Figure 4.1).

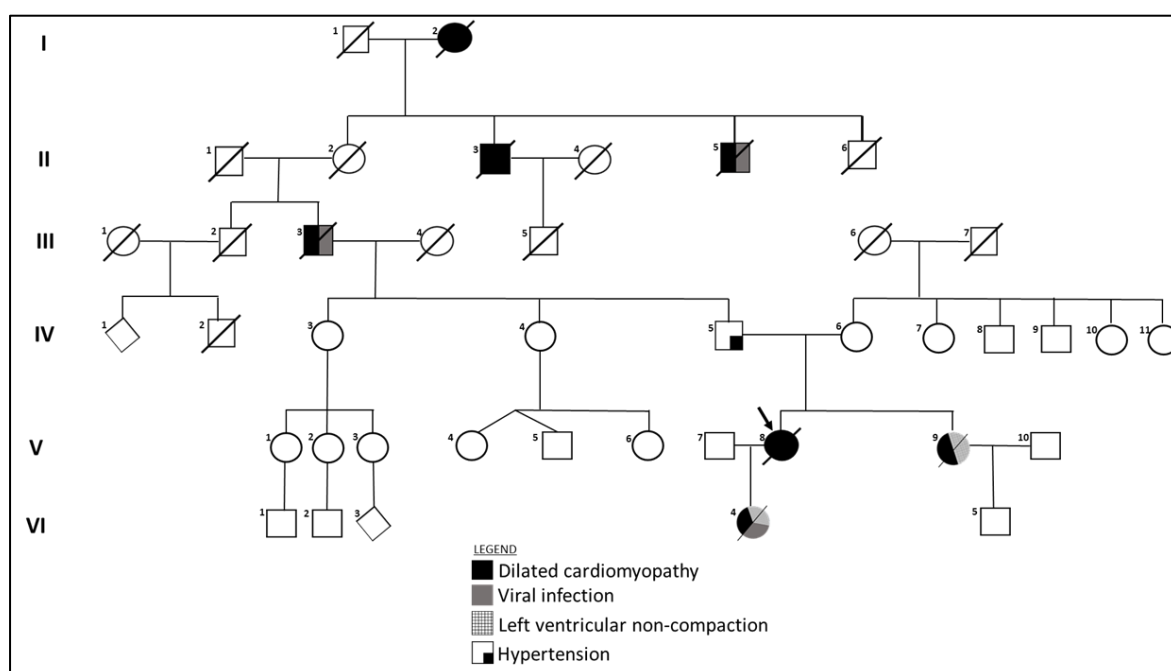
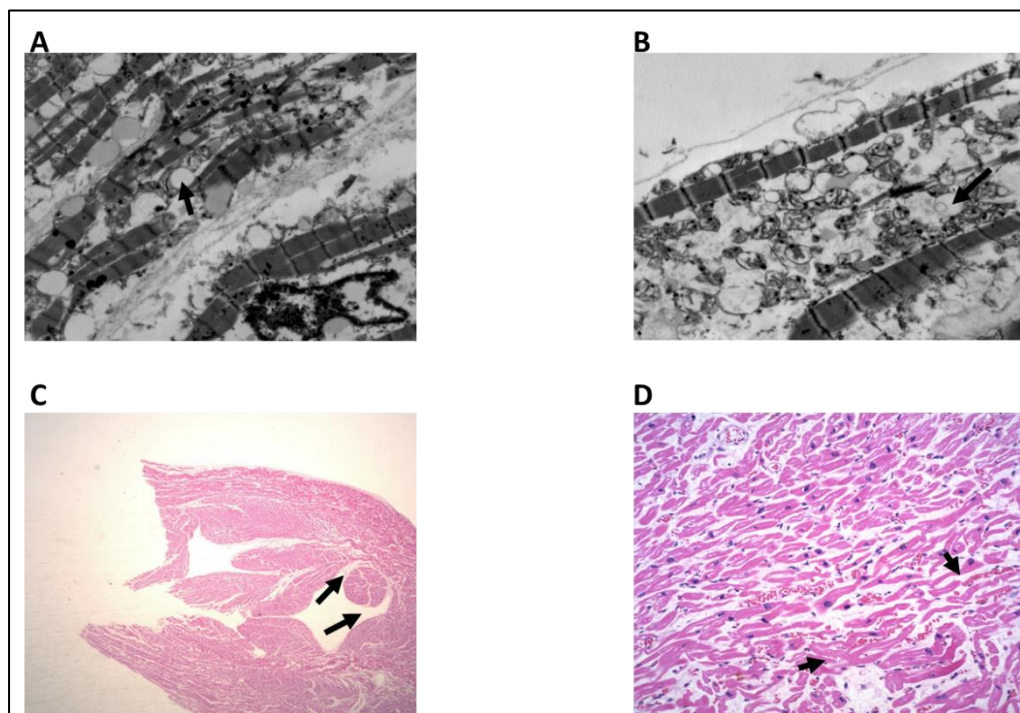


Figure 4.1: Pedigree of DCM family 343. Squares represent males, circles represent females. Shading in the symbols represent affected individuals. Crossed out squares and circles represent deceased individuals. Arrow points to the proband.

The proband V:8 was diagnosed with DCM at 28 years of age in 2010 with a LVEF of 20%. CMR performed in the United Kingdom was ambiguous about the diagnosis of LVNC. The proband's daughter (VI:4) presented at 9 years of age with HF and was diagnosed with severe DCM and a LVEF of 15%. This diagnosis was preceded by a viral infection necessitating admission to intensive care. The daughter died a week after admission. Histological evaluation of the

tissues of the deceased daughter revealed marked LV non-compaction with several microthrombi present in non-compacted areas. Most of the apex and lateral wall of the LV showed deep intertrabecular recesses and pronounced endocardial trabeculations. Pathologists observed aggregations of mitochondria consistent with abnormality or cytopathy of the mitochondria, this was related to conspicuous lipid droplets. (Figure 4.2) The proband's husband (V:7), who was in his 30's at the time of her diagnosis, did not have any features of DCM.

Figure 4.2: **Electron and light microscopy images of the heart of individual VI:4** A and B.) Electron microscopy images of the heart showing conspicuous lipid droplets indicated by the black arrows,



which are consistent with mitochondrial abnormality C.) Black arrows in histological findings show deep intertrabecular recesses consistent confirming non-compaction D.) Large pleomorphic nuclei on histology indicated by the black arrows consistent with cardiomyopathy

In 2010, the proband's father (IV:5, 54 years old), mother (IV:6, 53 years old) and sister (V:9, 30 years old), were screened and were found to be normal. Upon rescreening in 2018, the proband's sister (V:9) was found to have DCM and LVNC with an LVEF of 42%. The sister was receiving chemotherapy treatment prior to recruitment and her subsequent diagnosis with DCM and LVNC. Also noted was that the proband's father (IV:5) was being treated for hypertension with ACE inhibitors.

Interestingly, clinical records indicated that the proband's grandfather, individual III:3, had been diagnosed with myocarditis and then subsequently died at the age of 29. The grandfather's uncle, individual II:3, presented at age 59 with DCM with pronounced endomyocardial fibrosis of the LV, and had a subsequent heart transplant. Individual III:3's grandmother, I:2, presented with DCM aged 52 and died of HF. Additional family members were included for genetic testing of the candidate variant identified.

#### 4.5.2 Whole exome sequencing

Four individuals in family DCM 343 went for WES which included the proband (V:8), her affected sister (V:9), her father (IV:5) and her affected daughter (IV:4). A total of 70,341 variants in both the cardiac and candidate cardiomyopathy genes were identified in the affected individuals (Figure 4.3).

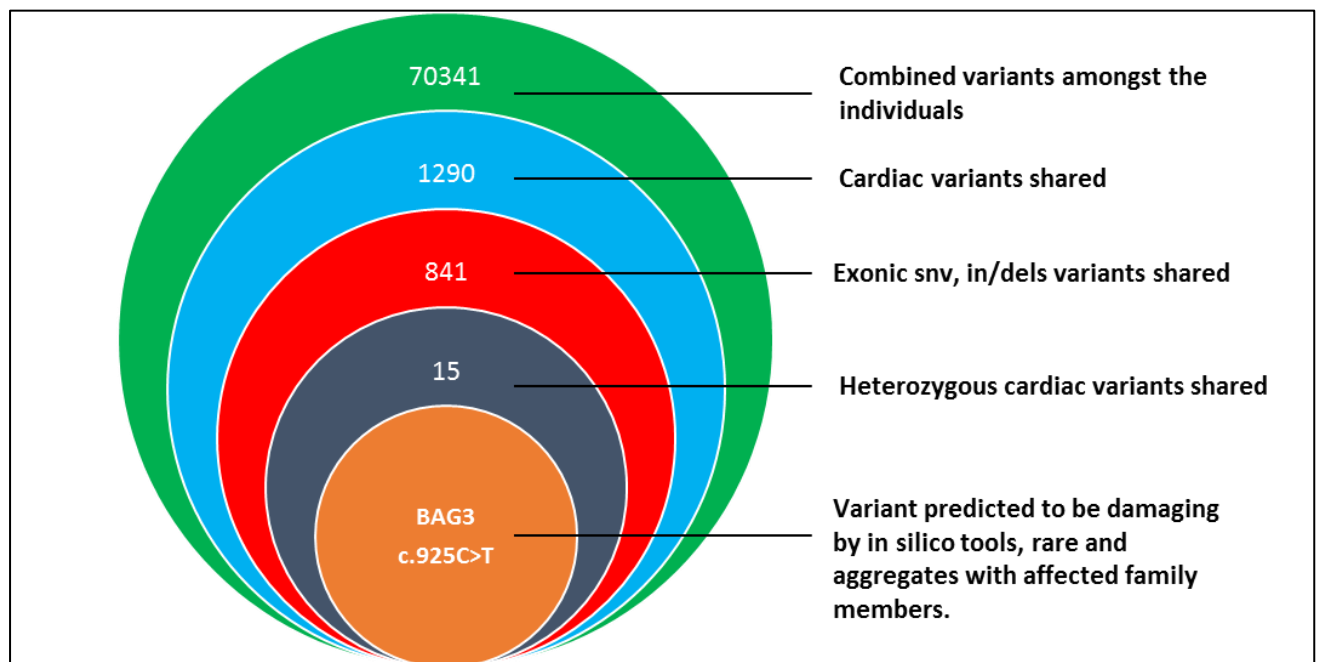


Figure 4.3: VEP data analysis of variants for family DCM343.

We applied filtering criteria discussed in Chapter 2 (section 2.4) and 1,290 variants were shared by the four affected family members, were found to occur in the known cardiac genes

panel (Appendix J). We looked for variants occurring in exonic regions (n=841), variants predicted to be damaging by *in silico* tools, population frequencies <1% or that were novel, and the variants that segregated with affected members of the family. Preliminary WES analysis showed that the four family members harboured a the *BAG3* c.925C>T (p.R309Ter) nonsense variant (Table 4.1.)

**Table 4.1: Variants identified after filtering of WES data for family DCM343**

| VARIANT IDENTIFIERS (CP) |                    |                |                    | POPULATION FREQUENCIES |          |               |      | ACMG Classification | PATHOGENICITY PREDICTION |             |        |          | PROTEIN EXPRESSION |           |
|--------------------------|--------------------|----------------|--------------------|------------------------|----------|---------------|------|---------------------|--------------------------|-------------|--------|----------|--------------------|-----------|
| Gene                     | Variant            | Protein        | rsNumber           | ALFA                   | 1000G    | gnomAD        | ExAC |                     | Varsome                  | MT          | SIFT   | PP-2     | Genecards          | Gtex      |
| <b>BAG3</b>              | <b>c.925C&gt;T</b> | <b>p.R309*</b> | <b>rs869248137</b> | <b>-</b>               | <b>-</b> | <b>0.0004</b> |      | <b>5</b>            | <b>P</b>                 | <b>DC 1</b> |        |          | <b>Yes</b>         | <b>42</b> |
| <i>PDE4DIP</i>           | c.5730C>A          | p.D1910E       | rs1613780          | 0.5                    | -        | -             | 0.5  | -                   | -                        | PL 1        | T 1    | B 0      | Yes                | 2         |
| <i>PDE4DIP</i>           | c.5599C>T          | p.R1867C       | rs1620560          | 0.28                   | -        | -             | 0.5  | -                   | -                        | P 0.94      | D 0.04 | PD 0.989 | Yes                | 2         |
| <i>PDE4DIP</i>           | c.5269G>A          | p.A1757T       | rs1628310          | 0.5                    | -        | -             | 0.5  | -                   | -                        | PL 1        | T 0.11 | B 0.031  | Yes                | 2         |
| <i>PDE4DIP</i>           | c.4511G>A          | p.R1504Q       | rs2762875          | 0.5                    | -        | -             | 0.5  | -                   | -                        | PL 1        | T 1    | B 0      | Yes                | 2         |
| <i>PDE4DIP</i>           | c.4075A>G          | p.K1359E       | rs1747958          | 0.36                   | -        | -             | 0.5  | -                   | -                        | PL 1        | T 0.38 | B 0.115  | Yes                | 2         |
| <i>PDE4DIP</i>           | c.3196G>A          | p.A1066T       | rs1698647          | 0.27                   | -        | -             | 0.5  | -                   | -                        | PL 0.69     | T 0.12 | PD 0.849 | Yes                | 2         |
| <i>PDE4DIP</i>           | c.3037T>A          | p.F1013I       | rs1698624          | 0.5                    | -        | -             | 0.5  | -                   | -                        | PL 1        | T 1    | B 0      | Yes                | 2         |
| <i>PDE4DIP</i>           | c.2042G>A          | p.R681H        | rs1629011          | 0.5                    | -        | -             | 0.5  | -                   | -                        | PL 0.91     | T 0.09 | B 0.444  | Yes                | 2         |
| <i>PDE4DIP</i>           | c.1864C>T          | p.R622Ter      | rs1778111          | 0.28                   | -        | -             | 0.5  | -                   | -                        | DC 1        | D 0.02 | PD 0.958 | Yes                | 2         |
| <i>PDE4DIP</i>           | c.1607G>C          | p.S536T        | rs1747930          | 0.5                    | -        | -             | 0.5  | -                   | -                        | DC 0.85     | D 0.01 | PD 0.958 | Yes                | 2         |
| <i>PDE4DIP</i>           | c.1445A>G          | p.H482R        | rs1698681          | 0.5                    | -        | -             | 0.5  | -                   | -                        | PL 1        | T 0.26 | B 0.02   | Yes                | 2         |
| <i>PDE4DIP</i>           | c.1229A>T          | p.E410V        | rs1061308          | 0.5                    | -        | -             | 0.5  | -                   | -                        | DC 1        | D 0    | PD 0.973 | Yes                | 2         |
| <i>PDE4DIP</i>           | c.824C>T           | p.S275L        | rs1359300          | 0.27                   | -        | -             | 0.5  | -                   | -                        | PL 1        | T 0.31 | B 0.031  | Yes                | 2         |
| <i>PDE4DIP</i>           | c.74G>T            | p.R25L         | rs1664022          | 0.5                    | -        | -             | 0.5  | -                   | -                        | DC 1        | D 0    | PD 0.999 | Yes                | 2         |

B; Benign; DC- disease causing; D-Deleterious/damaging; MT, MutationTaster; PD-probably damaging; PL-Polymorphism; Polyphen-2; PS-possibly damaging; T-Tolerated; VUS-variant of unknown significance

The *BAG3* c.925C>T variant passed all the filtering criteria, segregated with affected members of the family, and was earmarked for Sanger sequencing. The variant is an ACMG class 5 pathogenic variant and 6/8 prediction tools in Varsome predicted it to be pathogenic. The variant has a very rare population frequency of  $4 \times 10^{-5}$  according to gnomAD and was not reported in 1000Genomes, ExAC and ALFA population frequency databases.

### 4.5.3 Variant validation

The heterozygous *BAG3* c.925C>T mutation was validated through HRM and Sanger sequencing (Figure 4.4).

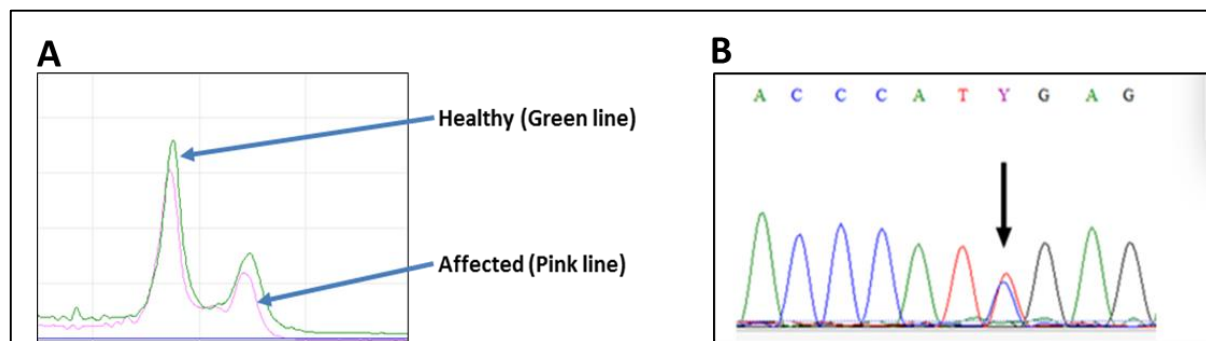


Figure 4.4: Validation of the *BAG3* c.925C>T variant in the proband of family DCM343. A.) High resolution melt analysis shows a deviation in the melt profiles when a mutation positive and DCM positive affected individual V:8 is compared against mutation negative healthy individual V:7. B.) Sanger sequencing results revealed the heterozygous C>T change in a mutation positive individual V:8. The presence of the two peaks indicate a C (blue) and a T (red) nucleotide at this locus.

When we looked at the conservation of the p.R309 residue across multiple species, we found that the amino acid is highly conserved across various species (Figure 4.5).

|                |     |                          |
|----------------|-----|--------------------------|
| H. sapiens     | 309 | RPQQPMTH <b>R</b> ETAPVS |
| P. troglodytes | 309 | RPQQPMTH <b>R</b> ETAPVP |
| M. mulatta     | 309 | RPQQPMTH <b>R</b> ESAPVP |
| F. catus       | 251 | RPQQPMTH <b>R</b> ESSPVP |
| M. musculus    | 315 | RPQ PMTH <b>R</b> EPPPV  |
| X. tropicalis  | 309 | PQ <b>R</b> ESSPKP       |

Figure 4.5: Conservation of the *BAG3* p.R309 residue across multiple species.

4.5.4 *BAG3* c.925C>T (p.R309Ter) protein structural prediction

To look for secondary structural changes caused by the *BAG3* c.925C>T variant we submitted our query sequence to PHYRE2. The results indicated that only residues 20 to 58 of the mutant *BAG3* protein were correctly aligned and confidently modelled. Our region of interest, residue 395, was not included among the 38 residues aligned with PHYRE2 and observed with CHIMERA.

4.5.5 *BAG3* IGV visualization

Manual inspection of *BAG3* through the IGV tool, we confirmed the *BAG3*c.925C>T variant in the affected proband, her unaffected father, affected sister and deceased daughter (Figure 4.6).

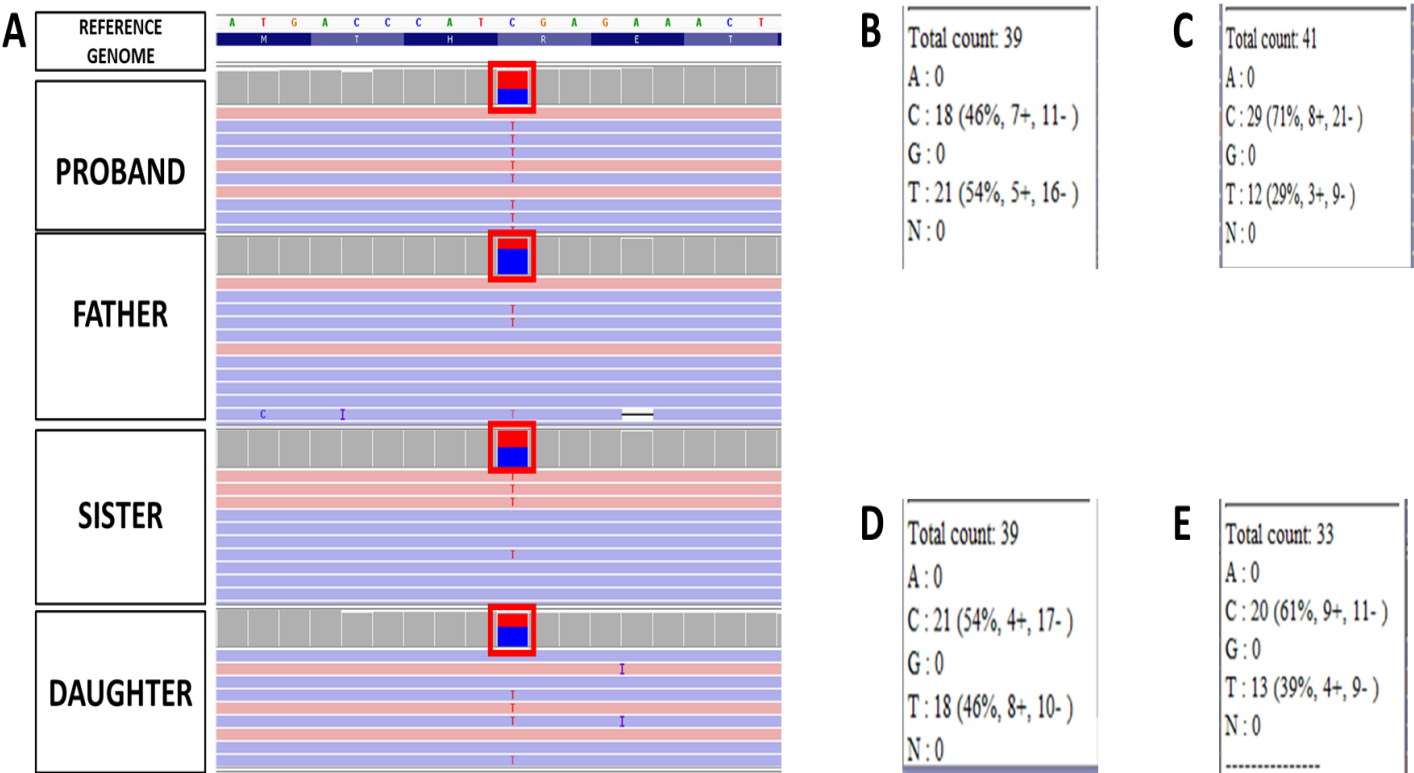


Figure 4.6: Visualization of the *BAG3* c.925C>T variant in the DCM343 family. A.) IGV shows a C>T change in *BAG3* at this locus in the proband, father, sister and daughter highlighted in red all with 99% confidence of base calling. B.) Allelic heterogeneity in the proband shows a total of 39 bases at this position with 18 C bases being 18 and T bases being 21 in number. C.) Allelic heterogeneity in the proband's father showing a total count of 41 bases with 29 of them being C and 12 being T. D.) Allelic heterogeneity in the proband's sister showing a total count of 39 bases with

21 of them being C and 18 being T E.) Allelic heterogeneity in the proband's daughter showing a total count of 33 bases with 20 of them being C and 13 being T.

#### 4.5.6 *BAG3* c.925C>T familial segregation

HRM analysis of the family was able to differentiate between the mutant *BAG3* of the proband (individual V:8) and the wildtype *BAG3*, in the unaffected (Figure 4.4). The results showed the proband shared this variant with her sister (individual V:9), father (individual IV:5) and daughter (individual VI:4). Sanger sequencing confirmed the presence of the shared *BAG3* c.925C>T variant in the proband (Figure 4.4). Additional family members screened included individual VI:5 (son to the proband's sister) who was found to harbour the *BAG3* mutation, however he has not been clinically evaluated yet. Individual IV:4 (the proband's paternal aunt) and her two daughters V:4 and V:6 was also found to be mutation positive while her son V:5 is mutation negative. Figure 4.7 shows the segregation pattern of the *BAG3* c.925C>T variant identified in this family. We were not able to clinically evaluate the mutation positive individuals due to the SARS-CoV-2 pandemic.

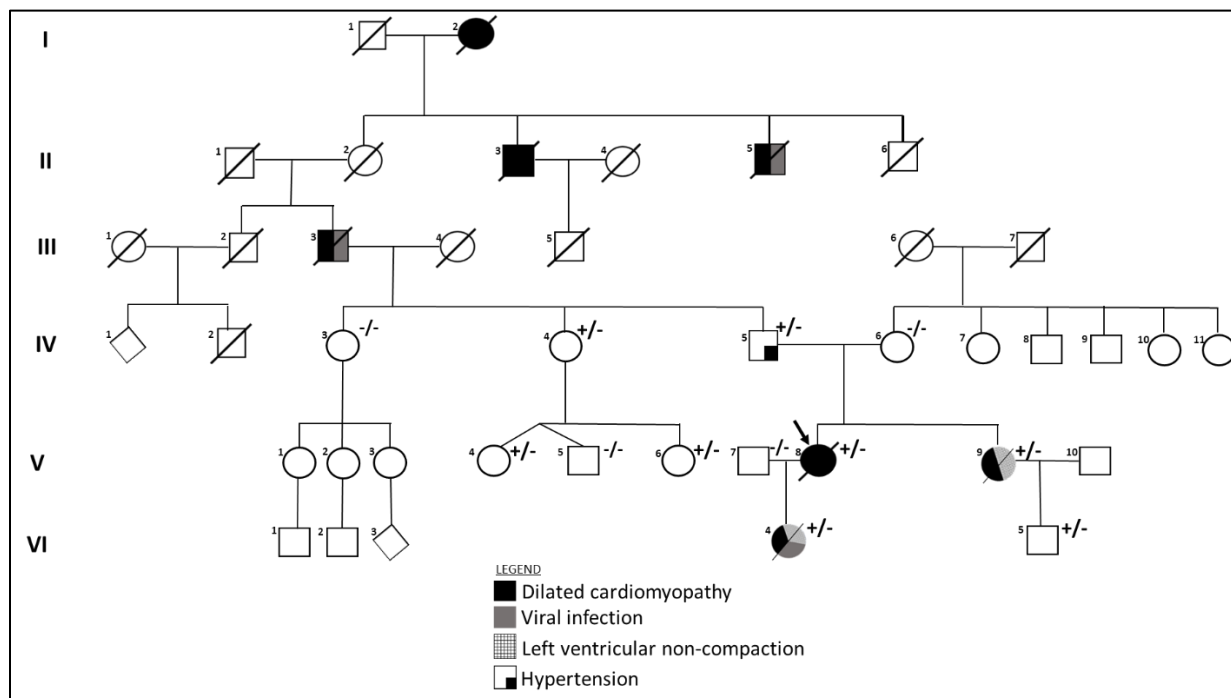


Figure 4.7: Segregation of *BAG3* c.925C>T (p.R309Ter) variant in family DCM343.

#### 4.5.7 *BAG3* targeted resequencing in an extended cardiomyopathy cohort

Targeted resequencing for 38 genes (Appendix K) was performed in the principal study called IMHOTEP, where we sequenced a cohort of 634 cardiomyopathy probands (DCM=479, HCM=69, ACM=78, RCM=39)(unpublished data). Targeted resequencing was done using the Haloplex Target Enrichment System Reagent Kit at the University of Manchester, Oxford Medical Laboratory (Chapter 2, section 2.8). We only extracted and analysed the sequencing data for *BAG3*, as the other 37 genes did not form part of this PhD study.

Analysis of this extended cardiomyopathy cohort identified the *BAG3* c.925C>T (p.R309\*) nonsense mutation in two unrelated probands, DCM181.1 and DCM307.1, both of these individuals were of mixed ancestry and was diagnosed at Groote Schuur Hospital.

DCM181.1 presented with HF symptoms at 59 years of age and was diagnosed with DCM (LVEDD 64mm, LVEF 18%). He subsequently had a transient ischaemic attack and was later diagnosed with diabetes and hypercholesterolaemia in his 70s. He initially showed improvement on medical therapy of both his HF symptoms and LV function (LVEF 38%) but at 1 year following diagnosis, he subsequently developed end-stage HF.

DCM307.1 presented with HF in the third trimester of her second pregnancy and the baby was subsequently delivered by caesarian section. She was initially diagnosed with PPCM with an acute dilated phenotype (LVEF 27%) and severe mitral regurgitation. Due to the severity of her presentation, she was referred for cardiac transplantation, however, she stabilized on medical therapy (LVEF 35%), despite being unable to tolerate beta-blockers due to underlying asthma. She was subsequently de-listed for transplantation.

#### 4.5.8 Modifiers

In the known cardiac gene list (cardiac genes and genes reported to be associated with cardiomyopathy) panel, we looked at all class 3 variants (variants of unknown significance) in order to determine if there were additional variants that could have a modifying effect on the phenotype. We identified the *RYR1* c.10616G>A(p.R3539H) variant (Table 4.3) and all other variants were excluded if they had a reported population frequency of >1% according to ExAC or gnomAD

The variant in *RYR1* c.10616G>A(p.R3539H) was only found in individual VI:4, the proband's deceased daughter; the *CRK* c.48\_49delinsCG(p.R17G) variant was only found in individual V:9 (proband's sister), the *NEB* c.16582C>G(p.D5528H) variant was found in individual VI:4 while the *NEB* c.10531C>A (p.A3511S) variant was found in individuals V:8 (proband), V:9 and VI:4. Segregation of the *RYR1* variant is shown in figure 4.8.

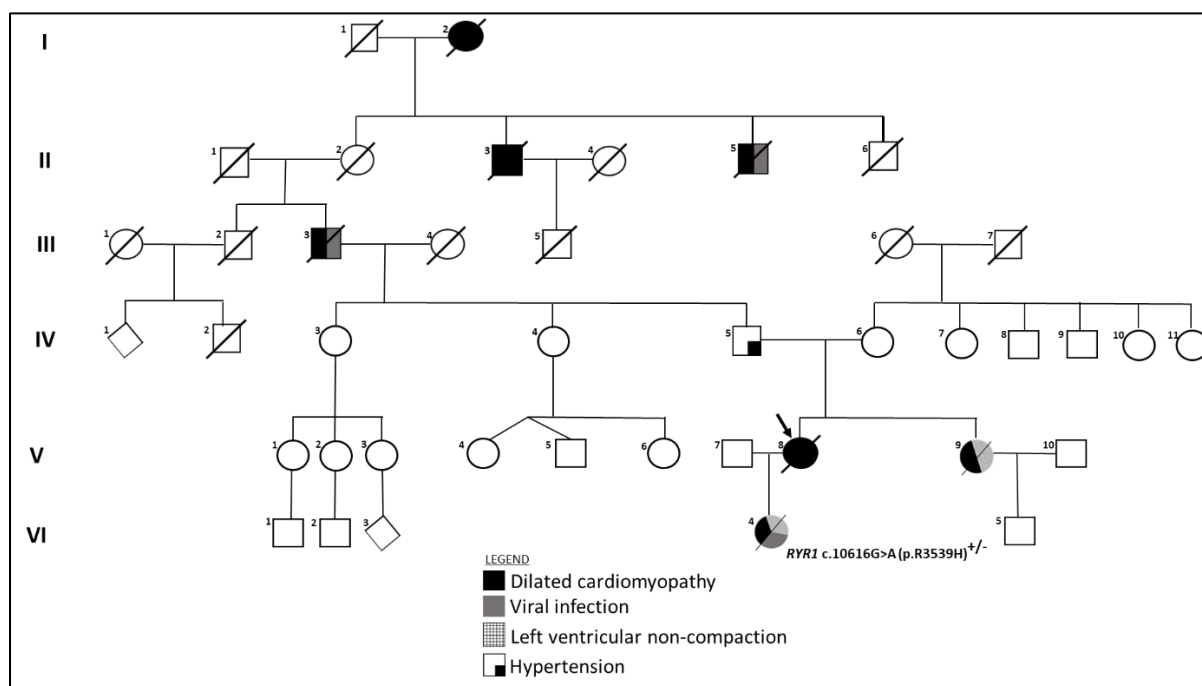


Figure 4.8: Segregation of class 3 variants identified in the *RYR1* cardiomyopathy gene.

In the novel cardiomyopathy genes list (possible novel genes), we identified one variant in the *CRK* gene and two variants in the *NEB* gene (Table 4.2). The *CRK* c.48\_49delinsCG(p.R17G) variant was only found in individual V:9 (proband's sister), the *NEB* c.16582C>G(p.D5528H) variant was found in individual VI:4 while the *NEB* c.10531C>A (p.A3511S) variant was found in individuals V:8 (proband), V:9 and VI:4. Segregation of the *CRK* and *NEB* variants are shown below in figure 4.8.

**Table 4.2: Modifier variants in family DCM 343**

| VARIANT IDENTIFIERS (NCP)             |                 |          |             | POPULATION FREQUENCIES |        |        |       | ACMG           | PATHOGENICITY PREDICTION |          |      |          | PROTEIN EXPRESSION |      |
|---------------------------------------|-----------------|----------|-------------|------------------------|--------|--------|-------|----------------|--------------------------|----------|------|----------|--------------------|------|
| Gene                                  | Variant         | Protein  | rsNumber    | ALFA                   | 1000G  | gnomAD | ExAC  | Classification | Varsome                  | MT       | SIFT | PP-2     | Genecards          | Gtex |
| <b>Known cardiac genes</b>            |                 |          |             |                        |        |        |       |                |                          |          |      |          |                    |      |
| <i>RYR1</i>                           | c.10616G>A      | p.R3539H | rs143987857 | 0.0024                 | 0.0006 | 0.002  | 0.002 | 4              | LP                       | DC_0.999 | D 0  | PD 0.999 | Yes                | 0.12 |
| <b>Candidate cardiomyopathy genes</b> |                 |          |             |                        |        |        |       |                |                          |          |      |          |                    |      |
| <i>CRK</i>                            | c.48_49delinsCG | p.R17G   | -           | -                      | -      | -      | -     | 3              | VUS                      | DC_0.999 | -    | -        | Yes                | 11   |
| <i>NEB</i>                            | c.16582C>G      | p.D5528H | rs201979610 | 0.003                  | 0.0002 | 0.002  | 0.003 | 3              | VUS                      | DC 0.999 | D 0  | PD 0.989 | Yes                | 0.22 |
| <i>NEB</i>                            | c.10531C>A      | p.A3511S | -           | -                      | -      | -      | -     | 3              | VUS                      | DC 0.999 | -    | PD 1     | Yes                | 0.22 |

B-Benign; DC- disease causing; D-Deleterious/damaging; MT-MutationTaster; PD-probably damaging; PL-Polymorphism; Polyphen-2; PS-possibly damaging; T-Tolerated; VUS-variant of unknown significance

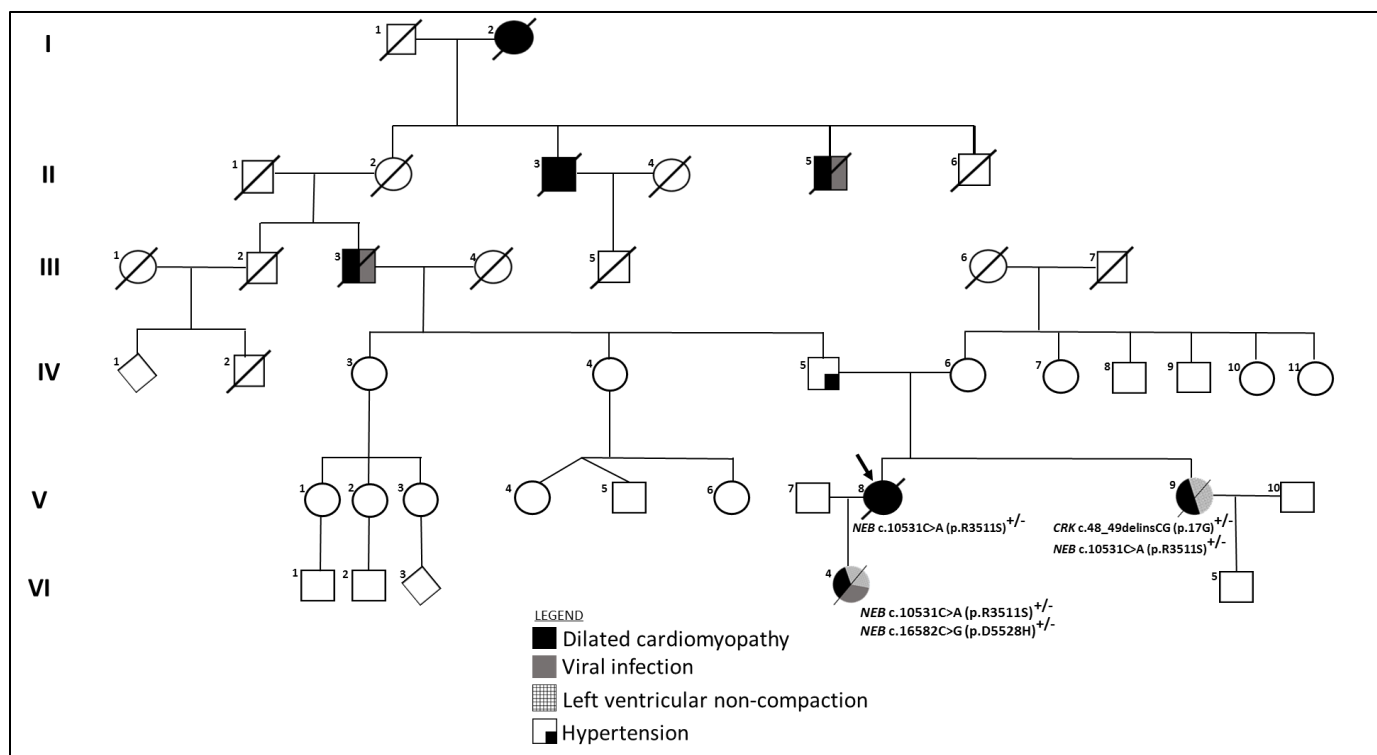


Figure 4.9: Segregation of variants identified in *CRK* and *NEB* genes.

## 4.6 Discussion

### 4.6.1 *BAG3* c.925C>T mutation

In this chapter we present a six-generation South African family with a generational history of HF characterised by a variable phenotype expression. We performed WES on four family members (IV:5, V:8, V:9 and VI:4) and found the *BAG3* c.925C>T variant to segregate with disease in all the affected family members.

*BAG3* codes for the Bcl-2-associated athanogene 3 protein (BAG3), which is 74kDa and 575 amino acid long and constitutes the BAG family of proteins.<sup>195</sup> The ATPase domain of the heat shock cognate (Hsc/Hsp) 70 was used as bait to originally discover *BAG3*.<sup>196</sup> The BAG family of proteins contains six member proteins and each family member contains at least 1 Bag domain that complexes to the ATPase domain of Hsc70/Hsp70 and is in proximity to the carboxy-terminal end

of the BAG protein.<sup>195,196</sup> Members of the BAG family are conserved and found in mammals (humans and mice), amphibians (*Xenopus laevis*), invertebrates (*Drosophila*, *Ciona intestinalis*, *Caenorhabditis elegans*), plants (*Arabidopsis thaliana*) and yeast (*Schizosaccharomyces pombe*, *Saccharomyces cerevisiae*). This evolutionary conservation highlights its significant biological role.<sup>197,200</sup>

In addition to the BAG domain, the BAG3 protein contains other domains which provide multiple binding sites for other proteins, thus pointing out the versatile capacity for BAG3 to participate in a range of major biological processes (Figure 4.10).<sup>195</sup>

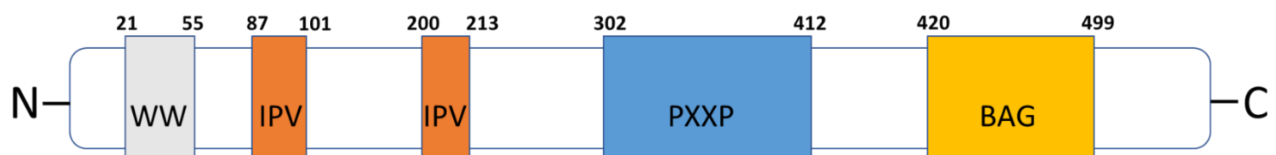


Figure 4.10: Structure of BAG3 protein and its binding domains WW, IPV, PXXP and BAG.<sup>195</sup>

The WW domain interacts with the PXXP region to alter BAG's three-dimensional structure thus ensuring correct conformation of the protein for its functions. From the WW domain, BAG3 forms the chaperone assisted selective autophagy (CASA) together with Hsp70 and HspB8. CASA is an autophagy pathway that is induced by tension and is crucial in muscle and immune cells for mechanotransduction. This autophagic pathway contributes to the cellular autophagic processes of maintaining protein homeostasis.<sup>201,202</sup> The two highly conserved IPV (Ile-Pro-Val) domains facilitate the binding of BAG3 to the heat shock proteins HspB6 and HspB8 to aid in the catabolic sequestration of cytoplasmic material in the autophagosome, and degraded after being shuttled to lysosomes.<sup>203</sup>

Isolation and concentration of substrates from other cytoplasmic components is needed for the autophagy of aberrant proteins and this sequestration and concentration takes place in the aggresome, which has elevated levels of autophagy.<sup>204</sup> This aggresome-autophagy pathway also contributes to standardised cell activity and protein homeostasis through the selective removal of damaged, misfolded, undegraded or ubiquitinated proteins.<sup>204</sup> Various studies have pointed towards the direct localisation of BAG3 and have shown that a downregulation of *BAG3* expression drastically decreased aggresome maturation thus underpinning the involvement of *BAG3* in protein equilibrium through the aggresome.<sup>205,206</sup> Accumulation of misfolded proteins and other protein untoward outcomes, termed as proteotoxicity, perturbs protein homeostasis which is implicated in the pathogenesis of many heart diseases, including DCM.<sup>207</sup>

The BAG domain confers protective properties to cardiomyocytes by coupling with the antiapoptotic protein Bcl-2 to inhibit cardiomyocyte apoptosis. Cardiomyocyte apoptosis is one of the contributing factors leading to LV dilation and HF which are pathological hallmarks for DCM.<sup>201,208</sup> The PXXP domain also binds to heat shock cognate 71 kDa protein (Hsc70) initiating a complex formation to reinforce myofibrillar composition with the actin capping protein, CapZ, in response to mechanical stress which is essential for the stability of the cardiac sarcomere. BAG3 ensures this proper interaction by influencing and enhancing the proper localisation and stabilisation of a *CapZ* protein, CapZ $\beta$ 1, to the Z-disk where it maintains the structural integrity of filamentous actin (F-actin) which is a contractile protein that aids in cardiac muscle function.<sup>195,196</sup>

Mutations in *BAG3* are responsible for about 2% of DCM cases and also appear to be associated with the variation in LVEF, LV end diastolic volume (LVEDV) and LV end systolic volume (LVESV).<sup>201</sup>

In the family under investigation, the *BAG3* mutation had a truncating effect and this mutation

affects the PXXP domain which is found in exon 4 of the BAG3 protein (Figure 4.10). Studies done to establish *BAG3* as a significant factor in the development and progress of HF have shown that in mice with a *BAG3* homozygous aberration, they displayed non-inflammatory myofibrillar degeneration which characterises fulminant myopathy, apoptotic attributes, Z-disk architectural disruption, during the initial post-natal period and died within a month.<sup>209</sup> BAG3 enhances the protein-protein interaction of Hsp70 and CapZ $\beta$ 1 binding without involving any other components and CapZ $\beta$ 1 prevents the dissociation of actin into actin monomers by binding to the barbed end of actin. CapZ $\beta$ 1 also links neighbouring sarcomeres, positions actin filaments in the z-disk by interacting with a protein nebulin and stabilises the Z-disk.<sup>210,212</sup> The role of BAG3 in increasing the distribution of CapZ $\beta$ 1 has been shown in *BAG3* homozygous deficient mice which display mislocalisation and low-level expression of CapZ. Reduced or no functional CapZ $\beta$ 1 promotes rapid destabilisation of myofibrillar structures under mechanical stress which leads to a degenerative disease state.<sup>212,213</sup>

BAG3 also stabilises the myofibrillar Z-disk by regulating the level of functional filamin, which is a dimeric actin that has its role of being a signalling hub for various proteins. This filamin clearance and regulation together with stabilisation of the sarcomere appears to occur through the previously mentioned CASA pathway and knockdown of this pathway leads to fragmentation of the Z-disk with accompanying pathologic manifestations in cardiac and skeletal muscle. The CASA pathway is stress related and one of its predominant substrates in cells undergoing mechanical stress is the filamin protein. Filamin transcription is stimulated by BAG3, when BAG3 uses its WW domain to bind to transcriptional activators.<sup>214</sup> Collectively, this filamin homeostatic balance contributes to get rid of misfolded and degraded proteins in the myocyte.<sup>195</sup>

The significant role of *BAG3* in the heart also extends to its modulation of cardiac contractility and reducing cardiac arrhythmias.<sup>195</sup> In neonatal rat ventricular cardiomyocytes, it has been shown that *Bag3* is significant in initiating a cardiac signal because a *Bag3* reduction led decreased signals in amplitude and frequency. In addition, knockdown of *Bag3* led to an adverse effect on synchrony of conduction, activation time and degree of connectivity, thus showing how *Bag3* is significant for signal propagation throughout the myocardium. On the other hand, overexpression of *Bag3* improved the activity of cardiomyocytes with respect to degree of connectivity and electrical signal amplitude. The deficiency of *Bag3* was further revealed to be linked to the appearance of cardiomyocyte aggregates that are self-reliant.<sup>201</sup> This highlights the important role *BAG3* plays in the regulation of cardiac electrophysiological activity through affecting the coupling of myocyte excitation-contraction.

The *BAG3* c.925C>T variant reported in our South African family is a previously reported nonsense mutation which was found in three French-Canadian DCM families.<sup>215</sup> In our family, carriers of this mutation displayed phenotypic variability and part of this variability may be due to LV reverse remodelling following treatment: V:8 showed improved cardiac function and her LVEF went from 20% to almost normal following HF therapy. Upon the death of her daughter (VI:4), the proband discontinued the prescribed medication and subsequently died. The proband's father, IV:5 is mutation positive but showed no signs of the disease, despite the penetrant nature of the variant in the family. It was noted that he was on antihypertensive medication to treat his high blood pressure. A similar observation was made by Chami et al, who reported that in one family in their study, the proband's 67 year-old father was mutation positive for the *BAG3* mutation but did not manifest DCM phenotype.<sup>215</sup> What was interesting was that he was also being treated for hypertension with angiotensin receptor blockers (ARBs).<sup>215</sup> ACE-is and ARBs are interchangeable

for the treatment of hypertensive patients and are similar in their effectiveness in lowering mortality in HF.<sup>216</sup> We hypothesise that this protection, in both cases, might be due to the ACE inhibitors the probands' fathers were being treated with.

Also noted in this family, was that viral/stressful events appear to precede the early onset of DCM. A previous study by Belkaya et al reported that 16% of children with acute myocarditis harbored rare biallelic nonsynonymous variants in *BAG3* and other genes implicated in cardiomyopathy.<sup>217</sup> They hypothesised that an anomaly in *BAG3* increases the risk of an individual to develop left ventricular dysfunction preceding a stressful episode. This work substantiates the observations made within our family and opens other avenues to explore such as administering  $\alpha$ -adrenergic receptor antagonists to mitigate the effects of disease onset and stress among patients with *BAG3* mutations.<sup>218</sup>

What is certain is that *BAG3* is gaining momentum in the cardiac research community, and in the future, we hope to address the fundamental questions in order to realise the utility of laboratory research in the clinical and therapeutic fields.<sup>218</sup>

#### 4.6.2 *BAG3* extended cardiomyopathy cohort screen

The extended cardiomyopathy cohort of 634 probands were screened using targeted resequencing and we found the previously identified *BAG3* c.925C>T (p.R309\*) in two unrelated South African probands with primary DCM. The possibility of a founder effect for this *BAG3* mutation is being explored as South Africa has a very colourful history and a unique genetic composition arising from Dutch exploration as from the late 16th century resulting in carriers emigrating into new regions and being founders of genetic risks as seen in South Africa where various heritable diseases have been attributed to early Dutch settlers.<sup>219,220</sup> This has resulted in

South Africa having a high frequency of founder effect variants. This constitutes the first report of multiple probands with a pathogenic *BAG3* variant in a South African cardiomyopathy cohort.

#### 4.6.3 Modifiers

In this family, the phenotypic variability amongst the individuals harbouring the *BAG3* c.925C>T variant led us to postulate that one of the reasons for this observed variability is the existence of genetic modifier/s in the family which would affect the manifestation of the primary *BAG3* c.925C>T mutation. We then extended our filtering criteria to find any variants in the ACMG class 3-5 categories (GVUS, LP, P)<sup>49</sup>, as potential genetic modifiers. We first explored the known cardiac panel but all the variants identified were occurred in the *PDE4DIP* gene and had population frequencies greater than 1%. The variants were excluded from further investigation as they were not considered to be variants of interest under the scope of genetic modifiers (Table 4.2). We then investigated variants in the novel candidate gene panel (possible novel genes) and found four variants of interest: *RYR1* c.10616G>A(p.R3539H), *CRK* c.48\_49delinsCG(p.R17G), *NEB* c.16582C>G p.D5528H and *NEB* c.10531C>A p.A3511S.

##### 4.6.3.1 *RYR1* c.10616G>A(p.R3539H)

The *Ryanodine Receptor 1* variant (*RYR1*), *RYR1* c.10616G>A(p.R3539H), was only found in VI:4 and has been previously reported (accession number rs143987857) in relation to malignant hyperthermia.<sup>221</sup>

There are only a few studies on *RYR1* and they have been done to address any cardiovascular defects due to abnormal forms of *RYR1* and their correlation with cardiovascular disease.<sup>222</sup> A

GWAS of LVH among Korean patients documented the association between mutations in *RYR1* and LVH.<sup>223</sup> Because genes involved in the cardiomyopathy spectrum have been shown to cause more than one cardiomyopathy, depending on the variant under observation, it is possible other *RYR1* mutations can lead to ventricular dilatation as observed in DCM.<sup>67</sup> It is noteworthy that the lack of a clear link between *RYR1* mutations and DCM is supported by the fact that *RYR1* is predominantly expressed in the skeletal muscle while its isoform *RYR2* is expressed in cardiac muscle. This means that rather than cause an overt DCM phenotype, patients with cardiac dysfunction will exhibit reduced exercise capacity due to the resultant skeletal muscle defect brought about by anomalies in *RYR1*.<sup>224,225</sup>

Even though the *RYR1* discovered is a class 4 variant (likely pathogenic), heart failure due to *RYR1* variants needs further exploration to support the modifying role of *RYR1* c.10616G>A(p.R3539H) in the HF manifested by individual VI:4.

#### 4.6.3.2 *CRK* c.48\_49delinsCG(p.R17G)

The *Proto-Oncogene, Adaptor Protein (CRK)* variant, *CRK* c.48\_49delinsCG variant, was only found in V:9. The Crk adaptor protein in mice has been studied and postulated to function in growth regulation, cell adhesion, cell migration, signal transduction and cell transformation. *Crk*<sup>-/-</sup> mice were generated and the muscular wall of their hearts was thinner and had fewer loosely packed cells. There was severe dilation in the embryonic heart ventricles of *Crk*<sup>-/-</sup> animals.<sup>226</sup> This thinning of the heart and dilatation of ventricular walls reported in the mice study above, mirrors the hallmark features of DCM which include ventricular dilatation that can translate to the extremely poorly developed and thin muscular wall of the heart pointing towards the possibility of *CRK* variant identified above exacerbating the cardiac phenotype caused by the *BAG3* variant.

The association of *CRK* mutations with poorly developed myocardium and thin-walled heart muscular walls points towards our identified *CRK c.48\_49delinsCG(p.R17G)* variant as a possibly acting as a modifier to the poor heart development as evidenced by non-compaction of the heart and reduced heart function of 42% possibly due to reduced wall thickness of member V:9. This gene and variant warrants further investigation.

#### 4.6.3.3 *NEB c.16582C>G p.D5528H*

The *Nebulin* (*NEB*) variant, *NEB c.16582C>G*, was only found in individual VI:4 and has been previously reported during genotype-phenotype correlations for nemaline myopathy with an accession number of rs201979610.227 Nemaline myopathy (NM) is a rare kind of congenital myopathy characterized by hypomyotonia and a varied degree of muscle weakness with the skeletal muscle mostly affected. It has been recognized that patients with a mutation in the myopalladin (*MYPN*) gene have DCM or HCM in addition to slow and progressive muscle weakness.<sup>228,229</sup>

Both *NEB* and *BAG3* work to stabilise the sarcomeric Z-disk which is vital for function of the cardiac muscle through ensuring the proper localization of CapZB1 which is responsible for positioning of sarcomeric actin filaments and linking adjacent sarcomeres.<sup>210,212,230</sup> Individual VI:4 had severe DCM compared to the other *BAG3* mutation positive members and the interaction of *NEB* with a known DCM causing *BAG3* gene, possibly supports the modifier role our identified class 3 *NEB* variant in exacerbate the observed DCM observed. This gene and variant warrants further investigation.

#### 4.6.3.4 *NEB* c.10531C>A p.A3511S

The *NEB* c.10531C>A variant was found in V:8, V:9 and VI:4 and hasn't been reported before. Even though this variant is a class 3 variant, all the evidence reporting the association of *NEB* and *BAG3* at the cardiac sarcomere suggest that this variant could possibly act as a modifier that could result in a variable cardiac phenotype due to the DCM phenotype diagnosed in the carriers who harbours the shared *BAG3* mutation.<sup>230</sup> This gene and variant warrants further investigation.

## 4.7 Conclusion

As a first for South Africa, we report a known *BAG3*c.925C>T stop variant that is responsible for DCM in a multi-generational South African family. We also report the first South African cohort screen for *BAG3*, where multiple DCM probands have been identified as having the same pathogenic *BAG3* variant. We hypothesise that this could be a founder variant however, this requires further investigation through haplotyping.

In addition to the *BAG3*c.925C>T, identified four variants that could act as possible modifiers. In the future, it will be important to study the modulating effect of ACE-is and/or ARBs on *BAG3*c.925C>T and potentially patients with pathogenic *BAG3* mutations as this could be used to personalise treatment regimens.

## CHAPTER 5 – INVESTIGATION OF *PKP2* AS THE DISEASE-CAUSING VARIANT IN THIS SOUTH AFRICAN ACM FAMILY

### 5.1 Introduction

The development of NGS technologies has broadened the molecular genetics of ACM through the rapid discovery of disease causing variants and led to the consensus that it is classically an autosomal dominant disease of the desmosome where cell adhesion occurs with heterozygous mutations in the desmosomal gene *PKP2* leading to aberrant splicing or a truncated protein being the most common.<sup>231-235</sup> Biallelic forms of inheritance have also been observed at a varied inheritance from 4% to 47% and these biallelic carriers have been reported to have a more severe arrhythmic phenotype in most studies.<sup>234,236,237</sup> Non desmosomal including *CDH2*, *RYR2*, *TGFB3*, *CTNNA*, *TTN*, *PLN*, *DES* and *LMNA*, have also been found in regulatory proteins and intermediate filaments and have been shown to cause the ACM in up to 10% of cases.<sup>48,238-241</sup>

In one of the most intensive ACM investigations involving 439 index cases, no pathogenic desmosomal variant was found in 37% of the cohort with only one-fifth of the individuals without an identifiable variant having evidence of familial disease.<sup>242</sup> In a recent meta-analysis, they also confirmed the observation that ACM cases without anomalies in the desmosomal genes have a lower prevalence of family history.<sup>243</sup> It remains to be determined if the observations of low prevalence of family disease among desmosomal-gene negative ACM cases is due to a primary monogenic form of the disease or whether their phenotypic observation represents an oligogenic inheritance characterised by low-penetrant gene variants and/or external factors that are either environmental or innate interacting to manifest the disease state.<sup>244</sup> It is to be noted though that

the cases without an identified variant and no family history were mostly seen in high-level endurance athletes which likely places exercise as a risk factor in such individuals.<sup>244</sup> ACM genetics is purported to extend beyond the monogenic inheritance due to reports of some patients having more than a single gene affect their phenotypic.<sup>51</sup> This myocardial disease is usually inherited in an autosomal dominant manner with variable penetrance.<sup>245,19,152</sup> Carrier relatives exhibit heterogenous phenotypes leading to familial being clinically underdiagnosed in many instances.<sup>140,246,247</sup>

In South Africa, no large-scale genetic studies have been conducted on ACM and the few studies coming from Africa are vastly inadequate and highlight an urgent need for this void to be filled in order to understand the genetic basis and offer effective treatment options to the patients. SSA could provide some very important insights into ACM and identify other possible disease mechanisms which might shed some light on the different pathways that lead to ACM.<sup>56</sup>

Further investigations are still needed to elucidate the molecular mechanisms underlying ACM pathogenesis to explain the cases without an identifiable cause and these can be extended into areas including disease models and in vivo drug testing.

In this chapter, we investigate a three-generation South African family diagnosed with ACM.

## 5.2 Aim

The aim of this study was to identify the disease-causing variant in a family with severe ACM and any variants that might modify the observed phenotype.

.

## 5.3 Objectives

1. To use the WES data to identify the disease-causing variants in a South African family with ACM.
2. To validate candidate variants any variants that modify the ACM phenotype in the family using HRM and Sanger sequencing.

## 5.4 Methods

### 5.4.1 Whole exome sequencing

Whole exome sequencing protocols and parameters are described in Chapter 2, section 2.4.

### 5.4.2 Targeted resequencing

Targeted resequencing was done using the Haloplex Target Enrichment System Reagent Kit at the University of Manchester, Oxford Medical Laboratory (Chapter 2, section 2.8). Targeted resequencing for 38 genes (Appendix K) was performed in the principal study called IMHOTEP where we sequenced a cohort of 634 cardiomyopathy probands (DCM=479, HCM=69, ACM=78, RCM=39)(unpublished data), which included the proband of ACM145 (III:3). We only extracted and analysed the sequencing data for this proband (ACM 145, III:3), as the other probands did not form part of this PhD study.

## 5.5 Results

### 5.5.1 Clinical history of the ACM145 family

A three-generation Caucasian South African family of Caucasian origin included a proband and his brother diagnosed with ACM at ages 19 and 23 years, respectively. Clinical evaluation and diagnosis were undertaken at the Cardiac Clinic at GSH. The proband (III:1) presented with cardiac arrest during a competitive sports event at age 19 years, requiring defibrillation for ventricular fibrillation. He was subsequently diagnosed with ACM, with major structural and repolarisation criteria, and minor depolarisation criteria. CMR showed a severely dilated RV (RVEDV/BSA 205ml/m<sup>2</sup>) with severely impaired function (RVEF 14%) and regional wall motion abnormalities, and marginally reduced LVEF of 50%. An implantable cardioverter defibrillator (ICD) was inserted, with subsequent shocks for ventricular tachycardia (VT). It was also noted that the proband had been diagnosed with thyrotoxicosis prior to the diagnosis of ACM.

Family screening revealed that the proband's sibling (III:2) was asymptomatic but clinically affected with dilated RV (RVEDV/BSA 125ml/m<sup>2</sup>) with reduced function (RVEF 41%) and LGE on CMR; he also had late potentials on SAECG and an increased PVC burden (9.65 %) on 24-hour Holter. The proband's parents were both screened and neither fulfilled the Task Force Criteria for ACM, however, the proband's father (II:4) had minor depolarisation changes on SAECG and LGE in the absence of structural or functional abnormalities on CMR, which may be early signs of disease. The proband's mother (II:3) had no cardiac abnormalities noted on electrocardiogram, SAECG, echocardiogram or CMR. No skin or hair abnormalities were noted in either the proband or the affected sibling. There was no reported family history of sudden cardiac death, cardiomyopathy or HF (Figure 5.1).

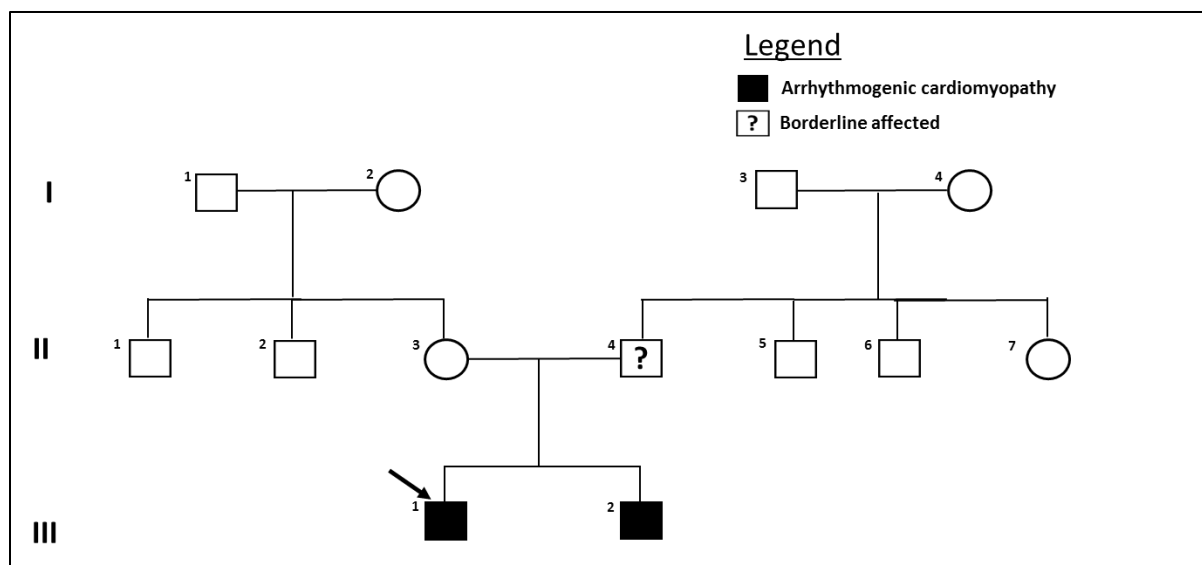


Figure 5.1: Pedigree of family ACM145. Squares represent males, circles represent females. Shaded symbols represent affected individuals. Arrow points towards the proband.

### 5.5.2 Whole exome sequencing of combined affected individuals in ACM145

As clinical phenotyping indicated that both brothers were affected with ACM, we analysed the data of both brothers to determine if they had a pathogenic variant in common that could explain their disease.

For WES data analysis, we filtered for variants occurring in both affected brothers (III:1 and III:2) respectively, using (1) variants that occurred in a panel of 188 genes known to cause or be associated with cardiomyopathy (Appendix J), and (2) variants that occurred in novel genes that are expressed in the heart. In these two scenarios, we explored small-scale genetic variants such as point variants (silent, missense or nonsense), deletions and insertions and we also attempted to determine the presence of CNVs within the affected individuals. The variants of interests were then validated through HRM and Sanger sequencing. The segregation pattern of the identified variants was then matched to the family members to determine a possible genotype-phenotype correlation of the variants to explain the disease described in the family.

### 5.5.2.1 Known cardiac panel screen

We performed WES on four family members, which included the proband (III:1), the proband's affected brother (III:2), his mother (II:3) and his father (III:4). WES data analysis for the known cardiomyopathy genes, for the two affected siblings, found 69,441 variants shared between the two. Filtering through gnomAD for variant frequency, identified 99 rare variants (<1% frequency), 25 of which were predicted to have a moderate or high impact on protein sequence (Figure 5.2).

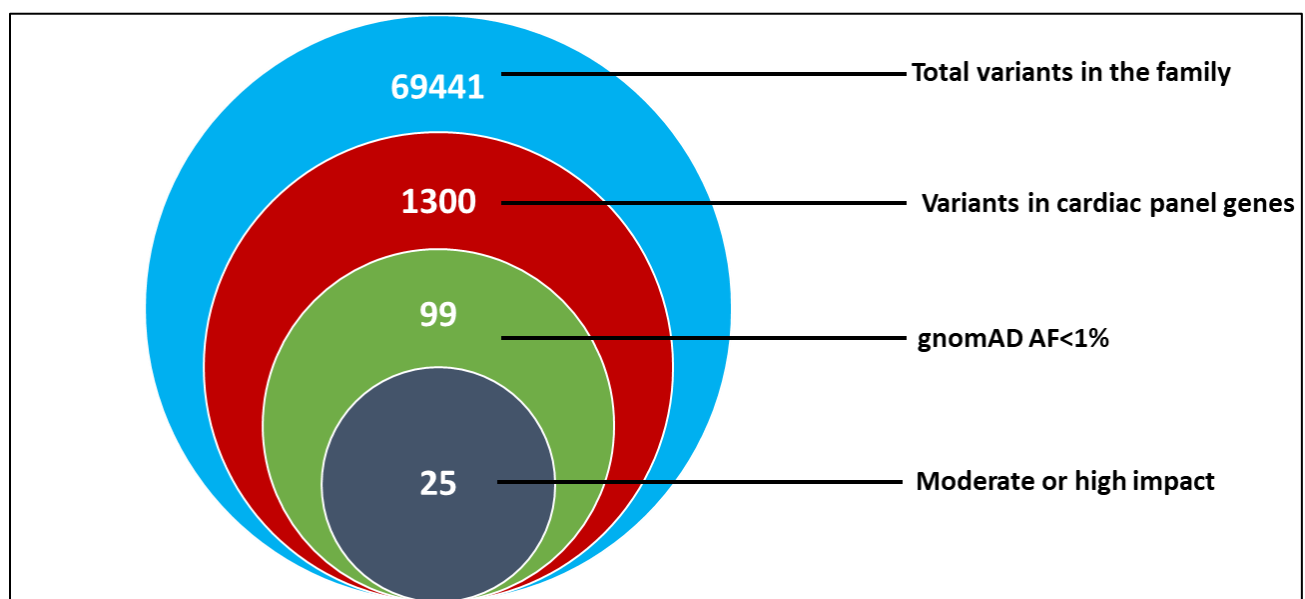


Figure 5.2: Filtering for WES data from the cardiac panel for ACM145 family.

For variants that did not have a frequency in the gnomAD databases, additional population frequency databases were used to confirm their frequencies include ALFA, 1000G and ExAC. The 25 variants comprised of 3 x class 1 variants, 17 x class 2 variants and 5 x class 3 variants. All the class 3 variants were predicted to have a frequency of >1% according to additional population frequency databases used (Table 5.1) and were excluded as variants of interest.

**Table 5.1: List of rare variants identified in ACM 145 family from the cardiac panel**

| VARIANT IDENTIFIERS (NCP) |                   |             | POPULATION FREQUENCIES |        |        |           |        | PATHOGENICITY PREDICTION |          |        |          |
|---------------------------|-------------------|-------------|------------------------|--------|--------|-----------|--------|--------------------------|----------|--------|----------|
| Gene                      | Variant           | Protein     | rsNumber               | ALFA   | 1000G  | gnomAD    | ExAC   | Varsome                  | MT       | SIFT   | PP-2     |
| <i>SVIL</i>               | c.5138C>T         | p.T1713I    | rs143178190            | 0.001  | 0.002  | 0.001176  | 0.001  | LB                       | PL 0.999 | D 0.05 | B 0.067  |
| <i>PDE4DIP</i>            | c.5730C>A         | p.D1910E    | rs1613780              | 0.5    | -      | -         | 0.5    | -                        | PL 0.999 | T1     | B 0      |
| <i>PDE4DIP</i>            | c.5599C>T         | p.R1867C    | rs1620560              | 0.28   |        | -         | 0.5    | -                        | DC 0.938 | D 0.01 | PD 0.989 |
| <i>PDE4DIP</i>            | c.5269G>A         | p.A1757T    | rs1628310              | 0.5    | -      | -         | 0.5    | -                        | PL 0.999 | T 0.51 | B 0.031  |
| <i>PDE4DIP</i>            | c.4793A>G         | p.H1598R    | rs1778155              | 0.4    | -      | -         | 0.4    | LB                       | PL 0.999 | T 0.86 | B 0.007  |
| <i>PDE4DIP</i>            | c.4511G>A         | p.R1504Q    | rs2762875              | 0.5    | 0.4    | -         | 0.5    | -                        | PL 0.999 | T 1    | B 0      |
| <i>PDE4DIP</i>            | c.4360A>G         | p.K1454E    | rs1778120              | 0.3    | -      | -         | 0.3    | US                       | PL 0.999 | T 0.29 | B 0.159  |
| <i>PDE4DIP</i>            | c.4075A>G         | p.K1359E    | rs1747958              | 0.36   | -      | -         | 0.5    | -                        | PL 0.999 | T 0.12 | B 0.115  |
| <i>PDE4DIP</i>            | c.3796A>G         | p.K1266Q    | rs12568796             | 0.26   | -      | -         | 0.26   | US                       | DC 0.693 | D 0.01 | PD 0.998 |
| <i>PDE4DIP</i>            | c.3196G>A         | p.A1066T    | rs1698647              | 0.26   |        | -         | 0.5    | -                        | PL 0.693 | D 0    | PD 0.849 |
| <i>PDE4DIP</i>            | c.3037T>A         | p.F1013I    | rs1698624              | 0.5    |        | -         | 0.5    | -                        | PL 0.999 | T 1    | B 0      |
| <i>PDE4DIP</i>            | c.2042G>A         | p.R681H     | rs1629011              | 0.5    |        | -         | 0.5    | -                        | PL 0.908 | T 0.09 | B 0.444  |
| <i>PDE4DIP</i>            | c.1864C>T         | p.R622Ter   | rs1778111              | 0.28   | 0.3    | -         | 0.4    | -                        | DC 1     | -      | -        |
| <i>PDE4DIP</i>            | c.1607G>C         | p.S536T     | rs1747930              | 0.5    | 0.5    | -         | 0.5    | -                        | DC 0.855 | D 0.02 | PD 0.958 |
| <i>PDE4DIP</i>            | c.1445A>G         | p.H482R     | rs1698681              | 0.26   |        | -         | 0.5    | -                        | PL 0.999 | T 0.18 | B 0.02   |
| <i>PDE4DIP</i>            | c.1229A>T         | p.Q410V     | rs1061308              | 0.5    |        | -         | 0.5    | -                        | DC 0.991 | D 0.01 | PD 0.973 |
| <i>PDE4DIP</i>            | c.824C>T          | p.S275L     | rs1359300              | 0.5    |        | -         | 0.5    | -                        | PL 0.995 | T 0.15 | B 0.031  |
| <i>PDE4DIP</i>            | c.74G>T           | p.R25L      | rs1664022              | 0.5    |        | -         | 0.5    | -                        | DC 0.999 | D 0.01 | PD 0.999 |
| <i>PDE4DIP</i>            | c.16A>G           | p.T6A       | rs111958438            | 0.001  | 0.003  | 0.002521  | 0.003  | LB                       | DC 0.999 | T 0.68 | B 0      |
| <i>KCNE1</i>              | c.112A>G          | p.S38G      | rs1805127              | 0.7    | 0.6    | -         | 0.7    | B                        | PL 0.999 | T 0.44 | B 0      |
| <i>TTN</i>                | c.47077G>A        | p.V15693I   | rs201717871            | 0.0007 | 0.0008 | 0.0007973 | 0.0008 | LB                       | DC 0.999 | -      | -        |
| <i>TTN</i>                | c.40498G>T        | p.V13500F   | rs201944202            | 0.0006 | 0.0005 | 0.0005967 | 0.0007 | LB                       | PL 0.926 | -      | -        |
| <i>TTN</i>                | c.33513_33515dup  | p.Q11172dup | rs368327166            | 0.006  | -      | -         | 0.006  | B                        |          | -      | -        |
| <i>TTN</i>                | c.20335A>T        | p.S6779C    | rs149470241            | 0.001  | 0.0002 | 0.0006025 | 0.0006 | LB                       | PL 0.799 | -      | -        |
| <i>LAMA4</i>              | c.848_849delinsAC | p.A283D     | rs71543223             | -      | -      | -         | -      | LB                       | PL 0.799 | T 1    | B 0      |

B; Benign; DC- disease causing; D-Deleterious/damaging; PD-probably damaging; PL-Polymorphism; PS-possibly damaging; T-Tolerated; VUS-variant of unknown significance

### 5.5.2.2 Novel candidate gene screen

When no variants were identified in the known cardiac gene panel, we explored the possibility of novel candidate genes.

A total of 68,141 variants were identified in genes that have not yet been associated with heart disease and 24,229 variants were found to be shared amongst the two affected brothers. Through gnomAD filtering, 4,761 variants were reported to have a frequency of <1% in the populations. These 4,761 variants were filtered to identify 489 variants which had a moderate to high impact on the protein sequence. The 489 variants were prioritised using the bioinformatics tools Toppgene, ENDEAVOUR and VarElect.

The top 25 variants from prioritisation were analysed and filtered according to the various filtering criteria mentioned in Chapter 2, section 2.7 (Figure 5.3).

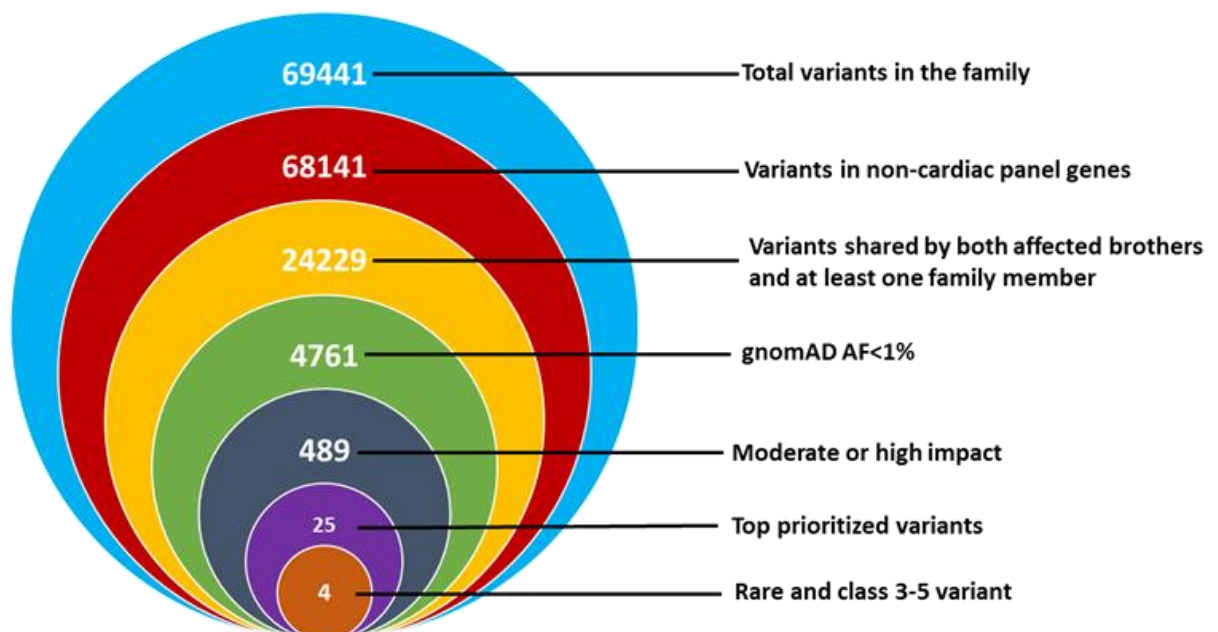


Figure 5.3: Filtering for WES data from the non-cardiac gene panel for ACM145 family.

After filtering, only 19 variants were shown to be expressed in the heart (Table 5.2). Of these, 15 variants were filtered out because they were classified as benign or likely benign according to ACMG variant classification or their variant effect was tolerated according to pathogenicity prediction. The remaining four variants were identified to be rare and classified according to the ACMG classification as either pathogenic (class 5), likely pathogenic (class 4) or a variant of unknown significance (class 3). The four variants were heterozygous *CFTR* c.1727G>C(p.G576A), heterozygous *CFTR* c.2002C>T(R668C), heterozygous *PKD1*c.6749C>T(p.T2250M) and homozygous *LRP1*c.9927\_9928del(p.G3311R fs). These four variants were further analysed through variant validation (Table 5.2).

**Table 5.2: Filtering for WES data from the novel candidate genes for ACM145 family**

| VARIANT IDENTIFIERS (NCP) |                   |                 |              | POPULATION FREQUENCIES |       |        |       | ACMG CLASSIFICATION | PATHOGENICITY PREDICTION |                |          |          | PROTEIN EXPRESSION |      |
|---------------------------|-------------------|-----------------|--------------|------------------------|-------|--------|-------|---------------------|--------------------------|----------------|----------|----------|--------------------|------|
| Gene                      | Variant           | Protein         | rsNumber     | ALFA                   | 1000G | gnomAD | ExAC  |                     | Varsome                  | MutationTaster | SIFT     | Polyphen | Genecards          | Gtex |
| <i>PLEC</i>               | c.7678G>A         | p.A2560T        | rs193257576  | 0.005                  | 0.001 | 0.003  | 0.003 | 2                   | LB                       | P 0.999        |          | B 0      | YES                | 159  |
| <i>VWF</i>                | c.8333G>A         | p.R2778Q        | rs150577615  | 0.0003                 |       | 0      | 0     | 2                   | LB                       | P 0.999        | T 0.78   | B 0      | YES                | 2    |
| <i>NOS2</i>               | c.3017G>T         | p.R1006L        | rs542585205  | 0                      | 0     | 0      | 0     | 2                   | LB                       | P 0.999        | D 0.04   | B 0.007  | YES                | 3    |
| <i>HLA-DRB1</i>           | c.752G>A          | p.R251K         | rs71547382   | 0.03                   | -     | 0.001  | 0.04  | 2                   | LB                       | P 0.898        | T 0.17   | B 0      | YES                | 13   |
| <i>PKD1</i>               | c.6749C>T         | p.T2250M        | rs139971481  | 0.003                  | 0.001 | 0.002  | 0.003 | 4                   | US/P                     | D 0.02         | PD 0.995 | D 0.02   | YES                | 0.35 |
| <i>KCNJ12</i>             | c.44C>T           | p.S15L          | rs1657738    | 0.25                   | -     | -      | -     | 1                   | US/B                     | DC 0.999       | D 0.04   | B 0.179  | YES                | 4    |
| <i>KCNJ12</i>             | c.167A>C          | p.E56Ala        | rs1714865    | -                      | -     | -      | -     | 1                   | US/B                     | DC 0.999       | T 0.81   | B 0.065  | YES                | 4    |
| <i>14KC NJ12</i>          | c.353_354delinsAC | p.R118H         | -            | -                      | -     | -      | -     | 3                   | US                       | P 0.999        | T 0.24   | B 0.005  | YES                | 4    |
| <i>KCNJ12</i>             | c.467C>T          | p.P156L         | rs1714864    | 0.5                    | -     | -      | 0.5   | 3                   | US                       | DC 0.999       | D 0.04   | PD 1     | YES                | 4    |
| <i>LRP1</i>               | c.9927_9928del    | G3311ArgfsTer26 | -            | -                      | -     | -      | -     | 4                   | P                        | DC 0.61        |          |          | YES                | 0.55 |
| <i>CREB BP</i>            | c.5836C>G         | p.P1946A        | rs1321085895 | 0                      | -     | -      | -     | 2                   | LB                       | P 0.999        | T 0.4    | B 0      | YES                | 0.73 |
| <i>CFTR</i>               | c.1727G>C         | p.G576A         | rs1800098    | 0.006                  | 0.002 | 0.005  | -     | 3                   | US                       | D 0.997        | T 0.15   | PD 0.639 | YES                | 3    |
| <i>CFTR</i>               | c.2002C>T         | p.R668C         | rs1800100    | 0.008                  | 0.006 | 0.006  | -     | 4                   | LP                       | DC 0.997       | D 0      | PD 1     | YES                | 3    |
| <i>DLC1</i>               | c.762_763delinsCG | p.N255delinsHD  | rs372924164  | -                      | -     |        | -     | 3                   | US                       | P 0.999        |          |          | YES                | 46   |
| <i>KRT14</i>              | c.188_189delinsAT | p.C63Y          | rs386797102  | -                      | -     | -      | -     | 1                   | B                        | P 0.999        | T 0.29   | B 0      | YES                | 251  |
| <i>PDCD6IP</i>            | c.1147G>A         | p.V383I         | -            | -                      | -     | -      | -     | 1                   | B                        | P 1            | T 0.06   | B 0.038  | YES                | 25   |
| <i>CACNA2D4</i>           | c.1927G>A         | p.D643N         | rs367750825  | 0                      | 0     | 0      | 0     | 2                   | LB                       | DC 0.796       | T 0.38   | PD 0.544 | YES                | 0.1  |
| <i>HIP1R</i>              | c.970G>C          | -               | -            | -                      | -     | -      | -     | 2                   | LB                       | P 0.999        | T 0.058  |          | YES                | 0.12 |

B; Benign; DC- disease causing; D-Deleterious/damaging; PD-probably damaging; PL-Polymorphism; PS-possibly damaging; T-Tolerated; VUS-variant of unknown significance

### 5.5.2.3 Variant validation

The *CFTR* c.1727G>C(p.G576A), *CFTR* c.2002C>T(R668C), *PKD1*c.6749C>T(p.T2250M) and *LRP1*c.9927\_9928del(p.G3311R fs) variants from the non-cardiac panel were validated through HRM and Sanger sequencing.

#### 5.5.2.3.1 *CFTR* c.1727G>C

HRM analysis of the *CFTR* variant in the four family members displayed different HRM profiles in the proband (III:2), proband's father (II:4) and proband's mother (II:3). Sanger sequencing of the samples revealed that the proband and his affected brother(III:1) were heterozygous for the *CFTR* c.1727G>C variant, the father was homozygous mutant, and the mother was homozygous wildtype (Figure 5.4)

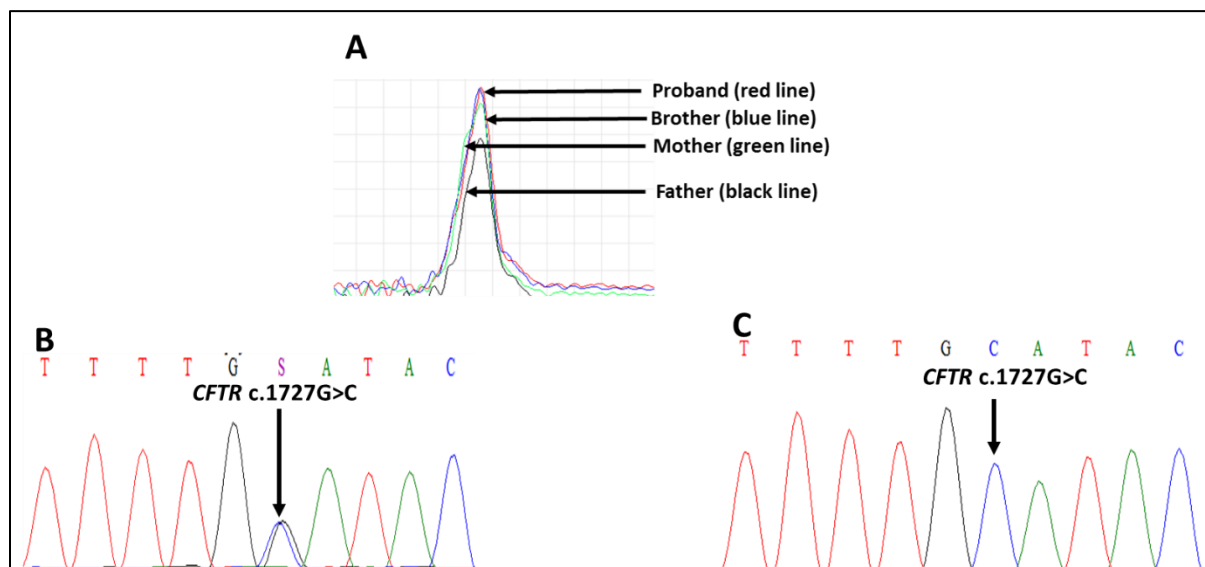


Figure 5.4: **Validation analysis of *CFTR* c.1727G>C variant in family ACM145** A) Melt-curve profiles of the family members. B) Sanger sequencing showed heterozygous G>C change in the proband and his brother. The presence of the two peaks indicates a C (blue) and a G (black) nucleotide at this locus. C) Sanger sequencing of the proband's dad revealed a homozygous G>C change as shown by the black arrow.

When we looked at the conservation of the *CFTR* p.G576A residue, we found that the *CFTR* c.1727G>C variant affected a highly conserved amino acid across various species (Figure 5.5)

|               |     |                  |
|---------------|-----|------------------|
| H. sapiens    | 576 | DADLYLLDSPFGYLDV |
| Pteroglodytes | 576 | DADLYLLDSPFGYLDV |
| M. mulatta    | 576 | DADLYLLDSPFGYLDV |
| M. musculus   | 576 | DADLYLLDSPFGYLDV |
| Ggallus       | 578 | DADLYLLDSPFGYLDV |

Figure 5.5: Amino acid conservation of the CFTR protein across multiple species.

Investigation of the secondary structural changes caused by the *CFTR* c.1727G>C (p.G576A) variant, through PHYRE2, showed that residues 1 to 406 of the mutant protein were correctly aligned and also included our region of interest (residue 576). Visualisation by Chimera identified the *CFTR* c.1727G>C(p.G576A) in the variant sequence(5.10A) when compared to the wild type (Figure 5.6).

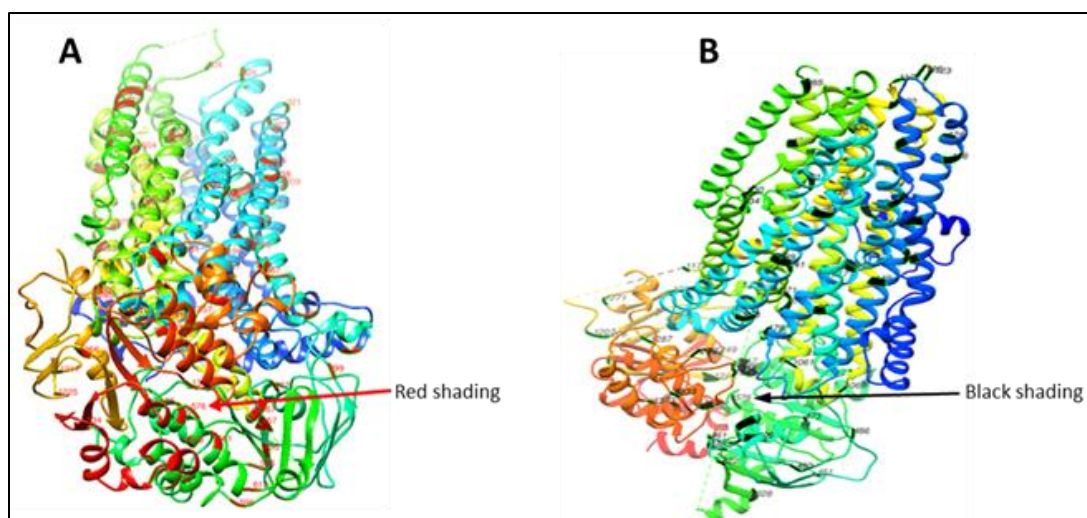


Figure 5.6: Chimera visualization of the variant identified at residue 576 of CFTR.

Site directed mutator (SDM) database predicted the outcome of the *CFTR* c.1727G>C(p.G576A) variant to cause a reduction in the stability of the resultant protein and damaging (Table 5.3).

**Table 5.3: SDM prediction outcome of *CFTR* c.1727G>C(p.G57A) variant**

| PDB file | Chain ID | Mutation                       | Predicted ddg | Outcome           |
|----------|----------|--------------------------------|---------------|-------------------|
| 6msm.pdb | A        | <i>CFTR</i> c.1727G>C (p.576A) | -0.33         | Reduced stability |

#### 5.5.2.3.2 *CFTR* c.2002C>T

HRM analysis and Sanger sequencing results of this *CFTR* variant revealed that the variant was heterozygous in the brothers, homozygous mutant in the father and wildtype in the mother (Figure 5.7).

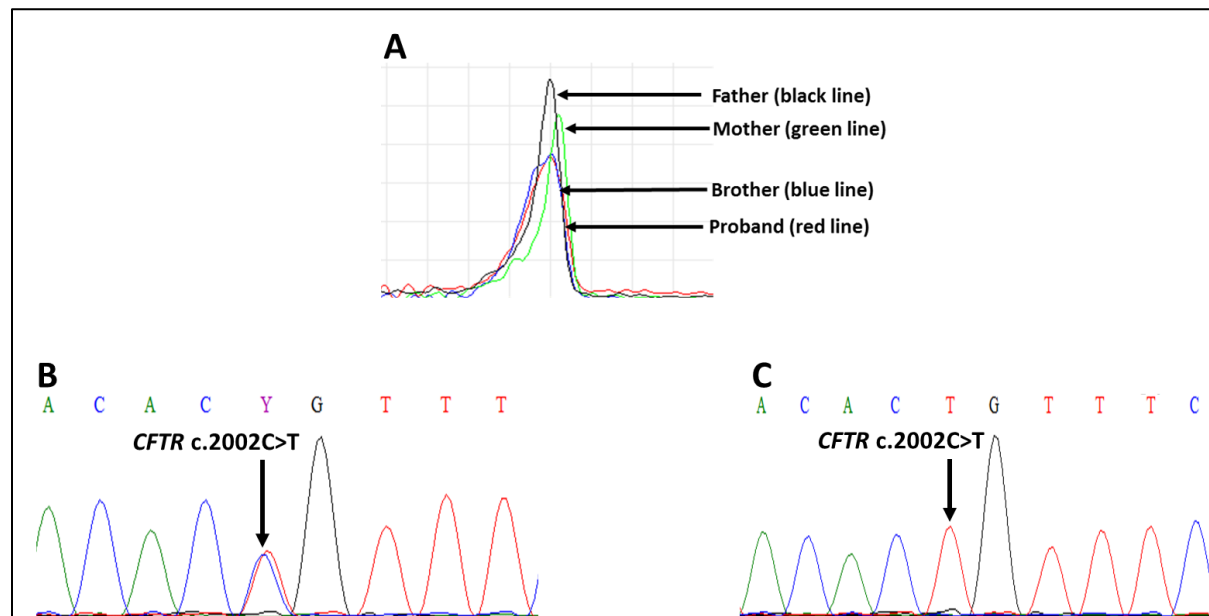


**Figure 5.7: Validation analysis of *CFTR* c.2002C>T variant in family ACM145** A.) Melt-curve profiles of the family members. B.) Sanger sequencing showed heterozygous C>T change in the proband and his brother. The presence of the two peaks indicate a C (blue) and a T (red) nucleotide at this locus. C.) Sanger sequencing of the proband's dad revealed a homozygous C>T change as shown by the black arrow.

When we looked at the conservation of the *CFTR* protein, we found that the *CFTR* c.1727G>C variant affected a highly conserved amino acid across various species (Figure 5.8).

|               |     |                                                                 |
|---------------|-----|-----------------------------------------------------------------|
| H. sapiens    | 668 | ILTETLH <span style="border: 2px solid red;">R</span> FSLEGDAPV |
| Pteroglodytes | 668 | ILTETLRRFSLEGDAPV                                               |
| Mmulatta      | 668 | ILTETLRRFSLEGDAPV                                               |
| Fcatus        | 648 | ILTETLRRFSLEGDAAV                                               |
| Mmusculus     | 668 | ILTETLRRFSVD-DSSAP                                              |
| Ggallus       | 670 | ILTETLRRFSFEGESMG                                               |
| Trubripes     | 667 | ILTETLRRVSVDETAGF                                               |
| Drerio        | 667 | ILTETLRRVSVDESAGM                                               |
| Xtropicalis   | 669 | ILTETLRRCSIDS                                                   |

Figure 5.8: Amino acid conservation of the CFTR protein across multiple species.

Secondary structural predictions for the *CFTR* c.2002C>T(R668C) variant to did not reveal any structural changes as the region of interest was not covered when alignment was done by PHYRE2. Investigation of CFTR amino acid 3D structure through Chimera revealed that our region of interest was among a number of 209 amino acids not included in the CFTR's PDB file designated for 3D visualisation (Figure 5.9).

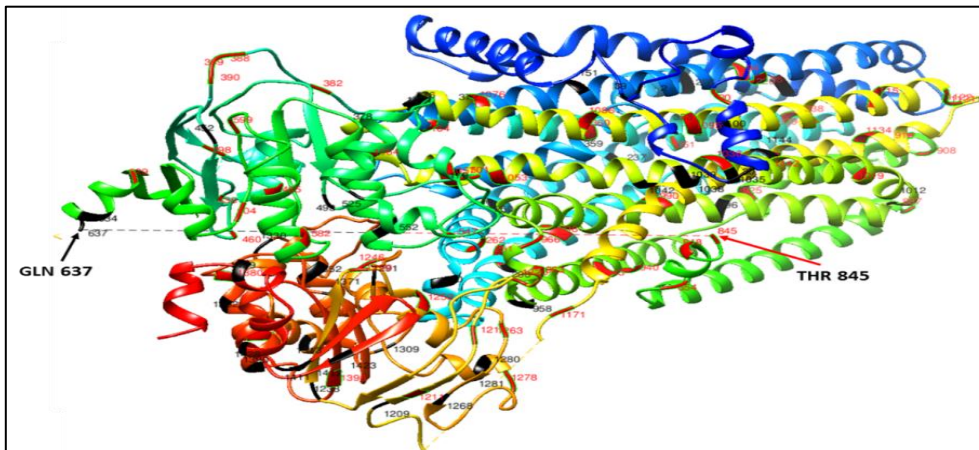


Figure 5.9: Visualization of CFTR's PDB file showed that our region of interest was not covered during alignment. The residues between Gln637 and Thr845 contain our residue of interest at position 668 but the 209 residues are undetected.

#### 5.5.2.3.3 *PKD1* c.6749C>T

HRM analysis of the *PKD1* variant in the family members showed that all the four family members had similar HRM melt profiles. Sanger sequencing identified a wild-type C variant in all the family

members instead of the predicted *PKD1* c.6749C>T variant in the two brothers and father as determined by WES (Figure 5.10). Sanger sequencing thus detected a false positive variant, and this variant was removed as a variant of interest.

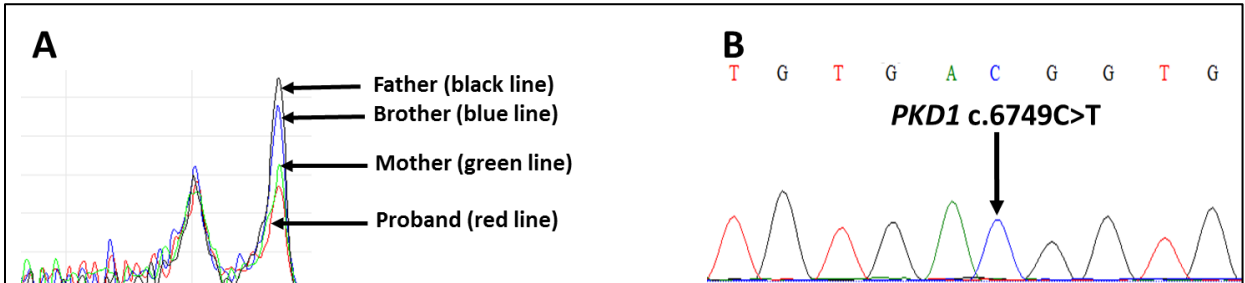


Figure 5.10: **Validation analysis of *PKD1* c.6749C>T variant in family ACM145** A.) Similar melt-curves among all family members. B.) Sanger sequencing of *PKD1* c.6749C>T in family ACM145.

When we looked at the conservation of the PKD1 protein, we found that the *PKD1* c.6749C>T variant affects a highly conserved amino acid across various species (Figure 5.11)

|             |      |                   |
|-------------|------|-------------------|
| H. sapiens  | 2250 | TPLTQSIQANVTVAPER |
| Mmusculus   | 2246 | TPLARSIQANVTVAAER |
| Ggallus     | 2204 | TPLSKSIQANVTMIPS  |
| Xtropicalis | 2165 | TPLSRNSLANVTMMPS  |

Figure 5.11: Amino acid conservation of the PKD1 protein across species.

Secondary structural alignment done by PHYRE2 did not reveal any structural changes as the region of interest was absent amongst all the 120 aligned sequences done with high confidence (>90%) as reported by PHYRE2 for this gene. Observation of the PKD1 PDB file through Chimera identified highly confident aligned sequences from residues TYR 3075 to PRO4120 (Figure 5.12)

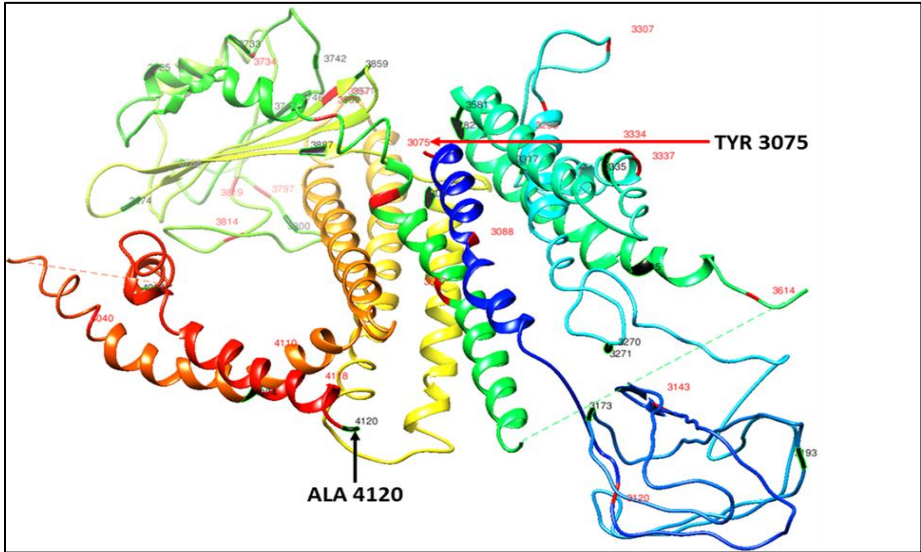


Figure 5.12: Residue coverage of PKD1 with successful 3D coverage with Chimera.

UNIPROT database revealed that the *PKD1* gene only has two PDB files available that cover residues 275-354 and 3049-4169 of the protein but not our residue of interest at position 2250 (Figure 5.13).

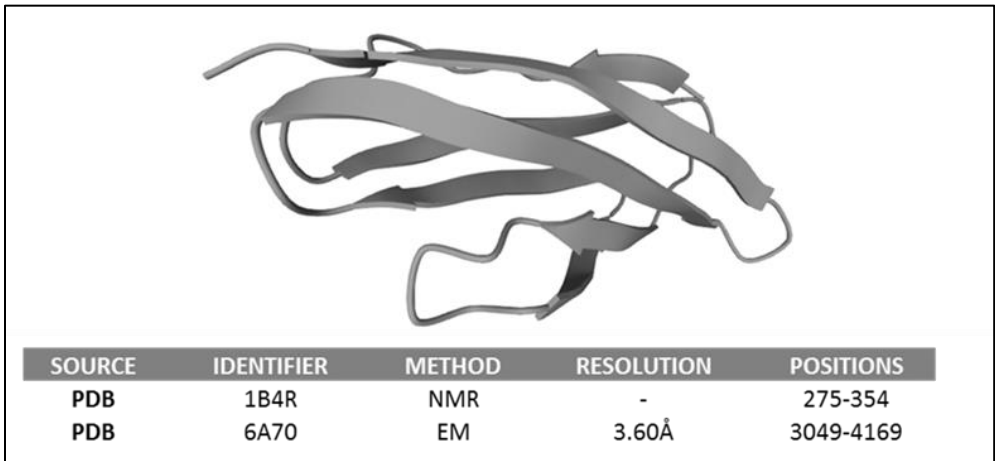


Figure 5.13:Residue coverage of PKD1 with successful 3D coverage as reported by UNIPROT.

There was no PDB file available that contained our region of interest for SDM to correctly assess the effect of protein structural change due to *PKD1* c.6749C>T(p.T2250M) residue.

5.5.2.3.4 *LRP1* c.9927\_9928delCG

HRM analysis of the *LRP1* variant in family ACM145 showed that all the four family members had similar melt profiles. Sanger sequencing showed wild-type CG in all the family members instead of the predicted CG deletion (Figure 5.14) Sanger sequencing thus detected a false positive variant, and this variant was removed as a variant of interest.

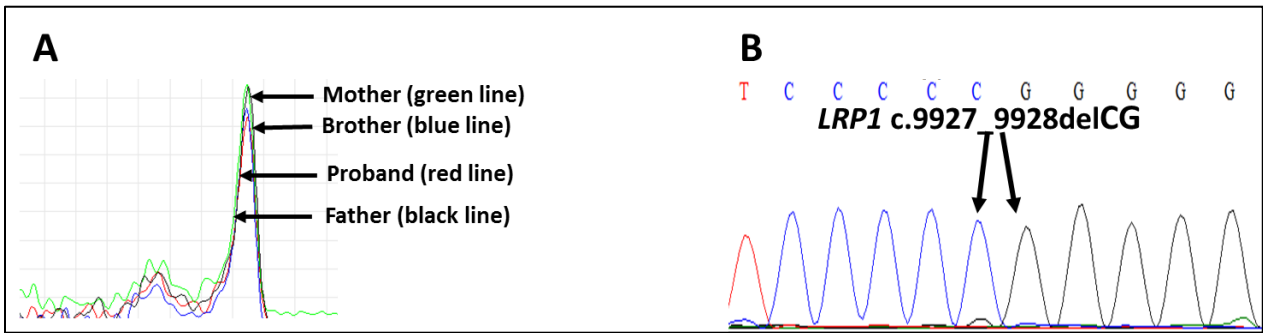


Figure 5.14: **Validation analysis of *LRP1* c.9927\_9928delCG in family ACM145** A.) Similar melt-curve among family members. B.) Sanger sequencing analysis of *LRP1* c.9927\_9928delCG in family ACM145.

When we looked at the conservation of the *LRP1* G3311R fs residue, we found that the *LRP1* c.9927\_9928delCG variant affects a highly conserved amino acid across various species (Figure 5.15)

|               |      |                  |
|---------------|------|------------------|
| Hsapiens      | 3308 | CSNLCLLSFGGGHKCA |
| Pteroglodytes | 3175 | CSNLCLLSFGGGHKCA |
| Fcatus        | 3307 | CSNLCLLSFGGGHKCA |
| Mmusculus     | 3309 | CSNLCLLSFGGGHKCA |

Figure 5.15: Amino acid conservation of the *LRP1* protein across multiple species.

Secondary structural alignment done by PHYRE2 did not reveal any structural changes as our region of interest was absent amongst all the 120 aligned sequences done with high confidence (>90%) as reported by PHYRE2 for this gene. Observation of the *LRP1* PDB file through Chimera identified highly confident aligned sequences from residues ALA933 to PHE1581 (Figure 5.16)

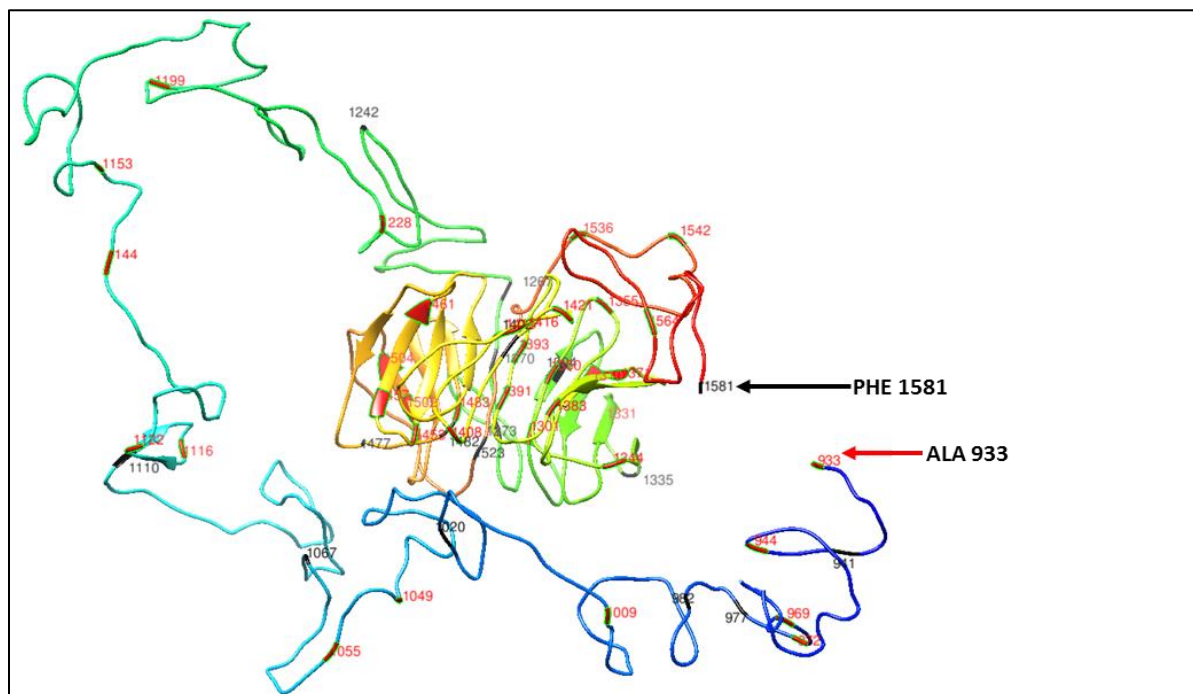


Figure 5.16: Residues coverage of LRP1 with successful 3D coverage with Chimera.

UNIPROT database revealed that the *LRP1* gene has seven PDB files available that do not cover our residues of interest at amino acid position 3311 of the protein (Table 5.4).

**Table 5.4: Residues coverage of *PKD1* with successful 3D coverage as reported by UNIPROT**

| SOURCE | IDENTIFIER | METHOD | RESOLUTION | CHAIN | POSITIONS   |
|--------|------------|--------|------------|-------|-------------|
| PDB    | 1CB        | NMR    |            | A     | 1059 - 1100 |
| PDB    | 1D2L       | NMR    |            | A     | 851 - 893   |
| PDB    | 1J8E       | X-ray  | 1.8Å       | A     | 1011 - 1054 |
| PDB    | 2FYJ       | NMR    |            | A     | 932 - 1013  |
| PDB    | 2FYL       | NMR    |            | A     | 932 - 1013  |
| PDB    | 2KNX       | NMR    |            | A     | 2770 - 2817 |
| PDB    | 2KNY       | NMR    |            | A     | 2770 - 2817 |

There was no PDB file available that contained our region of interest for SDM to correctly assess the effect of protein structural change due to the deletion discovered in *LRP1* after WES analysis.

Segregation of the variants identified is shown in figure 5.17.

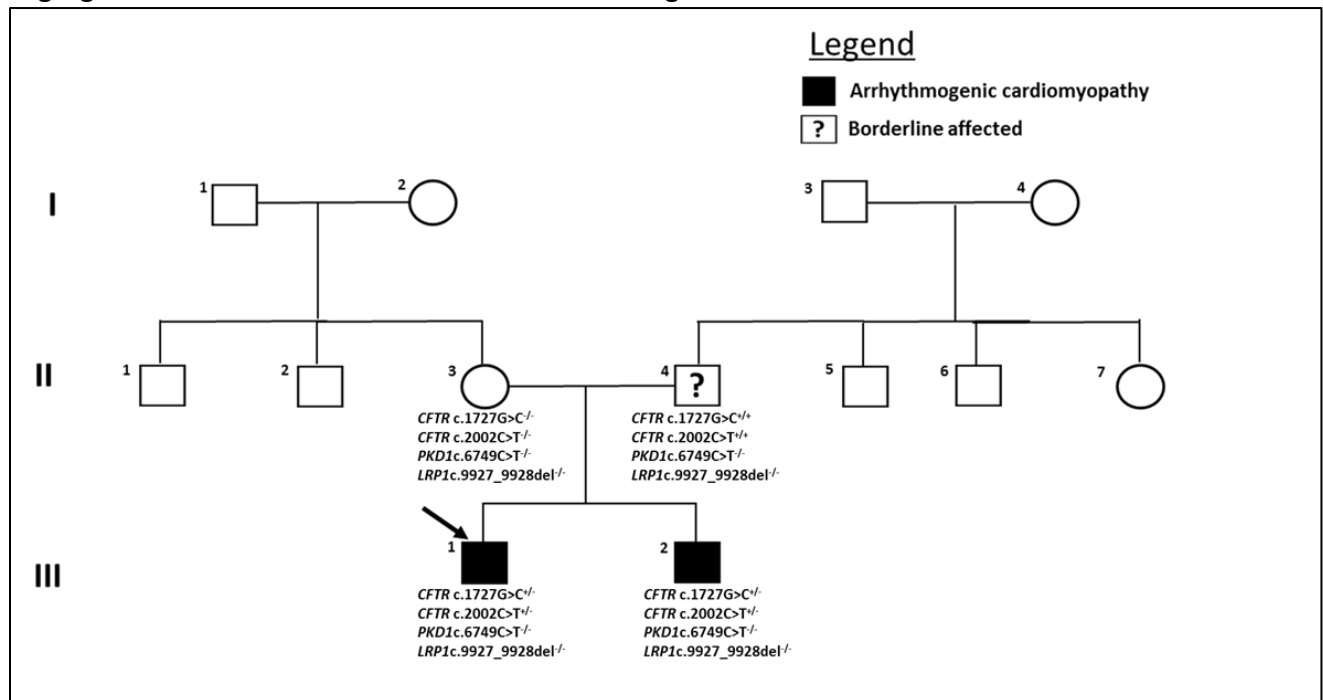


Figure 5.17: Pedigree of variants identified in family ACM145.

### 5.5.3 Whole exome sequencing of individual affected members of ACM145

As clinical phenotyping indicated that both brothers were affected with ACM, we analysed the data of both brothers to determine if they had a pathogenic variant in common that could explain their disease. The initial data analysis did not identify any common variant between the brothers. We then performed individual data analysis, followed by trio analysis, for each of the brothers to determine if there were any pathogenic variants (de novo or known) that could explain the disease.

### 5.5.3.1 Proband

For the proband (III:2), we identified 38785 variants from WES with 733 of them being variants from the known cardiac panel. We applied the various filtering criteria described in Chapter 2, section 2.7. Filtering using variant minor allele frequency resulted in 146 variants, of which 28 were predicted to have high or moderate impact on the resultant protein (Figure 5.18).

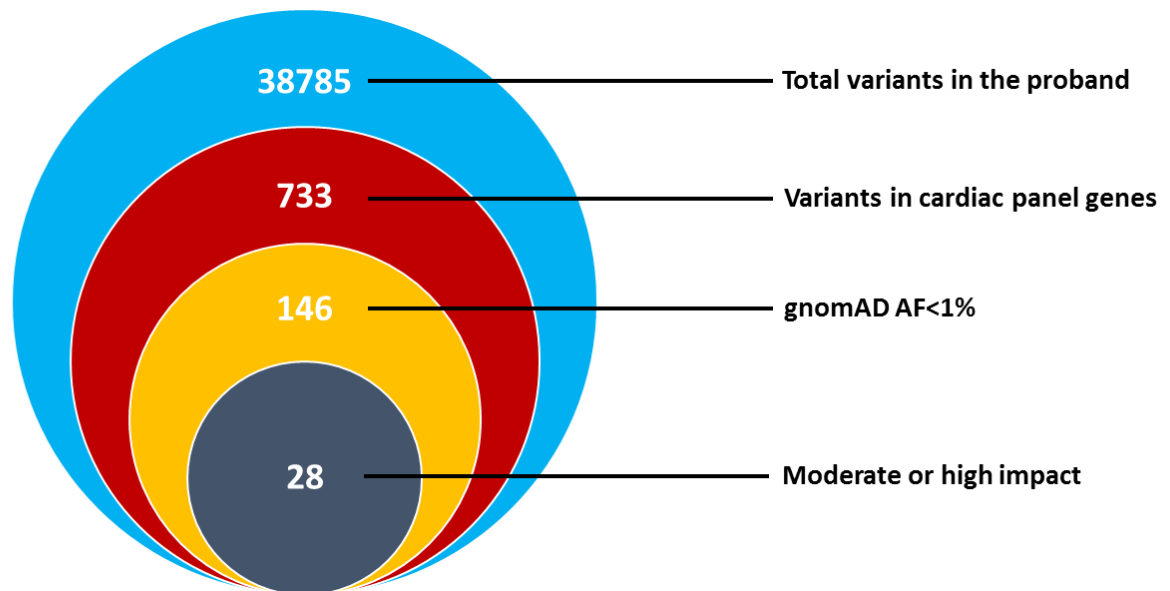
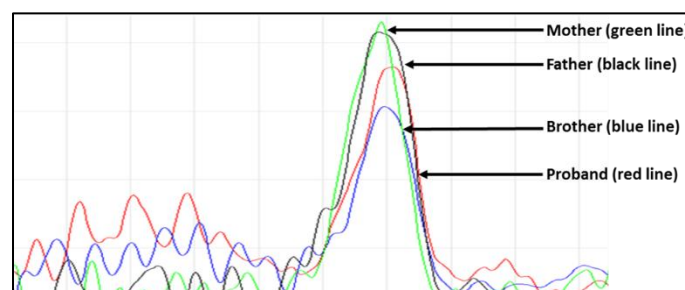


Figure 5.18: Filtering for WES data from the proband alone ACM145 family.

These 28 variants are described in Table 5.5. Twenty-seven variants were excluded as being (1) they all had a MAF of >1% according to the population databases or (2) they were synonymous. The only remaining variant of interest that met the criteria was the *ALPK3* c.2701C>T(p.Q901Ter) variant. The novel *ALPK3* c.2701C>T(p.Q901Ter) variant was identified in exon 6 of the *ALPK3* gene in the proband (III:2) of ACM145. The variant is classified as an ACMG class 4, likely



pathogenic variant. The variant was confirmed through HRM and segregation analysis found the proband's father to be positive for the variant as well (Figure 5.19).

Figure 5.19: Validation HRM analysis of *ALPK3* c.2701C>T variant in family ACM145.

**Table 5.5: Filtering for WES data in the proband**

| VARIANT IDENTIFIERS (NCP) |                     |           |             | POPULATION FREQUENCIES |       |        |       | ACMG CLASSIFICATION | PATHOGENICITY PREDICTION |          |         |          | PROTEIN EXPRESSION |      |
|---------------------------|---------------------|-----------|-------------|------------------------|-------|--------|-------|---------------------|--------------------------|----------|---------|----------|--------------------|------|
| Gene                      | Variant             | Protein   | rsNumber    | ALFA                   | 1000G | gnomAD | ExAC  |                     | Varsome                  | MT       | SIFT    | Polyphen | Genecards          | Gtex |
| <i>SVIL</i>               | c.5138C>T           | T 1713 I  | rs143178190 | 0.001                  | 0.002 | 0      | 0.001 | 2                   | LB                       | P 0.999  | T 0.144 | B 0.105  | YES                | 3    |
| <i>ALPK3</i>              | c.2701C>T           | p.Q901*   | -           | -                      | -     | -      | -     | 4                   | LP                       | DC 1     | -       | -        | YES                | 2    |
| <i>OBSCN</i>              | c.7865C>T           | p.S2622F  | rs112265836 | 0.002                  | 0.01  | 0.002  | 0.003 | 1                   | B                        | PL 0.999 | D 0.019 | B 0.03   | YES                | 20   |
| <i>PDE4DIP</i>            | c.6872_6873delinsAG | p.R2291Q  | -           | -                      | -     | -      | -     | -                   | -                        | -        | -       | B 0.002  | YES                | 2    |
| <i>PDE4DIP</i>            | c.5730C>A           | D1910E    | -           | 0.4                    | -     | -      | 0.5   | -                   | -                        | PL 0.999 | T 1     | B 0      | YES                | 2    |
| <i>PDE4DIP</i>            | c.5599C>T           | R1867C    | rs1620560   | 0.28                   | -     | -      | 0.5   | 3                   | VUS                      |          | D 0.04  | PD 1     | YES                | 2    |
| <i>PDE4DIP</i>            | c.5269G>A           | A1757T    | rs1628310   | 0.5                    | -     | -      | 0.5   | -                   | -                        |          | T 0.11  | B 0.025  | YES                | 2    |
| <i>PDE4DIP</i>            | c.4793A>G           | H1598R    | rs1778155   | 0.23                   | -     | -      | 0.3   | -                   | -                        | PL 0.997 | T 0.49  | B 0.001  | YES                | 2    |
| <i>PDE4DIP</i>            | c.4511G>A           | R1504Q    | rs2762875   | 0.5                    | -     | -      | 0.5   | -                   | -                        | PL 0.99  | T 1     | B 0      | YES                | 2    |
| <i>PDE4DIP</i>            | c. 4360A>G          | p.K1454E  | rs1778120   | 0.5                    | -     | -      | 0.5   | -                   | -                        | PL 0.99  | T 0.38  | PD 0.991 | YES                | 2    |
| <i>PDE4DIP</i>            | c.4075A>G           | p.K1359E  | rs1747958   | 0.27                   | -     | -      | 0.3   | -                   | -                        | PL 0.999 | D 0     | B 0.229  | YES                | 2    |
| <i>PDE4DIP</i>            | c.3796A>G           | p.K1266E  | rs12568796  | 0.33                   | -     | -      | 0.5   | 3                   | VUS                      | DC 0.69  | T 0.12  | PD 0.998 | YES                | 2    |
| <i>PDE4DIP</i>            | c.3196G>A           | p.A1066T  | rs1698647   | 0.23                   | -     | -      | 0.4   | 1                   | LB                       | PL 0.69  | T 1     | PD 0.954 | YES                | 2    |
| <i>PDE4DIP</i>            | c.3037T>A           | p.F1013I  | rs1698624   | 0.5                    | -     | -      | 0.5   | -                   | -                        | PL 0.999 | T 0.09  | B 0      | YES                | 2    |
| <i>PDE4DIP</i>            | c.2042G>A           | p.R681H   | rs1629011   | 0.5                    | -     | -      | 0.5   | -                   | -                        | PL 0.907 |         | PD 1     | YES                | 2    |
| <i>PDE4DIP</i>            | c.1864C>T           | p.R622*   | rs1778111   | 0.28                   | -     | -      | 0.5   | 3                   | VUS                      | DC 1     | D 0.01  | B 0      | YES                | 2    |
| <i>PDE4DIP</i>            | c.1607G>C           | p.S536T   | rs1747930   | 0.5                    | -     | -      | 0.5   | 3                   | VUS                      | DC 0.85  | T 0.26  | PD 0.998 | YES                | 2    |
| <i>PDE4DIP</i>            | c.1445A>G           | p.H482R   | rs1698681   | 0.5                    | -     | -      | 0.5   | -                   | -                        | PL 0.999 | T 0.26  | B 0.01   | YES                | 2    |
| <i>PDE4DIP</i>            | c.1229A>T           | p.E410V   | rs1061308   | 0.5                    | -     | -      | 0.5   | 2                   | LB                       | DC 0.99  | D 0     | PD 0.995 | YES                | 2    |
| <i>PDE4DIP</i>            | c.824C>T            | p.S275L   | rs1359300   | 0.27                   | -     | -      | 0.5   | -                   | -                        | PL 0.995 | T 0.31  | B 0.09   | YES                | 2    |
| <i>PDE4DIP</i>            | c.74G>T             | p.R25L    | rs1664022   | 0.5                    | -     | -      | 0.5   | 2                   | LB                       | DC 0.999 | T 0.38  | PD 1     | YES                | 2    |
| <i>PDE4DIP</i>            | c.16A>G             | p.T16A    | rs111958438 | 0.1                    |       | 0.2    | 0.03  | -                   | -                        | DC 0.999 | T 0     | B 0      | YES                | 2    |
| <i>KCNE1</i>              | c.112A>G            | p.S38G    | rs1805127   | 0.63                   | 0.67  | -      | 0.65  | 1                   | B                        | P 1      | T 0.4   | B 0.213  | YES                | 11   |
| <i>TTN</i>                | c.47077G>A          | p.V15693I | rs201717871 | 0.001                  | 0.002 | 0.001  | 0     | 1                   | B                        | DC 0.98  | T 0.5   | B 0.41   | YES                | 1056 |
| <i>TTN</i>                | c.40498G>T          | p.V13500F | rs149470241 | 0.001                  | 0.002 | 0.001  | 0     | 2                   | LB                       | DC 0.999 | D 0.033 |          | YES                | 1056 |
| <i>TTN</i>                | c.20335A>T          | p.S6779C  | rs149470241 | 0.001                  | 0.002 | 0.001  | 0     | 2                   | LB                       | PL 0.999 | D 0.033 |          | YES                | 1056 |
| <i>DNAJC19</i>            | c.181C>T            | p.R61W    | rs141007488 | 0                      | 0.01  | 0      | 0     | 2                   | LB                       | DC 1     | T 0.072 | B 0      | YES                | 5    |
| <i>LAMA4</i>              | c.848_849delinsAC   | p.A283D   | rs71543223  | -                      | -     | -      | -     | 2                   | LB                       | -        | -       | -        | YES                | 8    |

B; Benign; DC- disease causing; D-Deleterious/damaging; MT-MutationTaster; PD-probably damaging; PL-Polymorphism; PS-possibly damaging; T-Tolerated; VUS-variant of unknown significance

Investigation into the PHYRE2 and CHIMERA generated secondary structure caused by the *ALPK3* c.2701C>T variant were inconclusive as PHYRE2 had no alignment data for position 901, which is where our variant was located.

Conservation of the truncating *ALPK3* p.Q901\* residue revealed it to be partly conserved in only three species (Figure 5.20).

|             |     |                    |
|-------------|-----|--------------------|
| H. sapiens  | 901 | ADKGTQEDRRMQGEKGMQ |
| M. mulatta  | 907 | ADEGTQEDRRMQEETGMQ |
| M. musculus | 715 | PGD - QEE-- - -    |

Figure 5.20: Conservation of the *ALPK3* p.Q901\* residue across species.

#### 5.5.3.2 Proband's affected brother

Analysis of WES data from the proband's brother (III:1) revealed that he had 38801 total variants with 743 of them being variants filtered through genes in the cardiac panel. GnomAD allele frequency identified 150 variants with a MAF of <1% from which 30 of them were predicted to have a high or moderate impact on protein (Figure 5.21).

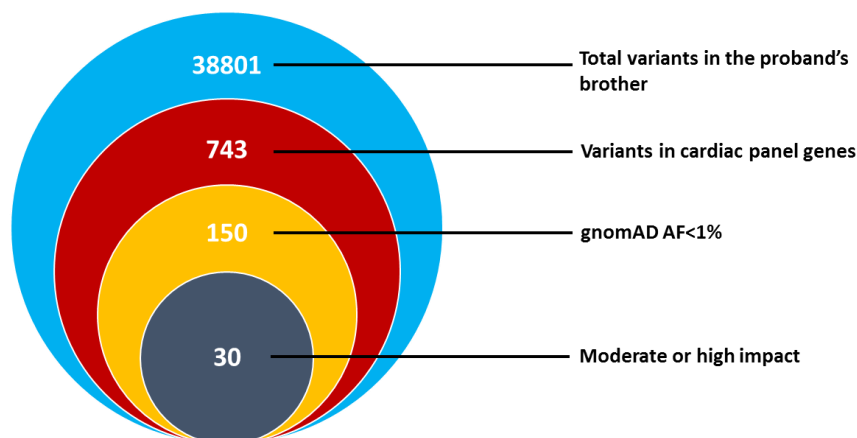


Figure 5.21: Filtering for WES data from the proband's brother alone ACM145 family.

Further investigation excluded all the 30 variants as variants because (1) they were either benign or likely benign according to the ACM classification or (2) they all had a MAF of >1% according to the population databases (Table 5.6).

**Table 5.6: Filtering for WES data in the proband's brother**

| VARIANT IDENTIFIERS (NCP) |                     |           |             | POPULATION FREQUENCIES |       |        |       | ACMG CLASSIFICATION | PATHOGENICITY PREDICTION |          |         |          | PROTEIN EXPRESSION |      |
|---------------------------|---------------------|-----------|-------------|------------------------|-------|--------|-------|---------------------|--------------------------|----------|---------|----------|--------------------|------|
| Gene                      | Variant             | Protein   | rsNumber    | ALFA                   | 1000G | gnomAD | ExAC  |                     | Varsome                  | MT       | SIFT    | PP-2     | Genecards          | Gtex |
| <i>SVIL</i>               | c.1264G>C           | p.V422L   | rs1247696   | 0.98                   | 0.99  | 0.98   | 1     | -                   | -                        | PL 1     | D 0.02  | B 0      | YES                | 3    |
| <i>SVIL</i>               | c.5138C>T           | p.T713I   | rs143178190 | 0.001                  | 0.002 | 0      | 0.001 | 2                   | LB                       | P 0.999  | T 0.144 | B 0.105  | YES                | 3    |
| <i>DSG2</i>               | c.1174G>A           | p.V392I   | rs193922639 | 0.002                  | 0.001 | 0      | 0.002 | 1                   | B                        | DC 0.08  | D 0.017 | B 0.001  | YES                | 12   |
| <i>DTNA</i>               | c.92G>A             | p.R31E    | rs202046233 | 0                      | 0     | 0      | 0     | 1                   | B                        | DC 0.999 | D 0     |          | YES                | 23   |
| <i>OBSCN</i>              | c.7865C>T           | p.S2622F  | rs112265836 | 0.002                  | 0.01  | 0.002  | 0.003 | 1                   | B                        | PL 0.999 | D 0.019 | B 0.03   | YES                | 20   |
| <i>PDE4DIP</i>            | c.6872_6873delinsAG | p.R2291Q  | -           | -                      | -     | -      |       | -                   | -                        | PL 0.999 | -       | -        | YES                | 2    |
| <i>PDE4DIP</i>            | c.5730C>A           | p.D1910Q  | rs1613780   | 0.4                    | -     | -      | 0.5   | -                   | -                        | PL 0.999 | T 1     | B 0      | YES                | 2    |
| <i>PDE4DIP</i>            | c.5599C>T           | R1867C    | rs1620560   | 0.23                   | -     | -      | 0.5   | 3                   | VUS                      | DC 0.94  | D 0.04  | PD 1     | YES                | 2    |
| <i>PDE4DIP</i>            | c.5269G>A           | p.A1757T  | rs1628310   | 0.5                    | -     | -      | 0.5   | -                   | -                        | PL 0.999 | T 0.11  | B 0.025  | YES                | 2    |
| <i>PDE4DIP</i>            | c.4793A>G           | p.H1598R  | rs1778155   | 0.23                   | -     | -      | 0.3   | -                   | -                        | PL 0.997 | T 0.49  | B 0.001  | YES                | 2    |
| <i>PDE4DIP</i>            | c.4511G>A           | p.R1504Q  | rs2762875   | 0.5                    | -     | -      | 0.5   | -                   | -                        | PL 0.99  | T 1     | B 0      | YES                | 2    |
| <i>PDE4DIP</i>            | c.4360A>G           | p.K1454G  | rs1778120   | 0.5                    | -     | -      | 0.5   | -                   | -                        | PL 0.99  | D 0.02  | PD 0.991 | YES                | 2    |
| <i>PDE4DIP</i>            | c.4075A>G           | p.K1359E  | rs1747958   | 0.27                   | -     | -      | 0.3   | -                   | -                        | PL 0.999 | T 0.38  | B 0.229  | YES                | 2    |
| <i>PDE4DIP</i>            | c.3796A>G           | p.K1266E  | rs12568796  | 0.33                   | -     | -      | 0.5   | 3                   | VUS                      | DC 0.69  | D 0     | PD 0.998 | YES                | 2    |
| <i>PDE4DIP</i>            | c.3196G>A           | p.A1066T  | rs1698647   | 0.23                   | -     | -      | 0.4   | 2                   | LB                       | PL 0.69  | T 0.12  | PD 0.954 | YES                | 2    |
| <i>PDE4DIP</i>            | c.3037T>A           | p.F1013I  | rs1698624   | 0.5                    | -     | -      | 0.5   | -                   | -                        | PL 0.999 | T 1     | B 0      | YES                | 2    |
| <i>PDE4DIP</i>            | c.2042G>A           | p.R681H   | rs1629011   | 0.2                    | -     | -      | 0.23  | -                   | -                        | PL 0.907 | T 0.09  | PD 1     | YES                | 2    |
| <i>PDE4DIP</i>            | c.1864C>T           | p.R622Ter | rs1778111   | 0.23                   | -     | -      | 0.5   | 3                   | VUS                      | DC 1     |         |          | YES                | 2    |
| <i>PDE4DIP</i>            | c.1607G>C           | p.S536T   | rs1747930   | 0.5                    | -     | -      | 0.5   | 3                   | VUS                      | DC 0.85  | D 0.01  | PD 0.998 | YES                | 2    |
| <i>PDE4DIP</i>            | c.1445A>G           | p.H482R   | rs1698681   | 0.3                    | -     | -      | 0.3   | -                   |                          | PL 0.999 | T 0.26  | B 0.01   | YES                | 2    |
| <i>PDE4DIP</i>            | c.1229A>T           | p.E410V   | rs1061308   | 0.5                    | -     | -      | 0.5   | 2                   | LB                       | DC 0.99  | D 0     | PD 0.995 | YES                | 2    |
| <i>PDE4DIP</i>            | c.824C>T            | p.S275L   | rs1359300   | 0.5                    | -     | -      | 0.5   | -                   | -                        | PL 0.995 | T 0.31  | B 0.09   | YES                | 2    |
| <i>PDE4DIP</i>            | c.74G>T             | p.R25L    | rs1664022   | 0.5                    | -     | -      | 0.5   | 2                   | LB                       | DC 0.999 | D 0     | PD 1     | YES                | 2    |
| <i>PDE4DIP</i>            | c.16A>G             | p.T6A     | rs111958438 | 0.23                   | -     | -      | 0.5   | -                   | -                        | DC 0.999 | T 0.41  | B 0      | YES                | 2    |
| <i>KCNE1</i>              | c.112A>G            | p.S38G    | rs1805127   | -                      | -     | -      | -     | 1                   | B                        | P 1      | T 0.04  | B 0      | YES                | 11   |
| <i>TTN</i>                | c.53012C>T          | p.A17671V | rs549478203 | 0                      | 0     | 0      | 0     | 3                   | VUS                      | PL 1     | T 0.368 | B 0      | YES                | 1056 |
| <i>TTN</i>                | c.47077G>A          | p.V15693I | rs201717871 | 0.001                  | 0     | 0      | 0     | 1                   | B                        | DC 0.97  | T 0.465 | B 0.213  | YES                | 1056 |
| <i>TTN</i>                | c.40498G>T          | p.V13500F | rs201944202 | 0.001                  | 0     | 0      | 0.001 | 2                   | LB                       | PL 0.93  | T 0.05  | B 0.41   | YES                | 1056 |
| <i>TTN</i>                | c.20335A>T          | p.S6779C  | rs149470241 | 0.001                  | 0     | 0      | 0.001 | 2                   | LB                       | PL 1     | D 0.03  | B 0      | YES                | 1056 |
| <i>LAMA4</i>              | c.848_849delinsAC   | p.A283D   | rs71543223  | -                      | -     | -      | -     | 2                   | LB                       | PL 0.799 | T 1     | B 0      | YES                | 8    |

B; Benign; DC- disease causing; D -Deleterious/damaging; MT-MutationTaster; PD-probably damaging; PL-Polymorphism; Polyphen-2; PS-possibly damaging; T-Tolerated;  
VUS-variant of unknown significance

### 5.5.4 Targeted resequencing

Targeted resequencing for 38 genes was performed in the principal study, called IMHOTEP, where we sequenced a cohort of 634 cardiomyopathy probands (DCM=479, HCM=69, ACM=78, RCM=39)(unpublished data), which included the proband of ACM145 (III:1). We only extracted and analysed the sequencing data for this proband (ACM 145, III:1), as the other probands did not form part of this PhD study. Targeted resequencing identified the *PKP2* c.2197\_2202delInsG and ALPK3 c.2701C>T(p.Q901Ter) variants.

#### 5.5.4.1 *PKP2* c.2197\_2202delInsG

Through targeted resequencing we identified the *PKP2* c.2197\_2202delInsG insertion/deletion in the affected proband (III:1) and his father (II:3) (Figure 5.22). This *PKP2* variant was validated using HRM and Sanger sequencing (Figure 5.22).

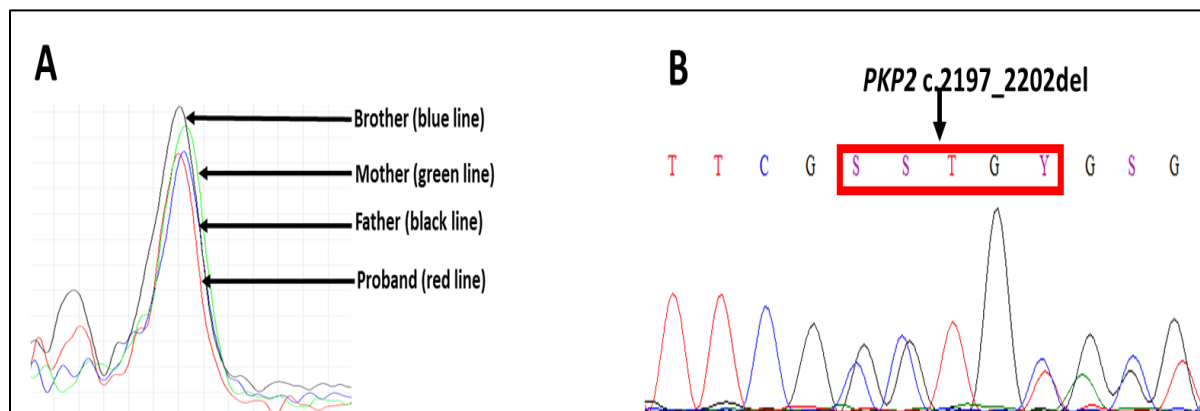


Figure 5.22: **Validation of *PKP2* c.2197\_2202del in family ACM145** A.) Melt-curve profiles of the family members. B.) Sanger sequencing showed the 2197\_2202 del in the proband and his father as shown by the area highlighted in red.

#### 5.5.4.2 *ALPK3 c.2701C>T(p.Q901Ter)*

The Starlim's report generated by the University of Oxford diagnostic laboratory, indicated that the variant had a high coverage of 858 and quality score of 2962, both of which are markedly above the threshold values for a true variant.

This variant was found in the WES data and was validated through HRM and Sanger sequencing in section 5.5.3 above. Targeted resequencing thus confirmed the WES finding.

### 5.5.5 Integrative Genomics Viewer

#### 5.5.5.1 *PKP2 c.2197\_2202delInsG*

As our previous WES data analysis had not identified this variant, the possibility existed that that the other family members could also have insertions or deletions that were not called by Ion Reporter's variant calling software.

Through the Integrative Genomics Viewer (IGV) tool, we manually investigated the IGV tracks for the desmosomal regions (which included *DSP*, *DSG2*, *PKP2*, *DSC2*, *JUP* as well as *CDH2*) for all the family members using the BAM files. IGV confirmed the presence of the *PKP2* c.2197\_2202delInsG variants in the affected proband (III:2) and his father (II:3) (Figure 5.23). The variant caller software for Haloplex was Software Platypus v 0.8.1, GATK v 3.8 had a set threshold of 400 and scores above this threshold were called as mutation positive.

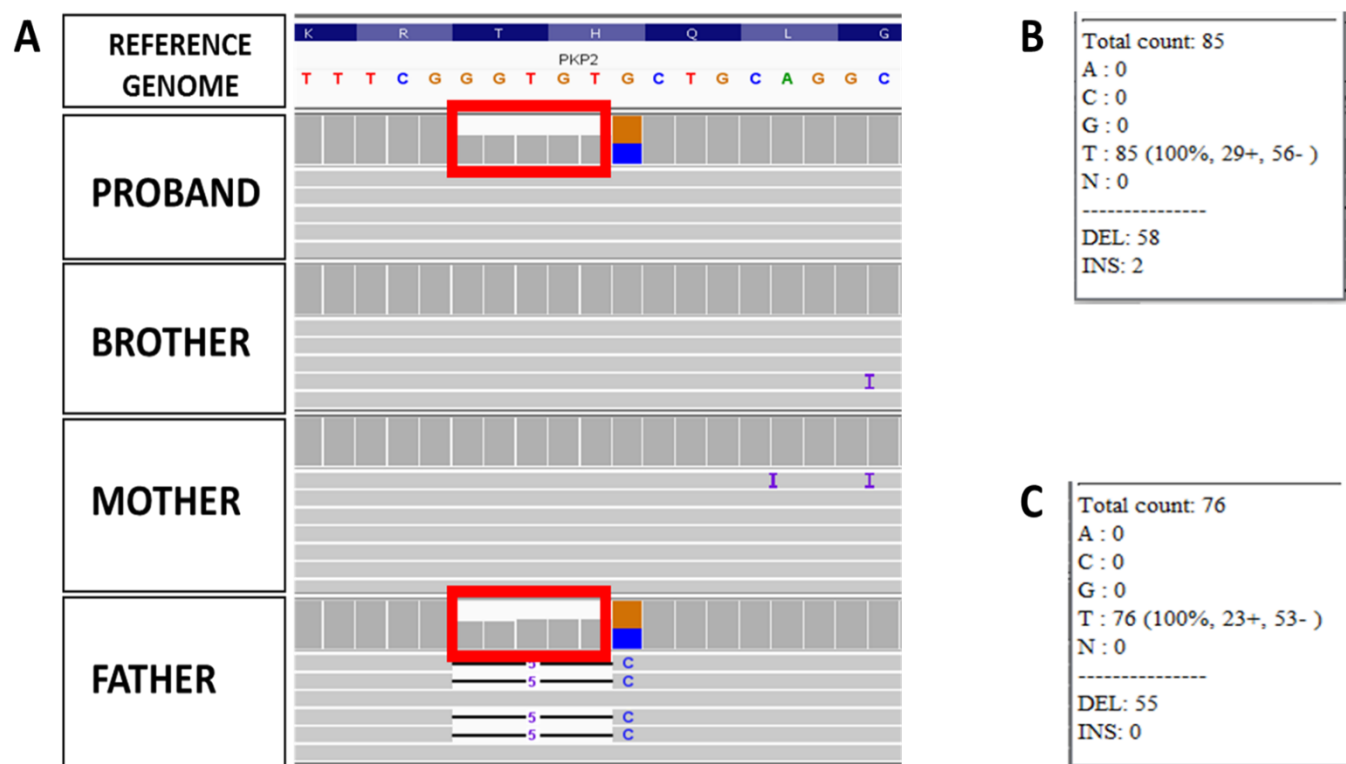


Figure 5.23: Visualization of the *PKP2* c.2197\_2022delInsG variant in the ACM145 family. A.) BAM files seen using IGV with the reference genome hg19 shows a deletion by the significant reduction in bar sizes for the respective bases highlighted in red for both the proband and father. B.) A default threshold of 20 identifies a deletion of 58 bases with 99% confidence in the proband. C.) A deletion of 55 bases in the father identified with a 99% confidence from the base calls at that locus.

The variant is rare according to population databases, predicted to be disease causing according to prediction databases and is a class 5 variant according to ACMG classification (Table 5.6). As observed from IGV, it was found in the proband and his father and was thus considered for variant validation through HRM and Sanger sequencing.

Sanger sequencing of the samples revealed that the proband and his father were heterozygous for the *PKP2* c.2197\_2202del variant, while the proband's affected brother and healthy mother were wildtype and did not harbor the *PKP2* deletion as previously shown in figure 5.23 above.

To look for secondary structural changes caused by the *PKP2* c.2197\_2202del variant we submitted our query sequence to PHYRE2 resulting in residues 345 to 619 being confidently modelled but not our residue at position 733 for us to ascertain the protein structural change by PHYRE2 and CHIMERA.

When we looked at the conservation of the *PKP2* p.H733fs residue, we found that the *PKP2* c.2197\_2202del variant affected a highly conserved amino acid across various species (Figure 5.24)

|               |     |   |   |   |   |   |   |   |   |
|---------------|-----|---|---|---|---|---|---|---|---|
| Hsapiens      | 733 | H | T | R | K | M | L | H | V |
| Pteroglodytes | 719 | H | T | R | K | M | L | H | V |
| Mmulatta      | 722 | H | T | R | K | M | L | H | V |
| Mmusculus     | 647 | H | T | R | K | M | L | H | V |
| Xtropicalis   | 669 | H | R | N | T | L | N | S | S |

Figure 5.24: Conservation of the *PKP2* p.H733fs residue across multiple species.

#### 5.5.5.2 *ALPK3* c.2701C>T(p.Q901Ter)

WES data showed that the *ALPK3* variant is also found in the father and IGV visualization confirmed the observation (Figure 5.25)

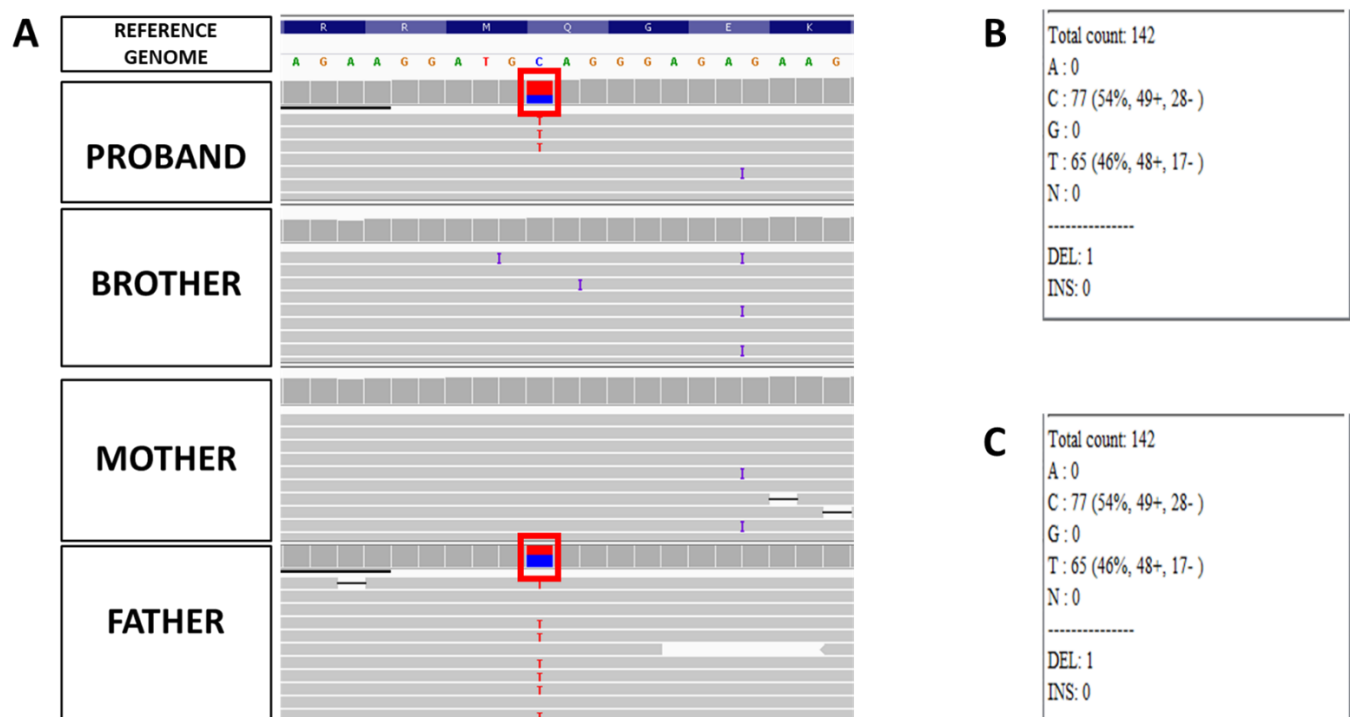


Figure 5.25: Visualisation of *ALPK3* c.2701C>T variant using IGV with the reference genome hg19. A.) The C>T change is highlighted in red for both the proband and father. B.) A default threshold of 20 in IGV identifies allelic heterogeneity at this locus in the proband showing a C>T change with 99% confidence base call C.) The allelic heterogeneity identified in the father indicates a total of 142 bases called at this locus with 77 being C (blue color) and 65 being T (red color).

### 5.5.6 Variant summary for ACM145

Following WES analysis and equally reliable quality checks from either IGV visualisation, Starlim's report from Haloplex NGS or Sanger sequencing we were able to identify a *PKP2* deletion which has been reported and a novel *ALPK3* variant, both of which were in the proband and his father. We were not able to identify a potential candidate variant from WES analysis of the brother. The segregation of the *PKP2* and *ALPK3* variants is as shown in the pedigree below (Figure 5.26)

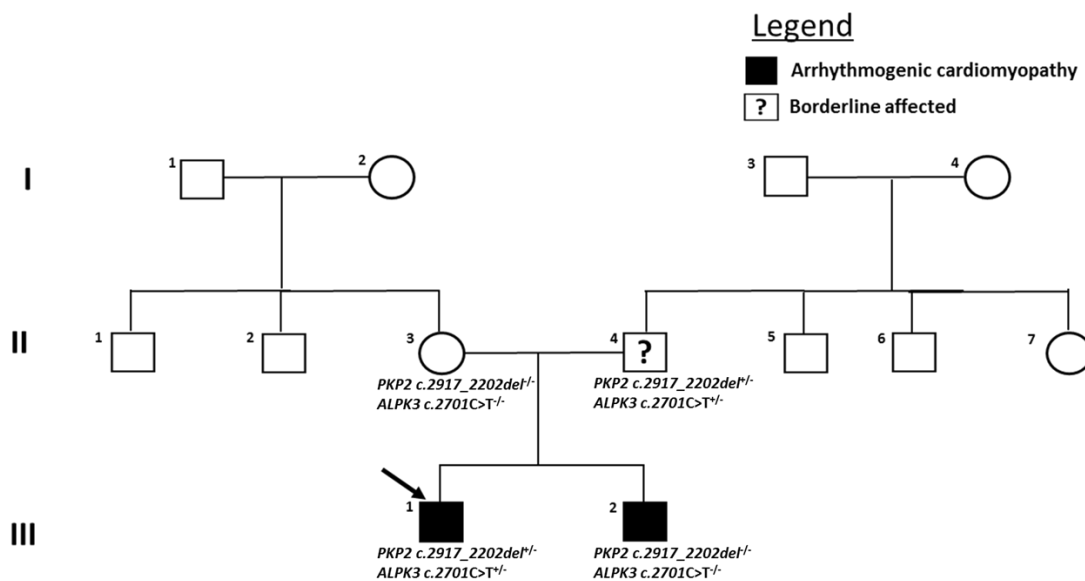


Figure 5.26: Segregation of variants identified after individual WES analysis from the brothers in family ACM145. Squares represent males, circles represent females. Shaded squares and circles represent affected individuals. Arrow points towards the proband.

## 5.6 Discussion

ACM is often described as an SD disease of the cardiac desmosome with variability in expression and accompanying factors.<sup>28,246</sup> We performed whole exome sequencing on four members of a Caucasian South African family where the two brothers had clinically been diagnosed with ACM.

### 5.6.1 Known cardiac gene panel

WES data analysis did not identify any pathogenic variants common between the two affected brothers. However, through additional investigation of all scenarios, targeted resequencing identified the pathogenic *PKP2* c.2197\_2202del (p.H733Afs\*8) variant and de novo variants analysis approach identified the likely pathogenic *ALPK3* c.2701C>T(p.Q901Ter) variant.

#### 5.6.1.1 PKP2 c.2197\_2202del(p.H733Afs\*8)

This variant was detected though targeted resequencing but was not identified through the WES data analysis. This could be due to the stringency of the IonReporter variant calling software that was used which resulted in the variant being eliminated.

The *PKP2* pathogenic deletion was found in the affected proband (III:1), and his borderline affected father (II:4). This *PKP2* variant has been previously reported and shown to be responsible for ACM in studies involving a US cohort and Polish patients whereby 3/56 ACM cases and 1/56 ACM cases were shown to harbor the variant respectively.<sup>250,251</sup>

PKP2 is an 100kDa protein that is found in all cell types with desmosomal junctions, such as the myocardium, and in the nucleus whereby its presence in the nucleus is independent of its desmosomal presence.<sup>252</sup> The PKP2 protein is made up of an N-terminal head domain containing the homologous region 2 (HR2), a series of eight 42-amino acid repeats (ARM) and a short C-terminal tail (Figure 5.27).<sup>252,253</sup>



Figure 5.27: Structure of PKP2 showing the N-terminal domain (N), ARM motifs (A) and C-terminal domain (C).<sup>252,253</sup>

PKP2 is ubiquitously expressed and this points to its broad cellular functions due to its ability to bind with numerous ligands from its ARM motifs.<sup>254</sup> It interacts with the desmosomal components DSC2, DSP, DSG2 and JUP to recruit and integrate them to the cell border during assembly of the desmosomes, which is vital for optimal cell adhesion in the heart.<sup>254,255</sup> It also regulates the expression of cardiac genes such as *connexin 42*, *RYR2*, *ANK2* and *calsequestrin-2* which represent calcium handling genes that play a key role in the normal heart electrical

conduction for a coordinated cardiac contraction.<sup>255,256</sup> In addition, PKP2 plays a role in the transcriptional modification of multiple molecular pathways through its interaction with  $\beta$ -catenin and PKC, highlighting an intracellular signalling node whose dysfunction is suggested to contribute to adipogenesis and fibrogenesis pathway.<sup>255,235</sup>

*PKP2* variants are the most common cause of ACM and they frequently result in truncated proteins due to mutations that are either nonsense, frameshift and splice site.<sup>235,257</sup> ACM cases that are linked to *PKP2* commonly manifest an orthodox ARVC phenotype characterised by limited involvement of the LV, although mild LV dysfunction and dilation later on in life have been documented.<sup>258,259,260</sup> It is also noteworthy that carriers of *PKP2* variants do not manifest cardiac-cutaneous disease.<sup>261</sup>

Studies done on various models have enlarged the scope of *PKP2* and its involvement in the ACM phenotype. Work done to study the effects of PKP2 was carried out by Cerrone and colleagues where they generated a KO mouse model with the ability to control the expression of *Pkp2* and showed that despite both ventricles losing PKP2, the RV was first to be affected indicating that predominance of RV involvement in ACM is possibly due to the interplay between functional requirements and structural composition of the RV instead of developmental attributes that favor one ventricle over the other during embryogenesis.<sup>255</sup> They also noted that mechanical dysfunction precedes fibrotic replacements of the cardiomyocyte indicating that fibrosis is reparative following loss of PKP2 expression.<sup>255</sup> This is in line with fibrotic changes observed in ACM at a later stage because the profibrotic and inflammatory pathways are upregulated due to transcriptional dysregulation.<sup>255,261-264</sup> The transcriptional dysregulation has also been shown to downregulate adipo-cytokine and insulin signaling genes that play a role in the metabolic facets of the disease.<sup>265</sup>

The *PKP2* variant identified in this family occurs at PKP2 ARM 6 and while the exact function of the individual PKP2 ARM motif is still to be elucidated, the ARM proteins are known to be involved in facilitating protein-protein interactions, transcriptional activation and nuclear localization.<sup>266</sup> This points to the significance of PKP2 in incorporating scaffolding proteins at the cell junction which is vital for the integrity of cell adhesion and intercellular communication among the cardiomyocytes.<sup>255</sup> A mutation at this arm can be a possible factor to abrogate the binding properties of PKP2 which occur at the grooves formed by the ARM domain, a key factor in ensuring proper desmosomal architecture for cardiac activity.<sup>266</sup> In addition, *PKP2* c.2197-2202del variant is also found next to other residues that have been reported to cause ACM within ARM6, underpinning a mutational area of interest in ARM 6 which up to date, contains four variants, which is the most number of variants causative for ACM amongst the ARM motifs (Figure 5.28).

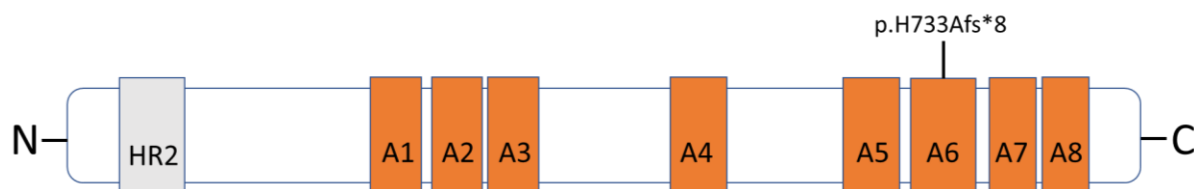


Figure 5.28: *PKP2* c.2197-2202del variant in PKP2 ARM 6 motif.

In this ACM family, both the proband and his father were mutation positive for the *PKP2* c.2197\_2202del variant however, the proband was much more severely affected while his father had more subtle disease. One explanation for this could be that ACM is known to result in phenotypic variability within members of the same family and incomplete penetrance which has been purported to be related to gender, epigenetic modifications or infections which can trigger disease.<sup>244,248</sup> However, within this family, another explanation could be that the phenotypic severity between the proband and his father could be attributed to the prior diagnosis of

thyrotoxicosis in the proband. It is known that some of the common manifestations of thyroid disease are due to the thyroid hormone exerting its effects on the cardiovascular system resulting in unconventional myocardial oxygen uptake, cardiac output and contractility.<sup>249</sup> We hypothesise that proband was already genetically predisposed to the cardiomyopathy due to the pathogenic *PKP2* variant and the diagnosis of thyrotoxicosis was thus an added risk factor that exacerbated an already predisposed individual; this resulted in the more severe cardiac phenotype seen in the proband.

Genetic analysis of the affected brother's DNA was negative for the *PKP2* c.2197\_2202del variant and no other plausible candidate variants were identified thus reinforcing the complexity of ACM heritability.<sup>51</sup> We will have to explore other techniques for this family, as WES or targeted sequencing might not have been able to detect the pathogenic variant/s.

#### 5.6.1.2 *ALPK3* c.2701C>T(p.Q901Ter)

Whole exome sequencing identified the novel truncating heterozygous *ALPK3* c.2701C>T(p.Q901Ter) variant in the proband and his affected father. The variant was found in exon 6 of the *ALPK3* gene which is between the two Ig-like domains and 3 exons away from the alpha kinase domain (Figure 5.30). *ALPK3* is localized in the nucleus and encodes for the  $\alpha$ -protein kinase 3 at chromosome 15 and is highly expressed in the heart early on during cardiogenesis and in adulthood.<sup>267,268</sup> The *ALPK3* protein is 1907 amino acids long with 14 exons and is made up of two immunoglobulin-like (Ig-like) domains and an alpha-kinase towards the C-terminal end (Figure 5.29).

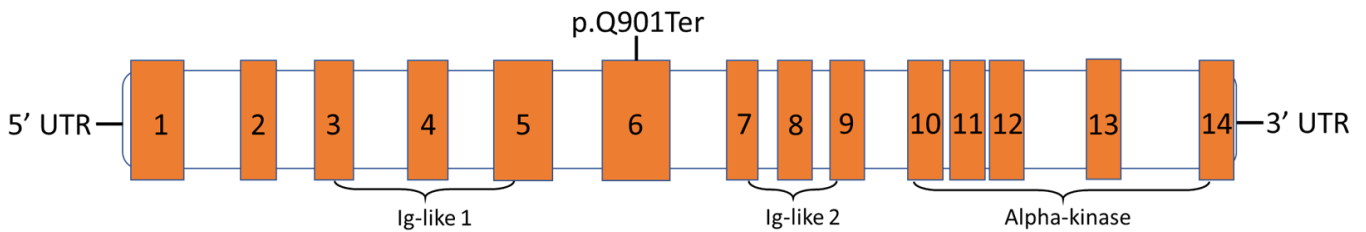


Figure 5:29: Location of the *ALPK3* variant in exon 6. The exons are shaded orange and numbered accordingly.

*ALPK3* is required for the normal formation of intercalated discs (IDs) and cardiomyofibril coordination thus highlighting its important role in ensuring proper electrical and mechanical coupling of cardiomyocytes for normal cardiac function.<sup>267,269</sup> It also plays a key role in the maturation of cardiomyocytes by regulating the transcription of crucial cardiac transcription factors such as HEY and HAND proteins, which are involved in pathways that are needed for the maturation of cardiomyocytes during the early stages of heart development.<sup>267,268,270</sup>

*ALPK3* variants were initially reported in paediatric forms of HCM in a recessive manner but they have subsequently been shown to be implicated in patients with other forms of cardiomyopathy including DCM, ACM and LVNC.<sup>271,272,273</sup> Patients with *ALPK3* variants manifest high phenotypic variability ranging from heart failure, ventricular arrhythmias with poor prognosis, patients requiring an ICD implantation as seen in our proband, mixed phenotype and sudden cardiac death.<sup>267,274</sup> Variations in the gene also cause differing ages of onset among mutation carriers with some carriers presenting at the paediatric stage with some being clinically diagnosed later in life although the age of presentation is reported to be between infancy and the fourth decade of life.<sup>275</sup> Despite carriers of *ALPK3* variants having a poor prognosis in the proceeding initial first weeks, numerous children survivors show a significant improvement of their cardiac function with time.<sup>271,275</sup> In addition, parents and relatives who only have one damaging *ALPK3* allele exhibit reduced penetrance, which was observed in a Dutch population with the age of diagnosis

ranging from 20 weeks gestation to 53 years whereby cardiac evaluations did not reveal any cardiomyopathy among relatives with loss of function mutations in *ALPK3*.<sup>271</sup>

Various reports have documented *ALPK3* variants among ACM patients, although these have been fewer compared to reports among HCM and DCM cases. In a multicentre study done to investigate pathogenicity of *ALPK3*, 4904 unrelated probands were recruited to the study from Denmark, Spain, Russia, the UK, Latvia, Argentina and Brazil of which 435 had ACM.<sup>272</sup> *ALPK3* truncating variants (*ALPK3tv*) were identified from 36 probands with one of the proband falling in the ACM cohort. The ACM case was a male diagnosed at 64 years with RV dilatation and impairment together with LVEF 66% on echo. He had an ICD implanted and on cardiac MRI he showed RVEF 32%, LVEF 49% with LGE inferior RV and anterior LV.<sup>272</sup> In another *ALPK3* study done to assess the clinical and genetic profile of *ALPK3* among 1548 patients in a Dutch and US cohort with adult-onset cardiomyopathy, four patients were discovered to have ACM and the age of diagnosis was ranging from 20 years to 55 years. All the four patients were Dutch and they exhibited three missense *ALPK3* variants, predicted to be VUS and one LoF *ALPK3* variant which was predicted to be pathogenic.<sup>271</sup>

Studies done on *ALPK3* have had a significant impact in elucidating how variations in this gene lead to cardiac malformations. In *Alpk3*<sup>-/-</sup> mice, the cardiomyocytes were shown to be thicker in most areas of the heart compared to their wild-type littermates with overt changes in the IDs and nuclei and myofibrils.<sup>269</sup> The IDs were convoluted, jagged and significantly impaired in number and they tended to be less distinct in contrast to WT which exhibited large numbers of IDs with a sharp and dense appearance. The cell nuclei of the cardiomyocytes contained large nucleoli and were lobulated, enlarged or rounded while the myofibrils were loosely arranged, overlapping and crossing instead of being tightly organised in parallel bundles as in the WT mice.

There were no deaths noted in the mice up to 12 months of age despite the cardiac anomalies observed.<sup>269</sup>

While the exact biology of these ALPK3 functional domains remains to be fully understood, it is known that *ALPK3* is in a lineage of protein kinases that identify sites of phosphorylation and effect their regulatory role in cellular functions such as cell adhesion, proliferation, ion homeostasis and cellular migration.<sup>276,277</sup> Exon 6 of *ALPK3* contains more truncating variants than all the other 13 exons of the gene (17/31, ~55%) representing a mutational hotspot for causing cardiomyopathy.<sup>272</sup> This truncation results in a complete deletion of the alpha-kinase domain of *ALPK3* which is thought to play key roles in cell adhesion and ion handling, which are affected in an ACM phenotype.<sup>269,271</sup> The causality of this *ALPK3* variant in ACM is also supported by the fact that patients with *ALPK3* variants have been shown to display conduction and rhythm anomalies similar to those observed in ACM which is likely due to a localisation of *JUP* to the intracellular and nuclear locations rather than the IDs leading to dysfunctional IDs which require *JUP* for their cardiac mechanotransduction.<sup>271</sup> This redistribution of *JUP* from IDs also leads to a repression of the Wnt/beta-catenin signal pathway resulting in myocyte apoptosis and fibrosis as seen in ACM cases.<sup>271,278</sup>

As these variants do not explain the disease phenotype and severity within this family, the possibility remains that a novel gene might be at the root of the cause of disease within this family. This led us to explore the entire WES data set for any possible novel candidate genes.

### 5.6.2 Possible novel candidate genes

Our filtering criteria for the possible novel cardiomyopathy genes resulted in the prioritization of two variants that were found to be rare with one of the variant being pathogenic and the other

being a variant of unknown significance (Figure 5.3). The two variants were *CFTR* c.1727G>C (p.G576A) and *CFTR* c.2002C>T(R668C).

#### 5.6.2.1 *CFTR* c.1727G>C(p.G576A)

This heterozygous *CFTR* variant was identified in the two brothers from WES data. WES data also showed a homozygous *CFTR* c.1727G>C variant in the father. This homozygous *CFTR* variant was classified as a variant of unknown significance according to ACM variant classification and had a population frequency of  $5 \times 10^{-3}$  according to gnomAD.

Sanger sequencing revealed that the two affected brothers harbor the heterozygous *CFTR* c.1727G>C variant while the father, who is borderline, has the homozygous *CFTR* c.1727G>C variation. This genotype observation does not support the *CFTR* c.1727G>C(p.G576A) genetic variant as the cause of ACM in this family as it would be expected the father who has this variant in its homozygous state to have a more severe disease phenotype than the brothers who are severely affected yet have the heterozygous form of the variant.

#### 5.6.2.2 *CFTR* c.2002C>T(p.R668C)

According to WES data, the *CFTR* c.2002C>T variant was identified in the two brothers in heterozygous state. and a homozygous state in the father. This homozygous *CFTR* variant was classified as a likely pathogenic variant according to ACMG variant classification and had a population frequency of  $6 \times 10^{-3}$  as reported by gnomAD.

Through Sanger sequencing, it was shown that the two affected brothers harbor the heterozygous *CFTR* c.2002C>T change while the father who is borderline has the homozygous *CFTR* c.2002C>T variation and thus this genotype observation through segregation analysis does not support the *CFTR* c.2002C>T(p.R668C) genetic variant as the cause of ACM in this family

because the brothers who have this *CFTR* variant in its heterozygous state manifest a more severe ACM phenotype compared to the father who has this variant in its homozygous state.

## 5.7 Conclusion

We present WES results in two brothers with confirmed ACM. Different NGS analysis approaches identified the previously reported pathogenic *PKP2* variant and a novel likely pathogenic *ALPK3* variant in the affected proband and his mildly affected father. To the best of our knowledge, we are the first to report the two variants in a South African and African cohort. However, no mutation was found in the affected brother. It is possible that the lack of a genetic diagnosis in the proband's brother can be due to the pathogenic variant being in other areas of the genome.

This would lead us to explore other regions such as intronic and regulatory regions as they were not studied in the scope of this thesis and form an active area that could be explored further. We could also explore copy number variations using the Cytoscan HD technologies. Some of the limitations in this study include putative disease-causing variants being initially missed by WES but identified through resequencing and even with resequencing no disease-causing variants were identified for the proband's brother who is clinically affected. These study limitations point towards the shortcomings of current sequencing technologies which still have some error margins due to the chemistry of the chips used.<sup>279</sup> With rapid technological advancements, these chemistries are being improved which would lead to reduced error margins and ensure consistently high confidence variant calling. In addition, our stringent filtering pipeline used for WES data analysis might have eliminated the disease-causing variants identified through resequencing and we thus aim to have less stringent filtering criteria in identifying candidate variants.

We used WES to unravel the genetic variant responsible for cardiomyopathy however, there are other techniques that can be used to discover plausible pathogenic variants within this family which include karyotyping, comparative genomic hybridization (CGH), DNA microarray and Multiplex ligation-dependent probe amplification (MLPA) and many more.<sup>280</sup>

## CHAPTER 6 - CONCLUSION AND FUTURE DIRECTIONS

### 6.1 Conclusion

A lifestyle shift in SSA has led to an epidemiological transition towards increase of NCDs such as CVD of which the cardiomyopathies present the greatest challenge.<sup>281</sup> However, research into this is a largely unexplored in Africa and previous work has highlighted the limited genetic studies conducted in the continent.<sup>56</sup> Through our research we hypothesized that NGS could be used to identify the genetic causes of disease in South African families with various cardiomyopathies. Through our aims and objectives we showed that NGS successfully identified novel and known causes of disease within these families.

In the first family described in this thesis, DCM 334, we describe a multi-generational family affected with severe DCM. WES was performed on two severely affected DCM family members and we identified the *DMPK* c.925C>T. To the best of our knowledge, we are the first to show that a rare variant in *DMPK* is associated with primary DCM in severely affected young patients. Through additional screening of our extended cardiomyopathy cohort, we identified the *DMPK* c.1477C>T variant in another proband but there have been conflicting reports on this variant. We propose that this variant could lead to DCM in an individual with additional risk factors and have classified this variant as a variant of unknown significance. To the best of our knowledge, we are the first to report *DMPK* as a primary cause of DCM.

In the second family investigated in this thesis, DCM343, we describe a three-generation family also affected with DCM. While this *BAG3* c.925C>T (p.R309Ter) variant has been reported in other families, to the best of our knowledge, this is the first identification of the mutation in South African DCM patients. We also report the first South African cohort screen for *BAG3*, where multiple DCM probands have been identified as having the same pathogenic *BAG3* variant. We

hypothesise that this could possibly be a founder variant in the South African population. This requires further investigation through haplotyping. Additional factors such as viral infections will be investigated as various family members' initial presentation was preceded by some form of infection. This could have also contributed to the variable phenotypic expression seen within not only this family, but other families as well. We also discuss modifiers and of studying the modulating effect of ACE-is and/or ARBs on *BAG3* c.925C>T and how patients with pathogenic *BAG3* mutations could possibly benefit from receiving personalised treatment regimens.

In the third family, we investigated a South African family affected with severe ACM. Targeted resequencing identified the *PKP2* c.2197\_2202delInsG indel as well a novel *ALPK3* c.2701C>T(p.Q901Ter) variant in the proband and his affected father. We were not able to detect any plausible candidate variants for the affected brother of the proband. Diagnostic yields of genetics in cardiomyopathies can be up to 60% and the remainder still need to be unraveled. This points towards other unexplored regions in the genome such as intronic regions which might give more insight on their role in causing cardiomyopathies.

Our research has contributed to the genetic landscape and has filled in some of the gaps in our knowledge for the different cardiomyopathies. SSA could provide some valuable insights into the cardiomyopathies and identify additional pathogenic mechanisms responsible for disease.

## 6.2 Future directions

The genetic approaches employed here have highlighted the role that Africa's genetic studies have in providing a significant contribution to cardiomyopathy cases. This also provides a platform for Africa to add onto the global knowledge of cardiomyopathy causing genes.

We intend to create a zebrafish model where we will remove the kinase domain of *dmpk* to determine if this domain would result in a primary cardiomyopathy, as was seen in our patient with the *DMPK* variant. The various models to recapitulate disease phenotype have increasingly been used and both cellular and animal models have shown their significant role in understanding the molecular mechanisms underlying the cardiomyopathies due to similar disease presentation to humans or similar conserved regions being affected. More use of these disease models will ensure translation to clinical utility of genetic data is incorporated more to aid in the realization of personalized medicine.

We intend to further explore the possibility of a *BAG3* founder variant. The two cases of *BAG3* variants seen in individuals with hypertension but did not display a DCM phenotype point towards a potential role of anti-hypertensives being protective against DCM manifestation in individuals with *BAG3* variants. This can further be extended to other cardiomyopathies and genes in a similar molecular pathway to *BAG3* because these cardiomyopathy causing genes have been shown to cause more than one disease phenotype.

And finally, in our ACM 145 family, we intend to explore other regions such as intronic and regulatory regions as they were not studied in the scope of this thesis and form an active area that could be explored further. We could also explore copy number variations using the Cytoscan HD technologies. Apart from regions unexplored during WES, it is also worthwhile to note that the current sequencing technologies still have some error margins which can come into play to impede the correct observation of a genetic cause of cardiomyopathy in individuals.

For the modifier variants identified, we intend to perform translational studies that include use of model organisms such as zebrafish and cellular models in order to identify and understand the molecular mechanisms that are responsible for disease and the variations in phenotype. This

would allow us to have a more tailored approach to patient management rather than a one size fits all for the patients.

The field of genetics has grown in leaps and bounds over the last few decades and it will be fascinating to see what the next few decades will bring.

## REFERENCES

1. Bygbjerg IC. Double burden of noncommunicable and infectious diseases in developing countries. *Science*. 2012 Sep 21;337(6101):1499-501. doi: 10.1126/science.1223466. PMID: 22997329.
2. "World Health Day 2022". 2022. WHO | Regional Office For Africa. URL: <https://www.afro.who.int/regional-director/speeches-messages/world-health-day-2022>.
3. Yach D, Kellogg M, Voute J. Chronic diseases: an increasing challenge in developing countries. *Trans R Soc Trop Med Hyg* 2005;99:321–4.
4. GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018 Nov 10;392(10159):1736-1788. doi: 10.1016/S0140-6736(18)32203-7. Epub 2018 Nov 8. Erratum in: *Lancet*. 2019 Jun 22;393(10190):e44. Erratum in: *Lancet*. 2018 Nov 17;392(10160):2170. PMID: 30496103; PMCID: PMC6227606.
5. A systematic analysis for the Global Burden of Disease Study 2017. Default results are deaths and DALYs for 2017 with trends since 1990 [Internet]. 2017 [cited March 20, 2019]. URL: Available at: <http://ghdx.healthdata.org/gbd-results-tool>.
6. Yuyun MF, et al. Cardiovascular Diseases in Sub-Saharan Africa Compared to High-Income Countries: An Epidemiological Perspective. *Global Heart*. 2020; 15(1): 15. DOI: <https://doi.org/10.5334/gh.403>
7. Mendis, S., Puska, P., Norrving, B., Organization, W., Federation, W., & Organization, W. (2011). *Global atlas on cardiovascular disease prevention and control*. World Health Organization. Retrieved from <https://apps.who.int/iris/handle/10665/44>
8. Damasceno A, Mayosi BM, Sani M, Ogah OS, Mondo C, Ojji D, et al. The causes, treatment, and outcome of acute heart failure in 1006 Africans from 9 countries. *Archives of internal medicine*. 2012; 172(18): 1386–94. DOI: <https://doi.org/10.1001/archinternmed.2012.3310>
9. Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, Chang AR, Cheng S, et al. Heart Disease and Stroke Statistics – 2018 Update: A Report From the American Heart Association. *Circulation*. 2018; 137(12): e67–e492. DOI: <https://doi.org/10.1161/CIR.0000000000000573>
10. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JGF, Coats AJS, et al. 2016 ESC Guidelines for the Diagnosis and Treatment of Acute and Chronic Heart Failure. *Revista espanola de cardiologia (English ed)*. 2016; 69(12): 1167. DOI: <https://doi.org/10.1016/j.rec.2016.11.005>
11. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Colvin MM, Drazner MH, Filippatos GS, Fonarow GC, Givertz MM, Hollenberg SM, Lindenfeld J, Masoudi FA, McBride PE, Peterson PN, Stevenson LW, Westlake C. 2017 ACC/AHA/HFSA focused update of the 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Failure Society of America. *Circulation*. 2017; 136:e137–e161
12. Eberly LA, Rusingiza E, Park PH, Ngoga G, Dusabeyezu S, Mutabazi F, et al. Understanding the Etiology of Heart Failure Among the Rural Poor in Sub-Saharan Africa: A 10-Year Experience From District Hospitals in Rwanda. *Journal of Cardiac Failure*. 2018; 24(12): 849–53. DOI: <https://doi.org/10.1016/j.cardfail.2018.10.002>
13. Sliwa, K., Wilkinson, D., Hansen, C., Ntyintyane, L., Tibazarwa, K., Becker, A., & Stewart, S. (2008). Spectrum of heart disease and risk factors in a black urban population in South Africa (the Heart of Soweto Study): a cohort study. *Lancet*, 371(9616), 915-922. doi: 10.1016/S0140-6736(08)60417-1
14. Khatibzadeh S, Farzadfar F, Oliver J, Ezzati M, Moran A. Worldwide risk factors for heart failure: A systematic review and pooled analysis. *Int J Cardiol*. 2012
15. Sliwa, K., Damasceno, A., & Mayosi, B. M. (2005). Epidemiology and etiology of cardiomyopathy in Africa. *Circulation*, 112(23), 3577-3583. doi:10.1161/CIRCULATIONAHA.105.542894

16. Akinkugbe OO, Nicholson GD, Cruickshank JK. Heart disease in blacks of Africa and the Caribbean. *Cardiovasc Clin.* 1991; 21: 377–391
17. Braunwald E.: Cardiomyopathies an overview. *Circ Res* 2017; 121: pp. 711-721
18. Yogasundaram, H., Alhumaid, W., Dzwiniel, T., Christian, S., & Oudit, G. (2021). Cardiomyopathies and Genetic Testing in Heart Failure: Role in Defining Phenotype-Targeted Approaches and Management. *Canadian Journal Of Cardiology*, 37(4), 547-559. doi: 10.1016/j.cjca.2021.01.016
19. Maron BJ, Towbin JA, Thiene G, Antzelevitch C, Corrado D, Arnett D, Moss AJ, Seidman CE, Young JB; American Heart Association; Council on Clinical Cardiology, Heart Failure and Transplantation Committee; Quality of Care and Outcomes Research and Functional Genomics and Translational Biology Interdisciplinary Working Groups; Council on Epidemiology and Prevention. Contemporary definitions and classification of the cardiomyopathies: an American Heart Association Scientific Statement from the Council on Clinical Cardiology, Heart Failure and Transplantation Committee; Quality of Care and Outcomes Research and Functional Genomics and Translational Biology Interdisciplinary Working Groups; and Council on Epidemiology and Prevention. *Circulation*. 2006 Apr 11;113(14):1807-16. doi: 10.1161/CIRCULATIONAHA.106.174287. Epub 2006 Mar 27. PMID: 16567565.
20. Elliott P., Andersson B., Arbustini E., et al: Classification of the cardiomyopathies: a position statement from the European Society of Cardiology working group on myocardial and pericardial diseases. *Eur Heart J* 2008; 29: pp. 270-276
21. "The Myocardium". Slideshare.Net, 2021, <https://www.slideshare.net/AdeboyeOluwajuyitan/the-myocardium>. Accessed 7 Nov 2021
22. Seferovi\_c PM, Polovina M, Bauersachs J, et al. Heart failure in cardiomyopathies: a position paper from the Heart Failure Association of the European Society of Cardiology. *Eur J Heart Fail* 2019;21:553-76.
23. Yogasundaram H, Putko BN, Tien J, et al. Hydroxychloroquineinduced cardiomyopathy: case report, pathophysiology, diagnosis, and treatment. *Can J Cardiol* 2014;30:1706-15
24. Yogasundaram H, Kim D, Oudit O, Thompson RB, Weidemann F, Oudit GY. Clinical features, diagnosis, and management of patients with Anderson-Fabry cardiomyopathy. *Can J Cardiol* 2017;33:883-97.
25. Wexler RK, Elton T, Pleister A, Feldman D. Cardiomyopathy: an overview. *Am Fam Physician*. 2009 May 1;79(9):778-84. PMID: 20141097; PMCID: PMC2999879.
26. Yogasundaram, H., Alhumaid, W., Dzwiniel, T., Christian, S., & Oudit, G. (2021). Cardiomyopathies and Genetic Testing in Heart Failure: Role in Defining Phenotype-Targeted Approaches and Management. *Canadian Journal Of Cardiology*, 37(4), 547-559. doi: 10.1016/j.cjca.2021.01.016
27. Fiedler A. Ueber akute interstitielle mvokarditis., zu Dresden R, ed. In: *Festschrift zur Feier des funfzigjahrigen Bestehens des Stadtkrankenhauses zu Dresden-Friedrichstadt*. Part 2. Dresden, Germany: Wilhelm Baensch; 1899:3–24
28. Salemi VMC, Mohty D, Altavila SLL, Melo MDT, Kalil Filho R, Bocchi EA. Insights into the Classification of Cardiomyopathies: Past, Present, and Future Directions. *Clinics (Sao Paulo)*. 2021;76:e2808 <https://doi.org/10.6061/clinics/2021/e2808>
29. Brigden W.: Uncommon myocardial diseases: the non-coronary cardiomyopathies. *Lancet* 1957; 273: pp. 1179-1184
30. Brandenburg RO, Chazov E, Cherian G, False AO, Grosogoeat Y, Kawai C, Loogen F, Marin Judez V, Orinius E, Goodwin JF, Olsen J, Oakley CM, Pisa Z. Report of the WHO/ISFC task force on the definition and classification of cardiomyopathies. *Br Heart J*. 1980; 44:672–673.
31. Goodwin J.F., Gordon H., Hollman A., et al: Clinical aspects of cardiomyopathy. *Br Med J* 1961; 1: pp. 69-79

32. Richardson P., McKenna W., Bristow M., et al: Report of the 1995 World Health Organization/International Society and Federation of Cardiology task force on the definition and classification of cardiomyopathies. *Circulation* 1996; 93: pp. 841-842
33. Arbustini E, Narula N, Tavazzi L, Serio A, Grasso M, Favalli V, Bellazzi R, Tajik JA, Bonow RO, Fuster V, Narula J. The MOGE(S) classification of cardiomyopathy for clinicians. *J Am Coll Cardiol*. 2014 Jul 22;64(3):304-18. doi: 10.1016/j.jacc.2014.05.027. Erratum in: *J Am Coll Cardiol*. 2014 Sep 16;64(11):1186. Bonow, Robert D [Corrected to Bonow, Robert O]. PMID: 25034069.
34. Arbustini E., Narula N., Dec G.W., et al: The MOGE(S) classification for a phenotype-genotype nomenclature of cardiomyopathy: endorsed by the World Heart Federation. *J Am Coll Cardiol* 2013; 62: pp. 2046-2072
35. Brandenburg RO, Chazov E, Cherian G, False AO, Grosogoeat Y, Kawai C, Loogen F, Marin Judez V, Orinius E, Goodwin JF, Olsen J, Oakley CM, Pisa Z. Report of the WHO/ISFC task force on the definition and classification of cardiomyopathies. *Br Heart J*. 1980; 44:672–673.
36. Ntusi NA. Cardiovascular magnetic resonance imaging in rheumatic heart disease. *Cardiovasc J Afr*. 2018 May/Jun;29(3):135-136. PMID: 30067271; PMCID: PMC6282423
37. Kakou-Guikahue M, N'Guetta R, Anzouan-Kacou JB, Kramoh E, N'Dori R, Ba SA, et al. Optimizing the management of acute coronary syndromes in sub-Saharan Africa: A statement from the AFRICARDIO 2015 Consensus Team. *Archives of cardiovascular diseases*. 2016; 109(6–7): 376–83. DOI: <https://doi.org/10.1016/j.acvd.2015.12.005>
38. Schamroth C. Management of acute coronary syndrome in South Africa: Insights from the ACCESS (Acute Coronary Events – a Multinational Survey of Current Management Strategies) registry. *Cardiovascular Journal of Africa*. 2012; 23(7): 365–70. DOI: <https://doi.org/10.5830/CVJA-2012-017>
39. Yao H, Ekou A, Hadeou A, N'Djessan JJ, Kouame I, N'Guetta R. Medium and long-term follow-up after ST-segment elevation myocardial infarction in a sub-Saharan Africa population: A prospective cohort study. *BMC Cardiovascular Disorders*. 2019; 19(1): 65. DOI: <https://doi.org/10.1186/s12872-019-1043-1>
40. Bonny A, Ngantcha M, Jeilan M, Okello E, Kaviraj B, Talle MA, Nel G, Marijon E, Sani MU, Yousef Z, Karaye KM, Touré IA, Awad MA, Millogo G, Kologo J, Kane A, Houndolo R, Dzudié A, Mbakwem A, Mayosi BM, Chin A. Statistics on the use of cardiac electronic devices and interventional electrophysiological procedures in Africa from 2011 to 2016: report of the Pan African Society of Cardiology (PASCAR) Cardiac Arrhythmias and Pacing Task Forces. *Europace*. 2018 Sep 1;20(9):1513-1526. doi: 10.1093/europace/eux353. PMID: 29309556; PMCID: PMC6123943.
41. Kaviarasan, V., Mohammed, V., & Veerabathiran, R. (2022). Genetic predisposition study of heart failure and its association with cardiomyopathy. *The Egyptian Heart Journal*, 74(1). doi: 10.1186/s43044-022-00240-6
42. Abdulrahim JW, Kwee LC, Alenezi F, et al. Identification of Undetected Monogenic Cardiovascular Disorders. *J Am Coll Cardiol*. 2020 Aug 18;76(7):797-808. doi: 10.1016/j.jacc.2020.06.037. PMID: 32792077; PMCID: PMC7428466. Jawan W. Abdulrahim, MD,a Lydia Coulter Kwee, PHD,a Fawaz Alenezi, MD, et al Identification of Undetected Monogenic Cardiovascular Disorders <https://doi.org/10.1016/j.jacc.2020.06.037>
43. Chong JX, Buckingham KJ, Jhangiani SN, et al. The genetic basis of Mendelian phenotypes: discoveries, challenges, and opportunities. *Am J Hum Genet* 2015;97:199–215.
44. Richard, P., Villard, E., Charron, P., & Isnard, R. (2006). The Genetic Bases of Cardiomyopathies. *Journal Of The American College Of Cardiology*, 48(9), A79-A89. doi: 10.1016/j.jacc.2006.09.014
45. Kamisago M, Sharma SD, DePalma SR, et al. Mutations in sarcomere protein genes as a cause of dilated cardiomyopathy. *N Engl J Med* 2000;343:1688 –96.

46. Villard E, Duboscq-Bidot L, Charron P, et al. Mutation screening in dilated cardiomyopathy: prominent role of the beta myosin heavy chain gene. *Eur Heart J* 2005;26:794–803.
47. Osterziel KJ, Perrot A. Dilated cardiomyopathy: more genes means more phenotypes. *Eur Heart J* 2005;26:751–4.
48. Mayosi BM, Fish M, Shaboodien G, et al. Identification of cadherin 2 (CDH2) mutations in arrhythmogenic right ventricular cardiomyopathy. *Circ Cardiovasc Genet* 2017;10:e001605.
49. Damrauer, S., & Weintraub, W. (2020). Testing for Not so Rare Monogenic Cardiovascular Diseases. *Journal Of The American College Of Cardiology*, 76(7), 809-811. doi: 10.1016/j.jacc.2020.06.060
50. Cahill, T., Ashrafian, H., & Watkins, H. (2013). Genetic Cardiomyopathies Causing Heart Failure. *Circulation Research*, 113(6), 660-675. doi: 10.1161/circresaha.113.300282
51. Xu T, Yang Z, Vatta M, Rampazzo A, Beffagna G, Pilichou K, Scherer SE, Saffitz J, Kravitz J, Zareba W, Danieli GA, Lorenzon A, Nava A, Bauce B, Thiene G, Basso C, Calkins H, Gear K, Marcus F, Towbin JA; Multidisciplinary Study of Right Ventricular Dysplasia Investigators. Compound and digenic heterozygosity contribute to arrhythmogenic right ventricular cardiomyopathy. *J Am Coll Cardiol*. 2010 Feb 9;55(6):587-97. doi: 10.1016/j.jacc.2009.11.020. Erratum in: *J Am Coll Cardiol*. 2010 Mar 30;55(13):1401. Pilichou, Kalliopi [corrected to Pilichou, Kalliopi]. PMID: 20152563; PMCID: PMC2852685.
52. Antzelevitch C. Molecular genetics of arrhythmias and cardiovascular conditions associated with arrhythmias. *Heart Rhythm*. 2004; 1 (5C): 42C–56C.
53. Ali J. Marian and Eugene Braunwald; Hypertrophic Cardiomyopathy; Genetics, Pathogenesis, Clinical Manifestations, Diagnosis, and Therapy. <https://doi.org/10.1161/CIRCRESAHA.117.311059> *Circulation Research*. 2017;121:749–770
54. Mouton JM, van der Merwe L, Goosen A, Revera M, Brink PA, Moolman-Smook JC, Kinnear C. MYBPH acts as modifier of cardiac hypertrophy in hypertrophic cardiomyopathy (HCM) patients. *Hum Genet*. 2016; 135:477–483. doi: 10.1007/s00439-016-1649-7.
55. Nyirenda MJ. Non-communicable diseases in sub-Saharan Africa: understanding the drivers of the epidemic to inform intervention strategies. *Int Health*. 2016 May;8(3):157-8. doi: 10.1093/inthealth/ihw021. PMID: 27178673.
56. Gasnat Shaboodien, Timothy F. Spracklen, Stephen Kamuli, Polycarp Ndibangwi, Carla Van Niekerk, Ntobeko A. B. Ntusi. Genetics of inherited cardiomyopathies in Africa.. <http://dx.doi.org/10.21037/cdt.2019.10.03>
57. Fish M., Shaboodien G., Kraus S., Sliwa K., Seidman C.E., Burke M.A., Crotti L., Schwartz P.J., Mayosi B.M. Mutation analysis of the phospholamban gene in 315 South Africans with dilated, hypertrophic, peripartum and arrhythmogenic right ventricular cardiomyopathies. *Sci. Rep*. 2016;6:22235. doi: 10.1038/srep22235
58. Spracklen, T. F., Keavney, B., Laing, N., Ntusi, N. and Shaboodien, G. "Modern Genomic Techniques In The Identification Of Genetic Causes Of Cardiomyopathy". *Heart*, 2022, pp. heartjnl-2021-320424. *BMJ*, doi:10.1136/heartjnl-2021-320424.
59. Watkins, D. A., Hendricks, N., Shaboodien, G., Mbele, M., Parker, M., Vezi, B. Z., . . . Africa, Arvc Registry of the Cardiac Arrhythmia Society of Southern. (2009). Clinical features, survival experience, and profile of plakophilin-2 gene mutations in participants of the arrhythmogenic right ventricular cardiomyopathy registry of South Africa. *Heart Rhythm*, 6(11 Suppl), S10-17. doi: 10.1016/j.hrthm.2009.08.018
60. Caviedes Bottner P, Córdova Fernández T, Larraín Valenzuela M; Cruces Romero Presentación de Casos Clínicos P. Dilated cardiomyopathy and severe heart failure An update for pediatricians. *Arch Argent Pediatr*. 2018;116:e421–8

61. Trivedi A., and Knight B.P.: ICD therapy for primary prevention in hypertrophic cardiomyopathy. *Arrhythm Electrophysiol Rev* 2016; 5: pp. 188-196
62. Calkins H., Corrado D., and Marcus F.: Risk stratification in arrhythmogenic right ventricular cardiomyopathy. *Circulation* 2017; 136: pp. 2068-2082
63. Ponikowski P, Voors AA, Anker SD, et al. 2016 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: the task force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur J Heart Fail* 2016;18:891–975
64. Job A. J. Verdonschot, Mark R. Hazebroek, James S. Ware, Sanjay K. Prasad, Stephane R. B. Heymans. 2019 Role of Targeted Therapy in Dilated Cardiomyopathy: The Challenging Road Toward a Personalized Approach 10.1161/jaha.119.012514
65. Helms, Adam S., Andrea D. Thompson, and Sharlene M. Day. 2022. "Translation Of New And Emerging Therapies For Genetic Cardiomyopathies". *JACC: Basic To Translational Science* 7 (1): 70-83. doi:10.1016/j.jacbts.2021.07.012.
66. Kieserman JM, Myers VD, Dubey P, Cheung JY, Feldman AM. Current Landscape of Heart Failure Gene Therapy. *J Am Heart Assoc.* 2019 May 21;8(10):e012239. doi: 10.1161/JAHA.119.012239. PMID: 31070089; PMCID: PMC6585330.
67. Antoine Bondue, Eloisa Arbustini, Anna Bianco, Michele Ciccarelli, Dana Dawson, Matteo De Rosa, Nazha Hamdani, Denise Hilfiker-Kleiner, Benjamin Meder, Adelino F Leite-Moreira, Thomas Thum, Carlo G Tocchetti, Gilda Varricchi, Jolanda Van der Velden, Roddy Walsh, Stephane Heymans, Complex roads from genotype to phenotype in dilated cardiomyopathy: scientific update from the Working Group of Myocardial Function of the European Society of Cardiology, Cardiovascular Research, Volume 114, Issue 10, 01 August 2018, Pages 1287–1303, <https://doi.org/10.1093/cvr/cvy122>
68. Daubeney P, Nugent A, Chondros P, et al. National Australian Childhood Cardiomyopathy Study. Clinical features and outcomes of childhood dilated cardiomyopathy: results from a national population-based study. *Circulation* 2006;114(24):2671-8
69. Naudi D, Rossano JW. Epidemiology and Cost of heart failure in children. *Cardiol Young* 2015;25(8):1460-8
70. Ntusi NBA, Mayosi BM. Epidemiology of heart failure in sub-Saharan Africa. *Expert Rev Cardiovasc Ther.* 2009;7(2):169–80
71. Reichart, D., Magnussen, C., Zeller, T., & Blankenberg, S. (2019). Dilated cardiomyopathy: from epidemiologic to genetic phenotypes. *Journal Of Internal Medicine*, 286(4), 362-372. doi: 10.1111/joim.12944
72. Bradlow, B. A., Zion, M. M., & Fleishman, S. J. (1964). Heart Disease in Africa, with Particular Reference to Southern Africa. *Am J Cardiol*, 13, 650-669
73. Cosnett, J. E. (1962). Heart disease in the Zulu: especially cardiomyopathy and cardiac infarction. *Br Heart J*, 24, 76-82.
74. Powell, S. J., & Wright, R. (1965). Cardiomyopathy in Durban. *S Afr Med J*, 39(42), 1062-1066.
75. McGlashan, N. D. (1988). Southern African cardiomyopathy in the Republic of South Africa, 1978-1980. *Afr J Med Med Sci*, 17(1), 33-46
76. Steenekamp, J. H., Simson, I. W., & Theron, W. (1992). Cardiovascular causes of death at Tshepong Hospital in 1 year, 1989-1990. A necropsy study. *S Afr Med J*, 81(3), 142-146
77. Ntusi, N. B., Badri, M., Gumedze, F., Wonkam, A., & Mayosi, B. M. (2011). Clinical characteristics and outcomes of familial and idiopathic dilated cardiomyopathy in Cape Town: a comparative study of 120 cases followed up over 14 years. *S Afr Med J*, 101(6), 399-404

78. Batchelder, K., & Mayosi, B. M. (2005). Pentoxifylline for heart failure: a systematic review. *S Afr Med J*, 95(3), 171-175
79. Bozkurt B, Colvin M, Cook J, Cooper LT, Deswal A, Fonarow GC, et al. Current diagnostic and treatment strategies for specific dilated cardiomyopathies: A scientific statement from the American Heart Association. *Circulation*. 2016;134:579–646. Doi: <https://doi.org/10.1161/CIR.0000000000000455>. A summary of knowledge regarding DCM, emphasizing recent developments in therapeutic approaches for cardiomyopathies
80. Bakalakos A, Risatos K, Anastasakis A. Current perspective on the diagnosis and management of dilated cardiomyopathy Beyond heart failure: a Cardiomyopathy Clinic Doctor's point of view. *Hellenic Journal of Cardiology*. 2018. <https://doi.org/10.1016/j.hjc/2018.05.008> Provides a comprehensive review of DCM diagnosis and management, including new classification concepts and identified genes by the American Heart Association
81. Herman DS, Lam L, Taylor MR, et al. Truncations of titin causing dilated cardiomyopathy. *N Engl J Med* 2012;366:619–28
82. Zimmerman RS, Cox S, Lakdawala NK, et al. A novel custom resequencing array for dilated cardiomyopathy. *Genet Med* 2010;12:268–78
83. Mestroni L, Brun F, Spezzacatene A, Sinagra G, Taylor MRG. Genetic causes of dilated cardiomyopathy. *Prog Pediatr Cardiol*. 2014;37:13–8. <https://doi.org/10.1016/j.ppedcard.2014.10.003>
84. Mestroni L, Rocco C, Gregori D, Sinagra G, Lenarda A, Miocic S. Familial dilated cardiomyopathy: evidence for genetic and phenotypic heterogeneity. *J Am Coll Cardiol*. 1999;34:181–90. [https://doi.org/10.1016/S0735-1097\(99\)00172-2](https://doi.org/10.1016/S0735-1097(99)00172-2)
85. Hershberger RE, Hedges DJ, Morales A (2013) Dilated cardiomyopathy: the complexity of a diverse genetic architecture. *Nat Rev Cardiol*; 10(9):531–47
86. Mazzarotto F, Tayal U, Buchan RJ, Midwinter W, Wilk A, Whiffin N, Govind R, Mazaika E, de Marvao A, Dawes TJW, Felkin LE, Ahmad M, Theotokis PI, Edwards E, Ing AY, Thomson KL, Chan LLH, Sim D, Baksi AJ, Pantazis A, Roberts AM, Watkins H, Funke B, O'Regan DP, Olivetto I, Barton PJR, Prasad SK, Cook SA, Ware JS, Walsh R. Reevaluating the Genetic Contribution of Monogenic Dilated Cardiomyopathy. *Circulation*. 2020 Feb 4;141(5):387-398. doi: 10.1161/CIRCULATIONAHA.119.037661. Epub 2020 Jan 27. PMID: 31983221; PMCID: PMC7004454.
87. Chin TK, Perloff JK, Williams RG, Jue K, Mohrmann R. Isolated noncompaction of left ventricular myocardium. A study of eight cases. *Circulation*. 1990 Aug;82(2):507-13. doi: 10.1161/01.cir.82.2.507. PMID: 2372897.
88. Dusek J, Ostádal B, Duskova M. Postnatal persistence of spongy myocardium with embryonic blood supply. *Arch Pathol*. 1975 Jun;99(6):312-7. PMID: 1147832.
89. Jenni R, Goebel N, Tartini R, Schneider J, Arbenz U, Oelz O. Persisting myocardial sinusoids of both ventricles as an isolated anomaly: echocardiographic, angiographic, and pathologic anatomical findings. *Cardiovasc Intervent Radiol*. 1986;9(3):127-31. doi: 10.1007/BF02577920. PMID: 3089618.
90. Ichida F. Left ventricular noncompaction - Risk stratification and genetic consideration. *J Cardiol*. 2020 Jan;75(1):1-9. doi: 10.1016/j.jjcc.2019.09.011. Epub 2019 Oct 17. PMID: 31629663.
91. Kovacevic-Preradovic T, Jenni R, Oechslin EN, Noll G, Seifert B, Attenhofer Jost CH. Isolated left ventricular noncompaction as a cause for heart failure and heart transplantation: a single center experience. *Cardiology*. 2009;112(2):158-64. doi: 10.1159/000147899. Epub 2008 Jul 25. PMID: 18663296.
92. Sedaghat-Hamedani, Farbod, Jan Haas, Feng Zhu, Christian Geier, Elham Kayvanpour, Martin Liss, and Alan Lai et al. 2017. "Clinical Genetics And Outcome Of Left Ventricular Non-Compaction Cardiomyopathy". *European Heart Journal* 38 (46): 3449-3460. doi:10.1093/eurheartj/ehx545.

93. Mazzarotto, Francesco, Megan H. Hawley, Matteo Beltrami, Leander Beekman, Antonio de Marvao, Kathryn A. McGurk, and Ben Statton et al. 2021. "Systematic Large-Scale Assessment Of The Genetic Architecture Of Left Ventricular Noncompaction Reveals Diverse Etiologies". *Genetics In Medicine* 23 (5): 856-864. doi:10.1038/s41436-020-01049-x.
94. Towbin JA, McKenna WJ, Abrams DJ, Ackerman MJ, Calkins H, Darrieux FCC, Daubert JP, de Chillou C, DePasquale EC, Desai MY, Estes NAM 3rd, Hua W, Indik JH, Ingles J, James CA, John RM, Judge DP, Keegan R, Krahn AD, Link MS, Marcus FI, McLeod CJ, Mestroni L, Priori SG, Saffitz JE, Sanatani S, Shimizu W, van Tintelen JP, Wilde AAM, Zareba W. 2019 HRS expert consensus statement on evaluation, risk stratification, and management of arrhythmogenic cardiomyopathy. *Heart Rhythm*. 2019 Nov;16(11):e301-e372. doi: 10.1016/j.hrthm.2019.05.007. Epub 2019 May 9. PMID: 31078652.
95. Corrado D, Link MS, Calkins H. Arrhythmogenic right ventricular cardiomyopathy. *N Engl J Med* 2017;376:61–72.
96. Baucé B et al. Clinical phenotype and diagnosis of arrhythmogenic right ventricular cardiomyopathy in pediatric patients carrying desmosomal gene mutations. *Heart Rhythm* 8, 1686–1695 (2011). [PMC free article] [PubMed] [Google Scholar]
97. Te Riele ASJM et al. Arrhythmogenic right ventricular dysplasia/cardiomyopathy in the pediatric population: clinical characterization and comparison with adult-onset disease. *JACC Clin. Electrophysiol* 1, 551–560 (2015).
98. Limongelli G, Rea A, Masarone D, et al. Right ventricular cardiomyopathies: a multidisciplinary approach to diagnosis. *Echocardiography* 2015; 32:75–94.
99. Sen-Chowdhry S et al. Left-dominant arrhythmogenic cardiomyopathy: an under-recognized clinical entity. *J. Am. Coll. Cardiol* 52, 2175–2187 (2008).
100. Martins D et al. Myocardial inflammation detected by cardiac MRI in arrhythmogenic right ventricular cardiomyopathy: a paediatric case series. *Int. J. Cardiol* 271, 81–86 (2018).
101. Sen-Chowdhry S et al. Clinical and genetic characterization of families with arrhythmogenic right ventricular dysplasia/cardiomyopathy provides novel insights into patterns of disease expression. *Circulation* 115, 1710–1720 (2007).
102. Gregoire FM, Smas CM, Sul HS. Understanding adipocyte differentiation. *Physiological Reviews*. 1998;78:783–809. doi: 10.1152/physrev.1998.78.3.783.
103. Austin K.M., Trembley M.A., Chandler S.F., Sanders S.P., Saffitz J.E., Abrams D.J., Pu W.T. Molecular mechanisms of arrhythmogenic cardiomyopathy. *Nat. Rev. Cardiol*. 2019;16:519–537. doi: 10.1038/s41569-019-0200-7.
104. Basso C, Baucé B, Corrado D, Thiene G. Pathophysiology of arrhythmogenic cardiomyopathy. *Nature Reviews. Cardiology*. 2012;9:223–233. doi: 10.1038/nrcardio.2011.173.
105. Sommariva E, Stadiotti I, Perrucci GL, Tondo C, Pompilio G. Cell models of arrhythmogenic cardiomyopathy: advances and opportunities. *Disease Models & Mechanisms*. 2017;10:823–835. doi: 10.1242/dmm.029363.
106. Nava A, Baucé B, Basso C, Muriago M, Rampazzo A, Villanova C, Daliento L, Buja G, Corrado D, Danieli GA, Thiene G. Clinical profile and long-term follow-up of 37 families with arrhythmogenic right ventricular cardiomyopathy. *J Am Coll Cardiol*. 2000;36:2226–2233
107. Corrado, D., Basso, C., Schiavon, M., & Thiene, G. (1998). Screening for hypertrophic cardiomyopathy in young athletes. *N Engl J Med*, 339(6), 364-369. doi: 10.1056/NEJM199808063390602
108. Tabib, A., Loire, R., Chalabreysse, L., Meyronnet, D., Miras, A., Malicier, D., . . . Bouvagnet, P. (2003). Circumstances of death and gross and microscopic observations in a series of 200 cases of sudden death associated with arrhythmogenic right ventricular cardiomyopathy and/or dysplasia. *Circulation*, 108(24), 3000-3005. doi: 10.1161/01.CIR.0000108396.65446.21

109. Cynthia A James, Petros Syrris, J Peter van Tintelen, Hugh Calkins. The role of genetics in cardiovascular disease: arrhythmogenic cardiomyopathy. *European Heart Journal*, Volume 41, Issue 14, 7 April 2020, Pages 1393–1400, <https://doi.org/10.1093/eurheartj/ehaa141>
110. Awad MM, Calkins H, Judge DP. Mechanisms of disease: molecular genetics of arrhythmogenic right ventricular dysplasia/cardiomyopathy. *Nat Clin Pract Cardiovasc Med* 2008;5:258–67.
111. Patel, Viraj & Asatryan, Babken & Siripanthong, Bhurint & Munroe, Patricia & Owens, Anjali & Lopes, Luis & Khanji, Mohammed & Protonotarios, Alexandros & Santangeli, Pasquale & Muser, Daniele & Marchlinski, Francis & Brady, Peter & Chahal, Anwar A. (2020). State of the Art Review on Genetics and Precision Medicine in Arrhythmogenic Cardiomyopathy. *International Journal of Molecular Sciences*. 21. 6615. 10.3390/ijms21186615.
112. Hershberger RE, Lindenfeld J, Mestroni L, Seidman CE, Taylor MRG, Towbin JA, Heart Failure Society of America. Genetic evaluation of cardiomyopathy—a Heart Failure Society of America practice guideline. *J Card Fail* 2009;15:83–97
113. Vincent AT, Derome N, Boyle B, Culley AI, Charette SJ. Next-generation sequencing (NGS) in the microbiological world: How to make the most of your money. *J Microbiol Methods*. 2017 Jul;138:60–71. doi: 10.1016/j.mimet.2016.02.016. Epub 2016 Mar 16. PMID: 26995332.
114. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 2005 Sep 15;437(7057):376–80. doi: 10.1038/nature03959. Epub 2005 Jul 31. Erratum in: *Nature*. 2006 May 4;441(7089):120. Ho, Chun He [corrected to Ho, Chun Heen]. PMID: 16056220; PMCID: PMC1464427.
115. Favalli V, Serio A, Grasso M, Arbustini E. Genetic causes of dilated cardiomyopathy. *Heart*. 2016;102:2004–14.
116. "An Overview Of Next-Generation Sequencing". 2022. *Genomics Research From Technology Networks*. <https://www.technologynetworks.com/genomics/articles/an-overview-of-next-generation-sequencing-346532>.
117. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet*. 2016 May 17;17(6):333–51. doi: 10.1038/nrg.2016.49. PMID: 27184599.
118. Buermans HP, den Dunnen JT. Next generation sequencing technology: Advances and applications. *Biochim Biophys Acta*. 2014;1842(10):1932–1941. doi: 10.1016/j.bbadis.2014.06.015.
119. Rothberg, J., Hinz, W., Rearick, T. *et al.* An integrated semiconductor device enabling non-optical genome sequencing. *Nature* **475**, 348–352 (2011). <https://doi.org/10.1038/nature10242>
120. Ronaghi M, Karamohamed S, Pettersson B, Uhlén M, Nyrén P. Real-time DNA sequencing using detection of pyrophosphate release. *Anal Biochem*. 1996 Nov 1;242(1):84–9. doi: 10.1006/abio.1996.0432. PMID: 8923969.
121. Slatko BE, Gardner AF, Ausubel FM. Overview of Next-Generation Sequencing Technologies. *Curr Protoc Mol Biol*. 2018 Apr;122(1):e59. doi: 10.1002/cpmb.59. PMID: 29851291; PMCID: PMC6020069.
122. Porreca, G. Genome sequencing on nanoballs. *Nat Biotechnol* **28**, 43–44 (2010).
123. Drmanac R, Sparks AB, Callow MJ, Halpern AL, Burns NL, Kermani BG, Carnevali P, Nazarenko I, Nilsen GB, Yeung G, Dahl F, Fernandez A, et al. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*. 2010 Jan 1;327(5961):78–81. doi: 10.1126/science.1181498. Epub 2009 Nov 5. PMID: 19892942.
124. Kumar KR, Cowley MJ, Davis RL. Next-Generation Sequencing and Emerging Technologies. *Semin Thromb Hemost*. 2019 Oct;45(7):661–673. doi: 10.1055/s-0039-1688446. Epub 2019 May 16. PMID: 31096307.

125. Bewicke-Copley F, Arjun Kumar E, Palladino G, Korfi K, Wang J. Applications and analysis of targeted genomic sequencing in cancer studies. *Comput Struct Biotechnol J*. 2019 Nov 7;17:1348-1359. doi: 10.1016/j.csbj.2019.10.004. PMID: 31762958; PMCID: PMC6861594.
126. Gorcenco S, Ilinca A, Almasoudi W, Kafantari E, Lindgren AG, Puschmann A. New generation genetic testing entering the clinic. *Parkinsonism Relat Disord*. 2020 Apr;73:72-84. doi: 10.1016/j.parkreldis.2020.02.015. Epub 2020 Mar 2. PMID: 32273229.
127. Fata F, Vecoli C, Foffa I, Andreassi MG. Next generation sequencing in cardiovascular diseases. *World J Cardiol*. 2012 Oct 26;4(10):288-95. doi: 10.4330/wjc.v4.i10.288. PMID: 23110245; PMCID: PMC3482622.
128. Ntusi N. A. B., Shaboodien G., Badri M., Gumede F. & Mayosi B. M. Clinical features, spectrum of causal genetic mutations, and outcome of hypertrophic cardiomyopathy in South Africans. *Cardiovasc J Afr* 2016;27:152-8.
129. Ramchand J, Wallis M, Macciocca I, Lynch E, Farouque O, Martyn M, Phelan D, Chong B, Lockwood S, Weintraub R, Thompson T, Trainer A, Zentner D, Vohra J, Chetrit M, Hare DL, James P. Prospective Evaluation of the Utility of Whole Exome Sequencing in Dilated Cardiomyopathy. *J Am Heart Assoc*. 2020 Jan 21;9(2):e013346. doi: 10.1161/JAHA.119.013346. Epub 2020 Jan 14. PMID: 31931689; PMCID: PMC7033851.
130. Minoche AE, Horvat C, Johnson R, et al. Genome sequencing as a first-line genetic test in familial dilated cardiomyopathy. *Genet Med*. 2019;21(3):650-662. doi:10.1038/s41436-018-0084-7
131. Duncker DJ, Bakkers J, Brundel BJ, Robbins J, Tardiff JC, Carrier L. Animal and in silico models for the study of sarcomeric cardiomyopathies. *Cardiovasc Res*. 2015 Apr 1;105(4):439-48. doi: 10.1093/cvr/cvv006. Epub 2015 Jan 18. PMID: 25600962; PMCID: PMC4375391.
132. Doppler SA, Carvalho C, Lahm H, Deutsch MA, Dreßen M, Puluca N, Lange R, Krane M. Cardiac fibroblasts: more than mechanical support. *J Thorac Dis* 2017; 9: S36-S51 [PMID: 28446967 DOI:10.21037/jtd.2017.03.122]
133. Lou W, Reynolds CA, Li Y, Liu J, Hüttemann M, Schlame M, Stevenson D, Strathdee D, Greenberg ML. Loss of tafazzin results in decreased myoblast differentiation in C2C12 cells: A myoblast model of Barth syndrome and cardiolipin deficiency. *Biochim Biophys Acta Mol Cell Biol Lipids* 2018; 1863: 857-865 [PMID: 29694924 DOI: 10.1016/j.BBALIP.2018.04.015]
134. Welser J. Primary Cells Versus Cell Lines [Internet]. ScienCell Research Laboratories. 2015; [cited 2018 Jun 30] Available from: <https://sciencellonline.com/blog/primary-cells-versus-cell-lines>
135. Shaw SY, Brettman AD. Phenotyping patient-derived cells for translational studies in cardiovascular disease. *Circulation* 2011; 124: 2444-2455 [PMID: 22125190 DOI: 10.1161/CIRCULATIONAHA.111.043943]
136. Nonaka, M. (2014). Experimental models of inherited cardiomyopathy and its therapeutics. *World Journal Of Cardiology*, 6(12), 1245. doi: 10.4330/wjc.v6.i12.1245
137. Houser SR, Margulies KB, Murphy AM, Spinale FG, Francis GS, Prabhu SD, Rockman HA, Kass DA, Molkentin JD, Sussman MA, Koch WJ; American Heart Association Council on Basic Cardiovascular Sciences, Council on Clinical Cardiology, and Council on Functional Genomics and Translational Biology. Animal models of heart failure: a scientific statement from the American Heart Association. *Circ Res*. 2012 Jun 22;111(1):131-50. doi: 10.1161/RES.0b013e3182582523. Epub 2012 May 17. Erratum in: *Circ Res*. 2012 Jul 6;111(2):e54. PMID: 22595296.
138. Duncker DJ, Bakkers J, Brundel BJ, Robbins J, Tardiff JC, Carrier L. Animal and in silico models for the study of sarcomeric cardiomyopathies. *Cardiovasc Res*. 2015 Apr 1;105(4):439-48. doi: 10.1093/cvr/cvv006. Epub 2015 Jan 18. PMID: 25600962; PMCID: PMC4375391.
139. Kittleson MD, Meurs KM, Harris SP. The genetic basis of hypertrophic cardiomyopathy in cats and humans. *Journal of Veterinary Cardiology*. 2015;17(Suppl 1):S53-S73

140. Ackerman MJ, Priori SG, Willems S, Berul C, Brugada R, Calkins H, et al. HRS/EHRA expert consensus statement on the state of genetic testing for the channelopathies and cardiomyopathies. *Europace*. 2011;13:1077–109.
141. M, Löbe, Meineke F, and Winter A. 2019. "Scenarios For Using Openclinica In Academic Clinical Trials". *Studies In Health Technology And Informatics* 258. <https://pubmed.ncbi.nlm.nih.gov/30942748/>.
142. Gessner, Sophia, Michael Storck, Stefan Hegselmann, Martin Dugas, and I&#241;aki Soto-Rey. 2017. "Automated Transformation Of CDISC ODM To Openclinica". IOS Press, 95-99. <https://ebooks.iospress.nl/publication/47514>.
143. Lin Liu, Yinhu Li, Siliang Li, Ni Hu, Yimin He, Ray Pong, Danni Lin, Lihua Lu, Maggie Law, "Comparison of Next-Generation Sequencing Systems", *BioMed Research International*, vol. 2012, Article ID 251364, 11 pages, 2012. <https://doi.org/10.1155/2012/251364>
144. Richards, Sue, Nazneen Aziz, Sherri Bale, David Bick, Soma Das, Julie Gastier-Foster, and Wayne W. Grody et al. 2015. "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". *Genetics In Medicine* 17 (5): 405-424. doi:10.1038/gim.2015.30.
145. McLaren, William, Laurent Gil, Sarah E. Hunt, Harpreet Singh Riat, Graham R. S. Ritchie, Anja Thormann, Paul Flicek, and Fiona Cunningham. 2016. "The Ensembl Variant Effect Predictor". *Genome Biology* 17 (1). doi:10.1186/s13059-016-0974-4.
146. 2022. *Clinicalgenome.Org*. [https://www.clinicalgenome.org/site/assets/files/2071/clingen\\_cmp\\_acmg\\_specifications\\_v1-1.pdf](https://www.clinicalgenome.org/site/assets/files/2071/clingen_cmp_acmg_specifications_v1-1.pdf).
147. Kelley, Lawrence A, Stefans Mezulis, Christopher M Yates, Mark N Wass, and Michael J E Sternberg. 2015. "The Phyre2 Web Portal For Protein Modeling, Prediction And Analysis". *Nature Protocols* 10 (6): 845-858. doi:10.1038/nprot.2015.053.
148. Neupane, Chetanath & Ramakrishna, Seeram & Nakarmi, Jeevan & Adhikari, Binod. (2020). Molecular Structure Properties of Heme Group in the Oxymyoglobin Protein (1mbo.pdb) Using PyMOL & UCSF Chimera. *Nanoarchitectonics*. 1. 87-92. 10.37256/nat.122020425.
149. Worth CL, Preissner R, Blundell TL. SDM--a server for predicting effects of mutations on protein stability and malfunction. *Nucleic Acids Res*. 2011 Jul;39(Web Server issue):W215-22. doi: 10.1093/nar/gkr363. Epub 2011 May 18. PMID: 21593128; PMCID: PMC3125769.
150. Jan MF, Tajik AJ. Modern Imaging Techniques in Cardiomyopathies. *Circ Res*. 2017 Sep 15;121(7):874-891. doi: 10.1161/CIRCRESAHA.117.309600. PMID: 28912188.
151. Marcus, F., McKenna, W., Sherrill, D., Basso, C., Bauce, B., & Bluemke, D. et al. (2010). Diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia: Proposed Modification of the Task Force Criteria. *European Heart Journal*, 31(7), 806-814. doi: 10.1093/eurheartj/ehq025
152. Hamid MS, Norman M, Quraishi A, Firoozi S, Thaman R, Gimeno JR, Sachdev B, Rowland E, Elliott PM, McKenna WJ. Prospective evaluation of relatives for familial arrhythmogenic right ventricular cardiomyopathy/dysplasia reveals a need to broaden diagnostic criteria. *J Am Coll Cardiol*. 2002 Oct 16;40(8):1445-50. doi: 10.1016/s0735-1097(02)02307-0. PMID: 12392835.
153. Miles C, Fanton Z, Tome M, Behr ER. Inherited cardiomyopathies. *BMJ*. 2019 May 2;365:l1570. doi: 10.1136/bmj.l1570. PMID: 31048310.
154. Steenekamp JH, Simson IW, Theron W. Cardiovascular causes of death at Tshepong Hospital in 1 year, 1989-1990. A necropsy study. *S Afr Med J*. 1992 Feb 1;81(3):142-6. PMID: 1734552.
155. Bakalakos A, Risatos K, Anastasakis A. Current perspective on the diagnosis and management of dilated cardiomyopathy Beyond heart failure: a Cardiomyopathy Clinic Doctor's point of view.

- Hellenic Journal of Cardiology. 2018. <https://doi.org/10.1016/j.hjc/2018.05.008> Provides a comprehensive review of DCM diagnosis and management, including new classification concepts and identified genes by the American Heart Association.
156. Atteya G, Lampert R. Sudden Cardiac Death in Genetic Cardiomyopathies. *Card Electrophysiol Clin*. 2017 Dec;9(4):581-603. doi: 10.1016/j.ccep.2017.07.009. PMID: 29173404.
  157. Zhang, Xinlin et al. "Next-Generation Sequencing Identifies Pathogenic And Modifier Mutations In A Consanguineous Chinese Family With Hypertrophic Cardiomyopathy". *Medicine*, vol 96, no. 24, 2017, p. e7010. Ovid Technologies (Wolters Kluwer Health), doi:10.1097/md.0000000000007010.
  158. Jonathan J. Magaña, Rocío Suárez-Sánchez, Norberto Leyva-García, Bulmaro Cisneros and Oscar Hernández-Hernández (2012). *Myotonic Dystrophy Protein Kinase: Structure, Function and Its Possible Role in the Pathogenesis of Myotonic Dystrophy Type 1*, *Advances in Protein Kinases*, Dr. Gabriela Da Silva Xavier (Ed.), ISBN: 978-953-51-0633-3, InTech,
  159. Kaliman P, Catalucci D, Lam JT, Kondo R, Gutiérrez JC, Reddy S, Palacín M, Zorzano A, Chien KR, Ruiz-Lozano P. Myotonic dystrophy protein kinase phosphorylates phospholamban and regulates calcium uptake in cardiomyocyte sarcoplasmic reticulum. *J Biol Chem*. 2005 Mar 4;280(9):8016-21. doi: 10.1074/jbc.M412845200. Epub 2004 Dec 13. PMID: 15598648.
  160. Konrad F Frank, Birgit Bölk, Erland Erdmann, Robert H.G Schwinger, *Sarcoplasmic reticulum Ca<sup>2+</sup>-ATPase modulates cardiac contraction and relaxation*, *Cardiovascular Research*, Volume 57, Issue 1, January 2003, Pages 20–27, [https://doi.org/10.1016/S0008-6363\(02\)00694-6](https://doi.org/10.1016/S0008-6363(02)00694-6)
  161. MacLennan, D., Kranias, E. Phospholamban: a crucial regulator of cardiac contractility. *Nat Rev Mol Cell Biol* 4, 566–577 (2003). <https://doi.org/10.1038/nrm1151>
  162. Dumas AR, Wisniewsky C, R. BK, Keurs HT, Fiszman MY, Schwartz K. The sarco(endo)plasmic reticulum Ca<sup>2+</sup>-ATPase gene is regulated at the transcriptional level during compensated left ventricular hypertrophy in the rat. *Life Sci* 1997; 320: 963-9.
  163. Simmerman, Heather & JONES, LARRY. (1998). Simmerman, H. K. & Jones, L. R. Phospholamban: protein structure, mechanism of action, and role in cardiac function. *Physiol. Rev.* 78, 921-947. *Physiological reviews.* 78. 921-47. 10.1152/physrev.1998.78.4.921.
  164. Fujii, J., Lytton, J., Tada, M., & MacLennan, D. (1988). Rabbit cardiac and slow-twitch muscle express the same phospholamban gene. *FEBS Letters*, 227(1), 51-55. doi: 10.1016/0014-5793(88)81412-1
  165. Abrol N, de Tombe PP, Robia SL. Acute inotropic and lusitropic effects of cardiomyopathic R9C mutation of phospholamban. *J Biol Chem*. 2015 Mar 13;290(11):7130-40. doi: 10.1074/jbc.M114.630319. Epub 2015 Jan 15. PMID: 25593317; PMCID: PMC4358133.
  166. Pall GS, Johnson KJ, Smith GL. Abnormal contractile activity and calcium cycling in cardiac myocytes isolated from DMPK knockout mice. *Physiol Genomics*. 2003 Apr 16;13(2):139-46. doi: 10.1152/physiolgenomics.00107.2002. PMID: 12595579.
  167. Wahbi, K., & Furling, D. (2020). Cardiovascular manifestations of myotonic dystrophy. *Trends In C Hong Jiang, Ami Mankodi, Maurice S. Swanson, Richard T. Moxley, Charles A. Thornton*, *Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons*, *Human Molecular Genetics*, Volume 13, Issue 24, 15 December 2004, Pages 3079–3088, <https://doi.org/10.1093/hmg/ddh327> *Cardiovascular Medicine*, 30(4), 232-238. doi: 10.1016/j.tcm.2019.06.001
  168. Mankodi A, Lin X, Blaxall BC, Swanson MS, Thornton CA. Nuclear RNA foci in the heart in myotonic dystrophy. *Circ Res*. 2005 Nov 25;97(11):1152-5. doi: 10.1161/01.RES.0000193598.89753.e3. Epub 2005 Oct 27. PMID: 16254211.
  169. Lee HC, Patel MK, Mistry DJ, Wang Q, Reddy S, Moorman JR, Mounsey JP. Abnormal Na channel gating in murine cardiac myocytes deficient in myotonic dystrophy protein kinase. *Physiol Genomics*. 2003 Jan 15;12(2):147-57. doi: 10.1152/physiolgenomics.00095.2002. PMID: 12454205.

170. Wansink, Derick G. et al. "Alternative Splicing Controls Myotonic Dystrophy Protein Kinase Structure, Enzymatic Activity, And Subcellular Localization". *Molecular And Cellular Biology*, vol 23, no. 16, 2003, pp. 5489-5501. American Society For Microbiology, doi:10.1128/mcb.23.16.5489-5501.2003.
171. Kaliman, Perla, and Esther Llagostera. "Myotonic Dystrophy Protein Kinase (DMPK) And Its Role In The Pathogenesis Of Myotonic Dystrophy 1". *Cellular Signalling*, vol 20, no. 11, 2008, pp. 1935-1941. Elsevier BV, doi:10.1016/j.cellsig.2008.05.005.
172. Harmon EB, Harmon ML, Larsen TD, Paulson AF, Perryman MB. Myotonic dystrophy protein kinase is expressed in embryonic myocytes and is required for myotube formation. *Dev Dyn*. 2008 Sep;237(9):2353-66. doi: 10.1002/dvdy.21653. PMID: 18729234.
173. A Role for Myotonic Dystrophy Protein Kinase in Synaptic Plasticity. Paul E. Schulz, Adeka D. McIntosh, Michael R. Kasten, Berend Wieringa, and Henry F. Epstein. 2003 <https://doi.org/10.1152/jn.00504.2002>
174. Grogan A, Kontrogianni-Konstantopoulos A. Unraveling obscurins in heart disease. *Pflugers Arch*. 2019 May;471(5):735-743. doi: 10.1007/s00424-018-2191-3. Epub 2018 Aug 11. PMID: 30099631; PMCID: PMC6397709.
175. Kontrogianni-Konstantopoulos A, Ackermann MA, Bowman AL, Yap SV, Bloch RJ (2009) Muscle giants: molecular scaffolds in sarcomerogenesis. *Physiol Rev* 89:1217–1267. 10.1152/physrev.00017.2009
176. Kontrogianni-Konstantopoulos A, Catino DH, Strong JC, Sutter S, Borisov AB, Pumpilin DW, Russell MW, Bloch RJ (2006) Obscurin modulates the assembly and organization of sarcomeres and the sarcoplasmic reticulum. *FASEB J* 20:2102–2111. 10.1096/fj.06-5761com
177. Kontrogianni-Konstantopoulos A, Jones EM, Van Rossum DB, Bloch RJ (2003) Obscurin is a ligand for small ankyrin 1 in skeletal muscle. *Mol Biol Cell* 14:1138–1148. 10.1091/mbc.E02-07-0411
178. Lange S, Ouyang K, Meyer G, Cui L, Cheng H, Lieber RL, Chen J (2009) Obscurin determines the architecture of the longitudinal sarcoplasmic reticulum. *J Cell Sci* 122:2640–2650. 10.1242/jcs.046193
179. Raeker MO, Bieniek AN, Ryan AS, Tsai HJ, Zahn KM, Russell MW (2010) Targeted deletion of the zebrafish obscurin a RhoGEF domain affects heart, skeletal muscle and brain development. *Dev Biol* 337:432–443. 10.1016/j.ydbio.2009.11.018
180. Marston, Steven. "Obscurin Variants And Inherited Cardiomyopathies". *Biophysical Reviews*, vol 9, no. 3, 2017, pp. 239-243. Springer Science And Business Media LLC, doi:10.1007/s12551-017-0264-8.
181. Manring HR, Carter OA, Ackermann MA. Obscure functions: the location-function relationship of obscurins. *Biophys Rev*. 2017 Jun;9(3):245-258. doi: 10.1007/s12551-017-0254-x. Epub 2017 Mar 29. PMID: 28510116; PMCID: PMC5498325.
182. Raeker MO, Bieniek AN, Ryan AS, Tsai HJ, Zahn KM, Russell MW (2010) Targeted deletion of the zebrafish obscurin a RhoGEF domain affects heart, skeletal muscle and brain development. *Dev Biol* 337:432–443. 10.1016/j.ydbio.2009.11.018
183. White CD, Erdemir HH, Sacks DB. IQGAP1 and its binding proteins control diverse biological functions. *Cell Signal*. 2012 Apr;24(4):826-34. doi: 10.1016/j.cellsig.2011.12.005. Epub 2011 Dec 11. PMID: 22182509; PMCID: PMC3268868.
184. Kim H, White CD, Sacks DB. IQGAP1 in microbial pathogenesis: Targeting the actin cytoskeleton. *FEBS Lett*. 2011;585:723–729.
185. Briggs MW, Sacks DB. IQGAP proteins are integral components of cytoskeletal regulation. *EMBO Rep*. 2003;4:571–574
186. Noritake J, Watanabe T, Sato K, Wang S, Kaibuchi K. IQGAP1: a key regulator of adhesion and migration. *J. Cell Sci*. 2005;118:2085–2092.

187. Schrick C, Fischer A, Srivastava DP, Tronson NC, Penzes P, Radulovic J. N-cadherin regulates cytoskeletally associated IQGAP1/ERK signalling and memory formation. *Neuron*. 2007;55:786–798.
188. Sbroggiò, Mauro et al. "IQGAP1 Regulates ERK1/2 And AKT Signalling In The Heart And Sustains Functional Remodelling Upon Pressure Overload". *Cardiovascular Research*, vol 91, no. 3, 2011, pp. 456-464. Oxford University Press (OUP), doi:10.1093/cvr/cvr103.
189. Bollen IAE, van der Velden J. The contribution of mutations in MYH7 to the onset of cardiomyopathy. *Neth Heart J*. 2017 Dec;25(12):653-654. doi: 10.1007/s12471-017-1045-5. PMID: 29052809; PMCID: PMC5691819.
190. McNally, E., Golbus, J., & Puckelwartz, M. (2013). Genetic mutations and mechanisms in dilated cardiomyopathy. *Journal Of Clinical Investigation*, 123(1), 19-26. doi: 10.1172/jci62862
191. Revera M, Van der Merwe L, Heradien M, Goosen A, Corfield VA, Brink PA, Moolman-Smook JC. Long-term follow-up of R403WMYH7 and R92WTNNT2 HCM families: mutations determine left ventricular dimensions but not wall thickness during disease progression. *Cardiovasc J Afr*. 2007 May-Jun;18(3):146-53. PMID: 17612745; PMCID: PMC4213759.
192. McKenna WJ, Maron BJ, Thiene G. Classification, Epidemiology, and Global Burden of Cardiomyopathies. *Circ Res*. 2017 Sep 15;121(7):722-730. doi: 10.1161/CIRCRESAHA.117.309711. PMID: 28912179.
193. Jordan E, Hershberger RE. Considering complexity in the genetic evaluation of dilated cardiomyopathy. *Heart*. 2021; 107:106–112. doi: 10.1136/heartjnl-2020-316658
194. Norton, N. et al. Evaluating pathogenicity of rare variants from dilated cardiomyopathy in the exome era. *Circ. Cardiovasc. Genet*. 5, 167–174 (2012).
195. Knezevic T, Myers VD, Gordon J, Tilley DG, Sharp TE 3rd, Wang J, Khalili K, Cheung JY, Feldman AM. BAG3: a new player in the heart failure paradigm. *Heart Fail Rev*. 2015 Jul;20(4):423-34. doi: 10.1007/s10741-015-9487-6. PMID: 25925243; PMCID: PMC4463985.
196. Behl C. Breaking BAG: The Co-Chaperone BAG3 in Health and Disease. *Trends Pharmacol Sci*. 2016 Aug;37(8):672-688. doi: 10.1016/j.tips.2016.04.007. Epub 2016 May 6. PMID: 27162137.
197. Takayama S, Reed JC. Molecular chaperone targeting and regulation by BAG family proteins. *Nat Cell Biol*. 2001 Oct;3(10):E237-41. doi: 10.1038/ncb1001-e237. PMID: 11584289.
198. Coulson M, Robert S, Saint R. Drosophila starvin encodes a tissue-specific BAG-domain protein required for larval food uptake. *Genetics* 2005; 171: 1799–1812.
199. Doukhanina EV, Chen S, van der Zalm E, Godzik A, Reed J, Dickman MB. Identification and functional characterization of the BAG protein family in Arabidopsis thaliana. *J Biol Chem* 2006; 281: 18793–18801.
200. Colinet H, Hoffmann A. Gene and protein expression of Drosophila Starvin during cold stress and recovery from chill coma. *Insect Biochem Mol Biol* 2010; 40: 425–428.
201. Liu L, Sun K, Zhang X, Tang Y, Xu D. Advances in the role and mechanism of BAG3 in dilated cardiomyopathy. *Heart Fail Rev*. 2021 Jan;26(1):183-194. doi: 10.1007/s10741-019-09899-7. PMID: 31808029.
202. Roperto S, Russo V, Rosati A, et al. Chaperone-assisted selective autophagy in healthy and papillomavirus-associated neoplastic urothelium of cattle. *Veterinary Microbiology*. 2018 Jul;221:134-142. DOI: 10.1016/j.vetmic.2018.06.013. PMID: 29981700.
203. Klionsky DJ, Baehrecke EH, Brumell JH, Chu CT, Codogno P, Cuervo AM, Debnath J, Deretic V, Elazar Z, Eskelinen EL, Finkbeiner S, Fueyo-Margareto J, Gewirtz D, Jäättelä M, Kroemer G, Levine B, Melia TJ, Mizushima N, Rubinsztein DC, Simonsen A, Thorburn A, Thumm M, Tooze SA. A comprehensive glossary of autophagy-related molecules and processes (2nd edition). *Autophagy*. 2011 Nov;7(11):1273-94. doi: 10.4161/auto.7.11.17661. Epub 2011 Nov 1. PMID: 21997368; PMCID: PMC3359482.

204. Johnston JA, Illing ME, Kopito RR (2002) Cytoplasmic dynein/ dynactin mediates the assembly of aggresomes. *Cell Motil Cytoskeleton* 53(1):26-38. <https://doi.org/10.1002/cm.10057>
205. Zhang X, Qian SB (2011) Chaperone-mediated hierarchical control in targeting misfolded proteins to aggresomes. *Mol Biol Cell* 22(18):3277-3288. <https://doi.org/10.1091/mbc.E11-05-0388>
206. Munoz-Moreno R, Barrado-Gil L, Galindo I, Alonso C (2015) Analysis of HDAC6 and BAG3-aggresome pathways in African swine fever viral factory formation. *Viruses* 7(4):1823 - 1831. <https://doi.org/10.3390/v7041823>
207. Mizushima W, Sadoshima J (2017) BAG3 plays a central role in proteostasis in the heart. *J Clin Invest* 127(8):2900-2903. <https://doi.org/10.1172/jci95839>
208. Zhang J, He Z, Xiao W, Na Q, Wu T, Su K, Cui X (2016) Overexpression of BAG3 attenuates hypoxia-induced cardiomyocyte apoptosis by inducing autophagy. *Cell Physiol Biochem* 39(2):491-500. <https://doi.org/10.1159/000445641>
209. Homma, S. et al. (2006) BAG3 deficiency results in fulminant myopathy and early lethality. *Am. J. Pathol.* 169, 761-773
210. Pappas CT, Krieg PA, Gregorio CC. Nebulin regulates actin filament lengths by a stabilization mechanism. *J Cell Biol.* 2010 May 31;189(5):859-70. doi: 10.1083/jcb.201001043. Epub 2010 May 24. PMID: 20498015; PMCID: PMC2878950.
211. Pappas CT, Bhattacharya N, Cooper JA, Gregorio CC. Nebulin interacts with CapZ and regulates thin filament architecture within the Z-disc. *Mol Biol Cell.* 2008 May;19(5):1837-47. doi: 10.1091/mbc.e07-07-0690. Epub 2008 Feb 13. PMID: 18272787; PMCID: PMC2366866.
212. Schafer DA, Hug C, Cooper JA. Inhibition of CapZ during myofibrillogenesis alters assembly of actin filaments. *J Cell Biol.* 1995 Jan;128(1-2):61-70. doi: 10.1083/jcb.128.1.61. PMID: 7822423; PMCID: PMC2120327.
213. Hart MC, Cooper JA. Vertebrate isoforms of actin capping protein beta have distinct functions In vivo. *J Cell Biol.* 1999 Dec 13;147(6):1287-98. doi: 10.1083/jcb.147.6.1287. PMID: 10601341; PMCID: PMC2168092.
214. Ulbricht A, Eppler FJ, Tapia VE, van der Ven PF, Hampe N, Hersch N, Vakeel P, Stadel D, Haas A, Saftig P, Behrends C, Fürst DO, Volkmer R, Hoffmann B, Kolanus W, Höfeld J. Cellular mechanotransduction relies on tension-induced and chaperone-assisted autophagy. *Curr Biol.* 2013 Mar 4;23(5):430-5. doi: 10.1016/j.cub.2013.01.064. Epub 2013 Feb 21. PMID: 23434281.
215. Chami N, Tadros R, Lemarbre F, Lo KS, Beaudoin M, Robb L, Labuda D, Tardif JC, Racine N, Talajic M, Lettre G. Nonsense mutations in BAG3 are associated with early-onset dilated cardiomyopathy in French Canadians. *Can J Cardiol.* 2014 Dec;30(12):1655-61. doi: 10.1016/j.cjca.2014.09.030. Epub 2014 Oct 2. PMID: 25448463.
216. Omboni S, Volpe M. Angiotensin Receptor Blockers Versus Angiotensin Converting Enzyme Inhibitors for the Treatment of Arterial Hypertension and the Role of Olmesartan. *Adv Ther.* 2019 Feb;36(2):278-297. doi: 10.1007/s12325-018-0859-x. Epub 2018 Dec 27. PMID: 30591990; PMCID: PMC6824372.
217. Belkaya S, Kontorovich AR, Byun M, Mulero-Navarro S, Bajolle F, Cobat A, Josowitz R, Itan Y, Quint R, Lorenzo L, Boucherit S, Stoven C, Di Filippo S, Abel L, Zhang SY, Bonnet D, Gelb BD, Casanova JL. Autosomal Recessive Cardiomyopathy Presenting as Acute Myocarditis. *J Am Coll Cardiol.* 2017 Apr 4;69(13):1653-1665. doi: 10.1016/j.jacc.2017.01.043. PMID: 28359509; PMCID: PMC5551973.
218. Myers VD, McClung JM, Wang J, Tahrir FG, Gupta MK, Gordon J, Kontos CH, Khalili K, Cheung JY, Feldman AM. The Multifunctional Protein BAG3: A Novel Therapeutic Target in Cardiovascular Disease. *JACC Basic Transl Sci.* 2018 Feb;3(1):122-131. doi: 10.1016/j.jacbs.2017.09.009. PMID: 29938246; PMCID: PMC6013050

219. Gordon DM, Beckers P, Castermans E, Neggers SJMM, Rostomyan L, Bours V, Petrossians P, Dideberg V, Beckers A, Daly AF. Dutch founder SDHB exon 3 deletion in patients with pheochromocytoma-paraganglioma in South Africa. *Endocr Connect*. 2022 Jan 31;11(1):e210560. doi: 10.1530/EC-21-0560. PMID: 34939938; PMCID: PMC8859937.
220. Zeegers MPA, Poppel van F, Vlietinck R, Spruijt L, Ostrer H. Founder mutations among the Dutch. *European Journal of Human Genetics* 2004 12 591–600. (<https://doi.org/10.1038/sj.ejhg.5201151>)
221. Hoppe K, Jurkat-Rott K, Kranepuhl S, Wearing S, Heiderich S, Merlak S, Klingler W. Relevance of pathogenicity prediction tools in human RYR1 variants of unknown significance. *Sci Rep*. 2021 Feb 9;11(1):3445. doi: 10.1038/s41598-021-82024-7. PMID: 33564012; PMCID: PMC7873245.
222. Cassandrini D, Trovato R, Rubegni A, Lenzi S, Fiorillo C, Baldacci J, Minetti C, Astrea G, Bruno C, Santorelli FM; Italian Network on Congenital Myopathies. Congenital myopathies: clinical phenotypes and new diagnostic tools. *Ital J Pediatr*. 2017 Nov 15;43(1):101. doi: 10.1186/s13052-017-0419-z. PMID: 29141652; PMCID: PMC5688763.
223. Hong, Kyung-Won et al. "Common Variants In RYR1 Are Associated With Left Ventricular Hypertrophy Assessed By Electrocardiogram". *European Heart Journal*, vol 33, no. 10, 2011, pp. 1250-1256. *Oxford University Press (OUP)*, doi:10.1093/eurheartj/ehr267.
224. Vassort, Guy, and Alain Lacampagne. "The Role Of Ryanodine Receptors And Consequences Of Their Alterations During Cardiac Insufficiency". *Experimental & Clinical Cardiology*, vol 10, no. 3, 2005, p. 196., <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2716251/>.
225. Priori, S. G. "Cardiac And Skeletal Muscle Disorders Caused By Mutations In The Intracellular Ca<sup>2+</sup> Release Channels". *Journal Of Clinical Investigation*, vol 115, no. 8, 2005, pp. 2033-2038. *American Society For Clinical Investigation*, doi:10.1172/jci25664.
226. Park TJ, Boyd K, Curran T. Cardiovascular and craniofacial defects in Crk-null mice. *Mol Cell Biol*. 2006 Aug;26(16):6272-82. doi: 10.1128/MCB.00472-06. PMID: 16880535; PMCID: PMC1592792.
227. Wallgren-Pettersson C, Pelin K, Nowak KJ, Muntoni F, Romero NB, Goebel HH, North KN, Beggs AH, Laing NG; ENMC International Consortium On Nemaline Myopathy. Genotype-phenotype correlations in nemaline myopathy caused by mutations in the genes for nebulin and skeletal muscle alpha-actin. *Neuromuscul Disord*. 2004 Sep;14(8-9):461-70. doi: 10.1016/j.nmd.2004.03.006. PMID: 15336686.
228. de Winter JM, Ottenheijm CAC. Sarcomere Dysfunction in Nemaline Myopathy. *J Neuromuscul Dis*. 2017;4(2):99-113. doi: 10.3233/JND-160200. PMID: 28436394; PMCID: PMC5467716.
229. Sewry CA, Laitila JM, Wallgren-Pettersson C. Nemaline myopathies: a current view. *J Muscle Res Cell Motil*. 2019 Jun;40(2):111-126. doi: 10.1007/s10974-019-09519-9. Epub 2019 Jun 21. PMID: 31228046; PMCID: PMC6726674.
230. Hishiya A, Kitazawa T, Takayama S. BAG3 and Hsc70 interact with actin capping protein CapZ to maintain myofibrillar integrity under mechanical stress. *Circ Res*. 2010 Nov 12;107(10):1220-31. doi: 10.1161/CIRCRESAHA.110.225649. Epub 2010 Sep 30. PMID: 20884878; PMCID: PMC2980587.
231. Quarta G, Muir A, Pantazis A, Syrris P, Gehmlich K, Garcia-Pavia P, Ward D, Sen-Chowdhry S, Elliott PM, McKenna WJ. Familial evaluation in arrhythmogenic right ventricular cardiomyopathy: impact of genetics and revised task force criteria. *Circulation*. 2011 Jun 14;123(23):2701-9. doi: 10.1161/CIRCULATIONAHA.110.976936. Epub 2011 May 23. PMID: 21606390.
232. Christensen AH, Benn M, Bundgaard H, Tybjaerg-Hansen A, Haunso S, Svendsen JH. Wide spectrum of desmosomal mutations in Danish patients with arrhythmogenic right ventricular cardiomyopathy. *J Med Genet*. 2010 Nov;47(11):736-44. doi: 10.1136/jmg.2010.077891. Epub 2010 Sep 23. PMID: 20864495.

233. Fressart V, Duthoit G, Donal E, Probst V, Deharo JC, Chevalier P, Klug D, Dubourg O, Delacretaz E, Cosnay P, Scanu P, Extramiana F, Keller D, Hidden-Lucet F, Simon F, Bessirard V, Roux-Buisson N, Hebert JL, Azarine A, Casset-Senon D, Rouzet F, Lecarpentier Y, Fontaine G, Coirault C, Frank R, Hainque B, Charron P. Desmosomal gene analysis in arrhythmogenic right ventricular dysplasia/cardiomyopathy: spectrum of mutations and clinical impact in practice. *Europace*. 2010 Jun;12(6):861-8. doi: 10.1093/europace/euq104. Epub 2010 Apr 16. PMID: 20400443.
234. Cox MG, van der Zwaag PA, van der Werf C, van der Smagt JJ, Noorman M, Bhuiyan ZA, Wiesfeld AC, Volders PG, van Langen IM, Atsma DE, Dooijes D, van den Wijngaard A, Houweling AC, Jongbloed JD, Jordaens L, Cramer MJ, Doevendans PA, de Bakker JM, Wilde AA, van Tintelen JP, Hauer RN. Arrhythmogenic right ventricular dysplasia/cardiomyopathy: pathogenic desmosome mutations in index-patients predict outcome of family screening: Dutch arrhythmogenic right ventricular dysplasia/cardiomyopathy genotype-phenotype follow-up study. *Circulation*. 2011 Jun 14;123(23):2690-700. doi: 10.1161/CIRCULATIONAHA.110.988287. Epub 2011 May 23. PMID: 21606396.
235. Stevens, Tyler L., Michael J. Wallace, Mona El Refaey, Jason D. Roberts, Sara N. Koenig, and Peter J. Mohler. 2020. "Arrhythmogenic Cardiomyopathy: Molecular Insights For Improved Therapeutic Design". *Journal Of Cardiovascular Development And Disease* 7 (2): 21. doi:10.3390/jcdd7020021.
236. Vite A, Radice GL. N-cadherin/catenin complex as a master regulator of intercalated disc function. *Cell Commun Adhes*. 2014 Jun;21(3):169-79. doi: 10.3109/15419061.2014.908853. Epub 2014 Apr 28. Erratum in: *Cell Commun Adhes*. 2014 Dec;21(6):291. PMID: 24766605; PMCID: PMC6054126.
237. Vermij SH, Abriel H, van Veen TA. Refining the molecular organization of the cardiac intercalated disc. *Cardiovasc Res*. 2017 Mar 1;113(3):259-275. doi: 10.1093/cvr/cvw259. PMID: 28069669.
238. Kostera-Pruszczyk, A., Pruszczyk, P., Kamińska, A., Lee, H., & Goldfarb, L. (2008). Diversity of cardiomyopathy phenotypes caused by mutations in desmin. *International Journal Of Cardiology*, 131(1), 146-147. doi: 10.1016/j.ijcard.2007.08.095
239. Quarta G, Syrris P, Ashworth M, Jenkins S, Zuborne Alapi K, Morgan J, Muir A, Pantazis A, McKenna WJ, Elliott PM. Mutations in the Lamin A/C gene mimic arrhythmogenic right ventricular cardiomyopathy. *Eur Heart J*. 2012 May;33(9):1128-36. doi: 10.1093/eurheartj/ehr451. Epub 2011 Dec 23. PMID: 22199124.
240. van der Zwaag PA, van Rijsingen IA, Asimaki A, Jongbloed JD, van Veldhuisen DJ, Wiesfeld AC, Cox MG, van Lochem LT, de Boer RA, Hofstra RM, Christiaans I, van Spaendonck-Zwarts KY, Lekanne dit Deprez RH, Judge DP, Calkins H, Suurmeijer AJ, Hauer RN, Saffitz JE, Wilde AA, van den Berg MP, van Tintelen JP. Phospholamban R14del mutation in patients diagnosed with dilated cardiomyopathy or arrhythmogenic right ventricular cardiomyopathy: evidence supporting the concept of arrhythmogenic cardiomyopathy. *Eur J Heart Fail*. 2012 Nov;14(11):1199-207. doi: 10.1093/eurjhf/hfs119. Epub 2012 Jul 20. PMID: 22820313; PMCID: PMC3475434.
241. Beffagna G, Occhi G, Nava A, Vitiello L, Ditadi A, Basso C, Baucé B, Carraro G, Thiene G, Towbin JA, Danieli GA, Rampazzo A. Regulatory mutations in transforming growth factor-beta3 gene cause arrhythmogenic right ventricular cardiomyopathy type 1. *Cardiovasc Res*. 2005 Feb 1;65(2):366-73. doi: 10.1016/j.cardiores.2004.10.005. PMID: 15639475.
242. Groeneweg JA, Bhonsale A, James CA, te Riele AS, Dooijes D, Tichnell C, Murray B, Wiesfeld AC, Sawant AC, Kassamali B, Atsma DE, Volders PG, de Groot NM, de Boer K, Zimmerman SL, Kamel IR, van der Heijden JF, Russell SD, Jan Cramer M, Tedford RJ, Doevendans PA, van Veen TA, Tandri H, Wilde AA, Judge DP, van Tintelen JP, Hauer RN, Calkins H. Clinical Presentation, Long-Term Follow-Up, and Outcomes of 1001 Arrhythmogenic Right Ventricular Dysplasia/Cardiomyopathy Patients and Family Members. *Circ Cardiovasc Genet*. 2015 Jun;8(3):437-46. doi: 10.1161/CIRCGENETICS.114.001003. Epub 2015 Mar 27. PMID: 25820315.

243. Xu Z, Zhu W, Wang C, Huang L, Zhou Q, Hu J, Cheng X, Hong K. Genotype-phenotype relationship in patients with arrhythmogenic right ventricular cardiomyopathy caused by desmosomal gene mutations: A systematic review and meta-analysis. *Sci Rep.* 2017 Jan 25;7:41387. doi: 10.1038/srep41387. PMID: 28120905; PMCID: PMC5264593.
244. Hoorntje ET, Te Rijdt WP, James CA, Pilichou K, Basso C, Judge DP, Bezzina CR, van Tintelen JP. Arrhythmogenic cardiomyopathy: pathology, genetics, and concepts in pathogenesis. *Cardiovasc Res.* 2017 Oct 1;113(12):1521-1531. doi: 10.1093/cvr/cvx150. PMID: 28957532.
245. Ellinor PT, MacRae CA, Thierfelder L. Arrhythmogenic right ventricular cardiomyopathy. *Heart Fail Clin.* 2010 Apr;6(2):161-77. doi: 10.1016/j.hfc.2009.12.003. PMID: 20347785.
246. Azaouagh A, Churzidse S, Konorza T, Erbel R. Arrhythmogenic right ventricular cardiomyopathy/dysplasia: a review and update. *Clin Res Cardiol.* 2011 May;100(5):383-94. doi: 10.1007/s00392-011-0295-2. Epub 2011 Mar 1. PMID: 21360243.
247. Saffitz JE. Arrhythmogenic cardiomyopathy: advances in diagnosis and disease pathogenesis. *Circulation.* 2011 Oct 11;124(15):e390-2. doi: 10.1161/CIRCULATIONAHA.111.064022. PMID: 21986778.
248. Gerull B, Heuser A, Wichter T, Paul M, Basson C, & McDermott D. et al. (2004). Mutations in the desmosomal protein plakophilin-2 are common in arrhythmogenic right ventricular cardiomyopathy. *Nature Genetics*, 36(11), 1162-1164. doi: 10.1038/ng1461
249. Klein I, Danzi S. Thyroid Disease and the Heart. *Curr Probl Cardiol.* 2016 Feb;41(2):65-92. doi: 10.1016/j.cpcardiol.2015.04.002. Epub 2015 Apr 23. PMID: 26792255.
250. Dalal D, Molin L, Piccini J, Tichnell C, James C, & Bomma C. et al. (2006). Clinical Features of Arrhythmogenic Right Ventricular Dysplasia/Cardiomyopathy Associated With Mutations in Plakophilin-2. *Circulation*, 113(13), 1641-1649. doi: 10.1161/circulationaha.105.568642
251. Biernacka E, Borowiec K, Franaszczyk M, Szperl M, Rampazzo A, & Woźniak O. et al. (2021). Pathogenic variants in plakophilin-2 gene (PKP2) are associated with better survival in arrhythmogenic right ventricular cardiomyopathy. *Journal Of Applied Genetics*, 62(4), 613-620. doi: 10.1007/s13353-021-00647-y
252. Steffen Neuber, Mario Mühmer, Denise Wratten, Peter J. Koch, Roland Moll, Ansgar Schmidt, "The Desmosomal Plaque Proteins of the Plakophilin Family", *Dermatology Research and Practice*, vol. 2010, Article ID 101452, 11 pages, 2010. <https://doi.org/10.1155/2010/101452>
253. Alcalde, Mireia, Oscar Campuzano, Paola Berne, Pablo García-Pavía, Ada Doltra, Elena Arbelo, and Georgia Sarquella-Brugada et al. 2014. "Stop-Gain Mutations In PKP2 Are Associated With A Later Age Of Onset Of Arrhythmogenic Right Ventricular Cardiomyopathy". *Plos ONE* 9 (6): e100560. doi:10.1371/journal.pone.0100560.
254. Chen, Xinyu, Stefan Bonn  , Mechthild Hatzfeld, Frans van Roy, and Kathleen J. Green. 2002. "Protein Binding And Functional Characterization Of Plakophilin 2". *Journal Of Biological Chemistry* 277 (12): 10512-10522. doi:10.1074/jbc.m108765200.
255. Cerrone, Marina, Jerome Montnach, Xianming Lin, Yan-Ting Zhao, Mingliang Zhang, Esperanza Agullo-Pascual, and Alejandra Leo-Macias et al. 2017. "Plakophilin-2 Is Required For Transcription Of Genes That Control Calcium Cycling And Cardiac Rhythm". *Nature Communications* 8 (1). doi:10.1038/s41467-017-00127-0.
256. Kim JC, P  rez-Hern  ndez M, Alvarado FJ, Maurya SR, Montnach J, Yin Y, Zhang M, Lin X, Vasquez C, Heguy A, Liang FX, Woo SH, Morley GE, Rothenberg E, Lundby A, Valdivia HH, Cerrone M, Delmar M. Disruption of Ca<sup>2+</sup> Homeostasis and Connexin 43 Hemichannel Function in the Right Ventricle Precedes Overt Arrhythmogenic Cardiomyopathy in Plakophilin-2-Deficient Mice. *Circulation.* 2019 Sep 17;140(12):1015-1030. doi: 10.1161/CIRCULATIONAHA.119.039710. Epub 2019 Jul 18. PMID: 31315456; PMCID: PMC6746608.

257. van der Zwaag PA, Jongbloed JD, van den Berg MP, van der Smagt JJ, Jongbloed R, Bikker H, Hofstra RM, van Tintelen JP. A genetic variants database for arrhythmogenic right ventricular dysplasia/cardiomyopathy. *Hum Mutat.* 2009 Sep;30(9):1278-83. doi: 10.1002/humu.21064. PMID: 19569224.
258. Herren T, Gerber PA, Duru F. Arrhythmogenic right ventricular cardiomyopathy/dysplasia: a not so rare "disease of the desmosome" with multiple clinical presentations. *Clin Res Cardiol.* 2009 Mar;98(3):141-58. doi: 10.1007/s00392-009-0751-4. Epub 2009 Feb 9. PMID: 19205777.
259. Gandjbakhch E, Redheuil A, Pousset F, Charron P, Frank R. Clinical Diagnosis, Imaging, and Genetics of Arrhythmogenic Right Ventricular Cardiomyopathy/Dysplasia: JACC State-of-the-Art Review. *J Am Coll Cardiol.* 2018 Aug 14;72(7):784-804. doi: 10.1016/j.jacc.2018.05.065. PMID: 30092956.
260. Bhonsale A, Groeneweg JA, James CA, Dooijes D, Tichnell C, Jongbloed JD, Murray B, te Riele AS, van den Berg MP, Bikker H, Atsma DE, de Groot NM, Houweling AC, van der Heijden JF, Russell SD, Doevendans PA, van Veen TA, Tandri H, Wilde AA, Judge DP, van Tintelen JP, Calkins H, Hauer RN. Impact of genotype on clinical course in arrhythmogenic right ventricular dysplasia/cardiomyopathy-associated mutation carriers. *Eur Heart J.* 2015 Apr 7;36(14):847-55. doi: 10.1093/eurheartj/ehu509. Epub 2015 Jan 23. PMID: 25616645.
261. Delmar M, McKenna WJ. The cardiac desmosome and arrhythmogenic cardiomyopathies: from gene to disease. *Circ Res.* 2010 Sep 17;107(6):700-14. doi: 10.1161/CIRCRESAHA.110.223412. PMID: 20847325.
262. Chen SN, Gurha P, Lombardi R, Ruggiero A, Willerson JT, Marian AJ. The hippo pathway is activated and is a causal mechanism for adipogenesis in arrhythmogenic cardiomyopathy. *Circ Res.* 2014 Jan 31;114(3):454-68. doi: 10.1161/CIRCRESAHA.114.302810. Epub 2013 Nov 25. PMID: 24276085; PMCID: PMC3946717.
263. Lombardi R, da Graca Cabreira-Hansen M, Bell A, Fromm RR, Willerson JT, Marian AJ. Nuclear plakoglobin is essential for differentiation of cardiac progenitor cells to adipocytes in arrhythmogenic right ventricular cardiomyopathy. *Circ Res.* 2011 Dec 9;109(12):1342-53. doi: 10.1161/CIRCRESAHA.111.255075. Epub 2011 Oct 20. PMID: 22021931; PMCID: PMC3237769.
264. Dubash AD, Kam CY, Aguado BA, Patel DM, Delmar M, Shea LD, Green KJ. Plakophilin-2 loss promotes TGF- $\beta$ 1/p38 MAPK-dependent fibrotic gene expression in cardiomyocytes. *J Cell Biol.* 2016 Feb 15;212(4):425-38. doi: 10.1083/jcb.201507018. Epub 2016 Feb 8. PMID: 26858265; PMCID: PMC4754716.
265. Kim C, Wong J, Wen J, Wang S, Wang C, Spiering S, Kan NG, Forcales S, Puri PL, Leone TC, Marine JE, Calkins H, Kelly DP, Judge DP, Chen HS. Studying arrhythmogenic right ventricular dysplasia with patient-specific iPSCs. *Nature.* 2013 Feb 7;494(7435):105-10. doi: 10.1038/nature11799. Epub 2013 Jan 27. PMID: 23354045; PMCID: PMC3753229.
266. Choi, Hee-Jung, and William I. Weis. 2005. "Structure Of The Armadillo Repeat Domain Of Plakophilin 1". *Journal Of Molecular Biology* 346 (1): 367-376. doi:10.1016/j.jmb.2004.11.048.
267. Almomani, R., Verhagen, J., Herkert, J., Brosens, E., van Spaendonck-Zwarts, K., & Asimaki, A. et al. (2016). Biallelic Truncating Mutations in ALPK3 Cause Severe Pediatric Cardiomyopathy. *Journal Of The American College Of Cardiology*, 67(5), 515-525. doi: 10.1016/j.jacc.2015.10.093
268. Hosoda, T., Monzen, K., Hiroi, Y., Oka, T., Takimoto, E., & Yazaki, Y. et al. (2001). A Novel Myocyte-specific Gene Midori Promotes the Differentiation of P19CL6 Cells into Cardiomyocytes. *Journal Of Biological Chemistry*, 276(38), 35978-35989. doi: 10.1074/jbc.m100485200
269. Van Sligtenhorst I, Ding ZM, Shi ZZ, Read RW, Hansen G, Vogel P. Cardiomyopathy in  $\alpha$ -kinase 3 (ALPK3)-deficient mice. *Vet Pathol.* 2012 Jan;49(1):131-41. doi: 10.1177/0300985811402841. Epub 2011 Mar 25. PMID: 21441111.

270. Çağlayan, A., Sezer, R., Kaymakçalan, H., Ulgen, E., Yavuz, T., & Baranoski, J. et al. (2017). ALPK3 gene mutation in a patient with congenital cardiomyopathy and dysmorphic features. *Molecular Case Studies*, 3(5), a001859. doi: 10.1101/mcs.a001859
271. Herkert, J., Verhagen, J., Yotti, R., Haghighi, A., Phelan, D., & James, P. et al. (2020). Expanding the clinical and genetic spectrum of ALPK3 variants: Phenotypes identified in pediatric cardiomyopathy patients and adults with heterozygous variants. *American Heart Journal*, 225, 108-119. doi: 10.1016/j.ahj.2020.03.023
272. Lopes, L., Garcia-Hernández, S., Lorenzini, M., Futema, M., Chumakova, O., & Zateyshchikov, D. et al. (2021). Alpha-protein kinase 3 truncating variants are a cause of autosomal dominant hypertrophic cardiomyopathy. *European Heart Journal*, 42(32), 3063-3073. doi: 10.1093/eurheartj/ehab424
273. Carlo, S., Rodríguez-Fernández, L., Benítez Ríos, F., Arciniegas-Medina, N., & Martínez-González, H. (2022). Genetic Evaluation of Late-Onset Hypertrophic Cardiomyopathy: An Autobiographical Case Report. *Cureus*. doi: 10.7759/cureus.23349
274. Jorholt, J., Formicheva, Y., Vershinina, T., Kiselev, A., Muravyev, A., & Demchenko, E. et al. (2020). Two New Cases of Hypertrophic Cardiomyopathy and Skeletal Muscle Features Associated with ALPK3 Homozygous and Compound Heterozygous Variants. *Genes*, 11(10), 1201. doi: 10.3390/genes11101201
275. Al Senaidi, K., Joshi, N., Al-Nabhani, M., Al-Kasbi, G., Al Farqani, A., Al-Thihli, K., & Al-Maawali, A. (2019). Phenotypic spectrum of ALPK3-related cardiomyopathy. *American Journal Of Medical Genetics Part A*. doi: 10.1002/ajmg.a.61176
276. Middelbeek J, Clark K, Venselaar H, Huynen MA, van Leeuwen FN. The alpha-kinase family: an exceptional branch on the protein kinase tree. *Cell Mol Life Sci*. 2010 Mar;67(6):875-90. doi: 10.1007/s00018-009-0215-z. Epub 2009 Dec 12. PMID: 20012461; PMCID: PMC2827801.
277. Drennan D, Ryazanov AG. Alpha-kinases: analysis of the family and comparison with conventional protein kinases. *Prog Biophys Mol Biol*. 2004 May;85(1):1-32. doi: 10.1016/S0079-6107(03)00060-9. PMID: 15050379.
278. Maione, A., Pilato, C., Casella, M., Gasperetti, A., Stadiotti, I., Pompilio, G., & Sommariva, E. (2020). Fibrosis in Arrhythmogenic Cardiomyopathy: The Phantom Thread in the Fibro-Adipose Tissue. *Frontiers In Physiology*, 11. doi: 10.3389/fphys.2020.00279
279. Ma, X., Shao, Y., Tian, L., Flasch, D., Mulder, H., & Edmonson, M. et al. (2019). Analysis of error profiles in deep next-generation sequencing data. *Genome Biology*, 20(1). doi: 10.1186/s13059-019-1659-6
280. Mahdih N, Rabbani B. An overview of mutation detection methods in genetic disorders. *Iran J Pediatr*. 2013 Aug;23(4):375-88. PMID: 24427490; PMCID: PMC3883366.
281. Mocumbi AO. Lack of focus on cardiovascular disease in sub-Saharan Africa. *Cardiovasc Diagn Ther*. 2012 Mar;2(1):74-7. doi: 10.3978/j.issn.2223-3652.2012.01.03. PMID: 24282699; PMCID: PMC3839174.
282. Cao G, Xuan X, Zhang R, Hu J, Dong H. Gene Therapy for Cardiovascular Disease: Basic Research and Clinical Prospects. *Front Cardiovasc Med*. 2021 Nov 5;8:760140. doi: 10.3389/fcvm.2021.760140. PMID: 34805315; PMCID: PMC8602679.
283. Ferreira, Vanessa & Piechnik, Stefan & Dall'Armellina, Erica & Karamitsos, Theodoros & Francis, Jane & Ntusi, Ntobeko & Holloway, Cameron & Choudhury, Robin & Kardos, Attila & Robson, Matthew & Friedrich, Matthias & Neubauer, Stefan. (2014). Native T1-mapping detects the location, extent and patterns of acute myocarditis without the need for gadolinium contrast agents. *Journal of Cardiovascular Magnetic Resonance*. 16. 10.1186/1532-429X-16-36.
284. Ali, Z., Karimi Galougahi, K., Maehara, A., Shlofmitz, R., Ben-Yehuda, O., Mintz, G., & Stone, G. (2017). Intracoronary Optical Coherence Tomography 2018. *JACC: Cardiovascular Interventions*, 10(24), 2473-2487. doi: 10.1016/j.jcin.2017.09.042

285. Intravascular Ultrasound (2022). *Circulation*. Retrieved from <https://www.ahajournals.org/doi/10.1161/01.CIR.103.4.604>
286. Mehrotra, S., Mishra, S., & Paramasivam, G. (2018). Imaging during percutaneous coronary intervention for optimizing outcomes. *Indian Heart Journal*, 70, S456-S465. doi: 10.1016/j.ihj.2018.08.012
287. Dilsivian, V., Gewirtz, H., Marwick, T., Kwong, R., Raggi, P., Al-Mallah, M., & Herzog, C. (2021). Cardiac Imaging for Coronary Heart Disease Risk Stratification in Chronic Kidney Disease. *JACC: Cardiovascular Imaging*, 14(3), 669-682. doi: 10.1016/j.jcmg.2020.05.035
288. Polygenic Risk Scores for Cardiovascular Disease: A Scientific Statement From the American Heart Association. (2022). *Circulation*. Retrieved from <https://www.ahajournals.org/doi/10.1161/CIR.0000000000001077#d1e573>
289. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, Natarajan P, Lander ES, Lubitz SA, Ellinor PT, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet*. 2018; 50:1219–1224. doi: 10.1038/s41588-018-0183-z
290. Pirruccello, J.P., Bick, A., Wang, M. *et al*. Analysis of cardiac magnetic resonance imaging in 36,000 individuals yields genetic insights into dilated cardiomyopathy. *Nat Commun* **11**, 2254 (2020). <https://doi.org/10.1038/s41467-020-15823-7>
291. Tadros R, Francis C, Xu X, Vermeer AMC, Harper AR, Huurman R, Kelu Bisabu K, Walsh R, Hoorntje ET, Te Rijdt WP, Buchan RJ, van Velzen HG, van Slegtenhorst MA, Vermeulen JM, Offerhaus JA, Bai W, de Marvao A, Lahrouchi N, Beekman L, Karper JC, Veldink JH, Kayvanpour E, Pantazis A, Baksi AJ, Whiffin N, Mazzarotto F, Sloane G, Suzuki H, Schneider-Luftman D, Elliott P, Richard P, Ader F, Villard E, Lichtner P, Meitinger T, Tanck MWT, van Tintelen JP, Thain A, McCarty D, Hegele RA, Roberts JD, Amyot J, Dubé MP, Cadrin-Tourigny J, Giraldeau G, L'Allier PL, Garceau P, Tardif JC, Boekholdt SM, Lumbers RT, Asselbergs FW, Barton PJR, Cook SA, Prasad SK, O'Regan DP, van der Velden J, Verweij KJH, Talajic M, Lettre G, Pinto YM, Meder B, Charron P, de Boer RA, Christiaans I, Michels M, Wilde AAM, Watkins H, Matthews PM, Ware JS, Bezzina CR. Shared genetic pathways contribute to risk of hypertrophic and dilated cardiomyopathies with opposite directions of effect. *Nat Genet*. 2021 Feb;53(2):128-134. doi: 10.1038/s41588-020-00762-2. Epub 2021 Jan 25. PMID: 33495596; PMCID: PMC7611259.
292. Villard E, Perret C, Gary F, Proust C, Dilanian G, Hengstenberg C, Ruppert V, Arbustini E, Wichter T, Germain M, Dubourg O, Tavazzi L, Aumont MC, DeGroot P, Fauchier L, Trochu JN, Gibelin P, Aupetit JF, Stark K, Erdmann J, Hetzer R, Roberts AM, Barton PJ, Regitz-Zagrosek V; Cardiogenics Consortium, Aslam U, Duboscq-Bidot L, Meyborg M, Maisch B, Madeira H, Waldenström A, Galve E, Cleland JG, Dorent R, Roizes G, Zeller T, Blankenberg S, Goodall AH, Cook S, Tregouet DA, Tiret L, Isnard R, Komajda M, Charron P, Cambien F. A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. *Eur Heart J*. 2011 May;32(9):1065-76. doi: 10.1093/eurheartj/ehr105. Epub 2011 Apr 1. PMID: 21459883; PMCID: PMC3086901.
293. Fatumo, S., Chikowore, T., Choudhury, A. *et al*. A roadmap to increase diversity in genomic studies. *Nat Med* **28**, 243–250 (2022). <https://doi.org/10.1038/s41591-021-01672-4>
294. Sen-Chowdhry S, Syrris P, Pantazis A, Quarta G, McKenna WJ, Chambers JC. Mutational heterogeneity, modifier genes, and environmental influences contribute to phenotypic diversity of arrhythmogenic cardiomyopathy. *Circ Cardiovasc Genet*. 2010 Aug;3(4):323-30. doi: 10.1161/CIRCGENETICS.109.935262. Epub 2010 Jun 22. PMID: 20570917.
295. Kopanos, C., Tsiolkas, V., Kouris, A., Chapple, C., Albarca Aguilera, M., Meyer, R., & Massouras, A. (2018). VarSome: the human genomic variant search engine. *Bioinformatics*, 35(11), 1978-1980. doi: 10.1093/bioinformatics/bty897

296. Schwarz, J., Rödelberger, C., Schuelke, M., & Seelow, D. (2010). MutationTaster evaluates disease-causing potential of sequence alterations. *Nature Methods*, 7(8), 575-576. doi: 10.1038/nmeth0810-575
297. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 2003 Jul 1;31(13):3812-4. doi: 10.1093/nar/gkg509. PMID: 12824425; PMCID: PMC168916.
298. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet.* 2013 Jan;Chapter 7:Unit7.20. doi: 10.1002/0471142905.hg0720s76. PMID: 23315928; PMCID: PMC4480630.
299. Jagadeesh KA, Wenger AM, Berger MJ, Guturu H, Stenson PD, Cooper DN, Bernstein JA, Bejerano G. M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. *Nat Genet.* 2016 Dec;48(12):1581-1586. doi: 10.1038/ng.3703. Epub 2016 Oct 24. PMID: 27776117.
300. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 2001 Jan 1;29(1):308-11. doi: 10.1093/nar/29.1.308. PMID: 11125122; PMCID: PMC29783.
301. Clarke, L., Zheng-Bradley, X., Smith, R., Kulesha, E., Xiao, C., & Toneva, I. et al. (2012). The 1000 Genomes Project: data management and community access. *Nature Methods*, 9(5), 459-462. doi: 10.1038/nmeth.1974
302. Karczewski KJ, Weisburd B, Thomas B, Solomonson M, Ruderfer DM, Kavanagh D, Hamamsy T, Lek M, Samocha KE, Cummings BB, Birnbaum D; The Exome Aggregation Consortium, Daly MJ, MacArthur DG. The ExAC browser: displaying reference data information from over 60 000 exomes. *Nucleic Acids Res.* 2017 Jan 4;45(D1):D840-D845. doi: 10.1093/nar/gkw971. Epub 2016 Nov 28. PMID: 27899611; PMCID: PMC5210650.
303. Birney E, Andrews TD, Bevan P, Caccamo M, Chen Y, Clarke L, Coates G, Cuff J, Curwen V, Cutts T, Down T, Eyraas E, Fernandez-Suarez XM, Gane P, Gibbins B, Gilbert J, Hammond M, Hotz HR, Iyer V, Jekosch K, Kahari A, Kasprzyk A, Keefe D, Keenan S, Lehvaslaiho H, McVicker G, Melsopp C, Meidl P, Mongin E, Pettett R, Potter S, Proctor G, Rae M, Searle S, Slater G, Smedley D, Smith J, Spooner W, Stabenau A, Stalker J, Storey R, Ureta-Vidal A, Woodwark KC, Cameron G, Durbin R, Cox A, Hubbard T, Clamp M. An overview of Ensembl. *Genome Res.* 2004 May;14(5):925-8. doi: 10.1101/gr.1860604. Epub 2004 Apr 12. PMID: 15078858; PMCID: PMC479121.
304. Kang P, Kalloniatis M, Doig GS. Using Updated PubMed: New Features and Functions to Enhance Literature Searches. *JAMA.* 2021;326(6):479–480. doi:10.1001/jama.2021.12021
305. Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., & Shad, S. et al. (2013). The Genotype-Tissue Expression (GTEx) project. *Nature Genetics*, 45(6), 580-585. doi: 10.1038/ng.2653
306. Eppig JT. Mouse Genome Informatics (MGI) Resource: Genetic, Genomic, and Biological Knowledgebase for the Laboratory Mouse. *ILAR J.* 2017 Jul 1;58(1):17-41. doi: 10.1093/ilar/ilx013. PMID: 28838066; PMCID: PMC5886341.
307. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry.* 2004 Oct;25(13):1605-1612. DOI: 10.1002/jcc.20084. PMID: 15264254.
308. Dagona, A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series.*
309. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res.* 2009 Jul;37(Web Server issue):W305-11. doi: 10.1093/nar/gkp427. Epub 2009 May 22. PMID: 19465376; PMCID: PMC2703978.
310. Stelzer, G., Plaschkes, I., Oz-Levi, D., Alkelai, A., Olender, T., & Zimmerman, S. et al. (2016). VarElect: the phenotype-based variation prioritizer of the GeneCards Suite. *BMC Genomics*, 17(S2). doi: 10.1186/s12864-016-2722-2

311. Tranchevent LC, Ardeshirdavani A, ElShal S, Alcaide D, Aerts J, Auboeuf D, Moreau Y. Candidate gene prioritization with Endeavour. *Nucleic Acids Res.* 2016 Jul 8;44(W1):W117-21. doi: 10.1093/nar/gkw365. Epub 2016 Apr 30. PMID: 27131783; PMCID: PMC4987917.
312. Owczarzy R, Tataurov AV, Wu Y, Manthey JA, McQuisten KA, Almabrazi HG, Pedersen KF, Lin Y, Garretson J, McEntaggart NO, Sailor CA, Dawson RB, Peek AS. IDT SciTools: a suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Res.* 2008 Jul 1;36(Web Server issue):W163-9. doi: 10.1093/nar/gkn198. Epub 2008 Apr 25. PMID: 18440976; PMCID: PMC2447751.
313. Robinson JT, Thorvaldsdóttir H, Wenger AM, Zehir A, Mesirov JP. Variant Review with the Integrative Genomics Viewer. *Cancer Res.* 2017 Nov 1;77(21):e31-e34. doi: 10.1158/0008-5472.CAN-17-0337. PMID: 29092934; PMCID: PMC5678989.
314. Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., & Madden, T. (2008). NCBI BLAST: a better web interface. *Nucleic Acids Research*, 36(Web Server), W5-W9. doi: 10.1093/nar/gkn201
315. Pinamonti, B., Abate, E., De Luca, A., Finocchiaro, G., & Korcova, R. (2019). Role of Cardiac Imaging: Echocardiography. Dilated Cardiomyopathy, 83-111. doi: 10.1007/978-3-030-13864-6\_7
316. Mandeş, L., Roşca, M., Ciupercă, D., & Popescu, B. (2020). The role of echocardiography for diagnosis and prognostic stratification in hypertrophic cardiomyopathy. *Journal Of Echocardiography*, 18(3), 137-148. doi: 10.1007/s12574-020-00467-9
317. Abbate R, Sticchi E, Fatini C. Genetics of cardiovascular disease. *Clin Cases Miner Bone Metab.* 2008 Jan;5(1):63-6. PMID: 22460848; PMCID: PMC2781194.
318. Candy GP, Skudicky D, Mueller UK, et al. Association of left ventricular systolic performance and cavity size with angiotensin-converting enzyme genotype in idiopathic dilated cardiomyopathy. *Am J Cardiol* 1999;83:740-4. 10.1016/S0002-9149(98)00981-3
319. Tiago AD, Badenhorst D, Skudicky D, et al. An aldosterone synthase gene variant is associated with improvement in left ventricular ejection fraction in dilated cardiomyopathy. *Cardiovasc Res* 2002;54:584-9. 10.1016/S0008-6363(02)00281-X
320. Badenhorst D, Norton GR, Sliwa K, et al. Impact of beta2-adrenoreceptor gene variants on cardiac cavity size and systolic function in idiopathic dilated cardiomyopathy. *Pharmacogenomics J* 2007;7:339-45. 10.1038/sj.tpj.6500426
321. Du Preez J, Matolweni LO, Greenberg J, et al. The alpha 2C Del322-325 adrenergic receptor polymorphism is not associated with heart failure due to idiopathic dilated cardiomyopathy in black Africans. *Cardiovasc J Afr* 2008;19:15-6.
322. Woodiwiss AJ, Badenhorst D, Sliwa K, et al. Beta1- and alpha2c-adrenoreceptor variants as predictors of clinical aspects of dilated cardiomyopathy in people of African ancestry. *Cardiovasc J Afr* 2008;19:188-93.
323. Mahjoub S, Mehri S, Bousaada R, et al. Association of ACE I/D polymorphism in Tunisian patients with dilated cardiomyopathy. *J Renin Angiotensin Aldosterone Syst* 2010;11:187-91. 10.1177/1470320310368874
324. Sayed S, Idriss NK, Blann A, et al. The Number of GT(n) Repeats in the Hemeoxygenase-1 Gene Promoter is Increased in Pediatric Heart Failure but is Unrelated to Renal, Antioxidant and Anti-inflammatory Markers. *Pediatr Cardiol* 2015;36:1204-11. 10.1007/s00246-015-1146-0
325. Adadi N, Radi FZ, Lahrouchi N, et al. Inherited dilated cardiomyopathy in a large Moroccan family caused by LMNA mutation. *Anatol J Cardiol* 2018;20:65-8
326. Heradien M, Revera M, van der Merwe L, et al. Abnormal blood pressure response to exercise occurs more frequently in hypertrophic cardiomyopathy patients with the R92W troponin T mutation than in those with myosin mutations. *Heart Rhythm* 2009;6:S18-24. 10.1016/j.hrthm.2009.07.020

327. Moolman JC, Brink PA, Corfield VA. Identification of a new missense mutation at Arg403, a CpG mutation hotspot, in exon 13 of the beta-myosin heavy chain gene in hypertrophic cardiomyopathy. *Hum Mol Genet* 1993;2:1731-2. 10.1093/hmg/2.10.1731
328. 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;6:197-8. 10.1002/humu.1380060219
329. Posen BM, Moolman JC, Corfield VA, et al. Clinical and prognostic evaluation of familial hypertrophic cardiomyopathy in two South African families with different cardiac beta myosin heavy chain gene mutations. *Br Heart J* 1995;74:40-6. 10.1136/hrt.74.1.40
330. Tesson F, Dufour C, Moolman JC, et al. The influence of the angiotensin I converting enzyme genotype in familial hypertrophic cardiomyopathy varies with the disease gene mutation. *J Mol Cell Cardiol* 1997;29:831-8. 10.1006/jmcc.1996.0332
331. Moolman JC, Corfield VA, Posen B, et al. Sudden death due to troponin T mutations. *J Am Coll Cardiol* 1997;29:549-55. 10.1016/S0735-1097(96)00530-X
332. Moolman-Smook JC, Mayosi B, Brink P, et al. Identification of a new missense mutation in MyBP-C associated with hypertrophic cardiomyopathy. *J Med Genet* 1998;35:253-4. 10.1136/jmg.35.3.253
333. Moolman-Smook JC, De Lange WJ, Bruwer EC, et al. 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;65:1308-20. 10.1086/302623
334. Andersen PS, Havndrup O, Bundgaard H, et al. Myosin light chain mutations in familial hypertrophic cardiomyopathy: phenotypic presentation and frequency in Danish and South African populations. *J Med Genet* 2001;38:E43. 10.1136/jmg.38.12.e43
335. Blair E, Redwood C, de Jesus Oliveira M, et al. Mutations of the light meromyosin domain of the beta-myosin heavy chain rod in hypertrophic cardiomyopathy. *Circ Res* 2002;90:263-9. 10.1161/hh0302.10453
336. Moolman-Smook J, De Lange W, Corfield V, et al. Expression of HCM causing mutations: lessons learnt from genotype-phenotype studies of the South African founder MYH7 A797T mutation. *J Med Genet* 2000;37:951-6. 10.1136/jmg.37.12.951
337. van der Merwe L, Cloete R, Revera M, et al. Genetic variation in angiotensin-converting enzyme 2 gene is associated with extent of left ventricular hypertrophy in hypertrophic cardiomyopathy. *Hum Genet* 2008;124:57-61. 10.1007/s00439-008-0524-6
338. Carstens N, van der Merwe L, Revera M, et al. Genetic variation in angiotensin II type 2 receptor gene influences extent of left ventricular hypertrophy in hypertrophic cardiomyopathy independent of blood pressure. *J Renin Angiotensin Aldosterone Syst* 2011;12:274-80. 10.1177/1470320310390725
339. Kassem H, Azer RS, Saber-Ayad M, et al. Early results of sarcomeric gene screening from the Egyptian National BA-HCM Program. *J Cardiovasc Transl Res* 2013;6:65-80. 10.1007/s12265-012-9425-0
340. Jaafar N, Gomez J, Kammoun I, et al. Spectrum of Mutations in Hypertrophic Cardiomyopathy Genes Among Tunisian Patients. *Genet Test Mol Biomarkers* 2016;20:674-9. 10.1089/gtmb.2016.018
341. Mouton JM, van der Merwe L, Goosen A, et al. MYBPH acts as modifier of cardiac hypertrophy in hypertrophic cardiomyopathy (HCM) patients. *Hum Genet* 2016;135:477-83. 10.1007/s00439-016-1649-7
342. Lahrouchi N, Behr ER, Bezzina CR. Next-Generation Sequencing in Post-mortem Genetic Testing of Young Sudden Cardiac Death Cases. *Front Cardiovasc Med* 2016;3:13. 10.3389/fcvm.2016.00013
343. Matolweni LO, Bardien S, Rebello G, 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. 10.1186/1471-2350-7-29

344. Choung HYG, Vyas M, Jacoby D, et al. Arrhythmogenic Right Ventricular Cardiomyopathy (ARVC) in a young female athlete at 36 weeks gestation: a case report. *Pathol Res Pract* 2017;213:1302-5. 10.1016/j.prp.2017.07.015
345. Mouton JM, Pellizzon AS, Goosen A, 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;26:63-9. 10.5830/CVJA-2015-019
346. Tayal, U., Prasad, S. & Cook, S.A. Genetics and genomics of dilated cardiomyopathy and systolic heart failure. *Genome Med* 9, 20 (2017). <https://doi.org/10.1186/s13073-017-0410-8>
347. Rosenbaum AN, Agre KE, Pereira NL. Genetics of dilated cardiomyopathy: practical implications for heart failure management. *Nat Rev Cardiol.* 2020 May;17(5):286-297. doi: 10.1038/s41569-019-0284-0. Epub 2019 Oct 11. PMID: 31605094.
348. Jordan E, Peterson L, Ai T, Asatryan B, Bronicki L, Brown E, Celeghin R, Edwards M, Fan J, Ingles J, James CA, Jarinova O, Johnson R, Judge DP, Lahrouchi N, Lekanne Deprez RH, Lumbers RT, Mazzarotto F, Medeiros Domingo A, Miller RL, Morales A, Murray B, Peters S, Pilichou K, Protonotarios A, Semsarian C, Shah P, Syrris P, Thaxton C, van Tintelen JP, Walsh R, Wang J, Ware J, Hersherberger RE. Evidence-Based Assessment of Genes in Dilated Cardiomyopathy. *Circulation.* 2021 Jul 6;144(1):7-19. doi: 10.1161/CIRCULATIONAHA.120.053033. Epub 2021 May 5. PMID: 33947203; PMCID: PMC8247549.
349. Ross, S.B., Singer, E.S., Driscoll, E. et al. Genetic architecture of left ventricular noncompaction in adults. *Hum Genome Var* 7, 33 (2020). <https://doi.org/10.1038/s41439-020-00120-y>
350. Vallverdú-Prats, M., Alcalde, M., Sarquella-Brugada, G., Cesar, S., Arbelo, E., & Brugada, J. et al. (2021). Update on Genes Associated with Arrhythmogenic Cardiomyopathy. *Cardiomyopathy - Disease Of The Heart Muscle*. doi: 10.5772/intechopen.95332
351. McNally, E., MacLeod, H., & Dellefave-Castillo, L. (2017). *Arrhythmogenic Right Ventricular Cardiomyopathy*. University Of Washington, Seattle. Retrieved from <https://www.ncbi.nlm.nih.gov/books/NBK1131/>
352. Gacita AM, McNally EM. Genetic Spectrum of Arrhythmogenic Cardiomyopathy. *Circ Heart Fail.* 2019 Mar;12(3):e005850. doi: 10.1161/CIRCHEARTFAILURE.119.005850. PMID: 30871346; PMCID: PMC6422169.
353. Lippi M, Chiesa M, Ascione C, Pedrazzini M, Mushtaq S, Rovina D, Riggio D, Di Blasio AM, Biondi ML, Pompilio G, Colombo GI, Casella M, Novelli V, Sommariva E. Spectrum of Rare and Common Genetic Variants in Arrhythmogenic Cardiomyopathy Patients. *Biomolecules.* 2022 Jul 28;12(8):1043. doi: 10.3390/biom12081043. PMID: 36008935; PMCID: PMC9405889.
354. Darwish RK, Haghighi A, Seliem ZS, El-Saiedi SA, Radwan NH, El-Gayar DF, Elfeel NS, Abouelhoda M, Mehaney DA. Genetic study of pediatric hypertrophic cardiomyopathy in Egypt. *Cardiol Young.* 2020 Dec;30(12):1910-1916. doi: 10.1017/S1047951120003157. Epub 2020 Oct 5. PMID: 33012304.
355. Mehaney DA, Haghighi A, Embaby AK, Zeyada RA, Darwish RK, Elfeel NS, Abouelhoda M, El-Saiedi SA, Gohar NA, Seliem ZS. Molecular analysis of dilated and left ventricular noncompaction cardiomyopathies in Egyptian children. *Cardiol Young.* 2022 Feb;32(2):295-300. doi: 10.1017/S1047951121002055. Epub 2021 May 26. PMID: 34036930.
356. Jaouadi H, Kraoua L, Chaker L, et al. Novel ALPK3 mutation in a Tunisian patient with pediatric cardiomyopathy and facio-thoraco-skeletal features. *Journal of Human Genetics.* 2018 Oct;63(10):1077-1082. DOI: 10.1038/s10038-018-0492-1. PMID: 30046096.

## APPENDICES

### Appendix A: List of genes implicated in dilated cardiomyopathy<sup>346-348</sup>

|        |       |        |        |         |        |
|--------|-------|--------|--------|---------|--------|
| MYH7   | DMD   | SCN5A  | RAF1   | DSG2    | PSEN2  |
| TNNT2  | VCL   | FLNC   | PRDM16 | DTNA    | SGCD   |
| TTN    | LMNA  | NEBL   | ZBTB17 | EYA4    | TNNI3K |
| TPM1   | TAZ   | NEXN   | TBX5   | GATAD1  | LRRC10 |
| MYBPC3 | RBM20 | CSRP3  | NKX2-5 | ILK     | NPPA   |
| TNNC1  | PLN   | TCAP   | GATA4  | LAMA4   | MIB1   |
| TNNI3  | DSP   | LDB3   | TBX20  | MYH6    | MYL3   |
| MYL2   | DSC-2 | CRYAB  | JPH2   | MYPN    | PDLIM3 |
| FHOD3  | PKP2  | BAG3   | ABCC9  | OBSCN   | PSEN1  |
| DES    | JUP   | ANKRD1 | CTF1   | PLEKHM2 |        |

### Appendix B: List of genes implicated in left ventricular non-compaction<sup>349</sup>

|        |        |        |        |         |
|--------|--------|--------|--------|---------|
| APOB   | SOS2   | ACTC1  | COL5A2 | MT-RNR1 |
| SNTA1  | ABCA1  | MYBPC3 | COL1A2 | MT-ND1  |
| TTN    | MAT2A  | CACNB2 | SCN5A  | MT-ATP6 |
| HCN4   | PCSK9  | SCN3B  | NOTCH1 | MT-RNR2 |
| RBM20  | RYR2   | MYH6   | MIB1   | COIII   |
| DSP    | LOX    | FBN2   | KCNH2  | MT-CYB  |
| LDLR   | TBX5   | COL5A1 | SOS2   | MT-ND4  |
| CAV3   | CBL    | LZTR1  | MT-ND5 | MT-ND4L |
| PRDM16 | NKX2-5 | PCSK9  | COII   |         |
| AKAP9  | KCNJ8  | TPM1   | NC5    |         |
|        |        |        |        |         |

### Appendix C: List of genes implicated in arrhythmogenic cardiomyopathy<sup>350-353</sup>

|        |        |
|--------|--------|
| PKP2   | TMEM43 |
| DSP    | TTN    |
| DSG2   | SCN5A  |
| DSC2   | RYR2   |
| JUP    | PLN    |
| CTNNA3 | TGFB3  |
| CDH2   | RBM20  |
| TJP1   |        |
| LMNA   |        |
| DES    |        |
| FLNC   |        |

## Appendix D: Ethics Annual Renewal for study HREC REF 198/2020





**FHS016: Annual Progress Report / Renewal**

|                                                                                                     |                        |                                                                                   |                    |
|-----------------------------------------------------------------------------------------------------|------------------------|-----------------------------------------------------------------------------------|--------------------|
| <b>HREC office use only (FWA00001637; IRB00001938)</b>                                              |                        |                                                                                   |                    |
| <b>This serves as notification of annual approval, including any documentation described below.</b> |                        |                                                                                   |                    |
| <input checked="" type="checkbox"/> Approved                                                        | Annual progress report | Approved until/next renewal date                                                  | 28/02/23           |
| <input type="checkbox"/> Not approved                                                               | See attached comments  |                                                                                   |                    |
| Signature Chairperson of the HREC/ Designee                                                         |                        |  | Date Signed 2/3/22 |

**Note:** Please email this form and supporting documents (if applicable) in a combined pdf-file to [hrec-enquiries@uct.ac.za](mailto: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>

|                              |
|------------------------------|
| Comments to PI from the HREC |
|------------------------------|

**Principal Investigator to complete the following:**

**1. Protocol information**

|                                                                                                                                               |                                                   |                                           |                                        |
|-----------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------|-------------------------------------------|----------------------------------------|
| Date (when submitting this form)                                                                                                              |                                                   |                                           |                                        |
| HREC REF Number                                                                                                                               | 198/2020                                          | Current Ethics Approval was granted until | 28 February 2022                       |
| Protocol title                                                                                                                                | The molecular genetics of familial cardiomyopathy |                                           |                                        |
| Protocol number (if applicable)                                                                                                               |                                                   |                                           |                                        |
| Are there any sub-studies linked to this study?                                                                                               |                                                   | <input type="checkbox"/> Yes              | <input checked="" type="checkbox"/> 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                                                                                                                        | Prof. Ntobeko Ntusi                               |                                           |                                        |

## Appendix E: Ethics Approval for study HREC REF 198/2020

16<sup>th</sup> February 2021

**HREC REF: 198/2020**

**Professor Ntobeko Ntusi**  
Department of Medicine  
J46, J – Floor, Old Main Building  
Groote Schuur Hospital  
Email: [Ntobeko.ntusi@uct.ac.za](mailto:Ntobeko.ntusi@uct.ac.za)

Dear Professor Ntobeko Ntusi

**PROJECT TITLE: THE MOLECULAR GENETICS OF FAMILIAL CARDIOMYOPATHY. SUB-STUDY 766/2014. PHD CANDIDATE MR S KAMULI**

Thank you for submitting your study to the Faculty of Health Sciences Human Research Ethics Committee.

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study, subject to the following:

1. The HREC require a study closure report (FHS010) for HREC REF NO. 174/2019, as we previously requested on the 3<sup>rd</sup> June 2019.
2. In Addition, while we acknowledge IMHOTEP informed consent forms dated August 2020, please submit the updated IMHOTEP informed consent forms once these has been finalised.

**Approval is granted for one year until the 28<sup>th</sup> February 2022.**

**This approval is subject to strict adherence to the HREC recommendations regarding research involving human participants during COVID-19, dated 17 March 2020 and 06 July 2020.**

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](http://www.health.uct.ac.za/fhs/research/humanethics/forms))

The HREC also acknowledges that *Mr S Kamuli* will be involved in the 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 also 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.**

HREC REF 198/2020 SC

## Appendix F: Informed DNA consent for genetic analysis



### REQUEST FOR MOLECULAR STUDIES (DNA)

Molecular Laboratory, Division of Cardiovascular Genetics, Hatter Institute  
4<sup>th</sup> Floor, Chris Barnard Building, UCT Medical School, Observatory, Cape Town, 7925  
Tel: +2721 4066615; Fax: +2721 4478789



IMHOTEP: HREC REF R050/2014

### CONSENT FOR PARTICIPATION IN GENETICS RESEARCH

IMHOTEP Primary and Secondary ID

|  |  |  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|--|--|
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |

Instructions: Blood should be drawn in plastic EDTA Tubes (purple top) or Paxgene Tubes. 20 ml of blood should be collected for adults. For children, 5-20ml depending on size of the child (maximum 3ml/kg). Please label blood tubes with barcoded sticker/study number. Place blood in the fridge at 4°C (or cooler box) until able to send to laboratory. Please DO NOT send specimens on ice or frozen.

#### Patient details: (or hospital sticker here)

Surname: \_\_\_\_\_ First Name(s): \_\_\_\_\_

Hospital folder #: \_\_\_\_\_ Sex: ☐ M ☐ F ☐ Date of Birth (DD/MM/YYYY): \_\_\_\_/\_\_\_\_/\_\_\_\_

Parent's/Guardian's name(s): \_\_\_\_\_

Patient address: \_\_\_\_\_

Contact numbers: \_\_\_\_\_

Email address: \_\_\_\_\_

Referring hospital/doctor: \_\_\_\_\_

Referring doctor's contact details: \_\_\_\_\_

#### Clinical information: (PLEASE COMPLETE A FAMILY PEDIGREE OVER THE PAGE)

PROBAND ☐ RELATIVE ☐ If relative, family relationship to proband: \_\_\_\_\_

New Family: Yes ☐ No ☐ If no, please indicate family name/ID \_\_\_\_\_

Family surname and ID: \_\_\_\_\_ 3 - \_\_\_\_\_

Ethnic Origin: Black African ☐ Mixed race ☐ Caucasian ☐ Asian ☐ Other \_\_\_\_\_

#### Provisional Diagnosis:

Clinically affected ☐ At Risk (unaffected clinically) ☐ Spouse ☐ Query ☐

| Morphofunctional phenotype                           | Specific diagnosis                                                                                                                                     |
|------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|
| <input type="checkbox"/> Dilated cardiomyopathy      | <input type="checkbox"/> Familial <input type="checkbox"/> Idiopathic <input type="checkbox"/> PPCM <input type="checkbox"/> Secondary (specify) _____ |
| <input type="checkbox"/> Hypertrophic cardiomyopathy | <input type="checkbox"/> HCM <input type="checkbox"/> HCM – phenocopy (specify) _____                                                                  |
| <input type="checkbox"/> ARVC                        | <input type="checkbox"/> Definite <input type="checkbox"/> Borderline <input type="checkbox"/> Possible <input type="checkbox"/> Unconfirmed           |
| <input type="checkbox"/> Restrictive cardiomyopathy  | <input type="checkbox"/> Familial <input type="checkbox"/> Idiopathic <input type="checkbox"/> Specific aetiology (specify) _____                      |
| <input type="checkbox"/> Unspecified cardiomyopathy  | <input type="checkbox"/> LVNC <input type="checkbox"/> Takusubu <input type="checkbox"/> Other (specify) _____                                         |

Additional disorders (apparent or previously treated): \_\_\_\_\_

Have samples from this patient been sent to a DNA lab before? ☐ YES ☐ NO ☐ Don't Know

If yes, please specify lab: \_\_\_\_\_

DATE OF SPECIMEN COLLECTION: \_\_\_\_/\_\_\_\_/\_\_\_\_, TIME: \_\_\_\_\_

#### For Laboratory use only:

DNA number \_\_\_\_\_ Vol. blood \_\_\_\_\_ (ml) Other: \_\_\_\_\_

Date received (DD/MM/YYYY) \_\_\_\_/\_\_\_\_/\_\_\_\_

#### SOP\_Informed Consent Procedures\_APPENDIX C(b)

#### Summary of broad consent for participation in genetics research in minors

I understand that we have agreed to **participate in genetics research** that will be conducted at the molecular genetics laboratory at the Hatter Institute for Cardiovascular Research in Africa, situated at the University of Cape Town in South Africa. This laboratory, in collaboration with other local and international laboratories, is dedicated to doing **research related to genetic causes of cardiovascular disease in people living in Africa**.

I understand that a **blood sample (5-25ml or 1-5 teaspoons of blood)** will be collected from my child and genetic material will be extracted from this sample for analysis. In some instances, other samples may be collected depending on the circumstances (please specify if collected): \_\_\_\_\_

I understand that a portion of this genetic material will be **stored** at the Hatter Institute for **additional cardiomyopathy-related research projects** approved by the University of Cape Town Human Research Ethics Committee (HREC). My child's genetic material/information, together with other relevant medical information, may be shared with other researchers and institutions involved in HREC approved studies but my personal identifying information will not be shared. I authorize that our doctors can provide relevant clinical information (medical records) to researchers.

I understand that the nature of research means that **we may or may not receive a result from studies performed on my DNA**. Although the laboratory will do its best to confirm that the findings related to my child's condition, results received from a research laboratory should be confirmed diagnostically. If a genetic cause for my child's condition is found, the researchers from the Hatter Institute will do their best to inform me of the results, either via my doctor, a genetics counsellor or in writing, depending on the available resources. If my contact details change, it is my responsibility to inform the laboratory. In the event that I am unavailable or incapacitated, **I do / do not (please delete where not applicable)** want my immediate family members to be informed of the results.

I understand that the genetics molecular laboratory is under obligation to **respect our confidentiality**. I understand that our participation in genetics research is entirely voluntary and that we may withdraw our consent at any time without it affecting our future medical care. **We have been provided with an information sheet explaining the risks and benefits of participating in genetics research, and the contact information for the laboratory.**

Participant/Guardian's name \_\_\_\_\_ Signature \_\_\_\_\_ Date \_\_\_\_\_

Study doctor/nurse/genetics counsellor \_\_\_\_\_ Signature \_\_\_\_\_ Date \_\_\_\_\_

Consent for DNA collection, storage and analysis (minors). Version 3.0. June 2020

## Appendix G: Storage and governance SOP for IMHOTEP



CARDIOVASCULAR GENETICS (CVG) LABORATORY

|                     |                                                                |                       |               |
|---------------------|----------------------------------------------------------------|-----------------------|---------------|
| <b>SOP TITLE:</b>   | <b>Governance, storage and disposal/destruction of samples</b> | <b>SOP#</b>           | S223-01       |
| <b>WRITTEN BY:</b>  | A/Prof Gasnat Shaboodien                                       | <b>VERSION</b>        | 1.4           |
|                     |                                                                | <b>SUPERSEDES</b>     |               |
|                     |                                                                | <b>LINKED DOCS</b>    |               |
| <b>APPROVED BY:</b> | HREC (2022/05/09)                                              | <b>EFFECTIVE DATE</b> | December 2018 |

1. Purpose: **Governance and storage of samples**
2. Applies to: All DNA samples in the Cardiovascular Genetics (CVG) Biorepository
3. Responsibilities: The Cardiovascular genetics laboratory in Cape Town

REFERENCE:

**Sheina Josias** (Academic Space Administrator)  
Faculty of Health Sciences  
University of Cape Town  
Tel: 021 406 6725  
[Sheina.josias@uct.ac.za](mailto:Sheina.josias@uct.ac.za)



## 1. Terms and Definitions

| ABBREVIATIONS/DEFINITIONS        | DESCRIPTION                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|----------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Biological Material</b>       | <p>'Biological material' means material from a human being Which includes blood and blood products, DNA, RNA, blastomeres, polar bodies, cultured cells, embryos, gametes, progenitor stem cells, saliva, sputum, gastric juices, urine, small tissue biopsies and growth factors (Adjusted from DoH, 2015).</p> <p>Biological material refers to a sample obtained from an individual human being, living or deceased, which provides biological information, including genetic information, about that individual (WMA, 2016).</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| <b>Donor</b>                     | 'Donor' means a person from whose body human biological material has been removed or withdrawn for the purpose of genetic testing, genetic training, genetic health research and therapeutics. <sup>1</sup>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| <b>Anonymous Samples or Data</b> | Data or biological specimens obtained by a researcher without any identifying information and without a link to a specific participant or donor.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| <b>Identifiers</b>               | Information that could be associated with a specific research participant such as name, address, medical folder number, phone or fax number or biometric identifier (e.g. finger print).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| <b>Coded Data or Samples</b>     | Identifiers have been replaced with a number, symbol or a letter and a key exists to decipher the code, allowing linkage of the code to a specific individual.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| <b>Broad Consent</b>             | <p>The donor permits use of data and biological materials for future studies, <i>subject only to further prior ethics review and approval by the HREC</i>. It allows for storage and possible future use for research purposes, even though the precise nature of future research may be unclear at the time.</p> <p>In the informed consent documentation, the <i>nature of future use should be described as fully as possible and stipulate that further ethics review by the HREC of a new study or sub-study is necessary</i>. Consent can then be <i>waived and the HREC becomes the advocate for the participant</i>. Permission may also be sought by <i>re-consent</i> if the planned research is outside the scope of the current consent (DoH, 2015).</p> <p>Note: See Appendix A, Point 4 for an example of a template for storage and future use of unused samples of biological materials (Ethics in Health Research: Principles, Processes and Structures, DoH, 2015).</p> |
| <b>Blanket Consent</b>           | Blanket consent refers to a process by which individuals donate their samples without any restrictions. This standard operating procedure does not support blanket consent for future storage of samples.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| <b>Tiered consent</b>            | The donor permits use of biological materials for the current study and chooses                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |

## Appendix H: Buffers and solutions

### 1 M Tris-HCl (pH 8.0):

4. 121.1 g Tris base (Glenthams Life Sciences, Corsham, UK) in 800 ml distilled water
5. Adjust the pH to 8.0 with concentrated HCl
6. Make to a total volume of 1 L with sterile distilled water

### 0.5 M EDTA (pH 8.0):

7. 186.1 g EDTA (Glenthams Life Sciences) in 800 ml distilled water
8. Adjust the pH to 8.0 with NaOH
9. Make to a total volume of 1 L with sterile distilled water

### 1X Tris/EDTA (TE) buffer:

- 10.1 ml 1 M Tris-HCl (final concentration: 10 mM)
- 11.0.2 ml 0.5 M EDTA (final concentration: 1 mM)
- 12.98.8 ml sterile distilled water

### Tris/Borate/EDTA (TBE) buffers

#### 10X TBE (stock):

- 13.121.1 g Tris (Glenthams Life Sciences) (final concentration: 1 M)
- 14.61.8 g Boric acid (AMRESCO, Solon, OH USA) (final concentration: 1 M)
- 15.7.4 g EDTA (Glenthams Life Sciences) (final concentration: 0.02 M)
16. Made to a total 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

### Agarose gel electrophoresis

#### Loading dye:

1 ml 5X Green GoTaq® Flexi buffer (Promega)

5 µl GelRed® nucleic acid gel stain (Biotium)

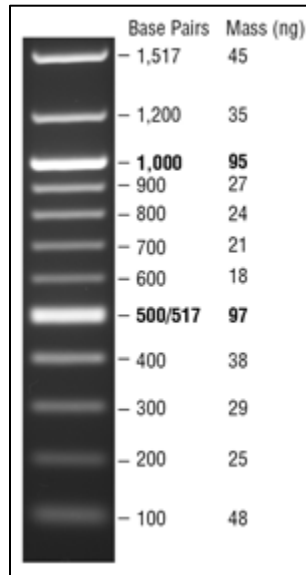
Mixed with DNA samples at a ratio of 3:6 (dye:sample) before loading

#### 1.5% agarose gel:

1.5 g SeaKem® LE agarose (Lonza, Basel, Switzerland)

100 ml 1X TBE

## Appendix I: New England Biolabs<sup>®</sup> 100bp DNA ladder



## APPENDIX J: CMO Genes

|                 |                |               |                |                |               |
|-----------------|----------------|---------------|----------------|----------------|---------------|
| <b>AARS2</b>    | <b>COA5</b>    | <b>GJA5</b>   | <b>KRAS</b>    | <b>PDLIM3</b>  | <b>TAZ</b>    |
| <b>ABCC9</b>    | <b>COX15</b>   | <b>GLA</b>    | <b>LAMA4</b>   | <b>PKP2</b>    | <b>TCAP</b>   |
| <b>ACADVL</b>   | <b>CRYAB</b>   | <b>GLB1</b>   | <b>LAMP2</b>   | <b>PLN</b>     | <b>TGFB3</b>  |
| <b>ACTA1</b>    | <b>CSRP3</b>   | <b>GPD1L</b>  | <b>LDB3</b>    | <b>PRKAG2</b>  | <b>TLN1</b>   |
| <b>ACTC1</b>    | <b>CTF1</b>    | <b>GUSB</b>   | <b>LMNA</b>    | <b>PSEN1</b>   | <b>TMEM43</b> |
| <b>ACTN2</b>    | <b>CTNNA3</b>  | <b>HCN4</b>   | <b>LMOD2</b>   | <b>PSEN1</b>   | <b>TMEM70</b> |
| <b>AGK</b>      | <b>DES</b>     | <b>HFE</b>    | <b>MAP2K1</b>  | <b>PSEN2</b>   | <b>TMOD1</b>  |
| <b>AGL</b>      | <b>DMD</b>     | <b>HRAS</b>   | <b>MAP2K2</b>  | <b>PTPN11</b>  | <b>TMPO</b>   |
| <b>AKAP1</b>    | <b>DMPK</b>    | <b>HRC</b>    | <b>MIB1</b>    | <b>RAF1</b>    | <b>TNNC1</b>  |
| <b>AKAP10</b>   | <b>DNAJC19</b> | <b>HSPB2</b>  | <b>MRPL3</b>   | <b>RANGRF</b>  | <b>TNNI3</b>  |
| <b>AKAP12</b>   | <b>DNM1L</b>   | <b>HSPB6</b>  | <b>MURC</b>    | <b>RBM20</b>   | <b>TNNT2</b>  |
| <b>AKAP13</b>   | <b>DOLK</b>    | <b>HSPB7</b>  | <b>MYBPC3</b>  | <b>RBM24</b>   | <b>TPM1</b>   |
| <b>AKAP5</b>    | <b>DPP6</b>    | <b>HSPB8</b>  | <b>MYH6</b>    | <b>RYR2</b>    | <b>TRDN</b>   |
| <b>AKAP6</b>    | <b>DMPK</b>    | <b>ILK</b>    | <b>MYH7</b>    | <b>SCN10A</b>  | <b>TRIM63</b> |
| <b>AKAP7</b>    | <b>DNAJC19</b> | <b>JPH2</b>   | <b>MYL2</b>    | <b>SCN1B</b>   | <b>TSFM</b>   |
| <b>AKAP9</b>    | <b>DNM1L</b>   | <b>JUP</b>    | <b>MYL3</b>    | <b>SCN2B</b>   | <b>TTN</b>    |
| <b>ALPK3</b>    | <b>DOLK</b>    | <b>KCNA5</b>  | <b>MYLK2</b>   | <b>SCN3B</b>   | <b>TTR</b>    |
| <b>ANK2</b>     | <b>DPP6</b>    | <b>KCND3</b>  | <b>MYOM1</b>   | <b>SCN4B</b>   | <b>TXNRD2</b> |
| <b>ANKRD1</b>   | <b>DSC2</b>    | <b>KCNE1</b>  | <b>MYOT</b>    | <b>SCN5A</b>   | <b>USP13</b>  |
| <b>ASPH</b>     | <b>DSG2</b>    | <b>KCNE1L</b> | <b>MYOZ2</b>   | <b>SCN7A</b>   | <b>UTRN</b>   |
| <b>ATP5E</b>    | <b>DSP</b>     | <b>KCNE2</b>  | <b>MYPN</b>    | <b>SCO2</b>    | <b>VCL</b>    |
| <b>BAG3</b>     | <b>DTNA</b>    | <b>KCNE3</b>  | <b>NDUFAF1</b> | <b>SDHA</b>    | <b>VIM</b>    |
| <b>BRAF</b>     | <b>EMD</b>     | <b>KCNE4</b>  | <b>NDUFS2</b>  | <b>SGCD</b>    | <b>XIRP1</b>  |
| <b>CACNA1C</b>  | <b>EYA4</b>    | <b>KCNH2</b>  | <b>NEBL</b>    | <b>SHOC2</b>   | <b>XK</b>     |
| <b>CACNA2D1</b> | <b>FHL1</b>    | <b>KCNIP2</b> | <b>NEXN</b>    | <b>SLC25A3</b> | <b>ZYX</b>    |
| <b>CACNA2D2</b> | <b>FHL2</b>    | <b>KCNJ11</b> | <b>NOS1AP</b>  | <b>SLC25A4</b> |               |
| <b>CACNA2D3</b> | <b>FKTN</b>    | <b>KCNJ2</b>  | <b>NPPA</b>    | <b>SLMAP</b>   |               |
| <b>CACNB2</b>   | <b>FLNC</b>    | <b>KCNJ4</b>  | <b>NRAS</b>    | <b>SNTA1</b>   |               |
| <b>CALM1</b>    | <b>FOXD4</b>   | <b>KCNJ8</b>  | <b>OBSCN</b>   | <b>SOS1</b>    |               |
| <b>CALR3</b>    | <b>FOXRED1</b> | <b>KCNK1</b>  | <b>PDE3A</b>   | <b>SPRED1</b>  |               |
| <b>CASQ2</b>    | <b>FXN</b>     | <b>KCNN2</b>  | <b>PDE4A</b>   | <b>SVIL</b>    |               |
| <b>CAV3</b>     | <b>GAA</b>     | <b>KCNN3</b>  | <b>PDE4B</b>   | <b>SYNE1</b>   |               |
| <b>CBL</b>      | <b>GATAD1</b>  | <b>KCNQ1</b>  | <b>PDE4D</b>   | <b>SYNE2</b>   |               |
| <b>CMYA5</b>    | <b>GJA1</b>    | <b>KLF10</b>  | <b>PDE4DIP</b> | <b>SYNM</b>    |               |

Appendix K: 38 gene panel used in Haloplex targeted analysis

|               |              |             |               |               |               |            |
|---------------|--------------|-------------|---------------|---------------|---------------|------------|
| <b>ACTC1</b>  | <b>CRYAB</b> | <b>DSP</b>  | <b>LAMP2</b>  | <b>PKP2</b>   | <b>TMEM43</b> | <b>TTR</b> |
| <b>ACTN2</b>  | <b>CSRP3</b> | <b>FHL1</b> | <b>LMNA</b>   | <b>PLN</b>    | <b>TNNC1</b>  | <b>VCL</b> |
| <b>ALPK3</b>  | <b>DES</b>   | <b>FHL2</b> | <b>MYBPC3</b> | <b>PRKAG2</b> | <b>TNNI3</b>  |            |
| <b>ANKRD1</b> | <b>DMD</b>   | <b>FLNC</b> | <b>MYH7</b>   | <b>RBM20</b>  | <b>TNNT2</b>  |            |
| <b>BAG3</b>   | <b>DSC2</b>  | <b>GLA</b>  | <b>MYL2</b>   | <b>SCN5A</b>  | <b>TPM1</b>   |            |
| <b>CDH2</b>   | <b>DSG2</b>  | <b>JUP</b>  | <b>MYL3</b>   | <b>TAZ</b>    | <b>TTN</b>    |            |

## Appendix L: Puregene DNA extraction protocol

### Gentra Puregene Blood Kit protocol – DNA purification from buffy coat

1. If the buffy coat preparation contains red blood cells, continue with step 2. Otherwise, pipet 3 ml Cell Lysis Solution into a 15 ml centrifuge tube, add 150–250  $\mu$ l sample, and continue with step 8.
2. Dispense 3 volumes RBC Lysis Solution into a 15 ml centrifuge tube (e.g., if processing 250  $\mu$ l buffy coat, dispense 750  $\mu$ l RBC Lysis Solution). Add 150–250  $\mu$ l 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 100–200  $\mu$ l of the residual liquid and the pellet.
6. Vortex the tube vigorously to resuspend the pellet in the residual liquid.
7. Add 3 ml 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 2 years at room temperature.
8. Optional: If RNA-free DNA is required, add 15  $\mu$ l RNase A Solution and mix by inverting 25 times. Incubate for 15 min at 37°C. Then incubate for 3 min on ice to quickly cool the sample. **(not done)**
9. Add 1 ml Protein Precipitation Solution and vortex vigorously for 20 s at high speed. **(1.5 ml Protein Precipitation Solution was added)**
10. Centrifuge for 5 min at 2000 x g. **(Samples were centrifuged for 10 min)**
11. Pipet 3 ml isopropanol into a clean 15 ml centrifuge tube and add the supernatant from the previous step by pouring carefully. **(6 ml 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 2 ml microcentrifuge tube)**
14. Carefully discard the supernatant and drain the tube by inverting on a clean piece of absorbent paper, taking care that the pellet remains in the tube.
15. Add 3 ml of 70% ethanol and invert several times to wash the DNA pellet. **(1 ml 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 300  $\mu$ l DNA Hydration Solution and vortex for 5 s at medium speed to mix. **(200  $\mu$ l DNA Hydration Solution was added)**
19. Incubate at 65°C for 1h to dissolve the DNA. **(not done)**

Incubate at room temperature overnight.