Genealogical Tracing of Founder Variants Linked to Cardiomyopathies in a South African Cohort.

By Abbey Houghton (HGHABB001)

Dissertation presented for the degree of

Master of Science in the Department of Medicine

University of Cape Town

February 2025

Supervisor: Associate Professor Gasnat Shaboodien

Co-Supervisor: Professor Ntobeko Ntusi

DECLARATION

I, *Abbey Houghton,* 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: 17/02/2025

Table of Contents

List of Figures

List of Tables

List of abbreviations

ACM	Arrhythmogenic Cardiomyopathy
ACMG	American College of Medical Genetics and Genomics
ACTN2	Alpha-actinin-2
ALVC	Arrhythmogenic Left Ventricular Cardiomyopathy
ARVC	Arrhythmogenic Right Ventricular Cardiomyopathy
BAG3	Bcl-2-associated anthanogene 2 protein
bp	Base Pair
CASSA	Cardiac Arrhythmia Society of Southern Africa
CHI	Cape Heart Institute
CMR	Cardiac Magnetic Resonance
cTnT	Cardiac Troponin T
CVG	Cardiovascular Genetics
DCM	Dilated Cardiomyopathy
DES	Desmin
dH2O	Distilled Water
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleotide Triphosphates
DRC/NGK	Dutch Reformed Church
DSC2	Desmocollin-2
dsDNA	Double-stranded DNA
DSG2	Desmoglein-2
DSP	Desmoplakin
ECG	Electrocardiogram
FHOD3	Formin Homology Domain Containing 3
FKTN	Fukutin
FLNC	Filamin C
GCP	Good Clinical Practice
gDNA	Genomic DNA
GISA	Genealogical Institute of South Africa
GNOMAD	Genome Aggregation Database
HCM	Hypertrophic Cardiomyopathy
HF	Heart Failure
HREC	Human Research Ethics Committee
HRM	High-Resolution Melt

HSP70	Heat shock protein 70
ICD	Implantable Cardioverter-Defibrillator
IMHOTEP	The Africa cardiomyopathy and myocarditis registry programme
JPH2	Junctophilin-2
JUP	Junction Plakoglobin
kbp	Kilobase pair
LMNA	Lamin A/C
LVEF	Left Ventricular Ejection Fraction
LVNC	Left Ventricular Non-Compaction Cardiomyopathy
LVOT	Left Ventricular Outflow Tract
MAF	Minor allele frequency
MgCl2	Magnesium Chloride
MYBPC3	Myosin Binding Protein C3
MYH7	Myosin Heavy Chain 7
NEXN	Nexilin F-actin binding protein
Ng	Nanogram
NKX2-5	NK2 Homeobox 5
P/LP	Pathogenic/Likely Pathogenic
PCR	Polymerase Chain Reaction
PKP2	Plakophilin-2
PLN	Phospholamban
RBM20	RNA binding motif protein 20
RCM	Restrictive Cardiomyopathy
SA	South Africa
SAGs	South African Genealogies
SCD	Sudden Cardiac Death
SCN5A	Sodium voltage-gated channel alpha subunit 5
SERCA2	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2
SNP	Single Nucleotide Polymorphism
SSA	Sub-Saharan Africa
STR	Short Tandem Repeat
TA	Annealing Temperature
TBE	Tris-Borate-EDTA
TFC	Task Force Criteria
TGFB3	Transforming Growth Factor Beta 3
TMEM43	Transmembrane Protein 43

TNNI3	Troponin I3
TNNT2	Troponin T2
TPM1	Tropomyosin 1
TTN	Titin
UCSC	University of California, Santa Cruz
UCT	University of Cape Town
uL	Microliter
VLC	Vinculin
VOC	Dutch East India Company
VT	Ventricular Tachycardia
VUS	Variant of uncertain significance
WES	Whole exome sequencing
WGS	Whole genome sequencing
βMHC	Beta-Myosin Heavy Chain

Acknowledgments

I would like to acknowledge and thank my supervisor, A/Prof. Gasnat Shaboodien and co-supervisor, Prof. Ntobeko Ntusi, of the Cardiovascular Genetics (CVG) group of the Cape Heart Institute (CHI), Department of Medicine at the University of Cape Town (UCT). Without their guidance, patience and support this work would not have been possible, thank you.

I would also like to extend a special thanks to many of the experts who enabled this research:

1. Dr. Tafadzwa Machipaisa for her prior works and guidance throughout this project
2. Dr. Sarah Krause and Mrs Nakita Lang for collection, summarising and relaying all the participant data, having constant open communication throughout this project has meant so much to me.
3. Prof. Gerhard Geldenhuys, an expert in Genealogy tracing, without his guidance, support, kindness and passion for genealogy tracing this project would not have been possible, and I am extremely grateful to have had such a wonderful teacher .
4. To the team at Inqaba for teaching me the skills I needed as well as their patience and guidance during troubleshooting.
5. To the probands and their families, their participation in this study regardless of the hardships makes me remember why I am so passionate about science, thank you for your contribution.
6. To the CVG team: Dr. Polycarp Ndibangwi, Lameez Pearce, Kayla Wagenaar and Joey Carollissen for their continued support, guidance

I would like to express my gratitude and thanks to my family and friends who have seen the highs and lows of the past two years. I am especially thankful to my parents, Adam and Kay for their unwavering love, support and encouragement, I hope I have made you both proud. Lastly, Daven, thank you for always encouraging me to do my best and for taking a special interest in listening to my presentations and rambling on about complicated and confusing concepts, which undoubtedly caused more confusion. This Master's wouldn't have been possible without them.

Abstract

Introduction:

Cardiomyopathies are a major cause of heart failure in South Africa, yet their prevalence, causes, and outcomes remain poorly understood. The IMHOTEP registry was established to clinically and genetically characterise affected patients and through this study several common pathogenic variants were identified, suggesting potential founder variants. This study aimed to genealogically trace the origins of these variants in the probands and their families.

Methods:

Participants were recruited from South African tertiary hospitals, with baseline data recorded at enrolment. Next generation sequencing identified several probands with possible founder variants which prompted haplotype construction and genealogical tracing. Variants were identified in the genes *PKP2*, *LMNA*, *BAG3* and *TTN* and three microsatellite markers, spanning the 5', intergenic and 3' regions, were designed for each of the genes of interest in order to perform haplotype analysis. Genealogical data was collected using online resources, genealogy databases, and the voter's roll.

Results:

Five *PKP2* c.1162C>T (p.Arg388Trp), *BAG3* c.925C>T (p.Arg309Ter), LMNA c.568C>T (p.Arg190Trp), *TTN* c.95008C>T (p.Arg31670Ter) and *TTN* c.87624C>G (p.Tyr29208Ter) variants were identified in 38 probands (12 probands with the *PKP2* variant; three probands with *BAG3* variant, three probands with *LMNA* variant and 20 probands with the two *TTN* variants). *PKP2* was traced to Afrikaner families of Dutch/French Huguenot descent from the 1600s. *BAG3* was linked to French Huguenot and Canadian ancestry. *LMNA* tracing was limited due to the number of families as well as limited Mixed-Ancestry family records. *TTN* variants had issues with haplotyping and genealogical challenges as well as limited records for Black African and Mixed-Ancestry families.

Conclusions:

This research significantly advances our understanding of cardiomyopathy genetics in South Africa. The identification of Dutch/French and French-Canadian ancestries associated with the *PKP2* and *BAG3* variants, respectively, provides valuable insights into the historical origins and population-specific genetic landscape of these cardiomyopathies. This ancestral information not only deepens our understanding of cardiomyopathy in the region but also has important implications for risk assessment and genetic counselling. By identifying common

progenitors, we can better target screening efforts and provide more accurate risk assessments for individuals carrying these founder variants.

Chapter 1: Introduction

1.1 Cardiomyopathies

Heart failure (HF) is the leading cause of mortality worldwide and was classified by the Global Burden of Disease as a pandemic in 2017 (3). The prevalence, incidence and outcomes of HF varies worldwide due to different geographical and socioeconomic factors (4).

On the African continent, the primary factors contributing to HF are cardiomyopathies, rheumatic heart disease, pericardial heart disease, and hypertension, with approximately 20-30% of HF cases in the region linked to cardiomyopathies (5, 6). A recent systematic review has highlighted the distinct epidemiological and clinical profile of heart failure in Sub-Saharan Africa (SSA), characterised by a predominance of dilated cardiomyopathy, hypertensive heart disease, and region-specific aetiologies such as rheumatic heart disease and peripartum cardiomyopathy (7). These findings emphasise the critical need for further investigation into the underlying mechanisms, particularly in the context of inherited cardiomyopathies, where understanding remains limited.

Cardiomyopathies are defined as disorders of the myocardium; therefore, the structure and/or function of the heart may be affected (8). They occur in the absence of congenital heart disease, coronary heart disease, hypertension, and valvular disease (9). The classification system of cardiomyopathies has evolved over the past few decades (10, 11) and five main types have been identified all having genetic components which are heritable: Dilated cardiomyopathy (DCM), Arrhythmogenic cardiomyopathy (ACM), Hypertrophic cardiomyopathy (HCM), Restrictive cardiomyopathy (RCM) and Left Ventricular Non-compaction cardiomyopathy (LVNC), Figure 1 (Adapted from (2)). However, there is conflicting information regarding LVNC which will be discussed in section 1.1.5.

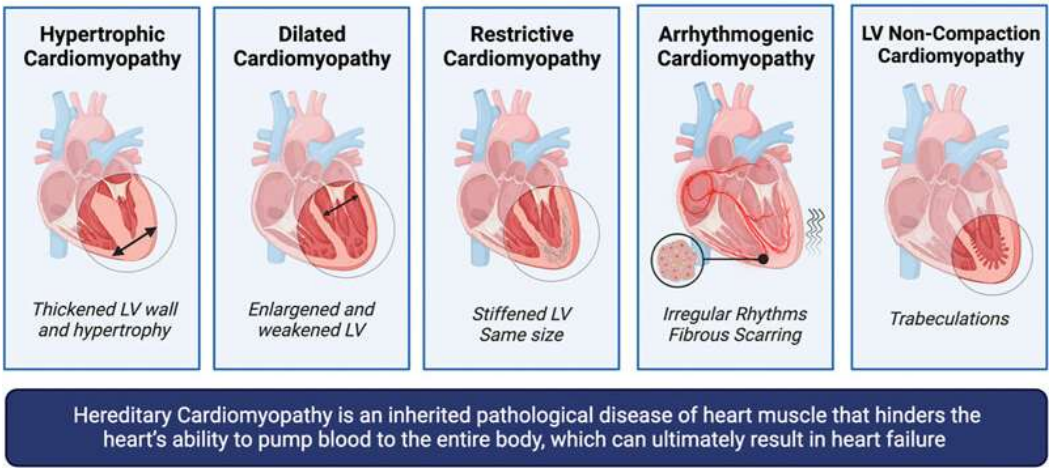


Figure 1: Visual representation of the five types of cardiomyopathies with each of their distinct characteristics highlighted (LV = Left Ventricle)(2)

1.1.1 Dilated Cardiomyopathy (DCM)

Dilated cardiomyopathy is defined as the enlargement and dilation of one or both ventricles resulting in decreased contractility of the heart (12) therefore, patients with an impaired ejection fraction of <40% are classified as having DCM (13).

1.1.1.1 Prevalence, Incidence, and Mortality

Dilated Cardiomyopathy accounts for 55% of cardiomyopathies, making it the most common cardiomyopathy (14), with a global incidence estimated to be 1:250 to 1:500 (15). Although there has been notable progression in HF therapies, the morality rate of DCM remains high, with 5-year survival rate of approximately 50% after diagnosis without heart transplantation (16), thus making DCM the leading cause of heart transplantation (17). Causes of death include SCD and HF progression due to pump failure (18).

1.1.1.2 Aetiology

The aetiology of DCM is heterogenous and causes include viral infections (such as Coxsackievirus B), chronic alcohol abuse, inherited metabolic disorders, nutritional deficiencies, and drug toxicity (12, 19). Additionally, autoimmune diseases and exposure to toxins like chemotherapy agents (e.g., doxorubicin) can induce DCM (20). In many cases, however, the cause remains idiopathic.

1.1.1.3 Diagnosis

Considering the above aetiology of DCM and the complex, multifactorial nature of this disease, the diagnosis usually employs a systematic approach. This allows clinicians to accurately identify and manage or treat this condition.

The first step of diagnosis is a clinical workup of the patient and patient's family history which is followed by a physical examination (21). A pedigree of three or more generations is constructed which focuses on premature cardiovascular-related deaths (i.e. HF or SCD) and other associated cardiac and non-cardiac phenotypes (22). An electrocardiogram (ECG) is done to detect signs of left ventricular hypertrophy, T-wave changes, left branch bundle block as well as interventricular conduction delay which may or may not be present in the early stages of DCM. In more advanced cases of DCM, the ECG may detect low voltage accompanied by low R-wave amplitudes which is indicative of myocardial fibrosis (23).

Thereafter, two imaging techniques can be used for diagnosis, the first is an echocardiogram which provides information about chamber dimension, morphology, severity of valve disease as well as systolic and diastolic function (24). The gold standard technique for imaging is cardiac magnetic resonance (CMR), it quantifies LVEF and volumes accurately and more precisely than the echocardiography (25).

If the patient has a positive family history of two or more first-degree relatives diagnosed with DCM or sudden cardiac death (SCD) <35 years old, genetic testing recommended to evaluate DCM and the underlying cause (26).

1.1.1.4 Genetics

Dilated cardiomyopathy (DCM) has a significant genetic component, with approximately 20%–35% of cases associated with certain variants (26). The genetics of DCM have significant overlap with other cardiomyopathies which may explain the wide heterogeneity observed in research (13, 27). Approximately 40% of DCM cases are associated with pathogenic or likely pathogenic (P/LP) genetic variants which occur in between 40-60 genes associated with DCM (28). A few of these genes are: *TTN*, *LMNA*, *BAG3, DES, FLNC, PLN, SCN5A, NEXN, MYH7*, *DSP*, and *TNNT2* (12, 29-31), as shown in Figure 2 (Adapted from (1)) and in Figure 3 (Adapted from (32)). The inheritance pattern is most often autosomal dominant; however, autosomal recessive and X-linked patterns have been reported (26).

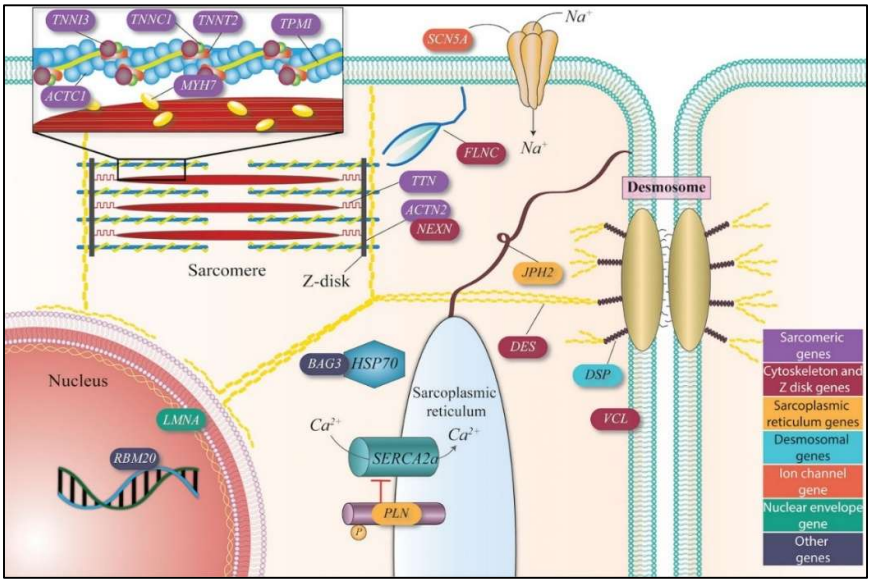


Figure 2: Schematic view of DCM-associated P/LP genes. The key on the right of the image shows the colour coding of genes based on the location of their resulting proteins in the context of cardiomyocytes(1)

Gene	Protein	Estimated contribution to total number of patients	Related mechanisms that trigger DCM	Clinical manifestation
TTN	Titin	25%	Haploinsufficiency: poison peptide/dominant negative mechanism; Disruption of cardiac metabolism and signalling, and loss of function.	Arrhythmias (atrial fibrillation and ventricular arrhythmias)
LMNA	Lamin A/C	6%	Mechanical hypotheses; Gene expression hypotheses; Cytotoxicity hypotheses	Conduction system disorders (sinus bradycardia, sinus node arrest with junctional rhythm, or heart block); arrhythmias (chronic atrial fibrillation/supraventricular arrhythmias/ventricular arrhythmias)
DSP	Desmoplakin	NA	Inflammation	Palmoplantar keratoderma; Fatal ventricular arrhythmia; Acute myocardial injury;
DES	Desmin	<1.6%	Affects the structure of the heart; Affects myocardial metabolism of glucose, fatty acids and amino acids	Neurologic symptoms (myopathy or muscle weakness); cardiac symptoms (heart block: Atrioventricular block, arrhythmia: atrial fibrillation); respiratory insufficiency
MYH7	Myosin 7	1-5.3%	Damage to the structural or functional integrity of the myotome	Atrial fibrillation; Atrial fibrillation
BAG3	BAG family molecular chaperone regulator 3	NA	cellular senescence	Progressive heart failure
FLNC	Filamin C	NA	Activation of platelet-derived growth factor receptor-α pathway	left ventricular dilatation with systolic dysfunction, myocardial fibrosis
PLN	Cardiac phospholamban	<1%	Defects in calcium metabolism	Early ventricular arrhythmia, end-stage heart failure
RBM20	RNA-binding protein 20	3%	Alters expression of protein subtypes that maintain muscle structure and heart function;	Severe impairment of cardiac systolic function; Cardiac arrhythmias (sustained ventricular arrhythmias/sudden cardiac death/atrial fibrillation)
SCN5A	Sodium channel protein type 5 subunit α	2–3%	Loss of functionality in the cardiac sodium channel; Disruption of the interaction between SCN5A and other constituents, leading to structural alterations and impairment of contraction.	Arrhythmias (atrial fibrillation/ventricular tachycardia), heart block.
TNNC1	Troponin C, slow skeletal and cardiac muscles	<1%	Change its binding affinity for Ca2+ or by altering the interaction of cTnC with its binding partner; reduced sensitivity of myofilaments to Ca2+	Early severe systolic heart failure
TNNT2	Troponin T, cardiac muscle	3%	XIN deficiency	Cardiac dilation, systolic dysfunction, cardiac fibrosis

Figure 3: Main genes associated with DCM and their diagnostic yield (32)

1.1.1.5 African Context

Not much is known about the precise burden of DCM in African populations. Although research on cardiomyopathies has increased in sub-Saharan Africa (SSA) (33-35), data on DCM prevalence, incidence, and genetics, particularly in South Africa, remain limited. Some

studies suggest a high prevalence of non-ischemic cardiomyopathy in SSA, potentially driven by untreated infections and underdiagnosed familial cases (34). More recently results released for the IMHOTEP Pilot Study, which is the largest cardiomyopathy cohort in Africa, showed that DCM is the dominant cardiomyopathy seen in SSA, with 72% of participants being affected by DCM, the majority of which are black Africans (36). Thus, highlighting the necessity for more research to be conducted into underrepresented SSA populations.

1.1.2 Arrhythmogenic Cardiomyopathy (ACM)

Arrhythmogenic cardiomyopathy (ACM) is characterized by the progressive replacement of myocardial tissue with fibrofatty tissue, primarily affecting the right ventricle but involvement of the left ventricle has been observed in some cases (37). This structural abnormality predisposes individuals to arrhythmias, which can lead to sudden cardiac death (SCD), particularly in young athletes (38).

1.1.2.1 Prevalence, Incidence, and Mortality

The estimated prevalence of ACM is 1:1000 to 1:5000 in the general population (39). It accounts for approximately 10%–20% of sudden cardiac deaths in individuals under 35 years of age, highlighting its significance among young athletes (37). ACM has a progressive course, and the mortality rate remains high due to fatal arrhythmias if not appropriately managed. Advances in the use of implantable cardioverter-defibrillators (ICDs) have improved survival in high-risk patients (40).

1.1.2.2 Aetiology

The aetiology of ACM is multifactorial, with a strong genetic predisposition. It has also been linked to environmental and lifestyle factors, such as strenuous physical activity, which exacerbates disease progression by causing mechanical stress on the heart. Inflammation and myocarditis are additional contributors to disease onset and progression (15).

1.1.2.3 Diagnosis

The diagnosis of ACM is multiparametric due to the complicated clinical presentation of this disease. There are six parameters set out by the international document for diagnosis of ACM: Morpho-functional abnormalities, structural myocyte abnormalities, ECG repolarization and depolarization abnormalities, ventricular arrhythmias and lastly family history and molecular genetics (41). In 2020, international experts proposed the Padua Criteria, which had been worked on by a multidisciplinary team over three decades (11). This became the updated diagnostic International Task Force (TF) Criteria, which addressed disparities between the 1994 and 2010 TF criteria and provided diagnostic criteria for the wide range of ACM phenotypes (37)(Appendix A), has shown below in Figure 4 (Adapted from (37)).

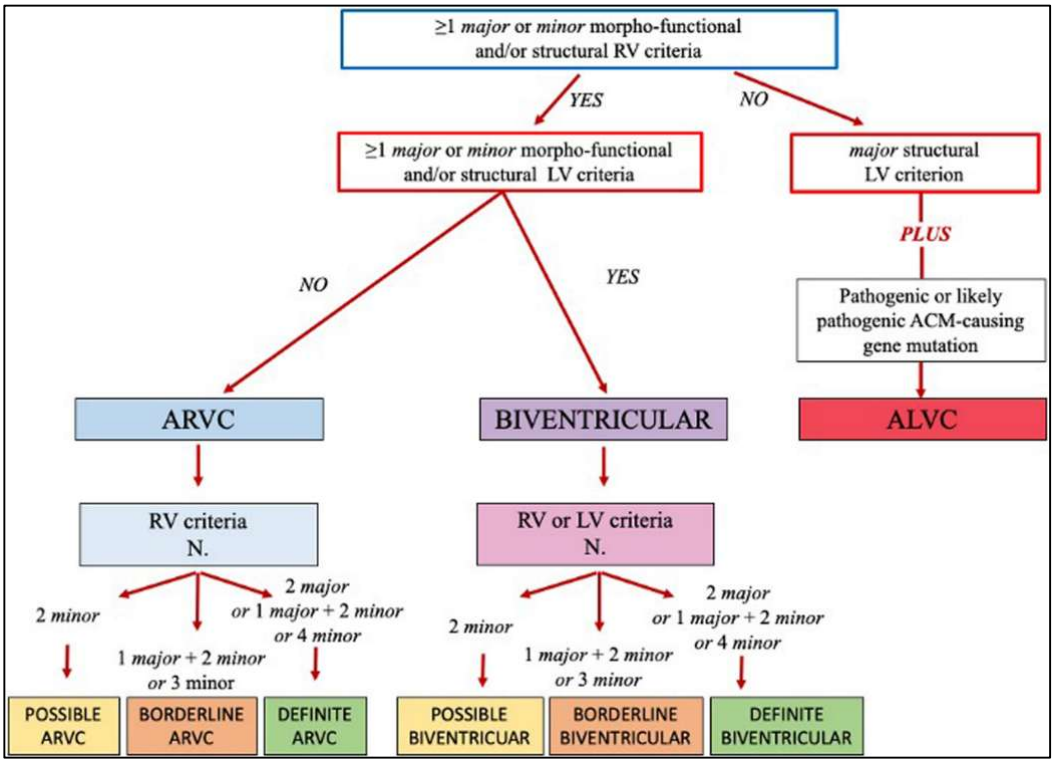


Figure 4: Diagnosis of ACM must have the criterion of either major or minor category I (morphological abnormalities) or II (structural abnormalities) to be fulfilled (37)

1.1.2.4 Genetics

Genetic variants associated with ACM are most often inherited in an autosomal dominant pattern with variable penetrance. ACM is a genetically heterogenous disorder with multiple genes linked to the condition. The most implicated genes are genes encoding desmosomal proteins, which play a role in cell-to-cell adhesion such as *PKP2* (plakophilin-2), *DSP* (desmoplakin), *JUP* (junction plakoglobin), and *DSG2* (desmoglein-2) (42). These variants

have been shown to disrupt the integrity of the myocardial tissue, making it prone to injury and fibrofatty replacement (Figure 5 (43) and Figure 6 (44)). Over the past few decades, with advancements as well as cost reduction in imaging techniques and next generation sequencing, researchers have identified non-desmosome genes can cause ACM. These genes include *FLNC (*filamin C*)(45), LMNA* (lamin A/C)(46), *PLN* (phospholamban)(47), *TMEM43* (transmembrane protein 43)(48), *DES* (desmin)(49), *TTN* (titin)(50), *SCN5A (51)* and *RBM20* (RNA binding motif protein 20)(52).

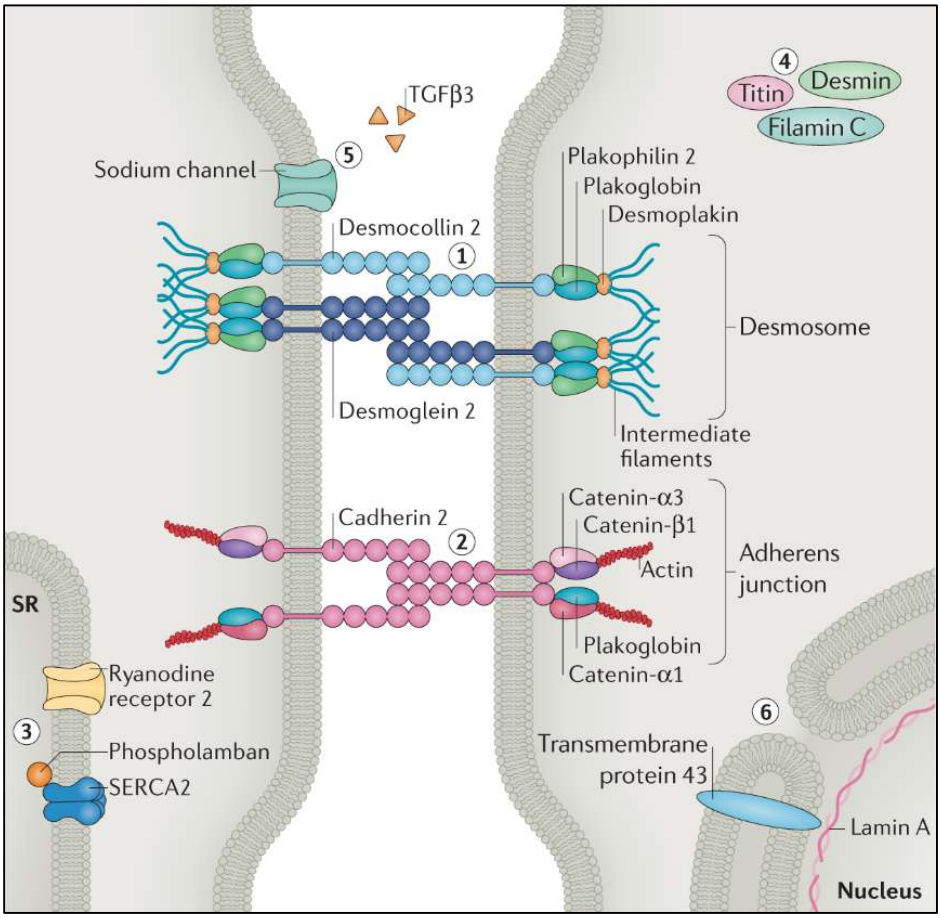


Figure 5: Schematic view of ACM-associated genes and the related cellular components. The genes and related cellular components are labelled as follows (1) Components of the desmosome, including Desmocollin-2 (DSC2), Desmoglein-2 (DSG2), Junction plakoglobin (JUP), Plakophilin-2 (PKP2) and Desmoplakin (DSP). (2) Components of the adherens junction, including Cadherin-2 (CDH2) and Catenin alpha-3 (CTNNA3). (3) Contributors to calcium handling, including Phospholamban (PLN) and Ryanodine receptor 2 (RYR2) located in the membrane of the sarcoplasmic reticulum (SR). (4) Intracellular structural proteins, including Desmin (DES), Titin (TTN) and Filamin-C (FLNC). (5) The Sodium channel Sodium voltage-gated channel subunit 5 (SCN5A) and Transforming growth factor beta-3 (TGFβ3). (6) Nuclear envelope proteins Transmembrane protein 43 (TMEM43) and Lamin A/C (LMNA). Sarcoplasmic/endoplasmic reticulum calcium ATPase 2 (SERCA2/ATP2A2) (43)

Gene	% of ARVC Caused by Pathogenic Variants in Gene	ClinGen Gene Validity Classification	Comment	Selected Allelic Disorders
PKP2	34%-74%	Definitive	May be more likely to cause VT	Brugada syndrome
DSG2	5%-26%	Definitive	Founder variant in population from East Asia	DCM
DSP	2%-39%	Definitive	• LV>RV involvement compared to *PKP2* • Can present as myocarditis or sarcoidosis-like	DCM
DSC2	1%-2%	Definitive	• ARVC can be nonsyndromic or associated with mild palmoplantar keratoderma & woolly hair. • Founder variant in Hutterite population	
JUP	0.5%-2%	Clinical Actionability	• Heterozygous variant causes nonsyndromic ARVC. • Biallelic variants cause Naxos disease (ARVC w/palmoplantar keratoderma & peculiar woolly hair).	
TMEM43	Rare	Definitive	Founder variant in population from Newfoundland, Canada	Auditory neuropathy; *TMEM43*-related myopathies
DES	Rare	Moderate	Founder variant in Dutch population	DCM: conduction system disease; myofibrillar myopathy; Kaiser-type neurogenic scapuloperoneal syndrome
PLN	Rare	Moderate		DCM HCM
Unknown	~40%			

Figure 6: Main genes associated with ACM and their diagnostic yield (44)

1.1.2.5 African Context

Not much is known about the prevalence and genetic patterns of ACM in African populations as only a few studies of small sizes have been conducted (14, 53, 54). There is limited data from sub-Saharan Africa, which is believed to be a direct result of suboptimal care. Approximately 15% of countries in SSA lack a trained cardiologist, one-third do not have pacemaker centres and over 50% lack coronary catheterisation laboratories (53). Only South Africa and a handful of North African countries can provide patients with complete cardiac services (55). The recent publication of the preliminary IMHOTEP results, which consisted of predominately South African participants, showed that 8.5% of participants are affected by ACM, with a higher prevalence in European patients (56). The increased proportion of European patients is hypothesised to stem directly from the historical disparities in access to health care (54). Further research is still needed to understand the disease's burden and clinical manifestations in African settings.

1.1.3 Hypertrophic Cardiomyopathy (HCM)

Hypertrophic cardiomyopathy is defined as left ventricular wall thickening with unknown secondary causes as well as a non-dilated left ventricle (57, 58). The thickening of the heart muscle can obstruct blood flow and impair relaxation, leading to diastolic dysfunction. HCM is a genetically and phenotypically complex heart disease however, early detection and treatment are vital to ensure patients' longevity (59).

1.1.3.1 1Prevalence, Incidence, and Mortality

HCM affects approximately 1:200 - 1:500 individuals globally, though many cases remain undiagnosed (59). The annual incidence of SCD in HCM patients is estimated to be 1% in high-risk populations, especially among athletes (60). With proper management, survival rates have improved; however, complications such as heart failure and arrhythmias remain significant contributors to morbidity (61, 62).

1.1.3.2 Aetiology

HCM is primarily a genetic disorder, although some cases can occur secondary to hypertension or inherited metabolic disorders. Other triggers include myocardial stress and systemic conditions, such as amyloidosis (58).

1.1.3.3 Diagnosis

Patients suspected to have HCM undergo several clinical assessments such as extensive personal and family history of SCD, HCM and HF, as well as a 12-lead ECG, echocardiogram assessment and a physical assessment (58). Patients often present with variable phenotype however, ventricular hypertrophy is the standardized manifestation of HCM and is often accompanied by myocyte hypertrophy and disarray, interstitial fibrosis and impaired ventricular filling. It is also common for patients with HCM to be asymptomatic. The frequently noted clinical manifestations include diastolic ventricular dysfunction, obstruction to the left ventricular outflow tract (LVOT), imbalanced supply and demand of myocardial oxygen supply and cardiac arrhythmias (57).

1.1.3.4 Genetics

HCM is most often inherited in an autosomal dominant pattern, with mutations in sarcomeric genes. HCM diagnostic gene panels (63, 64), vary widely in size, encompassing between 30 and 100 genes, with prominent inclusions such as *MYH7* (myosin heavy chain), *MYBPC3* (myosin-binding protein C) (65), *TNNT2* (cardiac troponin T), *TNNI3* (cardiac troponin I), *TPM1* (α-tropomyopsin) (66), *ACATC1* (Cardiac α-actin), *MYL2* (regulatory myosin light chain), *MYL3* (essential myosin light chain)(67) and *CSRP3* (cysteine and glycine-rich

protein 3)(68), Figure 7 (58) and Figure 8 (69). The most frequently detected variants associated with HCM are in *MYH7* and *MYBPC3*, these variants lead to abnormal myocardial thickening and disorganized cardiac muscle fibres (65).

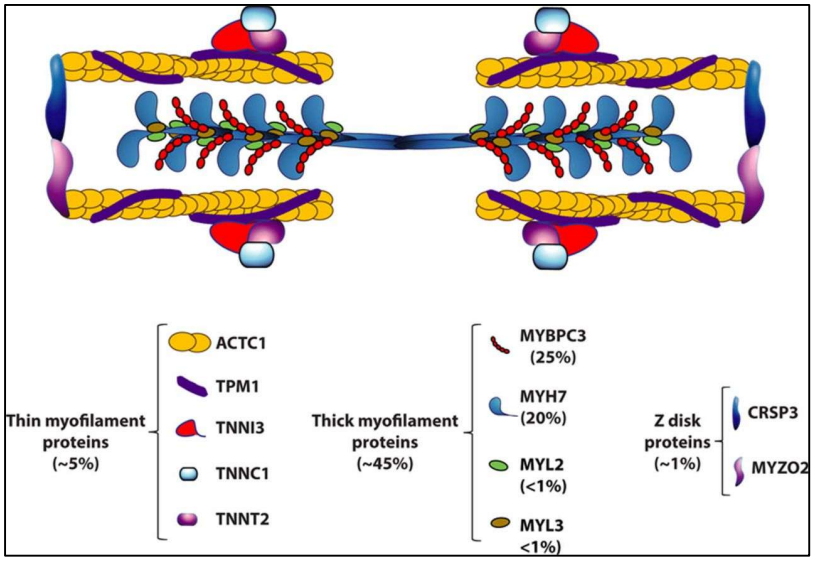


Figure 7: A schematic representation of the genes causing HCM. The key below the image represents the genes causing HCM and the location of the subsequent proteins. The percentages (%) are their population frequencies (59)

Location within the cell/function	Protein	Gene	Frequency within genotype-positive individuals	Level of evidence according to ClinGen and mode of inheritance
Sarcomere (contractile) proteins	Myosin-binding protein C	*MYBPC3*	40%–50%	Definitive (AD)
	Beta-myosin heavy chain	*MYH7*	35%–40%	Definitive (AD)
	Troponin T	*TNNT2*	7%–15%	Definitive (AD)
	Troponin I	*TNNI3*	5%	Definitive (AD)
	Tropomyosin	*TPM1*	3%	Definitive (AD)
	Regulatory myosin light chain	*MYL2*	1%–2%	Definitive (AD)
	Essential myosin light chain	*MYL3*	1%	Definitive (AD)
	Actin	*ACTC1*	1%	Definitive (AD)
	Troponin C	*TNNC1*	<1%	Moderate (AD)
Z-Disc proteins and other sarcomere associated	Alpha-actinin-2	*ACTN2*	<1%	Moderate (AD)
	Alpha-protein kinase 3	*ALPK3*	~2%	Definitive (AR). Recent evidence for AD inheritance for truncating variants
	Formin Homology 2 Domain Containing 3	*FHOD3*	1%–2%	Not curated (AD)
	Muscle LIM protein	*CSRP3*	<1%	Moderate (AD)
	Tripartite Motif Containing 63	*TRIM63*	Unknown	Moderate (AR)
	Filamin C	*FLNC*	<1%	Not curated for (isolated) HCM. Recent evidence for AD inheritance for missense variants
	Four-and-a-half LIM domain protein 1	*FHL1*	<1%	Not curated for (isolated) HCM (X-linked)
Calcium handling proteins	Phospholamban	*PLN*	<1%	Definitive (AD)
	Junctophilin 2	*JPH2*	Unknown	Moderate (AD)

Figure 8: Main genes associated with HCM and their diagnostic yield (70)

1.1.3.5 African Context

There have been several African studies published on HCM (70, 71). These studies highlight the misconception that HCM is rare in Black Africans, rather than the suspected decreased prevalence due to these patients being underrecognized, misdiagnosed and under-referred for therapies (71). Data published in 2021 shows that Black patients from SSA yield lower positive results from the known gene list associated with HCM, suggesting the underlying genetic mechanisms differ (72). Further research as well as improved interventions are vital to decrease disparities in HCM patients specifically Black African patients.

1.1.4 Restrictive Cardiomyopathy (RCM)

Restrictive cardiomyopathy (RCM) is defined as a myocardial disease of diastolic dysfunction due to increased myocardial stiffness (73), extremely enlarged atria and normal sized ventricles (74). Restrictive cardiomyopathy has a wide phenotypic range and is regarded as the least common cardiomyopathy. RCM can result from genetic or acquired predispositions or a combination of both. It is often associated with a coexisting persistent restrictive pathophysiology, diastolic dysfunction, non-dilated ventricles as well as atrial dilation regardless of ventricular wall thickness and systolic function (75).

1.1.4.1 Prevalence, Incidence, and Mortality

RCM is rare, accounting for only 2%–5% of cardiomyopathies. It is more commonly diagnosed in older adults, though it also affects children (76). The prognosis of RCM is poor, with high mortality due to progressive heart failure and arrhythmias if left untreated (74). Frequently, the only option for a patients' long-term survival is a heart transplant (62).

1.1.4.2 Aetiology

RCM can result from various underlying conditions, including infiltrative diseases like amyloidosis and sarcoidosis, or systemic disorders such as hemochromatosis (57) however, it can also be a familial disease caused by a single or multiple genetic variants (77). Idiopathic forms of RCM are also common.

1.1.4.3 Diagnosis

The diagnosis of RCM is complicated as there are four main causes of RCM which are described in Figure 9 (78).

Cause	Examples
Endomyocardial disorders	Carcinoid tumours Endocardial fibroelastosis Endomyocardial fibrosis (EMF) Hypereosinophilic syndrome (including Löffler syndrome) Metastatic cancer Medications (eg, hydroxychloroquine, ergotamine, methysergide) Radiation
Infiltrative disorders	Amyloidosis Sarcoidosis
Non-infiltrative disorders	Idiopathic restrictive cardiomyopathy Pseudoxanthoma elasticum Systemic sclerosis
Storage diseases	Fabry disease Glycogen storage diseases Hemochromatosis

Figure 9: The main causes and examples of what may contribute to the causes of RCM (79)

Several methods such as echocardiography (2-dimensional) and Doppler are used to determine diastolic dysfunction (73). Cardiac Magnetic resonance is also used in the diagnostic procedure which should be done on an individual basis, after a full patient clinical history is taken to ensure the procedure is safe. In rare cases endomyocardial biopsies are used to diagnose patients, however this is an invasive procedure and is normally avoided (73, 75).

1.1.4.4 1Genetics

Given that RCM is a rare cardiomyopathy, there have only been 19 genes identified as causative in patients with RCM (74), with significant genetic overlap with other cardiomyopathies. Most variants are inherited in an autosomal dominant mode of inheritance however, there have been reports of a recessive inheritance pattern (79) as well as X-linked inheritance (80). Familial RCM cases are linked to mutations in sarcomeric and cytoskeletal genes, including *TNNI3* (troponin I), *TNNT2* (troponin T), and *DES* (desmin) (73). The genetic overlap with other cardiomyopathies, such as HCM and DCM, complicates the diagnosis.

1.1.4.5 African Context

Not much is known about RCM in African populations. Diagnostic tools such as endomyocardial biopsy are rarely available, limiting comprehensive assessments. Through the IMHOTEP registry of South Africa, the largest cardiomyopathy registry in Africa, they reported that RCM was diagnosed in only 5.95% of participants of the total cardiomyopathy cohort, with most of the cases being recruited from Mozambique, making it the least frequent cardiomyopathy in SSA (56).

1.1.5 Left Ventricular Non-Compaction Cardiomyopathy (LVCN)

Left ventricular non-compaction is a heterogenous cardiomyopathy characterized by abnormal trabeculation and deep intertrabecular recesses in the left ventricle, resulting from incomplete myocardial development during embryogenesis (81, 82). LVNC can present with heart failure, arrhythmias, or thromboembolic events highlighting the heterogeneity of LVNC (83). There is controversy surrounding LVNC as in 2006, The American Heart Association (AHA) classified LVNC as a primary genetic cardiomyopathy while in 2008, the European Heart Association defined it as an unclassified cardiomyopathy (10). As of 2023 LVNC remained poorly defined, particularly when considering adults (84).

1.1.5.1 Prevalence, Incidence, and Mortality

The prevalence of LVNC is estimated at 0.12%–0.3% in the general population, though it may be underreported due to misdiagnosis (85). Studies conducted in Australia and the USA showed LVNC was the 3rd most frequent cardiomyopathy in children (86, 87). In adults, the prevalence of LVNC varied greatly across different populations which could be attributed to differential diagnostic techniques and therefore the characteristics of the cohort (83). LVNC can lead to severe heart failure, arrhythmias, or embolic stroke, contributing to significant morbidity and mortality (88).

1.1.5.2 Diagnosis

LVNC can manifest as HF, arrhythmias, and thromboembolic events (88). Historically LVNC was diagnosed using echocardiography, which was the only diagnostic technique (83). With

recent advancements in imaging techniques and modalities of diagnosis, it was recommended that multiple imaging techniques should be used to accurately diagnose a patient with LVNC (89). There are three main techniques currently used to confirm the diagnosis of LVNC: two-dimensional echocardiography, cardiac magnetic resonance imaging (CMR) and cardiac computed tomography (CT) (89).

1.1.5.3 Aetiology

LVNC is thought to result from abnormal embryonic development of the myocardium, leaving portions of the left ventricle non-compacted and spongy (88). The aetiology of LVNC remains largely unknown, however, it does have a large genetic component (90).

1.1.5.4 Genetics

LVNC is associated with mutations in genes involved in cardiac development and sarcomeric function, including *MYBPC3, TPM1, ACTC1, TNNT2, TNNI3, MYL2, MYL3*, *TTN* (titin), *MYH7*, and *TAZ* (tafazzin) (91, 92) (Figure 10 (93)). Inherited forms follow autosomal dominant, autosomal recessive, or X-linked patterns, depending on the gene involved (90).

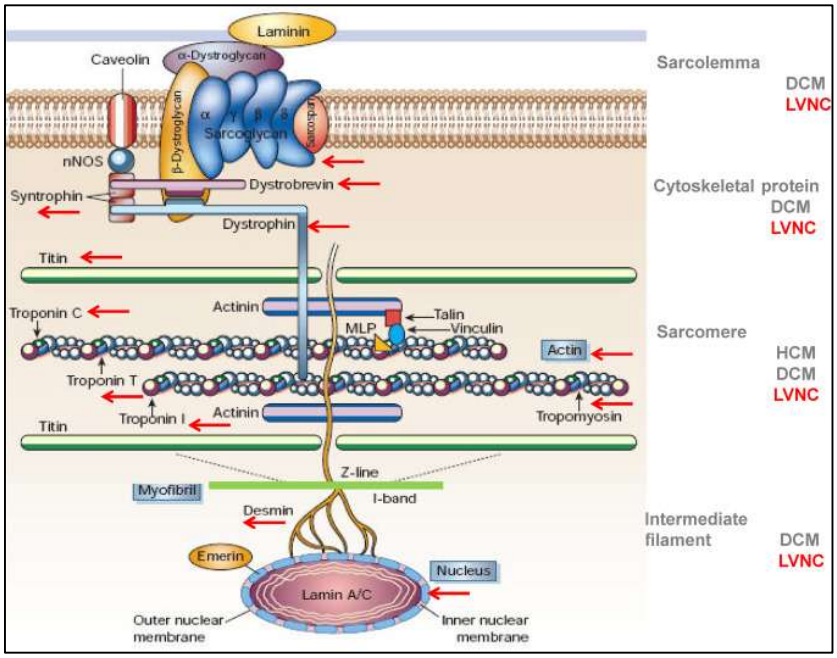


Figure 10: Schematic overview of the proteins in cardiomyocytes. Proteins and pathways involved in the development of cardiomyopathies. The cytoskeletal protein genes, dystrophin, and sarcoglycan in the sarcolemma, and the intermediate filament protein genes desmin and lamin A/C mutations are identified in DCM and LVNC. Sarcomere genes cause HCM, LVNC, or DCM (94)

1.1.5.5 African Context

Not much is known about LVNC in African populations, and the condition is likely underdiagnosed (14). A recent publication in Nigeria highlighted that LVNC is not uncommon in the SSA region, however, due to limited access to health care and technologies patients remain under or misdiagnosed (94).

1.2 The Founder Effect

Human migration, religion and culture have played a significant role in shaping population genetics. Present-day populations inherit variants from their ancestors, evolutionary factors such as migration, population admixture, selection pressure, effective population size as well as recombination rates influence the inheritance of these variants (95).

Founder variants, a subset of familial variants originating from a common ancestor, can offer valuable insights into disease mechanisms, influence disease expression, penetrance, and population-specific risk patterns and offer unique opportunities for genetic screening and preventative strategies within affected communities (96). Founder effects can influence the frequency of variants and can be affected by genetic drift and the bottleneck effect (97). A genetic bottleneck effect occurs when a large population is diminished due to environmental disasters, habitat loss, or predation resulting in a significant alternation in the variant frequencies (98). While the founder effect occurs when a small subset of a much larger population becomes isolated and forms a new, founding population. The new population suffers from a reduced gene pool due to endogamy, resulting in unfavourable variants being observed at higher frequencies than expected (99), as shown in Figure 11 (100).These are known as founder variants.

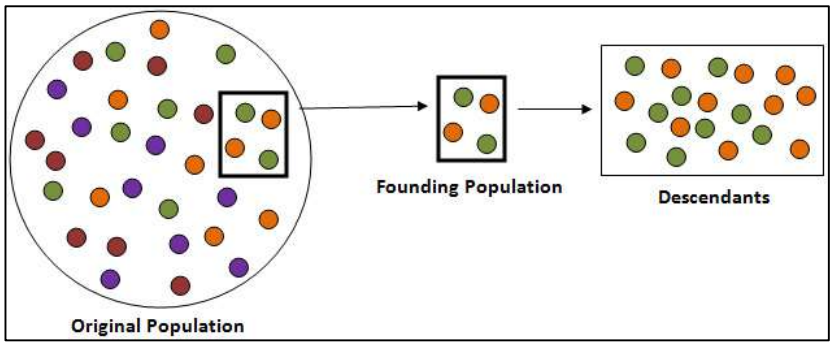


Figure 11: Visualization on the founder effect (101)

Founder variants offer valuable insights into human genetics and disease into (1) *Disease Prevalence and Risk*: Founder variants can significantly influence the prevalence of specific genetic diseases within a population. By studying these variants, we can better understand why certain conditions are more common in particular groups and identify individuals at higher risk (101, 102). (2) *Disease Mechanisms*: Founder variants can help pinpoint the specific genes and variants responsible for diseases. This knowledge can shed light on disease mechanisms and pathways, paving the way for targeted therapies (97). (3) *Genetic Screening and Prevention*: Identifying founder variants allows for targeted genetic screening programs within affected populations. Early detection can enable preventative measures and timely interventions, improving patient outcomes (97, 103). (4) *Population History and Migration*: Founder variants can provide clues about a population's history, migration patterns, and ancestral origins. They serve as genetic markers that reflect past demographic events and admixture (104). (5) *Pharmacogenomics*: Founder variants can influence drug response and efficacy. Understanding these variations can lead to personalized medicine approaches, tailoring treatments based on an individual's genetic makeup (97).

Examples of diseases caused by founder variants are the Ashkenazi Jewish population, where multiple founder variants have been shown to cause rare genetic diseases such as heritable breast cancer (105), Tay-Sachs disorder (106) as well as cardiomyopathies such as DCM and ACM (107). Similarly, the French-Canadian population of Quebec has founder variants in heritable breast cancer (108) and cardiomyopathies (DCM and HCM) (109). Thus, identifying founder variants would also be essential in risk stratification as it would allow for early intervention of heritable genetic diseases.

1.2.1 Founder Variants in Cardiomyopathies

One fascinating aspect of cardiomyopathy is the contribution of founder effects. Over the past three decades, substantial progress has been made in elucidating the genetic basis of cardiomyopathies with studies identifying patterns within specific populations that exhibit similar phenotypic presentations and share genetic variants or founder variants, which have subsequently been associated with various forms of cardiomyopathy (shown in Table 1).

Table 1: Founder Variants identified in the various Cardiomyopathies

CARDIOMYOPATHY	VARIANT	POPULATION	REFERENCE
ACM	*PKP2* c.1211dup	Dutch	https://doi.org/10.1007/s12471-023-01791-2 (110)
ACM	*TMEM43* c.1073C>T	UK and Canada	https://doi.org/10.1159/000517265 (48)
ACM	*DSC2* c.363G>T	Italian	https://doi.org/10.1093/cvr/cvu091.24 (111)
ACM	*PLN* c.40_42delAGA	Dutch	https://doi.org/10.1007/978-90-368-0705-0_11 (112)
ACM	*DSG2* c.1592T>G	East Asians	https://doi.org/10.1038/s41439-022-00206-9 (113)
ACM	*PKP2* c.775_776insG	Spain	https://www.preprints.org/manuscript/202305.0784 (114)
ACM	*DES* c.38C>T and *DES* c.1024A>G	Dutch	http://dx.doi.org/10.1007/s12471-011-0233-y (115)
ACM	*PKP2* c.235C>T	Dutch	https://doi.org/10.1007/s12471-010-0839-5 (116)
ACM	*DSC2* c.1660C>T	Hutterite (Canada)	https://doi.org/10.1161/CIRCGENETICS.113.000097 (111)
ACM (NAXOS)	*JUP* c.902A>G	French/Canadians	https://doi.org/10.1111/cge.12971 (117)
DCM	*FLNC* c.6864_6867dup	Dutch	https://doi.org/10.1016/j.hrthm.2023.08.003 (118)
DCM	*NKX2*-5 c.435C>G	USA	https://doi.org/10.1177/19253621241264857 (119)
DCM	*LMNA* c.1961dup	French	https://doi.org/10.1530/eje-21-0282 (120)
DCM	*MYH7* c.5754C >G	Dutch	https://doi.org/10.1007/s12471-017-1037-5 (121)
DCM	*TTN* c.59926+1G>A	Dutch	https://doi.org/10.1002/ejhf.1030 (122)
DCM	*TNNT2* c.629_631delAGA	Dutch	https://doi.org/10.1007/BF03091819 (123)
DCM	*PLN* c.25C>T	South Africa	https://doi.org/10.1038/srep22235 (124)
DCM	*BAG3* c.925C>T	French/Canadians	https://doi.org/10.1016/j.cjca.2014.09.030 (109)
DCM/ACM	*FLNC* c.3791-1G>C	Ashkenazi Jews	https://doi.org/10.1016/j.ijcard.2020.08.052 (107)
DCM/HCM	*FKTN* c.1371_1381dupTATCCAGTTAT	Qatar	https://doi.org/10.1002/mgg3.1709 (125)
HCM	*MYPBC3* c.927-2A>G	Iceland	https://doi.org/10.1161/CIRCULATIONAHA.114.011207 (126)
HCM	*MYPBC3* c.2373insG	Dutch	https://doi-org.ezproxy.uct.ac.za/10.1016/S0195-668X(03)00466-4 (127)
HCM	*MYBPC3* c.2149–1G>A	Spanish	https://doi.org/10.1136/openhrt-2021-001789 (128)
HCM	*MYH7* c.5135G>A	International cohort	https://doi.org/10.1007/s12471-023-01798-9 (129)
HCM	*FHOD3* c.1646+2T>C	Balkans	https://doi.org/10.1371/journal.pone.0294969 (130)
HCM	*TGFB3* c.787G>C	Belgian	https://doi.org/10.3389/fgene.2023.1251675 (131)

HCM	*MYBPC3* c.913_914del	Italian/Slovenian	https://doi.org/10.1007/s12265-024-10551-5 (132)
HCM	*MYBPC3* c.3330+2T > G	Armish/Swiss	https://doi.org/10.1016/j.ejmg.2022.104627 (133)
HCM	*TNNI3 c.574C>T*	Dutch	https://doi.org/10.1007/s12471-011-0135-z (134)
HCM	*TNNI3* c.433C>T	Dutch	https://doi.org/10.1007/s12471-011-0135-z (134)
HCM	*TPM1* c.523G>A	Finnish	https://doi.org/10.3109/07853890.2012.671534 (135)
HCM	*MYBPC3* c.3181C>T	Finnish	https://doi.org/10.3109/07853890.2012.671534 (135)
HCM	*βMHC* c.1207C>T	South Africa	https://doi.org/10.1086/302623 (70)
HCM	*βMHC* c.2389G>A	South Africa	https://doi.org/10.1086/302623 (70)
HCM	*cTnT* c.304C>T	South Africa	https://doi.org/10.1086/302623 (70)
HCM /RCM	*TNNI3* c.497C>T	Dutch	https://doi.org/10.1007/s12471-011-0135-z (134)

ACM: Arrhythmogenic Cardiomyopathy, DCM: Dilated Cardiomyopathy, HCM: Hypertrophic Cardiomyopathy, RCM: Restrictive Cardiomyopathy

Of the 36 founder variants that are associated with various cardiomyopathies, 17 were identified in HCM with overlap between DCM (*FKTN* c.1371_1381dupTATCCAGTTAT) and RCM (*TNNI3* p. Ser166Phe), 10 were associated with DCM and 11 associated with ACM. It should also be noted that 12 of these variants were identified in the Dutch population.

South Africa, aptly dubbed the "rainbow nation," boasts a rich tapestry of cultures, languages, and ethnicities woven together through a complex history of colonization. This unique demographic landscape has left an indelible mark on the nation's genetic makeup, particularly through founder effects introduced by early European settlers, including the Dutch and French Huguenots. These founder variants, passed down through generations, have become embedded within the genetic fabric of various South African populations, offering a powerful lens through which to explore the interplay of history, genetics, and disease.

Another distinct population is in India, which is considered a treasure for geneticists due to the large amount of human diversity, with more than 4500 distinct castes, tribes and religious groups (136). Research has shown that there are India-specific variations associated with several common diseases found in Indian populations. An investigation into the underlying genetics of 281 individuals identified 1,284 novel exonic variants, with approximately 29.2% of the variants being private to a single population (137).Many of these novel variants are likely a result of strict endogamy and founder effects, which have contributed to the genetic isolation of these groups. These private variants were largely missense and rare, with minor allele frequencies (MAF) typically between 2-5%, suggesting that they may be enriched for rare, population-specific traits. Importantly, 23 of the novel variants were predicted to be potentially deleterious, affecting protein stability and possibly playing a role in recessive disease expression (137). These findings highlight the wealth of knowledge which can be uncovered in underrepresented and isolated populations, as these variants may not be present in global datasets that are largely Eurocentric. The high level of homozygosity observed in these groups, driven by the founder effects and consanguinity, further facilitates the identification of disease-associated variants. (136, 137). Additionally, unique pharmacogenomic profiles were identified, revealing significant variability in drug response across different groups. This emphasizes the need for expanding genetic research in diverse, understudied populations to improve precision medicine efforts, disease risk predictions, and tailored treatment strategies

1.2.2 Founder Effects in Present-Day South Africa

In 1651 the Dutch East India Company (VOC) established a presence in the Cape of Good Hope and Jan van Riebeeck was given the command of the Dutch expedition to establish and fortify a provisioning station in the Cape to supply trade ships with fresh goods and water (138, 139).

The first white settlers arrived with Jan van Riebeeck in Table Bay and disembarked from three Dutch ships in 1652 (140), Figure 12. This led to the establishment of the Cape Colony which was governed by the VOC from 1652 to 1795. The Dutch Reformed Church (DRC/NGK) was established during this period and record keeping of all congregants began (141, 142). Baptisms, marriages, deaths and immigration of individuals were well documented and kept at churches.

During the period of the VOC rule, the importation of slaves to the Cape Colony began in 1654. Many of the slaves were later freed and married either the Dutch settlers or the native Hottentots/Khoi Khoi (143).

Figure 12: Map of Africa from the 15th – 17th century, with key. This map shows the movement of the French, English and Dutch to Africa. It also shows the trade routes with their main ports and therefore bases. We also see the slaves and their origins (https://gifex.com/detail-en/2010-01-12-11697/Africa-between-the-15th-and-17th-century.html)

The French were the next to settle after the Dutch. Following the Edict of Fontainebleau, also known as Revocation, in 1685 France experienced religious unrest and the Protestants were prohibited from practising their religion. As a result, 200 members of the French Calvinist

Protestants, also known as French Huguenots, fled their country and settled in the Cape Colony between 1688 and 1689 (144). They came to the Cape Colony as refugees and were similar to the Dutch in their religious beliefs. This resulted in admixture and led to the formation of the modern-day Afrikaner population.

In 1795, Revolutionary France invaded the Dutch Republic, leading to the collapse of the Dutch government and the flight of Prince William V of Orange to England. From exile, he issued the Kew Letters, asking the British to protect Dutch colonies from French occupation. That same year, the British occupied the Cape Colony for the first time to prevent it from falling into French hands. However, tensions between the Dutch and British in the Cape Colony persisted due to religious and colonial differences, as s result admixture between the two populations remained limited. In 1803, the Cape was returned to the Dutch under the Batavian Republic following the Treaty of Amiens, but the British reoccupied it in 1806 as war resumed in Europe. Around 5,000 British settlers arrived in Port Elizabeth (Gqeberha) in the 1820s, and 163 German settlers arrived in Natal in 1848.The German Settlers admixed with the Afrikaner population; however, it was rare for the British and Afrikaner populations to admix (145).

In present-day South Africa, and specifically in the Afrikaner population, some disorders occur at remarkably high frequencies. This may indicate that these disorders were brought to South Africa by the Dutch, French or German settlers. Due to the high prevalence of consanguineous marriages in the late 17th and early 18th centuries, it is possible that the disease incidence was increased (104). There are several founder studies done on disorders in the Afrikaner population which have shown founder effects are rife. Founder effects have been identified in diseases such as Parkinson's disease (104), pseudoxanthoma elasticum (PXE) (146), Long-QT syndrome (101), familial colonic polyposis (147), porphyria variegata (148), progressive familial heart block (102), Huntington's disease (HD) (149), osteogenesis imperfecta (150), schizophrenia (151), Fanconi's anaemia (152) and Bipolar disorder (153) through methods of genealogical tracing. While founder variants have been identified in pseudoxanthoma elasticum (PXE) (146), Long-QT syndrome (101), porphyria variegata (154), Huntington's disease (HD) (155) and Fanconi's anaemia (152) through molecular based studies.

We hypothesise that the high rate of founder effects and therefore founder variants in the Afrikaner population could be due to the founding population's geographic separation from

the original population, starting with the arrival of the Dutch in 1652. The genetic makeup of the current Afrikaner population is comprised of a third of Dutch, less than a third is German, and roughly a quarter is French, according to earlier descriptions of the admixture (156). Over the past three centuries, the Afrikaner population has spread throughout Africa, according to migration records, after initially settling in the Cape of Good Hope. The increased rates of intermarriage among the Afrikaner population today have helped to explain the genetic homogeneity we observe. Even though the percentage of European ancestry among modern Afrikaners varies, this population is still thought to be isolated and extremely homogeneous.

1.3 Rationale for the study

Previous work in our laboratory had identified the *PKP2* c.1162C>T variant as a founder variant (54) and a subsequent study (157) attempted to identify the founder couple or progenitors. However, the proposed minimum requirement of 10-12 unrelated probands to successfully trace a common founder or founder couple (104), was not met. At the time there were only 8 probands enrolled in the study and a single common founder or founder couple was not identified. Fortunately, new probands and their families were enrolled and were willing to participate in the continuation of the genealogical tracing of the *PKP2* variant. The total number of probands rose to 12 with the addition of four new families.

Recently, data from our parent IMHOTEP study, identified four additional recurring variants (*BAG3* c.925C>T, *LMNA* c.568C>T, *TTN* c.87624C>A and *TTN* c.95008C>T). These variant carriers appeared to also share the same ethnic origin which raised the possibility of additional founder variants.

We therefore hypothesised that these five recurring variants, that occurred in South African families of the same ethnicity and that are found in a confined geographic location, have a high likelihood of being founder variants.

1.3.1 Aim

To identify the possible founders of five variants identified in South African families with cardiomyopathy.

1.3.2 Objectives

I. To screen core family members of probands with one of the five variants of interest using high resolution melt (HRM) and validate using Sanger sequencing.

II. Design microsatellite markers and perform microsatellite analysis on the probands and their families with each of one of the five variants of interest.

III. Construction of haplotypes for the families and segregation patterns for the variants will be identified using the data collected from the microsatellite analysis.

IV. To genealogically trace potential founder variants found in the probands and their families using various vital statistical records to determine generational relationships.

Chapter 2: Methods and Materials

2.1 Study design:

This study was a research study where we investigate cases from The African Cardiomyopathy and Myocarditis Registry Program (IMHOTEP) (HREC 766/2014) for genetic mutations and perform genealogy tracing (Appendix B). It is a sub-study approved by the UCT Human Research Committee (HREC REF: 425/2023)(Appendix C) under the parent IMHOTEP registry.

Participants include probands and affected (or at risk) first-degree relatives, who have one of the disease-causing founder variants; *PKP2* c.1162C>T*, BAG3* c.925C>T, *LMNA* c.568C>T, *TTN* c.87624C>A and *TTN* c.95008C>T, which are associated with cardiomyopathies in South African families.

2.2 Study participants:

Informed consent for all participants of IMHOTEP was obtained by a genetic counsellor, registered nurse or clinician (Appendix D). There was no age restriction as cardiomyopathies affect adults and children. The inclusion of all age groups allows for researchers to determine disease severity and clinical outcomes which are affected by genetic and environmental factors. The inclusion and exclusion criteria of participants into IMHOTEP are outlined in Table 2.

Table 2: The inclusion and exclusion criteria of the IMHOTEP registry:

INCLUSION CRITERIA	EXCLUSION CRITERIA
• Patients reside in South Africa • Cases diagnosed at autopsy are included • Idiopathic cardiomyopathies • Familial cardiomyopathies • Neuromuscular disorders with cardiac involvement • Non-familial or secondary causes of cardiomyopathy • Myocarditis (infective/toxin/immune) • Human Immunodeficiency Virus (HIV) • Drugs/Toxins • Pregnancy • Endocrine • Nutritional • Obesity • Alcohol • Tachycardiomyopathy • Eosinophilic • Kawasaki disease • Tako Tsubo cardiomyopathy • Amyloidosis • Autoimmune • Endomyocardial fibrosis • Carcinoid heart disease • Radiation • Metastatic Cancer	• Systemic arterial hypertension (>160/100mmHg documented and confirmed at repeated measures) • Coronary artery disease (obstruction • >50% of the luminal diameter of a major branch) • Pericardial diseases • Congenital heart disease • Pulmonary disease with core pulmonale • Valvular heart disease, including Rheumatic Heart Disease

All research participants remained anonymous with numbers assigned through a two-level security system. Participants were coded with numbers, letters and abbreviations as well as a data matrix barcode which was used to decrypt the code for secure data storage.

2.3 Data collection:

All demographic and clinical data; date of birth, age of onset, sex, ethnicity, disease severity and contact information were obtained from the relevant clinical databases. The acquisition of all research data adhered to the standard operating procedures set out by IMHOTEP as well as guidelines set out by the University of Cape Town and Good Clinical Practice (GCP).

2.4 Research Workflow

After participants were enrolled and bloods were collected for DNA extraction, there were two phases of laboratory work which included (1) Genotyping, by screening participants via the variants (2) Haplotyping, which included developing haplotypes for each variant of interest and finally (3) Genealogy tracing where we attempt to find the progenitors. Below is a workflow diagram that is followed in Chapter 3 to 6.

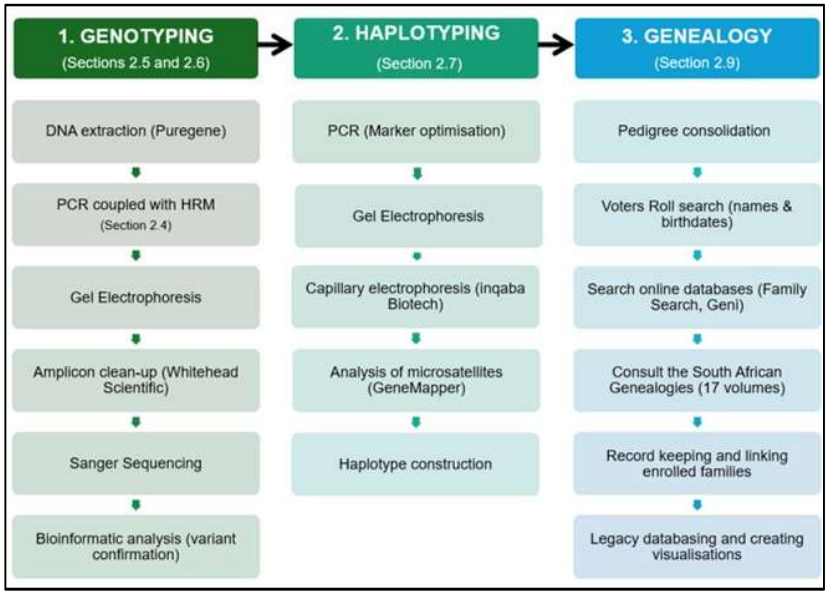


Figure 13: Research workflow showing the three phases of laboratory work, analysis and research done throughout this study

2.5 DNA extraction and storage:

2.5.1 Principles of DNA extraction

The extraction of DNA is a method used to purify DNA from a sample, such as blood. Therefore, DNA is separated from cell membranes, proteins and additional cellular components. This is achieved through lysing the cells and solubilizing DNA.

2.5.2 DNA extraction

Peripheral whole blood samples were previously collected from the participants via a venepuncture, where between 20-25ml and 5-25ml of blood was collected from adults and

children respectively. DNA was extracted using the PureGene Blood Kit protocol for white blood cell layers (Qiagen).

2.5.3 DNA quality and quantity checks

2.5.3.1 NanoDrop 2000™

Quality checks of the DNA was done prior to the DNA being used for experiments. The quality and quantity of genomic DNA (gDNA) was assessed through the calibrated NanoDrop 2000™ spectrophotometer v2.6 (Thermofisher). One microliter (uL) of homogenous raw or diluted DNA is used to test the samples concentrations and purity. The first ratio of absorbance we used to assess the purity of the DNA was the ratio between 260nm and 280nm (260/280), a ratio of ~1.8 is the universal standard for "pure" DNA, a ratio lower than 1.8 may indicate the presence of protein, phenol or other contaminants that absorb strongly at or around 280nm. Therefore, samples with a ratio between 1.6 – 2.0 were included. The second ratio is measured at 260nm and 230nm (260/230) and is a secondary measure of DNA purity, this ratio is usually higher than the 260/280 ratio and the universal standard for "pure" DNA is between 2.0 to 2.2. Any ratio below 2.0 indicates the presence of contaminants, which absorb at or around 230nm.

2.5.3.2 Agarose gels

The integrity of the DNA was checked using gel-electrophoresis. A one percent (1%) (w/v) agarose gel electrophoresis image is used to visualise if the samples were intact or damaged. Thus, 1g of agarose powder was weighed out and mixed with 100ml of 1X TBE buffer. It is heated until homogenized and left to cool to room temperature before it is poured into the casting tray and allowed to solidify. The gel is then placed in the gel tank and is emersed with 1X TBE. A ratio of 2uL:5uL, gDNA to loading dye (a combination of Gel Red and X5 Green GoTaq Flexi Buffer) is used to visualise the intactness. A 1 kbp (kilobase pair) ladder was used with the loading dye and the electrophoresis was programmed to run for 60 minutes at 110V. The gel was visualised using the iBright CL1500 imaging system (ThermoFisher Scientific).

2.5.4 DNA storage

After the gDNA quality and quantity were confirmed to be of good quality an aliquot of the raw gDNA was made into a working stock solution of 120ng/µl by diluting the samples with nuclease free water (BIO-TEK). The dilutions were used for targeted sequencing (if required) and microsatellite analysis. Extracted DNA samples were stored at the IMHOTEP biobank at -80°C.

2.6 Genetic Screening

2.6.1 Principles of Primers

Primers are short DNA oligonucleotide sequences, typically 18-25 base pairs in length. They are designed in pairs, with one serving as the forward primer and the other as the reverse primer, to bind to opposite strands of DNA and amplify a specific region of interest during a polymerase chain reaction (PCR) (158). The forward primer binds to the 3' end of the antisense strand (template strand) and directs amplification toward the 5' end, while the reverse primer binds to the 3' end of the sense strand and directs amplification toward the 5' end.

Primers require Taq polymerase (an enzyme), $MgCl_2$ (a catalyst), and deoxyribonucleotides (dNTPs: a mix of nucleotides; A, C, G, and T). During the extension phase, Taq polymerase synthesizes new DNA by adding dNTPs to the 3' end of each primer, elongating the complementary strand of DNA. This ensures precise and exponential amplification of the target region (158, 159).

The *PKP2*, *BAG3*, *LMNA* and *TTN* primers used in this study were designed by members of the Cardiovascular Genetics Laboratory at the Cape Heart Institute (CHI), Cardiovascular group and are presented in Table 3 to Table 6.

Table 3: Primer sequences used for the amplification of PKP2 exons 4 and 11.

***PKP2* Exon**	**Expected size (bp)**	**Ta, Annealing Temperature (°C)**	**Primer Name and Strand**	**Sequence (5'-3')**
4	291	55	*PKP2* Exon 4F	AGT ATT CGC TGA GTC GTC TCT
			PKP2 Exon 4R	GCA AAG TCA CCA TAA TAG AAG
11	219	55	*PKP2*_Exon11.2_New_F	ACA TCA GTG GCT CAG ACA G
			*PKP2*_Exon11.2_New_R	CCC ATT TCC AGT GCA TCT TTG TG

PKP2 Exon 4 and 11 were both screened, F= forward primer, R= reverse primer. (Adapted from: (157, 160)).

Table 4: Primer sequences used for the amplification of BAG3 exon 4

***BAG3* Exon**	**Expected size (bp)**	**Ta, Annealing Temperature (°C)**	**Primer Name and Strand**	**Sequence (5'-3')**
4	331	55	*BAG3* Exon 4F	TGG AAG CCT GAC TAA TGA TTG GG
			BAG3 Exon 4R	GCG GAT CAC TTG AAT TGG GAT G

Table 5: Primer sequences used for the amplification of LMNA exon 3

***LMNA* Exon**	**Expected size (bp)**	**Ta, Annealing Temperature (°C)**	**Primer Name and Strand**	**Sequence (5'-3')**
3	304	55	*LMNA* Exon 3F	CAG CAG CCC ACC TCT CAG C
			LMNA Exon 3R	GGA CTC AGG AAG GCG AGC

Table 6: Primer sequences used for the amplification of TTN exons 342 and 368.

***TTN* Exon**	**Expected size (bp)**	**Ta, Annealing Temperature (°C)**	**Primer Name and Strand**	**Sequence (5'-3')**
342 (*TTN* c.95008C>T)	368	55	*TTN* Exon 342F	TTG CCT ACT GAC AGC TCC AC
			TTN Exon 342R	GAT CAT TCT TTC TGC CCA CAC TG
TTN c.87624C>A	367	55	*TTN* Exon 368F	GAA ACA AGT ACT GCA GTG TGG AC
			TTN Exon 368R	ATG GTG TAG ATG GTG GTC CAG

2.6.2 Principles of Polymerase Chain Reaction

Polymerase Chain Reaction (PCR) is a technique used in molecular biology to amplify a specific region of DNA, generating millions to billions of copies. The key principles include DNA denaturation, amplification and extension (161). The denaturation causes the DNA to become single stranded by breaking the hydrogen bonds between complementary bases using high heat, 94-98°C.The annealing step allows primer binding. Thus, the temperature is lowered to around 50-65°C. Two primers, each complementary to a specific region flanking the target DNA, bind to the single-stranded templates. The extension step raises the temperature to 72°C, the optimal working temperature for Taq polymerase, a heat-stable DNA polymerase. This enzyme adds nucleotides (dNTPs) to the 3' end of each primer, extending the DNA strand in the 5' to 3' direction. These three steps are repeated for 25-40 cycles. Each cycle doubles the amount of target DNA; therefore, this process is exponential (161). The PCR conditions for all the primers for the genes of interest (*PKP2*, *BAG3*, *LMNA* and *TTN*) follow standard operating procedures of the Cardiovascular Genetics Laboratory, shown in Table 7 below.

Table 7: PCR conditions for amplification of genes of interest

** Indicates temperature changes as per primer set*

Condition	Temperature & time	
Initial denaturation	94°C - 4 minutes	
Denaturation Primer Annealing Template Elongation	94°C - 30 seconds * - 45 seconds 72°C - 50 seconds	33 cycles
Final elongation	72°c - 7 minutes	
Cooling (optional)	4°c - 30 minutes	

2.6.3 Principles of High-Resolution Melt

High Resolution Melt (HRM) analysis is a post-PCR technique used to detect DNA sequence variations. It does this by analysing the melting behaviour of double-stranded DNA (dsDNA). This process follows PCR amplification of a target DNA region in the presence of a fluorescent dye, (EvaGreen), which binds only to dsDNA and fluoresces upon binding. After amplification, the PCR product is gradually heated to induce denaturation of the dsDNA into single strands. As the temperature rises and the DNA denatures, the fluorescent dye is released, causing a measurable decrease in fluorescence, which is monitored in real-time (162).

Thereafter the fluorescence data is plotted against temperature to generate melting curves. Sharp melting peaks indicate homogenous sequences, while multiple peaks suggest heterogeneity (sequence variation).

For all the genes and variants of interest the same protocols were followed which are shown in Table 8 and Table 9, for the HRM master mix and conditions respectively.

Table 8: HRM Master Mix

Reagent (stock concentration)	Volume per reaction
Forward Primer (20 µM)	0.5 µl
Reverse Primer (20 µM)	0.5 µl
dNTPs (20 µM) (Bioline)	1 µl
GoTaq Polymerase (5 U/ ul) (Promega)	0.1 µl
GoTaq FlexiBuffer (5X) (Promega)	5 µl
$MgCl_2$ (25 mM) (Promega)	3 µl
EvaGreen dye (1X) (Anatech)	1 µl
DNA (50ng/ µl)	1 µl
Final Reaction Volume (made up to 25 µl with dH2O, distilled water)	**25 µl**

Table 9: HRM Conditions

CONDITION	TEMPERATURE & TIME	
Initial denaturation	95°C - 10 seconds	
Denaturation	95°C – 5 seconds	
Primer Annealing	55°C – 10 seconds	50 cycles
Template Elongation	72°C – 10 seconds	
High Resolution Melt	**72 - 95°C (0.1°C increments)**	

2.6.4 Principles of Sanger Sequencing

Sanger Sequencing is often also referred to as the "chain termination method", as it was designed to determine the sequence of nucleotide bases in a strand of DNA. It has become the 'gold standard' for DNA sequence validation as it has a 99.99% base accuracy (163, 164).

Sanger Sequencing therefore reads the DNA strand nucleotide by nucleotide. This is achieved by making use of template DNA (PCR products) and denaturing the DNA template into ssDNA (single stranded DNA). Sequencing reactions require components such as primers, DNA polymerase, 4 dNTPs (deoxynucleotides) and 4 ddNTPs (dideoxynucleotides). Dideoxynucleotides lack a 3' hydroxyl group which is needed for DNA synthesis; therefore, they result in a chain termination. Premature termination of the elongation process results in ssDNA molecules which vary in length by a single nucleotide. These molecules are then separated according to size using capillary electrophoresis(164, 165).

Dideoxynucleotides are fluorescently labelled, which allows the identification of the terminating nucleotide once the reaction is completed. Modern day techniques detect the various ddNTPs based on colour via laser firing. Each ddNTP has a different fluorescent colour which produces a chromatogram. Results are bioinformatically analysed which will display the presence or absence of a SNP (164).

Sanger Sequencing was used to investigate the PCR amplicons, in the samples of participants which showed a change in the HRM melting profile in comparison to controls. This method is used as a validation step and informs us if the proband and their immediate family members carry the same genetic variations (SNPs) in the same exons (157).

2.6.4.1 Sanger Sequencing Protocol

A clean-up reaction was done to purify the samples for Sanger sequencing. Table 10 shows the reagents and volumes required for the clean-up reaction (Whitehead Scientific).

Table 10: Clean-up reaction for sample purification

REAGENT	VOLUME PER REACTION
Nuclease free water (NFW)	13.9 µL
Shrimp alkaline	1 unit
Exonuclease	0.1 unit
PCR product	5
TOTAL	20

Each reaction tube was incubated in an Eppendorf thermal cycler at 37ºC for 60 minutes and at 75ºC for 15 minutes to digest any left-over products and inactivate the enzymes. During incubation, the exonuclease and phosphatase act to removed left-over primers and dNTPs.

The products of the clean-up were used to perform a Big-Dye v3.1 resequencing reaction. A 20 µL Big-Dye sequencing reaction was set up in a microtube (Whitehead Scientific) containing 15µL master mix as shown in Table 11.

Table 11: The reagents and volumes used per reaction for the Big-Dye reaction

REAGENT	VOLUME PER REACTION (µL)
Nuclease free water (NFW)	9
Buffer (5x)	4
Big-dye	2
Clean-up product	3
Primer (sense / antisense)	2
TOTAL	20

Each reaction tube was thereafter run in a Thermocycler using standard protocols, shown in Table 12.

Table 12: The PCR reaction protocol for the Big-Dye reaction

CONDITION	TEMPERATURE & TIME	
Initial denaturation	96 ºC for 5 mins	
Denaturation Primer annealing Elongation	96ºC for 30s 50ºC for 15s 60ºC for 4 mins	25 cycles
HOLD	4ºC for 30s	

The Big-Dye reaction products were used in sequencing electrophoresis-only reactions at the Stellenbosch University central analytical facilities (CAF) or Inqaba biotech to generate

electropherograms. The electropherograms were analysed to confirmed and validate the variants in Alamut Visual Plus version 1.5.1 or BioEdit sequence alignment editor version 7.2.6.1 (https://bioedit.software.informer.com/7.2/). Finally, the genotype positive results were correlated with the phenotype.

2.7 Haplotyping

2.7.1 Principles of Microsatellites

Microsatellites are short tandem repeats (STRs) and are commonly used for paternity testing (166). They are repeats of nucleotides (di-, tri-, tetra or penta- nucleotides repeated between eight and fifty times).

According to Mendelian inheritance patterns, each individual inherits one allele at a given autosomal marker from each parent, as only two alleles are present - one contributed by the mother and one by the father.

2.7.2 Primer design for microsatellite markers

For each single nucleotide polymorphism (SNP), in each gene of interest, three microsatellite markers were identified by making use of the UCSC Genome Browser (167). The microsatellite markers were selected based on the location around the SNP, therefore two flanking markers (5' and 3') as well as a middle marker were selected. Thereafter, primers were designed for each microsatellite marker, by using the online bioinformatic tool: IDT OligoAnalyser (168). The primers were designed according to the following parameters: between 18-25 base pairs long, good sequence variety (minimal sequence repeats and secondary structures), a melting temperature of between 45 – 65˚C, a GC content between 45 -65% and a G/C clamp located at the 3' end of the sequence. Additionally, each forward primer was designed with a tagging sequence in front of the primer sequence, these primers are fluorescently tagged. There are three M13 dyes that can be used for fluorescent tagging, and each represents a different fluorescent dye, FAM is blue, VIC is green and PET is red. The amplification of the microsatellites was performed by polymerase chain reaction (PCR) of varying conditions (Table 13). The polymerase chain reaction products were qualified and quantified using 1.5% agarose gel to ensure the correct product size was amplified.

Table 13: Polymerase chain reaction (PCR) procedure

Condition	Temperature (˚C)	Time	
Initial denaturation	94	4 min	
Denaturation	94	30 sec	35 cycles
Primer Annealing	*	45 sec	
Template elongation	72	1 min	
Final elongation	72	7 min	
Cooling	4	30 min	

**Indicates varying annealing temperatures per primer set.*

These products underwent capillary electrophoresis which was outsourced to Inqaba Biotechnology Industries. The data was visualised and compared against the GeneScan ROX-labelled GS500 internal size standard (Applied Biosystems). The microsatellite markers were analysed using GeneMarker HID V2.9.5, which allows the visualisation of markers through electrophentograms. There were two phases of analysis. First the pattern and heights of the microsatellites were analysed, this allowed us to exclude peaks which did not follow the generalised pattern for the marker and peaks which failed to reach a sufficient height. The second phase of analysis was the binning of the peaks which followed the pattern and achieved the greatest hights. Generally, if only one very large peak (double the height of any other) exists in a sample, this participant was considered homozygous. However, if two peaks with a small difference in height is evident, the participant was considered heterozygous.

Thereafter the microsatellite marker and genotypes were used for haplotype and segregation analysis.

2.7.2.1 Microsatellite markers for the PKP2 gene

The *PKP2* gene is found on chromosome 12, however this gene is not found in the forward orientation, it is coded in the reverse direction. Thus, the 5' end found in UCSC genome browser is the 3' end of the gene and vice versa.

Previous research in this laboratory used three microsatellite markers for the *PKP2* gene: *PKP2*_CA (tagged with FAM, for dinucleotide CA repeats), *PKP2*_D12S1692 (tagged with FAM, for dinucleotide CA repeats) and *PKP2*_TG (tagged with PET, for dinucleotide TG repeats). However, upon further analysis of the *PKP2*_CA marker with the newly enrolled probands and their families, this marker was found to be uninformative. A new marker, *PKP2*_AC, was designed (tagged with PET, for dinucleotide AC repeats). One family from

the previous research, ACM 8, was used as a positive control for the new probands and their families to ensure the haplotypes and subsequent haploblocks tracked with the old and new families. This will be presented in Chapter 3. For these three markers we used the Q5 High-Fidelity 2X Master Mix as it is known to have the highest fidelity amplicon with a high specificity and yield which was an important factor to consider when amplifying microsatellites which consist of several repeats of nucleotides. The three markers for *PKP2* and their conditions are presented in Table 14 to Table 19 below.

Table 14: PKP2 microsatellite markers and optimal PCR conditions

Microsatellite marker	Chromosomal position (bp)	Expected size (bp) and repeats	T_A, Annealing Temperature (˚C)	Fluorescent label and strand	Marker name and Sequences (5' to 3')
***PKP2*_TG**	33085200 - 33085832	341 TG	60	PET Forward	*PKP2* TG F GGGTCTGTATGTGTTTGCT
				None Reverse	*PKP2*_TG_R CATACACTCTCAAATAGAAATAGG
***PKP2*_D12S1692**	32987723 - 32987971 (Intron 7 of gene)	245-261 CA	60	FAM Forward	*PKP2*_D12S1692_F TTTGATTCCATACCCTCCT
				None Reverse	*PKP2*_D12S1692_R GCAGCAATTTCAGACTTCTC
***PKP2*_AC**	32324575 - 32325224	239 AC	61.8	PET Forward	*PKP2*_AC_F ATTGCTTGAACCCAGGAGGC
				None Reverse	*PKP2*_AC_R GAACGGAGGGAGAGGCAGA

Table 15: Polymerase chain reaction (PCR) reagents for PKP2_TG and PKP2_D12S1692, first polymerase chain reaction

Reagent	Volume (ul) n=1
dH_2O	10
Q5 Master Mix	12.5
Forward Primer	1
Reverse Primer	1
DNA (~100ng/ul)	0.5
Total	**25**

Table 16: Polymerase chain reaction (PCR) reagents for PKP2_TG, second polymerase chain reaction

Reagent	Volume (ul) n=1
dH_2O	9.5
Q5 Master Mix	12.5
M13 PET dye primer	1
Reverse Primer	1
1st PCR	1
Total	**25**

Table 17: Polymerase chain reaction (PCR)reagents for PKP2_ D12S1692, 2nd PCR

Reagent	Volume (ul) n=1
dH_2O	9.5
Q5 Master Mix	12.5
M13 FAM dye primer	1
Reverse Primer	1
1st PCR	1
Total	**25**

Table 18: Polymerase chain reaction (PCR) reagents for PKP2_ AC, first polymerase chain reaction

Reagent	Volume (ul) n=1
dH_2O	15.9
$MgCl_2$ (25Mm)	1
dNTPs (20mM)	1
GoTaq Flexi Buffer (5X)	5
GoTaq Polymerase (5U/ul)	0.1
Forward Primer (20uM)	0.5
Reverse Primer (20uM)	0.5
DNA (~100ng/ul)	1
Total	**25**

Table 19: Polymerase chain reaction (PCR) reagents for PKP2_ AC, second polymerase chain reaction

Reagent	Volume (ul) n=1
dH_2O	9.5
Q5 Master Mix	12.5
M13 PET dye primer	1
Reverse Primer	1
1st PCR	1
Total	**25**

2.7.2.2 Microsatellite markers for the BAG3 gene

The *BAG3* gene is located on chromosome 10. The three microsatellite markers for the *BAG3* gene were previously optimised and designed in our laboratory. These markers are as follows; *BAG3*_11Kb (tagged with FAM dye, for dinucleotide TG repeats), *BAG3*_AC (tagged with FAM dye, for dinucleotide AC repeats) and *BAG3*_TG marker (tagged with

FAM dye, for dinucleotide TG repeats). The *BAG3*_11Kb microsatellite marker underwent three rounds of PCR due to non-specific binding. The 1st PCR was used in the second PCR with a second set of primers (as seen in table 2.2 below) and the 3rd PCR used the products of the 2nd PCR and instead of the forward primer the PET dye was added, similarly to the other 2nd PCRs have been performed for the other markers. This is further discussed in Chapter 4. The three markers for *BAG3* and their conditions can be found in the Table 20 to Table 26 below.

Table 20: BAG3 microsatellite markers and optimal PCR conditions

Microsatellite marker	Chromosomal position (bp)	Expected size (bp) and repeats	T_A, Annealing Temperature (˚C)	Fluorescent label and strand	Marker name and Sequences (5' to 3')
***BAG3*_11KB_1**	121424305 - 121424939	TG	60	FAM Forward	*BAG3*_11KB_F1 ATGGTGGTGCACGCC TGTAATCC
				None Reverse	*BAG3*_11KB_R1 CCA TAT CAC TCC ACA CAG AGC CTC
***BAG3*_11KB_2**	121424305- 121424939	TG	60	FAM Forward	*BAG3*_11KB_F2 *GGAGGTGGAGGTTGCAGTGAGCCG*
				None Reverse	*BAG3*_11KB_R2 CCTAATTTATTGAACTGGTCTCC
***BAG3*_AC**	121377620 - 121378257	AC	60	FAM Forward	*BAG3*_AC_F AAGTCTGGTGTCTTGAACC
				None Reverse	*BAG3*_AC_R GTGAAGTAAGGTGGAAAGGA
BAG3_TG	121452313 - 121452963	TG	60	FAM Forward	*BAG3*_TG_F CTGCAAACTCATCAGGATG
				None Reverse	*BAG3_TG*_R CTCTTGAGCATCTGTCTGTT

Table 21: Polymerase chain reaction (PCR) reagents for BAG3_11KB, first polymerase chain reaction

Reagent	Volume (ul) n=1
dH_2O	10
One Taq Master Mix	12.5
Forward Primer (1)	1
Reverse Primer (1)	1
DNA (~100ng/ul)	0.5
Total	**25**

Table 22: Polymerase chain reaction (PCR) reagents for BAG3_11KB, second polymerase chain reaction

Reagent	**Volume (ul) n=1**
dH_2O	9.5
One Taq Master Mix	12.5
Forward Primer (2)	1
Reverse Primer (2)	1
1st PCR Product	1
Total	**25**

Table 23: Polymerase chain reaction (PCR) reagents for BAG3_11KB, third polymerase chain reaction

Reagent	**Volume (ul) n=1**
dH_2O	9.5
One Taq Master Mix	12.5
M13 FAM dye primer	1
Reverse Primer (2)	1
2nd PCR Product	1
Total	**25**

Table 24: Polymerase chain reaction (PCR) reagents for BAG3_AC and BAG3_TG, first polymerase chain reaction

Reagent	**Volume (ul) n=1**
dH_2O	10
Q5 Master Mix	12.5
Forward Primer	1
Reverse Primer	1
DNA (~100ng/ul)	0.5
Total	**25**

Table 25: Polymerase chain reaction (PCR) reagents for BAG3_AC, second polymerase chain reaction

Reagent	**Volume (ul) n=1**
dH_2O	9.5
Q5 Master Mix	12.5
M13 FAM dye primer	1
Reverse Primer	1
1st PCR	1
Total	**25**

Table 26: Polymerase chain reaction (PCR) reagents for BAG3_TG, second polymerase chain reaction

Reagent	**Volume (ul) n=1**
dH_2O	9.5
Q5 Master Mix	12.5
M13 PET dye primer	1
Reverse Primer	1
1st PCR	1
Total	**25**

2.7.2.3 Microsatellite markers for the LMNA gene

The *LMNA* gene is found on chromosome 1. The three markers selected for this gene can be found in Table x below, *LMNA*_GT (tagged with PET dye, for dinucleotide repeats GT),

*LMNA*_CT (tagged with PET dye, for dinucleotide repeats CT) and *LMNA*_AC (tagged with FAM dye, for dinucleotide repeats AC). Each marker underwent two PCRs, one to amplify the region of interest and the second to add the respective dye. The three markers for *LMNA* and their conditions can be found in the Table 27 to Table 31 below.

Table 27: LMNA microsatellite markers and optimal Polymerase Chain Reaction conditions

Microsatellite marker	Chromosomal position (bp)	Expected size (bp) and repeats	T_A, Annealing Temperature (˚C)	Fluorescent label and strand	Marker name and Sequences (5' to 3')
***LMNA*_GT**	155914772-155915414	334 GT	61.8	Forward PET	*LMNA*_GT_F CAACGGTGAAAGATGTGAGGA AG
				None Reverse	*LMNA*_GT_R ACAGATCCAATGAGAGATGCTGGT
***LMNA*_CT**	156055835 - 156056566	427 CT	60	Forward PET	*LMNA*_CT_F TCATTCAGCCAGCCACTTCCTA
				None Reverse	*LMNA*_CT_R ACAACTCCAGACACAGACAAGC
***LMNA*_AC**	156109602-156110346	439 AC	60	Forward FAM	*LMNA* AC F AAACGCTAAAGAGCCCTTGCC
				None Reverse	*LMNA*_AC_R CTGGTGGACGAGGGAAGTTAG

Table 28: Polymerase chain reaction (PCR) reagents for LMNA _GT, first polymerase chain reaction

Reagent	Volume (ul) n=1
dH_2O	14.9
$MgCl_2$ (25Mm)	2
dNTPs (20mM)	1
GoTaq Flexi Buffer (5X)	5
GoTaq Polymerase (5U/ul)	0.1
Forward Primer (20uM)	0.5
Reverse Primer (20uM)	0.5
DNA (~100ng/ul)	1
Total	**25**

Table 29: Polymerase chain reaction (PCR) reagents for LMNA_CT and LMNA_AC, first polymerase chain reaction

Reagent	Volume (ul) n=1
dH_2O	10
One Taq Master Mix	12.5
Forward Primer	1
Reverse Primer	1
DNA (~100ng/ul)	0.5
Total	**25**

Table 30: Polymerase chain reaction (PCR) reagents for LMNA _GT and LMNA_CT, second polymerase chain reaction

Reagent	Volume (ul) n=1
dH_2O	9.5
One Taq Master Mix	12.5
M13 PET dye primer	1
Reverse Primer	1
1st PCR	1
Total	**25**

Table 31: Polymerase chain reaction (PCR) reagents for LMNA _AC, second polymerase chain reaction

Reagent	Volume (ul) n=1
dH_2O	9.5
Q5 Master Mix	12.5
M13 FAM dye primer	1
Reverse Primer	1
1st PCR	1
Total	**25**

2.7.2.4 Microsatellite markers for the TTN gene

The *TTN* gene is found on chromosome 2 and is the largest gene in the human genome. Despite trying different Taq Polymerases, temperature gradients, $MgCl_2$ gradients and redesigning primers multiple times, one of the microsatellite markers for the *TTN* gene still requires further investigation and optimisation. It has multiple repeating nucleotides, and non-specific binding results from this, making it a very messy gene. Therefore, there were multiple bands on the gels and identifying the region of interest was extremely difficult. Two (*TTN*_TG and *TTN*_AC) of the three markers were successfully optimised, the third (*TTN*_TA) still requires further optimisation. Due to monetary and time constraints, this was not possible for this study. This is further discussed in Chapter 6. The three markers for *TTN* are presented in Table 32 below.

Table 32: TTN microsatellite markers and optimal PCR conditions

Microsatellite marker	Chromosomal position (bp)	Expected size (bp) and repeats	T_A, Annealing Temperature (°C)	Fluorescent label and strand	Marker name and Sequences (5' to 3')
***TTN*_TG**	179372428 - 179373166	496 TG	60	Forward PET	*TTN*_GT_F ACCGTTCCTTTACCTGGGATCT
				None Reverse	*TTN*_GT_R CCTCAGGGAGATATGGGGAAACA
***TTN*_TA**	179510532 - 179511174	411 TA	-	Forward -	*TTN* TA F CAC ATC TTC CTT AGG TGG AGC AG
				None Reverse	*TTN*_TA_R CTGTGGCTTGTCTGGATGCT
***TTN*_AC**	179785223- 179785656	411 AC	55.6	Forward -	*TTN*_AC_F TCATTCAGCCAGCCACTTCCTA
				None Reverse	*TTN*_AC_R GCATACTCTTCTGAACCTATGTCA

2.8 Familial segregation analysis

Haplotypes for each proband and their family members were constructed from both the genotype information and microsatellite markers. The haplotypes for each family for the specific variation in question was merged with their pedigree information which allowed patterns of inheritance to be determined when unknown. We were able to construct the assumed haplotypes of the previous generations were when enough information was provided, by the enrolled participants. The haploblock was construction based on the gene of interest, microsatellite markers and SNPs.

2.9 Genealogical tracing

Genealogy is defined as the studying of families, their histories and construction of pedigrees. Genealogical tracing was conducted using standard methods for genetic genealogical tracing, which includes interviews with the probands and their family members, searching of online data bases, making use of historical records for data through the Genealogical Institute of South Africa, state archives, church, death, birth, baptismal, immigration, slave and estate records were investigated to gather genealogical evidence (104, 169).

Recently, all the above-mentioned data was digitalised by the Cape Town City Archives and new records are uploaded and the data is indexed on an online platform, Family Search (https://www.familysearch.org/en/). This information is accessible to the general public and

has records dating back to the first French Huguenot refugees. However, to search an individual using this online tool with accuracy, there are a few fields required, such as their full first and last names, their date and place of birth and in some cases the name of their spouse or parents. Other resources such as the voters roll, published genealogies, as well as consulting Prof G Geldenhuys were used to resolve any discrepancies between records.
The Legacy v10.0 Software was used to track, store, manage, merge the immense amount of data collected while genealogically tracing each family. With each generation the number of individuals doubles, as a proband will have two parents, each of whom have two parents and so forth. This is usually the case with each generation expect in cases where there are consanguine marriages which was common when the early settlers, first came to South Africa. Information was gathered through these resources, family trees of several generations were created and compared to find a common founder for the variants of interest. Figure 14 below shows the methodology followed to complete genealogy tracing.

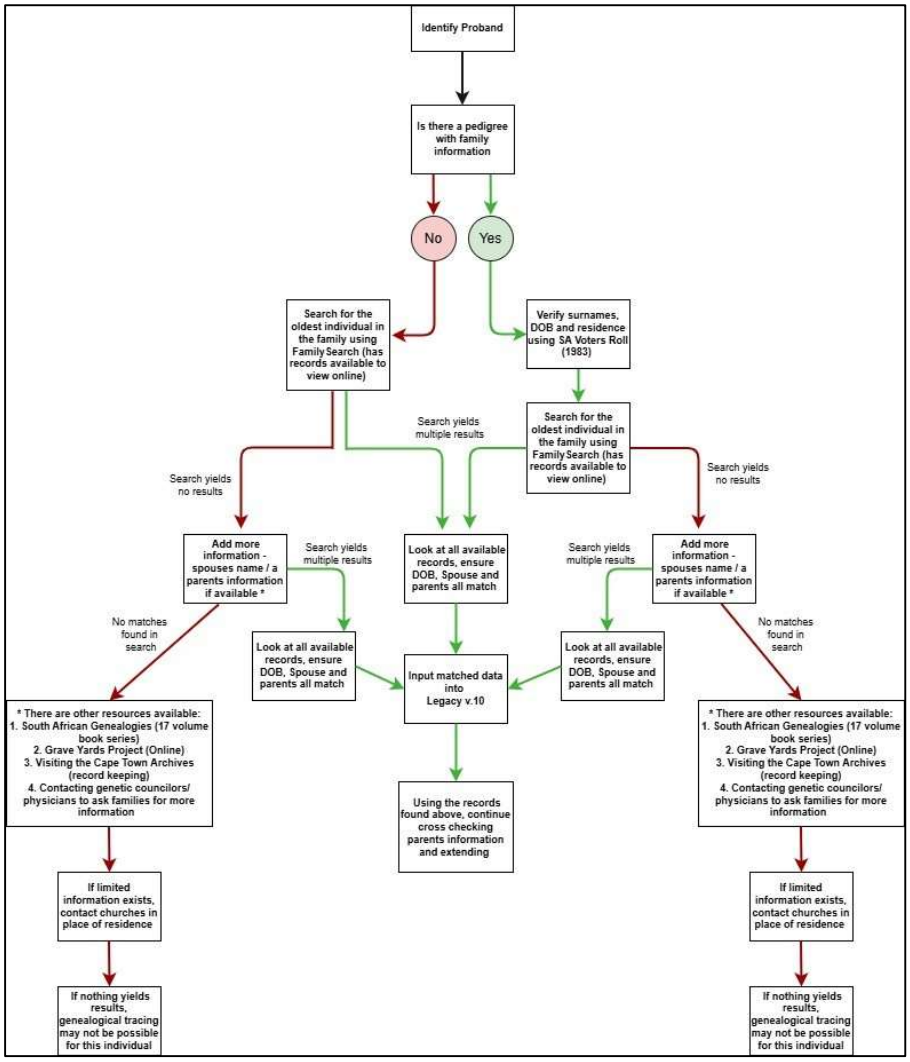


Figure 14: Flow chart demonstrating the methodology used to trace families and create genealogies

Chapter 3: *PKP2 c.1162* C>T pathogenic variant in South African Afrikaans Families with Arrhythmogenic Cardiomyopathy

3.1 Introduction

Arrhythmogenic Cardiomyopathy (ACM) is associated with five desmosomal genes; desmocollin-2 (*DSC2*), desmoglein-2 (*DSG2*), junctional plakoglobin (*JUP*), plakophilin-2 (*PKP2*) and desmoplakin (*DSP*), as mentioned in Chapter 1. However, *PKP2* has been identified as the most common genetic cause of ACM in European populations (46, 54).

In a South African setting, a programme was launched in 2004 by the Cardiac Arrhythmia Society of South Africa (CASSA) and the ARVC registry of South Africa, which was based at the Groote Schuur Hospital and Department of Medicine at the University of Cape Town. The main aim of the registry was to identify and understand the clinical features and molecular genetics of ACM (54).

In 2009, Watkins et al. reported on the genetic profile of 100 enrolled ARVC participants and found the *PKP2* gene had the highest number of disease-causing mutations (13.9%) of all the above-mentioned genes known to be associated with ACM (Figure 15, Adapted from (54)).

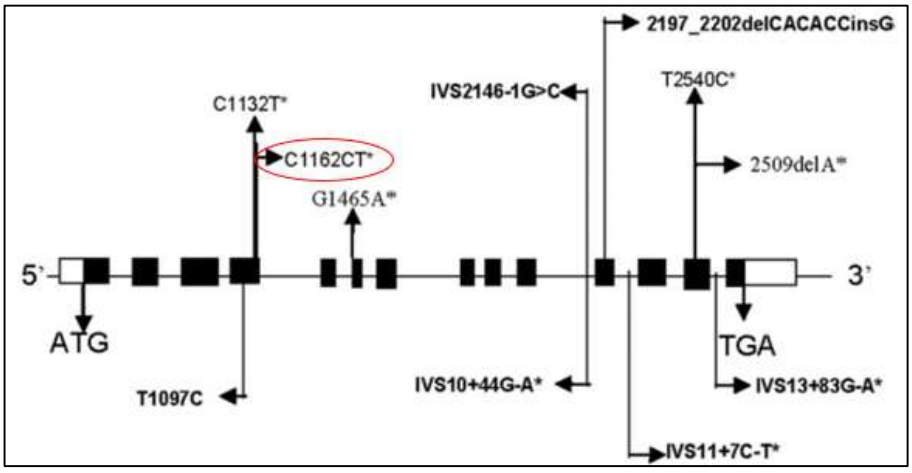


Figure 15: Schematic representation of PKP2 gene. The location of the founder PKP2 c.1162C>T variant is circled in red and is shown in the context of other variants. The black blocks represent the exons, above the exons are disease-causing mutations and below are polymorphisms

Extensive research have been done into the role the PKP2 protein has in cardiac function with the majority of cases reporting that pathogenic variants in the *PKP2* gene results in ACM

(170). PKP2 is mainly known for its mechanical functions between cardiac cells, providing lateral stabilising force that allows cell to cell adhesion (171). However, recent years and more research have provided evidence that PKP2 has a pleiotropic role as it is an important facilitator and regulator of a number of molecular pathways (172).

One key interaction involves PKP2 and the voltage-gated sodium channel (Nav1.5), which is essential for cardiac conduction. Studies have shown that reduced PKP2 levels lead to decreased Nav1.5 expression and sodium current density, impairing conduction and potentially contributing to arrhythmogenesis in the early stages of arrhythmogenic cardiomyopathy (ACM) (173). Additionally, PKP2 deficiency disrupts calcium handling, causing prolonged calcium transients and increased intracellular calcium accumulation, further predisposing the heart to ventricular arrhythmias (174).

Beyond sodium and calcium regulation, the loss of PKP2 also affects Cx43 expression at the intercalated disc, destabilising gap junctions and increasing myocardial membrane permeability (175). This instability facilitates mitochondrial Ca2+ overload via Cx43 hemichannels, triggering oxidative stress and activating arrhythmogenic pathways (176, 177).

Therefore, mutations in *PKP2* that disrupt the protein function initiate a cascade of disruptions of intracellular signalling, affecting of the electrophysiological stability and causing to arrhythmias (178). Notably, these arrhythmic events can occur before any structural damage is detectable, highlighting the critical role of *PKP2* in maintaining cardiac rhythm (179) (Figure 16 (170)).

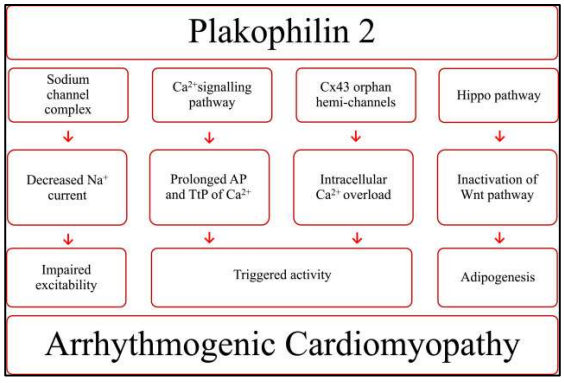


Figure 16: The pivotal role of PKP2 (171)

These findings contradict the current doctrine which showed that arrhythmias were the result of anatomical obstacles such as fibrofatty deposits and although this remains a prominent

cause of fatal arrhythmia, the various signalling cascade in which PKP2 is involved in can no longer be ignored.

Despite all this research, exact role or signalling pathway the variant of interest of this study, *PKP2* c.1162C>T (p.P388W), is yet to be determined, we know it has had a profound impact on many South African families. In Watkins et al, our research showed that many of the South African families harbouring the *PKP2* c.1162C>T (p.P388W) variant all self-identified as Afrikaners. Afrikaners are said to be Caucasian, Afrikaans-speaking South Africans, who have descended from Dutch and Huguenot settlers from the 17th century (104), more details on the Afrikaners are provided in Chapter 1, section 1.2.2. Watkins et al showed that all the probands and their families that harboured the *PKP2* c.1162C>T variant, had a common ancestor. Therefore, the original probands, and where possible, their extended relatives were clinically screened, genotyped for the *PKP2* c.1162C>T variant and then haplotyped w, where three microsatellites were designed which flanked the *PKP2* gene. The results showed that the same haplotype was shared across six Afrikaner families, which further strengthened the probability that these families shared a common ancestor, leading to a founder effect (157).

Through genealogy tracing an attempt was made to identify the original progenitors that brought the *PKP2* c.1162C>t to South Africa but, this proved to be difficult as an additional four families were needed to trace back to the 1600s (157), as per genealogy expert Professor Geldenhuys from the Genealogical Society of South Africa (GSSA) (104).

The establishment of the IMHOTEP registry in 2013 allowed the recruitment of five new ACM probands as well as their family members, bringing the total number of families to 12. This resulted in the current study where the aim is now to continue haplotype construction and genealogy tracing of the newly enrolled probands and their families to identify the original progenitors that brought the *PKP2* c.1162C>t to South Africa.

3.2 Methods

All methods are described in Chapter 2 as they are repeated across Chapters 3 to 6.

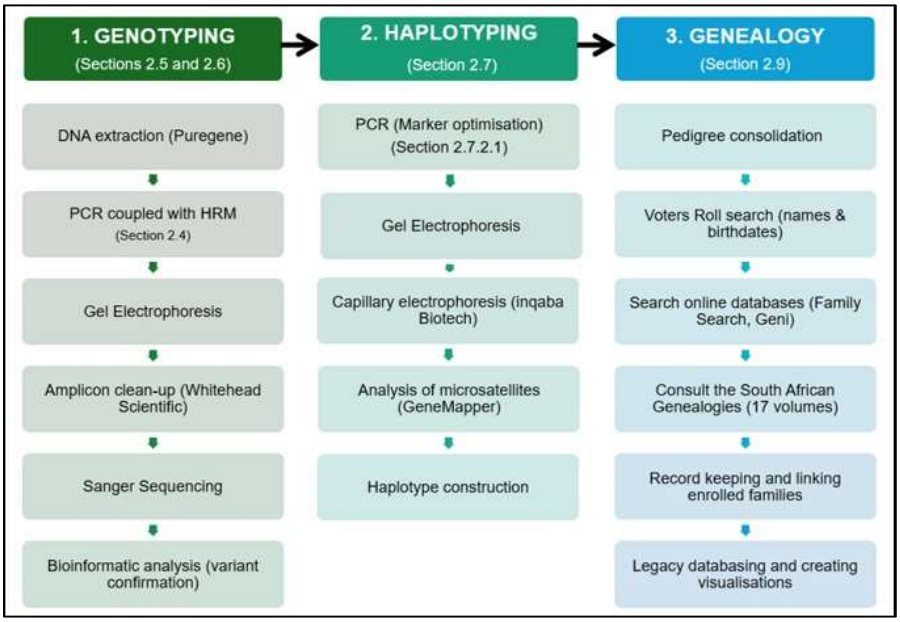


Figure 17: Workflow, adapted from Chapter 2, Section 2.3, to highlight the relevant sections detailing the methods associated with this chapter, specifically focusing on the PKP2 c.1162C>T variant.

3.2.1 Clinical

Patients with suspected ACM, were referred to the Cardiac Clinic at Groote Schuur Hospital. Once their diagnosis was confirmed they were enrolled into the IMHOPTEP registry the inclusion and exclusion criteria can be found in Chapter 2, section 2.2.

3.2.2 Genetics

Refer to the appropriate section in Chapter 2, section 2.4, Table 3. The probands were screened and were identified as such if they were positive for the *PKP2* c.1126 C>T variant. Thereafter, through continuous screening and recruitment, additional family members were enrolled into the registry by their clinician and/or genetic councillor.

3.2.3 Haplotype

A haplotype is a set of DNA variations, or alleles, inherited together on the same chromosome. These alleles are physically linked and tend to be passed down as a unit. Think of it as a specific combination of genetic variants on a single chromosome. Haplotypes are

essential for understanding genetic inheritance, mapping disease genes, and studying population history (180).

Note that haplotypes for the old families ACM 5, ACM 12, ACM 19, ACM 38, ACM 57 and ACM 71 (for convenience will be called group 1) were all constructed through previous research and were not haplotyped again; an additional six families were haplotyped for this project. We included one positive control (ACM8) from the previous research, for continuity between this project and the previous project. We successfully reproduced the red diseased haploblock for ACM8 (132, T, 262, 250) through use of the current microsatellites which confirmed that ACM8 could be used as a positive control for this project as it replicated the disease haplotypes of the previous project. For affected individuals where we had no DNA, we were able to infer (dotted borders around haplotypes) the haplotype if the proband's parents and/ or siblings' DNA were available.

Inferring haplotypes within a family becomes easier when you have DNA from parents or siblings. Parental DNA allows you to directly observe the haplotypes transmitted to their offspring. If you know the parents' genotypes, you can deduce the child's haplotypes with greater certainty (181). Similarly, sibling DNA can help resolve ambiguous haplotypes. By comparing the genetic variations shared between siblings, you can infer which segments of their chromosomes were inherited from each parent. This information is particularly useful in linkage studies, where researchers track the inheritance of specific haplotypes within families to identify genes associated with diseases (181, 182).

3.2.4 Genealogy

Genealogy tracing followed standard procedures (Figure 14, Chapter 2). The oldest enrolled family member was the starting point of genealogy tracing. The tracing and where possible the ancestral lineages were extended as far as records allowed.

3.3 Results

Twelve probands (ACM 163, ACM 12, ACM 38, ACM 154, ACM 19, ACM 8, ACM 5, ACM 57, ACM 157, ACM 158, ACM 161 and ACM 71) and 56 family members were enrolled.

The results reported in this section encompass the 12 families and will follow the order of clinical results, genetics screening for *PKP2* c.1162C>T per family, the haplotyping and finally the genealogy. Of these 12 probands and their families, 11 were self-identified Caucasian Afrikaners while one proband self-identified as mixed-race. This will be addressed in more detail per family, in the sections below.

The 12 families were screened: the seven previously published (54) old families were ACM 5, ACM 8, ACM 12, ACM 19, ACM 38 ACM 57, ACM 71 and the five newly recruited families were ACM 163, ACM 154, ACM 157, ACM 158 and ACM 161. The haplotypes were constructed using microsatellite markers combined with the *PKP2* c.1126C>T (p.Arg388Trp) variant to construct the haplotype.

One marker (PKP2_CA) was uninformative in the newly enrolled families (ACM 154, ACM 157, ACM 158, ACM 161 and ACM 163) and a new microsatellite marker was designed, PKP2_AC25.

Genealogy tracing followed standard procedures (Chapter 2, Figure 14). The oldest enrolled family member was the starting point of genealogy tracing.

The order of presentation of the families were from the most genealogically traced (with the most information) to the least traced (with the least information), thereby allowing us to eliminate potential progenitors through comparison of which progenitors were in common between the families. Every subsequent family allowed us to decrease the total number of progenitors; we started with 72 progenitors. The order was as follows: ACM 163, ACM 12, ACM 38, ACM 154, ACM 19, ACM 8, ACM 5, ACM 57, ACM 157. ACM 158, ACM 161 and ACM 71.

3.3.1 Family ACM 163 (new family)

3.3.1.1 Clinical history

The proband (ACM 163.1), her mother ACM 163.2 and her father ACM 163.3 were enrolled in the IMHOTEP registry. The proband was a marathon runner before her diagnosis of ACM at 53 years old when she first presented with tachycardia. Both her parents, ACM 163.2 (mother) and ACM163.3 (father), were screened and found to be clinically unaffected (Figure 18).

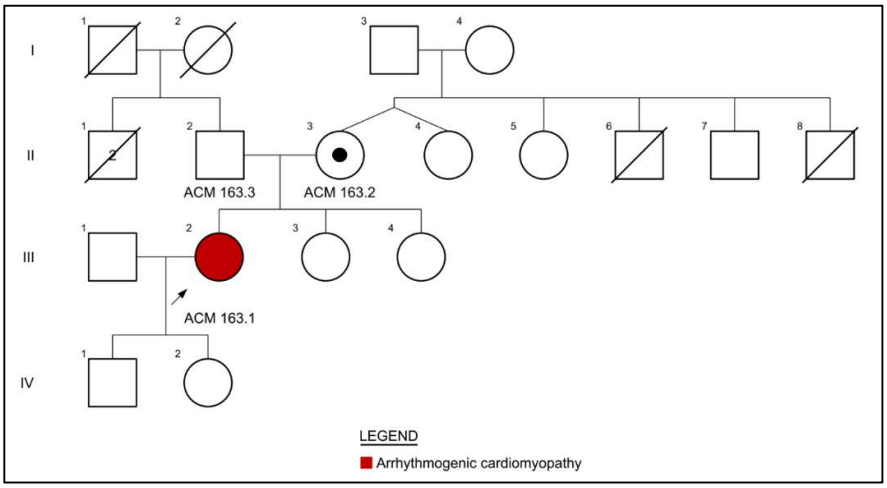


Figure 18: Pedigree of family ACM 163.The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals

3.3.1.2 Genetics: PKP2 c.1162C>T screening

In Figure 19 the *PKP2* c.1162C>T genotyping result is shown. The proband (ACM 163.1) was heterozygous for the *PKP2* variant. Her mother (ACM 163.2) was positive for the *PKP2* variant, showing inheritance through the probands maternal line, while her father (ACM 163.3) was negative.

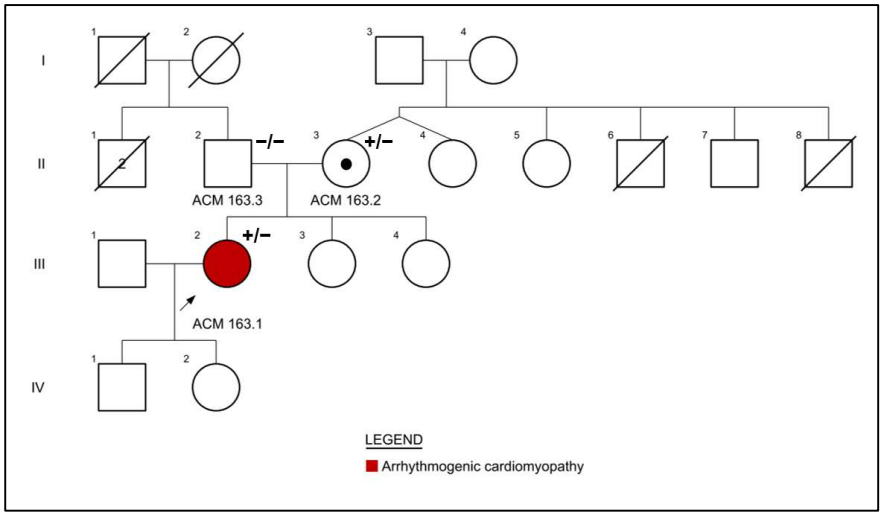


Figure 19: The genotyping results of ACM 163. The arrow indicates the proband, '+' indicates a positive genotype of the PKP2 variant while '−' indicates a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms

3.3.1.3 Haplotype analysis

Haplotyping was performed for two of the enrolled family members through a combination of microsatellites and the variant of interest (*PKP2* c.1162C>T). The proband's (ACM 163.1) microsatellite marker peaks are shown in Figure 20 below.

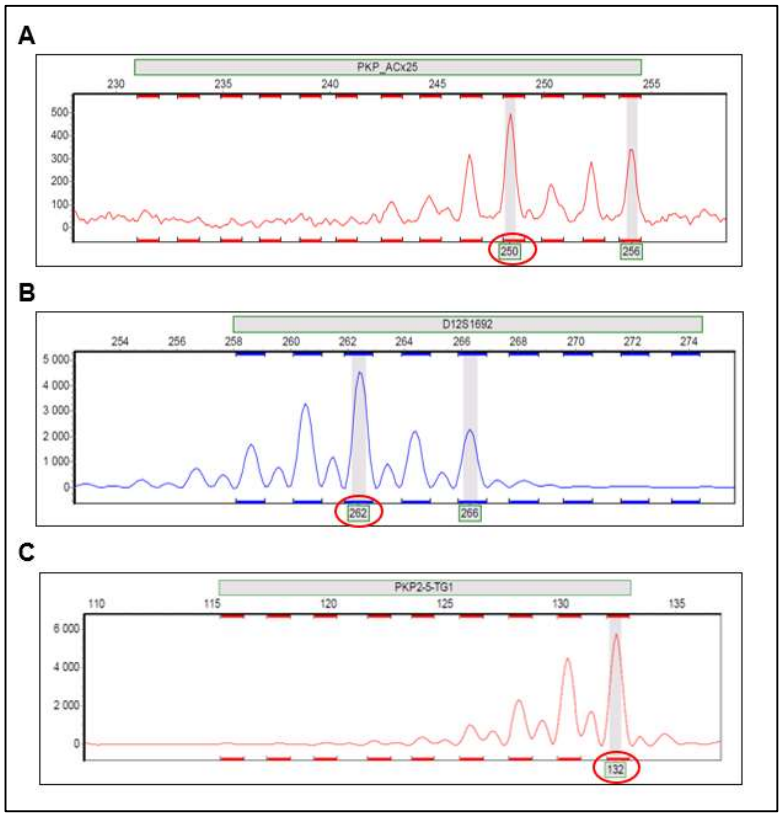


Figure 20: Microsatellite analysis of ACM 163.1. (A) PKP2_TG (132), the peak circled in red, 132, is the common haplotype peak for PKP2_TG. (B) D12S1692 (262), the peak circled in red, 262 is the common haplotype peak for D12S1692. (C) PKP2_AC (250), the peak circled in red, 250 bps, is the common haplotype peak for PKP2_AC.

The microsatellite analysis results were combined and used to construct haplotypes for ACM 163.1 (proband) and ACM 163.3 (the father), and while ACM 163.2 is genotype positive, her DNA was severely degraded, meaning we were unable to perform microsatellite analysis. We were however able to infer a haploblock (Figure 21). The proband has the disease-associated

haploblock, shaded in red, which is the inferred haploblock for her mother, ACM 163.2. ACM 163.3 is genotype and haplotype negative for the disease-associated haploblock.

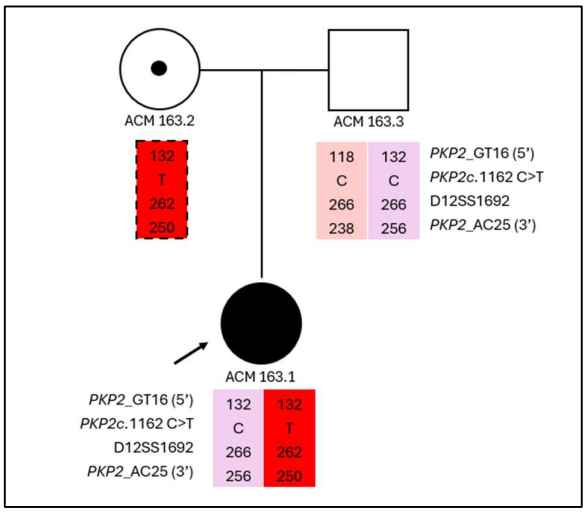


Figure 21: Haplotype analysis of family ACM 163. Circles indicate self-identified females, squares indicate self-identified males, black-filled symbols represent clinically affected individuals, symbols with periods in the centre represent individuals with the PKP2 variant but are not clinically affected, shapes with an oblique line through indicates deceased individuals shapes with an oblique line through indicates deceased individuals, haplotypes with dotted borders represent assumed haplotypes

3.3.1.4 Genealogy tracing

To trace this family's genealogy, a great deal of information had to be gathered and organised, including copies of birth, marriage, and death certificates; records in churches and graveyards; old articles, letters, photographs, and documents related to citizenship and naturalisation; searches of historical records, including land records and censuses; a backwards and forwards approach, starting with the proband.

The proband (ACM 163.1) and her intermediate family all self-identified as Caucasian Afrikaners which meant it would be easier to trace this family as the Afrikaners were known to have well-documented records in terms of birth, death and marriage records. This family represents the most completely genealogically traced family, with all lines extending back to the progenitors in the 1600s.

There was limited information provided on the pedigree during enrolment. A few names, dates of birth and locations of residence for the proband and her parents were provided. Fortunately, both the proband's mother (ACM 163.2) and father (ACM 163.3) provided the genetic councillor with well-documented family genealogies. As the *PKP2* c.1162C>T

variant was maternally inherited, we only extended the genealogy of ACM 163.2 (the proband's mother); she was also the oldest positive genotyped member of this family.

Both the paternal and maternal lines were extended for ACM 163.2, we started by verifying the full names and birth dates of both ACM 163.2's mother and father by making use of the 1984 voter's roll. Once both lineages were confirmed, the names and birth dates provided on the genealogy for previous generations were verified by making use of digitalised records (Baptism records, marriage records, death records etc) available from online databases (FamilySearch, Geni and Ancestry). This was done per individual to ensure a full and accurate genealogy was created. We found that the ACM 163 family currently resides in the Western Cape and have been residents since the arrival of their ancestors in the 17th century. Both the maternal and paternal lines of ACM 163.2 were traced back to their progenitors who came to South Africa in the 17th century. There was evidence of admixture between individuals of Dutch, French and German ancestry and many of the names are well-known progenitors of South Africa, the distributions of ancestry can be found Figure 22 below. There were instances of consanguineous marriages between cousins, dating back to the 18th century in both lineages.

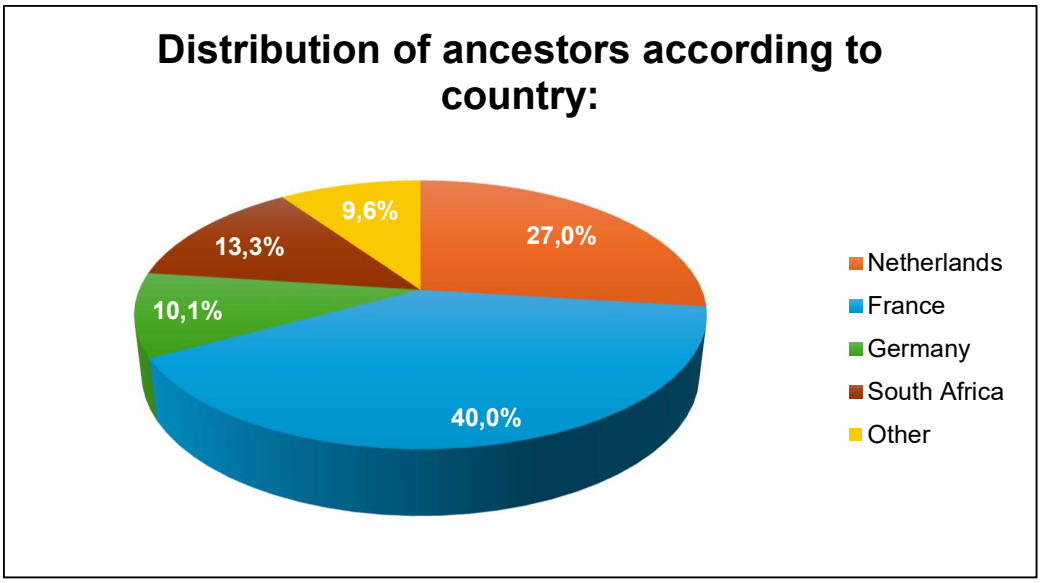


Figure 22: Distribution of ancestors according to country for ACM 163

All data was manually inputted, stored and analysed using Legacy v.10. We were able to identify 74 possible founder couples (progenitors) from this family by tracing back through 14 generations (shown in Figure 23).

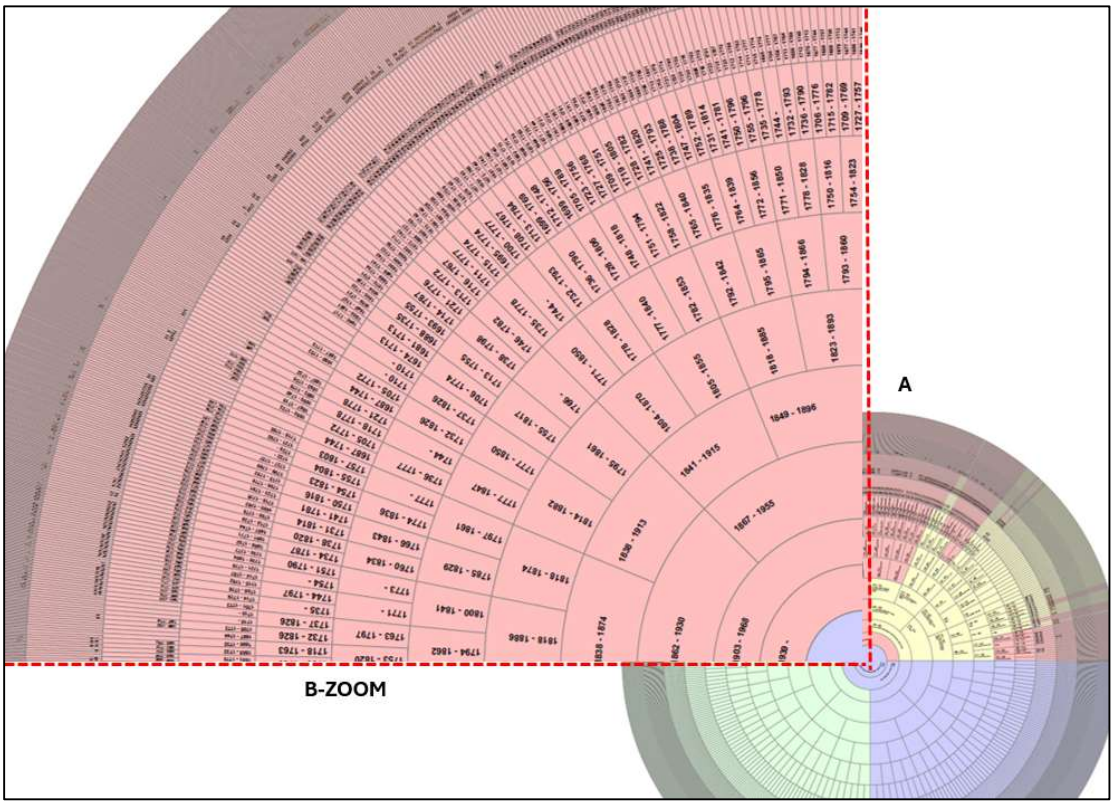


Figure 23: A fan diagram of the genealogy of Family ACM 163 is presented in A, with a zoomed-in section, B, with all names of ancestors removed (dates 1949-1643).

Below is a map showing a graphical representation (Figure 24) of the origins of the progenitors for family ACM 163, from generation 0 to 13, which roughly coincides with years 1949 to 1634, progenitors were found in many clusters in the Netherlands, present day Belgium, and more sparsely in France and Germany (B-zoom). The progenitors of ACM 163 immigrated to the Cape of Good Hope in the 16th century which is why we see the distinct clustering of the 9th to most recent generations located in the Western Cape (C-zoom), where this family still resides today.

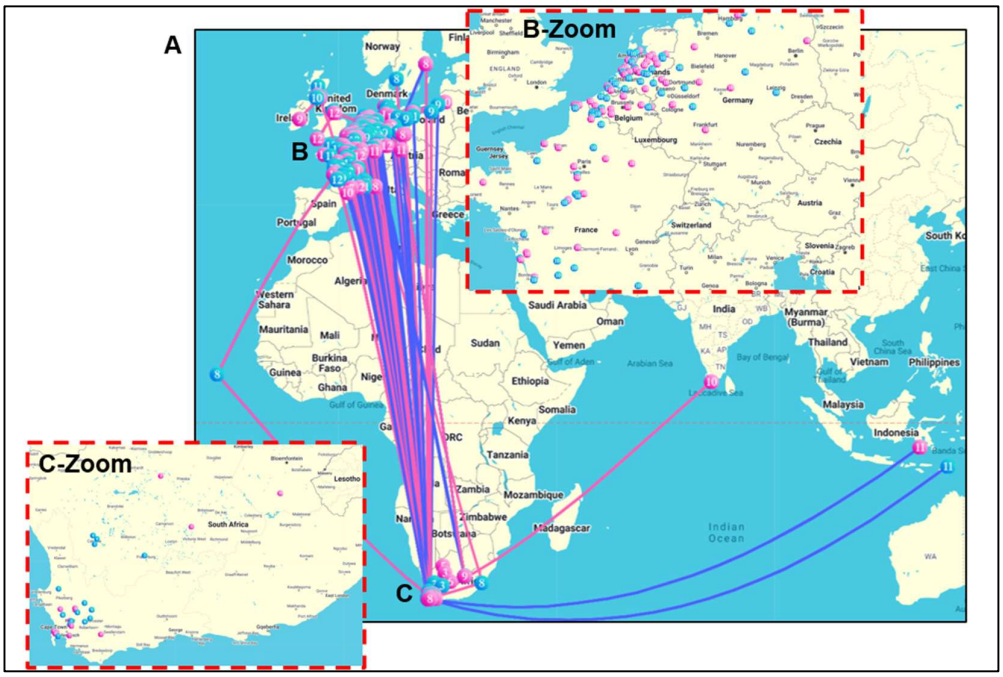


Figure 24: World map showing family ACM 163 through 12 generations and their global distribution. (A) World map with the origins of the progenitors of ACM 163. (B) Zoomed in image of where the majority of the progenitors came from in Europe (Netherlands, France and Germany). (C) Zoomed in image of the most recent generations of ACM 163 and their locations across South Africa. Pink dots represent the female ancestors, blue dots represent male ancestors, the numbers inside the dots indicate the generation of that individual, blue lines connect individuals from the paternal lineage and pink lines connect individuals from the maternal lineage (in the case of this family the paternal and maternal lineages of ACM 163.2) Generated with Rootmapper (https://rootsmapper.com/).

3.3.2 Family ACM 12 (old family)

3.3.2.1 Clinical history

Family ACM 12 is the first of three families where only the proband (ACM 12.1) was enrolled. The proband did not wish for the rest of his family to be recruited nor did he want their status to be known. ACM 12.1 was diagnosed at the age of 44 and was considered to have a mild phenotype, he also a self-identified Caucasian Afrikaner. Clinical notes recorded that the proband's father (I-1) passed away from a heart attack (Figure 25).

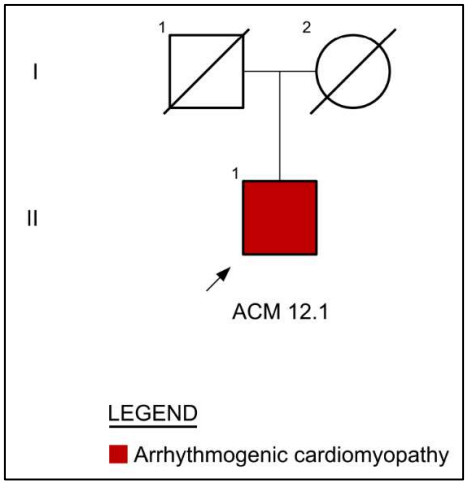


Figure 25: Pedigree of family ACM 12. The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals.

3.3.2.2 Genetics: PKP2 c.1162C>T screening

Genetic screening on the proband (ACM 12.1) found that he was heterozygous for the *PKP2* c.1162C>T variant (Figure 26).

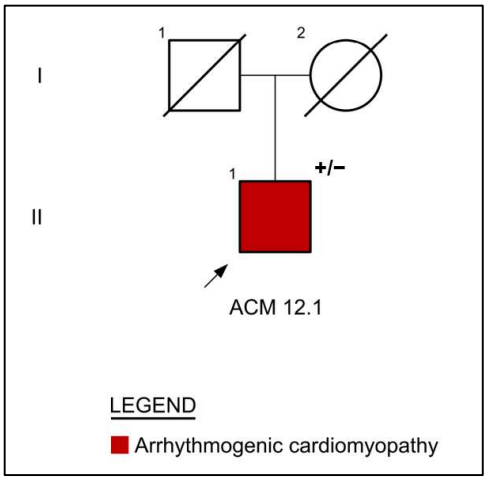


Figure 26: The genotyping results of ACM 12.The arrow indicates the proband, '+' indicates a positive genotype of the PKP2 variant while '–' indicate a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms

3.3.2.3 Haplotype analysis

The genetics results were followed up by haplotyping. The ACM12 family is an old family, and recruitment was ongoing as additional at-risk family members were being enrolled into the current study. New family members testing positive for the *PKP2* c.1162C>T variant was included in the haplotype screening. We included ACM 8 as a positive control as this individual had inherited the red diseased haploblock (132, T, 262, 250), seen in Figure 27, which contained the *PKP2* c.1162C>T variant as well microsatellites associated with disease.

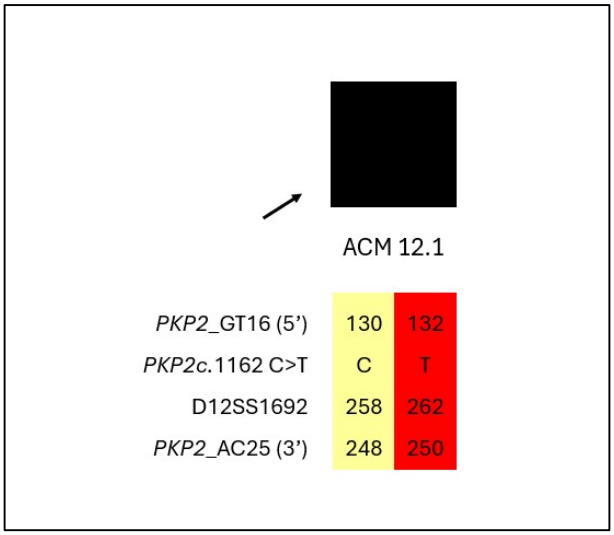


Figure 27: Haplotype analysis of family ACM 12. Circles indicate self-identified females, squares indicate self-identified males, black-filled symbols represent clinically affected individuals, symbols with periods in the centre represent individuals with the PKP2 variant but are not clinically affected, shapes with an oblique line through indicates deceased individuals, haplotypes with dotted borders represent inferred haplotypes.

3.3.2.4 Genealogy tracing

Despite only the proband, ACM 12.1 being enrolled into this study, this family is the second most genealogically complete family. This meant that there was a large amount of information that was gathered and organised, including copies of birth, marriage, and death certificates; records in churches and graveyards; old articles, letters, photographs, and documents related to citizenship and naturalisation; searches of historical records, including land records and censuses; a backwards and forwards approach, starting with the proband.

There was limited information on names, birth dates and places of residence of this ACM 12.1's family provided by the genetic counsellor. Being the only enrolled family member, the maternal and paternal lines of ACM 12.1's genealogy were extended, as he could have inherited the *PKP2* variant from either line. It was established that the proband resided in the Western Cape and he self-identified as a Caucasian Afrikaner. Once again, making this family

easier to trace as the Afrikaners are known to have well-documented records in terms of birth, death and marriage records.

Tracing began by confirming ACM 12.1's parents' names, birth dates and residences through the analysis of the 1984 South African Voters Roll. Both the paternal and maternal lineages were traced back to their progenitors, who arrived in the 17th century.

The data collected for the proband's paternal lineage showed this lineage mainly resided in the Western Cape. Evidence showed progenitors of French, Dutch, German and English origin were all found in this lineage.

The maternal lineage was traced back to German, Dutch and French ancestry. There was also evidence of consanguineous marriages in the 19th century and dating much further back in the 18th century, where cousins married. The global distributions of ancestry for family ACM 12 can be found in Figure 28 below.

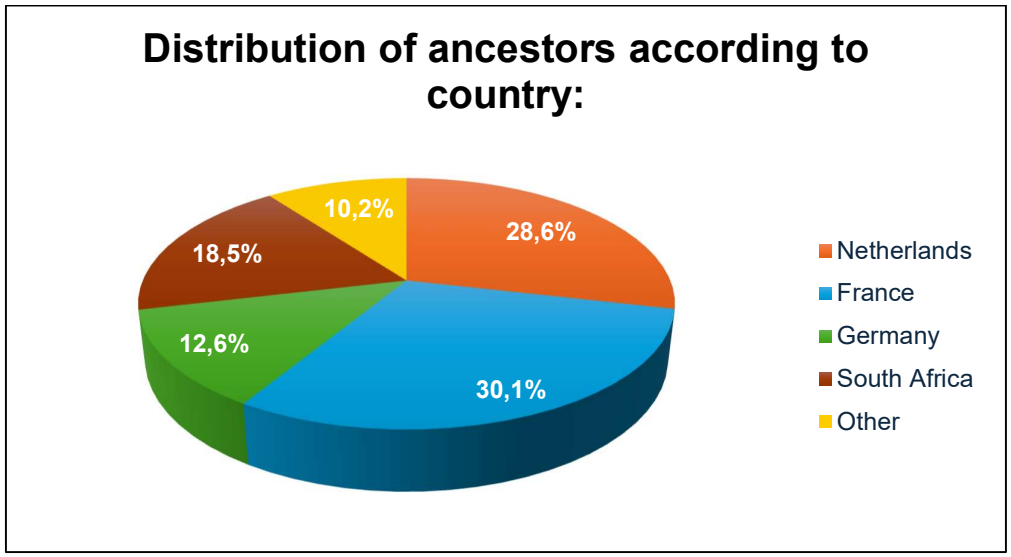


Figure 28: Distribution of ancestors according to country

All genealogical information was manually inputted, stored and analysed using Legacy v.10. This family was traced back through 12 generations (Figure 29).

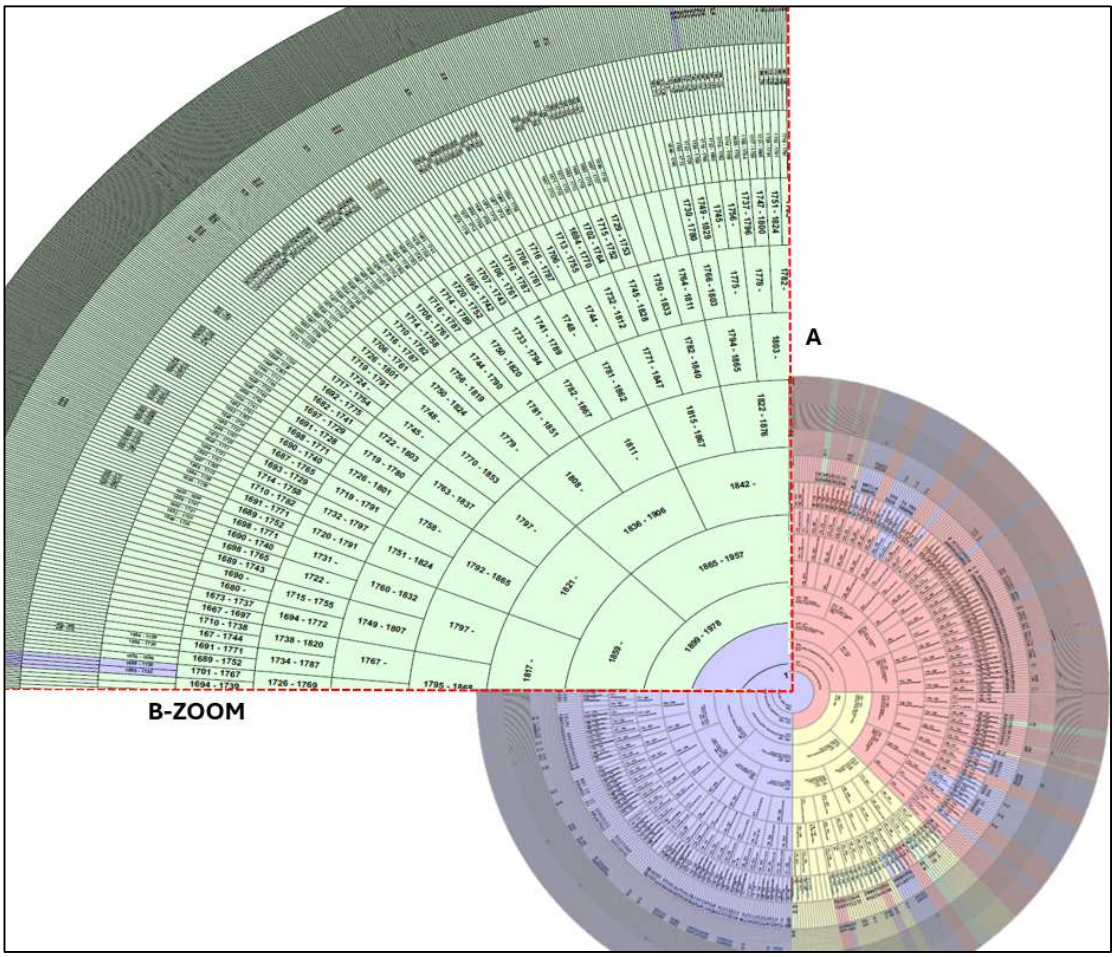


Figure 29: A fan diagram with the genealogy of Family ACM 12 presented in fan A, with a zoomed-in section, B, with all names of ancestors removed (dates 1953-1638)

When the tracing was complete for ACM 12, we combined the ancestral data with that of ACM 163 and found that we could reduce the of potential progenitors from 74 to 31, based on which progenitors were common to ACM 163 and ACM 12. Below is a graphical representation (Figure 30) of the origins of family ACM 12, from generation 0 to 12, which roughly coincides with years 1953 to 1638. Many of the ancestors that came to South Africa were found in the 10th generation, we observe a distinct clustering of progenitors in the Netherlands and present-day Belgium and were more sparsely populated in France and Germany (B-zoom). Migration of the French ancestors of ACM 12 to the Netherlands, occurred during the 11th and 12th generations, before the progenitors arrived in the Cape of Good Hope in the 16th century where majority of this family still resides today. There is also evidence that there were a few family members that moved further inland, in the 7th and 8th generations (C-zoom).

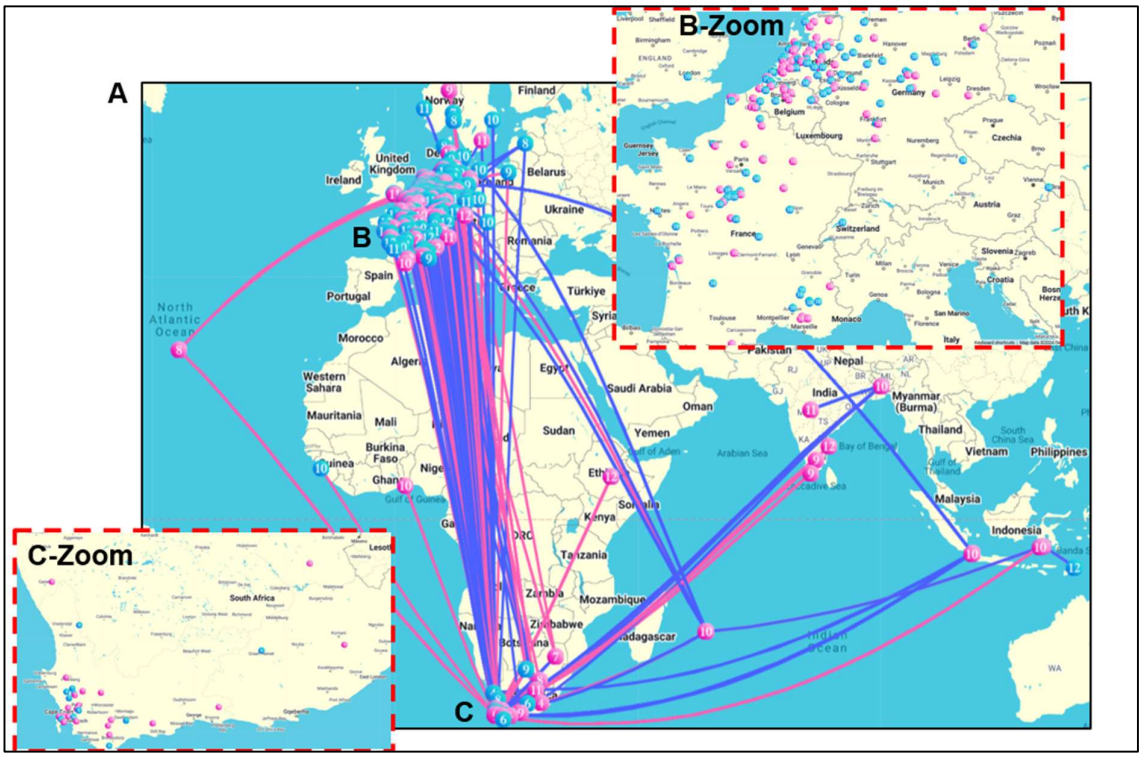


Figure 30: World map showing family ACM 12 through 12 generations and their global distribution. (A) World map with the origins of the progenitors of ACM 12. (B) Zoomed in image of where the majority of the progenitors came from in Europe (Netherlands, France and Germany). (C) Zoomed in image of the most recent generations of ACM 12 and their locations across South Africa. Pink dots represent the female ancestors, blue dots represent male ancestors, the numbers inside the dots indicate the generation of that individual, blue lines connect individuals from the paternal lineage and pink lines connect individuals from the maternal lineage (in the case of this family the paternal and maternal lineages of ACM 12.1) Generated with Rootmapper (https://rootsmapper.com/).

3.3.3 Family ACM 38 (old family)

3.3.3.1 Clinical history

The ACM 38 family had seven family members enrolled, however, only three were clinically screened and this included the proband (ACM 38.3) one of his siblings (ACM 38.5) and their mother (ACM 38.2). The proband (ACM 38.3) first presented with ACM at 9 years old and had a heart transplant at 16 years old.

The proband's brother (ACM 38.5) also had childhood onset of ACM at 8 years old and had a heart transplant the year of diagnosis. The probands mother, ACM 38.2 medical histories also show several miscarriages which may be related as she presented with a mild phenotype. Additionally, medical notes recorded she also has a mitral valve prolapse. It was also noted in the clinical report that the proband's sister (IV-2), passed away at 8 years old from SCD (suspected ACM); no records were available. The four remaining members of the family had

incomplete or unknown phenotypes (ACM 38.1, ACM 38.4, ACM 38.6 and ACM 38.7) (Figure 31).

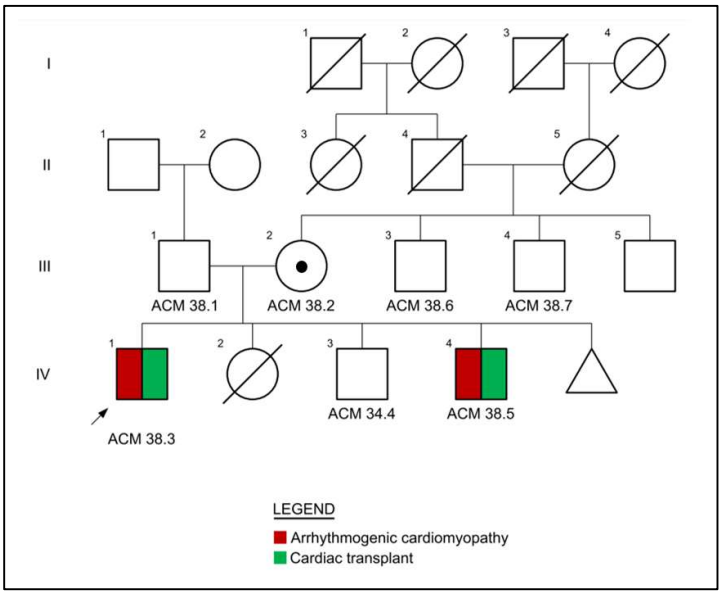


Figure 31: Pedigree of family ACM 38. The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals. The triangle indicates a spontaneous abortion

3.3.3.2 Genetics: PKP2 *c.1162C>T screening*

Despite only three members of the ACM 38 family being fully clinically screened, all seven members were genetically screened for the *PKP2* variant. The proband (ACM38.3), the affected brother (ACM38.5) and the mother (ACM38.2) were found to be heterozygous for the *PKP2* c.1162C>T variant. The remaining members, ACM 38.1, ACM 38.4, ACM 38.6 and ACM 38.7 were all negative (Figure 32). The variant was thus maternally inherited.

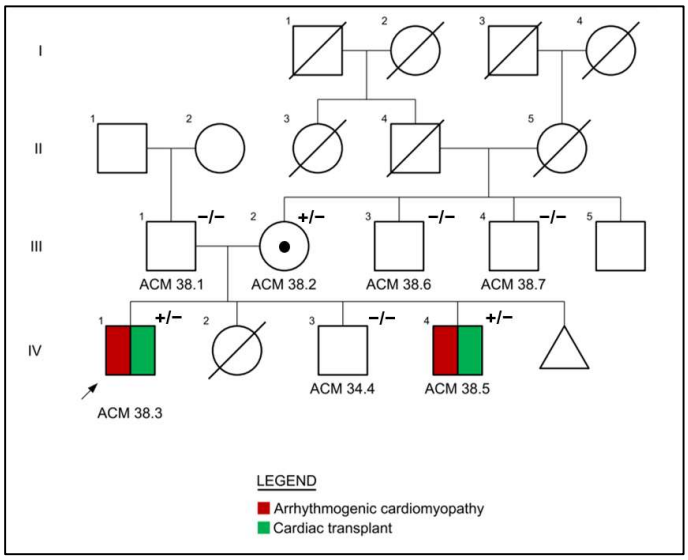


Figure 32: The genotyping results of ACM 38. The arrow indicates the proband, '+' indicates a positive genotype of the PKP2 variant while '–' indicates a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms. The triangle indicates a spontaneous abortion

3.3.3.3 Haplotype analysis

The genetics results were followed up by haplotyping. The ACM 38 family is an old family, and recruitment was ongoing as additional at-risk family members were being enrolled into the current study. New family members testing positive for the *PKP2* c.1162C>T variant was included in the haplotype screening. We included ACM 8 as a positive control as members of the ACM 38 family (ACM 38.3, ACM 38.2 and ACM 38.5) had inherited the red diseased haploblock (132, T, 262, 250) seen in Figure 31, which contained the *PKP2* c.1162C>T variant as well microsatellites associated with disease. We were able to infer the haplotypes if the grandparents, parents or sibling's DNA was available for screening. We constructed the haploblocks for all seven members of this family (Figure 33).

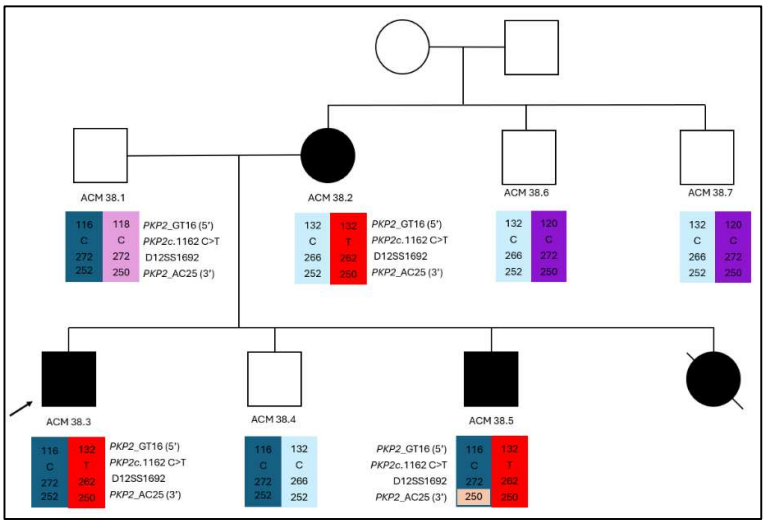


Figure 33: Haplotype analysis of family ACM 38. Circles indicate self-identified females, squares indicate self-identified males, black-filled symbols represent clinically affected individuals, symbols with periods in the centre represent individuals with the PKP2 variant but are not clinically affected, shapes with an oblique line through indicates deceased individuals, haplotypes with dotted borders represent inferred haplotypes

3.3.3.4 Genealogy tracing

Previous genealogically tracing on this family presented the most complete evidence to date as it was originally the most completely traced family in 2016 (157). The information of ACM 38 was matched to ACM 163 and ACM 12. Therefore, ACM 163, ACM 12 and ACM 38 represent the most completely traced families.

A great deal of information had to be gathered and organised, including copies of birth, marriage, and death certificates; records in churches and graveyards; old articles, letters, photographs, and documents related to citizenship and naturalisation; searches of historical records, including land records and censuses; a backwards and forwards approach, starting with the proband.

Pedigrees provided by the genetic councillor had limited information on names, dates of birth and locations pertaining to the proband, his parents and grandparents, it was noted that this family resided in the Western Cape and are a self-identified Caucasian Afrikaner family.

As the *PKP2* c.1162C>T variant was maternally inherited (ACM 38.2), we started extending both the paternal and maternal lines of ACM 38.2. First, we confirmed the full names, birth date and location for ACM 38.2 which was achieved by searching through the 1984 voter's roll. Thereafter, investigations, per individuals of the maternal and paternal lines of ACM

38.2 were conducted by investigating digitalised records (Baptism records, marriage records, death records etc) available from online databases (FamilySearch, Geni and Ancestry).

The paternal line was traced back to predominantly French Huguenot and Dutch couples. There was evidence of cousins marrying in the 18th century and thus the repetition of couples was noted in the data collected for this family. The data also showed there is English ancestry from the late 18th and early 19th centuries.

The maternal lineage of ACM 38.2 revealed English ancestry from the 19th century and German ancestry from the mid-18th century. The ancestors who traced back to the 17th century were mainly of French or Dutch origin. Marriages between cousins were recorded in the 18th century in this lineage. The global distributions of ancestry for family ACM 38 can be found in Figure 34 below.

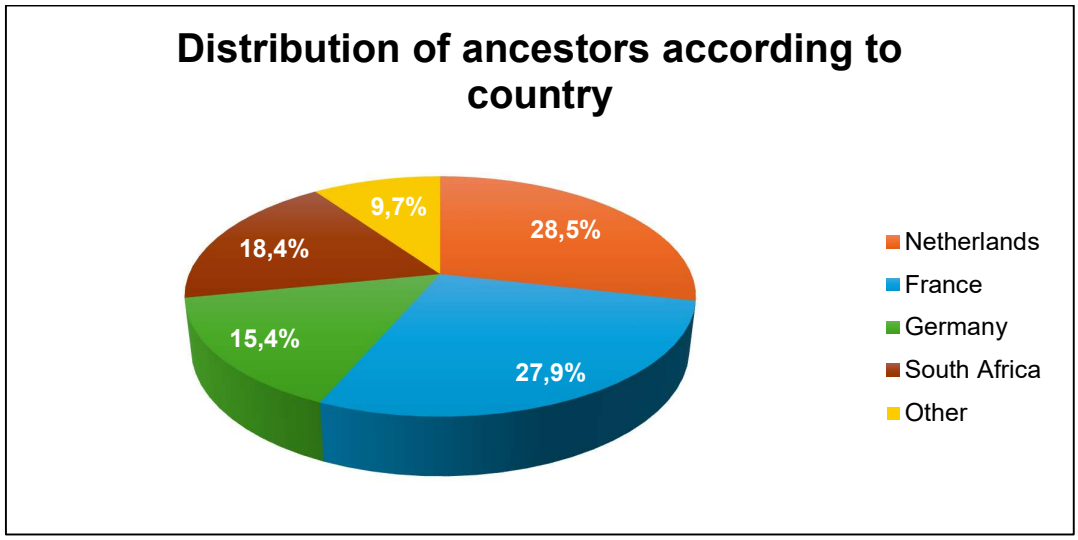


Figure 34: Distribution of ancestors according to country

All genealogical information was input manually, analysed and stored on Legacy v.10. We were able to trace this family back through 13 generations to the 17th century for both the paternal and maternal lineages of ACM 38.2 (Figure 35).

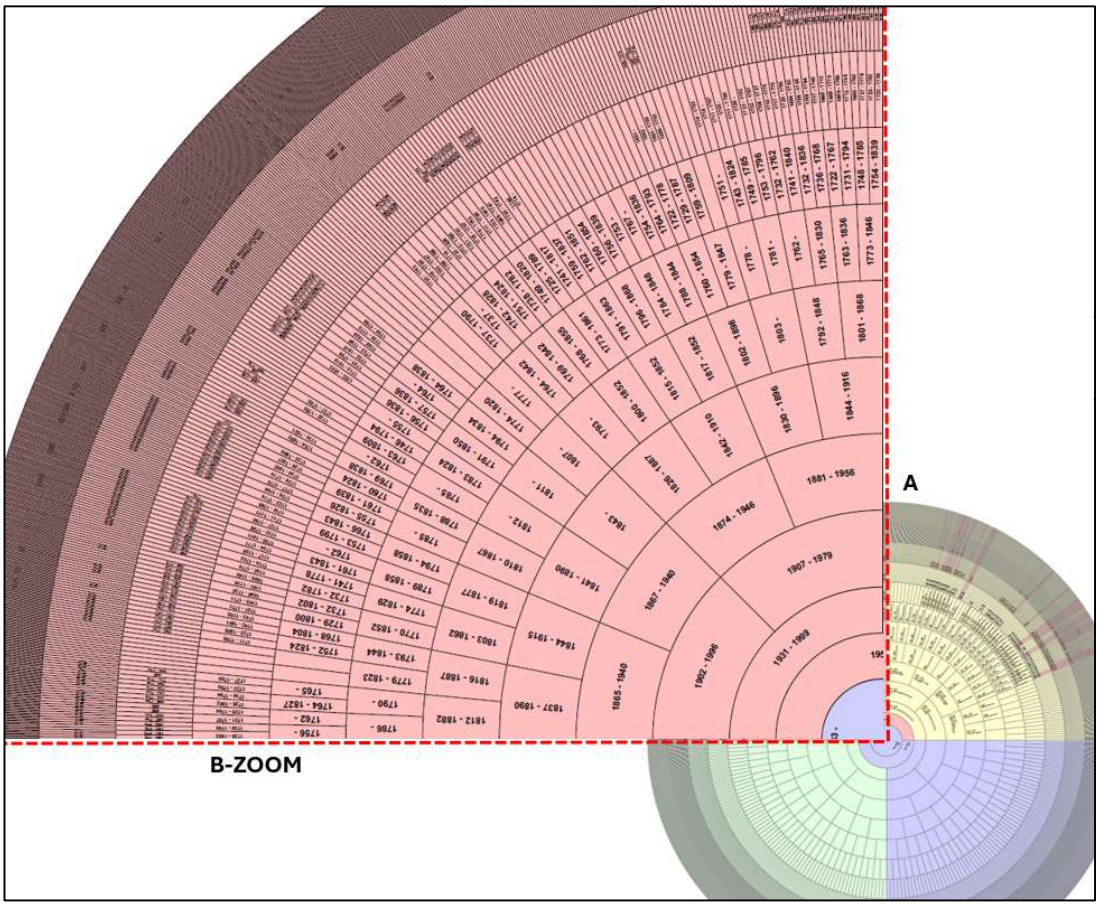


Figure 35: A fan diagram with the genealogy of Family ACM 38 presented in fan A, with a zoomed-in section, B, with all names removed of ancestors removed (dates 1983-1640)

When the genealogy tracing was complete for ACM 38, we combined their ancestral data with that of ACM 163 and ACM 12 and found that we could reduce the number of potential progenitors from 31 to 24, based on which progenitors were common to ACM 163, ACM 12 and ACM 38. Below is a graphical representation (Figure 36) of the origins of family ACM 38, from generation 0 to 13, which roughly coincides with years 1983 – 1640. Majority of the ancestors and progenitors of this family were found to be clustering in the Netherlands and present-day Belgium, these progenitors are from the 10th generation and immigrated to the Cape of Good hope in the 16th century (B-Zoom).The present day ACM 38 family still resides within the Western Cape region and have done since their progenitors arrived (C-zoom).

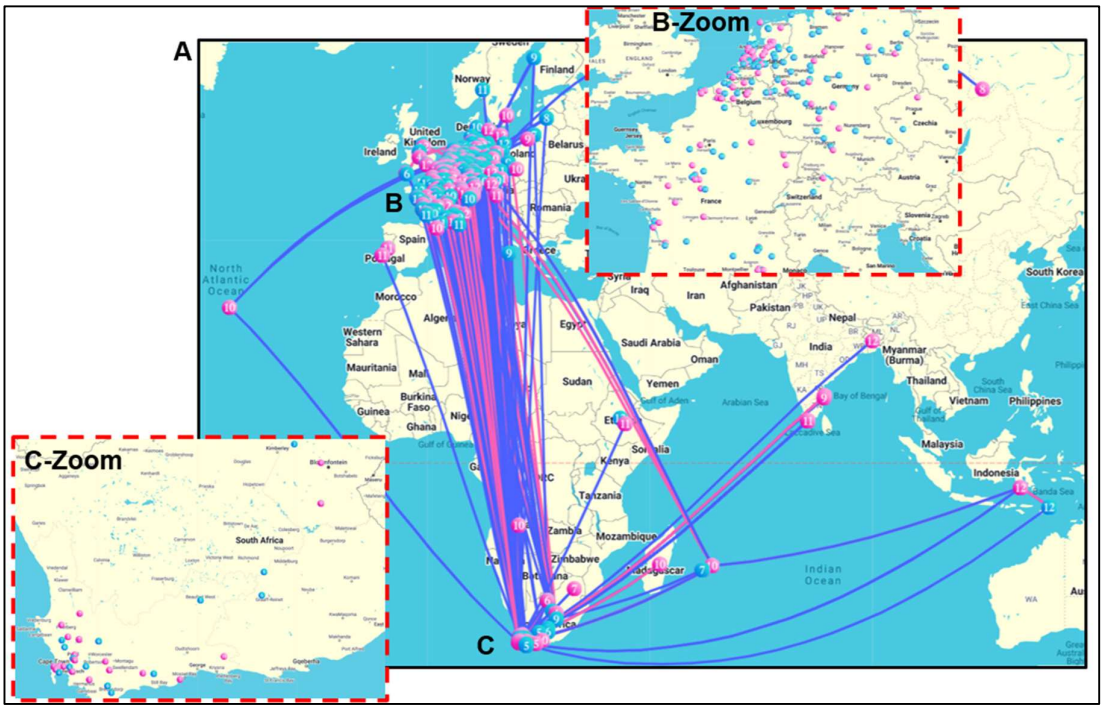


Figure 36: World map showing family ACM 38 through 12 generations and their global distribution. World map with the origins of the progenitors of ACM 38. (B) Zoomed in image of where the majority of the progenitors came from in Europe (Netherlands, France and Germany). (C) Zoomed in image of the most recent generations of ACM 38 and their locations across South Africa. Pink dots represent the female ancestors, blue dots represent male ancestors, the numbers inside the dots indicate the generation of that individual, blue lines connect individuals from the paternal lineage and pink lines connect individuals from the maternal lineage (in the case of this family the paternal and maternal lineages of ACM 38.5). Generated with Rootmapper (https://rootsmapper.com/).

3.3.4 Family ACM 154 (new family)

3.3.4.1 Clinical history

The proband (154.1), her daughter (ACM 154.2) and her parents (ACM 154.3 and ACM 154.4) were all clinically screened and self-identified as Caucasian Afrikaners. The proband was diagnosed with a mild form of ACM at 54 years old

Clinical notes for ACM 154.4, the probands mother, described her as experiencing palpitations in her twenties, however, clinical screening in 2018 concluded she was clinically unaffected. The probands daughter (ACM 154.2) and father (ACM 154.3) were both clinically unaffected (Figure 37).

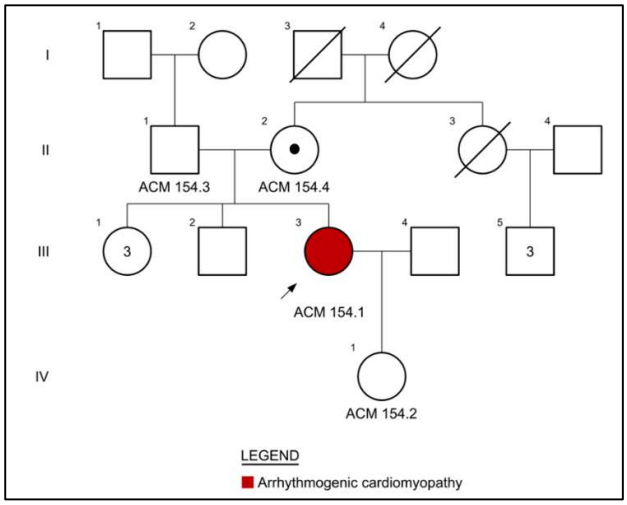


Figure 37: Pedigree of family ACM 154. The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals.

.

3.3.4.2 Genetics: PKP2 *c.1162C>T screening*

The four family members were enrolled and genetically screened. Both the proband (ACM 154.1) and her mother (ACM 154.4) were found to be heterozygous for the *PKP2* c.1162C>T variant while the proband's father (ACM 154.3) and daughter (ACM 154.2) were genotype negative (Figure 38).

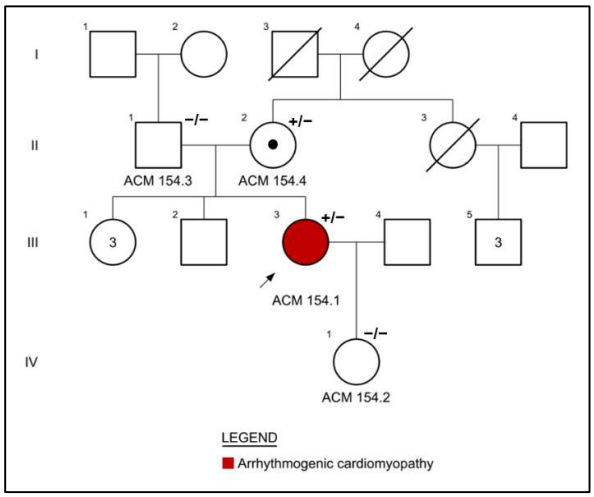


Figure 38: The genotyping results of ACM 154. The arrow indicates the proband, '+' indicates a positive genotype of the PKP2 variant while '−' indicates a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms

3.3.4.3 Haplotype analysis

The genetics results were followed up by haplotyping. Haplotyping was performed for all four of the enrolled family members by making use of the microsatellites and the variant of interest (*PKP2* c.1162C>T). The proband's (ACM 154.1) microsatellite marker peaks are shown in Figure 39 below.

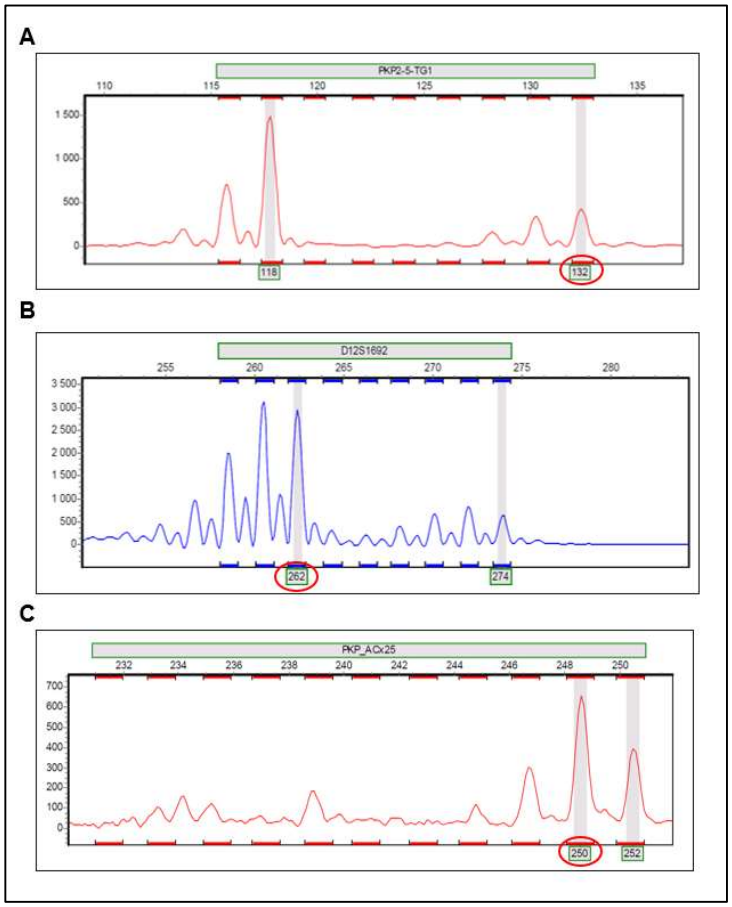


Figure 39: Microsatellite analysis of ACM 154.1. (A) PKP2_TG (132), the peak circled in red, 132, is the common haplotype peak for PKP2_TG. (B) D12S1692 (262), the peak circled in red, 262 is the common haplotype peak for D12S1692. (C) PKP2_AC (250), the peak circled in red, 250 bps, is the common haplotype peak for PKP2_AC.

The microsatellite analysis results were combined and used to construct haplotypes for the proband ACM 154.1, her daughter (ACM 154.2), and her father and mother respectively (ACM 154.3 and ACM 154.4) (Figure 40).

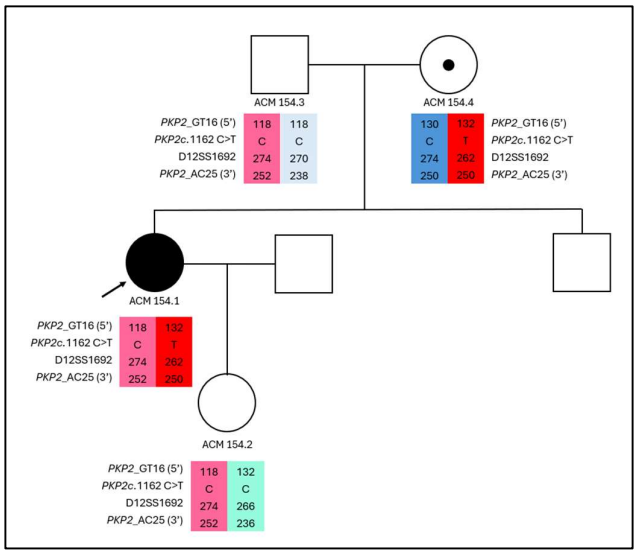


Figure 40: Haplotype analysis of family ACM 154. Circles indicate self-identified females, squares indicate self-identified males, black-filled symbols represent clinically affected individuals, symbols with periods in the centre represent individuals with the PKP2 variant but are not clinically affected, shapes with an oblique line through indicates deceased individuals, haplotypes with dotted borders represent inferred haplotypes.

The proband (ACM 154.1) was shown to have inherited the disease-associated haploblock, shaded in red, from her mother (ACM 154.4) as she too had the red shaded, disease-associated haploblock. ACM 154.2 and ACM 154.3 did not have this haploblock.

3.3.4.4 Genealogy tracing

Genealogy tracing of this newly enrolled family was made easier as the probands mother (ACM 154.4) had inherited records of the family history and a rough genealogy of her ancestors; this was provided to the family's genetic councillor. These records were thoroughly investigated and verified by gathering, comparing and organising all the information, including copies of birth, marriage, and death certificates; records in churches and graveyards; old articles, letters, photographs, and documents related to citizenship and naturalisation; searches of historical records, including land records and censuses; a backwards and forwards approach, starting with the proband.

The full names, birth dates and locations of the probands parents (ACM 154.3 and ACM 154.4), were confirmed by consulting the 1984 voter's roll as we knew this family resided in the Western Cape region. Thereafter, all information pertaining to this family in the provided

genealogy was confirmed through the local repositories of this region using online databases (FamilySearch, Geni and Ancestry).

As the *PKP2* c.1162C>T variant was maternally inherited, we started extending both the paternal and maternal lines of ACM154.4

The paternal line of ACM 154.4 was traced back to the 17th century with ancestors identified as French Huguenots, Dutch and German progenitors of South Africa. There were multiple duplicates identified due to consanguineous marriages between cousins in the 17th and 18th centuries.

The maternal lineage of ACM 154.4 was traced back to German ancestry in the mid-18th century. However, other lines extended further back into the 17th century with French, Dutch and German ancestry identified.

The global distributions of ancestry for family ACM 154 can be found in Figure 41 below.

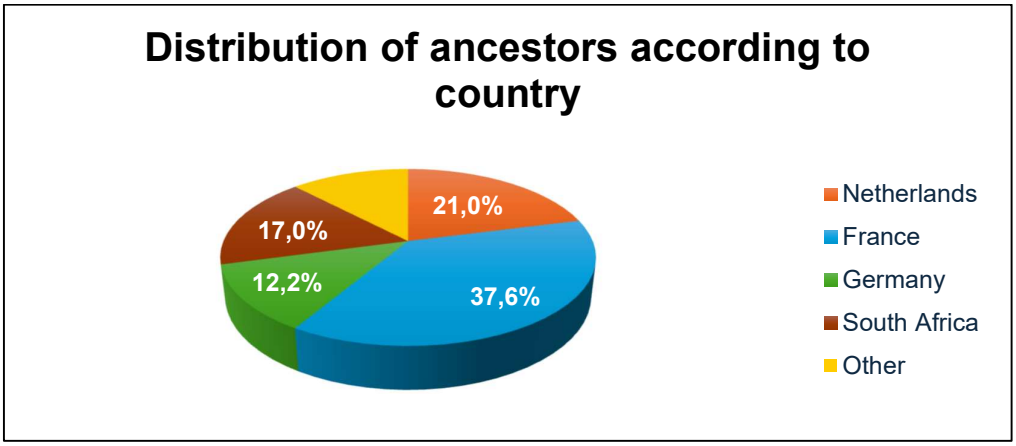


Figure 41: Distribution of ancestors according to country

Both the paternal and maternal lineages of ACM 154.4 were traced back through 13 generations to the 17th century and the progenitors of this family (Figure 42).

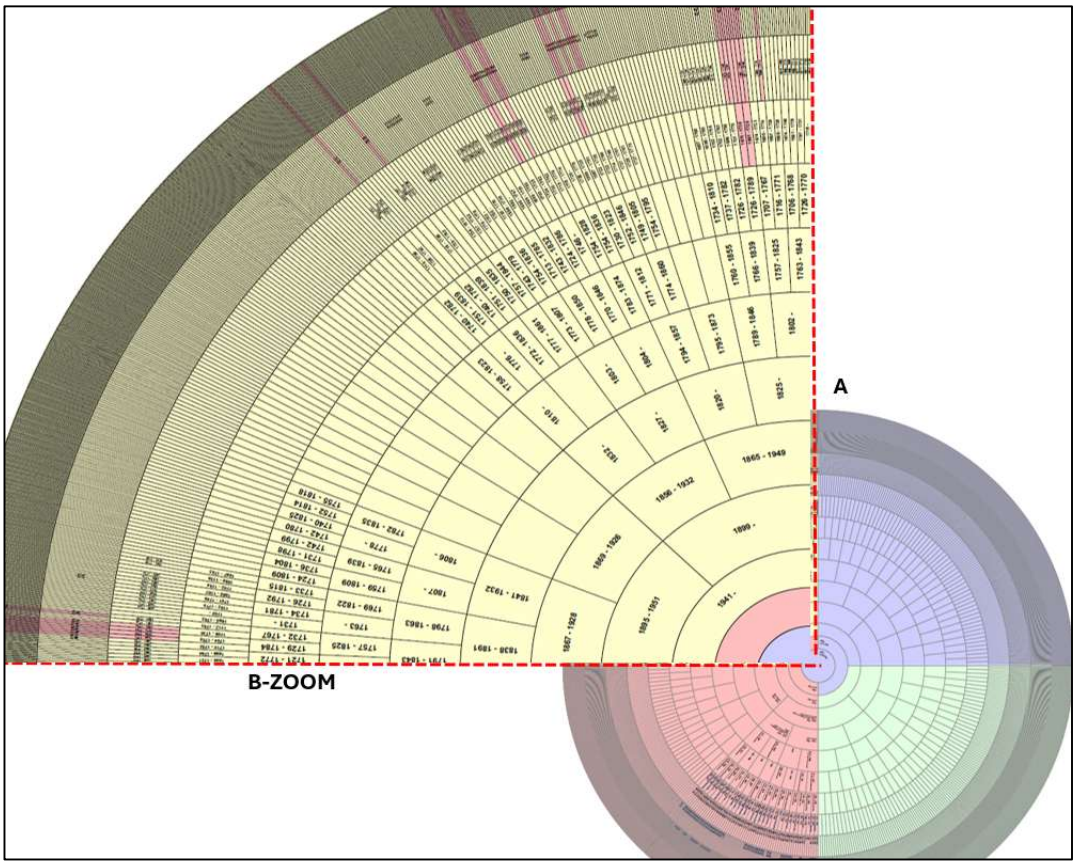


Figure 42: A fan diagram of the genealogy of Family ACM 154 is presented in A, with a zoomed-in section, B, with all names of ancestors removed (dates 2004-1638)

All genealogical information was manually inputted, analysed and stored on Legacy v.10. We combined the ancestral data of ACM 163, ACM 12, ACM 38 and ACM 154 and found that we could reduce the number of potential progenitors from 24 to 17, based on which progenitors were common to these four families. Below is a graphical representation (Figure 43) of the origins of family ACM 154, from generation 0 to generation 13, which roughly coincides with years 2004 – 1638. Most of the ancestors that immigrated to South Africa were found in the 10th generation. Upon comparison to the maps of the previous three families, we observed more ancestors in the French regions for ACM 154 and a less dense cluster in the Netherlands in the 10th generation (B-zoom). The present generation of ACM 154 still resides in the Western Cape; however, they did move further along the East Coast in recent generations (4th and 5th) (C-zoom).

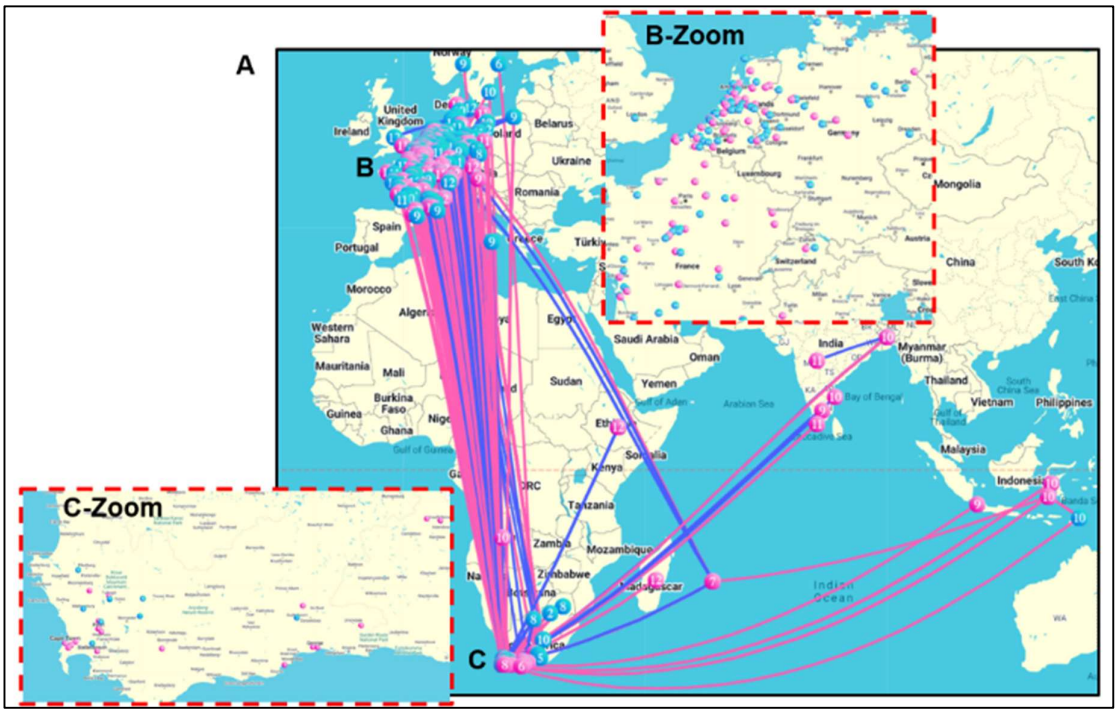


Figure 43: World map showing family ACM 154 through 12 generations and their global distribution. (A) World map with the origins of the progenitors of ACM 154. (B)Zoomed in image of where most of the progenitors came from in Europe (Netherlands, France and Germany)

3.3.5 Family ACM 19 (old family)

3.3.5.1 Clinical history

This family is the second largest enrolled family, with 13 family members enrolled in the IMHOTEP registry. They form part of the six families which were previously investigated (157) and this family self-identifies as Caucasian Afrikaners.

The proband, ACM 19.2, had childhood onset ACM at the age of 13 and by the age of 21 required a heart transplant. Similarly, her brother ACM 19.1 presented with ACM at the age of 14 and required a heart transplant a year after he was diagnosed. Unfortunately, he passed away due to complications. The proband's mother (ACM 19.6), presented with a very mild phenotype, and the probands cousin (ACM 19.13) presented with symptoms of ACM at the age of 32. The remaining eight members of this family were clinically screened and found to be asymptomatic (Figure 44).

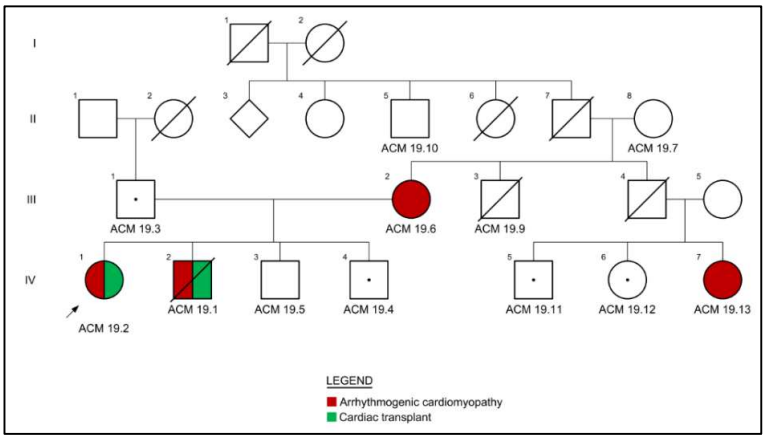


Figure 44: Pedigree of family ACM 19.The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals.

3.3.5.2 Genetics: PKP2 c.1162C>T screening

Genetic analysis of the proband revealed ACM 19.2 she was heterozygous for the*PKP2* variant c.1162C>T (Arg388Trp). Genetic evidence showed that the *PKP2* c.1162C>T variant was maternally inherited through the proband's mother (ACM 19.6). Even though we had no records of the proband's grandfather (II:7), we could infer that the *PKP2* c.1162C>T was inherited from the her grandfather as the proband's living grandmother (ACM19.7) tested negative for the *PKP2* c.1162C>T variant and (2) the proband's deceased uncle (III:4) (whom we had no data on) had three children (ACM19.11, ACM19.12 and ACM19.13) who had also inherited the *PKP2* c.1162C>T variant. The probands brothers (ACM 19.1 and ACM 19.4) were both heterozygotes for the *PKP2* variant. Both the probands father, ACM 19.3 and her remaining brother were negative for the *PKP2* variant (Figure 45).

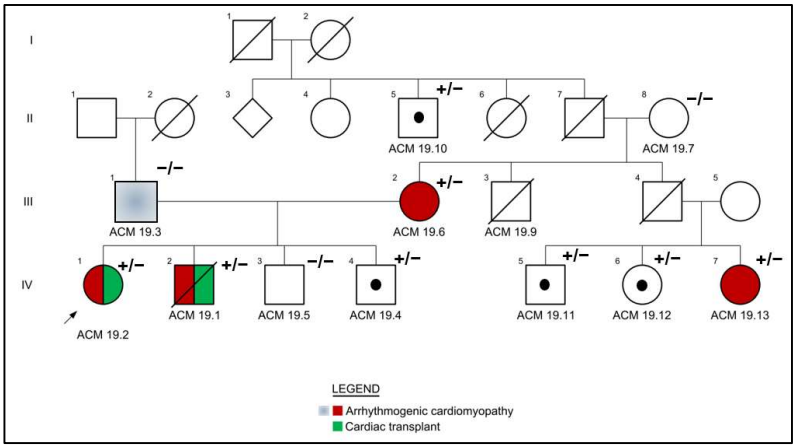


Figure 45: The genotyping results of ACM 19. The arrow indicates the proband, '+' indicates a positive genotype of the PKP2 variant while '–' indicates a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms

3.3.5.3 Haplotype analysis

The genetics results were followed up by haplotyping. The ACM19 family is an old family, and recruitment was ongoing as additional at-risk family members were being enrolled into the current study. New family members testing positive for the *PKP2* c.1162C>T variant was included in the haplotype screening. We included ACM 8 as a positive control as this individual had inherited the red diseased haploblock (132, T, 262, 250), seen in Figure 46, which contained the *PKP2* c.1162C>T variant as well microsatellites associated with disease. We were able to infer the haplotypes if the grandparents, parents or sibling's DNA was available for screening. Thus far, nine family members have the disease-associated haploblock, represented by the red block in Figure 46 below.

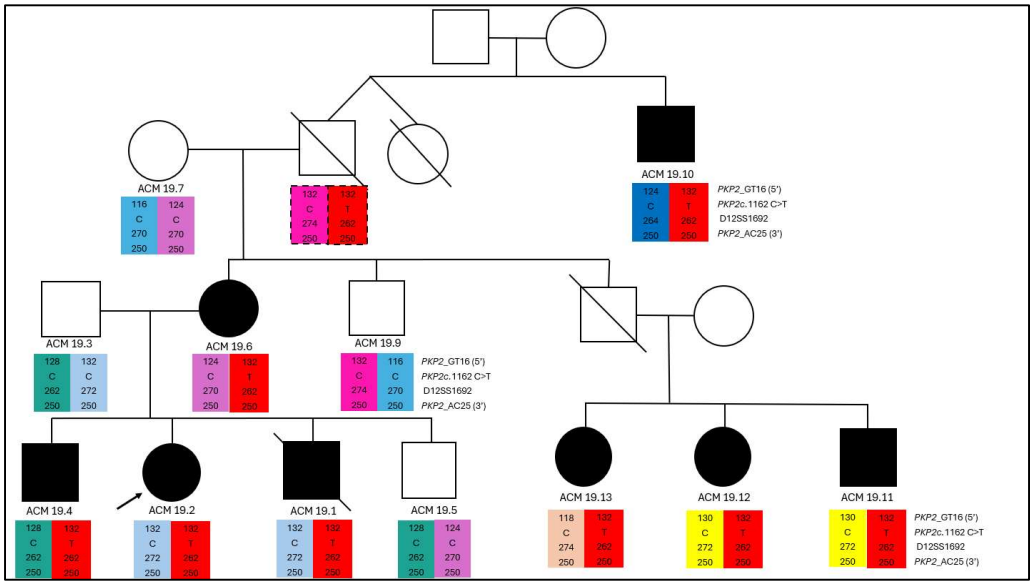


Figure 46: Haplotype analysis of family ACM 19. Circles indicate self-identified females, squares indicate self-identified males, black-filled symbols represent clinically affected individuals, symbols with periods in the centre represent individuals with the PKP2 variant but are not clinically affected, shapes with an oblique line through indicates deceased individuals, haplotypes with dotted borders represent inferred haplotypes.

3.3.5.4 Genealogy tracing

Previous work done to trace this family meant there was substantial knowledge on the genealogy of this family (157). However, due to software changes and the loss of the original genealogical records, we reconstructed the genealogy from the beginning. A great deal of information had to be gathered and organised: information on names, dates of birth and locations of the proband, her parents and grandparents from the genetic councillor, tracing began. The family self-identified as Caucasian Afrikaner and t was found that most of the ACM 19 family resided in the Northwest region of South Africa.

Genetic evidence showed that the *PKP2* c.1162C>T variant was maternally inherited through the proband's mother (ACM 19.6) and grandfather (II:7) therefore, the paternal line of ACM 19.6 was extended. Records show ACM 19.6's parents resided in the Free State of South Africa, while the paternal lineages parents were born in the Eastern Cape.

All information provided on this family's pedigree by the genetic counsellor (full names, birthdates and locations) were all verified by consulting the 1984 voter's role. Thereafter, after confirming this for ACM 19.6's father and his parents we began searching online databases (FamilySearch, Geni, Ancestry) for any information about ACM 19.6's paternal line. We were able to trace back to the 17th century to French, German and Dutch ancestry, with many of these lines residing in the old Cape Province region. Consanguineous marriages were also found in ACM 19, resulting in several duplicates being identified. The global distributions of ancestry for family ACM 19 can be found in Figure 47 below.

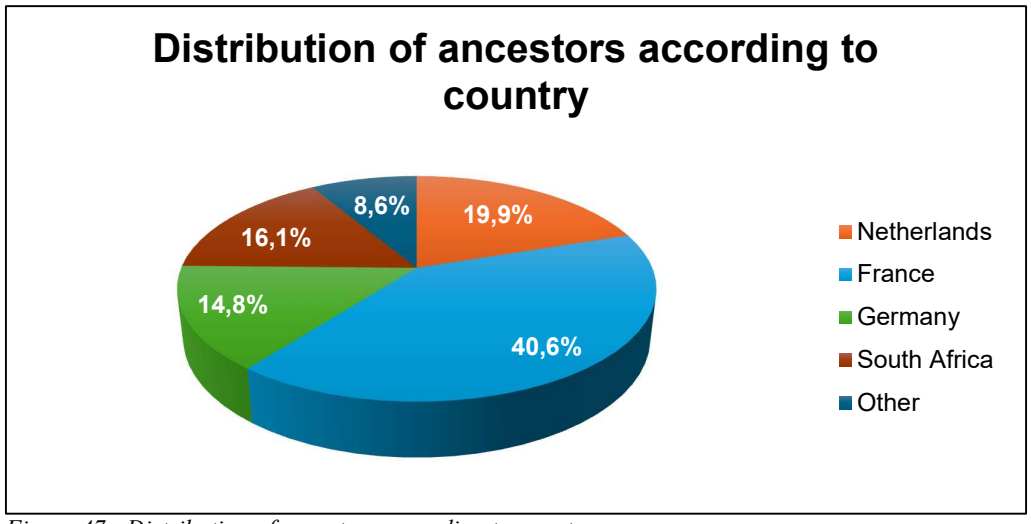


Figure 47: Distribution of ancestors according to country:

From the data collected on this family, we were able to create a genealogy of 12 generations tracing back all the way to the 17th century, to well-known progenitors of South Africa. All genealogical information was manually inputted, analysed and stored on Legacy v.10 (shown in Figure 48).

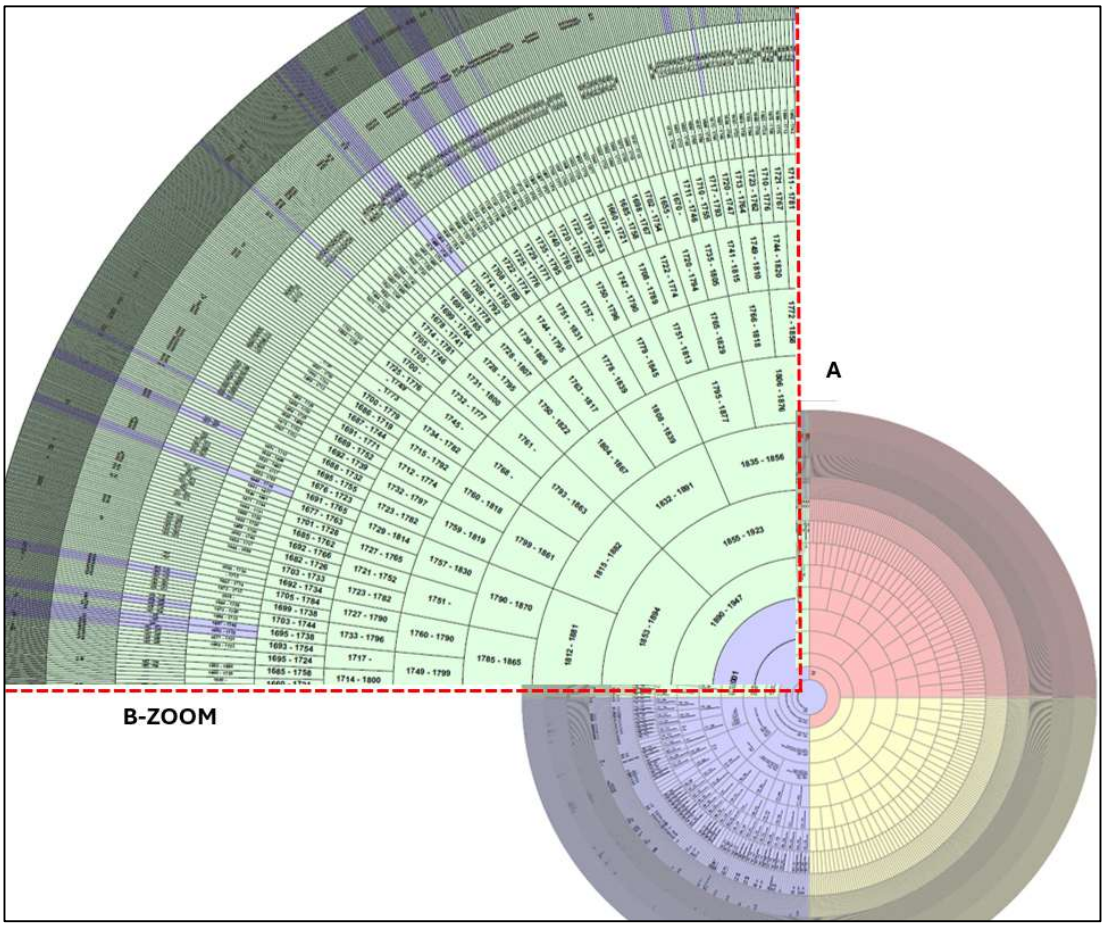


Figure 48: A fan diagram with the genealogy of Family ACM 19 presented in fan A, with a zoomed-in section, B, with all names removed of ancestors removed (dates 1956-1585)

We combined the ancestral data of ACM 163, ACM 12, ACM 38, ACM 154 and ACM 19 and found that we could reduce the number of potential progenitors from 17 to14, based on which progenitors were common between these five families. Below is a graphical representation (Figure 49) of the origins of family ACM 19, from generation 0 to 12, which roughly coincides with years 1956 to 1585. Majority of the ancestors that came to South Africa were found in the 10th generation and tended to cluster in the Netherlands and France (B-Zoom). This family has remained in the Western Cape since the arrival of their ancestors in the 16th century (C-zoom).

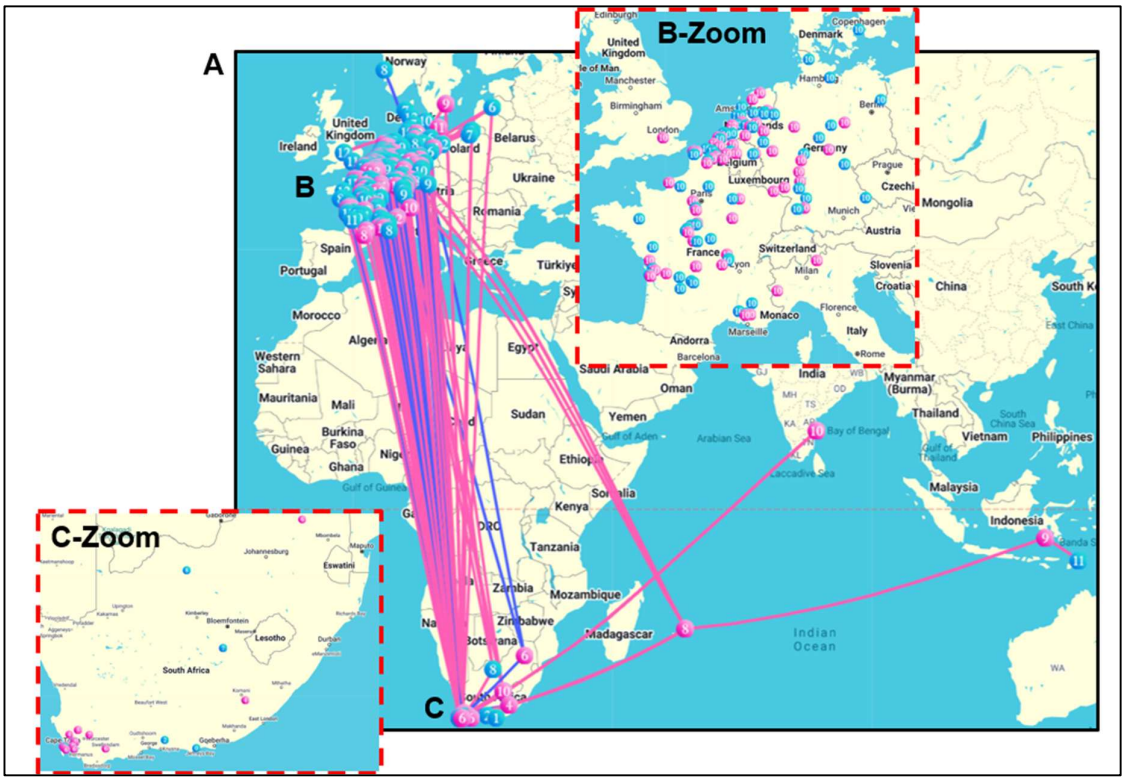


Figure 49: World map showing family ACM 19 through 12 generations and their global distribution. (A) World map with the origins of the progenitors of ACM 19. (B) Zoomed in image of where the majority of the progenitors came from in Europe (Netherlands, France and Germany). (C) Zoomed in image of the most recent generations of ACM 19 and their locations across South Africa. Pink dots represent the female ancestors, blue dots represent male ancestors, the numbers inside the dots indicate the generation of that individual, blue lines connect individuals from the paternal lineage and pink lines connect individuals from the maternal lineage (in the case of this family the paternal and maternal lineages of ACM 38.5). Generated with Rootmapper (https://rootsmapper.com/).

3.3.6 Family ACM 8 (old family)

3.3.6.1 Clinical history

Since the initial enrolment of this family in 2016, there have been additional members screened and added to the IMHOTEP registry. The total number of participants for this family has doubled from five to ten since the initial investigations to identify a common founder for the *PKP2* c.1162C>T variant (157).

The proband, ACM 8.3, presented with severe ACM at the age of 12 years old and received a heart transplant at 23 years old. Her brother (ACM 8.4) was also diagnosed with severe ACM in 2001 but died at the age of 14 due to complications of the disease. Screening of the proband's mother (ACM 8.2) found her to be clinically affected, however, her phenotype was mild. The father (ACM 8.1) was noted to be clinically unaffected. The probands remaining two siblings, ACM 8.9 and 8.11, also showed no signs of ACM. The proband's aunt (ACM

8.5) and her immediate family (ACM 8.6, 8.7 and 8.8) were all clinically screened and showed no signs of ACM (Figure 50).

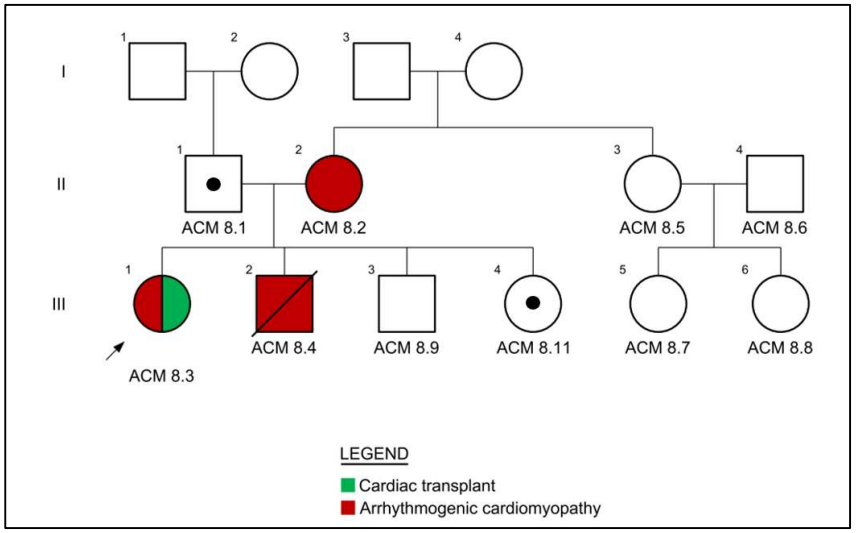


Figure 50: Pedigree of family ACM 8. The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals.

3.3.6.2 Genetics: PKP2 c.1162C>T screening

Genetic screening found the proband (ACM 8.3) and her brother (ACM 8.4) to be homozygous for the *PKP2* c.1162C>T variant. Both parents (ACM 8.1 and ACM 8.2), were found to be heterozygous for the *PKP2* variant. The proband's sibling (ACM 8.11) was found to be heterozygous for the PKP2 c.1162C>T variant but as had been previously indicated, showed no signs of disease. The other sibling, ACM 8.9, was found to be negative for the variant. The proband's aunt (ACM 8.5), her husband (ACM 8.6) and two daughters (ACM 8.7 and ACM 8.8) are all genotype negative (Figure 51) for the variant.

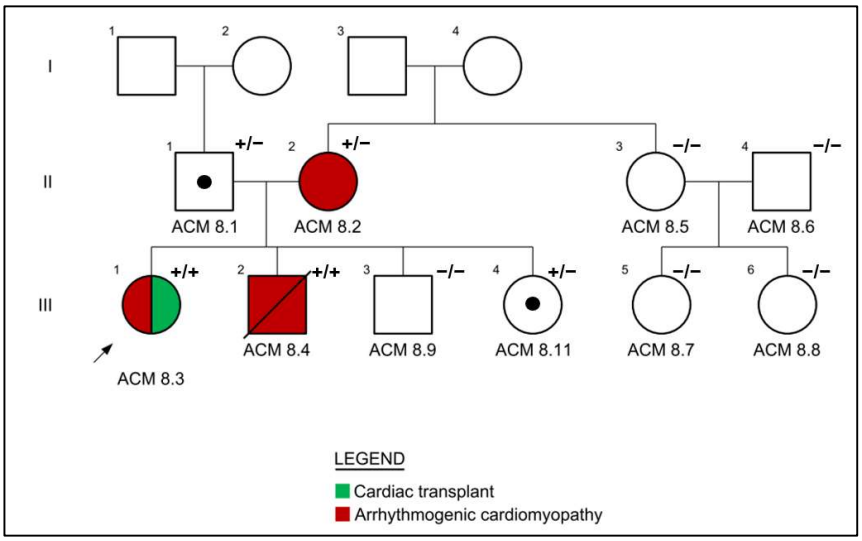


Figure 51: The genotyping results of family ACM 8. The arrow indicates the proband, '+' indicates a positive genotype of the PKP2 variant while '–' indicates a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms

3.3.6.3 Haplotype analysis

The genetics results were followed up by haplotyping. The ACM 8 family is an old family, and recruitment was ongoing as additional at-risk family members were being enrolled into the current study. New family members testing positive for the *PKP2* c.1162C>T variant was included in the haplotype screening. We used this proband as a positive control as this individual had inherited the red diseased haploblock (132,T,262,250), seen in Figure 52, which contained the *PKP2* c.1162C>T variant as well microsatellites associated with disease. We were able to infer the haplotypes if the grandparents, parents or sibling's DNA was available for screening Haplotyping was performed for all the enrolled family members by making use of the microsatellites and the variant of interest (*PKP2* c.1162C>T). The proband's (ACM 8.3) microsatellite marker peaks are shown in Figure 52 below.

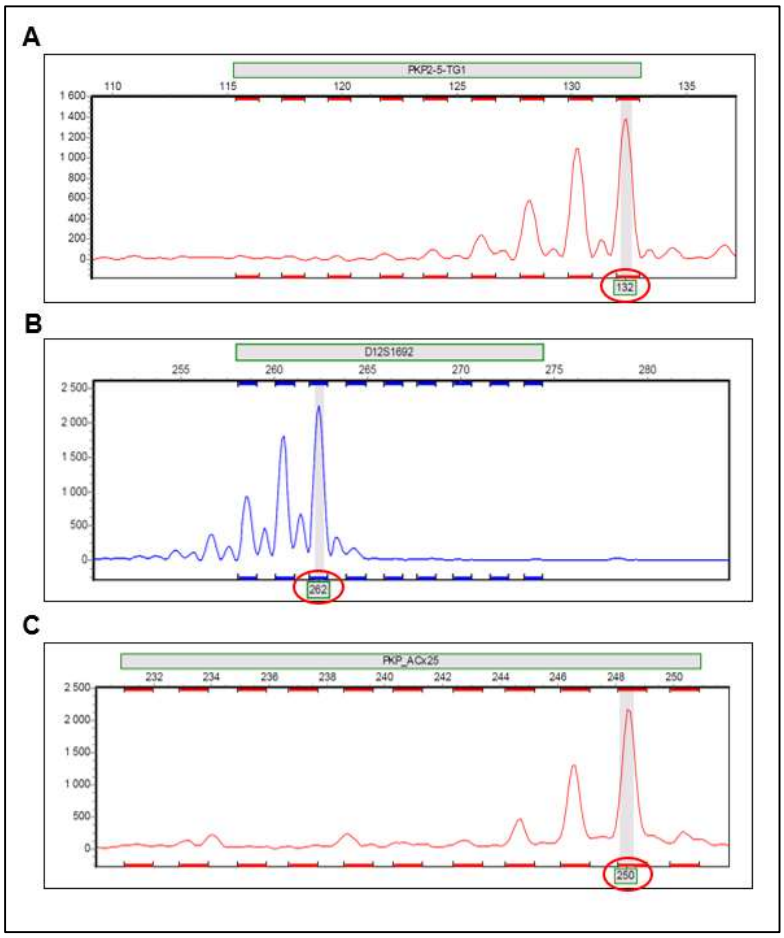


Figure 52: Microsatellite analysis of ACM 8.3. (A) PKP2_TG (132), the peak circled in red, 132, is the common haplotype peak for PKP2_TG. (B) D12S1692 (262), the peak circled in red, 262 is the common haplotype peak for D12S1692 (C) PKP2_AC (250), the peak circled in red, 250 bps, is the common haplotype peak for PKP2_AC.

Four of the 10 had the disease-associated haploblock, shown in the red block in Figure 53 below.

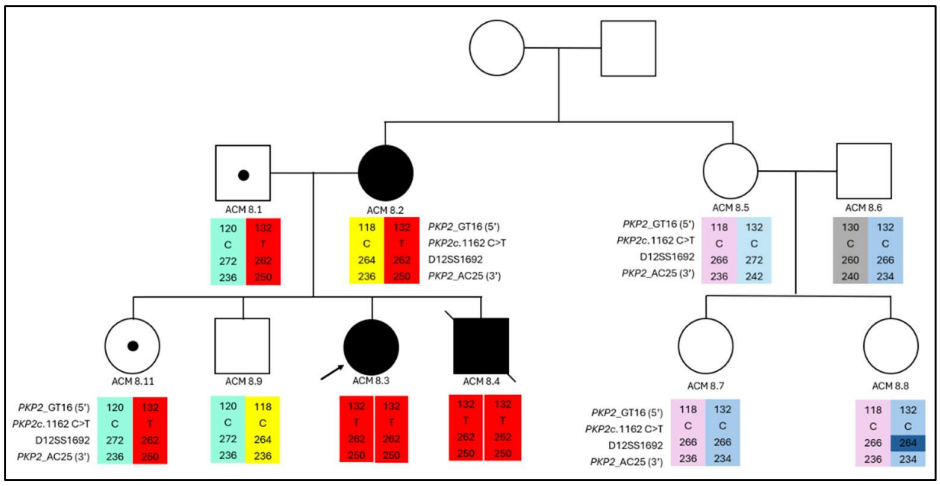


Figure 53: Haplotype analysis of family ACM 8. Circles indicate self-identified females, squares indicate self-identified males, black-filled symbols represent clinically affected individuals, symbols with periods in the centre represent individuals with the PKP2 variant but are not clinically affected, shapes with an oblique line through indicates deceased individuals, haplotypes with dotted borders represent inferred haplotypes.

3.3.6.4 Genealogy tracing

Even though this family had been previously investigated, not much data was available therefore, further research and data collection was required for this family.

As genetics indicated that the variant was paternally (ACM8.1) and maternally (ACM 8.2) inherited, we extended both lines genealogically. The pedigree provided by the genetic counsellor provided us with the birthdates, residences and full names of the proband and the proband's parents, we were able to confirm the parents' information by consulting the 1984 voter's role. We identified that this family resides in the Western Cape and readily available information was used to search the online databases (FamilySearch, Geni, Ancestry) for additional resources on this family.

Fist starting with ACM 8.1, we were able to trace his paternal line back to the late 18th century to English ancestry. Majority of this lineage resided in the Western Cape, however, there was migration up the coast to parts of the Eastern Cape. The maternal line of ACM 8.1 had a few lines missing due to limited records of full names, birthdates and death dates for several members of this extended linage. Despite this, we were still able to trace back through a number of lineages to the early 17th century which provided evidence of French, Dutch and German ancestry. Research also showed marriages between cousins in the early 18th century.

The maternal line of ACM 8.2 had several lineages missing; however, most of the traced lines were ancestors residing in the Western Cape and consanguineous marriages were evident

from the start of the 19th century. This lineage was traced back to French Huguenot and Dutch progenitors. The global distributions of ancestry for family ACM 8 can be found in Figure 54 below.

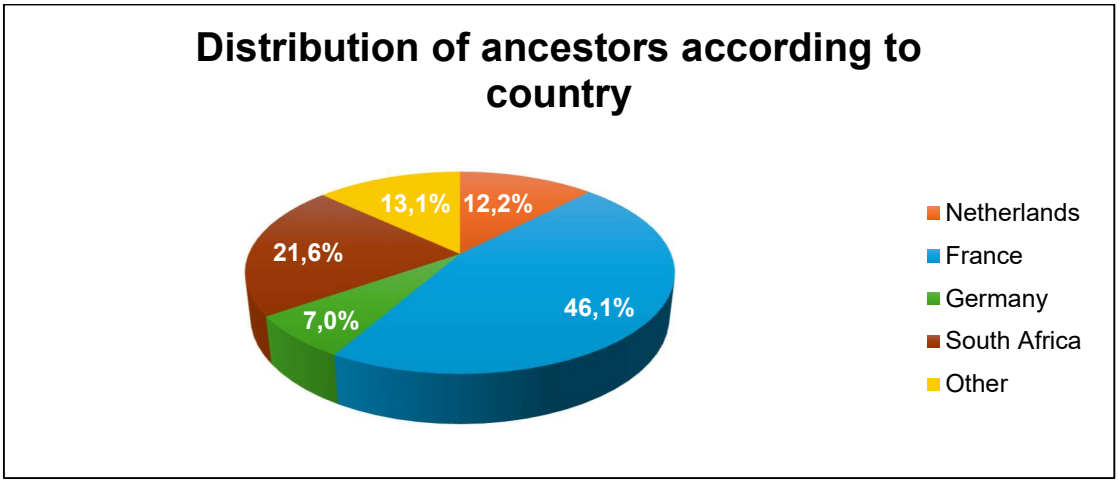


Figure 54: Distribution of ancestors according to country

All genealogical information was manually inputted, analysed and stored on Legacy v.10. The combined results of this family allowed us to trace back through 12 generations (Figure 55).

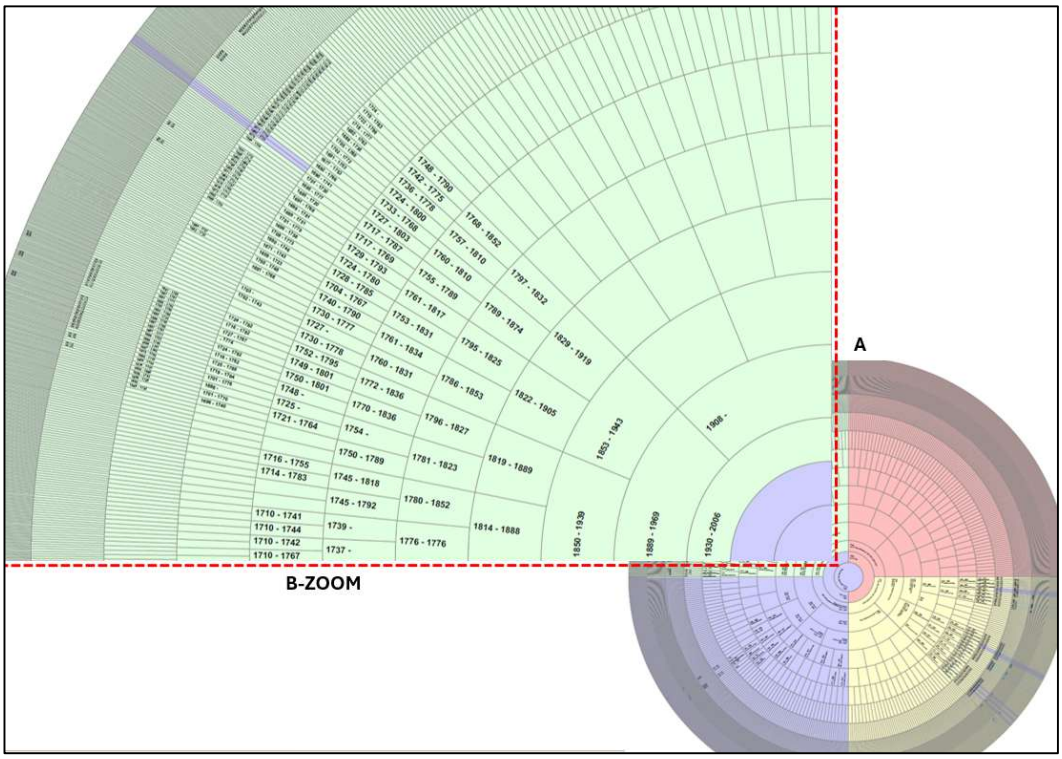


Figure 55: A fan diagram of the genealogy of Family ACM 8 is presented in A, with a zoomed-in section, B, with all names of ancestors removed (dates 1988-1643)

When the genealogical tracing was complete, we combined the ancestral data of ACM 163, ACM 12, ACM 38, ACM 154, ACM 19 and ACM 8 and found that we could not reduce the number of potential progenitors. The total of common progenitors remained at 14 for these seven families.

Below is a graphical representation (Figure 56) of the origins of family ACM 8, from generation 0 – 12, which roughly coincides with years 1988 – 1643. Many of the ancestors that came to South Africa were found in the 10th generation and clusters can be seen in certain locations in France as well as a more dense clustering in the Netherlands. Ancestors of this family were more sparely located in Germany (B-Zoom). The progenitors of the ACM 8 family immigrated to the Cape of Good Hope in the 16th century, and they have remained in the Western Cape region as seen by the dense cluster in C-zoom. However, we did identify a few more recent family members (8th generation) resided in land. This was likely due to the probands father's lineage having ancestors from England and therefore only arriving in South Africa much later (18th century) and at the port Qqeberha (C-zoom).

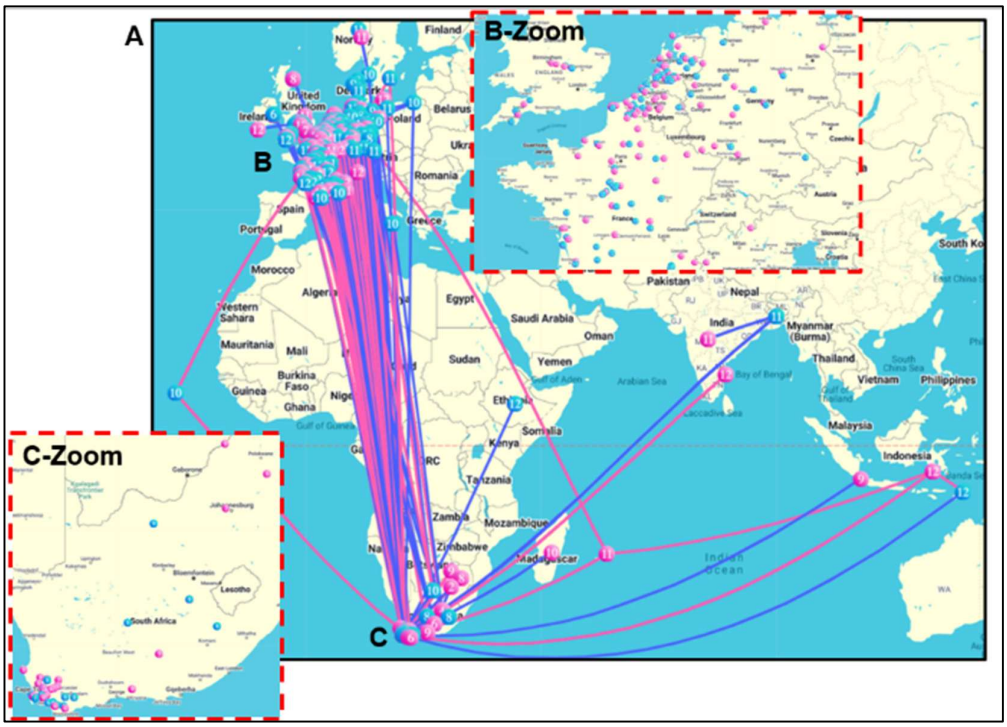


Figure 56: World map showing family ACM 8 through 12 generations and their global distribution. (A) World map with the origins of the progenitors of ACM 8. (B) Zoomed-in image of where majority of the progenitors came from in Europe (Netherlands, France and Germany). (C) Zoomed in image of the most recent generations of ACM 8 and their locations across South Africa. Pink dots represent the female ancestors, blue dots represent male ancestors, the numbers inside the dots indicate the generation of that individual, blue lines connect individuals from the paternal lineage and pink lines connect individuals from the maternal lineage (in the case of this family the paternal and maternal lineages of ACM 8.3). Generated with Rootmapper (https://rootsmapper.com/).

3.3.7 Family ACM 5 (old family)

3.3.7.1 Clinical history

This family is the largest enrolled family, with a total for 17 enrolled family members. The proband ACM 5.1, first presented with ACM at 44 years old with ventricular tachycardia. She was an endurance athlete, which has been shown to exacerbate the symptoms of ACM.

Sixteen additional family members were clinically phenotyped and of those who showed clinical symptoms (ACM 5.4, ACM 5.5, ACM 5.8, ACM 5.11, ACM 5.12) all had adult onset with mild to moderate symptoms, except for ACM 5.5 who presented during her adolescence. The pedigree for family ACM 5 is shown in Figure 57.

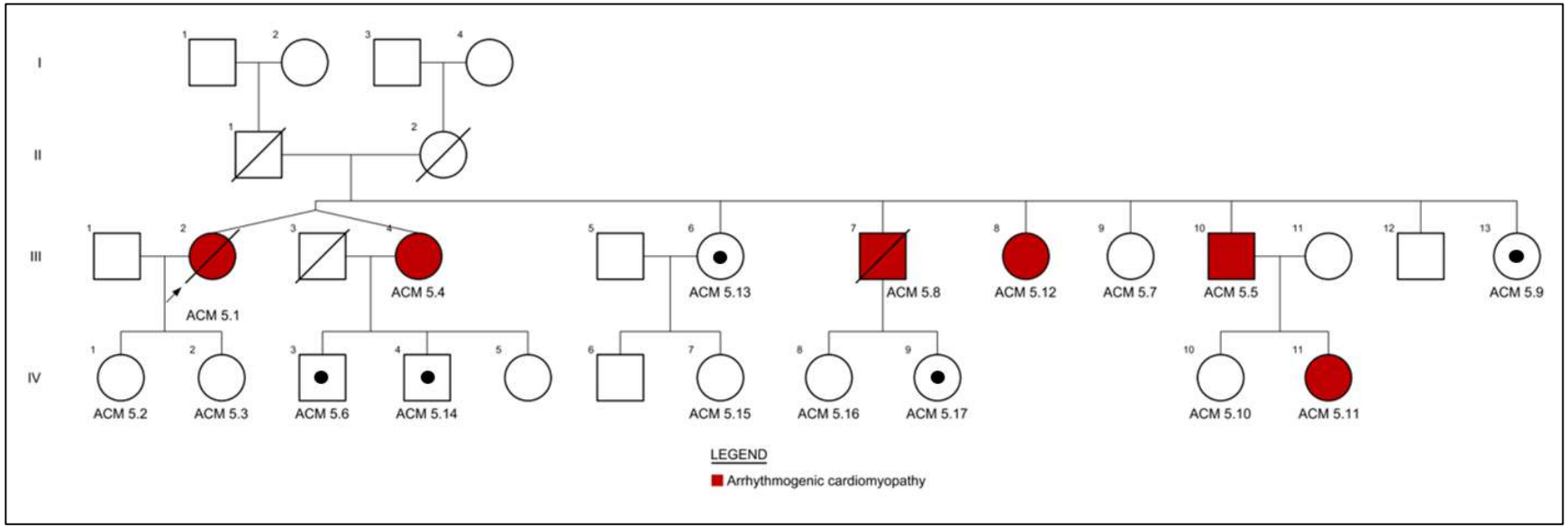


Figure 57: Pedigree of family ACM 5.The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals.

3.3.7.2 Genetics: PKP2 *c.1162C>T screening*

We genotyped 17 members of which 11 (ACM 5.1, ACM 5.4, ACM 5.5, ACM 5.6, ACM 5.8, ACM 5.9, ACM 5.11, ACM 5.12, ACM 5.13, ACM 5.14 and ACM 5.17) were found to be heterozygous for the *PKP2* c.1162C>T variant. The remaining six family members (ACM 5.2, ACM 5.3, ACM 5.7, ACM 5.10, ACM 5.15 and ACM 5.16) were all genotype negative, as shown in Figure 58 below

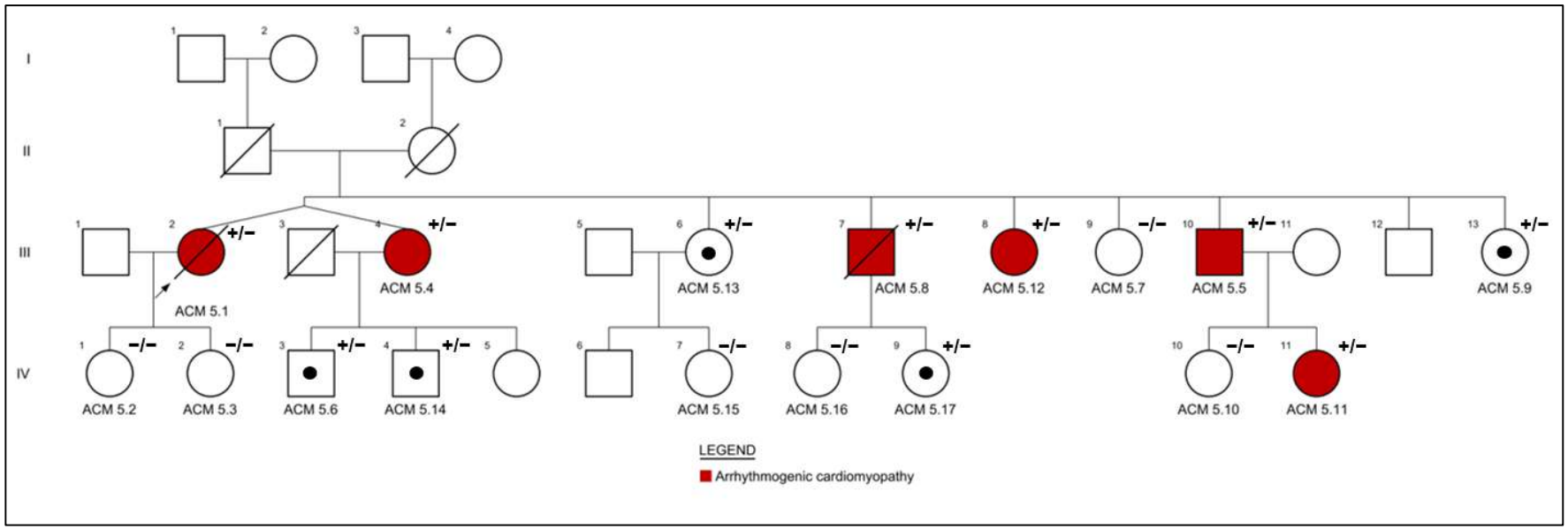


*Figure 58: The genotyping results of family ACM 5. The arrow indicates the proband, '+' indicates a positive genotype of the PKP2 variant while '–' indicates a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical sympt*oms

3.3.7.3 Haplotype analysis

The genetics results were followed up by haplotyping. This ACM 5 family is a continuation of previous research (54) with additional family members being recruited into the current study. New family members testing positive for the *PKP2* c.1162C>T variant (eg ACM 5.17) was included in the haplotype screening. We included ACM 8 as a positive control as this individual had inherited the red diseased haploblock (132, T,2 62, 250), seen in Figure X, which contained the *PKP2* c.1162C>T variant as well microsatellites associated with disease. We were able to infer the haplotypes from family members if the grandparents, parents or sibling's DNA was available for screening the microsatellite peaks of ACM 5.17 are shown in Figure 59 below.

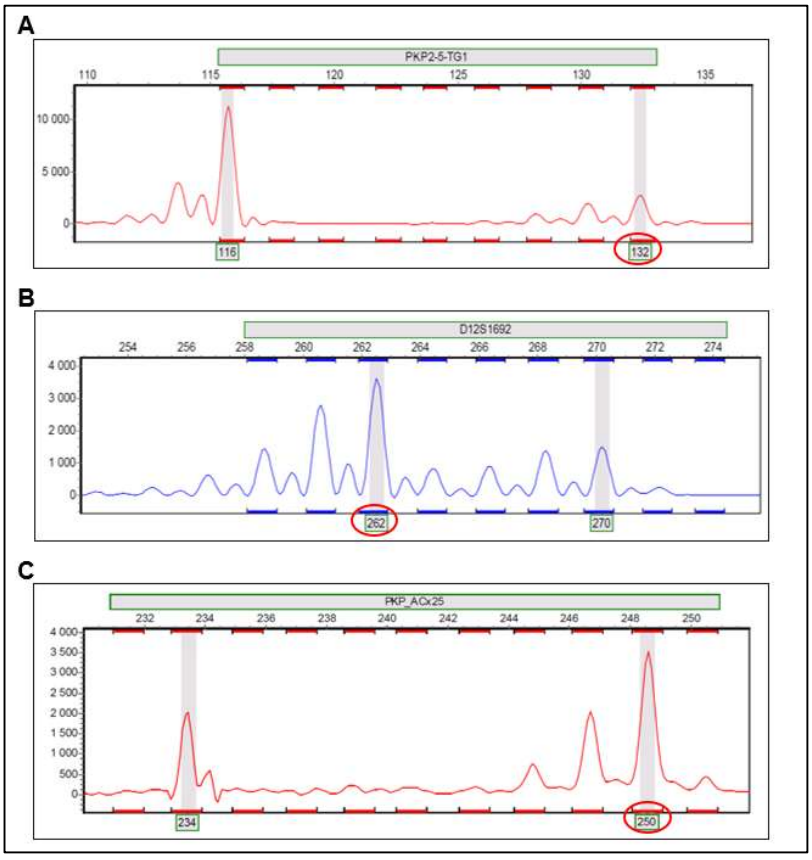


Figure 59: Microsatellite analysis of ACM 5.17. (A) PKP2_TG (132), the peak circled in red, 132, is the common haplotype peak for PKP2_TG. (B) D12S1692 (262), the peak circled in red, 262 is the common haplotype peak for D12S1692 (C) PKP2_AC (250), the peak circled in red, 250 bps, is the common haplotype peak for PKP2_AC.

The results of the haplotypes are shown in Figure 60 below. Eleven members of ACM 5 have the disease-associated haplotype indicated by the red block.

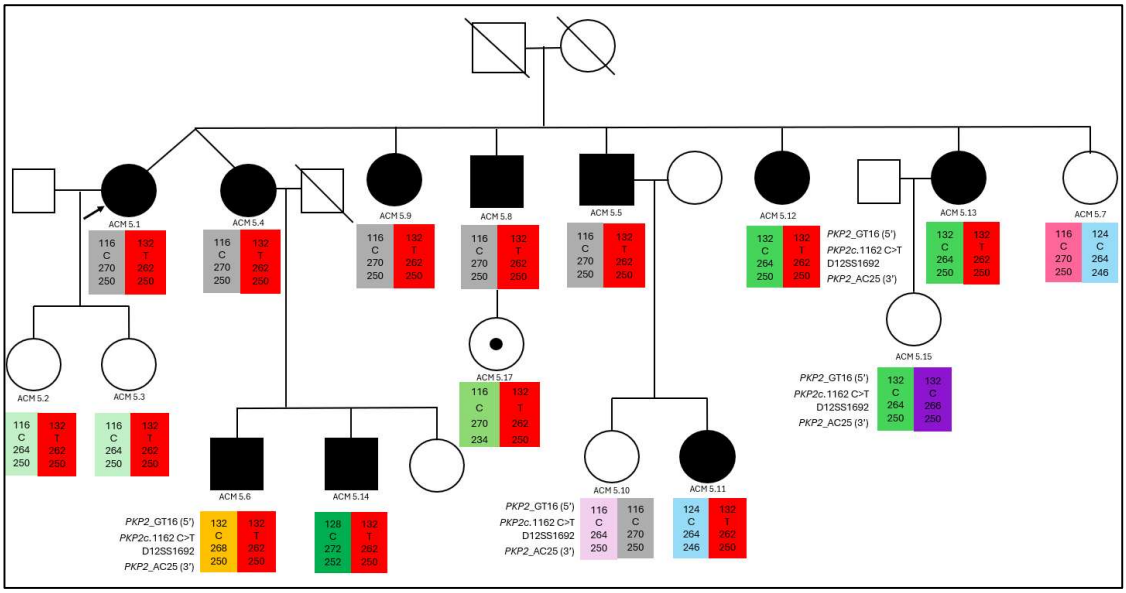


Figure 60: Haplotype analysis of family ACM 5. Circles indicate self-identified females, squares indicate self-identified males, black-filled symbols represent clinically affected individuals, symbols with periods in the centre represent individuals with the PKP2 variant but are not clinically affected, shapes with an oblique line through indicates deceased individuals, haplotypes with dotted borders represent inferred haplotypes

3.3.7.4 Genealogy tracing

Information for this family was provided by the genetic councillor, which included the names, dates of birth and residences of the proband's parents, grandparents and great-grandparents. They are self-identified Caucasian Afrikaners.

The proband of this family, ACM 5.1 and her twin sister were the oldest enrolled family members confirmed to harbour the *PKP2* variant and were the starting point of genealogy tracing. Both their maternal and paternal lines were investigated.

The proband maternal line was traced back to the late 17th century to German, French and Dutch progenitors. More recently the maternal line was shown to have English ancestry in the 19th century. There were multiple duplicates identified in this family due to consanguineous marriages in the late 17th and early 18th centuries between cousins.

The proband's paternal lineage was traced back to the early 18th century, to a German progenitor. There was also evidence of French and Dutch Progenitors dating back to the 17th century. The global distributions of ancestry for family ACM 5 can be found in Figure 61 below.

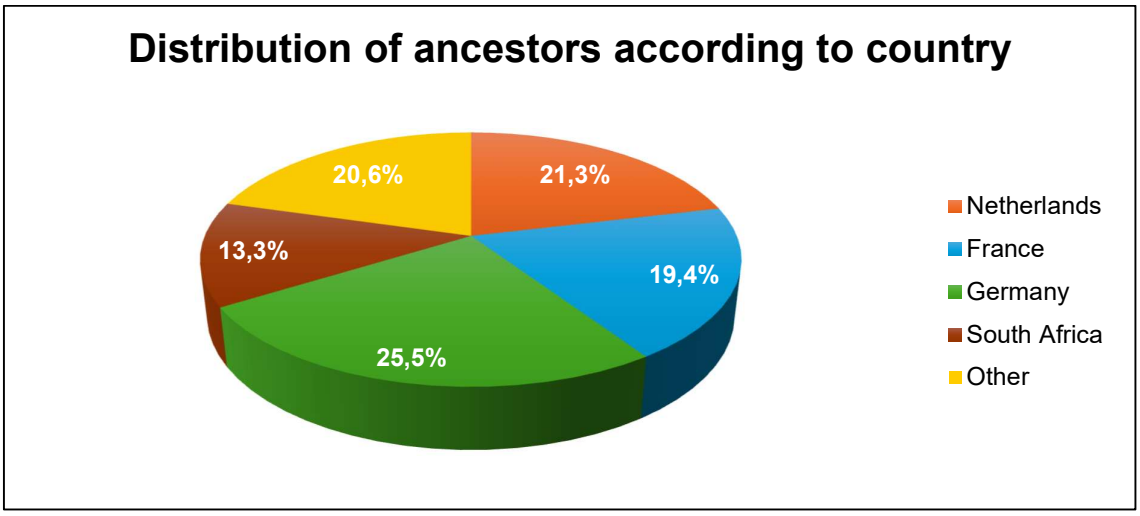


Figure 61: Distribution of ancestors according to country

All genealogical information was manually inputted, analysed and stored on Legacy v.10. this family was traced through 11 generations (shown in Figure 62).

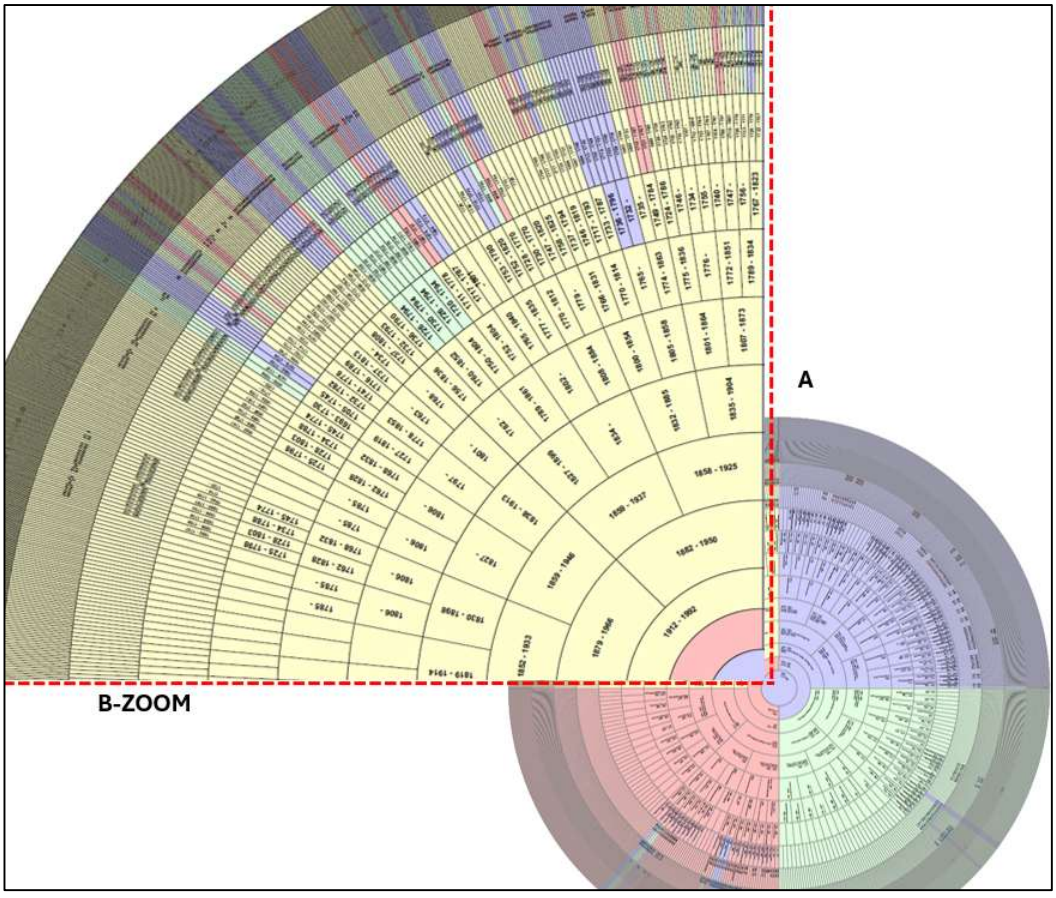


Figure 62: A fan diagram with the genealogy of Family ACM 5 presented in fan A, with a zoomed-in section, B, with all names removed of ancestors removed (dates 1958-1643)

When the genealogical tracing was complete for ACM5, we combined the ancestral data of ACM163, ACM12, ACM38, ACM154, ACM19, ACM8 and ACm5 and found that we could reduce the number of potential progenitors from 14 to 12.

Below is a graphical representation (Figure 63) of the origins of family ACM 5, from generation 0 – 12, which roughly coincides with years 1958 to 1643. Majority of the ancestors that came to South Africa were found in the 10th generation and this family had the most wide-spread 10th generation, there was a slight cluster of ancestors toward the Netherlands, but we also identified ancestors in the German, French, Belgium and Swiss regions in the 10th generation (B-Zoom). The progenitors of ACM 5 immigrated to the Cape of Good Hope in the 16th century where we found evidence of multiple generations to have remained. However, from the 6th generation we identified evidence of migration from the Western Cape to areas of the Free State and even further in land to Gauteng, where the present-day family resides (C-zoom).

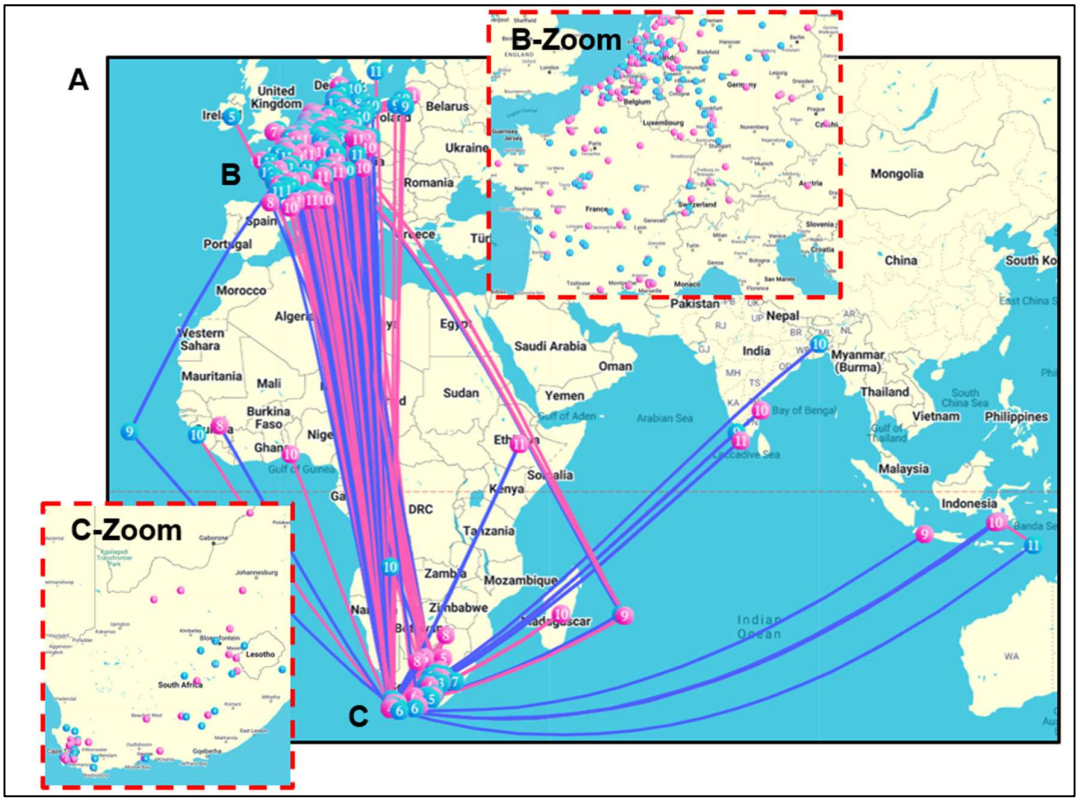


Figure 63: World map showing family ACM 5 through 11 generations and their global distribution. (A) World map with the origins of the progenitors of ACM 5. (B) Zoomed in image of where majority of the progenitors came from in Europe (Netherlands, France and Germany). (C) Zoomed in image of the most recent generations of ACM 5 and their locations across South Africa. Pink dots represent the female ancestors, blue dots represent male ancestors, the numbers inside the dots indicate the generation of that individual, blue lines connect individuals from the paternal lineage and pink lines connect individuals from the maternal lineage (in the case

of this family the paternal and maternal lineages of ACM 5.1). Generated with Rootmapper (https://rootsmapper.com/).

3.3.8 Family ACM 57 (old family)

3.3.8.1 Clinical history

This family represents one of three families where only the proband (ACM 57.1) was enrolled (Figure 64), she did not wish to disclose her family or have them recruited. She was a competitive marathon runner before her diagnosis at 46 years old. The initial presentation of the disease was ventricular tachycardia, she presents with a mild phenotype after decreasing her exercise load. The proband self-identified as a Caucasian Afrikaner.

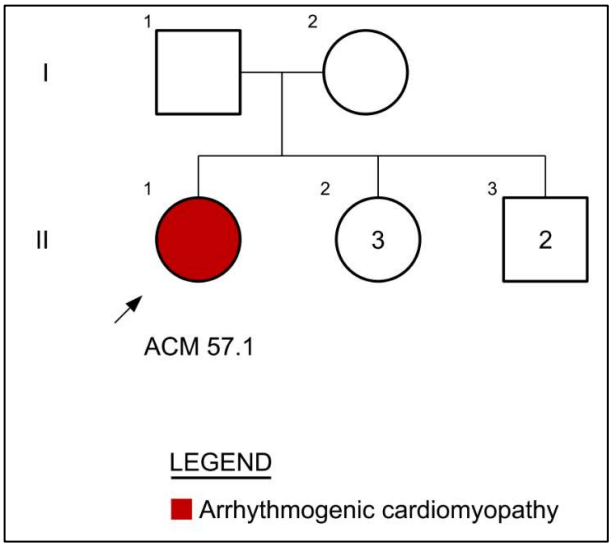


Figure 64: Pedigree of family ACM 57.The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals.

3.3.8.2 Genetics: PKP2 c.1162C>T screening

Genetic screening of the proband, ACM 57.1, found her to be heterozygous for the *PKP2* c.1162C>T variant (Figure 65).

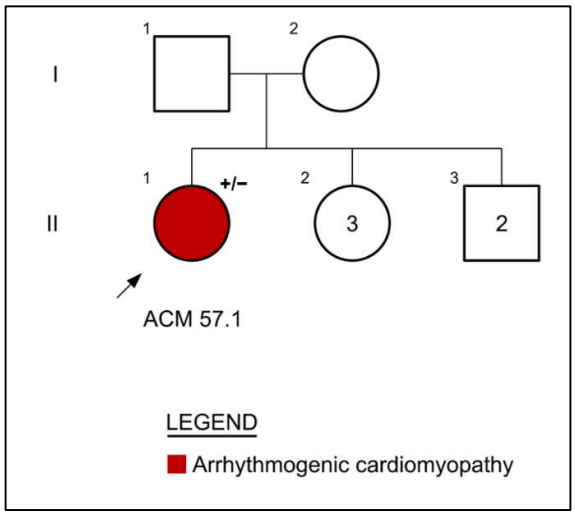


Figure 65: The genotyping results of family ACM 57. The arrow indicates the proband, '+' indicates a positive genotype of the PKP2 variant while '–' indicates a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms.

3.3.8.3 Haplotype analysis

The genetics results were followed up by haplotyping. The ACM57 family is a continuation of previous research (Watkins et al 2009) with additional family members being recruited into the current study. New family members testing positive for the *PKP2* c.1162C>T variant was included in the haplotype screening. We included ACM 8 as a positive control as this individual had inherited the red diseased haploblock (132, T, 262, 250), seen in Figure 59, which contained the *PKP2* c.1162C>T variant as well microsatellites associated with disease. We were able to infer the haplotypes from family members if the grandparents, parents or sibling's DNA was available for screening. Thus far, only the proband has the disease-associated haploblock, represented by the red block in Figure 66 below.

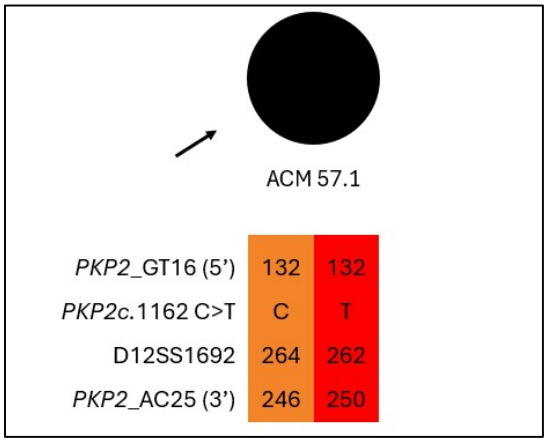


Figure 66: Haplotype analysis of ACM 57.1. Circles indicate self-identified females, squares indicate self-identified males, black-filled symbols represent clinically affected individuals, symbols with periods in the centre represent individuals with the PKP2 variant but are not clinically affected, shapes with an oblique line through indicates deceased individuals, haplotypes with dotted borders represent inferred haplotypes.

3.3.8.4 Genealogy tracing

There was limited information provided on the pedigree given to us by the genetic councillor regarding names dates of birth and residences of the ACM 57 family. Although this family was previously investigated, not much information was available and we started accumulating information on names, dates of birth and locations of the proband, her parents and grandparents.

Based on the information provided on the proband and her parents, we were able to confirm birthdates and residences by consulting the 1984 voter's role, where we found this family resides in the Western Cape region. As the proband (ACM 57.1) was the only enrolled participant investigations were carried out for both her paternal and maternal lineages. Searching online databases (FamilySearch, Geni and Ancestry) for the additional information on the probands parents and grandparents started through the appropriate regional records.

The paternal line of the proband was traced to the mid-17th century to Dutch, German and French descent. Well-researched and known French and Dutch couples, with well researched genealogies, were very frequent in this family and duplicates of these couples were evident from the 17th century due to marriages between cousins.

The maternal line of the proband was traced back to the mid-17th century to predominantly French and Dutch couples. Again, there was evidence of consanguineous marriages between

cousins from the late 17th century leading to several duplications. The global distributions of ancestry for family ACM 57 can be found in Figure 67 below.

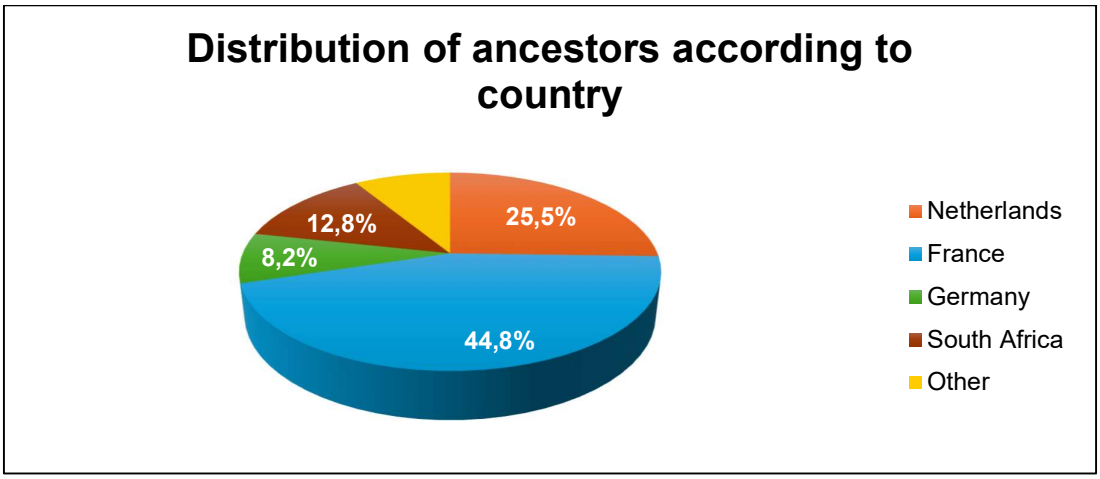


Figure 67: Distribution of ancestors according to country

All genealogical information was manually inputted, analysed and stored on Legacy v.10. This family was traced back through 10 generations (Figure 68).

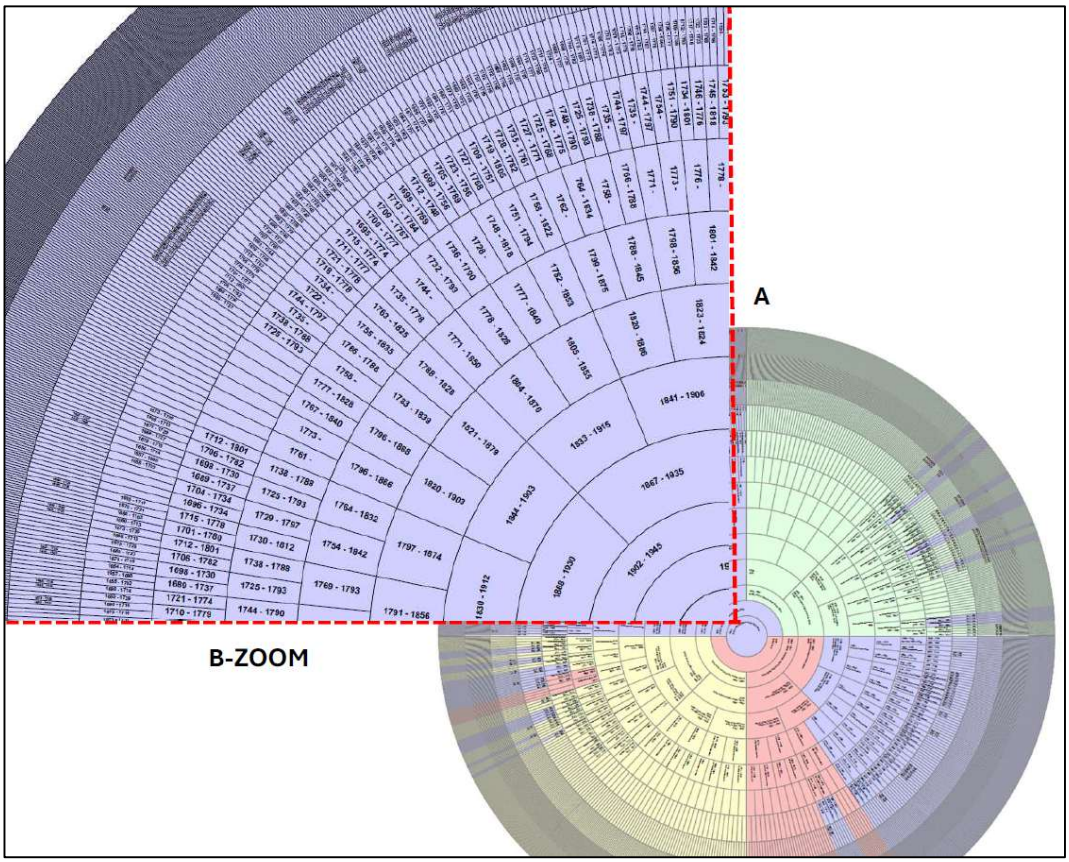


Figure 68: A fan diagram with the genealogy of Family ACM 57 presented in fan A, with a zoomed-in section, B, with all names of ancestors removed (dates 1962-1632)

Finally, once we had complied all possible records to create the genealogy, we compared it to the genealogies of ACM 163, ACM 12, ACM 38, ACM 154, ACM 19, ACM 8 and ACM 5, we were able to decrease the number of the possible progenitors from 12 to 8.

Below is a graphical representation (Figure 69) of the origins of family ACM 57. From generation 0 – 12, which roughly coincides with years 1962 – 1632. Majority of the ancestors that came to South Africa were found in the 10th generation, where majority clustered in the regions of the Netherlands and more sparsely throughout regions of France. We were also able to locate multiple ancestors in the region of Germany (B-Zoom). The progenitors of ACM 57 immigrated to the Cape of Good Hope in the 16th century where the present-day family still resides (C-zoom).

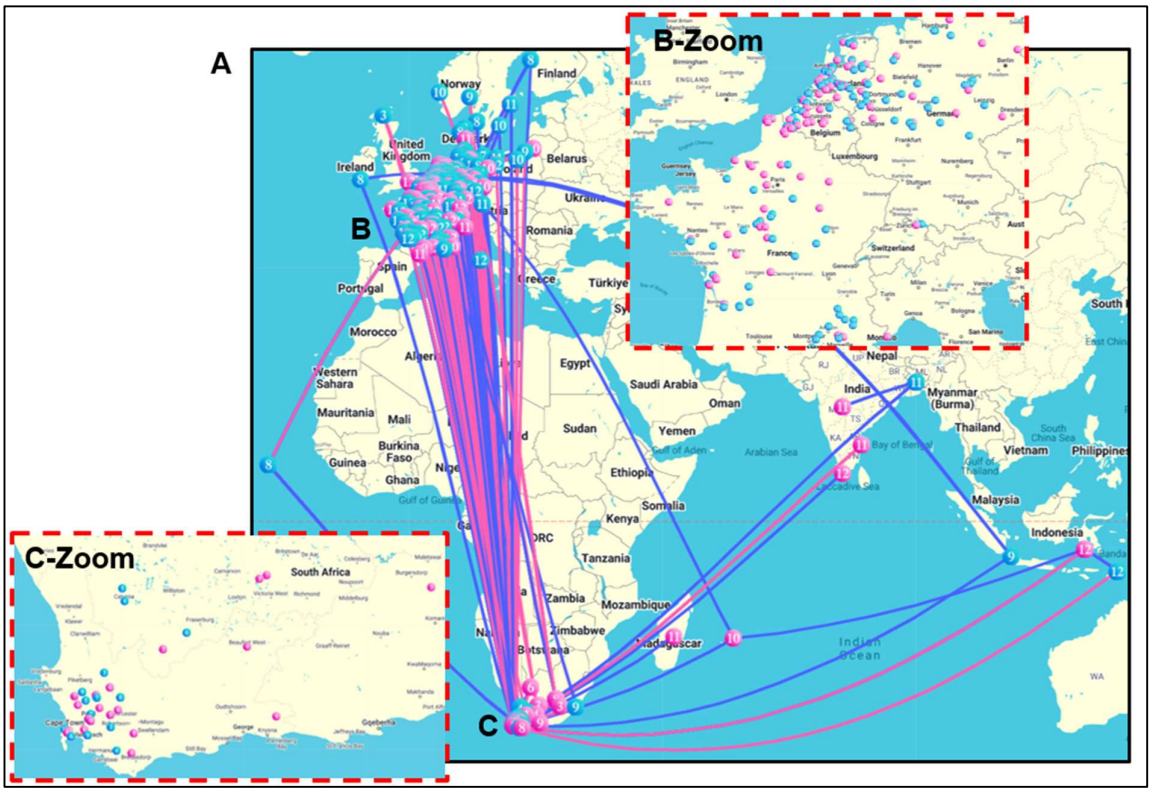


Figure 69: World map showing family ACM 57 through 12 generations and their global distribution. (A) World map with the origins of the progenitors of ACM 57. (B) Zoomed in image of where majority of the progenitors came from in Europe (Netherlands, France and Germany). (C)Zoomed in image of the most recent generations of ACM 57 and their locations across South Africa. Pink dots represent the female ancestors, blue dots represent male ancestors, the numbers inside the dots indicate the generation of that individual, blue lines connect individuals from the paternal lineage and pink lines connect individuals from the maternal lineage (in the case of this family the paternal and maternal lineages of ACM 57.1). Generated with Rootmapper (https://rootsmapper.com/).

3.3.9 Family ACM 157 (new family)

3.3.9.1 Clinical history

The proband, ACM 157.1 and his nuclear family (ACM 157.2, ACM 157.3 and ACM 157.4) were all enrolled in the IMHOTEP registry. Clinical onset of ACM 157.1 occurred at the age of 15, with severe right ventricular involvement leading to implantable cardioverter defibrillator (ICD) implantation. At 18, he developed amiodarone-induced thyrotoxicosis and right-sided heart failure. A heart transplant was performed the following year. The parents of the proband (ACM 157.3 and ACM 157.4) were clinically screened at 45 and 40 years respectively and were both found to be clinically unaffected. The proband's younger brother (ACM 157.2) was clinically screened at 14 years and was also found to be clinically unaffected (Figure 70).

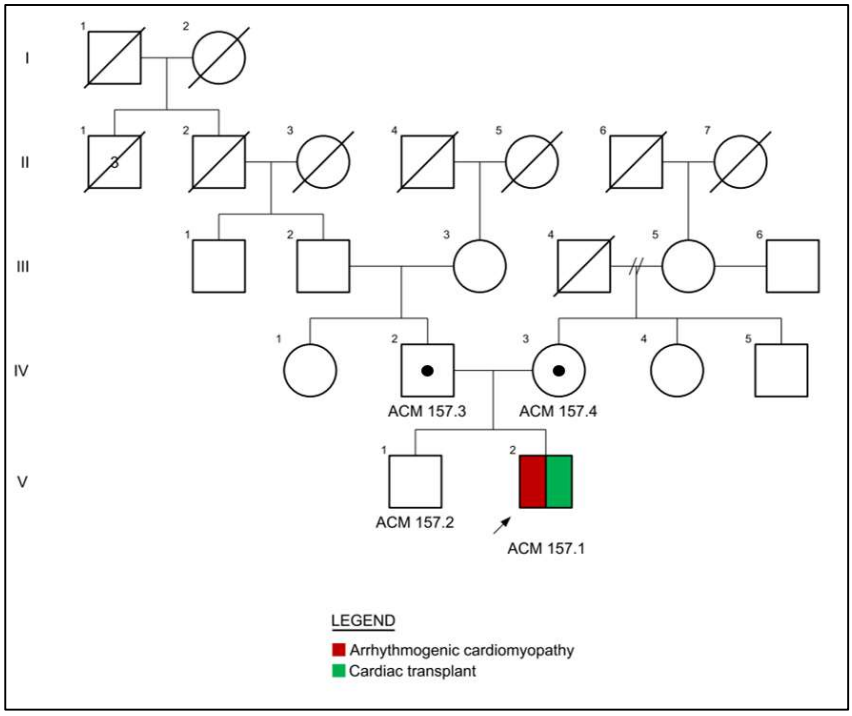


Figure 70: Pedigree of Family ACM 157. The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals.

3.3.9.2 Genetics: PKP2 c.1162C>T screening

Genetic screening of all four family members revealed the proband was homozygous for the *PKP2* c.1162C>T variant. He inherited a mutant allele from both his father (ACM 157.3) and mother (ACM 157.4), who were both found to be heterozygous for the *PKP2* variant. The proband's older sibling was, ACM 157.2, was genotype negative (Figure 71).

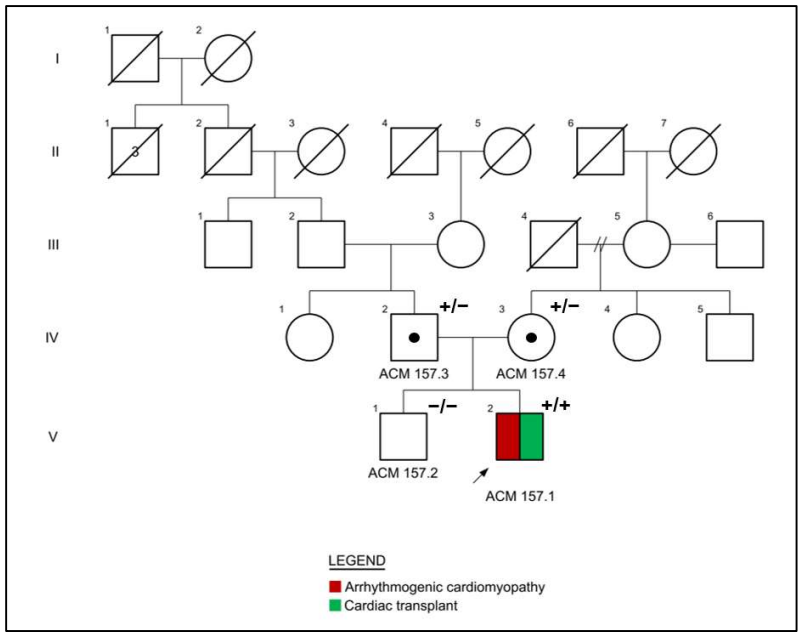


Figure 71: The genotyping results of family ACM 157. The arrow indicates the proband, '+' indicates a positive genotype of the PKP2 variant while '−' indicates a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms

3.3.9.3 Haplotype analysis

The genetics results were followed up by haplotyping. Haplotyping was carried out for each family member of ACM 157, through the combining the microsatellites and *PKP2* c.1162C>T. The microsatellite peaks for the proband (ACM 157.1) are shown in Figure 72 below.

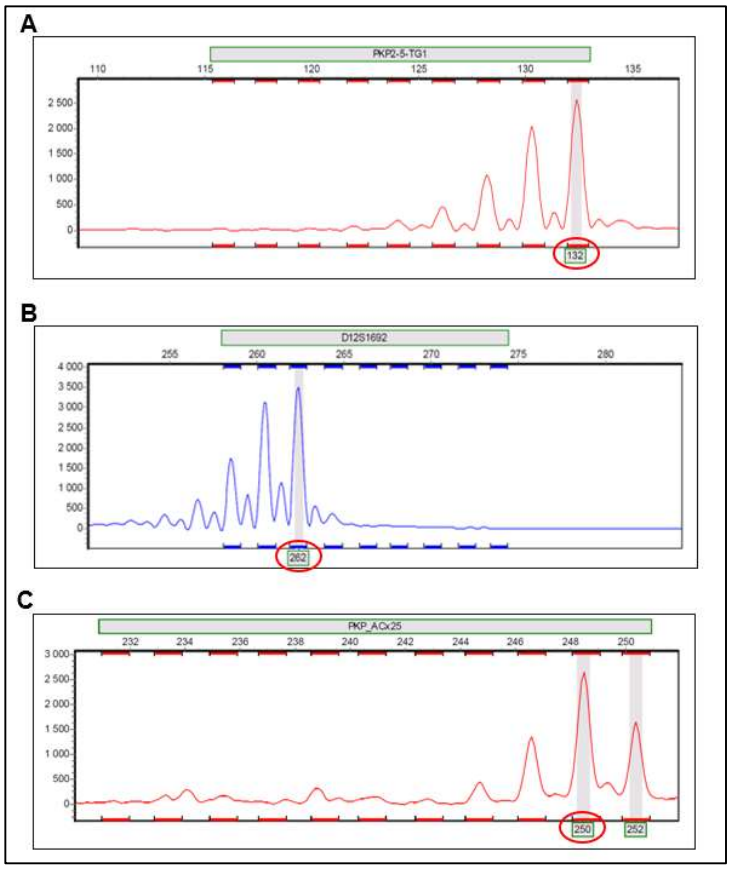


Figure 72: Microsatellite analysis of ACM 157.1. (A) PKP2_TG (132), the peak circled in red, 132, is the common haplotype peak for PKP2_TG. (B) D12S1692 (262), the peak circled in red, 262 is the common haplotype peak for D12S1692 (C) PKP2_AC (250), the peak circled in red, 250 bps, is the common haplotype peak for PKP2_AC.

The microsatellite analysis results were combined and used to construct haplotypes for ACM 157.1, ACM 157.2, ACM 157.3 and 157.4 (Figure 73), which showed the parents (ACM 157.3 and ACM 157.4) both had one disease associated haploblock, represented by the red block. They subsequently each passed their disease-associated haploblock down to the proband ACM 157.1, who homozygous for the disease-associated haploblock. While his brother, ACM 157.4 was genotype negative and therefore does not have the disease-associated haploblock.

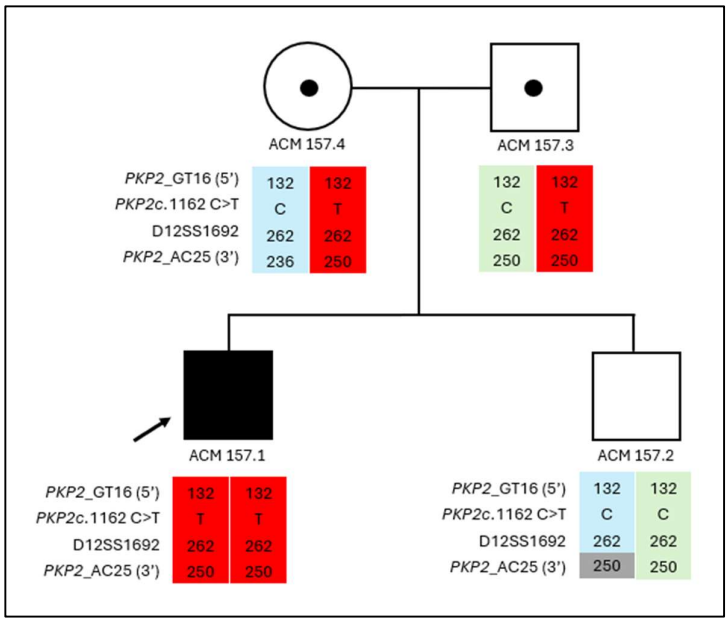


Figure 73: Haplotype analysis of family ACM 157. Circles indicate self-identified females, squares indicate self-identified males, black-filled symbols represent clinically affected individuals, symbols with periods in the centre represent individuals with the PKP2 variant but are not clinically affected, shapes with an oblique line through indicates deceased individuals, haplotypes with dotted borders represent inferred haplotypes

3.3.9.4 Genealogy tracing

There was limited information provided by the genetic councillor regarding names dates of birth and residences of the ACM 157 family. Some of the names on the pedigree provided were incomplete or nicknames were used. Fortunately, after the family physician spoke with the parents, more information was provided regarding their family history. We were able to confirm the probands (ACM 157.1) grandparents names and residences from both the maternal and paternal lineages. We identified that most of the living family members reside in the Gauteng region of South Africa, with the first of their ancestors migrating there from the Western Cape in the early 20th century. This family self-identified as a Caucasian Afrikaner family.

The probands paternal (ACM 157.3) and maternal (ACM 157.4) lineages were investigated as they are both carriers of the *PKP2* variant.

The ancestry of ACM 157.3 was only extended into the maternal line due to the limited names, dates and locations of his paternal grandparents. We attempted to contact the relevant Dutch Reformed Churches (NGK: Nederduitse Gereformeerde Kerk) and repositories in their supposed residence however, no records were found. The maternal lineage is predominantly

of Dutch with some German ancestors being recorded. There were consanguineous marriages recorded as recently as the 19th century.

The ancestry of ACM 157.4 was traced back through both her paternal and maternal lineages. However, due to limited information about her paternal lineage, we were unable to confidently trace this lineage passed the early 20th century, as it was noted by the family that the individuals of the paternal linage were born in Kenya. The maternal lineage of ACM 157.4 was successfully traced back to the late 17th century to French and Dutch progenitor couples. Consanguinity was evident in the 17th century. The global distributions of ancestry for family ACM 157 can be found in Figure 74 below.

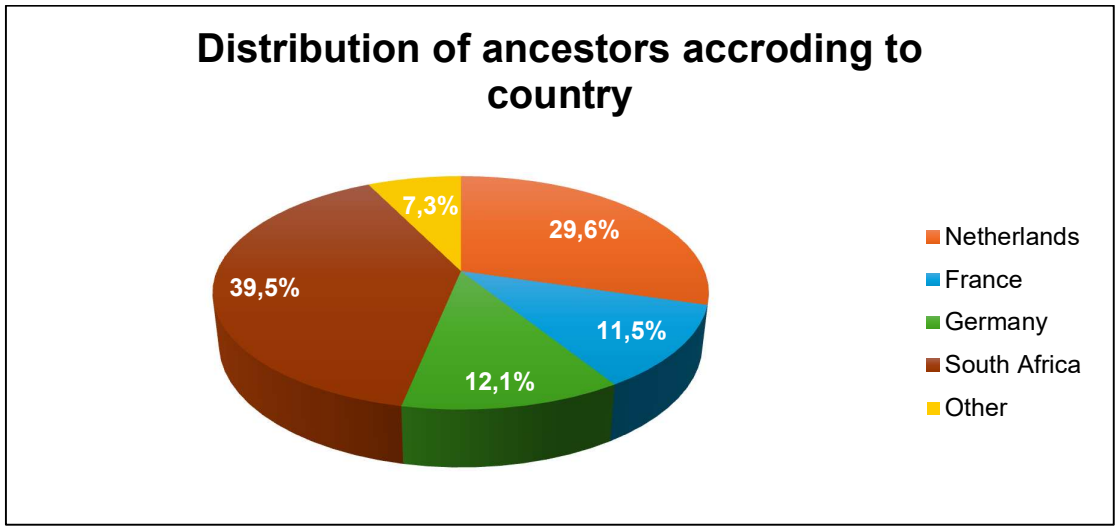


Figure 74: Distribution of ancestors according to country

This family was traced back through 14 generations to well-known French and Dutch progenitors (shown in Figure 75).

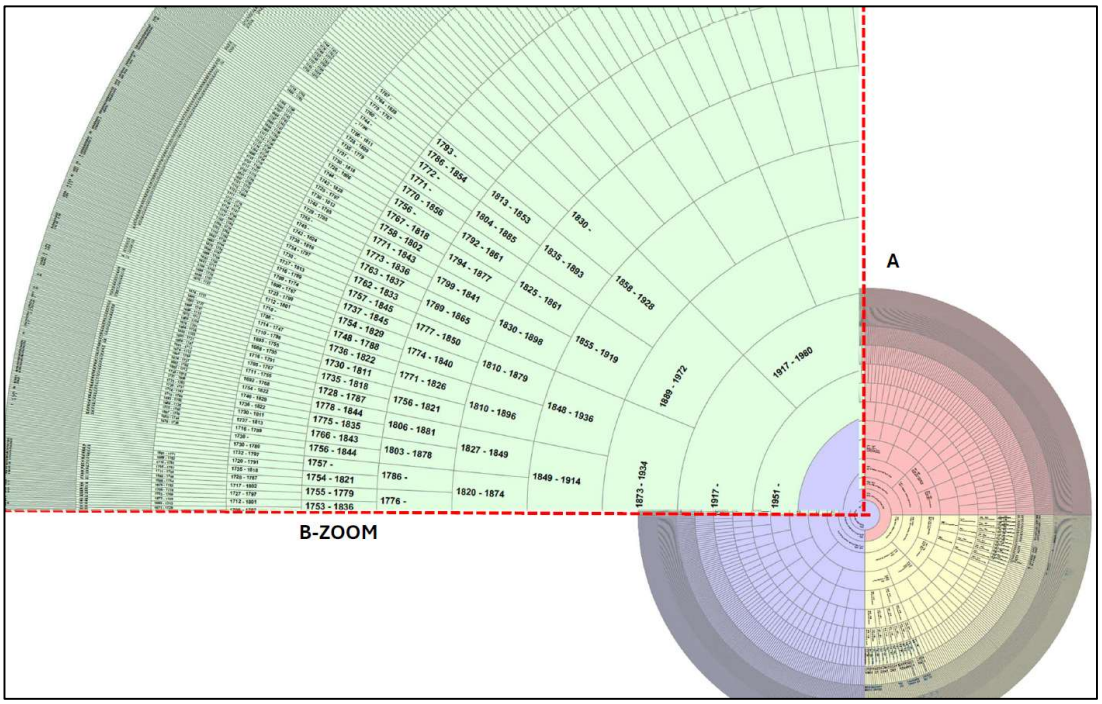


Figure 75: A fan diagram with the genealogy of Family ACM 157 presented in fan A, with a zoomed-in section, B, with all names removed of ancestors removed (dates 2002-1645).

Upon finalising the genealogy, we compared the genealogy of ACM 157 to that of ACM 163, ACM 12, ACM 38, ACM 154, ACM 19, ACM 8, ACM 5 and ACM 57 and found that the number of potential progenitor couples remained at 8; we were this not able to reduce this number.

Below is a graphical representation (Figure 76) of the origins of family ACM 157, from generation 0 to 12, which roughly coincides with years 2002 – 1645. Majority of the ancestors that came to South Africa were found in the 10th generation. Majority of the progenitors of ACM 157 were found to cluster in the Netherlands and more sparsely throughout the regions of present-day Belgium, Germany and France (B-Zoom). The progenitors immigrated to the Cape of Good Hope in the 16th century and multiple generations remained there (10th – 3rd). However, the most recent and enrolled individuals of this study reside in Gauteng and have for a couple of generations (C-zoom).

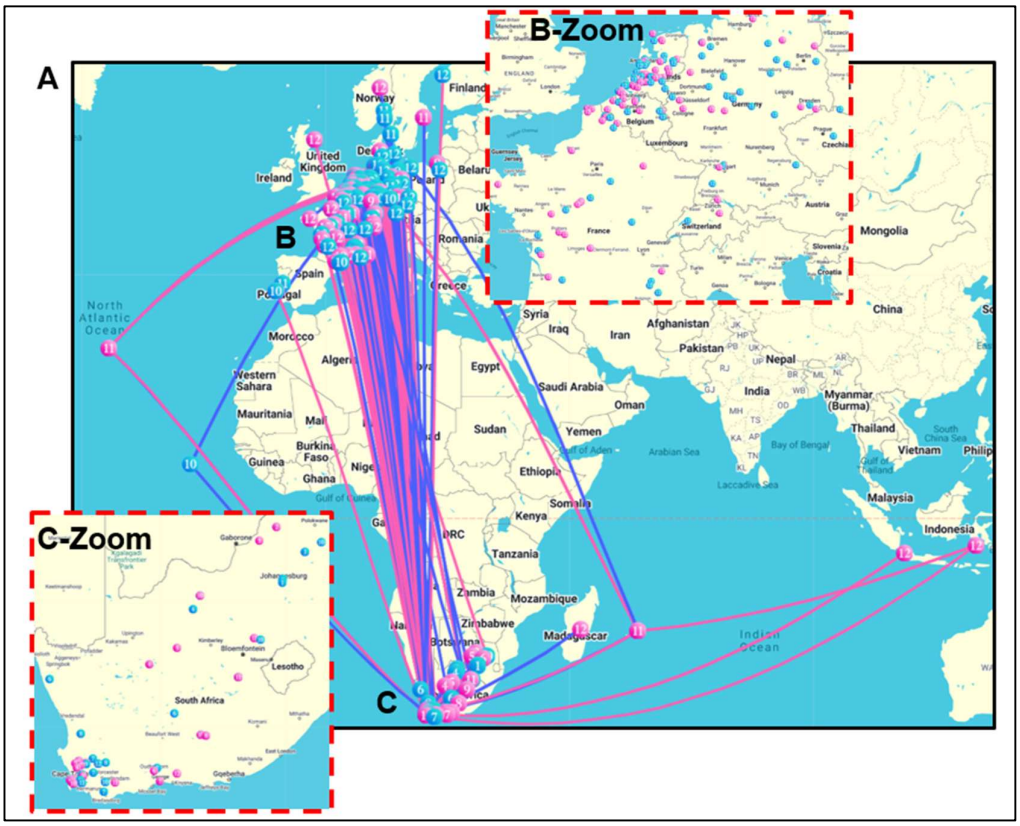


Figure 76: World map showing family ACM 157 through 12 generations and their global distribution. (A) World map with the origins of the progenitors of ACM 157. (B) Zoomed in image of where the majority of the progenitors came from in Europe (Netherlands, France and Germany). (C) Zoomed in image of the most recent generations of ACM 157 and their locations across South Africa. Pink dots represent the female ancestors, blue dots represent male ancestors, the numbers inside the dots indicate the generation of that individual, blue lines connect individuals from the paternal lineage and pink lines connect individuals from the maternal lineage (in the case of this family the paternal and maternal lineages of ACM 157.1). Generated with Rootmapper (https://rootsmapper.com/).

3.3.10 Family ACM 158 (new family)

3.3.10.1 Clinical history

This proband, ACM 158.1, was an endurance athlete when he was first diagnosed at 37 years, when the disease first presented as ventricular tachycardia. Following the diagnosis, an ICD was implanted. He has continued cycling and now presents with a mild phenotype. History taking by the genetic counsellor found the that the proband's cousin, IV-2, was believed to

have had a cardiac-related condition and proband's great uncle, II-4, was believed to have died at 50 years old due to a cardiac related event (Figure 77).

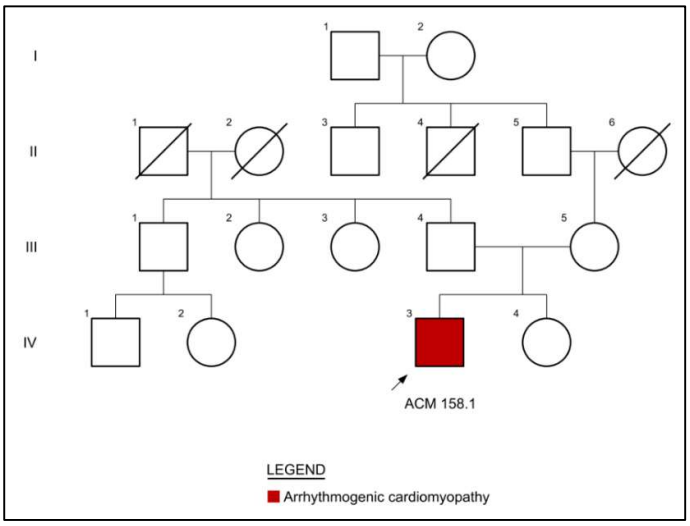


Figure 77: Pedigree of family ACM 158.The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals.

3.3.10.2 Genetics: PKP2 *c.1162C>T screening*

Only the proband, ACM 158.1, underwent genetic screening and his genotype result showed that he was heterozygous for the PKP2 c.1162C>T variant (Figure 78).

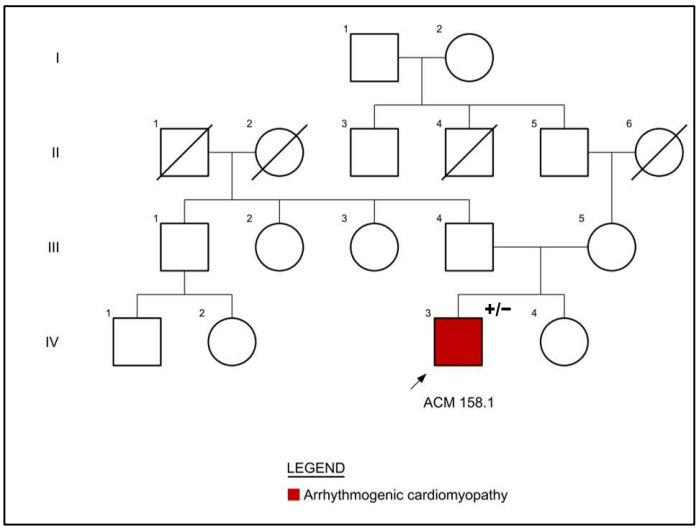


Figure 78: The genotyping results of family ACM 158. The arrow indicates the proband, '+' indicates a positive genotype of the PKP2 variant while '–' indicates a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms

3.3.10.3 Haplotype analysis

The genetics results were followed up by haplotyping. Haplotyping was carried out only for the proband, ACM 158.1, by making use of the microsatellites and the variant of interest (*PKP2* c.1162C>T). The resulting microsatellite peaks are shown below in Figure 79.

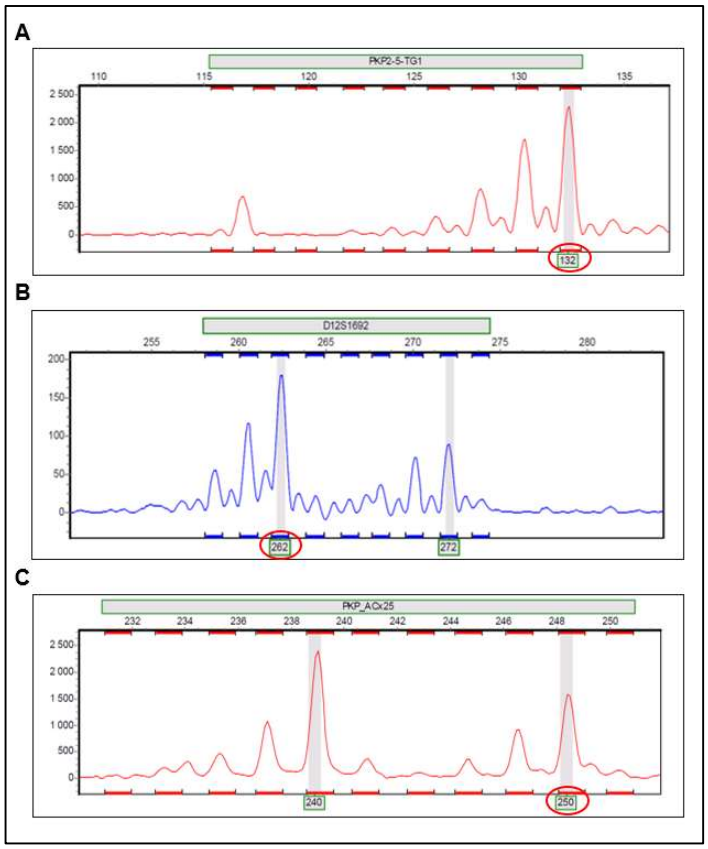


Figure 79: Microsatellite analysis of ACM 158.1. (A) PKP2_TG (132), the peak circled in red, 132, is the common haplotype peak for PKP2_TG. (B) D12S1692 (262), the peak circled in red, 262 is the common haplotype peak for D12S1692 (C) PKP2_AC (250), the peak circled in red, 250 bps, is the common haplotype peak for PKP2_AC

The microsatellite analysis results were combined and used to construct a haplotype for ACM 158.1 (Figure 80).

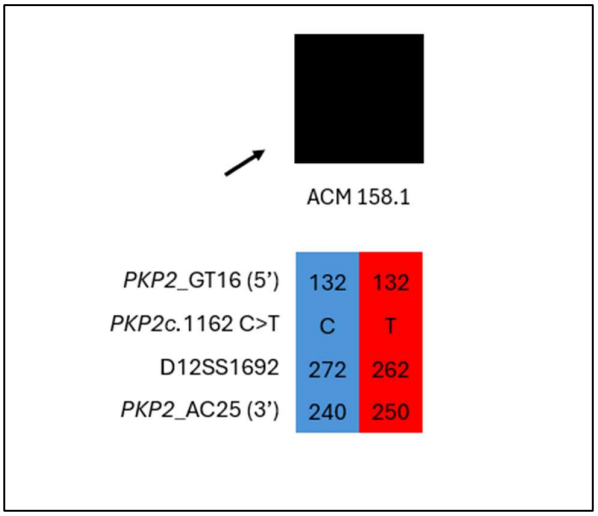


Figure 80: Haplotype analysis of family ACM 158. Circles indicate self-identified females, squares indicate self-identified males, black-filled symbols represent clinically affected individuals, symbols with periods in the centre represent individuals with the PKP2 variant but are not clinically affected, shapes with an oblique line through indicates deceased individuals, haplotypes with dotted borders represent inferred haplotypes.

3.3.10.4 Genealogy tracing

As ACM 158.1 was the only enrolled member of his family, there was limited information for names, dates of birth and locations of his immediate family recorded on the pedigree provided by the genetic counsellor. Therefore, names are residences were first confirmed by consulting the 1984 voter's roll. It was confirmed the proband resided in the Western Cape and his parents in Qqeberha.

To trace this family's genealogy, a great deal of information had to be gathered and organised, including copies of birth, marriage, and death certificates; records in churches and graveyards; old articles, letters, photographs, and documents related to citizenship and naturalisation; searches of historical records, including land records and censuses; a backwards and forwards approach was applied. All the relevant records for both the paternal and maternal lineages were search for using online databases (FamilySearch, Geni, Ancestry).

We started investigating the paternal line as this family came from a well-known French surname which has had extensive research done and has a South African Genealogy published with this surname. We were therefore able to extend the paternal lineage back to the 17th century to French, Dutch and German progenitors. There was evidence of consanguineous marriages in the late 17th century.

Initially, the maternal lineage was proving difficult to trace, as the maternal grandparents' surnames were Portuguese and there was limited information found on the online databases and in the Cape Town Archives. However, after consulting Professor Geldenhuys we were able to trace the maternal lineage back to the 17th century to French, Dutch German and Portuguese ancestry. The global distribution of ancestry for family ACM 158 can be found in Figure 81 below.

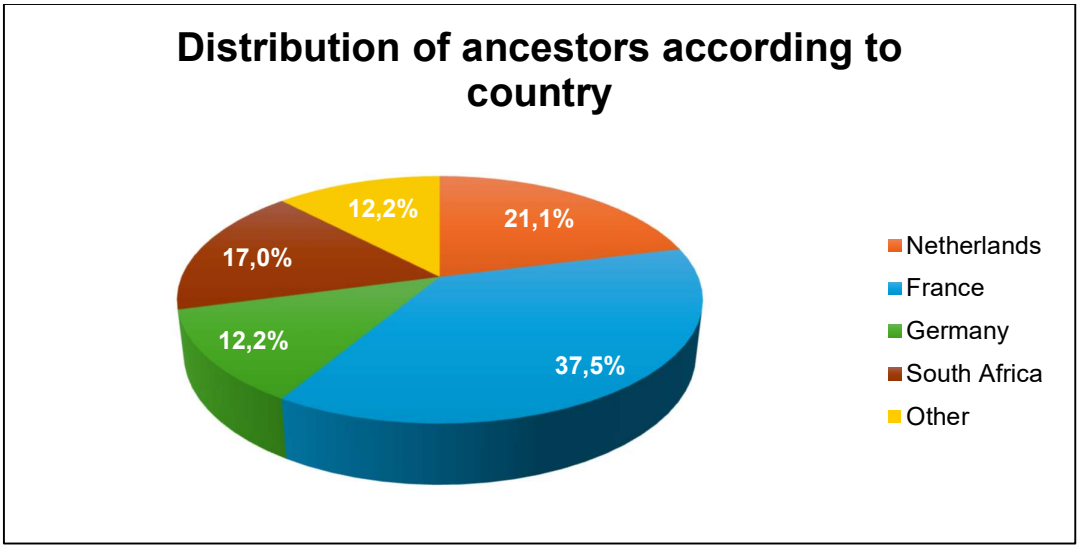


Figure 81: Distribution of ancestors according to country

All data was inputted manually, stored and analysed using Legacy v.10. Overall, we were able to trace this family back through 14 generations (shown in figure 82)

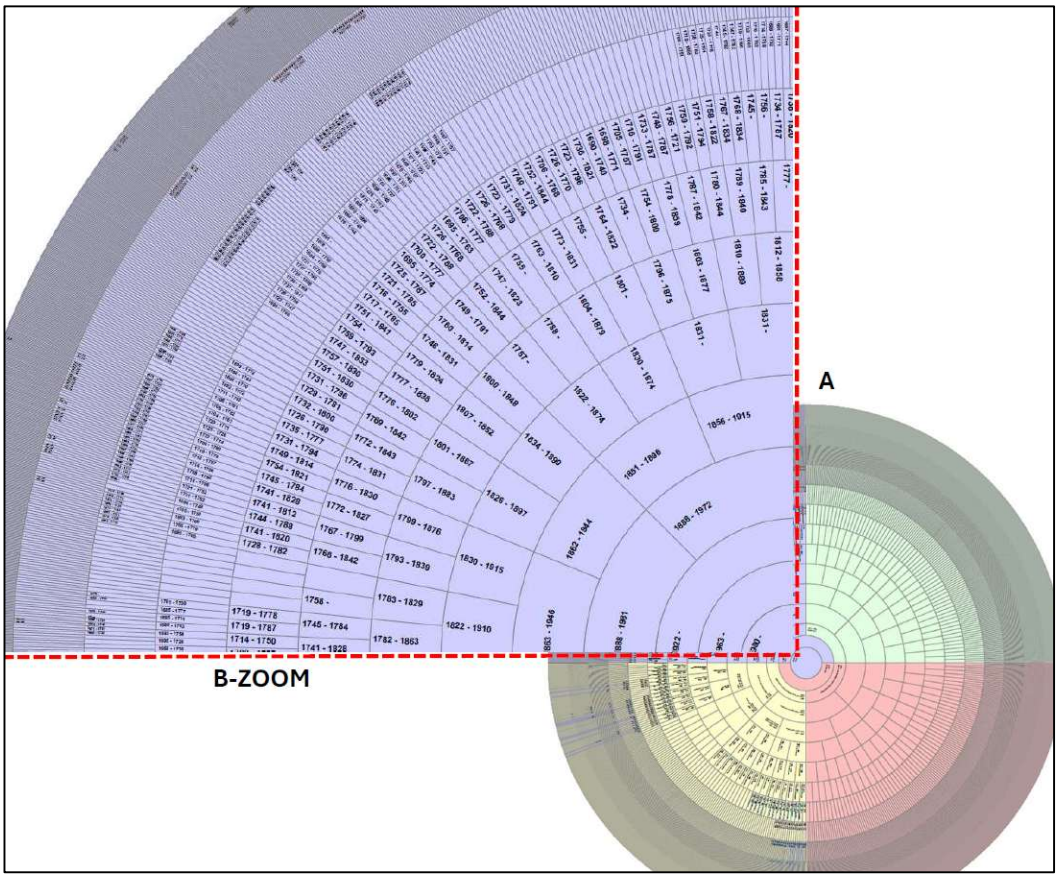


Figure 82: A fan diagram with the genealogy of Family ACM 158 presented in fan A, with a zoomed-in section, B, with all names removed of ancestors removed (dates 1980-1620)

We then compared the genealogy of ACM 158 to ACM 163, ACM 12, ACM 38, ACM 154, ACM 19, ACM 8, ACM 5, ACM 57 and ACM 157, where we found the number of potential founder couples remained at 8.

Below is a graphical representation (Figure 83) of the origins of family ACM 158, from generation 0 to13, which roughly coincides with years 1980 – 1620. Majority of the ancestors that came to South Africa were found in the 10th generation, where we can see distinct clustering of ancestors in the region of the Netherlands and present-day Belgium. There were also multiple ancestors located across the regions of France, Germany and Switzerland (B-Zoom).The progenitors of ACM 158 immigrated to the Cape of Good Hope in the 16th century, many of the ancestors remained in the Western Cape, however there was evidence of a few ancestors migrating further inland (7th – 2nd) toward Bloemfontein and Qqeberha (the probands parents), while the proband still resides in the Western Cape (C-zoom).

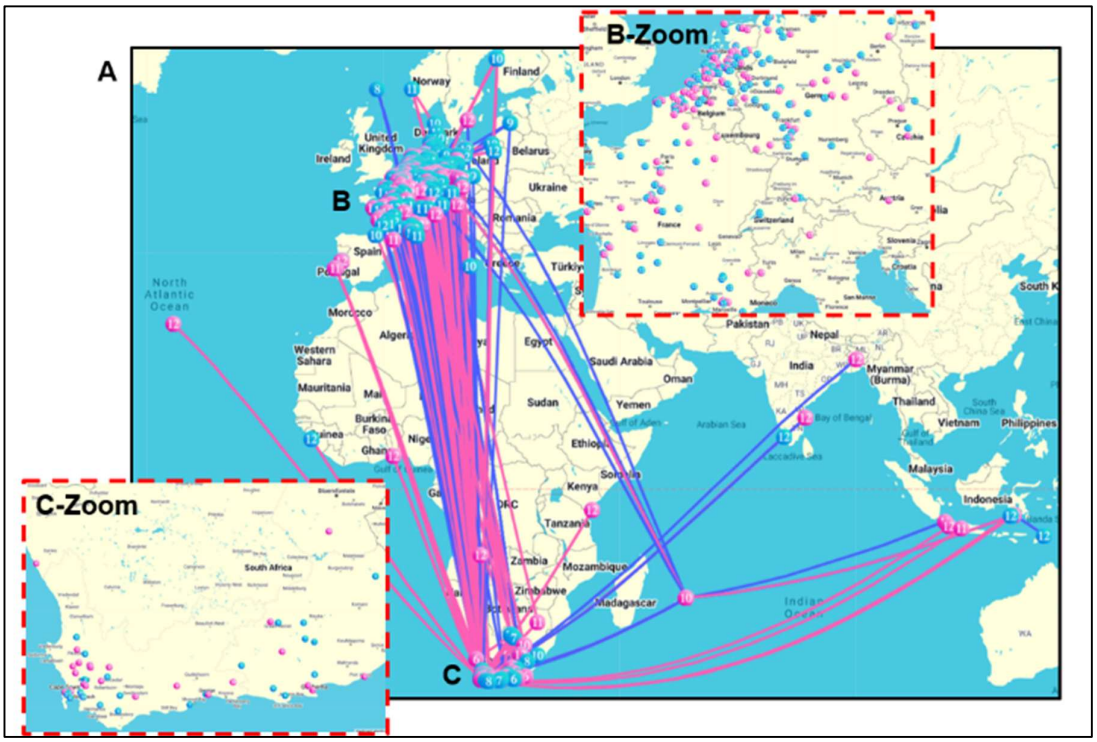


Figure 83: World map showing family ACM 158 through 12 generations and their global distribution. (A) World map with the origins of the progenitors of ACM 158. (B) Zoomed in image of where majority of the progenitors came from in Europe (Netherlands, France and Germany). (C)Zoomed in image of the most recent generations of ACM 158 and their locations across South Africa. Pink dots represent the female ancestors, blue dots represent male ancestors, the numbers inside the dots indicate the generation of that individual, blue lines connect individuals from the paternal lineage and pink lines connect individuals from the maternal lineage (in the case of this family the paternal and maternal lineages of ACM 158.1). Generated with Rootmapper (https://rootsmapper.com/).

3.3.11 Family ACM 161 (new family)

3.3.11.1 Clinical history

The proband (ACM 161.1) and her nuclear family (ACM 162.2, ACM 161.3 and ACM 161.4) were enrolled in the IMHOTEP registry. ACM 161.1, was one of the rare cases with disease onset at the young age of 12 years old, presenting with ventricular tachycardia and syncope. Fortunately, there were no signs of heart failure at diagnosis, therefore, an ICD was implanted, and all sporting activities were stopped as a precaution.

Clinical notes of the probands mother, ACM 161.3, stated she was previously diagnosed with heart failure of an unknown cause. However, she suffered from co-morbidities such as diabetes and was morbid obesity which could have contributed to the heart failure diagnosis. The father, ACM A161.4, was clinically unaffected, however, the proband's stepsister, ACMIII-2, from his first marriage, was described as having an unspecified arrhythmia in the

clinical notes. The youngest member of the family, ACM 161.2 was clinically unaffected (Figure 84).

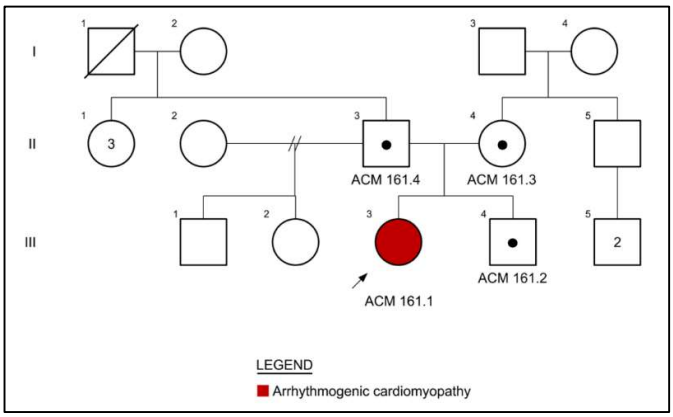


Figure 84: Pedigree of family ACM 161. The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals.

3.3.11.2 Genetics: PKP2 *c.1162C>T screening*

The DNA of the proband (ACM161.1) and her family (ACM162.2, ACM161.3 and ACM161.4) were genotyped the proband was found to be homozygous for the *PKP2* c.1162C>T variant while both parents were heterozygous carriers of the *PKP2* variant. The proband's brother, ACM 161.2, was also found to be homozygous for the *PKP2* variant (Figure 85).

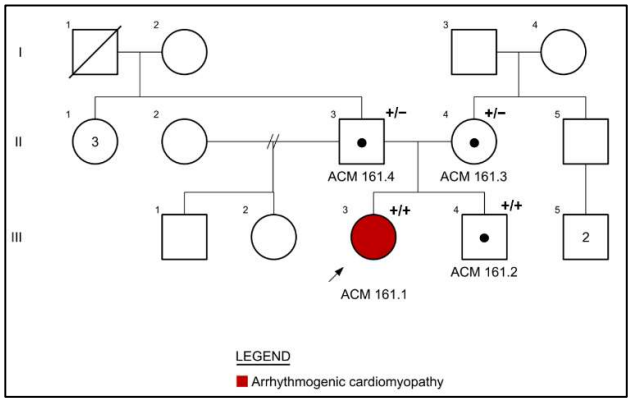


Figure 85: The genotyping results of family ACM 161. The arrow indicates the proband, '+' indicates a positive genotype of the PKP2 variant while '−' indicates a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms

3.3.11.3 Haplotype analysis

The genetics results were followed up by haplotyping. Haplotyping was performed for all four enrolled family members by making use of the microsatellites and the variant of interest (*PKP2* c.1162C>T). The proband's (ACM 161.1) microsatellite marker peaks are shown in Figure 86 below.

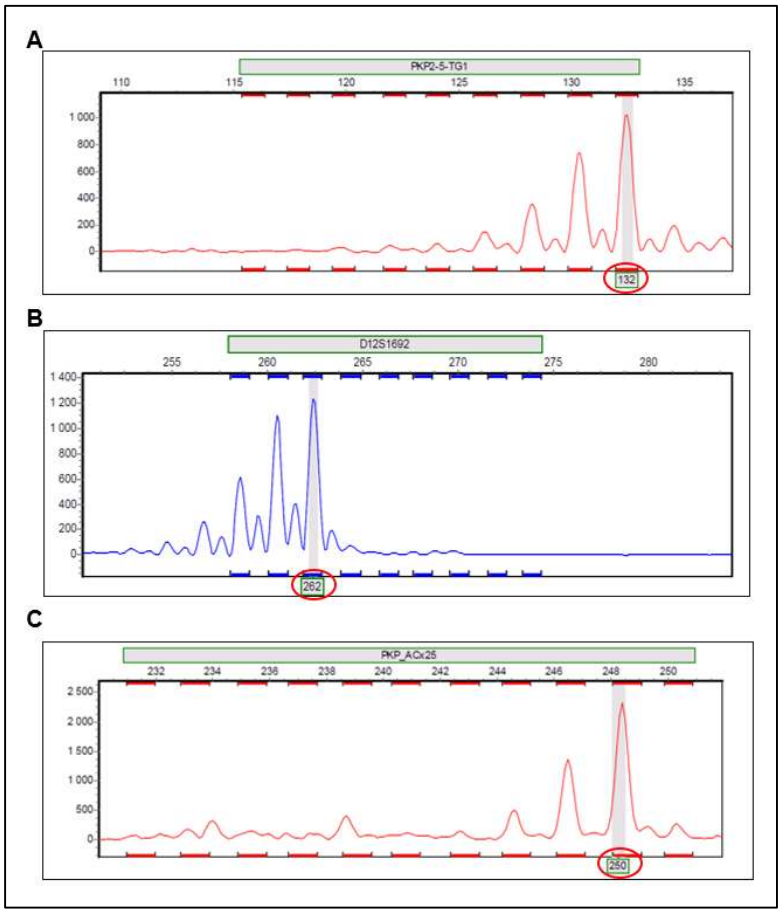


Figure 86: Microsatellite analysis of ACM 161.1. (A) PKP2_TG (132), the peak circled in red, 132, is the common haplotype peak for PKP2_TG. (B) D12S1692 (262), the peak circled in red, 262 is the common haplotype peak for D12S1692 (C) PKP2_AC (250), the peak circled in red, 250 bps, is the common haplotype peak for PKP2_AC

The microsatellite analysis results were combined and used to construct haplotypes for ACM 161.1, ACM 161.2, ACM 161.3 and ACM 161.4 (Figure 87), which showed the parents (ACM 161.3 and ACM 161.4) both had one disease associated haploblock, represented by the red block. They subsequently each passed their disease-associated haploblock down to their children, the proband ACM 161.1 and ACM 161.2, making them homozygous for the disease-associated haploblock.

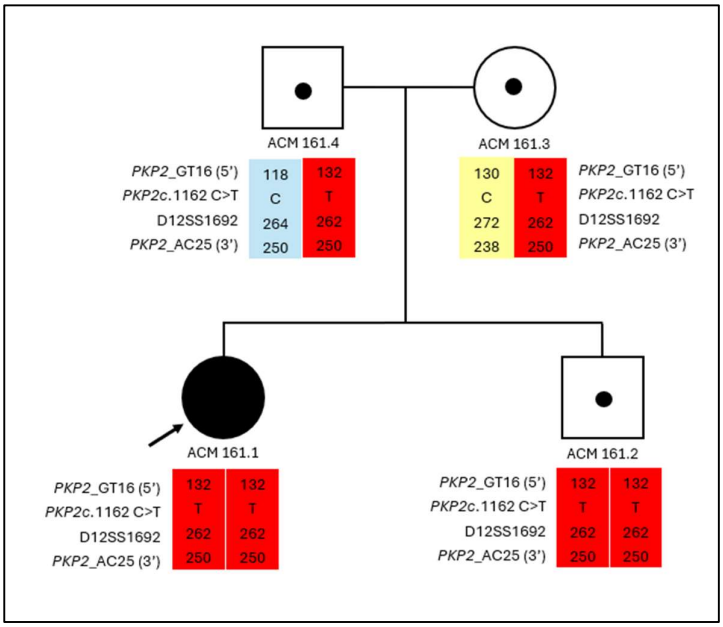


Figure 87: Haplotype analysis of family ACM 161. Circles indicate self-identified females, squares indicate self-identified males, black-filled symbols represent clinically affected individuals, symbols with periods in the centre represent individuals with the PKP2 variant but are not clinically affected, shapes with an oblique line through indicates deceased individuals, haplotypes with dotted borders represent inferred haplotypes

3.3.11.4 Genealogy tracing

With limited information regarding names, dates of birth and locations of the proband (ACM 161.1) and her immediate family provided on the pedigree by the genetic councillor, we first needed to verify the parents (ACM 161.3 and ACM 161.4) and their parents full names and residences by consulting the 1984 voter's roll. Where we identified they resided in Qqeberha.

Once the location was confirmed to trace this family's genealogy, a great deal of information had to be gathered and organised, including copies of birth, marriage, and death certificates; records in churches and graveyards; old articles, letters, photographs, and documents related to citizenship and naturalisation; searches of historical records, including land records and censuses; a backwards and forwards approach was applied. All the relevant records for both the paternal and maternal lineages were search for using online databases (FamilySearch, Geni, Ancestry).

Both the paternal and maternal lineages of both the probands father (ACM 161.3) and mother (ACM 161.4) needed to be extended as they are both carriers of the *PKP2* variant.

We began by extending the proband's father's (ACM 161.3) lineage, which we traced back to the early 18th century to German ancestry. Additional lines went back further to Dutch and French progenitors who arrived in South Africa in the mid-17th century. ACM 161.3's

maternal lineage was not extended further than the early 20th century, there were some inconsistencies with maiden names and birth dates which were not resolved through further investigations, such as contacting churches in the Qqeberha region and consulting published South African Genealogies.

The proband's mother's (ACM 161.4) lineage was traced back to the mid-17th century to French, Dutch and German Ancestry several duplicates were identified, due to consanguineous marriages between cousins. The global distributions of ancestry for family ACM 161 can be found in Figure 88 below.

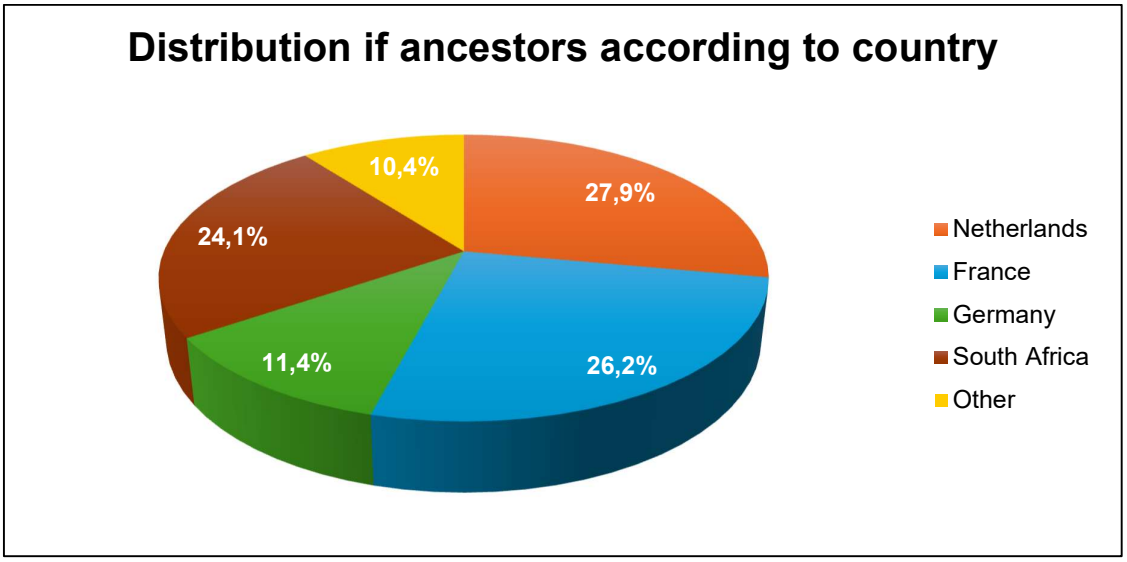


Figure 88: Distribution of ancestors according to country

All data was manually inputted, stored and analysed using Legacy v.10. This family was traced back through 12 generations (shown in Figure 89).

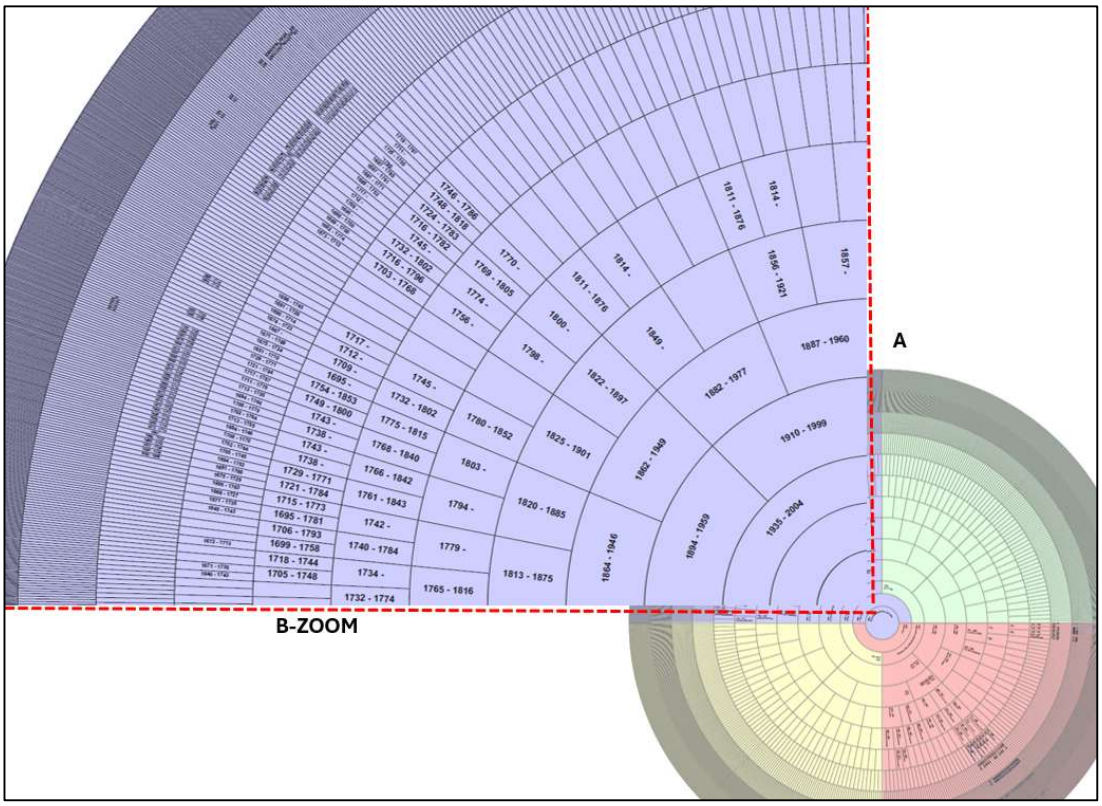


Figure 89: A fan diagram representing the genealogy of Family ACM 161 is presented in fan A, with a zoomed-in section, B, with all names of ancestors removed (dates 1980-1620)

Despite there being some inconsistencies and incomplete lineages for ACM 161, we were still able to trace back this family to several well-known South African progenitors. We therefore compared this family's genealogy to that of the 10 families in sections 3.3.1 – 3.3.10 but found we could not reduce the number of possible founder couples from 8.

Below is a graphical representation (Figure 90) of the origins of family ACM 161, from generation 0 – 13, which roughly coincides with years 1980 -1620.Majority of the ancestors that came to South Africa were found in the 10th generation, there was a distinct cluster of ancestors in the Netherlands and present-day Belgium. The ancestors of ACM 161 were fairly wide-spread across the regions of France, Germany and Switzerland (B-Zoom).The progenitors of ACM 161 immigrated to the Cape of Good Hope in the 16th century where multiple generations remained (10th – 6th generations), thereafter, there was migration inland from multiple ancestors where they settled in Bloemfontein and we find the most recent generations (4th to present day) (C-zoom).

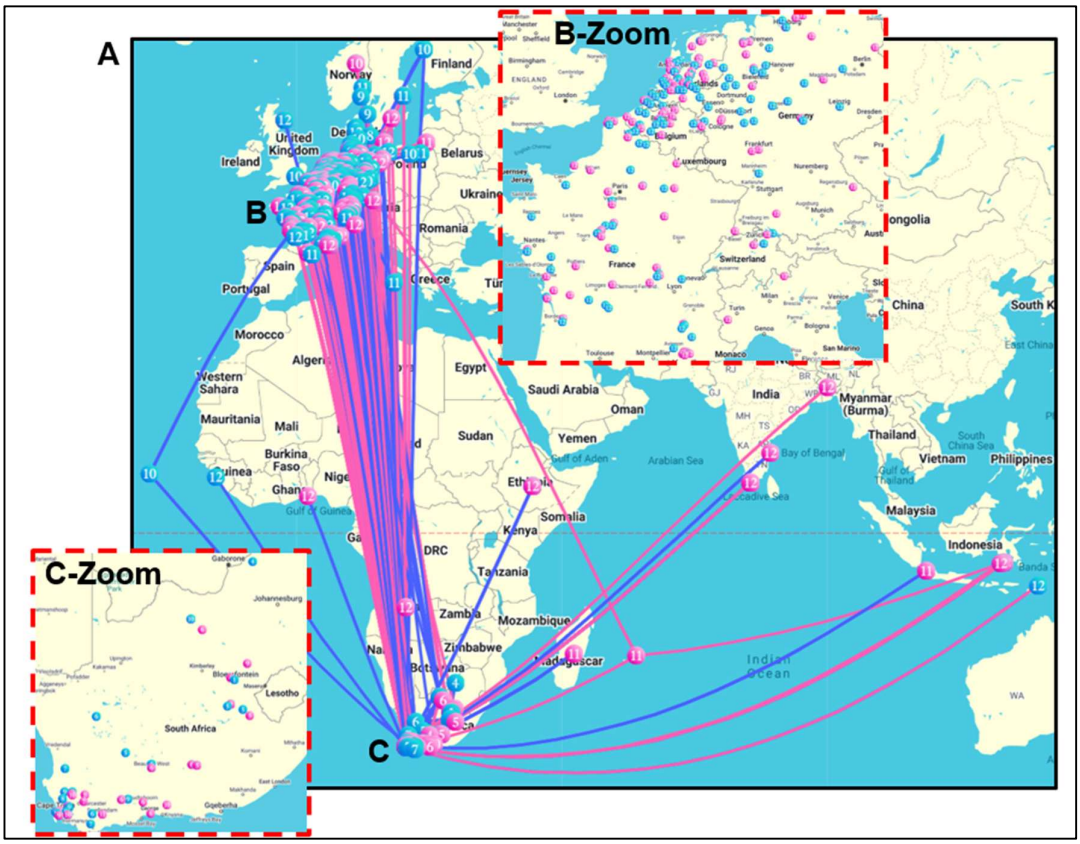


Figure 90: World map showing family ACM 161 through 12 generations and their global distribution. (A) World map with the origins of the progenitors of ACM 161. (B) Zoomed in image of where majority of the progenitors came from in Europe (Netherlands, France and Germany). (C) Zoomed in image of the most recent generations of ACM 161 and their locations across South Africa. Pink dots represent the female ancestors, blue dots represent male ancestors, the numbers inside the dots indicate the generation of that individual, blue lines connect individuals from the paternal lineage and pink lines connect individuals from the maternal lineage (in the case of this family the paternal and maternal lineages of ACM 161.1). Generated with Rootmapper (https://rootsmapper.com/).

3.3.12 Family ACM 71 (old family)

3.3.12.1 Clinical history

Family ACM 71 is the only family of Mixed-Ancestry included in this subsection of this study. Five members of this family were clinically screened (Figure 91) and enrolled into the IMHOTEP study. Only the proband, ACM 71.1, had clinical symptoms of ACM, and presented with a mild phenotype at the age of 38 years old. The additional enrolled family members (ACM 71.2, ACM 71.3, ACM 71.4 and ACM 71.5) were all asymptomatic.

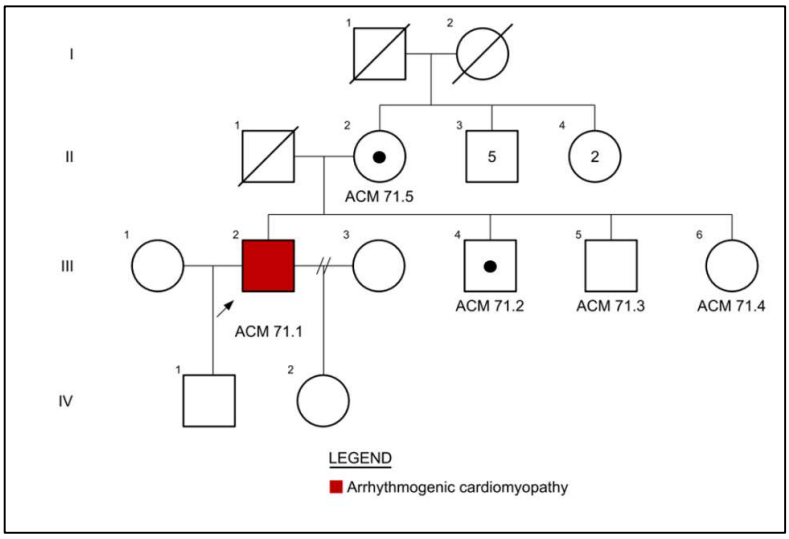


Figure 91: Pedigree of family ACM 71. The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals.

3.3.12.2 Genetics: PKP2 c.1162C>T screening

The proband, ACM 71.1, and four (ACM 71.2, ACM 71.3, ACM 71.4 and ACM 71.5) family members were genetically screened for the *PKP2* c.1162C>T variant. The proband (ACM71.1), his brother (ACM71.2) and mother (ACM17.5) were found to be heterozygous for the *PKP2* c.1162C>T variant. All other members were negative. The probands brother ACM 71.2 is also genotype positive, while the remaining siblings, ACM 71.3 and ACM 71.4 were both genotype negative, as shown in Figure 92.

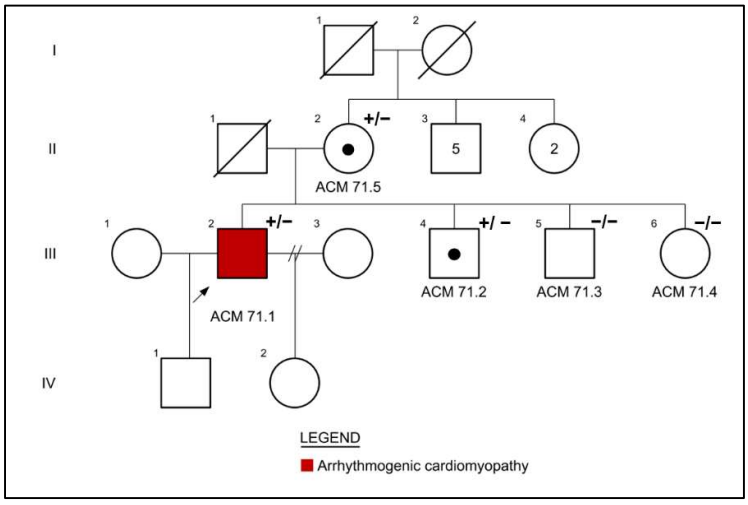


Figure 92: The genotyping results of family ACM 71. The arrow indicates the proband, '+' indicates a positive genotype of the PKP2 variant while '–' indicates a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms

3.3.12.3 Haplotype analysis

The genetics results were followed up by haplotyping. The ACM71 family is a continuation of previous research (Watkins et al 2009) with additional family members being recruited into the current study. New family members testing positive for the *PKP2* c.1162C>T variant was included in the haplotype screening. We included ACM 8 as a positive control as this individual had inherited the red diseased haploblock (132, T, 262, 250), seen in Figure 82, which contained the *PKP2* c.1162C>T variant as well microsatellites associated with disease. We were able to infer the haplotypes from family members if the grandparents, parents or sibling's DNA was available for screening. Thus far, three family members have the disease-associated haploblock, represented by the red block in Figure 93 below.

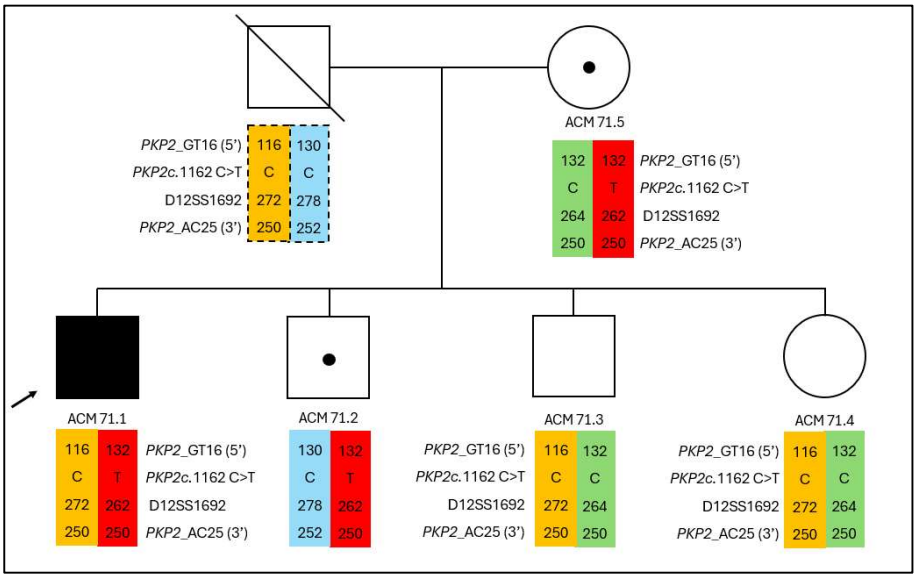


Figure 93: Haplotype analysis of family ACM 71. Circles indicate self-identified females, squares indicate self-identified males, black-filled symbols represent clinically affected individuals, symbols with periods in the centre represent carriers of the PKP2 variant but not clinically unaffected, shapes with an oblique line through indicates deceased individuals, haplotypes with dotted borders represent inferred haplotypes

3.3.12.4 Genealogy tracing

There was limited information provided for genealogical tracing for this family with the information provided only including names, estimated birth dates and residences of the proband (ACM 71.1), his parents and grandparents. Most of the family resided in the Western Cape. However, there were spelling errors on the records of family members due to direct translation from Afrikaans to English.

As the proband inherited the *PKP2* c.1162C>T variant maternally, we expanded both the paternal and maternal lines of the proband's mother (ACM 71.3). This family required alternative methods of research as they were of Mixed-Ancestry and the record-keeping was not as clear the Afrikaner population. Standard methods of genealogical tracing failed to provide results. The period proving most difficult to find information on was that of the Apartheid period. During this time, the government separated all records based on racial grouping. Individuals of some racial groupings often had incomplete or poorly recorded records. The Cape Town Archive does have some records of Mixed-Ancestry individuals however, these records only contain the surnames and no first names.

Extensive investigations were previously carried out by our cardiovascular genetics group, however they could not extend the maternal line of this family as there was a 100-year ban on the records and the search could not proceed past 1914 for the maternal lineage of this family.

However, it was found that the paternal lineage of ACM 71.5 was of Afrikaner descent. There was however no clear lineage, and research did not find any connections to any of the founding couples and was only extended back to the early 20th century.

3.4 Combined Haplotype Analysis and Genealogy Tracing Results

The 12 families are represented in Table 33 below, which summarises the number of family members screened in each family, the number of participants positive for the *PKP2* variant and of those positive, if they have a common haplotype.

Table 33: Study Participants screened for PKP2 c.1162C>T variant and common Haplotype

FAMILY ID	NO. OF STUDY PARTICIPANTS SCREENED	*PKP2* C.1162C>T VARIANT POSITIVE	COMMON HAPLOTYPE
ACM 5	17	11	11
ACM 8	9	5	5
ACM 12	1	1	1
ACM 19	12	8	8
ACM 38	7	3	3
ACM 57	1	1	1
ACM 71	5	3	3
ACM 154	4	2	2
ACM 157	4	3	3
ACM 158	1	1	1
ACM 161	4	4	4
ACM 163	3	2	1
TOTAL PARTICIPANTS	68	44	43

This allowed us to compare the progenitors of each family and narrow it down to eight founder couples (Table 34 below). To maintain anonymity, unique codes were assigned to each potential founder couple. These codes, along with their respective birth locations.

Table 34: A table showing the couple code and birth country of both the male and female counterparts (Key: Couple code (M x F): male x female initials, M: Male and F: Female)

Couple code (M x F)	Birth country (M)	Birth country (F)
LC x FM	France	France
FT x SS	France	France
HP x CD	France	France
PJ x SV	France	France
CM x CT	France	France
WSM x EC	Netherlands	Germany
CP x MF	France	France
JCS x GG	Netherlands	Netherlands
FS x HJ	Netherlands	Netherlands

3.5 Discussion

This study used a combination of methods and resources, clinical history, genetics, haplotype analysis and genealogy tracing, to show that 12 present-day South African families share a common ancestor that was responsible for introducing the *PKP2* C.1162C>T (p.Arg388Trp) variant, associated with ACM, into the population in the 17th century.

3.5.1 Phenotype-genotype

Through the IMHOTEP registry we identified 12 probands presenting with clinical symptoms of ACM, which was identified to be as a result of the *PKP2* c.1162C>T (p.Arg388Trp) variant. We also recruited and screened a total of 56 family members with clinical phenotyping and family histories revealing each family had multiple generations affected with ACM; all at-risk individuals are currently closely monitored by a clinical team. Variable penetrance within and between each family was attributed to factors such as lifestyle and zygosity status.

Of the 68 participants screened, 45 were genotype-positive for the *PKP2* variant and 32 showed varying clinical symptoms while the remaining 14 genotype positive participants were clinically asymptomatic. The absence of clinical manifestation in these participants could be attributed to a predominantly sedentary lifestyle, potentially mitigating an ACM phenotype. Based on these results we were able to identify two distinct phenotypic groups based on the participants' disease severity and the inheritance of one (heterozygote) or two (homozygous) copies of the mutant *PKP2* c.1162C>T allele, with homozygote *PKP*2 c.1162C>T carriers being more severely affected than the heterozygote carriers with a single faulty copy.

In the severely affected group, individuals inheriting two copies of a specific genetic mutation exhibited an extreme form of ACM, characterized by early onset and requiring interventions like heart transplants or ICDs. We were able to draw a genotype-phenotype correlation where we showed that individuals (probands and sometimes their family members) that had inherited two faulty copies of the mutant allele, were more severely affected in the majority of these cases. This group included families ACM 8, ACM 19, ACM 38, ACM 157 and ACM 161. We observed two affected children in the ACM 8 family where the proband required heart transplant in her early 20s and her younger brother died from a

sudden cardiac death at 14 years old. Both siblings were homozygous for the *PKP2* c.1161C>T variant. The proband in family ACM 157 (ACM157.1) was diagnosed at 15 years old and required an ICD a year later; he eventually received a heart transplant. Genetic Screening found he was also homogenous for the *PKP2* c.1162C>T variant. The ACM 161 family, the proband (ACM 161.1) was severely affected by ACM and had an ICD implanted at 15 years old. She was genetically screened and found to be homozygous for the *PKP2* variant. The proband's brother (ACM 161.2) was also found to be homozygous carrier but was still asymptomatic. This was not unexpected as he was still young; he receives regularly checkups and is closely monitored by the team of physicians. As a precaution, his sports activity is also monitored as it is known that endurance sports exacerbates ACM (183, 184) Similarly in the ACM 19 family, two severely affected children (ACM 19.2 and ACM 19.1) were clinically screened and diagnosed at ages 12 and 14 and required heart transplants at the ages of 15 and 17 years respectively. Unfortunately, the probands brother (ACM19.1 later suffered complications and died. Genetic screening revealed the siblings were compound heterozygotes, having both the *PKP2* c.1162C>T variant inherited from their mother and an additional *PKP2* c.2197-2202 Ins/Del which was inherited from their father (54). Research has shown there is increased penetrance in patients who are compound heterozygotes (having two variants in the same gene associated with the disease) and in digenic heterozygous patients (variants in two different genes associated with the disease) (185-187). The final family, ACM 38 had three severely affected children with the one affected dying at 8 years old, of a SCD before the family were enrolled into the study. The proband, ACM 38.3 presented with ACM pre-puberty and required a heart transplant at 16 years old, while his brother had a transplant at 8 years old. Genetic screening i found both affected to be heterozygous carriers for the *PKP2* variant, which was inherited from their mother, who has a history of several miscarriages. This heterozygous variant does not explain the disease severity and while incomplete penetrance could be a factor, we hypothesise that there could be a second, yet unknown, variant that would account for the disease severity seen within this family.

In the second group, the affected presented with mild to moderate disease with delayed (or late) onset with individuals who inherited one copy (heterozygote) of the *PKP2* c.1162C>T variant, appearing to be more mildly affected. These included family members from ACM 5, ACM 12, ACM 57, ACM 71, ACM 154, ACM 158 and ACM 163. The proband ACM 163.1, was a marathon runner prior to her diagnosis at 53 years old; proband ACM 12.1 was also a

marathon runner and presented at 45 years old; proband, ACM 57.1 also participated in endurance sport before her presentation at 42; proband, ACM 154.1 showed mild symptoms at the age of 52proband ACM 158.1 also participated in endurance sport and continued even after his with the aid of a ICD but did reduce his exercise load; ACM 5.1 was a competitive professional athlete before her diagnosis at 44 years old and finally, proband ACM 71.1 presented at 38, however, he was a manual labour on a wine farm and his clinical records showed he had hypertension and was a smoker.

Since the majority of the presenting/affected family members had a history of prolonged athletic involvement, we believe that lifestyle choices, such as endurance training, as well as the presence of two affected alleles (whether homozygous or digenic) may have exacerbated or altered the incomplete penetrance of ACM, as observed in other ACM studies (184, 188).

Historical death records indicate that many progenitors and their offspring had extended lifespans. When compared to their modern-day descendants, this observation might appear counterintuitive however, it is important to consider that many individuals during the 17th century employed enslaved labour (189), which significantly reduced their physical activity levels. This sedentary lifestyle contrasts with contemporary trends, where research highlights the long-term health benefits of physical activity (183, 184, 190), and the growing popularity of endurance sports in the 21st century. These factors likely contribute to the observed increase in mild to moderate phenotypic expression among middle-aged individuals in our cohort, particularly concerning arrhythmogenic cardiomyopathy (ACM).

Of interest to us was the *PKP2* c.1162C>T variant that was found in 12 families, eleven of which self-identified as Caucasian Afrikaners; the last one was of Mixed-Ancestry. These findings suggested these families might share a common ancestor that came to the Cape centuries ago and brought this *PKP2* c.1162C>T variant to South Africa. Previous research into the present-day Afrikaner population of South Africa has revealed previous founder variants such as HCM (70), pseudoxanthoma elasticum (PXE) (146), Long-QT syndrome (101), porphyria variegata (154), Huntington's disease (HD) (155) and Fanconi's anaemia (152). We then proceeded to construct haplotypes and genealogy tracing on each of the families with the *PKP2* c.1162C>T variant to determine if there was a common founder that could be traced back to when South Africa was colonised by the Europeans.

3.5.2 Haplotyping

Haplotypes were constructed to span the entire *PKP2* gene by selecting microsatellite markers that flanked the 5' (PKP2_TG), 3' (PKP2_AC) and intergenic region (D12S1692). We also combined this data with the *PKP2* c.1162C>T genotyping data and proceeded to construct the disease haploblock.

Previous research in our laboratory on seven families (ACM 5, ACM 8, ACM 12, ACM 19, ACM 38, ACM 57 and ACM 71) showed that the *PKP2* c.1162C>T variant was a founder variant (54), however the progenitor in the 1600s was not found as we had insufficient families to genealogically trace. To identify a founder, a minimum of 10-12 genealogically traced families is often recommended (104). In this study, we identified an additional five families (ACM 154, ACM 157, ACM 158, ACM 161 and ACM 163), bringing the total to 12 families. We used the previously available microsatellite markers (157, 160) and reconstructed the disease haplotypes for the old families. It must be noted that, through the IMHOTEP registry, both the old and new families, saw prospective recruitment of at-risk family members, so all 12 families data kept expanding.

We used ACM 8 as a positive control (for the old haploblock with the old microsatellite and as the new haploblock with the new 3' PKP2_AC microsatellite. The microsatellites markers were used to form haplotypes for all the families, and the results confirmed that all 12 families had the same disease haploblock, old and new. This provided further evidence that there was a common ancestor as the haplotype was identical by descent (IBD). The search continued and we used genealogical tracing to identify the progenitors couples who could have been responsible for introducing the pathogenic *PKP2* c.1162C>T variant into the South African population.

3.5.3 Genealogy

All participants who were screened and self-identified as Afrikaners had supporting genealogical evidence that suggested they were of Afrikaner descent. Our results showed that the haplotype data was strongly supported by the genealogical data for 11 of the 12 families.

For our study, many genealogies showed varying percentages of French, Dutch and German ancestry, however, for the progenitors dating back to the 17th century, we found mainly French and Dutch ancestors. In family ACM 158, the proband was male and came from a

well-known French Huguenot lineage, which could be traced through his present-day surname, this resulted in a significantly higher number of French ancestors in his lineage. Interestingly, ACM 12.1, was a male with the same French Huguenot surname as ACM 158.1, however, their most recent common ancestor arose in the early 18th century (over three centuries ago), through their paternal lineages. Both these families therefore have more French ancestry than Dutch, however they were still considered Afrikaner, as there are many other branches of their genealogies which showed Dutch and German ancestry as well as admixture between French, Dutch and German progenitors, which constitutes the general demographic make-up of present-day Afrikaners.

The Afrikaner population is believed to stem from a complex admixture between Dutch, French and German progenitors and there have been many investigations into this admixture. The first was by JA Heese, published in 1971, this book is an in-depth analysis of admixed marriages, which spanned across two centuries and it was concluded that the general demographic profile of the Afrikaner population was one-third Dutch ancestry, a quarter French and the remainder was attributed to German ancestry (156).

We were able to trace 11 families, both fully and partially and through a rigorous process of analysis and comparison between the 11 families, we identified eight common progenitors between these families. We believe there are two main explanations as to why our results pointed to the *PKP2* c.1162C>T variant originated from eight progenitor couples. Firstly, the extensive shared ancestry observed within and between these 11 families suggests the potential presence of a common genetic factor, such as a founder variant. While other explanations remain possible, the high degree of shared ancestry observed strengthens the hypothesis of a single ancestral origin for this variant. Secondly, there were three families (ACM 12, ACM 57 and ACM 158), where only the proband participated. This could have led to the inclusion of redundant lineages, which only increased the complexity of the genealogical data collected. Reducing one parent's lineage could have simplified the analysis, as the number of lineages increases exponentially with each generation (2^n).

Through our genealogical analysis we identified a higher proportion of French ancestry compared to Dutch ancestry among the Afrikaner families studied (ACM 163, ACM 12, ACM 154, ACM 19, ACM 8, ACM 57 and ACM 158). This pattern supports the hypothesis that the *PKP2* c.1162C>T variant may have been introduced to South Africa by a French

founder, which is further supported by genealogical results which show that both the male and female counterparts of six of the eight founder couples are French Huguenot progenitors.

It is important to consider the historical admixture between the Dutch and French. Upon their arrival in the Cape of Good Hope many French progenitors and their descendants intermarried with Dutch progenitors. When we delve even further into the history as to why they left Europe, we found that thousands of French Huguenots migrated to the Netherlands as early as 1620 as they were fleeing religious persecution (191). During this time, the Dutch provided refuge to the Huguenots, fostering integration and admixture, which may have been the beginning of the dissemination of the *PKP2* c.1162C>T variant into the population.

Jan van Riebeeck came to the Cape of Good Hope, in 1652, as described Chapter 1, section 1.2.2. The arrival of the founder couples in the Cape of Good Hope spanned a period of ~20 years 1657 to 1699, noting that many wives and their children only came later (FT x SS, WSM x EC and JCS x GG). The French Huguenot founder couples all immigrated to the Cape of Good Hope in 1688 onboard Dutch ships which departed from various ports around the Netherlands (LC x FM, FT x SS, HP x CD, PJ x SV, CM x CT and CP x MF), suggesting they had been in the Netherlands as refugees for some time prior to the voyage. Thereafter, the children of the French Huguenot, Dutch and German progenitors married, further propagating the *PKP2* c.1162C>T variant into the population, which at this stage was limited in number.

This historical context suggests that French and Dutch populations began intermingling long before arriving at the Cape of Good Hope. Given that French migration to the Netherlands began around 1620, it is plausible that the *PKP2* variant is significantly older than previously estimated.

As a side note, in 2016, we collaborated with a French researcher that had identified a French family with the *PKP2* c.1162C>T variant (unpublished data) and they wanted to determine if the French family and Afrikaner families shared a common haplotype. We constructed the haplotypes and analysed the data and although this family had the same *PKP2* variant, they did not have the same haplotype (unpublished data).

This non-identical haplotype further supports the hypothesis that the *PKP2* variant is much older than previously estimated, likely originating in Europe centuries ago. Over time, recombination events occurring across multiple generations may have led to the divergence

of the common haplotype observed in present-day Afrikaners from that found in present-day French individuals.

Using molecular dating methods such as the Gamma method (192), where hundreds of densely packed SNPs are screened for a small cohort individuals makes it is possible to date a rare variant. Future investigations into the dating of the *PKP2* c.1162C>T could provide valuable insight to the exact age of this pathogenic variant.

3.6 Conclusion

The identification of 12 families with the *PKP2* c.1162C>T pathogenic variant, with an identical haplotype within and between families as well as genealogical evidence that 11 of the 12 families are of Afrikaner descent with eight progenitor couples in common supports our hypothesis that the *PKP2* c.1162C>T variant is a founder variant introduced into the Afrikaner population over three centuries ago. Further investigation into the exact age of the variant could provide valuable insights into its origin and explain the loss of a common haplotype through geography and time.

Chapter 4: *BAG3* c.925C>T (p.Arg309Ter) pathogenic variant in South African Families with DCM

4.1 Introduction

Dilated cardiomyopathy (DCM) is the most common and genetically diverse cardiomyopathy. As described in Chapter 1, the main genes associated with DCM are sarcomere and cytoskeletal genes.

Over the past decade, researchers have identified that mutations in the *BAG3* gene is associated with dilated cardiomyopathy (193). The genetic structure and function of the *BAG3* gene has been well researched. The *BAG3* gene codes for Bcl-2-associated anthanogene 2 protein which plays an important role in maintaining the structural integrity of the sarcomere (194). The protein has four functional domains, namely the WW (Trp-Trp), two IPV (Ile-Pro-Val), proline-rich repeat (PXXP) and the BAG3 domain which interact with CapZ, the actin capping protein, (193) thereby providing an important role in maintaining the structural integrity of the sarcomere (194) (Figure 94, adapted from (195)) . BAG3 has a multifactorial role in cardiomyocytes as the four domains play vital roles in regulating macroautophagy by removing misfolded proteins for degradation (196). BAG3 also forms a complex with Hsp70 to protect antiapoptotic Bcl-2 proteins from degradation (197) and has a vital role in the excitation and contraction coupling mechanism by mediating Ca2+ release (198). This protein also acts as a quality control in mitochondrial regulation of endogenous and exogenous expression of major regulators of mitophagy (199).

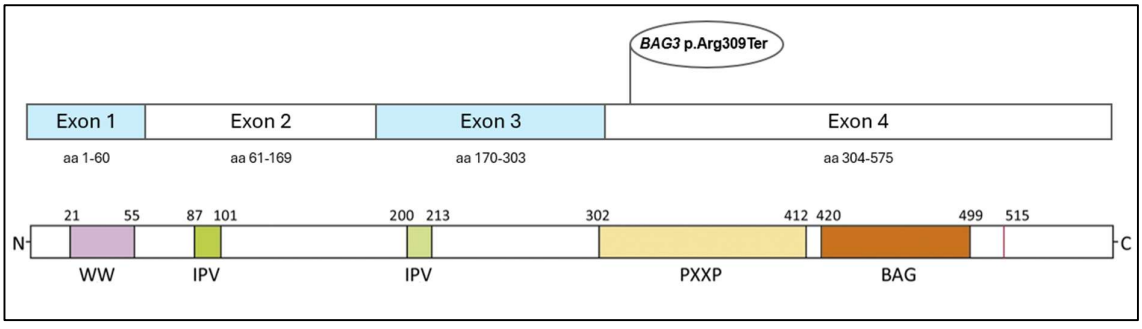


Figure 94: Schematic representation of BAG3 coding region and protein domains. The location of the BAG p.Arg309Ter variant is circled, it is shown in the context of the coding (exon) region (aa: amino acid). Below is the functional protein domains (196)

A study of 129 participants with *BAG3* variants across European centres found that patients with DCM caused by these variants had a high risk of early-disease onset, rapid progression to end-stage HF, and poor prognosis .Clinical endpoints for 30.1% of patients included

cardiac death, heart transplant, LV assist device, SCD and ventricular arrhythmia (193, 194). Additionally, variants in this gene have shown high penetrance within families and account for 15% of DCM cases in French-Canadians (109). *BAG3* variants present with a plethora of DCM phenotypes and age of onset ranges from 18 to 64 years old (200, 201). Despite the ongoing investigations into the *BAG3* gene, reports on founder variants in this gene are limited.

Through the IMHOTEP registry of South Africa, we found three probands to harbour the pathogenic *BAG3* c.925C>T (p.Arg309Ter) variant (202). This variant has a truncating effect in the PXXP domain, a protein-protein binding domain, of the BAG3 protein (Figure 97).

This nonsense variant was classified as pathogenic according to ACMG guidelines and research has shown that the *BAG3* c.925C>T variant was rare as it occurs at a population frequency of 0.0004% according to gnomAD. There have been no other reports of this variant in any other population frequency databases such as 1000Genomes, ExAC and ALFA. This variant was previously identified in three French-Canadian families (109).

As this variant was previously identified in a specific population, we hypothesised that all three South African probands shared a common ancestor. The aim of this chapter was to identify if a common haplotype was present within and between the three unrelated probands and their extended families. Additionally, where possible, we sought to trace each family's genealogy as far back as possible to identify the founder of the *BAG3* c.925C>T variant.

4.2 Methods

All methods were outlined in Chapter 2, the flow diagram below (Figure 95) represents the overall flow of chapters 3-6, which has been tailored to this chapter and refers to the relevant methods.

.

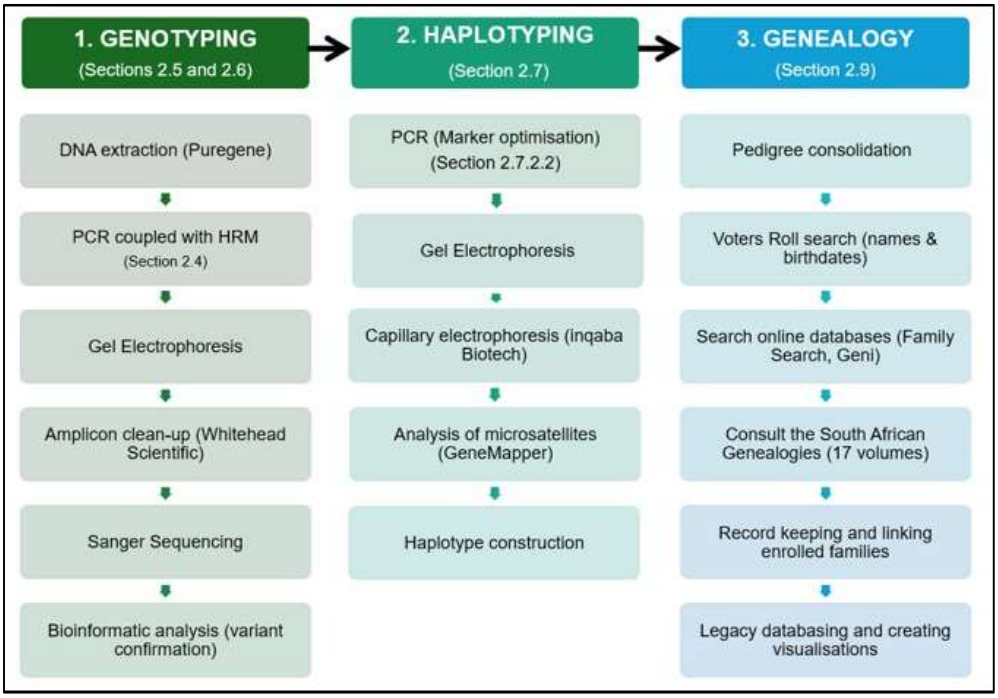


Figure 95: Workflow, adapted from Chapter 2, Section 2.3, to highlight the relevant sections detailing the methods associated with this chapter, specifically focusing on the BAG3 c.925C>T variant

4.2.1 Clinical

Patients with suspected DCM, were referred to the Cardiac Clinic at Groote Schuur Hospital. Once their diagnosis was confirmed they were enrolled into the IMHOPTEP registry the inclusion and exclusion criteria can be found in Chapter 2, section 2.2.

4.2.2 Genetics

The primer sequences and methodology for the *BAG3* c.925C>T variant can be found in Chapter 2, section 2.4, Table 4. The probands were screened and were identified as such if they were positive for the *BAG3* c.925C>T variant. Thereafter, through continuous screening and recruitment, additional family members were enrolled into the registry by their clinician and/or genetic councillor.

4.2.3 Haplotype

A haplotype is a set of DNA variations, or alleles, inherited together on the same chromosome. These alleles are physically linked and tend to be passed down as a unit. Think of it as a specific combination of genetic variants on a single chromosome. Haplotypes are

essential for understanding genetic inheritance, mapping disease genes, and studying population history (180).

Inferring haplotypes within a family becomes easier when the DNA of the parents or siblings are available as their DNA allows the direct observation of the haplotypes transmitted. If the parents' genotypes are known, the child's haplotypes can be deduced with greater certainty (181). Similarly, sibling DNA can help resolve ambiguous haplotypes. By comparing the genetic variations shared between siblings, segments of their chromosomes inherited from each parent can be inferred. This information is particularly useful in linkage studies, where researchers track the inheritance of specific haplotypes within families to identify genes associated with diseases (181, 182).

We made use of three microsatellite markers, BAG3_AC19 at the 5' end BAG3_17TG_11KB at the 3' end AND BAG3_25TG within the *BAG3* gene, thus spanning the entire gene. These markers were combined with the variant *BAG3* c.925C>T to construct haplotypes for the probands and their extended families.

4.2.4 Genealogy

Genealogy tracing followed standard procedures (Chapter 2, Figure 14). The oldest enrolled family member was the starting point of genealogy tracing. The tracing and where possible the ancestral lineages were extended as far as records allowed.

4.3 Results

Throughout this chapter we describe the clinical results, which was followed by the genetics, haplotyping and genealogy tracing of the families with the *BAG3* variant.

Patients showing clinical symptoms of DCM were referred to the Cardiac Clinic at Groote Schuur Hospital during the period 2012 to 2024. After informed consent was taken, they were enrolled into the IMHOTEP registry of South Africa. We identified three probands as well as 16 at-risk family members. Two of the probands and their family members self-identified as Mixed-Ancestry while the remaining proband and her family self-identified as Caucasian Afrikaners.

Genetic screening was initially performed on the three probands, thereafter 16 at-risk family members screened. We identified a total of 13 participants that were heterozygous the *BAG3* c.925C>T variant.

The three families had haplotypes constructed by making use of a combination of three microsatellite markers *BAG3*_11Kb, *BAG3*_AC and *BAG3*_TG marker and the *BAG3* c.925C>T variant. They therefore spanned across the entire *BAG3* gene as well as the 5' and 3' regions.

Haplotype analysis enabled the inclusion or exclusion of ancestral lines during genealogical tracing based on the known autosomal dominant inheritance pattern.

Genealogy tracing followed standard procedures (Chapter 2, Figure 14) with the oldest enrolled family member used as the starting point of genealogy tracing. The tracing were extended as far as records allowed. The families were presented in the order of most to least genealogically traced (DCM 343, DCM 181, DCM 307).

4.3.1. DCM 343

4.3.1.1 Clinical History

We enrolled and clinically screened a large six generation family with 12 at-risk members (Figure 96).

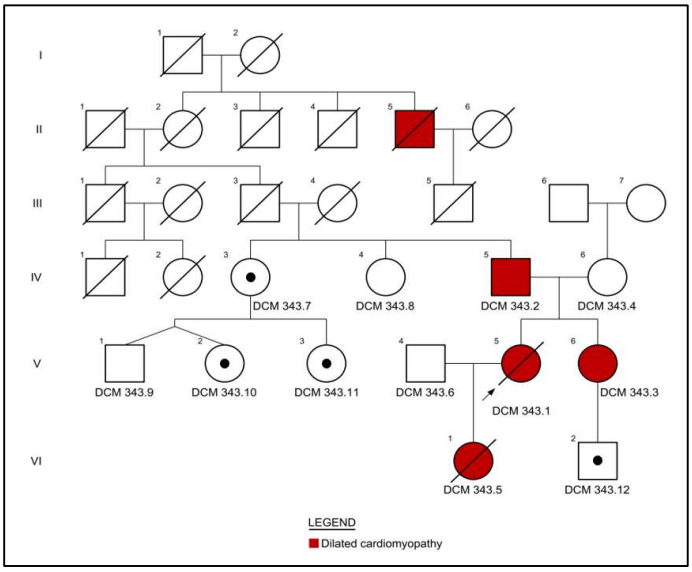


Figure 96: Pedigree of family DCM 343. The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals.

The proband (DCM 343.1) was first diagnosed at 28 years old. She acquired a severe viral infection and that went into HF with complications which included atrial fibrillation and severely impaired LVEF. She was successfully treated with ACE inhibitors and ARBs and made a full recovery. The probands daughter (DCM 343.5) was diagnosed with DCM-LVNC overlap at the age of nine. She presented with acute HF and cardiogenic shock following a viral gastroenteritis infection. DCM 343.5 passed away after suffering cardiac arrest and subsequent neurological damage. After the passing of her daughter the proband discontinued her medication and subsequently went into HF and passed away at the age of 36 years old. The proband's father (343.2) was diagnosed with DCM at the age of 65 years old however, clinical notes also show he was previously diagnosed with hypertension and atrial fibrillation for which he was taking prescribed ACE (angiotensin-converting enzyme) inhibitors. His symptoms of DCM were described as very mild and is currently doing well. The probands mother (DCM 343.4) and her husband (DCM 343.6) showed no clinical symptoms of DCM. The proband's eldest sister (DCM 343.3), was diagnosed with DCM-LVNC overlap at 37 years old and received the standard treatment; she made a fully cardiac recovery. Her son (DCM 343.12) was clinically unaffected at the time of screening.

Clinical investigations into the history found that the proband's paternal line had an extensive family history of heart problems. The proband's two paternal great-aunts (DCM 343.8 and

DCM 343.7) and their respective children were all found to be clinically unaffected, except for DCM 343.10 (the daughter of DCM343.7) who had recent-onset tachycardia and still requires further investigation. DCM 343.10 is one of a fraternal set of twins, and clinical investigations into the history also noted that the proband's paternal grandfather (III:3), was diagnosed with DCM at the age of 28 and died from HF at 29 years old (according to clinical records dating back to 1965) after acquiring the coxsackie viral infection which caused myocarditis. The clinical notes also showed that there were several relatives from the proband's paternal family that had heart-related issues, such as a great-uncle (III:1) that was diagnosed with DCM and later received a heart transplant (confirmed on pathology notes on the explanted heart). Additionally, the proband's great-great-great grandmother (I:2) from the paternal line, had died at the age of 52 due to HF and her death certificate showed the cause of death as "myocardial degradation". Also noted was a child death at the age of 4 due to pneumonia and subsequent HF, which was hypothesised to be related to a cardiomyopathy (II-4).

4.3.1.2 Genetics: BAG3 *c.925C>T screening*

We consented and genotyped three generations (IV-V) and 12 members of DCM 343, for the *BAG3* c.925C>T variant; all primer information was recorded in Chapter 2, Section 2.5, Table 4.

The proband (DCM 343.1), her daughter (DCM 343.5), her father (DCM 343.2), her sister (DCM 343.3) as well as her nephew (DCM 343.12) were all found to be heterozygous for the *BAG3* c.925C>T variant. We also found that one of the proband's paternal aunt's (DCM 343.7 and two of her children (DCM 343.10 and DCM 343.11) were heterozygous for the *BAG3* c.925C>T variant. The remaining enrolled family members were all genotype negative: the probands mother (DCM 343.4), the probands husband (DCM 343.6), the probands younger paternal aunt (DCM 343.8) and the probands cousin (DCM 343.9) (Figure 97).

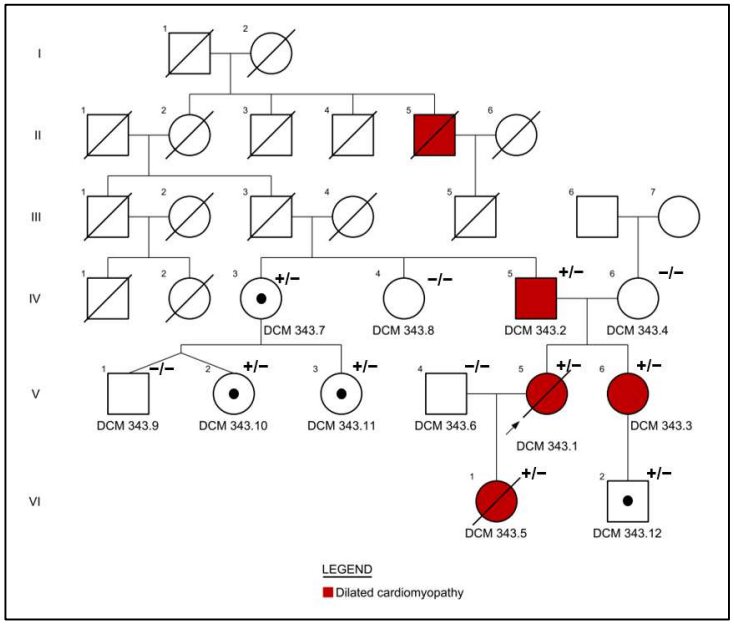


Figure 97: The genotyping results of DCM 343. The arrow indicates the proband, '+' indicates a positive genotype of the BAG3 variant while '−' indicates a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms

4.3.1.3 Haplotype Analysis

The genetics results were followed up by haplotyping. We prospectively recruited and enrolled all at-risk members of the DCM 343 family and members that tested positive for the *BAG3* c.925C>T variant was included in the haplotype screening. Through haplotyping we used a combination of the microsatellites described in Chapter 2, Section 3.5.3.2 and the variant of interest (*BAG3* c.925C>T). DCM 343.2's microsatellite marker peaks are shown in Figure 98 below.

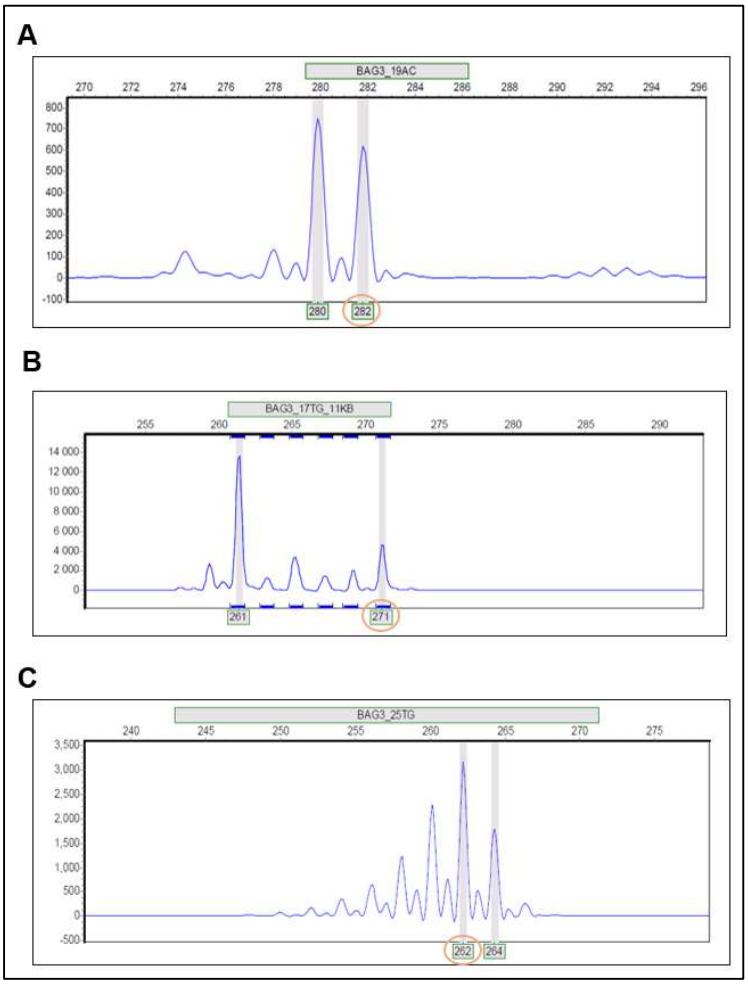


Figure 98: Microsatellite analysis of DCM 343.2. BAG3_19AC (282), the peak circled in brown, 282, is the common haplotype peak for BAG3_19AC. BAG3_17TG_11KB (271), the peak circled in brown, 271, is the common haplotype peak for BAG3_17TG_11KB. BAG3_25TG (262), the peak circled in brown, 262, is the common haplotype peak for BAG3_25TG.

The proband, DCM 343.1, her daughter (DCM 343.5), her sister (DCM 343.3), her nephew (DCM 343.12) and her father (DCM 343.2) were all positive for the *BAG3* c.925C>T variant all were shown to carry the disease associated haploblock, represented by the brown shaded

haploblock (262, 271, T, 262) (Figure 99). We could perform genotyping for DCM 343.8 however, microsatellite analysis was more difficult to perform due to the poor quality of the DNA As the genotyping showed that DCM 343.8 was negative for the *BAG3* variant, we were able to infer her haplotype through her sibling's marker information. Additionally, haplotype analysis showed that DCM 343.7 and her two daughters (DCM 343.10 and DCM 343.11) all also carried the disease associated haploblock.

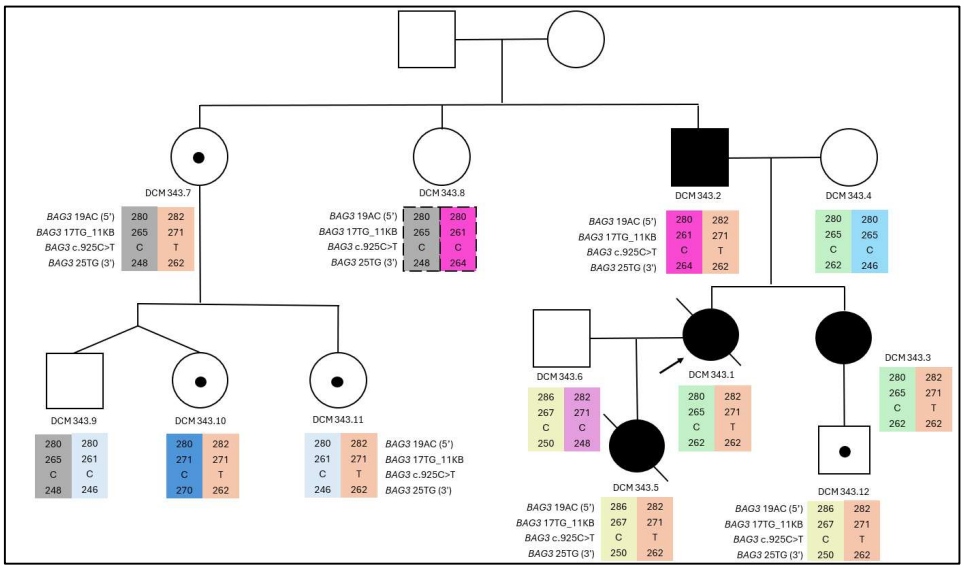


Figure 99: Haplotype analysis of family DCM 343. Circles indicate self-identified females, squares indicate self- identified males, black filled symbols represent clinically affected individuals, symbols with periods in the centre represent individuals with the BAG3 variant but are not clinically affected, shapes with an oblique line through indicates deceased individuals, haplotypes with dotted boarders represent assumed haplotypes from segregation analysis

4.3.1.4 Genealogy Tracing

To trace this family's genealogy, a substantial amount of information had to be gathered and organised, including copies of birth, marriage, and death certificates; records in churches and graveyards; old articles, letters, photographs, and documents related to citizenship and naturalisation; searches of historical records, including land records and censuses in a forward and reverse approach, starting with the proband.

The proband (DCM 343.1) and her immediate family all self-identified as Caucasian Afrikaners which meant it would be easier to trace this family as the Afrikaners are known to have well-documented records in terms of birth, death and marriage records. This family represented the most completely genealogically traced family, with all lines extended back to the progenitors.

With limited information provided by the genetic councillor pertaining to names, dates of birth and locations of the probands (343.1) family, we started by confirming names and dates of birth by making use of the Voter's Roll and various online resources. The proband and her immediate family resided in the Western Cape, this included her parents (DCM 343.2 and DCM 343.4) and her sister (DCM 343.3).

The genotype information showed the *BAG3* variant was paternally inherited by the proband and her sister therefore, the paternal and maternal lineages of the proband's father was extended. The paternal lineage was traced to the late 18th century English ancestry, with evidence suggesting that several progenitor couples originated from England and then immigrated to South Africa. There was also evidence of French-Canadian, Dutch, German and French descent in this line (Figure 100). The maternal lineage of DCM 343.2 was not traced past his mother as there were no available records to confirm her date of birth or place of birth.

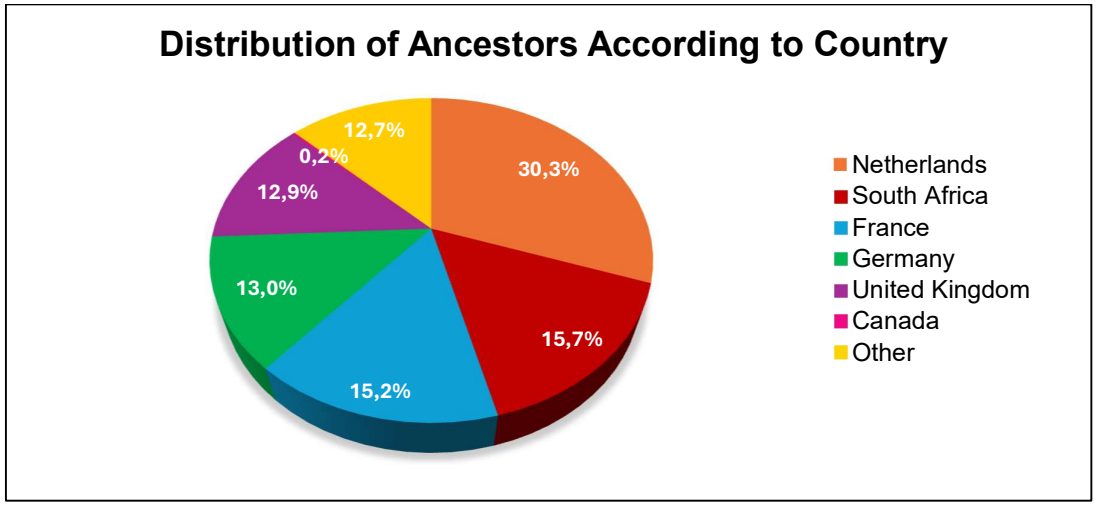


Figure 100: Distribution of ancestors according to country

All data was input manually, stored and analysed using Legacy v.10. The DCM 343 family's paternal lineage was traced back through 12 generations (Figure 101).

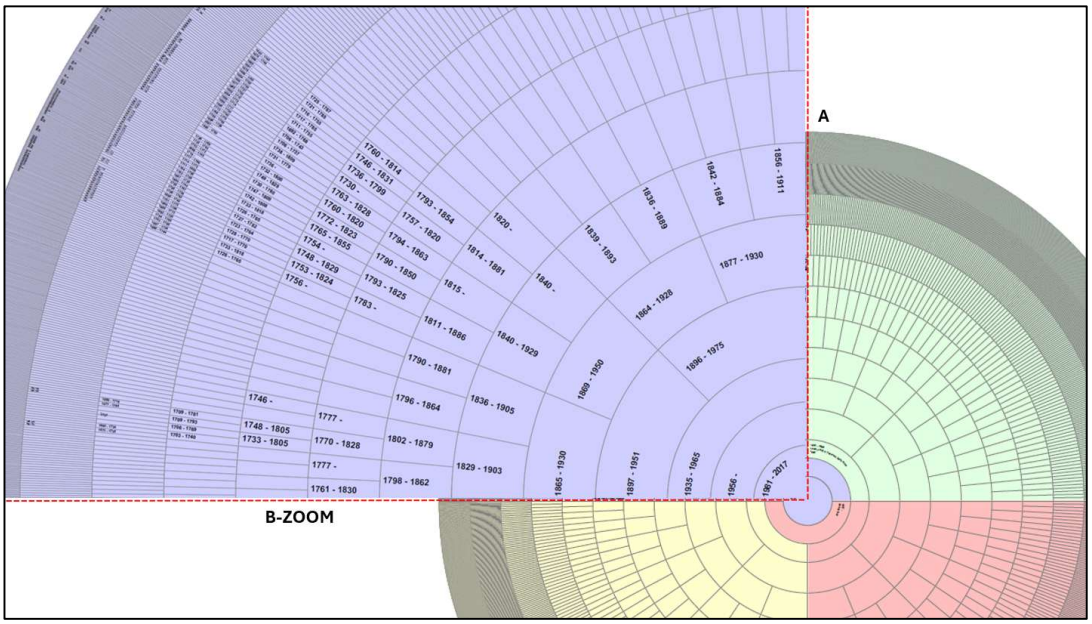


Figure 101: A fan diagram of the genealogy of Family DCM 343 presented in A, with a zoomed in section, B, with all names removed of ancestors removed (dates 1981-1614).

On completion of the ancestral tracing for DCM 343, we had identified 54 potential founders and were able to trace the progenitors to the 15th century. Below is a graphical representation (Figure 102) of the origins of family DCM 343. As can be seen, most of this family's ancestors that came to South Africa were traced to Europe in the 17th century to (generations 7 – 12), where we see distinct clustering of ancestors in France, Germany and the Netherlands (B-Zoom) (as discussed in Chapter 3). In image A we also see a small clustering of ancestors in Canada (generation 5), who came to South Africa. Majority of this family have remained in the Western Cape since the arrival of the progenitors (generation 11) of this family in the 16th century (C-Zoom).

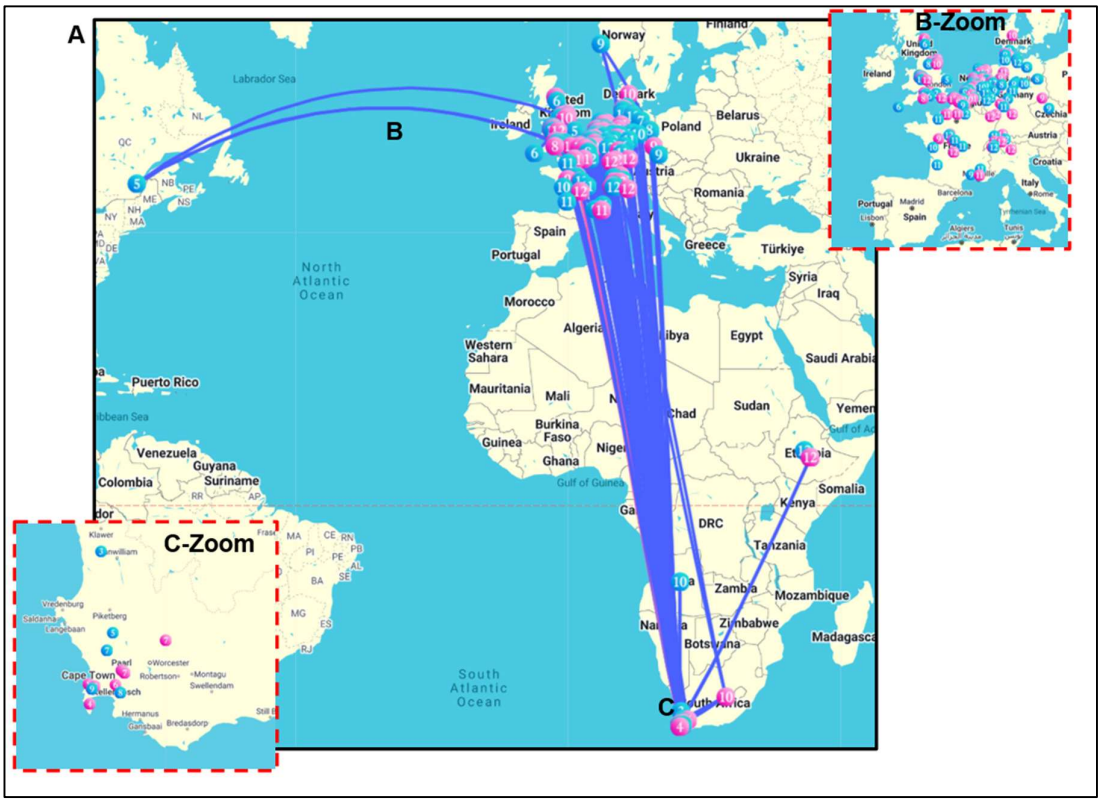


Figure 102: World map showing family DCM 343 through 12 generations and their global distribution. (A) World map with the origins of the progenitors of DCM 343. (B) Zoomed in image of where majority of the progenitors came from in Europe (Netherlands, France and Germany). (C) Zoomed in image of the most recent generations of ACM 343 and their locations across South Africa. Pink dots represent the female ancestors, blue does represent make ancestors, the numbers inside the dots indicate the generation of that individual, blue lines connect individuals from the paternal lineage and pink lines connect individuals from the maternal lineage (in the case of this family the paternal and maternal lineages of DCM 343.2) Generated with Rootmapper (https://rootsmapper.com/).

4.3.2. DCM 181

4.3.2.1 Clinical History

The proband (DCM 181.1) was first diagnosed with DCM at the age of 59 years old and had a left ventricular ejection fraction (LVEF) of 18%. He was subsequently diagnosed with type II diabetes, hypercholesteremia and transient ischemic attack (TIA), after being placed on medical treatment we observed improvements for his HF symptoms as well as LV function. However, a year later he developed end stage HF. The proband's daughter (DCM 181.2) and her two children (DCM 181.3 and 181.4) were also clinically screened and both DCM 181.2 and her youngest son, DCM 181.4, were noted to be clinically affected; DCM 181.3 was clinically unaffected (Figure 103). Clinical investigations into the history noted that there was a significant family history of HF and suspected DCM as the proband's younger sister (II-3) was diagnosed with DCM and later died of HF. Additionally, the daughters, DCM 181.2 and

III-2, as well as one of his sons (III-3) were diagnosed with DCM; unfortunately, III-3 also had a history of drug abuse and died at the age of 51. His death was suspected to be due to a combination of drug abuse and his DCM diagnosis.

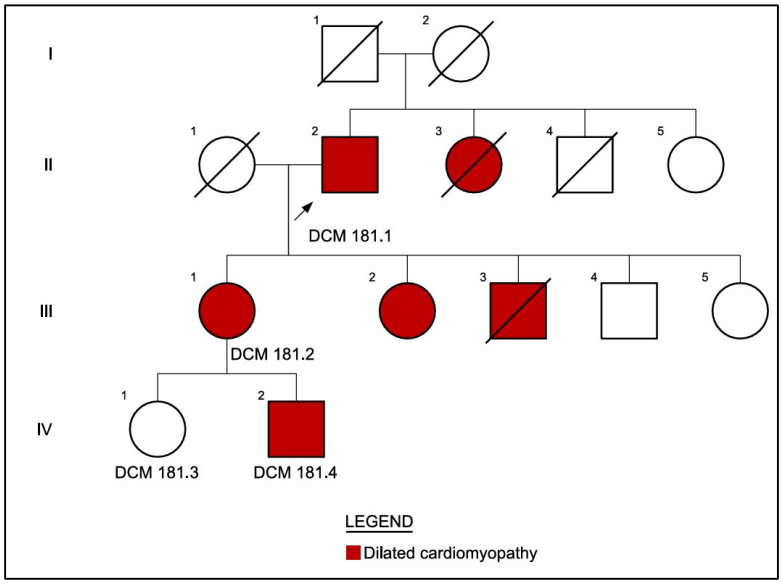


Figure 103: Pedigree of family DCM 181. The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals.

4.3.2.2 Genetics: BAG3 *c.925C>T screening*

The four enrolled family members of DCM 181 were consented and genetically screened for the *BAG3* c.925C>T variant; no other DNA samples were available for testing.

All primer information was recorded in Chapter 2, Section 2.5, Table 4. The proband (DCM 181.1), his daughter (DCM 181.8) and grandson (DCM181.4) were all found to be heterozygous for the *BAG3* c.925C>T variant, while his granddaughter (DCM 181.3) was genotype negative (Figure 104).

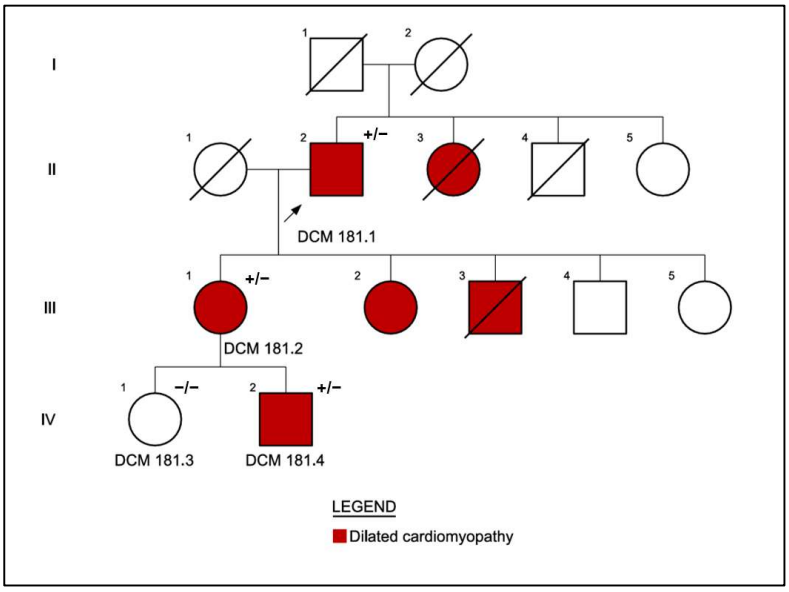


Figure 104: the genotyping results of DCM 181. The arrow indicates the proband, '+' indicates a positive genotype of the BAG3 variant while '−' indicates a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms

4.3.2.3 Haplotype Analysis

The genetics results were followed up by haplotyping. We prospectively recruited and enrolled available at-risk members of the DCM 181 family that tested positive for the *BAG3* c.925C>T variant and included them for haplotype screening. We made use of a combination of the *BAG3* microsatellites and the *BAG3* c.925C>T variant to produce the diseased haplotype. DCM 181.1's microsatellite marker peaks are shown in Figure 105 below.

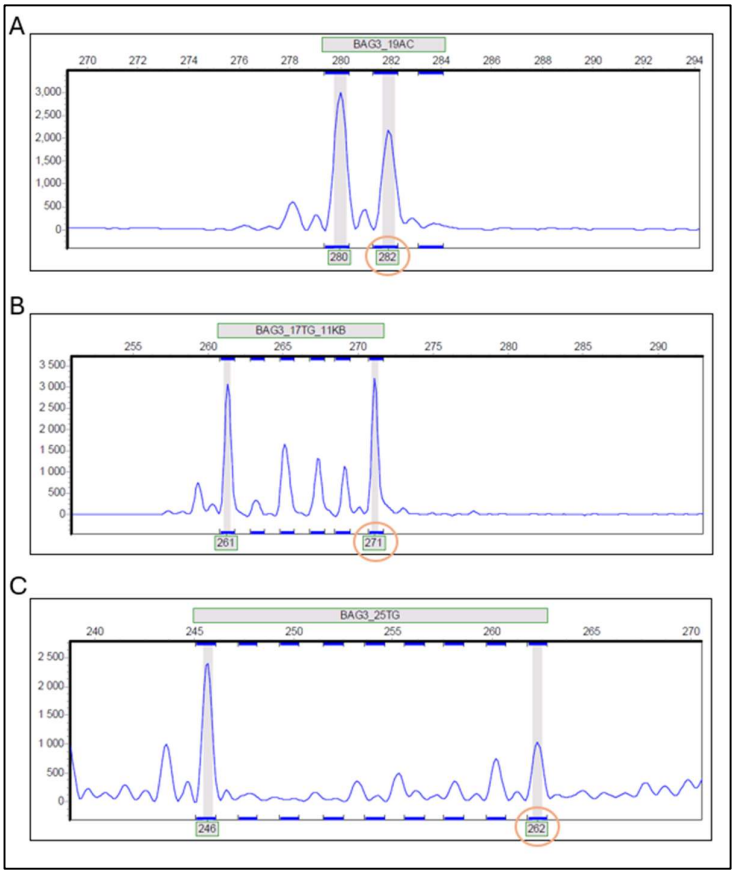


Figure 105: Microsatellite analysis of DCM 181.1.

(A) BAG3_19AC (282), the peak circled in brown, 282, is the common haplotype peak for BAG3_19AC. (B) BAG3_17TG_11KB (271), the peak circled in brown, 271, is the common haplotype peak for BAG3_17TG_11KB. (C) BAG3_25TG (262), the peak circled in brown, 262, is the common haplotype peak for BAG3_25TG

The proband, DCM 181.1, was shown to have the disease associated haplotype, represented by the brown haploblock (Figure 106). This haploblock matched the haploblock identified in family DCM 343 above. DCM 181.2 and DCM 181.4 both display the same disease associated haploblock.

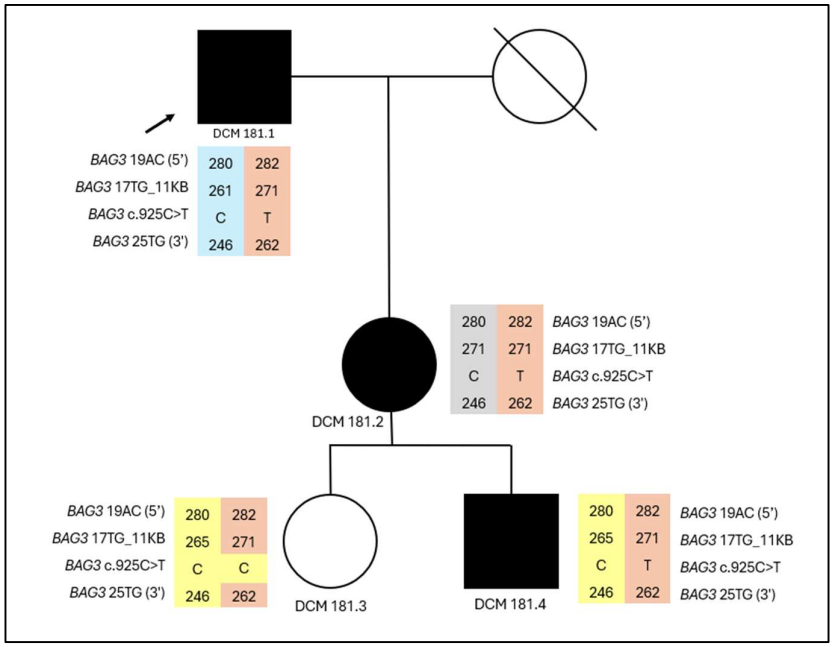


Figure 106: Haplotype analysis of family DCM 181. Circles indicate self-identified females, squares indicate self- identified males, black filled symbols represent clinically affected individuals, symbols with periods in the centre represent individuals with the BAG3 variant but are not clinically affected, shapes with an oblique line through indicates deceased individuals, haplotypes with dotted boarders represent assumed haplotypes from segregation analysis

4.3.2.4 Genealogy Tracing

To trace this family's genealogy, a substantial amount of information had to be gathered and organised, including copies of birth, marriage, and death certificates; records in churches and graveyards; old articles, letters, photographs, and documents related to citizenship and naturalisation; searches of historical records, including land records and censuses; a backwards and forwards approach, starting with the proband.

The proband and his enrolled family all self-identified as Mixed-Ancestry, which we anticipated would be challenging due to the previous issues we encountered in tracing the lineage of the ACM 71 Mixed-Ancestry family in Chapter 3.

There was limited information provided for genealogical tracing for this family, the information provided by the genetic councillor included some names, estimated birth dates and residences of some members of this family. This family resided in the Western Cape and is of Mixed-Ancestry.

The genotype information showed that the proband, DCM 181.1, was the eldest member of this family to have the *BAG3* variant. Therefore, both the paternal and maternal lines of DCM

181.1 were extended. This family required alternative methods of research as they were of Mixed-Ancestry and using standard methods of genealogical tracing failed to provide results. The most difficult period to find information on was the Apartheid era as the South African government had separated all records based on racial grouping and some individuals of some racial groupings often had incomplete or poorly recorded records. The Cape Town Archive did have some records of Mixed-Ancestry individuals however, these records only contained the surnames and no first names.

Unfortunately, we found ourselves in a similar situation as that in Chapter 3 (ACM 71 family), where genealogical tracing could not be extended any further than the early 20th century.

4.3.3 DCM 307

4.3.3.1 Clinical History

We enrolled and clinically screened three members of the DCM 307 family which included the proband (DCM 307.1) and her two daughters (DCM 307.3 and 307.4); all self-identified as Mixed-Ancestry.

The proband (DCM 307.1) was originally diagnosed with peripartum cardiomyopathy (PPCM) with an acute dilated phenotype as well as severe mitral valve regurgitation, clinical notes also described her experiencing trouble breathing and hot flushes during her 2nd pregnancy. She was referred for a cardiac transplantation, however, she stabilised on medical therapy and was delisted for transplantation. Her eldest daughter (DCM 307.3) was diagnosed with DCM at the age of 19 with her symptoms mirroring those of her mother; this included breathing problems and swollen legs. The proband's youngest daughter (DCM 307.4) was examined and showed no clinical evidence of DCM. Investigations into the family history by the counsellor indicated the family had a history of cardiac related problems. The proband's (DCM 307.1) father (II-2) passed away from a heart attack at the age of 59 and his brother (II-1) also had a history of cardiac related problems and had a heart transplant at Groote Schuur Hospital (GHS). The proband's grandmother (I-2) was suspected of having DCM as she passed away shortly after the birth of her second son due to cardiac related problems. The probands mother (II-3) suffered from hypertension and angina pectoris and at the age of 64 passed away from a heart attack. Hypertension is a prevalent heart

condition on the maternal line of the proband's family. Family records indicate that one of the proband's aunts and an uncle were also affected by hypertension (Figure 107).

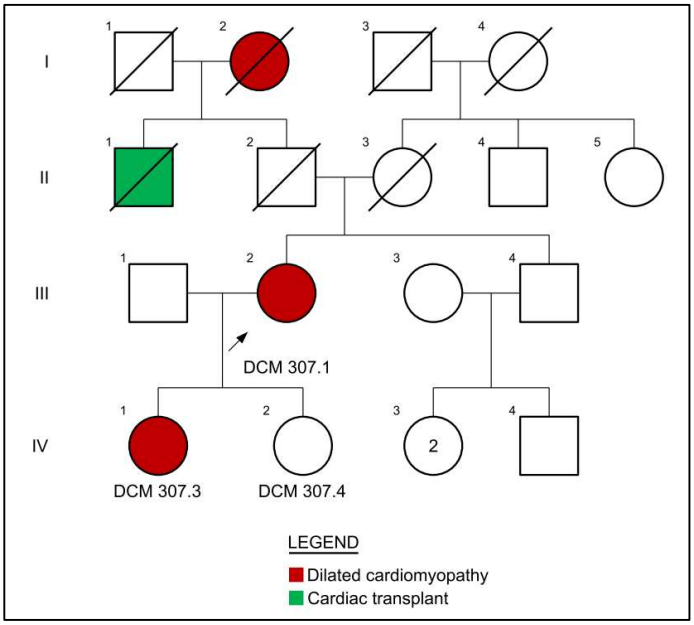


Figure 107: Pedigree of family DCM 307. The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals.

4.3.3.2 Genetics: BAG3 *c.925C>T screening*

The three enrolled family members of DCM 307 were genetically screened for the *BAG3* c.925C>T variant, all primer information was recorded in Chapter 2, Section 2.5, Table 4. The proband (DCM 307.1) and her eldest daughter (DCM 307.3) were both found to be heterozygous for the *BAG3* variant, while the probands youngest daughter (DCM 307.4) was genotype negative (Figure 108).

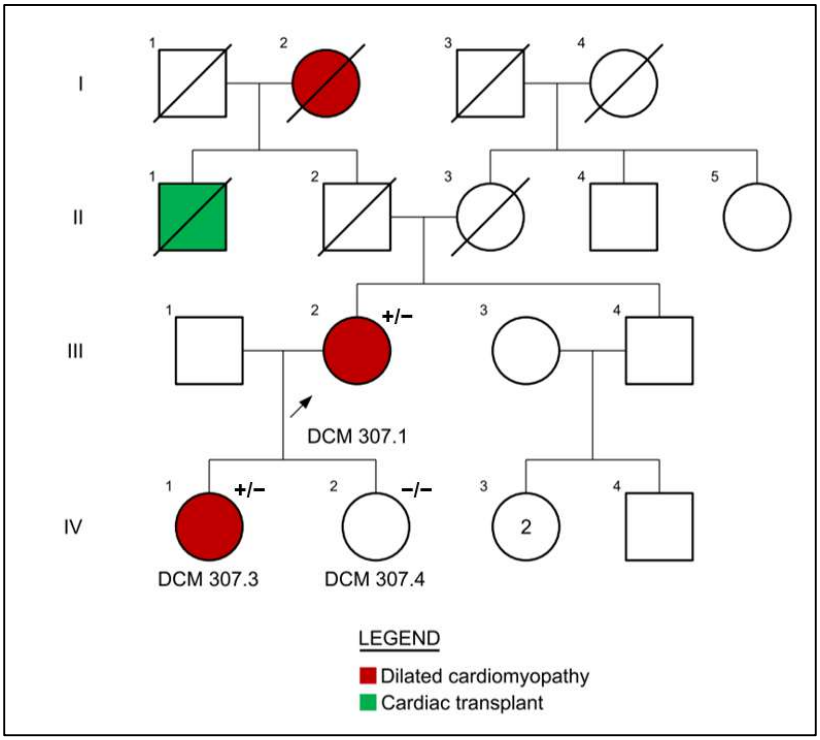


Figure 108: The genotyping results of DCM 181. The arrow indicates the proband, '+' indicates a positive genotype of the BAG3 variant while '−' indicates a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms

4.3.3.3 Haplotype Analysis

The genetics results were followed up by haplotyping. We prospectively recruited and enrolled available at-risk members of the DCM 307 family that tested positive for the *BAG3* c.925C>T variant and included them for haplotype screening. Haplotyping was performed for all the enrolled family members by making use of the microsatellites and the variant of interest (*BAG3* c.925C>T). DCM 307.1's microsatellite marker peaks are shown in Figure 109 below.

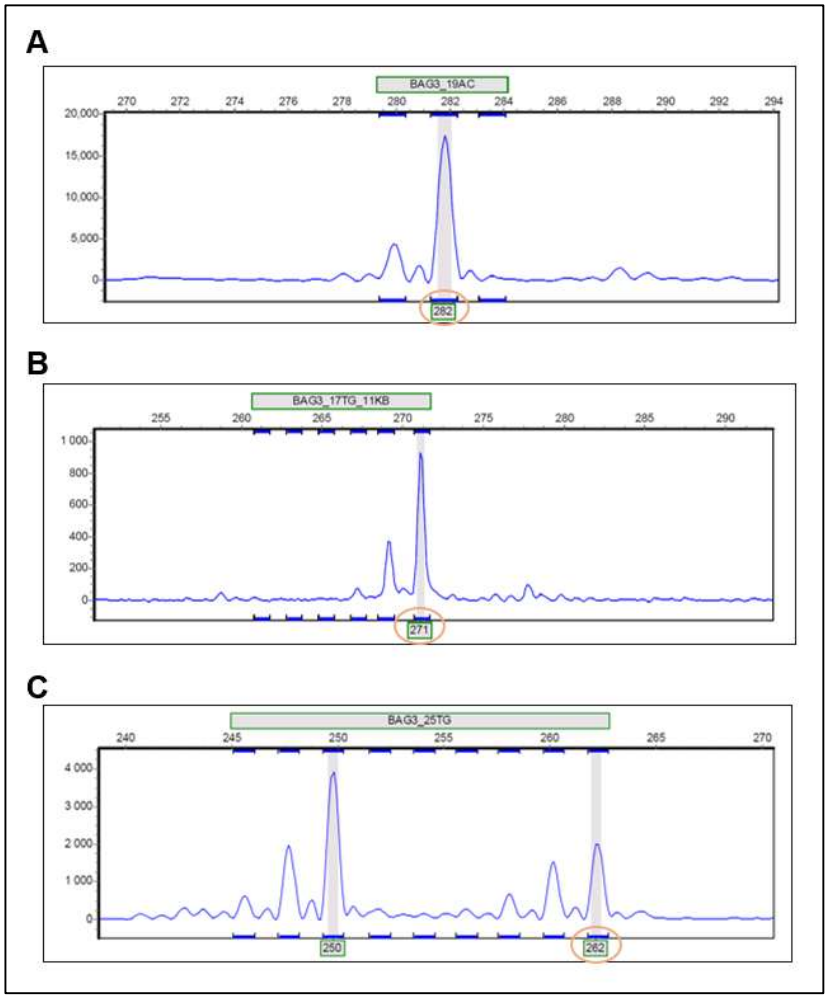


Figure 109: Microsatellite analysis of DCM 307.1. (A) BAG3_19AC (282), the peak circled in brown, 282, is the common haplotype peak for BAG3_19AC. (B) BAG3_17TG_11KB (271), the peak circled in brown, 271, is the common haplotype peak for BAG3_17TG_11KB. (C) BAG3_25TG (262), the peak circled in brown, 262, is the common haplotype peak for BAG3_25TG

The proband, DCM 307.1 is shown in Figure 110 to have the disease associated haplotype, represented by the brown haploblock. This haploblock matches the haploblock identified in family DCM 343 and DCM 181 above. DCM 307.3 has the same disease associated haploblock while DCM 307.4 was negative for the *BAG3* variant and we were able to infer her haplotype as we had the DNA of sister and mother.

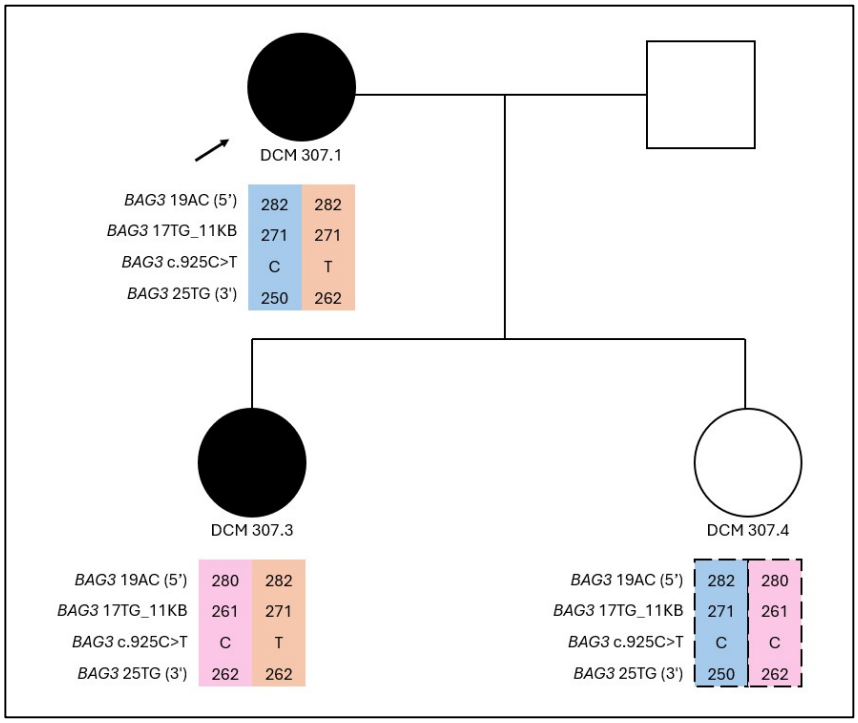


Figure 110: Haplotype analysis of family DCM 307. Circles indicate self-identified females, squares indicate self-identified males, black-filled symbols represent clinically affected individuals, symbols with periods in the centre represent individuals with the BAG3 variant but are not clinically affected, shapes with an oblique line through indicates deceased individuals,haplotypes with dotted borders represent assumed haplotypes from segregation analysis

4.3.3.4 Genealogy Tracing

To trace this family's genealogy, a substantial amount of information had to be gathered and organised, including copies of birth, marriage, and death certificates; records in churches and graveyards; old articles, letters, photographs, and documents related to citizenship and naturalisation; searches of historical records, including land records and censuses; a backwards and forwards approach, starting with the proband.

The proband and his enrolled family all self-identified as Mixed-Ancestry, which we anticipated would be challenging due to the previous issues we encountered in tracing the lineage of the ACM 71 Mixed-Ancestry family in Chapter 3.

There was limited information provided for genealogical tracing of this family, the information provided by the genetic councillor included names, estimated birth dates and residences of some members of this family. This family resided in the Western Cape.

The genotype information showed that the proband, DCM 307.1, was the eldest member of this family to have the *BAG3* variant. Therefore, both the paternal and maternal lines of DCM 307.1 were extended. This family required alternative methods of research as they are of Mixed-Ancestry and using standard methods of genealogical tracing failed to provide results. The most difficult period to find information on was the Apartheid era as the South African government had separated all records based on racial grouping and some individuals of some racial groupings often had incomplete or poorly recorded records. The Cape Town Archive did have some records of Mixed-Ancestry individuals however, these records only contained the surnames and no first names.

We encountered a recurrence of the limitations observed in DCM 181 and in Chapter 3 with the ACM 71 lineage, where genealogical tracing was constrained to the early 20th century. The three families are represented in table 4.1 below, which summarises the number of family members screened in each family, the number of participants positive for the *BAG3* variant and of those positive, if they have a common haplotype. This would allow us to compare the progenitors of each family and narrow it down to possible progenitor couples (Table 35 below).

Table 35: Study Participants screened for BAG3 c.925C>T variant and common Haplotype

Family ID	No. of study participants screened	*BAG3* c.925C>T variant Positive	Common Haplotype
DCM 343	12	8	8
DCM 181	4	3	3
DCM 307	3	2	2
Total Participants	19	13	13

4.4 Discussion

Through the IMHOTEP registry of South Africa, three probands presenting with clinical symptoms of DCM and through previous unpublished research were found to carry the *BAG3* c.925C>T variant. The clinical phenotypes and family histories revealed each family had multiple generations affected by DCM and possibility that they all shared a common progenitor couple, hundreds of years ago, was hypothesised. This study used a combination of methods and resources such as clinical history, haplotype analysis and genealogy tracing to show that these families.

4.4.1 Genotype-phenotype

In this chapter, we presented three probands (DCM 343.1, DCM 181.1 and DCM 307.1) along with an additional 10 family members carrying the *BAG3* c.925C>T variant, which is known to cause DCM. All 13 genotype-positive participants were heterozygous for the variant, and we observed varying phenotypic expressions both within and between the three families.

Upon further investigation, the varying phenotypes and disease severity among participants with the *BAG3* variant were identified to be the main drivers that may have influenced this variation, possibly triggered by viral infections or underlaying comorbidities.

The DCM 343 family had extensive clinical records spanning six generations. Although there was significant variation in phenotype within this family, a common pattern emerged, multiple family members died from SCD. It was suspected that most of the individuals from the proband's paternal lineage, where the variant was inherited, developed DCM that may have been triggered by viral infections. The proband (DCM 343.1), was diagnosed with DCM following a viral infection, that led to HF. Her daughter (DCM 343.5) presented with HF and cardiogenic shock after a viral gastroenteritis and the probands grandfather was noted as having the coxsackie viral infection prior to his DCM diagnosis. These three members of DCM 343 all experienced severe DCM following the viral infections. *BAG3,* as previously discussed, plays a crucial role in the structure and function of cardiomyocytes and viral infections can act as a trigger for disease onset or worsen pre-existing DCM, which results in a rapid decline of heart function (193). This mechanism likely explains the severe presentations observed in these participants.

The presence of underlaying comorbidities, such as hypertension, was also previously associated with disease severity (203) and it was found that the treatment of hypertension with ACE inhibitors or ARBs (Angiotensin II receptor blockers) in the French-Canadian population, provided a protective effect against disease progression in participants with the *BAG3* variant. The proband's father had been diagnosed with hypertension and was on antihypertensive medication (ACE inhibitors) for years before to his genetic status was known. Despite carrying the pathogenic variant, he only exhibited mild DCM symptoms and remains stable. Once placed on treatment, the proband (DCM 343.1) showed evidence of complete cardiac remodelling and her sister, who was started on treatment after her diagnosis,

achieved full cardiac recovery. Similar improvements were observed in the probands DCM 181.1 and DCM 307.1, both probands showed significant cardiac function improvement after the initiation of medical therapy, with DCM 307.1 even being removed from the cardiac transplantation list. These finding align with those of Chami et al., where the proband's father, despite carrying the same *BAG3* variant, remained asymptomatic while on antihypertensive medication (109). Research has shown that antihypertensive medication such as ACE inhibitors and ARBs can slow disease progression and improve HD symptoms in patients at risk of developing DCM (204).

Despite other members of the DCM 343, DCM 181 and DCM 307 families being heterozygous for the *BAG3* variant, many remain asymptomatic. This could be attributed to the absence of a viral trigger or untreated comorbidities that could have increase stress on the heart. Additionally, younger patients, such as DCM 343.12, may not yet exhibit symptoms, as DCM can have an age-dependant penetrance (205). Therefore, continuous monitoring and annual clinical screenings remain essential for early detection, disease prevention and management.

The occurrence of the *BAG3* variant, across multiple generations in three unrelated families from the Western Cape suggests a common ancestral origin. This led us to hypothesise that this variant may have been introduced to South Africa by a single founder several centuries ago. To explore this possibility, we attempted to construct haplotypes and investigated the genealogical histories of these families to determine whether they share a common founder.

4.4.2 Haplotype analysis

Haplotypes were constructed to span the entire *BAG3* gene by selecting microsatellite markers that flanked the 5’ (BAG3_19AC), 3’ (BAG3_25TG) and within the gene (BAG3_17TG_11kb) and were combined with the *BAG3* c.925C>T variant to form a haploblock. All positively screened participants for the *BAG3* c.925C>T variant exhibited a shared haplotype; therefore, 13 participants had the same variant and an identical haplotype. This finding supports the hypothesis that the variant is a founder variant inherited from a common ancestor, as the haplotype was identical by descent (IBD). Genealogical research was then used to try identify if a common progenitor was responsible for introducing the pathogenic *BAG3* c.925C>T variant into the South African population.

4.4.3 Genealogy

Of the three enrolled families, we were able to partially trace the genealogy of one family (DCM 343) carrying the *BAG3* c.925C>T variant. This family exhibited ancestral origins from English, Dutch, French and French-Canadian lineages. The most noteworthy was the ancestor of French-Canadian descent, as this variant was first identified as a founder variant in a French-Canadian population of Quebec (109). Several parallels were observed between the South Africa DCM 343 family and the French-Canadian family carrying the *BAG3* variant, most notably the use of antihypertensive medication, which was found in two patients from the separate studies and appeared to slow the progression if DCM.

Much like the South African Afrikaner population, the French-Canadian population has been found to harbour several founder variants due to migration patterns. The migration of the French Calvinists to Canada during the French regime during the 16th and 17th centuries (206) led to the establishment of certain genetic traits within this population. Research shows that approximately 14 rare autosomal dominant and autosomal recessive diseases occur at higher frequencies among French-Canadians (117, 207-219). This observation led to the hypothesis that the *BAG3* variant may have been introduced to both South Africa and Canada through French Huguenot migration.

In the genealogy of DCM 343, an ancestor who lived in Canada five generations ago was identified, and their descendants subsequently immigrated to South Africa. It is possible that this ancestor was the source of this variant in South Africa, though an earlier introduction via the French Huguenots in the 17th century cannot be ruled out. Interestingly, this variant was first reported in 2011 in two related European patients by researchers in France (220). The researchers who identified the variant in the French-Canadian families could not exclude the possibility that the European family and French-Canadian families shared a common ancestor (109). Therefore, the possibility remains that European, French-Canadian, and South African families carrying the *BAG3* variant may share a common lineage. Further collaboration with researchers who identified the variant, continued recruitment of new patients potentially harbouring the *BAG3* variant, and analysis of additional at-risk family members could provide evidence supporting a common founder.

The prevalence of the *BAG3* variant in South Africa could be significantly higher that currently recognised, particularly among individuals of Mixed-Ancestry, a population

frequently faces limited access to adequate healthcare services (221). The other two enrolled families (DCM 181 and DCM 307) were of Mixed-Ancestry, but tracing their genealogy proved significantly more challenging. One of the greatest difficulties was the historical disruption in record-keeping during the Apartheid era, when the South African government segregated all records by racial classification. Many records for individuals of certain racial groupings were incomplete or poorly maintained. The Cape Town Archive contained some records of mixed-ancestry individuals, but these typically listed only surnames without accompanying first names, making it difficult to establish family connections.

Beyond historical record limitations, tracing Mixed-Ancestry families was further complicated by variations in the spelling of names and surnames, repetition of names within the same family tree, children born out of wedlock, and cultural practices in which close family friends informally raised children who were considered "cousins" despite lacking genetic relation. Additionally, reluctance to share information and a lack of knowledge about family history further hindered genealogical tracing. Consulting specialists in this field, such as Mr. Aubrey Springveldt, a researcher and specialist genealogist who helps Cape families accurately trace their family histories, may allow for further tracing, but this approach would likely require direct interviews with family members, which may not be well received.

It is important to note that identifying the founder of a variant typically requires genealogical data from approximately 10-12 families (104). Given that only three families have been enrolled in this study, the current dataset is insufficient to definitively trace the *BAG3* variant to a common founder. Future efforts should focus on expanding recruitment and collaboration with specialists in genetic genealogy to improve tracing accuracy and establish potential ancestral links.

4.5 Conclusion

Despite the limitations faced in genealogically tracing, the shared haplotype, location of the families and the knowledge that one family is Caucasian, and two are of Mixed-Ancestry provide strong evidence that this variant is a possible founder variant however, additional families with this variant is required to definitively answer this question.

The continued recruitment, screening and investigation into this founder *BAG3* c.925C>T variant and all other *BAG3* variants in a South African context is crucial to the diagnosis, treatment and survival of patients harbouring this highly pathogenic variant.

Chapter 5: *LMNA* c.568 C>T (p.Arg190Trp) in South African families with DCM

5.1 Introduction

As previously described (Chapter 1 and Chapter 4), DCM has the most genetic heterogeneity of all the cardiomyopathies, with over 60 genes implicated (26), including *LMNA*, which accounts for 5.5% of DCM cases (28).

The *LMNA* gene consists of 12 exons coding for Lamin A/C proteins, which are essential structural components of the nuclear envelope (NE) (222). They aid in maintaining the nuclear structure and play an important role in DNA replication, Chromatin organisation and gene regulation (223-225). Pathogenic variants of *LMNA* have been shown to lead to severe cardiac remodelling which results in heart failure, additionally, *LMNA* variants have been linked to mechanotransduction disruption by altering nuclear integrity and therefore changing the response to extracellular matrix signals (226, 227). Variants in the *LMNA* are also known to have a pleiotropic effect (228) and have been associated with several disease phenotypes such as Hutchison-Gilford Progeria Syndrome (229), atypical progeroid syndromes (APS) (230), striatal muscle diseases (231), partial lipodystrophy (231) as well as dilated cardiomyopathy (232) and are often referred to as laminopathies.

Of particular interest was a founder variant, *LMNA*c.1961dup (p.Thr655fs), that was identified in the French population (120) and which caused familial partial lipodystrophy type 2 (FPLD2) and severe manifestations of DCM (233).

Through the parent IMHOTEP study, three probands were identified with the *LMNA* c.568C>T (p.Arg190Trp)variant in intermediate filament rod domain of the lamin A/C proteins (Figure 111, adapted from (234)).

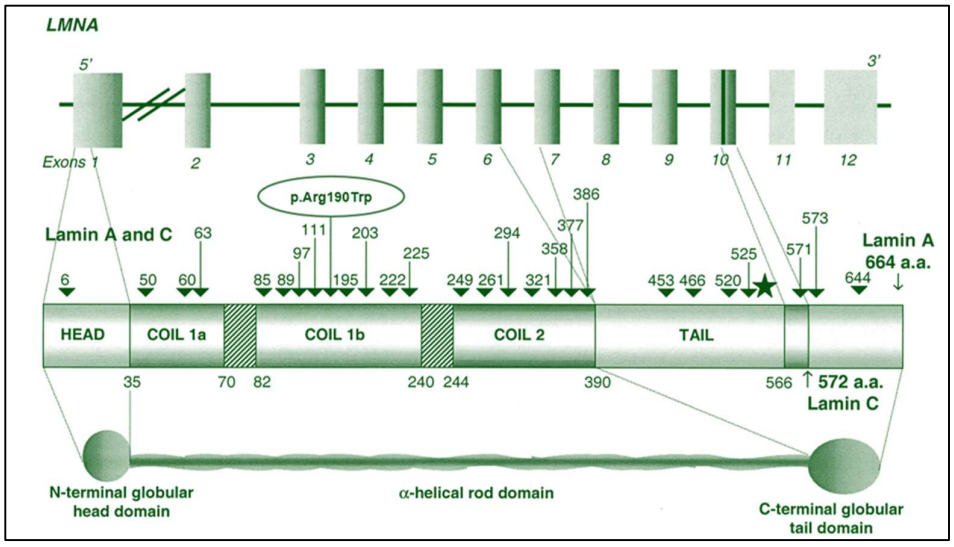


Figure 111: Schematic representation of LMNA gene and Lamin A/C protein. The location of the LMNA p.Arg190Trp variant is circled in red, shown in the context of the protein structure it affects (coil 1b). The arrows along the protein structures indicate the variants causing DCM, their location and the aa position is indicated by the corresponding number (234)

A computational prediction suggested that the *LMNA* c.568C>T variant may have deleterious impact on protein structure and function of Lamin A and C (internally defined REVEL score threshold >= 0.7,(235)). Thereafter, functional studies were conducted which suggest this variant may disrupt the nuclear envelope by affecting viscoelasticity of lamin A (236, 237). It was classified as pathogenic by the ACMG guidelines and was found to have a minor allele frequency (MAF) of 0.00001 according to gnomAD. Through prospective recruitment and screening of family members of the three probands, we identified additional relatives who also carry this variant.

The aim of this chapter was to identify if the three probands and their affected family members carried the LMNA c.568C>T variant and if these positive individuals shared a common haplotype and therefore a common progenitor.

5.2 Methods

All methods were outlined in Chapter 2, the flow diagram below represents the overall flow of chapters 3-6, which has been tailored to this chapter and refers to the relevant methods (Figure 112).

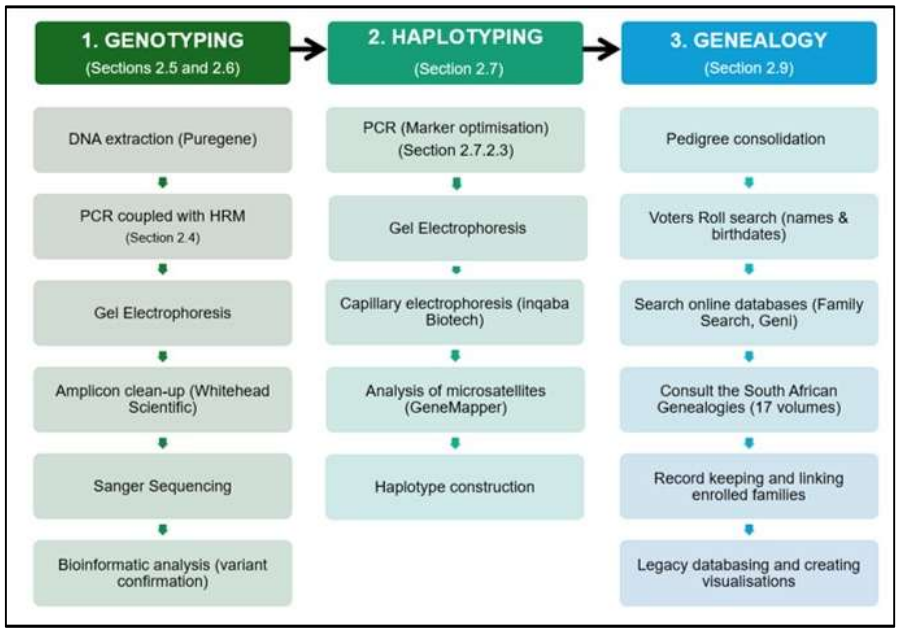


Figure 112: Workflow, adapted from Chapter 2, Section 2.3, to highlight the relevant sections detailing the methods associated with this chapter, specifically focusing on the LMNA c.568C>T variant

5.2.1 Clinical

Probands with suspected DCM, were referred to the Cardiac Clinic at Groote Schuur Hospital for screening and was then enrolled into the IMHOPTEP registry. The inclusion and exclusion criteria can be found in Chapter 2, section 2.2.

5.2.2 Genetics

Refer to the appropriate section in Chapter 2, section 2.4, Table 5. The probands were screened using Haloplex targeted resequencing and family members were screened using Sanger sequencing to determine if they were positive or negative for the *LMNA* c.568C>T variant. All family members that consented had their blood drawn, DNA extracted and genotyped.

5.2.3 Haplotype

A haplotype is a set of DNA variations, or alleles, inherited together on the same chromosome. These alleles are physically linked and tend to be passed down as a unit. Think of it as a specific combination of genetic variants on a single chromosome. Haplotypes are

essential for understanding genetic inheritance, mapping disease genes, and studying population history (180).

Inferring haplotypes within a family becomes easier when you have DNA from parents or siblings. Parental DNA allows you to directly observe the haplotypes transmitted to their offspring. If you know the parents' genotypes, you can deduce the child's haplotypes with greater certainty (181). Similarly, sibling DNA can help resolve ambiguous haplotypes. By comparing the genetic variations shared between siblings, you can infer which segments of their chromosomes were inherited from each parent. This information is particularly useful in linkage studies, where researchers track the inheritance of specific haplotypes within families to identify genes associated with diseases (181, 182).

We made use of three microsatellite markers, LMNA_GT21 at the 5' END LMNA_AC22 at the 3' end AND LMNA_CT16 within the *LMNA* gene, thus spanning the entire gene. These markers were used in combination with the *LMNA* c.568C>T variant to construct haplotypes for the probands and their extended families. Haplotype analysis enabled the inclusion or exclusion of ancestral lines during genealogical tracing based on the known autosomal dominant inheritance pattern

5.2.4 Genealogy

Genealogy tracing followed standard procedures (Figure 14, Chapter 2). The oldest enrolled family member was the starting point of genealogy tracing. The ancestral lineages were extended as far as records allowed.

5.3 Results

Between 2012 and 2024, patients with Dilated Cardiomyopathy (DCM) symptoms were referred to the Cardiac Clinic at Groote Schuur Hospital. After providing informed consent, they were enrolled in the IMHOPTEP registry. All three probands and their families self-identified as Mixed-Ancestry.

Genetic screening was conducted on the three probands, and at-risk family members were invited to join the genetics sub-study. A total of six participants were identified for analysis.

Haplotypes for the three families were constructed using a combination of three microsatellite markers (LMNA_GT21, LMNA_AC22, LMNA_CT16) and the BAG3 c.925C>T variant, covering the entire LMNA gene and its 5' and 3' regions. This analysis helped to identify or exclude ancestral lines based on the autosomal dominant inheritance pattern.

Genealogy tracing followed established procedures, starting with the oldest enrolled family member. The ancestral lineages were extended based on available records, and families were presented in order from most to least genealogically traced (DCM 809, DCM 781, DCM 569).

5.3.1. DCM 809

5.3.1.1 Clinical History

The proband, DCM 809.1, was first diagnosed with DCM at 33 years old. Also noted was that the proband was previously treated for renal failure and had been admitted into a rehabilitation centre for alcohol abuse.; his nephew, DCM 615.2, was later diagnosed with a DCM/LVCN overlap at the age of 35 years. There appeared to be a significant family history of HF in this family. All the probands siblings (III-2, III-3, III-4), suffered from heart failure. His brother (III-2) had an ICD implanted but still suffered from heart failure. Additionally, he was also diagnosed with brain cancer 42 years old. The probands sister (III-3) had a pacemaker implanted however, she died at 43 years old from HF. Lastly, the youngest sister (III-4) died in her 30s from tuberculosis (TB) related lung failure. However, clinical notes also show she was listed for a heart transplant and may have been in HF.

The cardiac-related matters appear to have originated from the probands maternal lineage as the probands mother (II-2) had died from HF in her 50s. Additionally, she had three siblings who also suffered from HF (II-4, II-5, and II-6), two of whom received heart transplants (II-4 and II-5) (Figure 113).

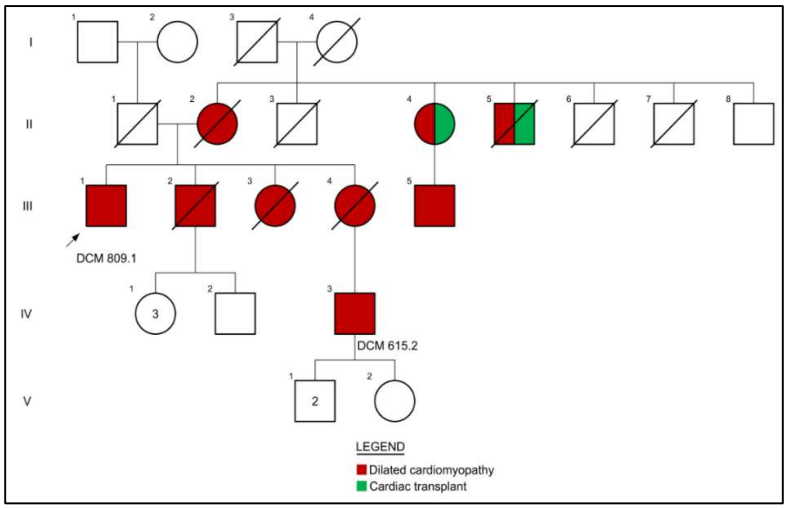


Figure 113: Pedigree of family DCM 809. The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals.

5.3.1.2 Genetics: LMNA c.568C>T Screening

The proband, DCM 809.1, and his nephew, DCM 615.2 were genetically screened for the *LMNA* c.568C>T variant and both were found to be heterozygous (Figure 114). All primer information was recorded in Chapter 2, Section 2.5, Table 5. No other family members were available for screening.

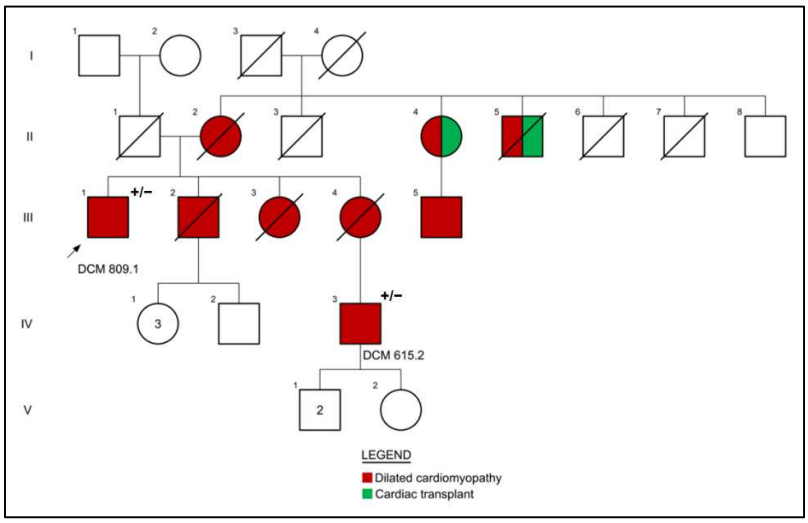


Figure 114: the genotyping results of DCM 809. The arrow indicates the proband, '+' indicates a positive genotype of the LMNA variant while '–' indicates a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms.

5.3.1.3 Haplotype Analysis

The haplotyping followed the genetics results. The DCM 809 continued prospectively and at-risk family members were recruited into the current study. New family members testing positive for the *LMNA* c.568C>T variant was included in the haplotype screening. Haplotyping was performed for all the enrolled family members by making use of a combination of microsatellites and the *LMNA* variant of interest (*LMNA* c.568C>T). DCM 809.1's microsatellite marker peaks are shown in Figure 115 below.

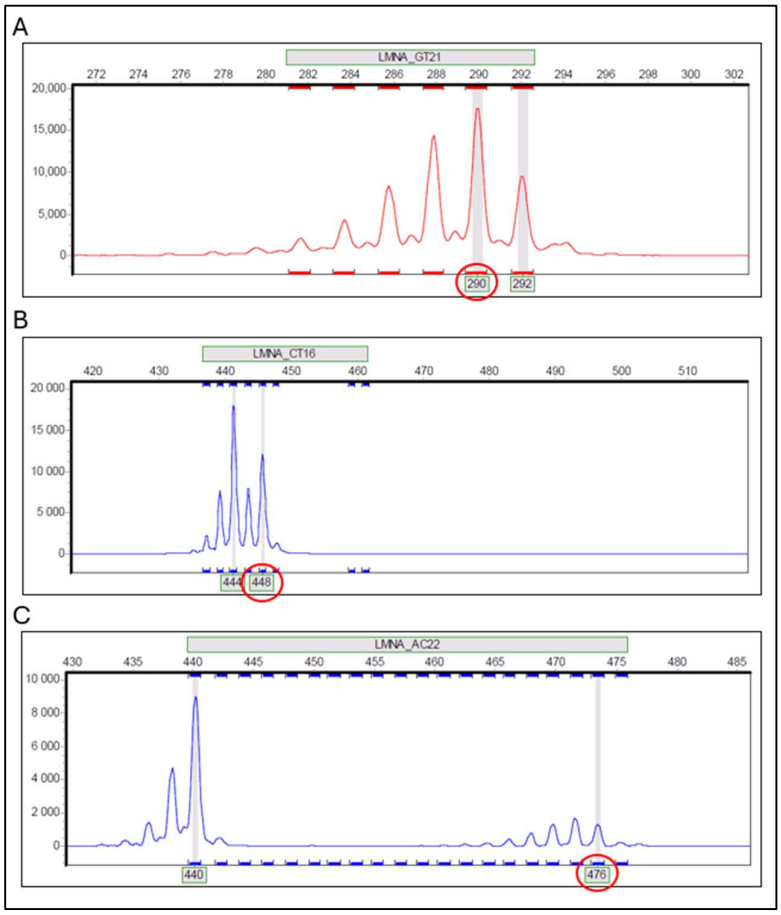


Figure 115: Microsatellite analysis of DCM 809.1. (A) LMNA_GT21 (282), the peak circled in red, 290, is the common haplotype peak for LMNA_GT21. (B) LMNA_CT16 (271), the peak circled in red, 448, is the common haplotype peak for LMNA_CT16. (C) LMNA_AC22 (476), the peak circled in red, 476, is the common haplotype peak for LMNA_AC22

The proband, DCM 809.1 was shown to have the disease associated haploblock, represented by the red haploblock in Figure 116 below. His nephew (DCM 615.2) also had the disease associated haploblock, represented by the red haploblock.

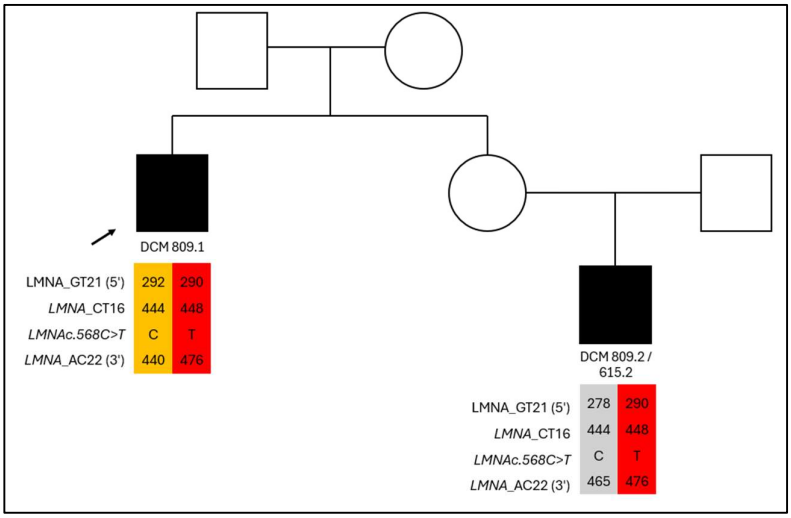


Figure 116: Haplotype analysis of family DCM 809. Circles indicate self-identified females, squares indicate self- identified males, black filled symbols represent clinically affected individuals, symbols with periods in the centre represent individuals with the LMNA variant but are not clinically affected, shapes with an oblique line through indicates deceased individuals, haplotypes with dotted boarders represent assumed haplotypes from segregation analysis.

5.3.1.4 Genealogical tracing

To trace this family's genealogy, a great deal of information had to be gathered and organised, including copies of birth, marriage, and death certificates; records in churches and graveyards; old articles, letters, photographs, and documents related to citizenship and naturalisation; searches of historical records, including land records and censuses; a backwards and forwards approach, starting with the proband.

There was limited information provided for genealogical tracing for this family, the information provided on the clinical notes, by the genetic councillor included names, estimated birth dates and residences of some members of this family. This family resided in the Western Cape and was of Mixed-Ancestry.

As the proband mother had died of HF, we extended the family from the maternal line. Despite the family's pedigree extending back through five generations, there was missing or wrong spelling of surnames, thus hindering genealogical tracing. Alternative methods of genealogical research were used as this family is of Mixed-Ancestry and using standard methods of genealogical tracing failed to provide results. During apartheid, the government separated all records based on racial grouping. Individuals of some racial groupings were often had incomplete or poorly recorded records. The Cape Town Archive possesses certain records pertaining to individuals of Mixed-Ancestry; however, these records are limited to

surnames without accompanying given names. Consequently, efforts to obtain detailed information on members of this family from that historical period yielded inconclusive results.

5.3.2 DCM 781

5.3.2.1 Clinical History

Three members of the DCM 781 family were clinically screened and included the proband (DCM 781.1), her proband's brother (DCM 781.2) and the proband's daughter (DCM 781.3). The proband first presented with DCM at 45 years old and eight years after the initial diagnosis, in 2023, she received a heart transplant. The proband's brother (DCM 781.2) was also found to be clinically affected by DCM, but her daughter DCM (781.3) was clinically unaffected.

Investigations into the clinical history of the family found a significant history of cardiac-related conditions in the proband's maternal lineage. The probands mother (III-4), died in her 60s due to suspected cardiac related problems. Two of the mother's seven siblings (III-5 and III-6) died due to sudden cardiac death, while two others were clinically diagnosed with dilated cardiomyopathy (DCM). Notably, III-5, who had a clinical diagnosis of DCM, died while undergoing a heart transplant. Clinical notes of the probands maternal grandfather (II-3) describe him experiencing "cardiac thrombosis" and subsequently passed away due to SCD. The proband's first cousins, once removed (III-1 and III-2) were also diagnosed with DCM, and both had subsequent heart transplants at GSH (Figure 117). This strengthened the hypothesis that the *LMNA* variant was inherited from the proband's maternal line.

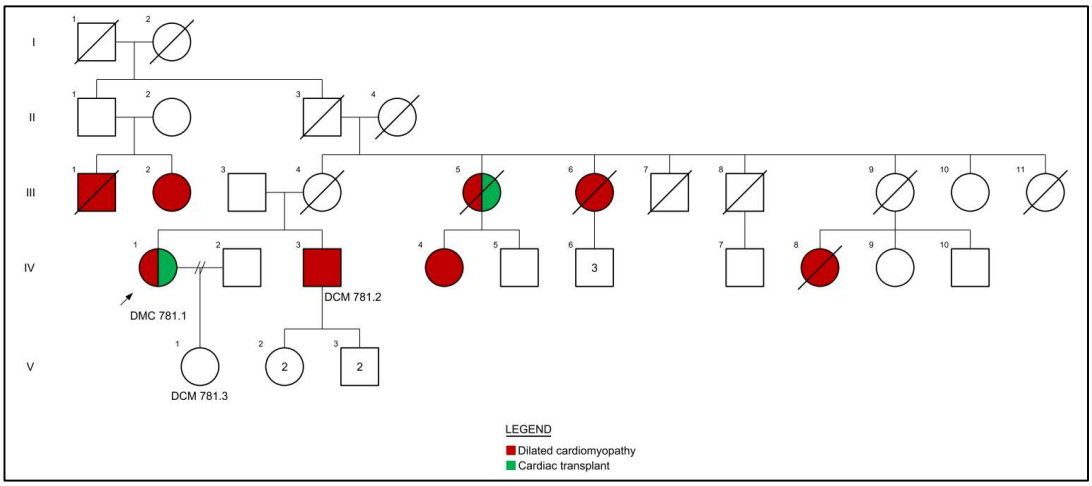


Figure 117: Pedigree of family DCM 781. The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals.

5.3.2.2 Genetics: LMNA c.568C>T screening

We consented and genotyped the proband (DCM 781.1), her brother DCM 781.2 and her daughter DCM 781.2. Both the proband and her brother were found to be heterozygous for the *LMNA* c.568C>T variant while the probands daughter was genotype negative (Figure 118).

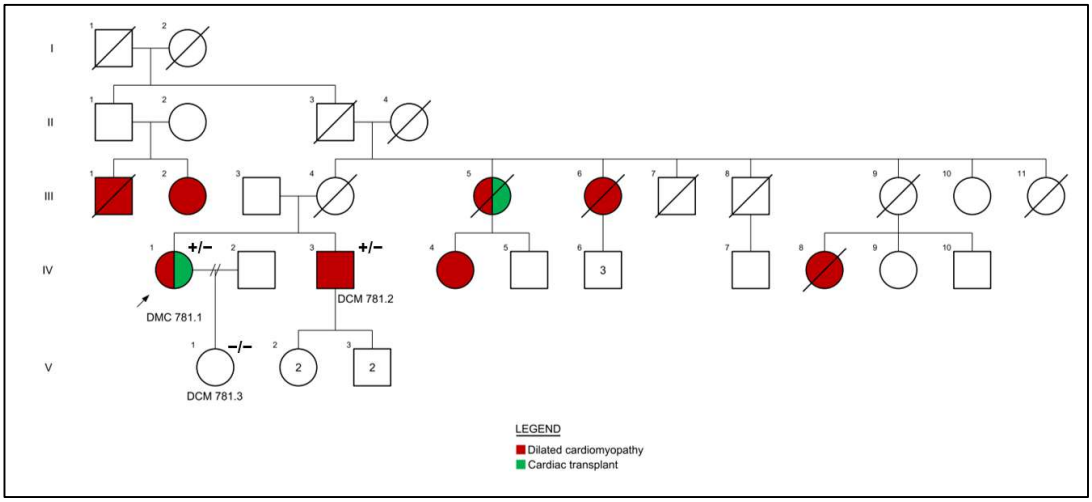


Figure 118: The genotyping results of DCM 781. The arrow indicates the proband, '+' indicates a positive genotype of the LMNA variant while '−' indicates a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms.

5.3.2.3 Haplotype Analysis

Haplotyping was performed for all the enrolled family members by making use of a combination of the LMNA microsatellites and the variant of interest (*LMNA* c.568C>T). DCM 809.1's microsatellite marker peaks are shown in Figure 119 below.

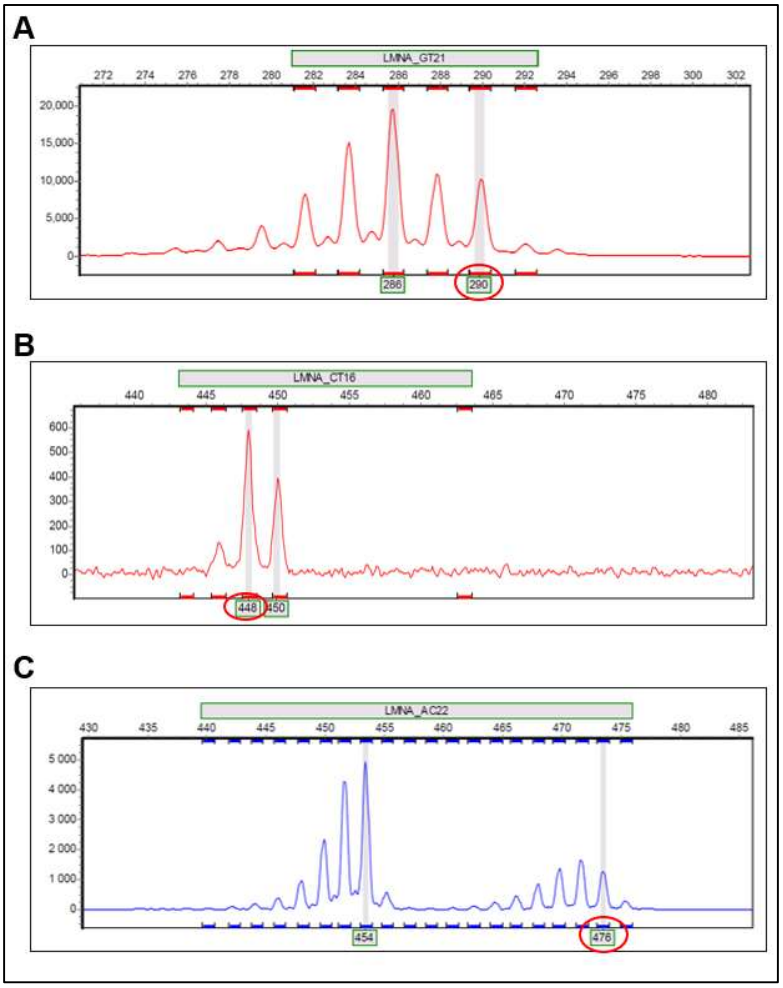


Figure 119: Microsatellite analysis of DCM 781.1. (A) LMNA_GT21 (282), the peak circled in red, 290, is the common haplotype peak for LMNA_GT21. (B) LMNA_CT16 (271), the peak circled in red, 448, is the common haplotype peak for LMNA_CT16. (C) LMNA_AC22 (476), the peak circled in red, 476, is the common haplotype peak for LMNA_AC22.

The proband, DCM 781.1, is show in Figure 120 to have the disease associated haploblock represented by the red haploblock (290,448, T, 476). This haploblock is identical to that found in the DCM 809 family. The proband's brother, DCM 781.2, also shared an identical haplotype and therefore had the disease associated haploblock. The probands daughter, DCM 781.3 did not have the disease associated haploblock as shown below.

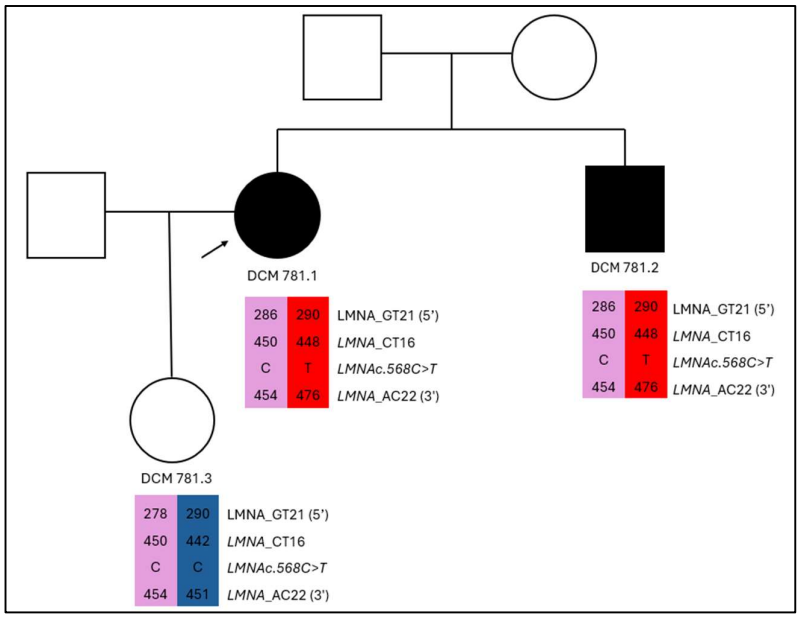


Figure 120: Pedigree of family DCM 781. Circles indicate self-identified females, squares indicate self-identified males, black filled symbols represent clinically affected individuals, symbols with periods in the centre represent individuals with the LMNA variant but are not clinically affected, shapes with an oblique line through indicates deceased individuals, haplotypes with dotted boarders represent assumed haplotypes from segregation analysis.

5.3.2.4 Genealogical Tracing

To trace this family's genealogy, a great deal of information had to be gathered and organised, including copies of birth, marriage, and death certificates; records in churches and graveyards; old articles, letters, photographs, and documents related to citizenship and naturalisation; searches of historical records, including land records and censuses; a backwards and forwards approach, starting with the proband.

There was limited information provided for genealogical tracing for this family. The information provided on clinical notes from the genetic councillor included names, estimated birth dates and residences of some members of this family. This family resided in the Western Cape and is of Mixed-Ancestry.

As the *LMNA* variant was maternally inherited, we extended both the paternal and maternal lines of DCM 781.1. Despite the family's pedigree extending back through five generations, there was missing or wrong spelling of surnames, thus impeding genealogical tracing. We used alternative methods of genealogical research, as this family is of Mixed-Ancestry. Unfortunately, during the apartheid era, records were separated by racial groupings by the government. Individuals of some racial groupings often had incomplete or poorly recorded records. Although the Cape Town Archive possesses certain records pertaining to individuals of Mixed-Ancestry, these records are limited to surnames without accompanying given names. Consequently, efforts to obtain detailed information on members of this family from that historical period yielded inconclusive results.

5.3.3 DCM 569

5.3.3.1 Clinical History

Two members were enrolled within this family. The proband (DCM 569.1), who was first diagnosed with DCM at the age of 38 and was subsequently referred to GHS for a heart transplant. Unfortunately, she passed away before receiving the transplant. The proband's second cousin (DCM 576.1), was enrolled much later and was diagnosed with DCM at the age of 41; he underwent a heart transplant at 47. Investigations into the clinical history of the family found a notable history of cardiac disease. The proband's mother (III-2) passed away in her 40s due to unspecified cardiac problems, while the proband's maternal grandmother (II-4) was diagnosed with unspecified cardiac disease in her early 30s and died at the age of 38. Additionally, both of the proband's sisters (IV-2 and IV-3) required heart transplants due to unspecified cardiac-related issues. The proband's second cousin, DCM 576.1 also had a family history of cardiac disease with his mother (III-6) diagnosed with unspecified heart problems in her early 50s; she died from a suspected heart attack in her 60s. DCM 576.1's maternal grandmother (II-6) was diagnosed with a cardiac condition in her 40s and died at the age of 47 due to unspecified heart problems (Figure 121). This strengthened the hypothesis that the *LMNA* variant was inherited through the proband's maternal line.

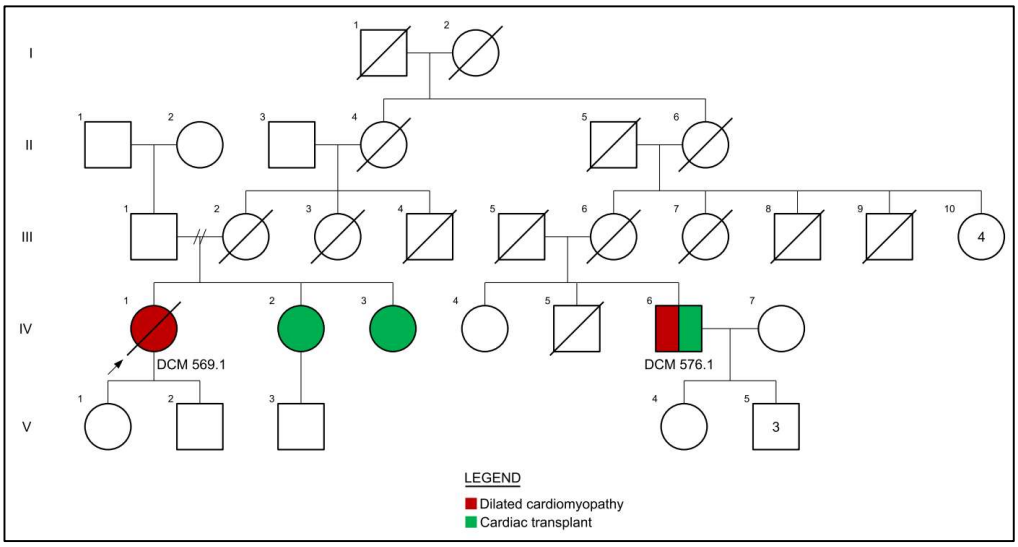


Figure 121: Pedigree of family DCM 569. The arrow indicates the proband, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms, shapes with an oblique line through indicates deceased individuals.

5.3.3.2 Genetics: LMNA c.568C>T screening

We consented and screened the two enrolled family members for DCM 569. The proband (DCM 569.1) and her second cousin (DCM 576.1) were found to be heterozygous for the *LMNA* c.568C>T variant (Figure 122). All primer information was recorded in Chapter 2, Section 2.5, Table 5.

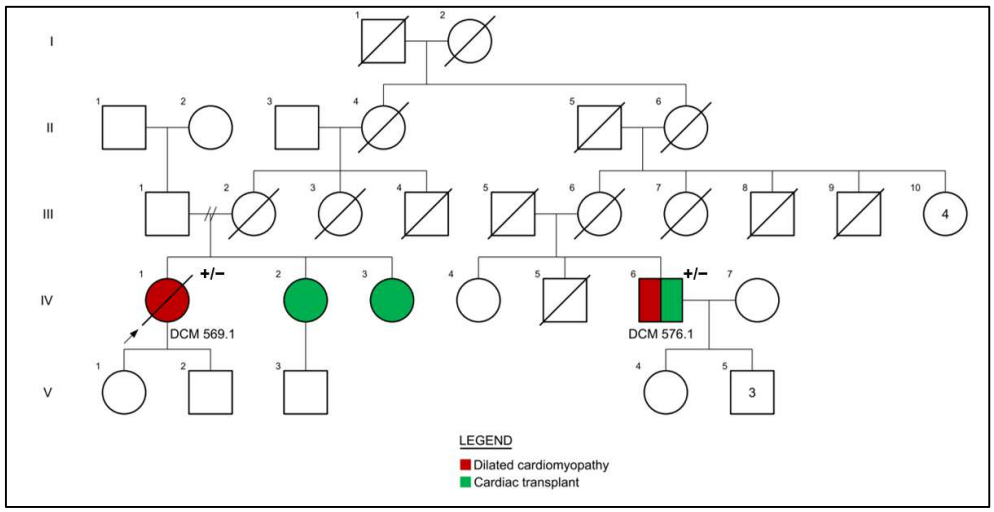


Figure 122: The genotyping results of DCM 781. The arrow indicates the proband, '+' indicates a positive genotype of the LMNA variant while '–' indicates a negative genotype, the period in the centre of a circle (self-identified female) or square (self-identified male) indicates a carrier with no clinical symptoms

5.3.3.3 Haplotype Analysis

Haplotyping was performed for all the enrolled family members by making use of a combination of LMNA microsatellites and the variant of interest (*LMNA* c.568C>T). DCM 569.1's microsatellite marker peaks are shown in Figure 123 below.

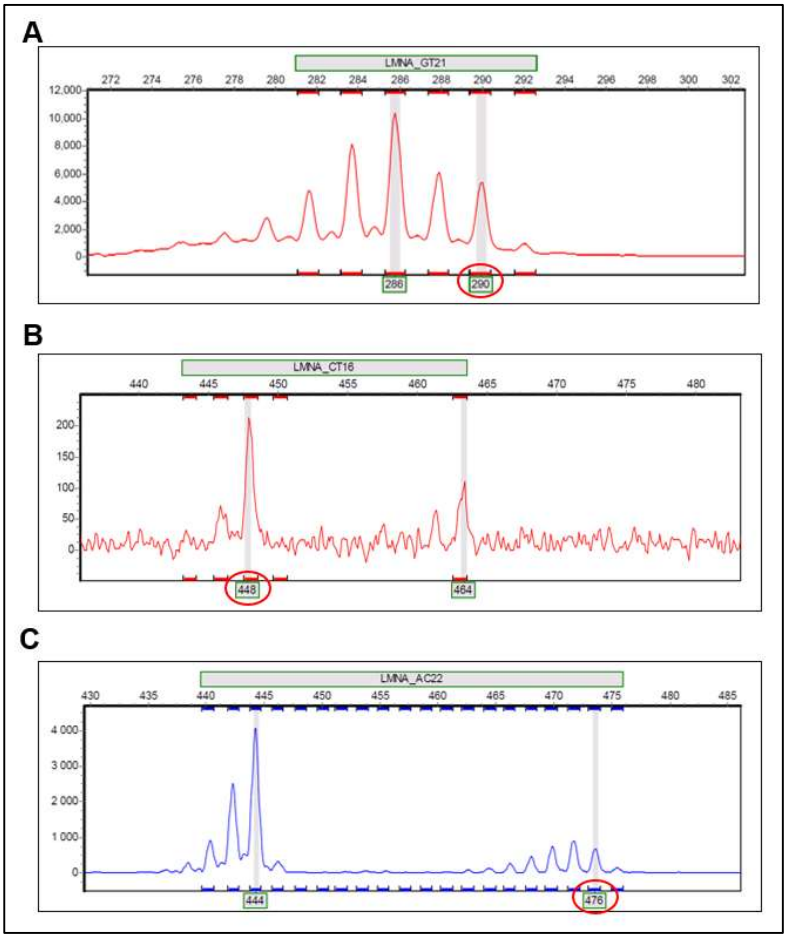


Figure 123: Microsatellite analysis of DCM 569.1. (A) LMNA_GT21 (282), the peak circled in red, 290, is the common haplotype peak for LMNA_GT21. (B) LMNA_CT16 (271), the peak circled in red, 448, is the common haplotype peak for LMNA_CT16. (C) LMNA_AC22 (476), the peak circled in red, 476, is the common haplotype peak for LMNA_AC22

The proband (DCM 569.1) and her cousin (DCM 567.1) both have the disease associated haploblock (290, 448,T, 476) as shown in Figure 124 below (the red haploblock), which is identical to the haploblock found in both the DCM 809 and DCM 781 families above.

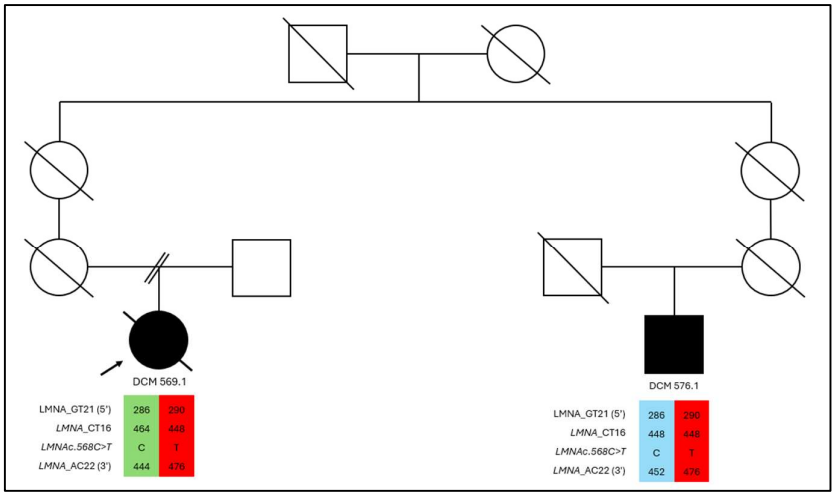


Figure 124: Haplotype analysis of family DCM 569. Circles indicate self-identified females, squares indicate self- identified males, black filled symbols represent clinically affected individuals, symbols with periods in the centre represent individuals with the LMNA variant but are not clinically affected, shapes with an oblique line through indicates deceased individuals, haplotypes with dotted boarders represent assumed haplotypes from segregation analysis

5.3.3.4 Genealogical Tracing

Limited information was provided for genealogical tracing of this family. The information provided by the genetic councillor included names, estimated birth dates and residences of some members of this family. This family resided in the Western Cape and is of Mixed-Ancestry.

The proband was the oldest *LMNA* positive member of this family and we therefore extended both the paternal and maternal lines. Despite the family's pedigree extending back through five generations, there was missing or wrong spelling of surnames, thus impeding genealogical tracing. Due this family descending from Mixed-Ancestries we made use of alternative methods of genealogical tracing. The racial separation of records during the apartheid period and poor or incomplete record keeping makes it difficult to accurately identify the ancestors of Mixed-Ancestry individuals. The Cape Town Archive possesses certain records pertaining to individuals of Mixed-Ancestry however, these records are limited to surnames. Consequently, efforts to obtain detailed information on members of this family from that historical period yielded inconclusive results.

Due to the lack of available records for the Mixed-Ancestry population group, we encountered f the same limitations in these three families as we did in Chapter 3 with the ACM 71 Mixed-Ancestry lineage, and Chapter 4 with the DCM 181 and DCM 307 Mixed-Ancestry lineages, where genealogical tracing was constrained to the early 20th century.

The three families are represented in table X below and which summarises the number of family members screened in each family, the number of participants positive for the *LMNA* variant and if they shared a common haplotype. This allowed us to compare the progenitors of each family and narrow it down to possible founder progenitor couples (Table 36 below).

Table 36: Study Participants screened for LMNA c.568C>T variant and common Haplotype

Family ID	No. of study participants screened	*LMNA* c.568C>T variant Positive	Common Haplotype
DCM 809	2	2	2
DCM 781	3	2	2
DCM 569	2	2	2
Total Participants	7	6	6

5.4 Discussion

This study used a combination of methods and resources such as clinical history, genetics, haplotype analysis and genealogical tracing to show that positively screened participants from three families in South Africa not only share the *LMNA* c.568C>T (p.Arg388Trp) variant associated with DCM, but they also share an identical haplotype.

5.4.1 Genotype-Phenotype

Genetics found three probands and an additional three family members heterozygous for the *LMNA* c.568C>T variant. All six participants presented with serve DCM at a very early age. In the DCM 809 family, the proband DCM 809.1 and his nephew DCM 615.2 were diagnosed at 33 and 35 years old respectively. The proband in DCM 781.1 was first diagnosed at 45 years old and received a heart transplant 8 years later. The proband (DCM 569.1) of the DCM 569 family was diagnosed at 38 and referred to GSH for a heart transplant but unfortunately died before receiving a heart. The proband's second cousin (DCM 576.1), was subsequently diagnosed at 41 and received a heart transplant 6 years later.

Previous research in our laboratory (unpublished data) also reported this *LMNA* c.568C>T variant in the IMHOTEP DCM cohort. Of the seven *LMNA* c.568C>T participants, 50% required heart transplants (DCM 781.1, DCM 569.1 and DCM 576.1). A detailed examination of the familial medical histories also revealed documented evidence of heart transplant procedures in a further five individuals from these families. We observed parallels in our

cohort with findings from studies conducted in Norway, where approximately 20% of participants harbouring various *LMNA* variants were reported to have undergone heart transplants (238).

The participants in this study with the *LMNA* c.568C>T variant show complete cardiac penetrance with the inheritance of one mutated allele leading to severe, early onset DCM. The participants in this study carrying the *LMNA* c.568C>T variant exhibited complete cardiac penetrance, with inheritance of a single mutated allele leading to severe, early-onset dilated cardiomyopathy (DCM). This is likely due to the critical physiological role of lamin A/C, the protein encoded by *LMNA*.

LMNA variants have been extensively documented in the literature and are associated with a spectrum of disorders known as laminopathies. Among these, pathogenic *LMNA* variants linked to DCM are particularly severe, as they frequently lead to progressive cardiac remodelling and HF. As discussed in Section 5.1, lamin A/C proteins are essential for nuclear envelope integrity, playing key roles in DNA replication, chromatin organization, and gene regulation (223, 225). Disruptions to these functions in cardiomyocytes almost invariably result in cardiac dysfunction and HF.

5.4.2 Haplotype Analysis Discussion

Haplotypes were constructed to span the entire *LMNA* gene by selecting three informative microsatellite markers that flanked the 5' (LMNA_21), 3' (LMNA_AC22) and within the gene (LMNA_CT15) and combined with the *LMNA* c.568C>T variant to form a haploblock. The positively genotyped participants for the *LMNA* c.568C>T variant, all shared an identical haplotype (represented by the red haploblock in the haplotype analysis sections), this supports our hypothesis that there is a common progenitor as the haplotype is identical by descent (IBD).

5.4.3 Genealogy Discussion

The genealogical analysis could not extend beyond the data presented in the pedigrees provided by the genetic counsellor for any of the families. This limitation arises from the Mixed-Ancestry of all three enrolled families, presenting challenges comparable to those encountered in the ACM 71 family (Chapter 3) and the DCM 181 and DCM 307 families

(Chapter 4). Neither traditional methods nor alternative methods yielded any results, due to several reasons.

One of the greatest difficulties was the historical disruption in record-keeping during the Apartheid era, when the South African government segregated all records by racial classification. Many records for individuals of certain racial groupings were incomplete or poorly maintained. The Cape Town Archive contained some records of Mixed-Ancestry individuals, but these typically listed only surnames without accompanying first names, making it difficult to establish family connections.

Beyond historical record limitations, tracing Mixed-Ancestry families was further complicated by variations in the spelling of names and surnames, repetition of names within the same family tree, children born out of wedlock, and cultural practices in which close family friends informally raised children who were considered "cousins" despite lacking genetic relation. Additionally, reluctance to share information and a lack of knowledge about family history further hindered genealogical tracing. Consulting specialists in this field, such as Mr. Aubrey Springveldt, a researcher and genealogist specialising in training people of the Cape to accurately trace their family histories, may allow for further tracing, but this approach would likely require direct interviews with family members, which may not be well received.

It is important to note that identifying the founder of a variant typically requires genealogical data from approximately 10-12 families (104). We are confident, that with the enrolment of an additional seven to nine families we would be able to identify a common founder for this highly pathogenic *LMNA* c.568C>T variant.

5.5 Conclusion

In the current study we have shown three unrelated families, who self-identified as Mixed-Ancestry, not only present with a severe form of DCM due to the pathogenic *LMNA* c.568C>T variant but also share an identical haplotype. We are confident through the pipeline we have created and the continued recruitment, screening and investigation into the *LMNA* c.568C>T variant, that a common founder can be identified.

The identification of just one patient with a pathogenic variant such as, *LMNA* c.568C>T could benefit many South African families as the screening of at-risk family members could aid in the early diagnosis, prevention and treatment of DCM thus prolonging their lives.

Chapter 6: *TTN* Variants in South African Families with DCM

6.1 Introduction

Dilated cardiomyopathy (DCM) is the most common and genetically diverse form of cardiomyopathy, as discussed in Chapters 1, 4, and 5. Over 40 genes have been implicated in DCM, with the primary contributors being sarcomere and cytoskeletal genes. The prevalence of DCM varies across ethnic and geographic populations (239) and familial DCM is reported in approximately 30–50% of cases (240).

In recent years, the titin (*TTN*) gene has been a major focus of research due to its strong association with DCM (241). Titin, encoded by TTN, is the largest human protein, with a gene size of approximately 100 kb, making it one of the largest in the human genome. It plays a crucial role in maintaining passive myocyte stiffness (242). The *TTN* gene is located on chromosome 2 and contains 364 exons (243), therefore titin is a complex multidomain protein and consists of four main structural and functional domains, the N-terminal, the I-band, the A-band and the C-terminal M-line. These domains are arranged to enable the modulation of titin expression and turnover through the tyrosine kinase domain which is as follows: the N-terminal Z-line serves as an anchor for the sarcomeric Z-disk, while the I-band contributes to the protein's elastic properties, the A-band plays a role in stabilizing the thick filament, and the C-terminal M-line region overlaps in an antiparallel orientation with the C-terminus of another titin molecule (244). Structurally, two titin molecules span the sarcomere, anchoring at the Z-line and M-line (Figure 125), making titin an essential component of sarcomere formation (241).

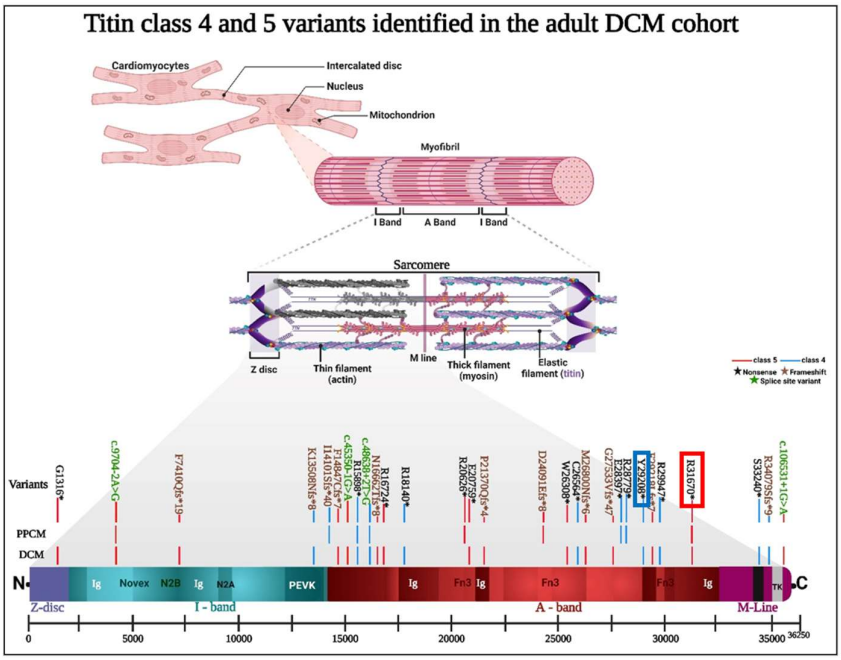


Figure 125: Schematic representation of the sarcomere with class 4 and 5 TTN variants: TTNc.87623A>T (p.Tyr29208Phe) variant highlighted by the blue box and TTN c.95008C>T (p.Arg31670Ter) highlighted by the red box (Adapted from unpublished Thesis, PN).

Although next-generation sequencing (NGS) has facilitated the sequencing of *TTN* (245), the immense size of this gene presents ongoing challenges for studying disease-related variants in isolation. Notwithstanding this difficulty, *TTN* has been recognized as a hotspot for pathogenic variants, resulting in a broad spectrum of disease phenotypes (243). In particular, truncating (nonsense) variants within *TTN* is the leading cause of familial DCM (246), further emphasizing its critical role in disease pathology.

Through the IMHOTEP registry of South Africa, previous unpublished research had identified two *TTN* variants within the hotspot: *TTN* c.95008C>T (p.Arg31670Ter) in the 368th exon and *TTN* c.87624C>A (p.Tyr29208Ter) in the 367th exon of the *TTN* gene. This study showed that the *TTN* c.95008C>T nonsense variant was detected in 14 probands and was located in the A-band of the TTN protein, resulting in a premature translational stop signal thereby truncating the TTN protein. This rare variant (rs1322596650) was reported in the GnomAD population database with a MAF score of A=0.000004 (1/248798) and according to the ACMG guidelines, was classified as likely pathogenic by Labcorp (formerly Invitae), through the ASPREE project (247)) and can be seen in Figure 125, in the red box (R30029*).

The *TTN* c.87624C>A variant, detected in 6 probands, was a nonsense variant, and was present in the A-band of the TTN protein. Although this variant was not predicted to result in nonsense-mediated decay, it did create a truncated TTN protein (76). There was no evidence of this variant in population databases, however, it has been previously reported by Labcorp (formerly Invitae) and has an ACMG classification of likely pathogenic. This variant can be observed in Figure 128, in the blue box (Y29208*).

6.2 Methods

All methods were outlined in Chapter 2, the flow diagram below represents the overall flow of chapters 3-6, which has been tailored to this chapter and refers to the relevant methods.

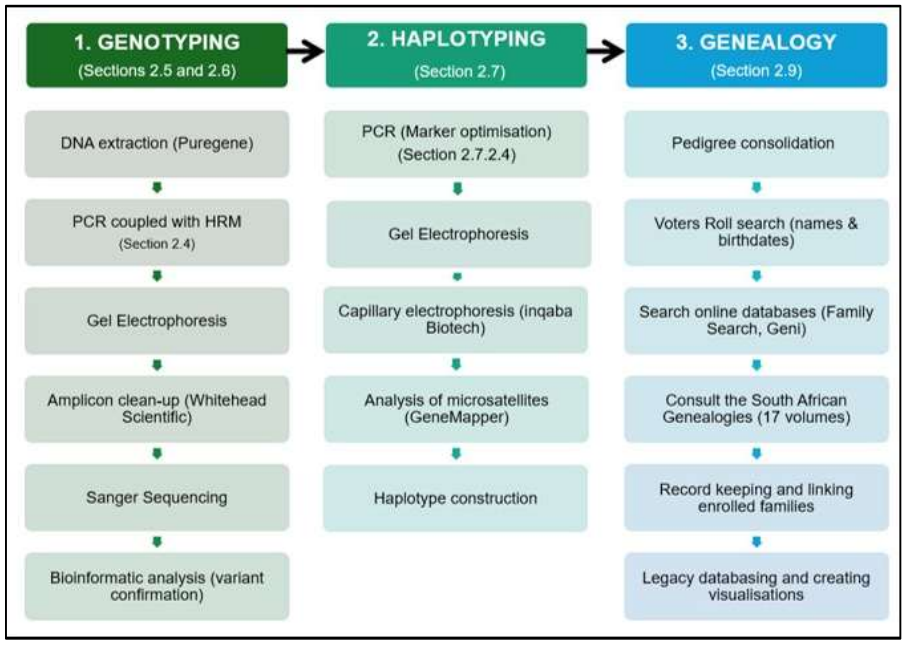


Figure 126: Workflow, adapted from Chapter 2, Section 2.3, to highlight the relevant sections detailing the methods associated with this chapter, specifically focusing on the TTN variants

6.2.1 Clinical

Patients with suspected DCM, were referred to the Cardiac Clinic at Groote Schuur Hospital and after their diagnosis was confirmed, they were enrolled into the IMHOTEP registry as per Chapter 2, section 2.2.

6.2.2 Genetics

The primer sequences and methodology for the *TTN* c.95008C>T (p.Arg31670Ter) and *TTN* c.87624C>A (p.Tyr29208Ter) variants can be found in Chapter 2, section 2.4, Table 6. The probands were genotyped and 14 probands were found to be heterozygous for the *TTN* c.95008C>T (p.Arg31670Ter) variant and six probands were heterozygous for the *TTN* c.87624C>A (p.Tyr29208Ter) variant. No additional family members were enrolled.

6.2.3 Haplotype

As in the previous chapters, haplotype analysis was conducted to facilitate variant tracing, elucidate inheritance patterns, enable disease mapping, and investigate familial history. A haplotype comprises a set of co-inherited alleles on a single chromosome. These linked alleles are typically transmitted as a unit (180).

Inferring haplotypes within a family becomes easier when the DNA from parents or siblings is available. Parental DNA allows the direct observation of the haplotypes transmitted to their offspring. Deducing the child's haplotypes can be done with greater certainty, if the parental haplotypes are already known (181). Similarly, sibling DNA can help resolve ambiguous haplotypes. By comparing the genetic variations shared between siblings, one can infer which segments of their chromosomes were inherited from each parent. This information is particularly useful in linkage studies, where researchers track the inheritance of specific haplotypes within families to identify genes associated with diseases (181, 182).

The same microsatellite markers were used for both variants as they occurred in the *TTN* gene. As described in Chapter 2 section 2.6.2.4, we used three microsatellite markers, TTN_TG at the 5’ end, TTN_AC at the 3’ end and TTN_TA within the *TTN* gene, thus spanning the entire gene.

6.2.4 Genealogy

Genealogy tracing followed standard procedures (Chapter 2, Figure 14), with the oldest enrolled family member used to commence genealogy tracing and where possible, the ancestral lineages were extended as far as records allowed.

6.3 Results

As we encountered many issues in our attempt to genealogically trace these families, we elected to present this chapter in a slightly different format to the previous chapters, to prevent repetition and redundancies.

Patients showing clinical symptoms of DCM were referred to the Cardiac Clinic at Groote Schuur Hospital during the period 2012 to 2024 and after informed consent was taken, they were enrolled into the IMHOTEP registry of South Africa. Attempts were made to prospectively recruit the family members of the probands but this was not successful and no additional family members were recorded. Clinical investigations into the family history were sparse and made mode of inheritance and genealogy tracing difficult.

6.3.1 Clinical History

We investigated 20 probands with *TTN* variants. Baseline characteristics indicated 13 of the probands were male and seven were female. The age of presentation varied with three probands (DCM 630.1, DCM 626.1 and DCM 662.1) having a childhood onset, the remaining 17 probands all presented with moderate DCM during adulthood with an age range from 20 to 64 years old. Clinical notes of DCM 622.1, DCM 626.1, DCM 497.1 and LVN 10.1 all showed an DCM/LVNC overlap. It was found in the clinical notes of half (10) of the probands that they had indicated a positive family history of cardiac related problems however, clinical investigations into the family history were sparse.

6.3.2 Genetics

Genotyping identified 14 probands with the *TTN* c.95008C>T variant and six probands with the *TTN* c.87624C>A variant.

Genotyping identified 14 probands (DCM 333.1, DCM 452.1, PCM 1.1, DCM 552.1, DCM 590.1, DCM 630.1, DCM 760.1, DCM 666.1, DCM 653.1, DCM 661.1, DCM 737.1, DCM 524.1, DCM 764.1 and LVN 10.1) that were all heterozygous for the *TTN* c.95008C>T variant. These probands self-identified as Black-Africans. Based on these findings, a common founder was suspected.

Genotyping identified six probands (DCM 622.1, DCM 659.1, DCM 497.1, DCM 555.1, DCM 626.1 and DCM 757.1) that were all heterozygous for the *TTN* c.87624C>A variant. Three of the probands self-identified as Caucasians (DCM 659.1, DCM 555.1 and DCM 757.1), two probands self-identified as Black-Africans (DCM 626.1 and DCM 622.1) and one proband (DCM 497.1) self-identified as Mixed-ancestry. Based on these findings, a common founder was suspected.

6.3.3 Haplotype analysis of South African Families with *TTN* c.95000C>T and *TTN* c.87624C>A variants

Clinical investigations into the family history of these 20 probands were sparse and made the mode of inheritance, construction of haplotypes and genealogy tracing difficult.

We designed three microsatellite markers which encompassed the 5', 3' and intergenic regions of the *TTN* gene. For each primer pair, we initially used a subset of the participants and a genotype negative control to optimise the primers and to test if the markers proved to be informative for haplotype analysis as described in Chapter 2 and section 6.2.3 above.

For the 5' region, we designed primers for the TTN_TG repeat, which had 19 repeats. This marker was successfully amplified and haplotyped for all 20 probands that had the two *TTN* variants.

For the 3' marker, we selected the TTN_AC marker, which had 17 repeats. Amplification was successful and even though there was some non-specific binding, the expected 410bp band was the brightest band on the gel (Figure 127) and could be used for microsatellite analysis. Amplicons were created for each of the probands harbouring the two *TTN* variants: six probands for *TTN* c.95008C>T and 14 for *TTN* c.87624C>A variants respectively.

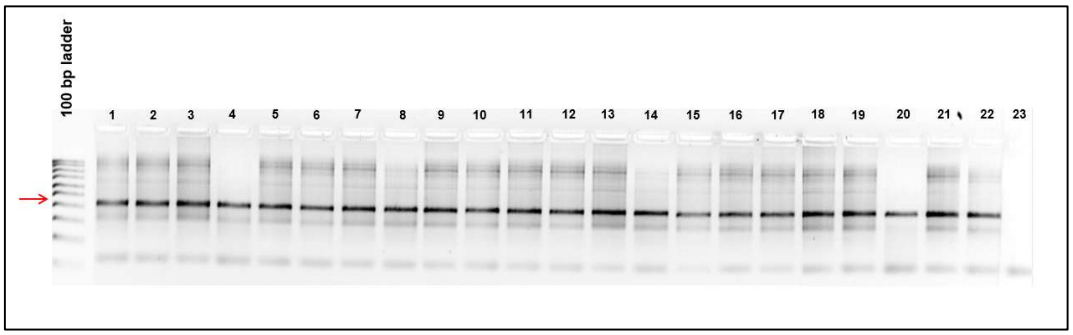


Figure 127: Agarose gel (1.5%) of the PCR products of the TTN_AC. The 3' microsatellite marker is indicated by the red arrow, 100bp ladder was used, lane one is the genotype negative control, lanes 2-22 are the participants samples, 2=DCM 764.1, 3 = DCM 497.1,4=DCM 590.1, 5=DCM 452.1, 6=PCM 1.1 , 7=DCM 333.1, 8=DCM LVN 10.1, 9=DCM 661.1,10=DCM 552.1, 11=DCM 630.1, 12=DCM 622.1, 13=DCM 757.1,14=DCM 555.1, 15=DCM 757.2, 16=DCM 737.1,17=DCM 524.1, 18=DCM 626.1, 19=DCM 666.1, 20=DCM 659.1, 21=DCM 653.1, 22=DCM 760.1 and lane 23 was the negative control.

Even though we had designed four sets (two we designed in-house, one primer pair was a published primer set, and one was designed by a company's technicians) all failed, and we were unable to amplify the amplicon. One final attempt resulted in a minor successful amplification of the TTN_TA marker which had 21 repeats and an expected band size of ~411bp. However, this fragment still needed to be optimised due to the double banding on the agarose gel, shown by the red arrow (Figure 128); unfortunately, time-constraints did not allow us to troubleshoot. We could thus not continue with further downstream experiments such as the addition of the fluorescent dye to the PCR products or capillary electrophoresis.

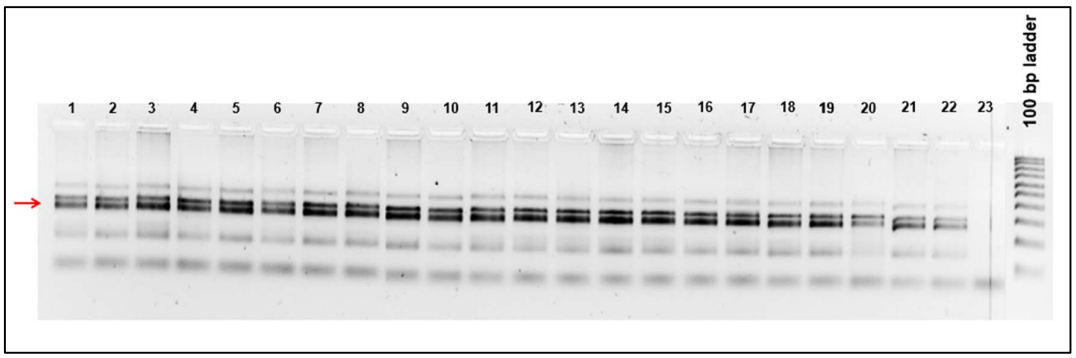


Figure 128: Agarose gel (1.5%) of the PCR products of the TTN_TA intergenic marker indicated by the red arrow, 100bp ladder was used, lane one is the genotype negative control, lanes 2-22 are the participants samples, 2=DCM 764.1, 3 = DCM 497.1,4=DCM 590.1, 5=DCM 452.1, 6=PCM 1.1 , 7=DCM 333.1, 8=DCM LVN 10.1, 9=DCM 661.1,10=DCM 552.1, 11=DCM 630.1, 12=DCM 622.1, 13=DCM 757.1,14=DCM 555.1, 15=DCM 757.2, 16=DCM 737.1,17=DCM 524.1, 18=DCM 626.1, 19=DCM 666.1, 20=DCM 659.1, 21=DCM 653.1, 22=DCM 760.1and lane 23 was the negative control

In summary, we were able to optimise two of the three markers for the *TTN* gene, the 5' marker was completed for all 20 probands and the 3' marker was completed for two probands, DCM 764.1 and DCM 497.1, with the *TTN* c.95008C>T and *TTN* c.87624C>A variants respectively. The two successfully optimised markers for DCM 764.1 and DCM 497.1 are shown in Figure 129 and Figure 130 respectively.

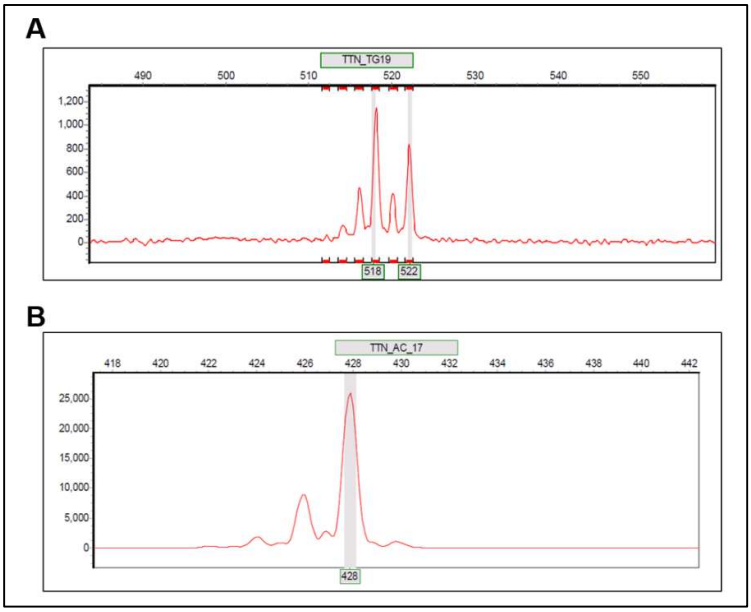


Figure 129: The electropherogram for DCM 764.1, for the binned peaks of the (A) 5' marker, TTN_TG. (B) 3' marker, TTN_AC

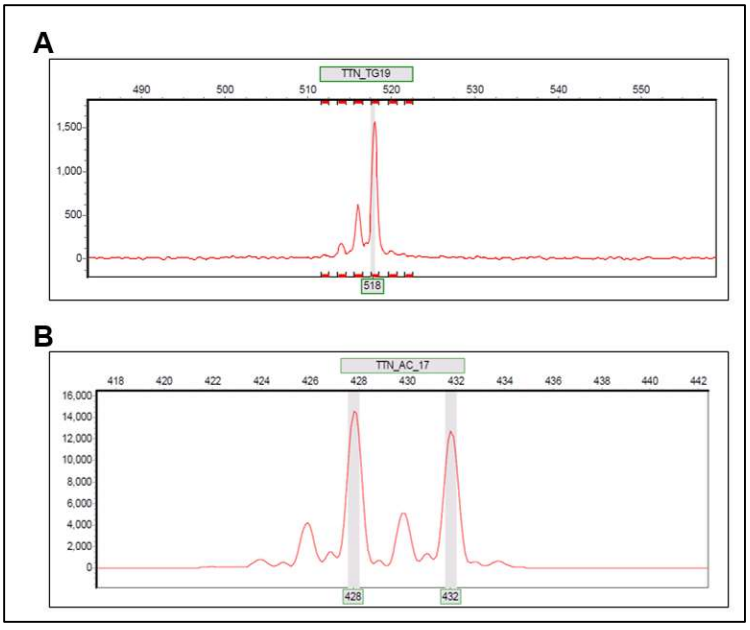


Figure 130: The electropherogram for DCM 497.1, for the binned peaks of (A) 5' marker, TTN_TG. (B) 3' marker, TTN_AC

6.3.4 Genealogy Tracing of South African Families with *TTN* c.95000C>T and *TTN* c.87624C>A variants

In previous chapters (Chapters 3 and 4), the standard methods used for genealogical tracing in this study were highly effective for tracing the ancestry of Caucasian Afrikaners in South Africa, largely due to the extensive documentation from the arrival of Dutch settlers to the present day. However, tracing the genealogies of Mixed-ancestry families (Chapters 3, 4, and 5) presented more challenges, and similar, if not more complex, obstacles arose when working with Black African probands carrying *TTN* variants. This will be further elaborated on in the discussion, genealogy (section 6.4.3) section.

6.4 Discussion

We identified two likely pathogenic truncating variants, *TTN* c.95008C>T and *TTN* c.87624C>A, in the *TTN* gene in two large cohorts, 14 and six probands respectively. Truncating variants in the *TTN* gene have been well researched and identified as the primary genetic cause of familial DCM (246).

6.4.1 Phenotype-Genotype

6.4.1.1 TTN c.95008C>T (p.Arg31670Ter) variant

The *TTN* c.95008C>T (p.Arg31670Ter) variant was classified as a likely pathogenic variant (LP) and was identified in 14 Black African DCM probands enrolled into the IMHOTEP registry. All 14 probands were heterozygous for the *TTN* c.95008C>T variant, with a significant age range between the probands, with the earliest onset being three years old in DCM 630.1 and the oldest presentation at 64 years old (DCM 653.1). Seven probands had indicated a positive family history of cardiac related problems, however, no further information was provided. The probands all presented with DCM before the age of 65 which contrasts the data published in the ASPREE study (247), where all the participants with this variant were all asymptomatic with ages >70 years old. It has been suggested in unpublished data from the CVG group that Black Africans may have DCM disease modifiers resulting in the manifestation of DCM.

Despite the dispersion of these individuals across South Africa, we found that all the probands with this *TTN* variant were Black African. Research conducted in the USA found that the incidence of DCM in Black African Americans was higher than that of White Americans (35). Their research also showed that Black patients present with increased familial risk and worse outcomes, which was attributed to differences in genetic architecture and a lack of clinical data to represent this ethnic group. A similar race disparity exists in Sub-Saharan Africa and therefore South Africa. Incidence is higher, outcomes and survival are worse and the referrals for heart transplants are fewer in Black African patients (248), highlighting the large knowledge gap that still exists for these patients.

This variant has also been found in asymptomatic older adults of European descent with a gnomAD minor allele frequency (MAF) of 0.000008882 (247). However, as previously mentioned, European and Black African genetic architecture differs and the lack of clinical data for Black Africans still poses a large knowledge gap (35).

6.4.1.2 TTN c.87624C>A (p. Tyr29208Ter) variant

The *TTN* c.87624C>A (p. Tyr29208Ter) variant has been reported previously as likely pathogenic and was identified in six probands in the IMHOTEP study. Research has shown that patients with this variant have an early onset mild form of DCM (249). We observed a similar trend in the six probands enrolled in this study. Additionally, four of the probands had indicated a positive family history of cardiac problems, however, no further detail was provided. The variant occurred in three Caucasian probands, two black African probands and a proband of Mixed-Ancestry. We noted that the two Black African probands, DCM 622.1 and DCM 626.1, both showed early onset of DCM at ages 3 and 11 respectively. While the Caucasian (DCM 659, DCM 555 and DCM 757) and Mixed-ancestry (DCM 497) probands only presented later in life, and with a mild DCM phenotype. These findings correlate with those discussed above, where researchers found differing phenotypes in individuals of European descent and African Americans due to their distinct differences in genetic architecture and the possibility of modifiers (35, 248, 250).

Truncating *TTN* variants and homozygous *TTN* variants are associated with different clinical outcomes and locations in the TTN protein. Truncating variants located in the A band have been linked to DCM (241) accounting for approximately 15% of all DCM cases and 25% of severe, end-stage, and familial cases (251); these variants lead to reduced protein production and are strongly associated with DCM. In contrast, homozygous or compound heterozygous *TTN* variants are associated with skeletal myopathy (252), often presenting as severe congenital muscle diseases that may or may not include cardiomyopathy. These variants are primarily located in the M-line exons (exons 359-362) (253). Researchers have noted that disease severity, penetrance and outcome is highly dependent on the type of variant, certain environmental factors as well as the distinct population or ethnicity of the patient with the *TTN* variant (251).

6.4.2 Haplotype-Analysis

We designed a total of eight primer pairs for six different microsatellite markers spanning the *TTN* gene. One primer pair was successfully designed for one microsatellite marker for the 5' end of the gene. Three primer pairs were designed for two different microsatellite markers for

the 3' end of the *TTN* gene. Lastly, four primer pairs were designed for three different microsatellite markers for the intronic marker, none of which were successful.

Haplotype analysis was not possible for either of the variants presented in this chapter, *TTN* c.95008C>T (p.Arg31670Ter) and *TTN* c.87624C>A (p. Tyr29208Ter).

Despite our best efforts, the optimisation of the intergenic microsatellite marker, TTN_TA was not possible during the laboratory phase of this study. As described in Chapter 2 and section 6.2.3 of this chapter a lot of optimisations of multiple microsatellite markers was required as amplification was problematic. We attempted to troubleshoot with temperature and $MgCl_2$ gradients, various Taq's, which included high-fidelity polymerases which are recommended as they minimize amplification errors , but still we still had difficulty in amplifying the intergenic marker.

The problems we encountered in the amplification of the TTN microsatellites could be explained by several challenges. *TTN* is the largest gene in the human genome and it was a huge feat when next-generation sequencing allowed for the complete sequencing of this gene (246). Another challenge when sequencing regions of the *TTN* gene, is its complexity. There are a high number of repetitive sequences further complicating sequencing and annotation (254), and when considering that microsatellite markers are short tandem repeats of repeated nucleic acid sequences, this further complicates optimisation for markers as highly repetitive content can lead to PCR slippage and therefore errors in amplification (255). We believe this may have what we observed with the intergenic TTN_TA marker. The regions surrounding this marker are highly repetitive, which caused non-specific binding and amplification of the incorrect region, in other words, slippage of the primers.

Recently research has shown that implementing long-range PCR or third-generation sequencing can vastly improve accuracy (254), however, this does come at a cost and budget restraints with all the prior primer design and optimisation already done as well as limited time remaining for the laboratory work, we could not invest anymore funds or time into the optimisation of the *TTN* markers.

We believe given more time and resources, this intergenic TTN_TA microsatellite marker could be optimised, and haplotypes could be constructed for the probands with these variants. As, the partial haplotype data of the 5' and 3' markers and the successful optimisation of these markers lays a strong foundation for future analyses and further development of this research.

6.4.3 Genealogy

Genealogical tracing of the *TTN* c.95008C>T (p.Arg31670Ter) and *TTN* c.87624C>A (p.Tyr29208Ter) variants was challenging due to incomplete haplotype construction. Without haplotype information to guide the tracing process, we lacked a key tool that had been instrumental in earlier chapters.

For the *TTN* c.95008C>T variant, five of the probands were found to reside in the Western Cape region, five in Bloemfontein, one in Qqeberha, one in the Eastern Cape and one in the Free State.

The variant *TTN* c.87624C>A (p. Tyr29208Ter) was found in six probands, three Caucasians, two black African and one of Mixed-Ancestry, Despite the various ancestries observed, there was a distinct clustering of ethnicity and location. We found that four of the six probands resided in the Western Cape, these individuals were Caucasian or of mixed ancestry and the Black Africans resided in Qqeberha. The only outlier was a Caucasian male proband residing in Qqeberha, however, migration from the Western Cape to the East Coast is common as previously seen and described in Chapter 3.

As only the probands for both *TTN* variants were enrolled we did not have a family history to assist in tracing the origin of the variant in this unique South African sub-population.

In Chapters 3, 4, and 5, haplotype construction helped identify which lineages to extend by tracing the oldest enrolled family member carrying the disease-associated haploblock. Furthermore, the identical haplotypes observed in these previous studies provided evidence of a common founder for those variants. However, the absence of complete genealogy records in Black African populations necessitated exploring alternative tracing methods, as the standard techniques used for Afrikaner populations yielded no results due to multiple factors.

As discussed in earlier chapters, the South African government historically separated records by racial group, leading to incomplete or poorly maintained records for Black African and Mixed-Ancestry individuals, which applied to the probands carrying the *TTN* variants. This has been a significant challenge for genealogical tracing efforts in South Africa.

Present-day South Africa, known as the "rainbow nation", is a product of its rich cultural diversity and complex history. This diversity stems from multiple waves of migration and colonization that began approximately 1,700 years ago, when farmers and metalworkers

travelled from northern Africa to South Africa, what is now termed the unofficial colonization of South Africa (256). Over time, various clans developed distinct identities, traditions, and hierarchical socio-political structures (257). The arrival of Europeans, the official colonisation of South Africa, brought significant changes to indigenous populations, including the dispossession of the KhoiKhoi and San peoples and the introduction of slavery. Records from the slave trade, which began in 1658 with shiploads of slaves from Dahomey and Angola (258), were poorly maintained. Between 1658 and 1808, an estimated that approximately 63000 slaves from various regions along the Indian Ocean rim were brought to the Cape (259). This influx of diverse groups led to a highly heterogeneous population, with lasting impacts on South Africa's societal structure (260).

Significant cultural and language barriers also posed challenges to tracing genealogies. A retired genealogist (private correspondence) who spent nearly two decades integrating into a Cape Province tribe gained the tribe's trust, learned their language, and became familiar with their traditions. South African tribes have unique practices, including naming ceremonies that vary significantly across tribes and regions (261-263).This diversity in cultural practices emphasizes the need for in-depth knowledge of African naming traditions, which would take years to acquire. Without such expertise, tracing genealogies for the probands and their families became even more difficult.

Although, we had 14 probands for the one variant, and the estimated number required to successfully trace back to a common founder is 10-12 (104), the clinical pedigrees lacked crucial information across multiple generations. In most instances, the names and birthplaces of the previous generation were unknown. This may be attributed to language barriers between patients and clinicians, long travel times and costs preventing family members from accompanying the proband to the clinic, and poor patient follow-up (264). Additionally, many South Africans, especially those in rural areas, still face limited access to healthcare, further complicating genealogical tracing efforts.

Despite these limitations, significant efforts were made, and the information gathered provided valuable insights. While we were unable to extend genealogies beyond the probands due to the limited available records, the work done in this study lays the groundwork for potential future advancements in tracing these family histories.

6.5 Conclusion

Genealogical tracing of Black African families in South Africa presents significant challenges due to historical factors such as forced migrations, the transatlantic slave trade, and the destruction or absence of comprehensive genealogical records. Cultural diversity and variations in naming conventions further complicate these efforts. To overcome these obstacles, future research should incorporate interdisciplinary approaches. Collaborations with cultural experts, linguists, and indigenous knowledge holders can provide invaluable insights into clan dynamics, oral histories, and traditional naming practices. Integrating these qualitative data sources with genetic analyses and historical records may offer a more nuanced and comprehensive understanding of familial lineages and population histories within South Africa.

Continued research and optimization of microsatellites within and surrounding the *TTN* gene should remain a priority, as this would facilitate haplotype construction, providing evidence that the variants identified in the unique South African population are founder variants.

Thus, the continuation of this study, through the pipeline we have set-up and presented throughout this thesis, by recruiting and genetic screening additional probands and family members presenting with DCM and carrying *TTN* variants through the IMHOTEP registry could offer deeper insights into the origin of these variants.

Chapter 7: Conclusion of Founder variants in South African Families

7.1 Conclusion

This project formed part of the larger IMHOTEP registry which is the largest cardiomyopathy registry in Africa. Throughout this thesis we provided strong evidence for the presence of multiple founder variants within South African populations, highlighting the significance of genealogical tracing, genetic screening, and haplotype analysis in understanding the origins and clinical impact of these variants. We were able to genealogically trace eleven of the families with the *PKP2* variant to their Dutch and French Huguenot progenitors who arrived in South Africa in the 1600's. We were also able to trace a three-generation family with the *BAG3* variant to their progenitors and evidence suggests this variant is of French-Canadian ancestry. Our research enhances the understanding of cardiomyopathies in South Africa and we believed we have laid the groundwork and set up pipeline for future work to successfully identify and genealogically trace unrelated families who harbour identical pathogenic genetic variants.

7.2 Study limitations

A significant limitation of this study is that most participants were recruited from government and tertiary hospitals across South Africa, which primarily service urban and peri-urban areas. This recruitment strategy may have led to an underestimation of the prevalence of these variants in the South African population. Additionally, the absence of comprehensive historical records for Mixed-Ancestry and Black African participants, combined with cultural and language barriers, hindered accurate and complete genealogical tracing of these families. Furthermore, the lack of resources to optimise the *TTN* microsatellite markers posed a challenge in constructing haplotypes for the 20 enrolled families with the two TTN variants, subsequently affecting the ability to trace their genealogical histories effectively.

7.3 Future recommendations

Continued recruitment of probands and additional family members from South Africa's unique population remains a top priority for this study. This is particularly critical for the *BAG3* and *LMNA* variants, where a common haplotype has been identified but further recruitment of seven to nine additional probands and their families is necessary to confirm

common founders. Additionally, further optimisation of the *TTN* microsatellite markers and subsequent haplotype construction will provide essential insights to support ongoing investigations into the origins of the two *TTN* variants. Additionally, the collaboration with Mixed-Ancestry and Black-African genealogists could aid in the genealogical tracing of these families.

References:

1. Antonopoulos AS, Xintarakou A, Protonotarios A, Lazaros G, Miliou A, Tsioufis K, et al. Imagenetics for Precision Medicine in Dilated Cardiomyopathy. Circulation. 2024;17(2).
2. Huang H, Verma J, Mok V, Bharadwaj HR, Alrawashdeh MM, Aratikatla A, et al. Exploring Health Care Disparities in Genetic Testing and Research for Hereditary Cardiomyopathy: Current State and Future Perspectives. Global Medical Genetics. 2024;11(1).
3. Bikbov B, Purcell CA, Levey AS, Smith M, Abdoli A, Abebe M, et al. Global, regional, and national burden of chronic kidney disease, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet. 2020;395(10225):709-33.
4. Damasceno AMDP, Cotter GMD, Dzudie AMD, Sliwa KMDP, Mayosi BMMD. Heart Failure in Sub-Saharan Africa: Time for Action. Journal of the American College of Cardiology. 2007;50(17):1688-93.
5. Agbor VN, Essouma M, Ntusi NAB, Nyaga UF, Bigna JJ, Noubiap JJ. Heart failure in sub-Saharan Africa: A contemporaneous systematic review and meta-analysis. International journal of cardiology. 2018;257:207-15.
6. Sliwa K, Mayosi BM. Recent advances in the epidemiology, pathogenesis and prognosis of acute heart failure and cardiomyopathy in Africa. Heart (British Cardiac Society). 2013;99(18):1317-22.
7. Siddikatou D, Mandeng Ma Linwa E, Ndobo V, Nkoke C, Mouliom S, Ndom MS, et al. Heart failure outcomes in Sub-Saharan Africa: a scoping review of recent studies conducted after the 2022 AHA/ACC/HFSA guideline release. BMC Cardiovascular Disorders. 2025;25(1).
8. Ormerod JOM, Yavari A. Cardiomyopathies. Medicine (Abingdon 1995, UK ed). 2022;50(8):492-506.
9. Kraus SM, Shaboodien G, Francis V, Laing N, Cirota J, Chin A, et al. Rationale and design of the African Cardiomyopathy and Myocarditis Registry Program: The IMHOTEP study. International journal of cardiology. 2021;333:119-26.
10. Maron BJ, Towbin JA, Thiene G, Antzelevitch C, Corrado D, Arnett D, et al. Contemporary Definitions and Classification of the Cardiomyopathies. Circulation. 2006;113(14):1807-16.
11. Corrado D, Perazzolo Marra M, Zorzi A, Beffagna G, Cipriani A, Lazzari MD, et al. Diagnosis of arrhythmogenic cardiomyopathy: The Padua criteria. International Journal of Cardiology. 2020;319:106-14.
12. Heymans S, Lakdawala NK, Tschöpe C, Klingel K. Dilated cardiomyopathy: causes, mechanisms, and current and future treatment approaches. The Lancet (British edition). 2023;402(10406):998-1011.
13. Ferreira A, Ferreira V, Antunes MM, Lousinha A, Pereira-Da-Silva T, Antunes D, et al. Dilated Cardiomyopathy: A Comprehensive Approach to Diagnosis and Risk Stratification. Biomedicines. 2023;11(3):834.
14. Shaboodien G, Spracklen TF, Kamuli S, Ndibangwi P, Niekerk CV, Ntusi NAB. Genetics of inherited cardiomyopathies in Africa. Cardiovascular diagnosis and therapy. 2020;10(2):262-78.
15. El Hadi H, Freund A, Desch S, Thiele H, Majunke N. Hypertrophic, Dilated, and Arrhythmogenic Cardiomyopathy: Where Are We? Biomedicines. 2023;11(2):524.

16. Bozkurt B, Coats AJS, Tsutsui H, Abdelhamid CM, Adamopoulos S, Albert N, et al. Universal definition and classification of heart failure: a report of the Heart Failure Society of America, Heart Failure Association of the European Society of Cardiology, Japanese Heart Failure Society and Writing Committee of the Universal Definition o. European Journal of Heart Failure. 2021;23(3):352-80.
17. Lund LH, Edwards LB, Dipchand AI, Goldfarb S, Kucheryavaya AY, Levvey BJ, et al. The Registry of the International Society for Heart and Lung Transplantation: Thirty-third Adult Heart Transplantation Report—2016; Focus Theme: Primary Diagnostic Indications for Transplant. The Journal of heart and lung transplantation : the official publication of the International Society for Heart Transplantation. 2016;35(10):1158-69.
18. Halliday BP, Gulati A, Ali A, Newsome S, Lota A, Tayal U, et al. Sex- and age-based differences in the natural history and outcome of dilated cardiomyopathy. European Journal of Heart Failure. 2018;20(10):1392-400.
19. Hershberger RE, Givertz MM, Ho CY, Judge DP, Kantor PF, McBride KL, et al. Genetic Evaluation of Cardiomyopathy—A Heart Failure Society of America Practice Guideline. Journal of cardiac failure. 2018;24(5):281-302.
20. Bondue A, Arbustini E, Bianco A, Ciccarelli M, Dawson D, De Rosa M, et al. 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. 2018;114(10):1287-303.
21. Japp AG, Gulati A, Cook SA, Cowie MR, Prasad SK. The Diagnosis and Evaluation of Dilated Cardiomyopathy. Journal of the American College of Cardiology. 2016;67(25):2996-3010.
22. Rapezzi C, Arbustini E, Caforio ALP, Charron P, Gimeno-Blanes J, Helio T, et al. Diagnostic work-up in cardiomyopathies: bridging the gap between clinical phenotypes and final diagnosis. A position statement from the ESC Working Group on Myocardial and Pericardial Diseases. European Heart Journal. 2013;34(19):1448-58.
23. Grimm W, Christ M, Bach J, MüLler H-H, Maisch B. Noninvasive Arrhythmia Risk Stratification in Idiopathic Dilated Cardiomyopathy. Circulation. 2003;108(23):2883-91.
24. Donal E, Delgado V, Bucciarelli-Ducci C, Galli E, Haugaa KH, Charron P, et al. Multimodality imaging in the diagnosis, risk stratification, and management of patients with dilated cardiomyopathies: an expert consensus document from the European Association of Cardiovascular Imaging. European Heart Journal - Cardiovascular Imaging. 2019;20(10):1075-93.
25. Patel AR, Kramer CM. Role of Cardiac Magnetic Resonance in the Diagnosis and Prognosis of Nonischemic Cardiomyopathy. JACC: Cardiovascular Imaging. 2017;10(10):1180-93.
26. Bui QM, Ding J, Hong KN, Adler EA. The Genetic Evaluation of Dilated Cardiomyopathy. Structural Heart. 2023;7(5):100200.
27. Hazebroek MR, Krapels I, Verdonschot J, Van Den Wijngaard A, Vanhoutte E, Hoos M, et al. Prevalence of Pathogenic Gene Mutations and Prognosis Do Not Differ in Isolated Left Ventricular Dysfunction Compared With Dilated Cardiomyopathy. Circulation: Heart Failure. 2018;11(3):e004682.
28. Jordan E, Peterson L, Ai T, Asatryan B, Bronicki L, Brown E, et al. Evidence-Based Assessment of Genes in Dilated Cardiomyopathy. Circulation. 2021;144(1):7-19.
29. Mcnally EM, Mestroni L. Dilated Cardiomyopathy. Circulation Research. 2017;121(7):731-48.

30. Wilde AAM, Semsarian C, Márquez MF, Sepehri Shamloo A, Ackerman MJ, Ashley EA, et al. European Heart Rhythm Association (EHRA)/Heart Rhythm Society (HRS)/Asia Pacific Heart Rhythm Society (APHRS)/Latin American Heart Rhythm Society (LAHRS) Expert Consensus Statement on the State of Genetic Testing for Cardiac Diseases. Heart Rhythm. 2022;19(7):e1-e60.
31. Rosenbaum AN, Agre KE, Pereira NL. Genetics of dilated cardiomyopathy: practical implications for heart failure management. Nature Reviews Cardiology. 2020;17(5):286-97.
32. Wang S, Zhang Z, He J, Liu J, Guo X, Chu H, et al. Comprehensive review on gene mutations contributing to dilated cardiomyopathy. Frontiers in Cardiovascular Medicine. 2023;10.
33. Tsabedze N, Ramsay M, Krause A, Wells Q, Mpanya D, Manga P. The genetic basis for adult-onset idiopathic dilated cardiomyopathy in people of African descent. Heart Failure Reviews. 2023;28(4):879-92.
34. Fundikira LS, Chillo P, Mutagaywa R, Kamuhabwa A, Kwesigabo G, Asselbergs FW, et al. Risk Factors and Prevalence of Dilated Cardiomyopathy in Sub-Saharan Africa: A Systematic Review. Global Heart. 2022;17(1):76.
35. Jordan E, Kinnamon DD, Haas GJ, Hofmeyer M, Kransdorf E, Ewald GA, et al. Genetic Architecture of Dilated Cardiomyopathy in Individuals of African and European Ancestry. JAMA. 2023;330(5):432.
36. Kraus SM, Cirota J, Pandie S, Thomas K, Thomas M, Makotoko M, et al. Etiology and Phenotypes of Cardiomyopathy in Southern Africa. JACC. 2024:100952.
37. Corrado D, Zorzi A, Cipriani A, Bauce B, Bariani R, Beffagna G, et al. Evolving Diagnostic Criteria for Arrhythmogenic Cardiomyopathy. Journal of the American Heart Association. 2021;10(18).
38. Di Gioia CRT, Giordano C, Cerbelli B, Pisano A, Perli E, De Dominicis E, et al. Nonischemic left ventricular scar and cardiac sudden death in the young. Human Pathology. 2016;58:78-89.
39. Towbin JA, Mckenna WJ, Abrams DJ, Ackerman MJ, Calkins H, Darrieux FCC, et al. 2019 HRS expert consensus statement on evaluation, risk stratification, and management of arrhythmogenic cardiomyopathy: Executive summary. Heart Rhythm. 2019;16(11):e373-e407.
40. Bosman LP, Verstraelen TE, Van Lint FHM, Cox MGPJ, Groeneweg JA, Mast TP, et al. The Netherlands Arrhythmogenic Cardiomyopathy Registry: design and status update. Netherlands Heart Journal. 2019;27(10):480-6.
41. Graziano F, Zorzi A, Cipriani A, De Lazzari M, Bauce B, Rigato I, et al. The 2020 “Padua Criteria” for Diagnosis and Phenotype Characterization of Arrhythmogenic Cardiomyopathy in Clinical Practice. Journal of Clinical Medicine. 2022;11(1):279.
42. Bhonsale A, Groeneweg JA, James CA, Dooijes D, Tichnell C, Jongbloed JDH, et al. Impact of genotype on clinical course in arrhythmogenic right ventricular dysplasia/cardiomyopathy-associated mutation carriers. European Heart Journal. 2015;36(14):847-55.
43. Austin KM, Trembley MA, Chandler SF, Sanders SP, Saffitz JE, Abrams DJ, et al. Molecular mechanisms of arrhythmogenic cardiomyopathy. Nature Reviews Cardiology. 2019;16(9):519-37.
44. Arrhythmogenic Right Ventricular Cardiomyopathy Overview [Internet]. 2023 [cited 15/01/2025]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1131/.

45. Brun F, Gigli M, Graw SL, Judge DP, Merlo M, Murray B, et al. FLNC truncations cause arrhythmogenic right ventricular cardiomyopathy. Journal of Medical Genetics. 2020;57(4):254-7.
46. Bhonsale A, Te Riele ASJM, Sawant AC, Groeneweg JA, James CA, Murray B, et al. Cardiac phenotype and long-term prognosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia patients with late presentation. Heart Rhythm. 2017;14(6):883-91.
47. Hof IE, Van Der Heijden JF, Kranias EG, Sanoudou D, De Boer RA, Van Tintelen JP, et al. Prevalence and cardiac phenotype of patients with a phospholamban mutation. Netherlands Heart Journal. 2019;27(2):64-9.
48. Baskin B, Skinner JR, Sanatani S, Terespolsky D, Krahn AD, Ray PN, et al. TMEM43 mutations associated with arrhythmogenic right ventricular cardiomyopathy in non-Newfoundland populations. Human Genetics. 2013;132(11):1245-52.
49. Lorenzon A, Beffagna G, Bauce B, De Bortoli M, Li Mura IEA, Calore M, et al. Desmin Mutations and Arrhythmogenic Right Ventricular Cardiomyopathy. The American Journal of Cardiology. 2013;111(3):400-5.
50. Taylor M, Graw S, Sinagra G, Barnes C, Slavov D, Brun F, et al. Genetic Variation in Titin in Arrhythmogenic Right Ventricular Cardiomyopathy–Overlap Syndromes. Circulation. 2011;124(8):876-85.
51. Te Riele ASJM, Agullo-Pascual E, James CA, Leo-Macias A, Cerrone M, Zhang M, et al. Multilevel analyses of SCN5A mutations in arrhythmogenic right ventricular dysplasia/cardiomyopathy suggest non-canonical mechanisms for disease pathogenesis. Cardiovascular Research. 2017;113(1):102-11.
52. Parikh VN, Caleshu C, Reuter C, Lazzeroni LC, Ingles J, Garcia J, et al. Regional Variation in RBM20 Causes a Highly Penetrant Arrhythmogenic Cardiomyopathy. Circulation: Heart Failure. 2019;12(3).
53. Bonny A, Ngantcha M, Scholtz W, Chin A, Nel G, Anzouan-Kacou J-B, et al. Cardiac Arrhythmias in Africa: Epidemiology, Management Challenges, and Perspectives. Journal of the American College of Cardiology. 2019;73(1).
54. Watkins DA, Hendricks N, Shaboodien G, Mbele M, Parker M, Vezi BZ, et al. Clinical features, survival experience, and profile of plakophylin-2 gene mutations in participants of the Arrhythmogenic Right Ventricular Cardiomyopathy Registry of South Africa. Heart rhythm. 2009;6(11):S10-S7.
55. Talle M, Bonny A, Scholtz W, Chin A, Nel G, Karaye K, et al. Status of cardiac arrhythmia services in Africa in 2018: a PASCAR Sudden Cardiac Death Task Force report. Cardiovascular Journal of Africa. 2018;29(2):115-21.
56. Kraus SM, Samuels P, Jermy S, Laing N, Van der Wall M, September U, et al. Clinical and cardiovascular magnetic resonance profile of cardiomyopathy patients from South Africa: Pilot of the IMHOTEP study. International journal of cardiology. 2024;399:131767-.
57. Elliott PM, Anastasakis A, Borger MA, Borggrefe M, Cecchi F, Charron P, et al. 2014 ESC guidelines on diagnosis and management of hypertrophic cardiomyopathy: The task force for the diagnosis and management of hypertrophic cardiomyopathy of the European Society of Cardiology (ESC). European heart journal. 2014;35(39):2733-79.
58. Marian AJ, Braunwald E. Hypertrophic Cardiomyopathy. Circulation Research. 2017;121(7):749-70.
59. Maron BJ, Desai MY, Nishimura RA, Spirito P, Rakowski H, Towbin JA, et al. Diagnosis and Evaluation of Hypertrophic Cardiomyopathy. Journal of the American College of Cardiology. 2022;79(4):372-89.

60. Ommen SR, Mital S, Burke MA, Day SM, Deswal A, Elliott P, et al. 2020 AHA/ACC Guideline for the Diagnosis and Treatment of Patients With Hypertrophic Cardiomyopathy. Circulation. 2020;142(25).
61. Jacobsen M-HB, Petersen JK, Modin D, Butt JH, Thune JJ, Bundgaard H, et al. Long term mortality in patients with hypertrophic cardiomyopathy – A Danish nationwide study. American heart journal plus : cardiology research and practice. 2023;25:100244.
62. Depasquale EC, Nasir K, Jacoby DL. Outcomes of adults with restrictive cardiomyopathy after heart transplantation. The Journal of Heart and Lung Transplantation. 2012;31(12):1269-75.
63. Ingles J, Goldstein J, Thaxton C, Caleshu C, Corty EW, Crowley SB, et al. Evaluating the Clinical Validity of Hypertrophic Cardiomyopathy Genes. Circulation: Genomic and Precision Medicine. 2019;12(2).
64. Hespe S, Waddell A, Asatryan B, Owens E, Thaxton C, Adduru M-L, et al. Genes Associated With Hypertrophic Cardiomyopathy: A Reappraisal by the ClinGen Hereditary Cardiovascular Disease Gene Curation Expert Panel. Journal of the American College of Cardiology. 2025/02/25;85(7).
65. Walsh R, Thomson KL, Ware JS, Funke BH, Woodley J, Mcguire KJ, et al. Reassessment of Mendelian gene pathogenicity using 7,855 cardiomyopathy cases and 60,706 reference samples. Genetics in Medicine. 2017;19(2):192-203.
66. Thierfelder L, Watkins H, MacRae C, Lamas R, McKenna W, Vosberg H-P, et al. α-tropomyosin and cardiac troponin T mutations cause familial hypertrophic cardiomyopathy: A disease of the sarcomere. Cell. 1994;77(5):701-12.
67. Poetter K, Jiang H, Hassanzadeh S, Master SR, Chang A, Dalakas MC, et al. Mutations in either the essential or regulatory light chains of myosin are associated with a rare myopathy in human heart and skeletal muscle. Nature genetics. 1996;13(1):63-9.
68. Geier C, Gehmlich K, Ehler E, Hassfeld S, Perrot A, Hayess K, et al. Beyond the sarcomere: CSRP3 mutations cause hypertrophic cardiomyopathy. Human molecular genetics. 2008;17(18):2753-65.
69. Lopes LR, Ho CY, Elliott PM. Genetics of hypertrophic cardiomyopathy: established and emerging implications for clinical practice. European Heart Journal. 2024;45(30):2727-34.
70. Moolman-Smook JC, De Lange WJ, Bruwer ECD, Brink PA, Corfield VA. The Origins of Hypertrophic Cardiomyopathy–Causing Mutations in Two South African Subpopulations: A Unique Profile of Both Independent and Founder Events. The American Journal of Human Genetics. 1999;65(5):1308-20.
71. Sliwa K, Damasceno A, Mayosi BM. Epidemiology and Etiology of Cardiomyopathy in Africa. Circulation. 2005;112(23):3577-83.
72. Ntusi NAB, Sliwa K. Associations of Race and Ethnicity With Presentation and Outcomes of Hypertrophic Cardiomyopathy. Journal of the American College of Cardiology. 2021;78(25):2573-9.
73. Muchtar E, Blauwet LA, Gertz MA. Restrictive Cardiomyopathy: Genetics, Pathogenesis, Clinical Manifestations, Diagnosis, and Therapy. Circulation research. 2017;121(7):819-37.
74. Brodehl A, Gerull B. Genetic Insights into Primary Restrictive Cardiomyopathy. Journal of Clinical Medicine. 2022;11(8):2094.
75. Rapezzi C, Aimo A, Barison A, Emdin M, Porcari A, Linhart A, et al. Restrictive cardiomyopathy: definition and diagnosis. European Heart Journal. 2022;43(45):4679-93.

76. Roberts A, Ware J, Herman D, Schafer S, Mazzarotto F, Baksi J, et al. 163 Integrated allelic, transcriptional, and phenotypic dissection of the cardiac effects of titin variation in health and disease. Heart. 2015;101(Suppl 4):A93.1-A.
77. Yang Z, Chen J, Li H, Lin Y. Genotype-Phenotype Associations with Restrictive Cardiomyopathy Induced by Pathogenic Genetic Mutations. Reviews in cardiovascular medicine. 2022;23(6).
78. Gowda SA-O, Ali HA-O, Hussain IA-O. Overview of Restrictive Cardiomyopathies. (1947-6108 (Electronic)).
79. Brodehl A, Pour Hakimi SA, Stanasiuk C, Ratnavadivel S, Hendig D, Gaertner A, et al. Restrictive Cardiomyopathy is Caused by a Novel Homozygous Desmin (DES) Mutation p.Y122H Leading to a Severe Filament Assembly Defect. Genes. 2019;10(11):918.
80. Peled Y, Gramlich M, Yoskovitz G, Feinberg MS, Afek A, Polak-Charcon S, et al. Titin Mutation in Familial Restrictive Cardiomyopathy. International journal of cardiology. 2014;171(1):24-30.
81. Gallucci AE, Grewal MR, Alexander BT, Heyer AM, Diaz YM. Left Ventricular Non-compaction Cardiomyopathy: The Key to Its Diagnosis and Implications for Management. Cureus. 2023.
82. Towbin JA. Inherited Cardiomyopathies. Circulation Journal. 2014;78(10):2347-56.
83. Cuevas J, Ptaszynski R, Cigarrán H, Calvo J, Martín M. Left ventricular noncompaction cardiomyopathy: Recent advances. Kardiologia Polska. 2022;80(5):529-39.
84. Arbelo E, Protonotarios A, Gimeno JR, Arbustini E, Barriales-Villa R, Basso C, et al. 2023 ESC Guidelines for the management of cardiomyopathies. European Heart Journal. 2023;44(37):3503-626.
85. Ross SB, Jones K, Blanch B, Puranik R, Mcgeechan K, Barratt A, et al. A systematic review and meta-analysis of the prevalence of left ventricular non-compaction in adults. European Heart Journal. 2020;41(14):1428-36.
86. Pignatelli RH, Mcmahon CJ, Dreyer WJ, Denfield SW, Price J, Belmont JW, et al. Clinical Characterization of Left Ventricular Noncompaction in Children. Circulation. 2003;108(21):2672-8.
87. Nugent AW, Daubeney PEF, Chondros P, Carlin JB, Cheung M, Wilkinson LC, et al. The Epidemiology of Childhood Cardiomyopathy in Australia. New England Journal of Medicine. 2003;348(17):1639-46.
88. Recent advancements in the molecular genetics of left ventricular noncompaction cardiomyopathy. Clinica Chimica Acta. 2017/02/01;465.
89. Bennett CE, Freudenberger R. The Current Approach to Diagnosis and Management of Left Ventricular Noncompaction Cardiomyopathy: Review of the Literature. Cardiology Research and Practice. 2016;2016:1-7.
90. Towbin JA, Jefferies JL. Cardiomyopathies Due to Left Ventricular Noncompaction, Mitochondrial and Storage Diseases, and Inborn Errors of Metabolism. Circulation Research. 2017;121(7):838-54.
91. Towbin JA, Beasley G. Left Ventricular Noncompaction and Vigorous Physical Activity. Journal of the American College of Cardiology. 2020;76(15):1734-6.
92. Hoedemaekers YM, Caliskan K, Michels M, Frohn-Mulder I, Van Der Smagt JJ, Phefferkorn JE, et al. The Importance of Genetic Counseling, DNA Diagnostics, and Cardiologic Family Screening in Left Ventricular Noncompaction Cardiomyopathy. Circulation: Cardiovascular Genetics. 2010;3(3):232-9.
93. Ichida F. Left ventricular noncompaction − Risk stratification and genetic consideration −. Journal of Cardiology. 2020;75(1):1-9.

94. Ogah OS, Iyawe EP, Orimolade OA, Okwunze K, Okeke M, Babatunde A, et al. Left ventricular noncompaction in Ibadan, Nigeria. The Egyptian Heart Journal. 2023;75(1).
95. Andrés AM, Nowick K. Editorial overview: Genetics of human evolution: The genetics of human origins. Current Opinion in Genetics & Development. 2014/12/01;29.
96. Founder Effect - an overview | ScienceDirect Topics.
97. Jain A, Sharma D, Bajaj A, Gupta V, Scaria V. Founder variants and population genomes—Toward precision medicine: Elsevier; 2021. 121-52 p.
98. Nei M, Maruyama T, Chakraborty R. THE BOTTLENECK EFFECT AND GENETIC VARIABILITY IN POPULATIONS. Evolution. 1975;29(1):1-10.
99. Slatkin M. A Population-Genetic Test of Founder Effects and Implications for Ashkenazi Jewish Diseases. The American Journal of Human Genetics. 2004;75(2):282-93.
100. Biology O. Biological Processes and Systems: Founder Effect. Texas Education Agency.
101. Brink PA, Schwartz PJ. Of founder populations, long QT syndrome, and destiny. Heart Rhythm. 2009;6(11):S25-S33.
102. Brink AJ, Torrington M. Progressive familial heart block - two types. South African medical journal. 1977;52(2):53-9.
103. Buchanan AH, Lester Kirchner H, Schwartz MLB, Kelly MA, Schmidlen T, Jones LK, et al. Clinical outcomes of a genomic screening program for actionable genetic conditions. Genetics in Medicine. 2020;22(11):1874-82.
104. Geldenhuys G, Glanzmann B, Lombard D, Boolay S, Carr J, Bardien S. Identification of a common founder couple for 40 South African Afrikaner families with Parkinson's disease. S Afr Med J. 2014;104(6):413-9.
105. Levy-Lahad E, Catane R, Halle D, Eisenberg S, Kaufman B, Hornreich G, et al. Founder BRCA1 and BRCA2 mutations in Ashkenazi Jews in Israel: Frequency and differential penetrance in ovarian cancer and in breast-ovarian cancer families. American journal of human genetics. 1997;60(5):1059-67.
106. Frisch A, Colombo R, Michaelovsky E, Karpati M, Goldman B, Peleg L. Origin and spread of the 1278insTATC mutation causing Tay-Sachs disease in Ashkenazi Jews: genetic drift as a robust and parsimonious hypothesis. Human Genetics. 2004;114(4):366-76.
107. Liebman SW, Palaganas H, Kobany H. A founder mutation in FLNC is likely a major cause of idiopathic dilated cardiomyopathy in Ashkenazi Jews. International journal of cardiology. 2021;323:124.
108. Tonin PN, Mes-Masson A-M, Futreal PA, Morgan K, Mahon M, Foulkes WD, et al. Founder BRCA1 and BRCA2 Mutations in French Canadian Breast and Ovarian Cancer Families. The American Journal of Human Genetics. 1998;63(5):1341-51.
109. Chami N, Tadros R, Lemarbre F, Lo KS, Beaudoin M, Robb L, et al. Nonsense Mutations in BAG3 are Associated With Early-Onset Dilated Cardiomyopathy in French Canadians. The Canadian journal of cardiology. 2014;30(12):1655-61.
110. Bos TA, Piers SRD, Wessels MW, Houweling AC, Hoedemakers YM, Bosman LP, et al. Arrhythmogenic cardiomyopathy phenotype associated with the pathogenic founder variant c.1211dup in PKP2. European heart journal. 2023;44(Supplement_2).
111. Gerull B, Kirchner F, Chong JX, Tagoe J, Chandrasekharan K, Strohm O, et al. Homozygous Founder Mutation in Desmocollin-2 (DSC2) Causes Arrhythmogenic Cardiomyopathy in the Hutterite Population. Circulation: Cardiovascular Genetics. 2013;6(4):327-36.

112. van der Zwaag PA, van Rijsingen IAW, de Ruiter R, Nannenberg EA, Groeneweg JA, Post JG, et al. Recurrent and founder mutations in the Netherlands—Phospholamban p.Arg14del mutation causes arrhythmogenic cardiomyopathy. Netherlands heart journal. 2013;21(6):286-93.
113. Murakami H, Tanimoto Y, Tanimoto K, Inoue S, Ishikawa T, Makita N, et al. Arrhythmogenic right ventricular cardiomyopathy in a Japanese patient with a homozygous founder variant of DSG2 in the East Asian population. Human Genome Variation. 2022;9(1).
114. Robles-Mezcua A, Ruíz-Salas A, Medina-Palomo C, Robles-Mezcua M, Díaz-Expósito A, Ortega-Jiménez MV, et al. The Novel Variant NP_00454563.2 (p.Glu259Glyfs*77) in Gene PKP2 Associated with Arrhythmogenic Cardiomyopathy in 8 Families from Malaga, Spain. Genes. 2023;14(7):1468.
115. Van Spaendonck-Zwarts KY, Van Der Kooi AJ, Van Den Berg MP, Ippel EF, Boven LG, Yee W-C, et al. Recurrent and founder mutations in the Netherlands: the cardiac phenotype of DES founder mutations p.S13F and p.N342D. Netherlands Heart Journal. 2012;20(5):219-28.
116. Van Der Zwaag PA, Cox MGPJ, Van Der Werf C, Wiesfeld ACP, Jongbloed JDH, Dooijes D, et al. Recurrent and founder mutations in the Netherlands. Netherlands Heart Journal. 2010;18(12):583-91.
117. Marino TC, Maranda B, Leblanc J, Pratte A, Barabas M, Dupéré A, et al. Novel founder mutation in French-Canadian families with Naxos disease. Clinical Genetics. 2017/10/01;92(4).
118. Schoonvelde SAC, Ruijmbeek CWB, Hirsch A, van Slegtenhorst MA, Wessels MW, von der Thüsen JH, et al. Phenotypic variability of filamin C–related cardiomyopathy: Insights from a novel Dutch founder variant. Heart rhythm. 2023;20(11):1512-21.
119. Scibetta N, Sampson BA, Tang Y, Gelb BD, Nicholas Scibetta BAS, Yingying Tang, Bruce D. Gelb. Sudden Death in Pediatric Patient With Dilated Cardiomyopathy Due to Founder Variant in NKX2-5: Case Report. Academic Forensic Pathology. 2024-07-21;14(3).
120. Treiber G, Flaus Furmaniuk A, Guilleux A, Medjane S, Bonfanti O, Schneebeli S, et al. A recurrent familial partial lipodystrophy due to a monoallelic or biallelic LMNA founder variant highlights the multifaceted cardiac manifestations of metabolic laminopathies. European Journal of Endocrinology. 2021/10/01;185(4).
121. Van Der Linde IHM, Hiemstra YL, Bökenkamp R, Van Mil AM, Breuning MH, Ruivenkamp C, et al. A Dutch MYH7 founder mutation, p.(Asn1918Lys), is associated with early onset cardiomyopathy and congenital heart defects. Netherlands Heart Journal. 2017;25(12):675-81.
122. Hoorntje ET, Van Spaendonck-Zwarts KY, Te Rijdt WP, Boven L, Vink A, Van Der Smagt JJ, et al. The first titin (c.59926 + 1G > A) founder mutation associated with dilated cardiomyopathy. European Journal of Heart Failure. 2018;20(4):803-6.
123. Otten E, Lekanne dit Deprez RH, Weiss MM, van Slegtenhorst M, Joosten M, van der Smagt JJ, et al. Recurrent and founder mutations in the Netherlands: mutation p.K217del in troponin T2, causing dilated cardiomyopathy. Netherlands Heart Journal 2010 18:10. 2010/10;18(10).
124. Fish M, Shaboodien G, Kraus S, Sliwa K, Seidman CE, Burke MA, et al. Mutation analysis of the phospholamban gene in 315 South Africans with dilated, hypertrophic, peripartum and arrhythmogenic right ventricular cardiomyopathies. Scientific Reports. 2016;6(1):22235.
125. Al-Shafai KN, Al-Hashemi M, Manickam C, Musa R, Selvaraj S, Syed N, et al. Genetic evaluation of cardiomyopathies in Qatar identifies enrichment of pathogenic

sarcomere gene variants and possible founder disease mutations in the Arabs. Molecular Genetics & Genomic Medicine. 2021;9(7).
126. Adalsteinsdottir B, Teekakirikul P, Maron BJ, Burke MA, Gudbjartsson DF, Holm H, et al. Nationwide Study on Hypertrophic Cardiomyopathy in Iceland. Circulation. 2014;130(14):1158-67.
127. Alders M, Jongbloed R, Deelen W, van den Wijngaard A, Doevendans P, Ten Cate F, et al. The 2373insG mutation in the MYBPC3 gene is a founder mutation, which accounts for nearly one-fourth of the HCM cases in the Netherlands. European heart journal. 2003;24(20):1848-53.
128. Méndez I, Fernández AI, Espinosa MÁ, Cuenca S, Lorca R, Rodríguez JF, et al. Founder mutation in myosin-binding protein C with an early onset and a high penetrance in males. Open Heart. 2021;8(2):e001789.
129. Marsili L, Van Lint FHM, Russo F, Van Spaendonck-Zwarts KY, Ader F, Bichon M-L, et al. MYH7 p.(Arg1712Gln) is pathogenic founder variant causing hypertrophic cardiomyopathy with overall relatively delayed onset. Netherlands Heart Journal. 2023;31(7-8):300-7.
130. Vodnjov N, Toplišek J, Maver A, Čuturilo G, Jaklič H, Teran N, et al. A novel splice-site FHOD3 founder variant is a common cause of hypertrophic cardiomyopathy in the population of the Balkans–A cohort study. PLOS ONE. 2023;18(12):e0294969.
131. Perik M, Govaerts E, Laga S, Goovaerts I, Saenen J, Van Craenenbroeck E, et al. Variable clinical expression of a Belgian TGFB3 founder variant suggests the presence of a genetic modifier. Frontiers in Genetics. 2023;14.
132. Vodnjov N, Maver A, Teran N, Peterlin B, Toplišek J, Writzl K. Clinical Outcome of Hypertrophic Cardiomyopathy in Probands with the Founder Variant c.913_914del in MYBPC3: A Slovenian Cohort Study. Journal of Cardiovascular Translational Research. 2024.
133. Redin C, Pavlidou DC, Bhuiyan Z, Porretta AP, Monney P, Bedoni N, et al. The Amish NM_000256.3:c.3330+2T>G splice variant in MYBPC3 associated with hypertrophic cardiomyopathy is an ancient Swiss mutation. European Journal of Medical Genetics. 2022/12/01;65(12).
134. Van Den Wijngaard A, Volders P, Van Tintelen JP, Jongbloed JDH, Van Den Berg MP, Lekanne Deprez RH, et al. Recurrent and founder mutations in the Netherlands: cardiac Troponin I (TNNI3) gene mutations as a cause of severe forms of hypertrophic and restrictive cardiomyopathy. Netherlands Heart Journal. 2011;19(7-8):344-51.
135. Jääskeläinen P, Heliö T, Aalto-Setälä K, Kaartinen M, Ilveskoski E, Hämäläinen L, et al. Two founder mutations in the alpha-tropomyosin and the cardiac myosin-binding protein C genes are common causes of hypertrophic cardiomyopathy in the Finnish population. Annals of Medicine. 2013;45(1):85-90.
136. Mastana SS. Unity in diversity: an overview of the genomic anthropology of India. Annals of human biology. 2014;41(4).
137. Machha P, Gopalan A, Elangovan Y, Chandra Mouli Veeravalli S, Tej Sowpati D, Thangaraj K. Endogamy and high prevalence of deleterious mutations in India: evidence from strong founder events. Journal of Genetics and Genomics. 2025;52(4).
138. South African History Online [Internet]2011/10/17. Available from: https://www.sahistory.org.za/people/jan-van-riebeeck.
139. Guelke L. Freehold Farmers and Frontier Settlers, 1657–1780 | 10 | Historiograph. London: Routledge; 2019.

140. Swanson MW. South African Revision - The Making of Modern South Africa: Conquest, Segregation and Apartheid. By Nigel Worden. Oxford and Cambridge MA: Blackwell Publishers, 1994. The Journal of African History. Cambridge Core. The Journal of African History. 1994;36(1):573-4.
141. Saunders C, Southey N. A dictionary of South African history. 2nd ed. Cape Town: David Philip; 2001.
142. Burrows EH. A history of medicine in South Africa up to the end of the nineteenth century. Cape Town: A.A. Balkema; 1958.
143. Elphick R. en R. Shell. 1989. Intergroup relations: Khoikhoi, settlers, slaves and free blacks, 1652–1795. Elphick en Giliomee (reds). 1989.
144. Bruin Kd. From Viticulture to Commemoration: French Huguenot Memory in the Cape Colony (1688-1824). Journal of the Western Society for French History. 2021;47:21.
145. Van Niekerk JP. The first British occupation of the Cape of Good Hope and two prize cases on joint capture in the High Court of Admiralty. Fundamina : a journal of legal history. 2005;2005(11):155-82.
146. Le Saux O, Beck K, Sachsinger C, Treiber C, Göring H, Curry K, et al. Evidence for a founder effect for pseudoxanthoma elasticum in the Afrikaner population of South Africa. Human Genetics. 2002;111(4-5):331-8.
147. Torrington M, Louw J. Familial polyposis of the colon-a condition illustrating the importance of genealogy in medicine. Familia. 1968;5(2):42-7.
148. Dean G. The porphyrias: a story of inheritance and environment: Pitman Medical; 1971.
149. Hayden MR, Hopkins HC, Macrea M, Beighton PH. The origin of Huntington's chorea in the Afrikaner population of South Africa. South African medical journal. 1980;58(5):197-200.
150. Knoll D, de Vries W, De Wet W. Genealogie van'n familie met osteogenesis imperfecta: Die stamregister van Pieter Willem Adriaan Piek. Familia. 1988;25:46-64.
151. Karayiorgou M, Torrington M, Abecasis GR, Pretorius H, Robertson B, Kaliski S, et al. Phenotypic characterization and genealogical tracing in an Afrikaner schizophrenia database. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics. 2004/01/01;124B(1).
152. Tipping AJ, Pearson T, Morgan NV, Gibson RA, Kuyt LP, Havenga C, et al. Molecular and genealogical evidence for a founder effect in Fanconi anemia families of the Afrikaner population of South Africa. Proceedings of the National Academy of Sciences. 2001;98(10):5734-9.
153. Engelbrecht H-R, Dalvie S, Agenbag G, Stein DJ, Ramesar RS. Whole-exome sequencing in an Afrikaner family with bipolar disorder. Journal of affective disorders. 2020;276:69-75.
154. Warnich L, Kotze MJ, Groenewald IM, Groenewald JZ, van Brakel MG, van Heerden CJ, et al. Identification of Three Mutations and Associated Haplotypes in the Protoporphyrinogen Oxidase Gene in South African Families with Variegate Porphyria. Human Molecular Genetics. 1996;5(7).
155. Scholefield J, Greenberg J. A common SNP haplotype provides molecular proof of a founder effect of Huntington disease linking two South African populations. European Journal of Human Genetics. 2007;15(5):590-5.
156. Heese JA. Die herkoms van die Afrikaner 1657-1867. Kaapstad: Balkema; 1971.
157. Machipisa T. Preliminary genealogical evidence for the Plakophilin-2 gene, PKP2 c.1162C>T founder mutation in cases with Arrhythmogenic Right Ventricular

Cardiomyopathy (ARVC). http://hdl.handle.net/11427/23420: University of Cape Town; 2016.
158. Templeton N. The polymerase chain reaction. History, methods, and applications. Diagnostic molecular pathology : The American Journal of Surgical Pathology, part B. 1992;1(1):58-72.
159. Bessmen M, Kornberg A, Lehman I, Simms E. Enzymic synthesis of deoxyribonucleic acid. Preparation of substrates and partial purification of an enzyme from Escherichia coli. The Journal of biological chemistry. 1958;21(1):197-8.
160. Mbele M. Molecular genetics of arrhythmogenic right ventricular and dilated cardiomyopathy in South Africans. http://hdl.handle.net/11427/18611: University of Cape Town; 2014.
161. Waters DLE, Shapter FM. The Polymerase Chain Reaction (PCR): General Methods. Methods in Molecular Biology2014. p. 65-75.
162. Erali M, Wittwer CT. High resolution melting analysis for gene scanning. Methods. 2010;50(4):250-61.
163. Crossley BM, Bai J, Glaser A, Maes R, Porter E, Killian ML, et al. Guidelines for Sanger sequencing and molecular assay monitoring. Journal of Veterinary Diagnostic Investigation. 2020;32(6):767-75.
164. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences. 1977;74(12):5463-7.
165. Sanger F, Coulson A. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. Journal of molecular biology. 1975;94(3).
166. Schlötterer C. Evolutionary dynamics of microsatellite DNA. Chromosoma. 2000;109(6):365-71.
167. Nassar LR, Barber GP, Benet-Pagès A, Casper J, Clawson H, Diekhans M, et al. The UCSC Genome Browser database: 2023 update. Nucleic Acids Research. 2023;51(D1):D1188-D95.
168. Owczarzy R, Tataurov AV, Wu Y, Manthey JA, Mcquisten KA, Almabrazi HG, et al. IDT SciTools: a suite for analysis and design of nucleic acid oligomers. Nucleic Acids Research. 2008;36(Web Server):W163-W9.
169. Merwe Jvd, editor ARCHIVES FOR THE PROMOTION OF HUMAN RIGHTS, TRANSPARENCY AND GOOD GOVERNANCE. 2016 National Archives Week; 2016 9-13 May 2016; National Archives.
170. Costa S, Cerrone M, Saguner AM, Brunckhorst C, Delmar M, Duru F. Arrhythmogenic cardiomyopathy: An in-depth look at molecular mechanisms and clinical correlates. Trends in Cardiovascular Medicine. 2021/10/01;31(7).
171. Delmar M, Mckenna WJ. The Cardiac Desmosome and Arrhythmogenic Cardiomyopathies. Circulation Research. 2010;107(6):700-14.
172. Agullo-Pascual E, Cerrone M, Delmar M. Arrhythmogenic cardiomyopathy and Brugada syndrome: Diseases of the connexome. FEBS Letters. 2014;588(8):1322-30.
173. Sato PY, Musa H, Coombs W, Guerrero-Serna G, PatiñO GA, Taffet SM, et al. Loss of Plakophilin-2 Expression Leads to Decreased Sodium Current and Slower Conduction Velocity in Cultured Cardiac Myocytes. Circulation Research. 2009;105(6):523-6.
174. Cerrone M, Montnach J, Lin X, Zhao Y-T, Zhang M, Agullo-Pascual E, et al. Plakophilin-2 is required for transcription of genes that control calcium cycling and cardiac rhythm. Nature Communications. 2017;8(1).
175. Kim J-C, Pérez-Hernández M, Alvarado FJ, Maurya SR, Montnach J, Yin Y, et al. Disruption of Ca2+i Homeostasis and Connexin 43 Hemichannel Function in the

Right Ventricle Precedes Overt Arrhythmogenic Cardiomyopathy in Plakophilin-2–Deficient Mice. Circulation. 2019;140(12):1015-30.
176. Kaplan SR, Gard JJ, Protonotarios N, Tsatsopoulou A, Spiliopoulou C, Anastasakis A, et al. Remodeling of myocyte gap junctions in arrhythmogenic right ventricular cardiomyopathy due to a deletion in plakoglobin (Naxos disease). Heart Rhythm. 2004/05/01;1(1).
177. Gambardella J, Sorriento D, Ciccarelli M, Del Giudice C, Fiordelisi A, Napolitano L, et al. Functional Role of Mitochondria in Arrhythmogenesis. Advances in Experimental Medicine and Biology. 2017.
178. 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. Circulation Research. 2014-January-31;114(3).
179. Sen-Chowdhry S, Morgan RD, Chambers JC, McKenna WJ. Arrhythmogenic Cardiomyopathy: Etiology, Diagnosis, and Treatment. Annual Review of Medicine. 2010;61:233-53.
180. Crawford DC, Bhangale T, Li N, Hellenthal G, Rieder MJ, Nickerson DA, et al. Evidence for substantial fine-scale variation in recombination rates across the human genome. Nature Genetics. 2004;36(7):700-6.
181. Alper CA, Larsen CE. Pedigree-Defined Haplotypes and Their Applications to Genetic Studies. Methods in Molecular Biology: Springer New York; 2017. p. 113-27.
182. Roach JC, Glusman G, Smit AFA, Huff CD, Hubley R, Shannon PT, et al. Analysis of Genetic Inheritance in a Family Quartet by Whole-Genome Sequencing. Science. 2010;328(5978):636-9.
183. Levine BD. Can Intensive Exercise Harm the Heart?: The Benefits of Competitive Endurance Training for Cardiovascular Structure and Function. Circulation (New York, NY). 2014;130(12):987-91.
184. James CA, Bhonsale A, Tichnell C, Murray B, Russell SD, Tandri H, et al. Exercise Increases Age-Related Penetrance and Arrhythmic Risk in Arrhythmogenic Right Ventricular Dysplasia/Cardiomyopathy–Associated Desmosomal Mutation Carriers. Journal of the American College of Cardiology. 2013;62(14):1290-7.
185. Xu T, Yang Z, Vatta M, Rampazzo A, Beffagna G, Pillichou K, et al. Compound and Digenic Heterozygosity Contributes to Arrhythmogenic Right Ventricular Cardiomyopathy. Journal of the American College of Cardiology. 2010;55(6):587-97.
186. Rigato I, Bauce B, Rampazzo A, Zorzi A, Pilichou K, Mazzotti E, et al. Compound and Digenic Heterozygosity Predicts Lifetime Arrhythmic Outcome and Sudden Cardiac Death in Desmosomal Gene–Related Arrhythmogenic Right Ventricular Cardiomyopathy. Circulation: Cardiovascular Genetics. 2013;6(6):533-42.
187. Marcus FI, Edson S, Towbin JA. Genetics of Arrhythmogenic Right Ventricular Cardiomyopathy. Journal of the American College of Cardiology. 2013;61(19):1945-8.
188. Gasperetti A, James CA, Carrick RT, Protonotarios A, Te Riele ASJM, Cadrin-Tourigny J, et al. Arrhythmic risk stratification in arrhythmogenic right ventricular cardiomyopathy. Europace. 2023;25(11).
189. Nugent P. Children of Bondage: A Social History of the Slave Society at the Cape of Good Hope, 1652–1838. Blackwell Publishers; 1996. p. 387-8.
190. Haskell WL, Lee IM, Pate RR, Powell KE, Blair SN, Franklin BA, et al. Physical activity and public health: Updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. Circulation. 2007;116(9):1081-93.

191. Cilliers J, Mariotti M. The shaping of a settler fertility transition: eighteenth- and nineteenth-century South African demographic history reconsidered. European Review of Economic History. 2018.
192. Gandolfo LC, Bahlo M, Speed TP. Dating Rare Mutations from Small Samples with Dense Marker Data. Genetics. 2014;197(4):1315-27.
193. Domínguez F, Cuenca S, Bilińska Z, Toro R, Villard E, Barriales-Villa R, et al. Dilated Cardiomyopathy Due to BLC2-Associated Athanogene 3 (BAG3) Mutations. Journal of the American College of Cardiology. 2018;72(20):2471-81.
194. Qu HQ, Feldman AM, Hakonarson H. Genetics of BAG3: A Paradigm for Developing Precision Therapies for Dilated Cardiomyopathies. Journal of the American Heart Association. 2022;11(23).
195. Behl C. Breaking BAG: The Co-Chaperone BAG3 in Health and Disease. Trends in Pharmacological Sciences. 2016;37(8):672-88.
196. Stürner E, Behl C. The Role of the Multifunctional BAG3 Protein in Cellular Protein Quality Control and in Disease. Frontiers in Molecular Neuroscience. 2017;10.
197. Kögel D, Linder B, Brunschweiger A, Chines S, Behl C. At the Crossroads of Apoptosis and Autophagy: Multiple Roles of the Co-Chaperone BAG3 in Stress and Therapy Resistance of Cancer. Cells. 2020;9(3):574.
198. Bers DM. Cardiac excitation–contraction coupling. Nature. 2002;415(6868):198-205.
199. Tahrir FG, Knezevic T, Gupta MK, Gordon J, Cheung JY, Feldman AM, et al. Evidence for the Role of BAG3 in Mitochondrial Quality Control in Cardiomyocytes. Journal of Cellular Physiology. 2017;232(4):797-805.
200. Norton N, Li D, Rieder J, Mark, Siegfried D, Jill, Rampersaud E, Züchner S, et al. Genome-wide Studies of Copy Number Variation and Exome Sequencing Identify Rare Variants in BAG3 as a Cause of Dilated Cardiomyopathy. The American Journal of Human Genetics. 2011;88(3):273-82.
201. Knezevic T, Myers VD, Gordon J, Tilley DG, Sharp TE, Wang J, et al. BAG3: a new player in the heart failure paradigm. Heart Failure Reviews. 2015;20(4):423-34.
202. Kamuli SN. The molecular genetics of familial cardiomyopathy. http://hdl.handle.net/11427/37985: University of Cape Town; 2023.
203. Drazner MH. The Progression of Hypertensive Heart Disease. Circulation. 2011;123(3):327-34.
204. Tong X, Shen L, Zhou X, Wang Y, Chang S, Lu S. Comparative Efficacy of Different Drugs for the Treatment of Dilated Cardiomyopathy: A Systematic Review and Network Meta-analysis. Drugs in R&D. 2023;23(3):197-210.
205. Tayal U, Ware JS, Lakdawala NK, Heymans S, Prasad SK. Understanding the genetics of adult-onset dilated cardiomyopathy: what a clinician needs to know. European Heart Journal. 2021;42(24):2384-96.
206. Henige D. A population history of North America. [Review of: Haines, M.R. and Steckel, R.H., ed. A population history of North America. New York: Cambridge U. Pr., 2000]. The Journal of interdisciplinary history. 2002;32(3).
207. Bouchard JP, Barbeau A, Bouchard R, Bouchard RW. Autosomal Recessive Spastic Ataxia of Charlevoix-Saguenay. Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques. 1978;5(1):61-9.
208. Gauvreau C, Brisson J-D, Dupré N. Hereditary Motor and Sensory Neuropathy with Agenesis of the Corpus Callosum: University of Washington, Seattle, Seattle (WA); 1993 1993.

209. Debray F-G, Morin C, Janvier A, Villeneuve J, Maranda B, Laframboise R, et al. LRPPRC mutations cause a phenotypically distinct form of Leigh syndrome with cytochrome c oxidase deficiency. Journal of Medical Genetics. 2011-03-01;48(3).
210. Rozen R, Braekeleer MD, Daigneault J, Ferreira-Rajabi L, Gerdes M, Lamoureux L, et al. Cystic fibrosis mutations in French Canadians: Three CFTR mutations are relatively frequent in a Quebec population with an elevated incidence of cystic fibrosis. American Journal of Medical Genetics. 1992/02/01;42(3).
211. Grompe M, St-Louis M, Demers S, al-Dhalimy M, Leclerc B, Tanguay R. A single mutation of the fumarylacetoacetate hydrolase gene in French Canadians with hereditary tyrosinemia type I. The New England journal of medicine. 1994;331(6).
212. Plante M, Claveau S, Lepage P, Lavoie È-M, Brunet S, Roquis D, et al. Mucolipidosis II: a single causal mutation in the N-acetylglucosamine-1-phosphotransferase gene (GNPTAB) in a French Canadian founder population. Clinical Genetics. 2008/03/01;73(3).
213. Scriver CR. VITAMIN D DEPENDENCY. Pediatrics. 1970/03/01;45(3).
214. Normand T, Bergeron J, Fernandez-Margallo T, Bharucha A, Ven Murthy MR, Julien P, et al. Geographic distribution and genealogy of mutation 207 of the lipoprotein lipase gene in the French Canadian population of Québec. Human Genetics 1992 89:6. 1992/08;89(6).
215. Vohl M, Moorjani S, Roy M, Gaudet D, Torres A, Minnich A, et al. Geographic distribution of French-Canadian low-density lipoprotein receptor gene mutations in the Province of Quebec. Clinical genetics. 1997;52(1).
216. Yotova V, Labuda D, Zietkiewicz E, Gehl D, Lovell A, Lefebvre J, et al. Anatomy of a founder effect: myotonic dystrophy in Northeastern Quebec. Human genetics. 2005;117.
217. De Braekeleer M. Hereditary disorders in Saguenay-Lac-St-Jean (Quebec, Canada). Human heredity. 1991;41(3).
218. Levesque S, Morin C, Guay S-P, Villeneuve J, Marquis P, Yik WY, et al. A founder mutation in the PEX6 gene is responsible for increased incidence of Zellweger syndrome in a French Canadian population. BMC Medical Genetics 2012 13:1. 2012-08-15;13(1).
219. Bchetnia M, Tremblay M-L, Leclerc G, Dupérée A, Powell J, McCuaig C, et al. Expression signature of epidermolysis bullosa simplex. Human Genetics 2011 131:3. 2011-08-30;131(3).
220. Villard E, Perret C, Gary F, Proust C, Dilanian G, Hengstenberg C, et al. A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. European heart journal. 2011;32(9).
221. Mooney GH, McIntyre DE. South Africa: a 21st century apartheid in health and health care? The Medical Journal of Australia. 2008/12/01;189(11).
222. Nishiuchi S, Makiyama T, Aiba T, Nakajima K, Hirose S, Kohjitani H, et al. Gene-Based Risk Stratification for Cardiac Disorders in LMNA Mutation Carriers. Circulation: Cardiovascular Genetics. 2017;10(6).
223. Dechat T, Pfleghaar K, Sengupta K, Shimi T, Shumaker DK, Solimando L, et al. Nuclear lamins: major factors in the structural organization and function of the nucleus and chromatin. Genes & Development. 2008;22(7):832-53.
224. Shimi T, Pfleghaar K, Kojima S-I, Pack C-G, Solovei I, Goldman AE, et al. The A- and B-type nuclear lamin networks: microdomains involved in chromatin organization and transcription. Genes & Development. 2008;22(24):3409-21.
225. Carmosino M, Torretta S, Procino G, Gerbino A, Forleo C, Favale S, et al. Role of nuclear Lamin A/C in cardiomyocyte functions. Biology of the Cell. 2014;106(10):346-58.

226. Veltrop RJA, Kukk MM, Topouzidou K, Didden L, Muchir A, Van Steenbeek FG, et al. From gene to mechanics: a comprehensive insight into the mechanobiology of LMNA mutations in cardiomyopathy. Cell Communication and Signaling. 2024;22(1).
227. Cattin M-E, Muchir A, Bonne G. 'State-of-the-heart' of cardiac laminopathies. Current Opinion in Cardiology. May 2013;28(3).
228. Pessente GDA, Sacilotto L, Calil ZO, Olivetti NQS, Wulkan F, Oliveira TGMD, et al. Effect of Occurrence of Lamin A/C (LMNA) Genetic Variants in a Cohort of 101 Consecutive Apparent "Lone AF" Patients: Results and Insights. Frontiers in Cardiovascular Medicine. 2022;9.
229. Gonzalo S, Kreienkamp R, Askjaer P. Hutchinson-Gilford Progeria Syndrome: A premature aging disease caused by LMNA gene mutations. Ageing Research Reviews. 2017;33:18-29.
230. Yukina M, Nuralieva N, Sorkina E, Troshina E, Tiulpakov A, Belaya Z, et al. Atypical progeroid syndrome (p.E262K LMNA mutation): a rare cause of short stature and osteoporosis. Endocrinology, Diabetes & Metabolism Case Reports. 2021;2021.
231. Fernandez-Pombo A, Diaz-Lopez EJ, Castro AI, Sanchez-Iglesias S, Cobelo-Gomez S, Prado-Moraña T, et al. Clinical Spectrum of LMNA-Associated Type 2 Familial Partial Lipodystrophy: A Systematic Review. Cells. 2023;12(5):725.
232. Parks SB, Kushner JD, Nauman D, Burgess D, Ludwigsen S, Peterson A, et al. Lamin A/C mutation analysis in a cohort of 324 unrelated patients with idiopathic or familial dilated cardiomyopathy. American Heart Journal. 2008;156(1):161-9.
233. Haas J, Frese KS, Peil B, Kloos W, Keller A, Nietsch R, et al. Atlas of the clinical genetics of human dilated cardiomyopathy. European Heart Journal. 2015;36(18):1123-35.
234. Taylor MRG, Fain PR, Sinagra G, Robinson ML, Robertson AD, Carniel E, et al. Natural history of dilated cardiomyopathy due to lamin A/C gene mutations. Journal of the American College of Cardiology. 2003;41(5):771-80.
235. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, Mcdonnell SK, Baheti S, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. The American Journal of Human Genetics. 2016;99(4):877-85.
236. Banerjee A, Rathee V, Krishnaswamy R, Bhattacharjee P, Ray P, Sood AK, et al. Viscoelastic Behavior of Human Lamin A Proteins in the Context of Dilated Cardiomyopathy. PLoS ONE. 2013;8(12):e83410.
237. Pritha Bhattacharjee AB, Amrita Banerjee, Dipak Dasgupta, and Kaushik Sengupta. Structural Alterations of Lamin A Protein in Dilated Cardiomyopathy: American Chemical Society; 2013.
238. Hasselberg NE, Haland TF, Saberniak J, Brekke PH, Berge KE, Leren TP, et al. Lamin A/C cardiomyopathy: young onset, high penetrance, and frequent need for heart transplantation. European Heart Journal. 2018;39(10):853-60.
239. Myers MC, Wang S, Zhong Y, Maruyama S, Bueno C, Bastien A, et al. Prevalence of Genetically Associated Dilated Cardiomyopathy: A Systematic Literature Review and Meta-Analysis. Cardiology Research. 2024;15(4):233-45.
240. Sweet ME, Taylor MR, Mestroni L. Diagnosis, prevalence, and screening of familial dilated cardiomyopathy. Expert Opinion on Orphan Drugs. 2015;3(8):869-76.
241. Herman DS, Lam L, Taylor MRG, Wang L, Teekakirikul P, Christodoulou D, et al. Truncations of Titin Causing Dilated Cardiomyopathy. New England Journal of Medicine. 2012;366(7):619-28.
242. Lewinter MM, Granzier HL. Titin Is a Major Human Disease Gene. Circulation. 2013;127(8):938-44.

243. Jolfayi AG, Kohansal E, Ghasemi S, Naderi N, Hesami M, Mozafarybazargany M, et al. Exploring TTN variants as genetic insights into cardiomyopathy pathogenesis and potential emerging clues to molecular mechanisms in cardiomyopathies. Scientific Reports. 2024;14(1).
244. Chauveau C, Rowell J, Ferreiro A. A Rising Titan: TTN Review and Mutation Update. Human Mutation. 2014;35(9):1046-59.
245. Merlo M SG, Carniel E, Slavov D, Zhu X, Barbati G, Spezzacatene A, Ramani F, Salcedo E, Di Lenarda A, Mestroni L, Taylor MR. Poor prognosis of rare sarcomeric gene variants in patients with dilated cardiomyopathy. Clinical and translational science. 2013 Dec;6(6).
246. Tharp CA, Haywood ME, Sbaizero O, Taylor MRG, Mestroni L. The Giant Protein Titin's Role in Cardiomyopathy: Genetic, Transcriptional, and Post-translational Modifications of TTN and Their Contribution to Cardiac Disease. Frontiers in Physiology. 2019;10.
247. Lacaze P, Sebra R, Riaz M, Ingles J, Tiller J, Thompson BA, et al. Genetic variants associated with inherited cardiovascular disorders among 13,131 asymptomatic older adults of European descent. Genomic Medicine. 2021;6(1).
248. Ntusi NAB, Sliwa K. Impact of Racial and Ethnic Disparities on Patients With Dilated Cardiomyopathy. Journal of the American College of Cardiology. 2021;78(25):2580-8.
249. Felkin LE, Walsh R, Ware JS, Yacoub MH, Birks EJ, Barton PJR, et al. Recovery of Cardiac Function in Cardiomyopathy Caused by Titin Truncation. JAMA cardiology. 2016 May 1;1(2).
250. Santiago CF, Huttner IG, Fatkin D. Titin-related Cardiomyopathy: Is it a Distinct Disease? Current Cardiology Reports. 2022;24(9):1069-75.
251. Ware JS, Cook SA. Role of titin in cardiomyopathy: from DNA variants to patient stratification. Nature Reviews Cardiology. 2018;15(4):241-52.
252. Savarese M, Vihola A, Oates EC, Barresi R, Fiorillo C, Tasca G, et al. Genotype–phenotype correlations in recessive titinopathies. Genetics in Medicine. 2020;22(12):2029-40.
253. Savarese M, Sarparanta J, Vihola A, Udd B, Hackman P. Increasing Role of Titin Mutations in Neuromuscular Disorders. Journal of Neuromuscular Diseases. 2016;3(3):293-308.
254. Perrin A, Van Goethem C, Thèze C, Puechberty J, Guignard T, Lecardonnel B, et al. Long-Reads Sequencing Strategy to Localize Variants in TTN Repeated Domains. The Journal of Molecular Diagnostics. 2022;24(7):719-26.
255. Putman AI, Carbone I. Challenges in analysis and interpretation of microsatellite data for population genetic studies. Ecology and Evolution. 2014;4(22):4399-428.
256. Oliver E, Oliver WH. The Colonisation of South Africa: A unique case. HTS Teologiese Studies / Theological Studies. 2017;73(3).
257. Giliomee H, Mbenga B, Nasson B. New history of South Africa. Second edition ed. Cape Town, South Africa: Tafelberg; 2022.
258. Schoeman K. Early slavery at the Cape of Good Hope, 1652-1717. 1st ed. Pretoria: Protea Book House; 2007.
259. Shell RC-H. Children of bondage : a social history of the slave society at the Cape of Good Hope, 1652-1838. Johannesburg: Witwatersrand University Press; 1994.
260. Groenewald G. Slaves and Free Blacks in VOC Cape Town, 1652–1795. History Compass. 2010/09/01;8(9).

261. Guma M. The Cultural Meaning of Names among Basotho of South Africa – DOAJ. Nordic Journal of African Studies. 2001/12/01;10(3).
262. Suzman SM. Names as pointers: Zulu personal naming practices. Language in society. 1994;23(2):253-72.
263. Makgopa M, Rugwiji TT. Naming Patterns among Northern Sotho, Xitsonga and Tshivenda Communities of SA: Religio-cultural Perspective. Studia Historiae Ecclesiasticae. 2022;47(3).
264. Frost L, Jenkins LS, Emmink B. Improving access to health care in a rural regional hospital in South Africa: Why do patients miss their appointments? African Journal of Primary Health Care & Family Medicine. 2017;9(1).

Appendices:

Appendix A

Table 1. Comparison of 2010 TF Criteria and 2020 International Criteria for Diagnosis of ARVC (37)

Category	2010 TF criteria	2020 International criteria
I. Global or regional dysfunction and structural alteration	*Major* *By 2D echocardiogram:* ● Regional RV akinesia, dyskinesia, or aneurysm *and* one of the following (end diastole) - PLAX RVOT ≥32 mm (corrected for body size [PLAX/BSA] ≥19 mm/m^2) - PSAX RVOT ≥36 mm (corrected for body size [PSAX/BSA] ≥21 mm/m^2) - Fractional area change ≤33% *By MRI:* ● Regional RV akinesia or dyskinesia or dyssynchronous RV contraction *and* one of the following: -	*Major* *By 2D echocardiogram, CMR, or angiography:* ● Regional RV akinesia, dyskinesia, or bulging *plus* 1 of the following: ● Global RV dilatation (increase of RV EDV according to the imaging test specific nomograms for age, sex, and BSA) *or* ● Global RV systolic dysfunction (reduction of RV EF according to the imaging test specific nomograms for age and sex) *Minor*

Category	2010 TF criteria	2020 International criteria
	Ratio of RV end-diastolic volume to BSA: ≥110 mL/m^2 (male) or ≥100 mL/m^2 (female) - *or* RV ejection fraction ≤40% *By RV angiography:* ● Regional RV akinesia, dyskinesia, or aneurysm *Minor* *By 2D echocardiogram:* ● Regional RV akinesia or dyskinesia *and* one of the following (end diastole): - PLAX RVOT ≥29–<32 mm; (corrected for body size [PLAX/BSA] ≥16–<19 mm/m^2) - PSAX RVOT ≥32–<36 mm; (corrected for body size [PSAX/BSA] ≥18–<21 mm/m^2) - or fractional area change >33%–≤40%	*By 2D echocardiogram, CMR, or angiography:* ● Regional RV akinesia, dyskinesia or aneurysm of RV free wall

Category	2010 TF criteria	2020 International criteria
	By MRI: ● Regional RV akinesia or dyskinesia or dyssynchronous RV contraction *and* one of the following: - Ratio of RV end-diastolic volume to BSA ≥100 to <110 mL/m^2 (male) or ≥90 to <100 mL/m^2 (female) - *or* RV ejection fraction >40% to ≤45%	
II. Tissue characterization	*Major* *By EMB* ● Residual myocytes <60% by morphometric analysis (or 50% if estimated), with fibrous replacement of the RV free wall myocardium in ≥1 sample, with or without fatty replacement of tissue on endomyocardial biopsy *Minor* *By EMB* ● Residual myocytes 60% to 75% by morphometric analysis (or 50–65% if	*Major* *By CE-CMR:* ● Transmural LGE (stria pattern) of ≥1 RV region(s) (inlet, outlet, and apex in 2 orthogonal views) *Major* *By EMB (limited indications):* ● Fibrous replacement of the myocardium in ≥1 sample, with or without fatty tissue

Category	2010 TF criteria	2020 International criteria
	estimated), with fibrous replacement of the RV free wall myocardium in ≥1 sample, with or without fatty replacement of tissue on endomyocardial biopsy	
III. Repolarization abnormalities	*Major* ● Inverted T waves in right precordial leads (V1, V_2, and V_3) or beyond in individuals >14 y of age (in the absence of complete RBBB QRS ≥120 ms) *Minor* ● Inverted T waves in leads V_1 and V_2 in individuals >14 y of age (in the absence of complete RBBB) or in V4, V5, or V6 ● Inverted T waves in V_1, V_2, V_3, and V_4 in individuals >14 y of age in the presence of complete RBBB	*Major* ● Inverted T waves in right precordial leads (V_1,V_2, and V_3) or beyond in individuals with complete pubertal development (in the absence of complete RBBB) *Minor* ● Inverted T waves in leads V1 and V2 in individuals with completed pubertal development (in the absence of complete RBBB) ● Inverted T waves in V1, V2, V3 and V4 in individuals with completed pubertal development in the presence of complete RBBB

Category	2010 TF criteria	2020 International criteria
IV. Depolarization and conduction abnormalities	*Major* ● Epsilon wave (reproducible low-amplitude signals between end of QRS complex to onset of the T wave) in the right precordial leads (V1 to V3) *Minor* ● Late potentials by SAECG in ≥1 of 3 parameters in the absence of a QRS duration of ≥110 ms on the standard ECG - Filtered QRS duration (fQRS) ≥114 ms - Duration of terminal QRS <40 μV (low-amplitude signal duration) ≥38 ms - Root-mean-square voltage of terminal 40 ms ≤20 μV ● Terminal activation duration of QRS ≥55 ms measured from the nadir of the S wave to the end of the QRS, including R', in V1, V2, or V3, in the	*Minor* ● Epsilon wave (reproducible low-amplitude signals between end of QRS complex to onset of the T wave) in the right precordial leads (V1 to V3) ● Terminal activation duration of QRS ≥55 ms measured from the nadir of the S wave to the end of the QRS, including R', in V1, V2, or V3 (in the absence of complete RBBB)

Category	2010 TF criteria	2020 International criteria
	absence of complete right bundle-branch block	
V. Arrhythmias	*Major* ● Nonsustained or sustained ventricular tachycardia of left bundle-branch block morphology with superior axis (negative or indeterminate QRS in leads II, III, and aVF and positive in lead aVL) *Minor* ● Nonsustained or sustained ventricular tachycardia of RV outflow configuration, left bundle-branch block morphology with inferior axis (positive QRS in leads II, III, and aVF and negative in lead aVL) or of unknown axis ● >500 ventricular extrasystoles per 24 h (Holter)	*Major* ● Frequent ventricular extrasystoles (>500 per 24 h), non-sustained or sustained ventricular tachycardia of LBBB morphology* *Minor* ● Frequent ventricular extrasystoles (>500 per 24 h), non-sustained or sustained ventricular tachycardia of LBBB morphology with inferior axis (“RVOT pattern”)
VI. Family history/genetics	*Major* ● ACM confirmed in a first-degree relative who meets diagnostic criteria ●	

Category	2010 TF criteria	2020 International criteria
	ACM confirmed pathologically at autopsy or surgery in a first-degree relative ● Identification of a pathogenic or likely pathogenetic ACM mutation in the patient under evaluation *Minor* ● History of ACM in a first-degree relative in whom it is not possible or practical to determine whether the family member meets diagnostic criteria ● Premature sudden death (<35 y of age) due to suspected ACM in a first-degree relative ● ACM confirmed pathologically or by diagnostic criteria in second-degree relative	

Cut-off values of EDV and EF of the 2020 International criteria for RV dilatation and systolic dysfunction, respectively, are reported in Table 3. ACM indicates arrhythmogenic cardiomyopathy; BSA, body surface area; CMR, cardiac magnetic resonance; EDV, end diastolic volume; EF, ejection fraction; EMB, endomyocardial biopsy; ITF, International Task Force; LBBB, left bundle-branch block; LGE, late gadolinium enhancement; LV, left ventricle; MRI, magnetic resonance imaging; PLAX, parasternal long axis; PSAX, parasternal short axis; RBBB, right bundle-branch block; SAECG, signal-averaged ECG; RV, right ventricle; and RVOT, right ventricular outflow tract.

*The morphology of “Major” ventricular arrhythmias is LBBB with a QRS axis other than inferior (ie, intermediate or superior).

Appendix B

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee

Office G50- Old Main Building
Groote Schuur Hospital
Observatory 7925
Email: hrec-enquiries@uct.ac.za
Website: www.health.uct.ac.za/fhs/research/humanethics/forms

05th February 2021

HREC/REF: 766/2014

Prof N Ntusi
Department of Medicine
J-46.53
Old Main Building
Groote Schuur Hospital
Email: ntobeko.ntusi@uct.ac.za

Dear Prof Ntusi

Project Title: RATIONALE, DESIGN AND IMPLEMENTATION OF THE AFRICAN CARDIOMYOPATHY AND MYOCARDITIS REGISTRY (ACMR) (incorporating the following studies: A CLINICAL AND GENETIC STUDY OF FAMILIAL DILATED CARDIOMYOPATHY IN SOUTH AFRICA HREC REF 197/96, THE ARRHYTHMOGENIC RIGHT VENTRICULAR CARDIOMYOPATHY REGISTRY OF SOUTH AFRICA HREC REF 047/2003)(PhD candidate Dr S Kraus)

Thank you for your email to the Faculty of Health Sciences Human Research Ethics Committee (HREC) dated 25th December 2020.

The HREC has granted annual approval until 12th February 2021.

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please quote the HREC REF in all your correspondence.

Yours sincerely

Singed by an official

PP **PROFESSOR M BLOCKMAN**
CHAIRPERSON,
FACULTY OF HEALTH SCIENCE HUMAN RESEARCH ETHICS COMMITTEE

Appendix C

UNIVERSITY OF CAPE TOWN

FACULTY OF HEALTH SCIENCES
Human Research Ethics Committee

FHS016: Annual Progress Report / Renewal

HREC office use only (FWA00001637; IRB00001938)			
This serves as notification of annual approval, including any documentation described below.			
☑ Approved	Annual progress report	Approved until/next renewal date	30.06.2025
☐ Not approved	See attached comments		
Signature Chairperson of the HREC/ Designee		Date Signed	3/4/2024

Note: Please email this form and supporting documents (if applicable) in a combined pdf-file to hrec-enquiries@uct.ac.za.
Please clarify your plan for research-related activities during COVID-19 lockdown.
Please use the latest form found on our website:
http://www.health.uct.ac.za/fhs/research/humanethics/forms

HUMAN RESEARCH ETHICS COMMITTEE
- 2 APR 2024
HEALTH SCIENCES FACULTY
UNIVERSITY OF CAPE TOWN

Comments to PI from the HREC

Principal Investigator to complete the following:

1. Protocol information

Date (when submitting this form)	02/04/2024		
HREC REF Number	425/2023	Current Ethics Approval was granted until	30/06/2024
Protocol title	Genealogical Tracing of Founder Variants Linked to a South African Cohort (Sub-Study Linked to HREC: 766/2014)		
Protocol number (if applicable)			
Are there any sub-studies linked to this study?		☐ Yes	✓No
If yes, could you please provide the HREC Reference number for all sub-studies? **Note**: A separate FHS016 must be submitted for each sub-study.			
Principal Investigator	A/Prof Gasnat Shaboodien		

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

Department / Office Internal Mail Address	Gasnat.shaboodien@uct.ac.za

1.1 Does this protocol receive US Federal funding?	☐ Yes	✓ No
1.2 If the study receives US Federal Funding, does the annual report require full committee approval? **Note:** Any annual approvals for **Full Committee** review MUST be submitted on the monthly HREC submission dates. (Please send electronic copy for full committee review to hrec-submission@uct.ac.za)	☐ Yes	✓ No

If yes in 1.2 please complete section 1.3 below for invoicing purposes

1.3 Ethics Renewal Fee

Please **(tick ✓)**appropriate box for billing purposes:

Submission Type	***Description***	***New fee (Vat Incl.)***	***tick ✓***
Research funded solely from UCT departmental/ divisional/group budget	Annual evaluation of research progress report for re-certification	**R0,00**	✓
Non-sponsored student research for degree purposes at UCT/Other Universities & Colleges	Annual evaluation of research progress report for re-certification	**R0,00**	☐
Annual re-certification / Progress report (FHS016 Form)	Clinical Trial & International Grant Funded Research - Annual evaluation of research progress report for re-certification for Full Committee Approval	**R7000,00**	☐
Annual re-certification / Progress report (FHS016 Form)	Clinical Trial & International Grant Funded Research - Annual evaluation of research progress report for re-certification for Expedited review	**R3 710.00**	☐
Annual re-certification / Progress report (FHS016 Form)	National grant funded research - Annual evaluation of research progress report for re-certification for Full Committee Approval	**R6000.00**	☐
Annual re-certification / Progress report (FHS016 Form)	National Grant funded research for Annual evaluation of research progress report for re-certification for Expedited review	**R1 500,00**	☐

NB: Protocols funded by UCT (e.g. departmental funding / student research) and by certain grant funding organizations (e.g. MRC, NRF, CANSA,) are exempt from these charges.

Please provide details for Invoicing, either complete section 1 or 2 :

1. Invoice billing – Directly to Sponsor

Sponsor's name	
Billing Address of Sponsor:	
Vat Number:	

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

Contact person	
Telephone number	
Email Address	
2. **Internal Journal Billing:**	
Fund Number:	
Cost Centre Number:	
Account Holder Name:	
Division of Account Holder:	

2. List of documentation for approval

NA

3. Protocol status (tick ✓)

☐	Open Enrolment
☐	Closed to enrolment (tick ✓)
☐	Research-related activities are ongoing
☐	Research-related activities are complete, long-term follow-up only
☐	Research-related activities are complete, data analysis only
☐	Main study is complete but sub-study research-related activities are ongoing
☐	Study is closed → Please submit a Study Closure Form (FHS010)

4. Enrolment

Number of participants enrolled to date	100
Number of participants enrolled, since last HREC Progress report (continuing review)	100
Additional number of participants still required	0

5. Refusals

Total number of refusals (participants invited to join the study, but refused to take part)	0

6. Cumulative summary of participants

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

Total number of participants who provided consent	**100**
Number of participants determined to be ineligible (i.e. after screening)	0
Number of participants currently active on the study	NA
Number of participants completed study (without events leading to withdrawal)	NA
Number of participants withdrawn at participants' request (i.e. changed their mind)	NA
Number of participants withdrawn by PI due to toxicity or adverse events	NA
Number of participants withdrawn by PI for other reasons (e.g. pregnancy, poor compliance)	NA
Number of participants lost to follow-up. Please comment below on reasons for loss of follow-up.	NA
Number of participants no longer taking part for reasons not listed above. Please provide reasons below:	NA

7. Progress of study

Please provide a brief summary of the research to date including the overall progress and the progress since the last annual report as well as any relevant comments/issues you would like to report to the HREC:
To date 100 participants have been enrolled: • Microsatellite primer optimisation is complete and haplotype analysis is underway for all participants • Genealogical tracing of participants is underway and some families have been completed

8. Protocol violations and exceptions (tick ✓ all that apply)

✓	No prior violations or exceptions have occurred since the original approval
☐	Prior violations or exceptions have been reported since the last review and have already been acknowledged or approved
☐	Unreported minor violations that have occurred since the last review, as well as significant deviations not yet reported, are attached for review

9. Amendments (tick ✓ all that apply)

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

✓	No Prior amendments have been made since the original approval
☐	Prior amendments have been reported since the last review and have already been approved
☐	New protocol changes/ amendments are requested as part of this continuing review (See note below)

Note: If new protocol changes are being requested in this review, please complete an amendment form (FHS006).
Specific changes in the amended protocol and consent/assent forms must be **bolded**, *italicised* or tracked and all changes must include a rationale.

10. Adverse events

10.1 Please provide below or attach a narrative summary of serious adverse events and/ or unanticipated problems since the last progress report. Please indicate changes made to the protocol and informed consent document(s) as a result (if not already reported to the HREC). Please comment on whether causality to any study procedure or intervention could be established.
None

10.2 Have participants received appropriate treatment/ follow-up/ referral when indicated (e.g. in the case of abnormal or incidental clinical findings, distress or anxiety)?		
☐ Yes	☐ No	✓ Not applicable
If yes, please describe:		

11. Summary of Monitoring and Audit Activities (tick ✓)

11.1 Was this study monitored or audited by an external agency (e.g. SAHPRA, FDA)?		
☐ Yes	☐ No	✓ Not applicable

11.2 Did a Data and Safety Monitoring Board publish a report?		
☐ Yes	☐ No	✓ Not applicable

11.3 If yes, please identify the agency and attach a summary of the findings.					
Agency Name		Report attached	☐ Yes	☐ No	☐ Not applicable
		DSMB report attached	☐ Yes	☐ No	☐ Not applicable

11.4 Has there been any agency, institutional or other inquiry into non-compliance in this study, or any finding of non-compliance concerning a member of the research team?	
☐ Yes	✓No

If yes, please explain:

12. Level of risk (tick ✓)

12.1 In light of your experience of this research, please indicate whether the level of risk to participants has:	
☐	Increased
☐	Decreased
✓	Shown no change
If there has been a change, please explain:	

12.2 Please provide a narrative summary of recent relevant literature that may have a bearing on the level of risk.

13. Insurance

Please confirm that valid no fault insurance is still in place? **(tick ✓)**			
☐ Yes	☐ No		✓ Not Applicable – N/A
If yes, please complete the following:			
Insurer's name:			
Policy no.		*Coverage Period:	
For UCT sponsored studies please liaise the Insurance office via fhs.sponsorship@uct.ac.za regarding the required documentation and information required obtain a renewed UCT No-fault Insurance Certificate.			

14. Statement of conflict of interest

Has there been any change in the conflict of interest status of this protocol since the original approval? (tick ✓)	
☐ Yes	✓ No
If yes, please explain and if necessary, attach a revised conflict of interest statement (Section #7 in the New Protocol Application Form FHS013):	

15. Signature

My signature certifies that the above is complete and correct.			
Signature of PI		Date	02/04/2024

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

Appendix D

Molecular Laboratory, Division of Cardiovascular Genetics, Hatter Institute
4th Floor, Chris Barnard Building, UCT Medical School, Observatory,
Cape Town, 7925. Tel: +2721 4066615; Fax: +2721 4478789
IMHOTEP PCC. Tel: 021 404 7674/7673, email: sarah.kraus@uct.ac.za

institutions involved in approved studies, personal identifying information (name, contact numbers, address etc) will not be shared. It is now common for genetic information to be shared with researchers around the world. The benefit is that many researchers can use the same information to answer different research questions. We are not always able to do all the tests on the DNA at our laboratories as we may not have the necessary equipment or expertise. For this reason, we sometimes have to ask other institutions and researchers in other parts of the world to help us.

Will you get to know your results?

Due to the nature of research, you may or may not receive a result from studies performed on your DNA. Although the laboratory will do its best to confirm that the findings relate to the heart condition, results received from a research laboratory should be confirmed diagnostically with a confirmatory test. If a genetic cause for the heart condition is found, the researchers from the Hatter Institute will do their best to inform you of the results, either via your doctor, a genetics counsellor or in writing, depending on the available resources. If your contact details change, it is your responsibility to inform the laboratory. It is important to tell the doctors whether you want your results to be shared with your family in the event that you are unavailable or incapacitated. We will only be investigating genetic causes of cardiomyopathy. Therefore, we will not be looking for genetic mistakes that cause other medical conditions. If we do happen to find something unexpected in a participant's DNA that is important to a participants health or well-being, we will do our best to inform them.

What are the risks of being involved in genetics studies and what about confidentiality?

The prick of the needle to take blood can be uncomfortable, but it is not dangerous to take blood. It is possible that someone could find out that you have participated in this study. We don't know if this is likely to happen, but we will do our very best to make sure that this does not happen. While a genetic result can be helpful in guiding your medical care, it can be difficult learning that you have a genetic diagnosis and this information may potentially affect your family relationships, lead to stigmatisation or loss of insurability. We will do our best to make sure you receive the help you need if such difficulties arise. Importantly, we are not required to share your results with anyone without your permission. The genetics laboratories are under obligation to respect and protect your confidentiality and your personal information will not be shared and your identity will not be published in any reports.

Are there any benefits to participating in this project?

You will not necessarily benefit from participating in this project, and we cannot offer you anything to thank you for participating. Your participation in this project may help us understand why some people develop cardiomyopathy and others do not. This could help us treat other people in the future. If we do find a 'spelling error' (mutation), we may be able to look for the same mutation in other members of your family, and find out who may be a risk of developing the same heart condition.

Has this research been approved?

All studies have to be approved by an ethics committee of the University of Cape Town. IMHOTEP and research looking at the genetic causes of cardiomyopathies has been approved (HREC REF: 766/2014). The full name of the committee is 'Faculty of Health Sciences Human Research Ethics Committee'. An ethics committee is composed of a group of people who decide whether the study is risky, and whether people should be allowed to participate in it. These people are there to protect research participants like yourself. You can contact the ethics committee if you have any concerns about this study. They are very friendly people and would welcome your questions. Please contact 021 406 6338, 021 406 6626 or 021 406 6496 if there is anything you would like to discuss with the ethics committee.

Do I have to take part and what if I change my mind?

You do not have to take part in this study. You can decide to participate or not. If you decide not to participate then you and your family will still be treated in the same way. You can also decide at any point in time that you do not want the blood samples to be used any longer. Please contact the molecular genetics laboratory at the Hatter Institute for Cardiovascular Research in Africa if you have any questions or want to withdraw. You can contact us on 021 406 6615. Alternatively, you can write to us and fax the letter to 021 447 8789 or post it to the Molecular Laboratory, Division of Cardiovascular Genetics, 4th Floor, Chris Barnard Building, University of Cape Town Medical School, Observatory 7925. We will destroy your blood samples and genetic material and information. However, once we have shared your genetic information with other researchers and published the results, we cannot destroy the research that has already been done.

If you have additional questions that have not been answered, please contact the IMHOTEP Project Coordinating Centre on 021 404 7674/7673 or sarah.kraus@uct.ac.za, and either a study doctor or a genetics counsellor will contact you to address your questions.

Summary of consent for participation in genetics research for adults	*Please circle your answer*

Molecular Laboratory, Division of Cardiovascular Genetics, Hatter Institute
4th Floor, Chris Barnard Building, UCT Medical School, Observatory,
Cape Town, 7925. Tel: +2721 4066615; Fax: +2721 4478789
IMHOTEP PCC. Tel: 021 404 7674/7673, email: sarah.kraus@uct.ac.za

I understand that I 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 I will not be compensated financially for participating in this research.	Yes / No
I understand that a blood sample (5-25ml or 1-5 teaspoons of blood) will be collected from me and my genetic material will be extracted from this sample for analysis. In some instances, other samples (skin cells, saliva, tissue, serum) may be collected depending on the circumstances (specify if collected):	Yes / No
I understand that these samples, including a portion of my genetic material will be stored at the Hatter Institute for additional/future cardiomyopathy-related research projects approved by the University of Cape Town Human Research Ethics Committee (HREC).	Yes / No
I authorize that my doctors can provide relevant clinical information (medical records) to researchers and that this information will be recorded anonymously on the IMHOTEP Registry database.	Yes / No
I understand that my genetic material/information, together with other relevant medical information, may be shared anonymously with other collaborating researchers and institutions involved in HREC approved studies but my personal identifying information will not be shared.	Yes / No
I understand that the nature of research means that I 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 condition, results received from a research laboratory should be confirmed diagnostically. If a genetic cause for my 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. I understand that my results will not be shared with anyone else (including family members, employers, insurers) unless I provide permission to do so. If my contact details change, it is my responsibility to inform the laboratory.	Yes / No
In the event that I am unavailable or incapacitated, I do want my immediate family members to be informed of the results. Specify family member if necessary	Yes / No
I understand, that depending on the results, it may be necessary to ask my family to be included in this research. I am willing to be part of a family genetics study.	Yes / No
I understand that the genetics molecular laboratory is under obligation to respect my confidentiality and my medical and genetic information will be kept private. I understand that my participation in genetics research is entirely voluntary and that I may withdraw my consent at any time without it affecting my future medical care.	Yes / No
I have been provided with an information sheet explaining the risks and benefits of participating in genetics research, and contact information for the laboratory.	Yes / No

_____	_____	_____
Participants name	Signature	Date
_____	_____	_____
Study doctor/nurse/genetics counsellor	Signature	Date

NB: Please complete the information over the page. This information is necessary for the laboratory to process the specimen. This form will be filed securely in the laboratory in order to protect the participants confidentiality.

Molecular Laboratory, Division of Cardiovascular Genetics, Hatter Institute
4th Floor, Chris Barnard Building, UCT Medical School, Observatory,
Cape Town, 7925. Tel: +2721 4066615; Fax: +2721 4478789
IMHOTEP PCC. Tel: 021 404 7674/7673, email: sarah.kraus@uct.ac.za

Instructions: Blood should be drawn in plastic EDTA Tubes (purple top) or Paxgene Tubes. 20ml of blood in total is required for adults (2-4 tubes). For children, 5-20ml depending on size (maximum 3ml/kg). Please label blood tubes with barcoded sticker. 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.

Place Barcoded Sticker here

DATE OF SPECIMEN COLLECTION: _____/_____/_____, TIME: _____

_ 3 _
3 _ .

Participant details

Surname:	**Name:**	
Hospital No.:	**Biological sex:** M ☐ F ☐	**DOB:**
Address:		
Contact numbers:		
Email address:		
Referring hospital/doctor:		
Referring doctor's contact details:		

Clinical and family information: Please complete family pedigree below

Participant: PROBAND ☐ RELATIVE ☐ If relative, family relationship to proband: _____

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

Family surname: _____ **Family ID** 3 _

Ethnic Origin: Black African ☐ Mixed race ☐ Caucasian ☐ Asian ☐ Other _____

Provisional Diagnosis:

Clinical status: **Clinically affected ☐** **At Risk (unaffected clinically) ☐** **Spouse ☐** **Query ☐**

Morphofunctional phenotype	**Specific diagnosis**
☐ Dilated cardiomyopathy	☐Familial ☐Idiopathic ☐PPCM ☐Secondary (specify)
☐ Hypertrophic cardiomyopathy	☐HCM ☐HCM – phenocopy (specify)
☐ ARVC	☐Definite ☐Borderline ☐Possible ☐Unconfirmed
☐ Restrictive cardiomyopathy	☐Familial ☐Idiopathic ☐Specific aetiology (specify)
☐ Unspecified cardiomyopathy	☐LVNC ☐Takusubu ☐Other (specify)
Additional disorders:	

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

If yes, please specify lab: _____

Family Pedigree (or attach separately):

Molecular Laboratory, Division of Cardiovascular Genetics, Hatter Institute
4th Floor, Chris Barnard Building, UCT Medical School, Observatory,
Cape Town, 7925. Tel: +2721 4066615; Fax: +2721 4478789
IMHOTEP PCC. Tel: 021 404 7674/7673, email: sarah.kraus@uct.ac.za

For Laboratory use only:
DNA number _____Vol. blood _____(ml) Other: _____
Date received (DD/MM/YYYY) _____/_____/_____

IMHOTEP (HREC REF: 766/2014)
Illustrated Assent Form for Children

Study number ☐☐☐ ☐ ☐☐☐☐☐

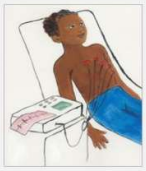
Doctors and nurses are doing a study on the hearts of children

Cardiomyopathy is a problem that affects the heart muscle

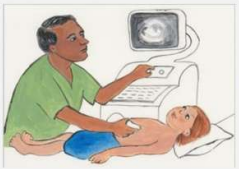
Do you have any questions for the doctor or nurse?

We would like to ask you questions about your health

We would also like to take some blood. The prick of the needle will cause some pain but it will be over quickly



We would like to do special tests. This test helps us to make a drawing of the heartbeat. This won't hurt at all.



We would like to take a video of your heart. This won't hurt at all

If you agree to be part of our study, you will be helping doctors learn more about the problem with your heart and how to treat children with the same heart problem

Write your name _____ *Investigators name and signature.* _____ *Date* _____

IMHOTEP Project Coordinating Centre
J52 UCT Clinical Research Unit, Old Main Building,
Groote Schuur Hospital, Cape Town, Tel: 0214067674

IMHOTEP: The African Cardiomyopathy and Myocarditis Registry Program
Project Coordinating Centre (PCC), University of Cape Town
J52, Old Main Building, Groote Schuur Hospital, Cape Town, 7925, South Africa
Tel: +2721 4047674. Email: sarah.kraus@uct.ac.za. UCT HREC Tel: +2721 4066492

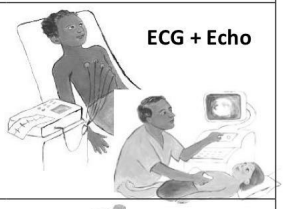
We are doctors and nurses and we are doing a study on the hearts of children with cardiomyopathy. Cardiomyopathy is a problem that affects the heart muscle, making the heart tired and unable to pump blood properly. If you agree to take part in this study, you will be helping doctors to learn more about the problem with your heart, and how to treat other children with the same heart problem.

If you agree to be in our heart study, what will you be asked to do?

You may tick the box ☑ if you understand and are happy to take part:

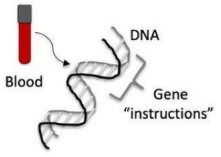
☐ We would like to ask you some questions about your health and listen to your heart. This will be very quick and will not hurt at all.

☐ We would like to do some special tests where we take a picture (ECG) and a video ("Echo") of your heart by using special machines. These are the same machines that your doctor may have used before to look at your heart. This will not hurt at all.



☐ We would also like to take some blood. The prick of the needle will cause some pain but it will be over quickly. We will take only a few teaspoons of blood (maximum five teaspoons) depending on how big you are.

☐ Your blood will be sent to a special laboratory where other doctors and scientists working with your doctors in this study, will be able to do special tests on your blood. In your blood, you have DNA that contains a special set of 'instructions' for your heart, and by reading these 'instructions', we may find out why you (or someone in your family) are sick. We will keep your blood safely stored in the laboratory freezer.



☐ Sometimes other people in your family can get the same heart problem, so we may invite your family to join the study. Would this be okay?

☐ Your parents have given us permission to ask you to join the study. You are allowed to say that you don't want to be in the study or have blood taken. Nobody will be angry with you if you say no and you may also change your mind at any time. Before you decide, you can ask questions, or talk to your mother and father about this. By ticking the boxes and writing your name (or stamping your fingerprint) on this form, you are saying YES to being in our study.

Write your name (or fingerprint)	*Investigators name and signature*	*Date*

☐☐☐ _ ☐ _ ☐☐☐☐☐

IMHOTEP: The African Cardiomyopathy and Myocarditis Registry Program
Project Coordinating Centre (PCC), University of Cape Town
J52, Old Main Building, Groote Schuur Hospital, Cape Town, 7925, South Africa
Tel: +2721 4047674. Email: sarah.kraus@uct.ac.za. UCT HREC Tel: +2721 4066492

For the investigator to complete to confirm the minors choice: ☐☐☐ _ ☐ _ ☐☐☐☐☐

- ☐ The minor has agreed to have a blood sample taken
- ☐ The minor has refused to have a blood sample taken, but agreed to be in the registry and have an alternative DNA sample taken (sputum, buccal swab)
- ☐ The minor has refused to have any DNA sample taken but has agreed to be in the registry
- ☐ Blood sample (or alternative DNA sample) not taken for another reason (i.e. not necessary)

Specific questions asked by the minor	Answers from the investigator

If you have any additional questions, you may contact

Name _____

Contact information for investigators	
Site investigator	
Site investigator contact number	
Site HREC number	
Site HREC Contact number	

IMHOTEP: The African Cardiomyopathy and Myocarditis Registry Program
Project Coordinating Centre (PCC), University of Cape Town
J52, Old Main Building, Groote Schuur Hospital, Cape Town, 7925, South Africa
Tel: +2721 4047674. Email: sarah.kraus@uct.ac.za. UCT HREC Tel: +2721 4066492

Assent to participate in IMHOTEP Registry and Repository for adolescents

My parents (guardian) have agreed to let me participate in a study of cardiomyopathy and/or myocarditis. I understand that cardiomyopathy is heart muscle problem that may lead to heart failure (where the heart becomes tired and unable to pump blood properly) and/or arrhythmias (where the heart beats too fast or too slow). Myocarditis is inflammation (injury) of the heart muscle that can lead to cardiomyopathy. **IMHOTEP** is registry and repository, that collects information and blood samples from patients with these heart conditions, and is approved by the Human Research Ethics Committee (HREC) at my hospital, and other hospitals. I have been invited to participate because I (or someone in my family) have been diagnosed with cardiomyopathy or myocarditis. I understand that I will be asked questions about my health and I will be examined by a doctor, and that study investigators will collect information (e.g. results from tests) from my medical records. I understand that I may need to have additional tests done. **These tests will only be done if necessary to confirm I have cardiomyopathy, or to help in the treatment of my condition**. These extra tests will not be painful and will not cost any money. I understand that we will not be given money for being in this study.

I agree to participate the IMHOTEP Medical Registry	**Yes / No**

In addition, the study doctors would like to take a blood sample from me (about 5 teaspoons). Blood is taken by pricking the skin with a needle which may be painful but will be over quickly. This blood will be used to test my DNA for genetic mistakes that may cause cardiomyopathy. The doctors have explained that in my blood, I have DNA that contains a special set of 'instructions' for my heart, and by reading these 'instructions', they may find out more about what causes cardiomyopathy. **Before my blood can be taken, I must agree, and my parents (guardian) must sign another consent form that protects the use of my DNA by study investigators under the rules of the Human Research Ethics Committee.** My blood sample will be stored (kept safe for many years) in the Cardiovascular Genetics Laboratory at the Hatter Institute in Cape Town, and will be used by scientists in IMHOTEP research studies. The DNA analysis (reading of 'instructions') will be done at this laboratory and other laboratories that are working for IMHOTEP. The doctors have explained that there are some risks and benefits of research to me and have given me more information to read with my parents (guardian).

I agree to having a blood sample taken that will be stored for use in genetics research	**Yes / No**

I also agree to be followed up by the study investigators to see if there are any changes in how I feel, my medication, or whether I have needed more tests, or had any complications or procedures. I understand that my doctor will follow me up as usual and I will not need to attend extra appointments if I join this study. The study doctors may contact my doctor(s) to get my medical information.

I agree to be followed-up by IMHOTEP and researchers may look at my medical records	**Yes / No**

I understand that cardiomyopathy may be genetic. Genetic cardiomyopathies are caused by changes (mutations) in genes within our DNA that result in abnormalities in the heart muscle cells. These genetic heart conditions can be inherited and may affect many individuals within a family. For this reason, other people in my family may also be asked to join the study. The study doctors will explain more about this to me if necessary.

I am willing to be in a study with my family	**Yes / No**

I understand that my participation in this study is entirely voluntary (my choice) and even although my parents (guardian) have agreed, I can say no. I can also change my mind about being in the study at any time, and I will not get into trouble. My information will be kept private (confidential), and will only be used for research about cardiomyopathies and myocarditis. Information about me will be recorded on a computer that is protected by a special password. My name will not be used so no one will know I am in the study. All my test results will be given to my doctor(s) and my parents. I may choose to keep some information I tell the doctor private, even from my parents.

I understand that I can choose to be in this study or not, and that I can change my mind	**Yes / No**

☐☐☐ - ☐ - ☐☐☐☐☐

Participants initials _____ Investigators initials _____

IMHOTEP: The African Cardiomyopathy and Myocarditis Registry Program
Project Coordinating Centre (PCC), University of Cape Town
J52, Old Main Building, Groote Schuur Hospital, Cape Town, 7925, South Africa
Tel: +2721 4047674. Email: sarah.kraus@uct.ac.za. UCT HREC Tel: +2721 4066492

Summary of consent continued (Page 3-4)

☐☐☐ _ ☐ _ ☐☐☐☐☐
Study number

Everything has been explained to me and I have been allowed to ask questions. My parents (guardian) and I have also been given extra information about cardiomyopathies, and the risks and benefits of participating in this study that we can read through together.

If I have more questions, I may ask my parents, my doctor, or the study doctor/nurse

_____at_____.

☐ I am willing to participate in IMHOTEP and give a blood sample for genetics research

☐ I am willing to participate in IMHOTEP but I DO NOT want to give a blood sample

☐ I DO NOT want to be in the IMHOTEP study at all even although my parents (guardian) have agreed

_____	_____	_____
Minor participant name and surname	*Minor participant signature*	*Date*
_____	_____	_____
Investigator name and surname	*Investigator signature*	*Date*

Contact information for investigators	
Site investigator	
Site investigator contact number	
Site HREC number	
Site HREC Contact number	