Validation of the *PARVA* c.392A>T variant in a South African family with severe Arrhythmogenic Right Ventricular Cardiomyopathy

Stephen Kamuli

Thesis presented for the Degree of

MASTER OF SCIENCE

In the Department of Medicine

UNIVERSITY OF CAPE TOWN

Supervisor: Dr Gasnat Shaboodien

Co-supervisor: Professor Bongani Mayosi

August 2016

DECLARATION

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

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

Signature: signature removed

Date: 18 August 2016

TABLE OF CONTENTS

ABSTRACT

Introduction

Cardiomyopathy is an endemic disease in Africa that is a major contributor to the clinical syndrome of heart failure. The various forms of cardiomyopathies pose a great challenge in Africa for many reasons, including the difficulty of diagnosis and the scarcity of interventions such as heart transplantations in resource-poor environments. The aetiology of the cardiomyopathies had been unknown but various genetic abnormalities associated with cardiomyopathy have been unraveled. A previous whole exome sequencing project conducted in the United Kingdom (UK) had identified parvin alpha (*PARVA*) as a candidate gene in a South African family, ACM 8, with several members affected with arrhythmogenic right ventricular cardiomyopathy (ARVC).

Hypothesis: We hypothesize that *PARVA* harbors novel genetic mutations that cause ARVC and other forms of cardiomyopathy.

Aim: To screen the *PARVA* gene for mutations in a large panel of probands with ARVC and other cardiomyopathies and to validate the whole exome sequencing results obtained in the UK on a different sequencing platform.

Methods and Results

We investigated the ACM 8 family with three affected individuals (two severely affected children and the mother) for whom the genetic cause of the disease was unknown. Genetic analysis was previously performed at Newcastle in the UK using whole exome sequencing on an Illumina platform. In this analysis, the *PARVA* c.392A>T variant was identified as a possible cause of ARVC in this family. We expanded on this work by using high resolution melt (HRM) analysis and Sanger sequencing to screen all the available ACM 8 family members to determine segregation of the *PARVA* c.392A>T variant within this family. We observed that the phenotypic variability seen within this family cannot be explained by the *PARVA* c.392A>T variant alone and called into question the causative role of *PARVA* within this family We also screened the cardiomyopathy cohort consisting of 180 probands diagnosed with ARVC, dilated

cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM) and restrictive cardiomyopathy (RCM) in the Cardiovascular Genetics Laboratory. No definitive evidence of pathogenic *PARVA* variants was found any of the cardiomyopathy probands screened. We subsequently performed whole exome sequencing on this family to validate the UK findings (Ion Torrent platform). We found that both affected individuals were homozygous for the *PKP2* c.1162C>T mutation. *PKP2* is a gene known to cause ARVC, and the c.1162C>T mutation has been described as a founder mutation for autosomal dominant ARVC families of Afrikaner decent in South Africa.

Conclusion

While this study set out to validate the whole exome sequencing experiments conducted in family ACM 8 in the UK, we instead found the causal variant to be the previously reported *PKP2* c.1162C>T mutation. We also explored the possibility of *PARVA* as a causal gene for ARVC but no pathogenic *PARVA* mutations were identified.

ACKNOWLEDGEMENTS

- To my supervisor, Dr Gasnat Shaboodien, for all her vital and consistent advice, assistance, enthusiastic guidance and support throughout this project

- To my co-supervisor, Professor Bongani Mayosi, for his motivation, shared wisdom and opportunity to join this team of exceptional individuals

- To Timothy Spracklen for his insightful comments, timely responses and expertise that contributed significantly in the whole project

- To all the other members of the Cardiovascular Genetics Team at the Hatter Institute for Cardiovascular Research in Africa, including Dr Gaurang Deshpande, Mr. Babu Muhamed, Mrs. Lameez Pearce, Dr. Maryam Fish, Ms. Janine Saaiman and Ms. Tafadzwa Machipisa for all their time and making the time at the lab well spent and enjoyable.

- To all the other members of the Hatter Institute for Cardiovascular Research in Africa including Professor Karen Sliwa, for all their time and making the student office a really conducive place to work.

- To the H3Africa Consortia (PROJECT: RHDGen) for funding my studies.

- To my mother and brother for always being there for me more so spiritually throughout the course of this project.

- Above all I would like to thank The Almighty GOD for the gift of life and sound health during the course of undertaking this master's course

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

LIST OF TABLES

Chapter 1

Chapter 2

Chapter 3

LIST OF ABBREVIATIONS AND SYMBOLS

ACM	Arrhythmogenic right ventricular cardiomyopathy
ACTC1	Alpha-actin cardiac muscle 1
ACTN	Alpha-actinin
AKAP9	A-Kinase Anchoring Protein 9
AKT	Protein kinase B
ALK	Anaplastic lymphoma receptor tyrosine kinase
AMPK	5^{l}-activated AMP protein kinase
ANK3	Ankyrin 3 Node Of Ranvier
ANKRD1	Ankyrin repeat domain-containing protein 1
ARVC	Arrhythmogenic right ventricular cardiomyopathy
ATP	Adenosine triphosphate
βMHC	β-Myosin Heavy Chain
BP	Base pair
Ca	Calcium
CACNA1C	Calcium Voltage-Gated Channel Subunit Alpha1 C
CASSA	Cardiac Arrhythmia Society of Southern Africa
CdGAP	Rho GTPase activating protein 31
CH	Calponin homology
CPGR	Centre for Proteomic and Genomic Research
CRP	C-reactive protein

CSRP3	Cysteine and glycine-rich protein 3
cTnT	Cardiac troponin T gene
DCM	Dilated cardiomyopathy
ddNTPs	Dideoxynucleotide triphosphates
DES	Desmin
DLEC1	Deleted in lung and esophageal cancer 1
DMD	Dystrophin
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates
DSC2	Desmocollin-2
dsDNA	Double stranded DNA
DSG2	Desmoglein-2
DSP	Desmoplakin
ECG	Electrocardiography
EDTA	Ethylenediaminetetraacetic acid
EMF	endomyocardial fibrosis
EVS	Exome Variant Server
gDNA	genomic DNA
GRIA3	Glutamate receptor ionotropic AMP3
GSK3	Glycogensynthase kinase3

GTPases	Guanine triphosphatases
HCM	hypertrophic cardiomyopathy
HFE	Hemochromatosis
HLA-DR1	Human Leukocyte Antigen - antigen D Related
HMGXB3	HMG-box containing 3
HRM	High Resolution Melting
ICDs	Implantable cardioverter-defibrillators
ILK	Integrin linked kinase
JNK	Jun Nterminal kinase
JUP	Plakoglobin
LAMP2	lysosome-associated membrane protein 2
LMNA	Lamin A/C
LV	left ventricular
LVH	left ventricular hypertrophy
LVNC	Left ventricular non-compaction
MAF	Minor allele frequency
MRI	Magnetic resonance imaging
MYBPC3	Myosin-binding protein C, cardiac-type
MYH7	Myosin-7
MYL2	Myosin Light Chain 2

MYL3	Myosin light chain 3
MYOZ2	Myozenin-2
NEURL	Neuralized homolog
NGS	Next generation sequencing
OBSCN	Obscurin
PARVA	Parvin alpha
PCR	Polymerase chain reaction
PDE4DIP	Phosphodiesterase 4D Interacting Protein
PINCH	Particularly interesting new cysteine-histidine-*rich* protein
PIP	Parva ILK PINCH complex
PKB	Protein kinase B
PKP2	Plakophilin 2
PLN	Phospholamban
RBM20	RNA Binding Motif Protein 20
RCM	restrictive cardiomyopathy
RV	Right ventricular
SAM	Sequence alignment map
SCD	Sudden cardiac death
SCN5A	Sodium Voltage-Gated Channel Alpha Subunit 10
SCN10A	Sodium Voltage-Gated Channel Alpha Subunit 10

SMAD5	SMAD family 5
SNP	Single Nucleotide Polymorphisms
SVIL	Supervillin
SYNE	Spectrin Repeat Containing Nuclear Envelope Protein 1
SYNM	Synemin
TAZ	Tafazzin
Tcap	Titin cap
TGFß3	Transforming growth factor beta-3
Tm	Melting temperature
TNF-α60	Tumor necrosis factor alpha60
TNNI3	Troponin I, cardiac muscle
TNNT2	Cardiac Troponin T type 2
TPM1	Tropomyosin 1
TRPS 1	Tichorhinophalangeal syndrome I
TTN	Titin
UBR4	Ubiquitin protein ligase E3 component N-recognin 4
µl	Microlitre
US	United States
VCL	Vinculin
WPBTS	Western Province Blood Transfusion Service

WPW Wolf-Parkinson-White syndrome

ZNF141 Zinc finger protein 141

Nucleic acid codes

A Adenine

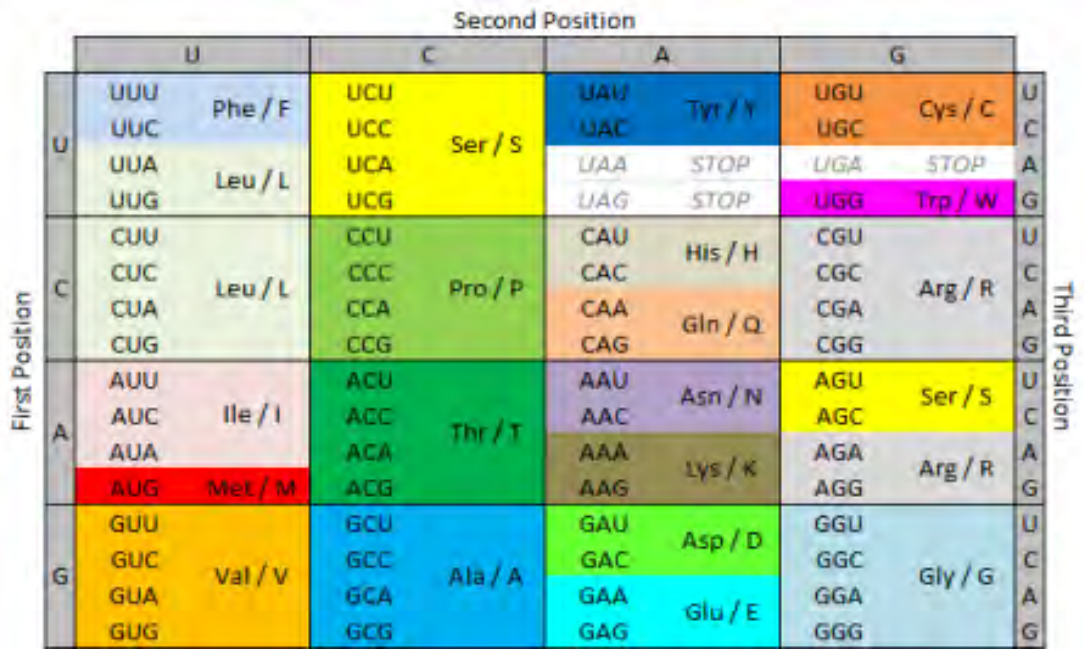
G Guanine

C Cytosine

T Thymine

Amino acid codes

Indicated below are the names, symbols and properties of the amino acids (from https://sarcasticresonance.wordpress.com/2013/10/)



CHAPTER 1: INTRODUCTION

1.1. Structure and function of the heart

The human heart is an organ that pumps blood throughout the body via the circulatory system, supplying oxygen and nutrients to the tissues, while removing carbon dioxide and other wastes (URL1). A healthy, adult heart is usually the size of an average clenched adult fist, but may vary due to a person's age, size, and the condition of their heart. The heart is comprised of four chambers, namely the right ventricle, left ventricle, right atrium and left atrium (Figure 1.1) (URL2). Surrounding the heart is a fibrous enclosure called the pericardium which holds the heart in place while allowing it to move as it beats. The wall of the heart is comprised of three layers: the inner endocardium, the middle myocardium, and the outer epicardium, with the myocardium being the thickest layer. The cardiac muscle is a specialized muscle which makes up the myocardium. It is composed of cardiomyocytes, which are specialized muscle cells that contract akin to other muscle cells, but unlike other muscle cells they generate and conduct electricity for contraction and coordination of the heart through the conduction system (URL3).

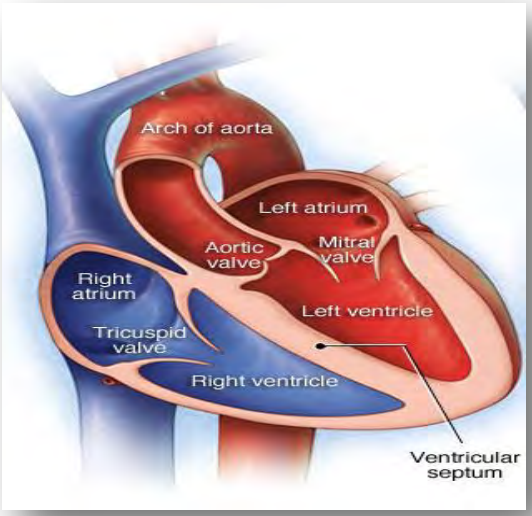


Figure 1.1: Structure of the heart. The blue shading indicates deoxygenated blood received from the rest of the body while the red shading indicates oxygenated blood received from the lungs and pumped to the rest of the body (Adapted from URL4).

Blood from the body enters the heart via the superior and inferior venae cavae which empty oxygen-poor blood into the right atrium. Contraction of the atrium causes blood to flow through the tricuspid valve into the right ventricle. The tricuspid valve shuts when the right ventricle contracts in systole, thus preventing blood from flowing back into the right atrium. When the ventricle contracts, blood is ejected into the pulmonary artery through the pulmonic valve and flows to the lungs, where it is oxygenated. The oxygenated blood then returns to the heart through the pulmonary veins where it is emptied into the left atrium. As this atrium contracts, blood flows from the left atrium into the left ventricle via the open mitral valve. Contraction of the left ventricle leads to ejection of blood into the systemic circulation via the aorta (URL2).

Compared to the atria, the ventricles are the more muscular pumping chambers that eject blood into the pulmonary or systemic circulation via the large arteries. The left atrium and the left ventricle have more muscle in their walls than the right side of the heart and are larger in size than the corresponding chambers on the right. This is because blood from the left side of the heart must be pumped to the rest of the body, while blood from the right side goes to the pulmonary circulation (URL2).

1.2. Cardiomyopathy

Cardiomyopathy is defined as a myocardial disorder in which the heart muscle is structurally and functionally abnormal, in the absence of coronary artery disease, hypertension, valvular disease and congenital heart disease sufficient to cause the observed myocardial abnormality (Elliott, Andersson et al. 2008).

According to the European Society of Cardiology classification system for cardiomyopathies, the cardiomyopathies are grouped into specific morphological and functional phenotypes whereby each phenotype is sub-classified into familial and non-familial forms (Elliott, Andersson et al. 2008, Arbustini, Narula et al. 2014). The main types of cardiomyopathy presented in this scheme are dilated cardiomyopathy (DCM), arrhythmogenic right ventricular cardiomyopathy

(ARVC), hypertrophic cardiomyopathy (HCM), restrictive cardiomyopathy (RCM), and unclassified cardiomyopathies.

In Africa, cardiomyopathy is an endemic disease that is a major contributor to the clinical syndrome of heart failure (Mayosi 2007). The cardiomyopathies pose a great challenge in Africa for many reasons, including the difficulty of diagnosis, which often requires specialized cardiological examinations and the scarcity of interventions such as heart transplantations in resource-poor environments. The high mortality related to these often irreversible heart muscle disorders, together with their high prevalence in societies still plagued by poverty and famine contributes to their burden (Akinkugbe, Nicholson et al. 1991).

1.2.1 Dilated Cardiomyopathy (DCM)

DCM is a heart muscle disease that is defined by the presence of left ventricular dilatation and left ventricular systolic dysfunction in the absence of abnormal loading conditions (hypertension, valve disease) or coronary artery disease sufficient to cause global systolic impairment (Figure 1.2).

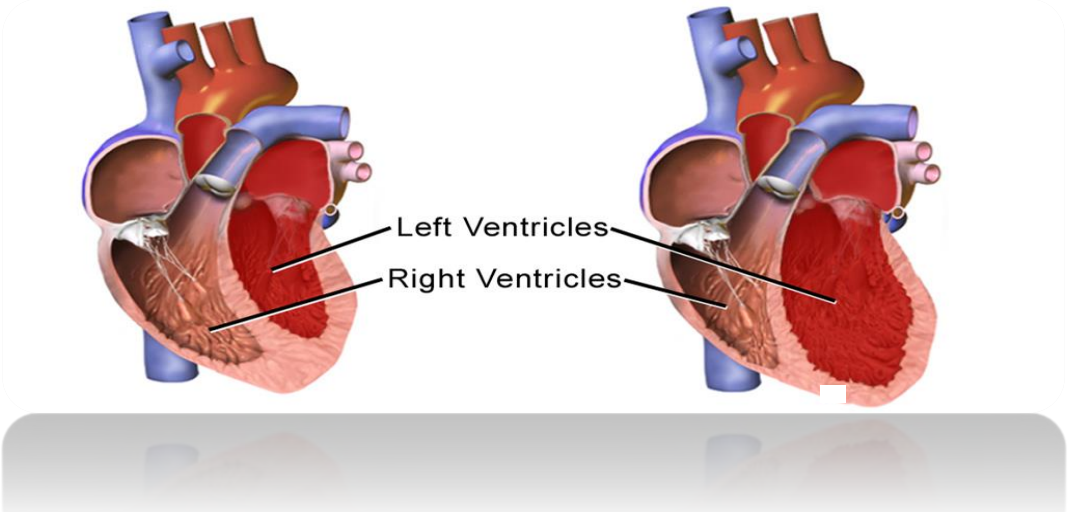


Figure 1.2: Comparison of a normal heart (left) with a DCM heart (right). A cross-section of each heart is presented, showing the ventricles (indicated) and the atria. The left ventricle is enlarged in DCM, and the left ventricular wall is stretched (URL5)

Classically, the disease starts in the left ventricle, where the heart muscle begins to dilate, stretch, and become thinner. Consequently, the inside of the chamber enlarges. This tends to advance to the right ventricle and later to the atria. Right ventricular dilation and dysfunction may be present but are not necessary for diagnosis (Elliott, Andersson et al. 2008). Dilation of the chambers causes the heart muscle to contract abnormally and the heart cannot pump blood effectively. This weakening of the heart causes a predisposition to progressive heart failure which has a 4 year mortality of 34% after onset of symptoms (Felker, Thompson et al. 2000). DCM has been described to have distinctive clinical features that contrast with other forms of heart disease despite its unknown aetiology. Many patients exhibit emboli formation and subsequent infarctions, mostly in the viscera (Cosnett 1962, Cosnett and Pudifin 1964). In early paediatric reports, DCM was also shown to present with similar clinical and pathological manifestations as in adults (Stein, Shnier et al. 1964).

Strict diagnostic criteria are lacking (Silverman and Aksut, 2015) but the diagnosis of DCM is made where there is evidence of dilation and impaired contraction of the left and/or both ventricles with left ventricular ejection fraction less than 40 percent or fractional shortening less than 25 percent (Richardson, McKenna et al. 1996). The disease is considered idiopathic if primary and secondary causes of heart disease such as myocarditis and coronary artery disease are excluded by evaluation including physical examination, careful medical history, laboratory testing results, echocardiography and coronary angiography (to exclude greater than 50 percent obstruction of one or more coronary arteries) (Elliott 2000) (Cosnett 1962, Cosnett and Pudifin 1964).

Reliable data on the incidence and prevalence of dilated cardiomyopathy are difficult to come by (Rakar, Sinagra et al. 1997) with the only formal estimate of DCM prevalence being a study that was conducted in Olmsted County, Minnesota from 1975 to 1984 that estimated its prevalence as of 1985 at 1:2,700 (Codd, Sugrue et al. 1989). This estimate was twice the estimated prevalence of HCM which was 1:5000 from the same cohort during the study period (Codd, Sugrue et al. 1989). Hershberger et al (Hershberger and Morales 1993) suggested the

frequency of DCM is equal to the established frequency of hypertrophic cardiomyopathy (1:500) or even greater at 1:250 (Hershberger and Morales 1993). In Africa, prior myocarditis, malnutrition, excessive alcohol consumption, subsequent myocarditis and heredity have been proposed as causes, and some or all of these factors may play a role in disease expression. Mechanical effects of fibrosis and the immunological response to myocardial damage are also likely contributory factors to progression of the disease (Watkins and Mayosi 2009).

In contrast to more industrialised regions, there have been no population-based studies conducted to assess the prevalence of DCM in Africa and thus its true burden in Africa is unknown (Sliwa, Damasceno et al. 2005). Despite this, it is widely accepted that cardiomyopathy is pervasive, particularly among black Africans (Bradlow, Zion et al. 1964). Studies conducted in South African populations revealed that in a clinical series, DCM accounted for 11.6% to 37.5% of diagnoses of heart disease (Cosnett 1962, McGlashan 1988) and 15.4% to 48% of heart failure admissions (Powell and Wright 1965, Sliwa, Wilkinson et al. 2008). In a necropsy series, DCM was the cause of death in 14.1% to 17% of cases of heart disease (Isaacson 1977, Steenekamp, Simson et al. 1992) and 12.7% of deaths from heart failure (Kallichurum 1969). In a study conducted by Ntusi et al, it was shown that DCM mortality is still relatively high despite modern medical and surgical intervention (Ntusi, Badri et al. 2011), and that some of the therapies being implemented such as the administration of digoxin drugs to the DCM patients, have not improved survival rates among patients. In contrast, these therapies seem to have a mortality risk associated with them (Rathore, Curtis et al. 2003).

Cardiac transplantation is still a standard therapeutic option for advanced heart failure and was pioneered in South Africa (Beck 1978). Other therapies that have been proposed include palliative pericardiotomy for refractory heart failure (Lewis, Barnard et al. 1973) as well as treatment with thiamine and nicotinamide (Seftel 1973). The investigation of pentoxifylline for heart failure was perhaps the most standout input of South African investigators in the DCM treatment; this drug suppresses the immune response to DCM, particularly Tumor necrosis factor alpha60 (TNF-α60) and C-reactive protein (CRP) (Sliwa, Woodiwiss et al. 2004). While pentoxifylline has shown promise, a systematic review of these studies – which have been the

only randomised studies on the drug – concluded that although there was a trend towards reduced mortality, the results were not statistically significant, and larger trials were needed (Batchelder and Mayosi 2005).

1.2.2 Arrhythmogenic Right Ventricular Cardiomyopathy (ARVC)

ARVC is defined by the presence of right ventricular dysfunction which may be global or regional, with or without left ventricular disease, in the presence of histological evidence for the disease or electrocardiographic abnormalities in accordance with published criteria (McKenna, Thiene et al. 1994). Unlike DCM, HCM, and RCM, ARVC is defined histologically by the presence of progressive replacement of right ventricular myocardium with adipose and fibrous tissue (Figure 1.3). This fibrofatty replacement is often confined to a 'triangle of dysplasia' comprising the right ventricular inflow, outflow, and apex. While these pathologic abnormalities can result in functional and morphological right ventricular abnormalities, they also produce a DCM phenotype when they manifest in the left ventricle. They can also be present in the absence of clinically detectable structural changes in either ventricle.

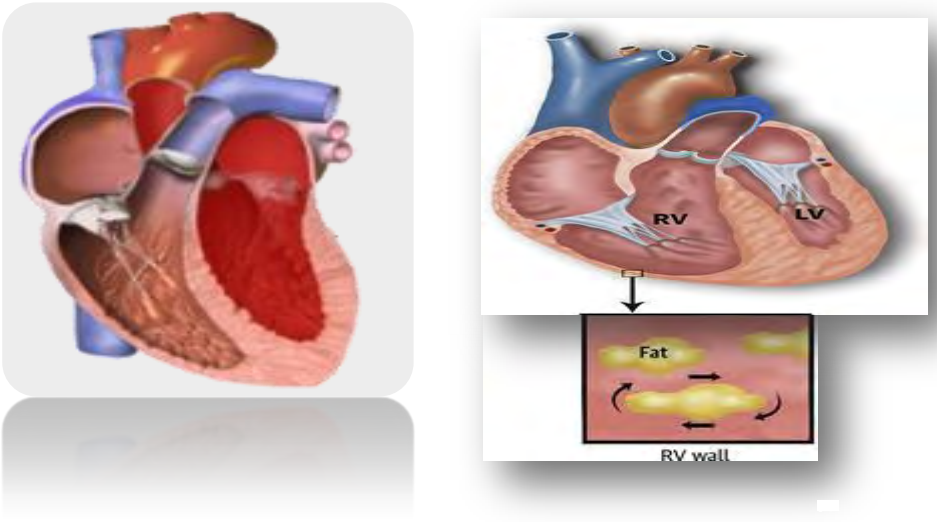


Figure 1.3: Diagram comparing a normal heart on the left with an ARVC heart on the right. The ARVC heart has fibrofatty replacement as sites of myocyte detachment (Link, Laidlaw et al. 2014).

ARVC usually starts causing symptoms by the third decade of life, with males manifesting symptoms of the disease more often than females (Corrado and Thiene 2006). Common presentations among patients include palpitations, dizziness and syncope. About a third of South African patients also experience chest pain, which makes ARVC an important consideration for clinicians evaluating young people with chest pain (Hendricks, Watkins et al. 2010).

The original 1994 International Task Force Criteria (TFC) for the clinical diagnosis of ARVC were based on structural, histological, electrocardiography (ECG), arrhythmic, and familial features of the disease (McKenna, Thiene et al. 1994). Abnormalities were subdivided into major and minor groups according to the specificity of their association with ARVC. Furthermore, the 1994 criteria focused on right ventricular (RV) disease manifestations and stipulated the absence of or only mild left ventricular (LV) involvement because of the need to exclude common disorders such as ischemic heart disease and dilated cardiomyopathy (Marcus, McKenna et al. 2010). Thus, a definite diagnosis of ARVC, according to the 1994 task force criteria (McKenna, Thiene et al. 1994) was made based on two major criteria or one major and two minor criteria or four minor criteria from different categories. Borderline diagnosis of ARVC was dependent on one major criterion and one minor or three minor criteria from different categories while a possible diagnosis was called in the event of one major criterion or two minor criteria from different categories. ARVC diagnosis can be a challenge because its symptoms can be discreet for years during its earlier phases, which means that a close follow up with high index of suspicion and repeat evaluations is needed.

Even in the South African cohort, who typically represented with advanced disease, 6% of subjects had a normal resting ECG at their last follow-up visit. Among the symptomatic patients, 12% of them had no abnormalities at all on cardiac imaging and their diagnosis was based on family history, electrocardiographic abnormalities and molecular genetic tests; nevertheless, this is not a benign cardiomyopathy (Hendricks, Watkins et al. 2010) .

An ARVC study conducted in South Africa by the ARVC Registry in South Africa gave practical implications of the disease (Watkins, Hendricks et al. 2009). Several Registry participants had

died by the end of the follow-up period (a median of around eight years). The annual mortality in the study was 2.8%, and the five-year cumulative mortality was 10%. On average, patients died two decades earlier than patients in France (Hulot, Jouven et al. 2004), thus the outcomes were worse compared to other parts of the world. There are several explanations for this, the most likely of which is that implantable cardioverter-defibrillators (ICDs), which are life-saving for ARVC patients, are under-utilized in South Africa. Furthermore, patients cannot always access the correct diagnosis and management of their disease.

Although ARVC is uncommon, with an estimated prevalence of 1 in 5000 individuals, it is a frequent cause of sudden death in young people accounting for 11% of all cases and 22% of cases among athletes (Corrado, Basso et al. 1998, Tabib, Loire et al. 2003). In South Africa, the disease is present in all ethnic and racial groups and according to a study by Watkins et al. the cohort, consisting of 50 patients with definite ARVC, had the following ethnic variation: 80% white (40 patients), 10% mixed ancestry (5 patients), 8% black African (4 patients), and 2% Indian (1 patient) (Watkins, Hendricks et al. 2009).

In two out of every three patients who died in the aforementioned ARVC study, arrhythmias were the cause of death, and to date no Registry participant who has received an ICD has died. From a practical outlook, these data advocate for the early use of ICDs in patients with syncope and sustained ventricular tachycardia as a way to prevent early mortality. Any patients with these symptoms and an ARVC diagnosis should be given top priority to be evaluated by a cardiologist (Hendricks, Watkins et al. 2010).

The natural history of ARVC is poorly understood and thus it is difficult to define a definite therapeutic algorithm. The treatment is essentially observational and is therefore based on presentation. For primary prevention in high-risk patients and secondary prevention of sudden cardiac death in patients with sustained ventricular tachycardia or ventricular fibrillation, an ICD is recommended. The goal of therapy is to reduce the frequency and severity of arrhythmias for symptomatic relief. Beta-blockers are generally considered the first line of drug therapy. Some of the antiarrhythmic agents include flecainide, propafenone, sotalol, and amiodarone, alone or in combination. Sotalol is the most effective drug for inducible or noninducible ventricular

tachycardia. In young patients amiodarone may not be convenient because of side effects associated with long-term use (Wichter, Borggrefe et al. 1992).

Pharmacotherapy is also used as an adjunct to ICD in selected patients with frequent life-threatening arrhythmias. Lifestyle modification in form of avoiding vigorous exercise in affected persons is recommended as activity can trigger arrhythmias. It is also advisable to avoid cardiac stimulants such as caffeine and pseudoephedrine (URL6).

1.2.3. Hypertrophic Cardiomyopathy (HCM)

HCM is characterised macroscopically by left ventricular hypertrophy (LVH) in the absence of a discernible cause, such as hypertension or aortic stenosis (Figure 1.4) and myocyte and myofibrillar disarray for histological assays (Mayosi 2005).

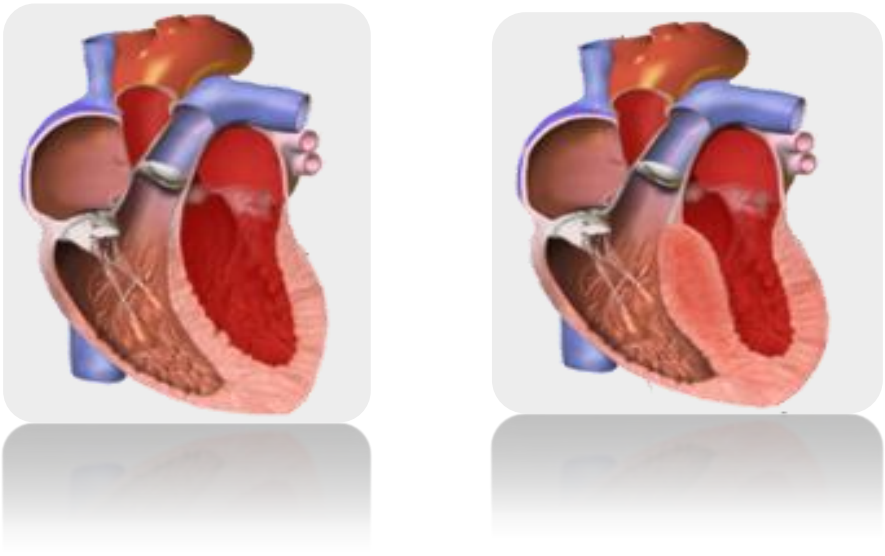


Figure 1.4: Comparison of a normal heart (left) with a HCM heart (right). In the cross sectional view above, the HCM heart presents with increased left ventricular wall thickness (URL13).

With the LVH there is an enlargement of the heart muscle notably at the interventricular septum leading to cardiac dysfunction (Green, Wakimoto et al. 2016). The LVH occurs in a non-dilated ventricle in the absence of other cardiac or systemic disease that might lead to the observed magnitude of increased LV wall thickness, such as pressure overload caused by

persisting hypertension or storage disorders including amyloidosis and Fabry disease (Cirino and Ho 1993). The clinical manifestations of HCM range from asymptomatic LVH, progressive heart failure to sudden cardiac death (SCD), and vary even amongst individuals within the same family (Cirino and Ho 1993).

The diagnosis of HCM depends on an abnormal thickening of the heart, but the earliest signs of disease are hyperdynamic contraction and impaired relaxation (Green, Wakimoto et al. 2016). Its diagnosis is most often established with noninvasive cardiac imaging, including echocardiography and/or cardiac magnetic resonance imaging (cardiac MRI) (Cirino and Ho 1993). The diagnosis can also be made by pathognomonic histopathologic findings in cardiac tissue, including myocyte disarray and fibrosis (Cirino and Ho 1993). In a familial involvement HCM diagnosis can be made through family history and molecular genetic testing, focusing on genes that encode different components of the sarcomere (Cirino and Ho 1993).

In general medical practice the disease may affect as many as 1 in 500 individuals in the general population (Mayosi 2005) with a prediction of approximately 600,000 persons with HCM in the United States (US) making it one of the most common monogenic cardiovascular disorders (Cirino and Ho 1993). It is the third most common form of cardiomyopathy in Ghana, after DCM and endomyocardial fibrosis (EMF) (Amoah and Kallen 2000). In Ethiopia, it accounts for 34% of all cardiomyopathies diagnosed at echocardiography (Abegaz 1990).

The aim of treatment is to relieve symptoms and prevent SCD in people at high risk. The severity of symptoms determines the specific treatment to be administered. Some of the available treatment options include medication to slow the heart rate and control heart rhythm, septal myectomy and ablation to improve the symptoms and ICD to continuously monitor and restore heart rhythm. The ICD option is applicable to patients who are at risk of SCD because of abnormal heart rhythm (URL4).

1.2.4. Restrictive Cardiomyopathy (RCM)

RCM is a rare disease of the myocardium in which the principal abnormality is diastolic dysfunction—specifically, restricted ventricular filling. Patients typically have increased left ventricular diastolic stiffness, which may lead to diastolic heart failure (Figure 1.5).

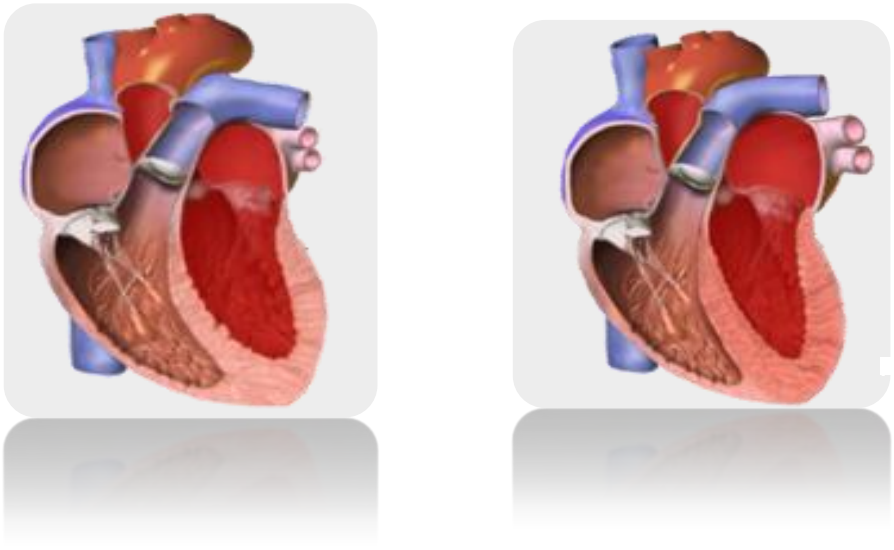


Figure 1.5: A diagrammatic comparison of a normal heart, on the left, with a restrictive cardiomyopathy heart. The restrictive cardiomyopathy heart has stiff ventricles but not necessarily thickened (URL14).

The left ventricle cannot fill sufficiently at normal diastolic pressures, leading to reduced cardiac output due to reduced left ventricular filling volume. Systolic function usually remains normal, at least early in the disease. Wall thickness may be increased secondary to myocardial infiltration (e.g., in amyloidosis) or the left ventricular wall is stiff due to fibrosis (e.g. EMF).

The clinical presentation of RCM may be confused with that of constrictive pericarditis because patients might actually have RCM while presenting with symptoms typical of constrictive pericarditis (URL6). In respect to history and clinical profile, the two may be indistinguishable. In addition, both conditions can coexist in the same patient; for example, radiation therapy affects the myocardium as well as the pericardium. The distinction between the two conditions may be made on echocardiography, cardiac catheterization and endomyocardial biopsy (URL6).

RCM accounts for approximately 5% of all cases of primary heart muscle disease (URL6) and is difficult to define because restrictive ventricular physiology is present in a wide range of different pathologies. The exact prevalence of RCM is unknown but it is probably the least common type of cardiomyopathy (Elliott, Andersson et al. 2008). In Africa, EMF, a form of RCM is a relatively common cause of heart failure in equatorial Africa (Silverman and Aksut 2015).

The cause of RCM determines its progression and treatment is often unsatisfactory. It may be idiopathic, familial, or result from a systemic disorder, such as amyloidosis. Diuretics, vasodilators, angiotensin-converting enzyme inhibitors, and anticoagulation may be indicated for managing restrictive cardiomyopathy (URL6). Prognosis is generally poor with progressive deterioration. Patients who are refractory to supportive therapy usually die of low-output cardiac failure unless cardiac transplantation is performed.

1.3. Molecular genetics of cardiomyopathy

The aetiology of the cardiomyopathies had been unknown but various genetic abnormalities associated with cardiomyopathy have been unraveled. Candidate gene approaches which focused on the genes encoding proteins which interact with products of the previously identified disease genes have been successful in identifying novel disease genes (Kimura 2016). As shown in Table 1.1 many different disease genes been identified to date. The overlapping of disease genes for different clinical types is noteworthy (McNally, Golbus et al. 2013).

Table 1.1: Disease genes for hereditary cardiomyopathy (McNally, Golbus, & Puckelwartz, 2013)

Condition	Gene
HCM/DCM/LVNC	Myosin-7 *(MYH7)*
DCM/HCM/ARVC	Titin *(TTN)*
HCM/DCM/RCM	Cardiac Troponin T type 2 *(TNNT2)*
DCM	Titin cap *(Tcap)*
DCM/HCM	Alpha-actinin *(ACTN)*
ARVC/DCM	Desmoplakin *(DSP)*
DCM/LVNC	Lamin A/C (*LMNA)*
DCM/LVNC	Tafazzin *(TAZ)*
DCM/HCM	Cysteine and glycine-rich protein 3 (*CSRP3*)
DCM/HCM/LVNC	Alpha-actin cardiac muscle 1 (*ACTC1)*
RCM/DCM/HCM	Troponin I, cardiac muscle *(TNNI3*)

LVNC-Left ventricular non-compaction

1.3.1 DCM

Only a minority (about 30%) of patients with DCM have evidence of familial segregation (Grunig, Tasman et al. 1998). Moreover, DCM is a feature of syndromic disorders, often with accompanying skeletal and limb-girdle myopathies. It was this co-existence of DCM and skeletal myopathy in Duchenne's and Becker's muscular dystrophies that lead to the discovery of dystrophin defects as a cause of X-linked DCM without obvious skeletal involvement, (Franz, Muller et al. 2000). This co-existence also lead to the consequential speculation that familial DCM is a 'cytoskeletopathy' (Towbin 1998).

This proposal was also reinforced by additional studies in other skeletal myopathy DCM phenocopies, which implicated more proteins that make up the cytoskeleton, the internal structure of the cell. It was not only the internal cytoskeleton that was responsible, because

proteins that form part of the extracellular matrix function in cell-to-cell contact at myocyte junctions have also been implicated in DCM. Proteins such as lamin A/C and emerin that stabilize the membrane around the cellular nucleus and those that connect these elements were also found to be defective in patients with dilated hearts (Figure1.6)

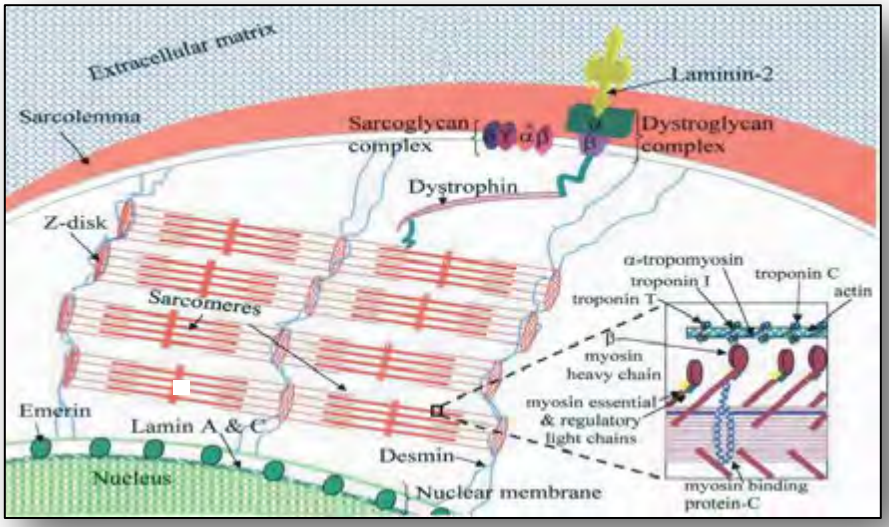


Figure 1.6: Cellular localization of proteins involved in cardiomyopathy (Towbin and Bowles 2002)

Up to one-third of individuals who have DCM inherit it from their parents but often its cause is unknown. Of the known cases, mutations in cytoskeletal, intercalated disc protein, sarcomeric protein, Z-band and nuclear membrane genes are responsible for autosomal dominant forms of the disease (Table 1.2). Patients with mitochondrial cytopathies and inherited metabolic disorders such as haemochromatosis may also have DCM (American Heart Association 2016).

Table 1.2: Genes associated with DCM

Condition	Gene	Proportion (%) of inherited cardiomyopathy caused by pathogenic variants of genes
Autosomal dominant DCM	Titin (*TTN*) (Gerull, Gramlich et al. 2002)	10-22
	Prelamin A/C (*LMNA)* (Fatkin, MacRae et al. 1999)	6
	Myosin-7 (Kamisago, Sharma et al. 2000)	4.2
	Myosin-6 (Carniel, Taylor et al. 2005)	4.3
	Sodium channel protein type 5 subunit alpha (*SCN5A*) (McNair, Ku et al. 2004)	2-4
	Myosin-binding protein C, cardiac-type (*MYBPC3*) (Daehmlow, Erdmann et al. 2002)	2-4
	Troponin T, cardiac muscle (*TNNT2*) (Hanson, Jakobs et al. 2002)	2.9
	Ankyrin repeat domain-containing protein 1 (*ANKRD1*) (Duboscq-Bidot, Charron et al. 2009)	2.2
	Vinculin (*VCL)* (Olson, Illenberger et al. 2002)	1
X-linked	Dystrophin (*DMD*) (Towbin, Hejtmancik et al. 1993)	Still Unknown
	Tafazzin (*TAZ*) (Bione, D'Adamo et al. 1996)	Still Unknown
Autosomal recessive DCM	Troponin I, cardiac muscle (*TNNI3*) (Murphy, Mogensen et al. 2004)	<1

The discovery that a number of cytoskeletal proteins form substrates for proteases expressed by viruses known to cause cardiac dilation (Shoeman, Kesselmier et al. 1991, Badorff, Berkely et al. 2000) may indicate that the causative principle involved in DCM may be instability at any structural point throughout the integrated organization of the cardiac syncytium.

Several of the sarcomeric protein-encoding genes originally implicated in HCM (Section *1.3.3*) have also been found to be defective in some DCM cases, making the etiological distinction between these two disorders unclear (Olson, Michels et al. 1998, Gerull, Gramlich et al. 2002). It may be that these particular mutations cause DCM rather than HCM because of their different involvement in the functional domains of these proteins, or simply because the sarcomere itself inherently also forms part of the integrated internal structure of these cells.

To add to this complexity, DCM is seemingly also not purely a disease of cell architecture, but that energy metabolism could play a major role as well, as is now postulated for HCM. It has long been known that mitochondrial DNA defects are associated with DCM (Grasso, Diegoli et al. 2001, Khogali, Mayosi et al. 2001). Morphometric studies of mitochondria in cardiac biopsies from DCM and HCM has produced data which suggested that the mitochondria of DCM hearts showed decreased activity, while those from HCM hearts showed increased activity (Tashiro, Masuda et al. 1990). An establishment of whether these mitochondrial defects are the cause of the cardiac phenotype, or a consequence of it, has been difficult to prove.

In Africa, several gene association studies have been conducted, which suggest that heredity may be influential in the susceptibility to DCM in Africans (Mayosi 2007). An association with Human Leukocyte Antigen - antigen D Related (HLA-DR1) has been reported in South African patients which suggest that genetically determined immune-response factors have a role in the pathogenesis of some people with DCM (Maharaj and Hammond 1990). A common mitochondrial DNA polymorphism c.16189T>C has also been found to be a genetic risk factor for DCM in a South African cohort, with a population attributable risk of 6% (Khogali, Mayosi et al. 2001). These genetic associations suggest that they may represent genetic risk factors for DCM worldwide as they have been replicated in other populations (Marin-Garcia, Zoubenko et al. 2002, Rodriguez-Perez, Fragoso et al. 2007). Mutation screening studies in South African patients with idiopathic and familial DCM have identified families with early onset DCM caused by a known mutation in the troponin T gene (c.471C>T) (Mayosi 2007) as well as in the Phospholamban (*PLN)* gene (Fish et al 2016), but did not show mutations in the cardiac and skeletal actin genes (Mayosi, Khogali et al. 1999).

1.3.2 ARVC

A familial background has been demonstrated in over 50% of ARVC cases (Corrado and Thiene 2006). The disease is usually inherited as an autosomal dominant trait with incomplete penetrance and variable expression. In 1994 the first chromosomal locus (14q23-q24) was associated with ARVC after clinical evaluation of a large Italian family (Rampazzo, Nava et al. 1994). Afterwards, linkage analysis provided evidence for genetic heterogeneity with sequential discovery of several ARVC loci on chromosomes 1, 2, 3, 6, 10, 12, and 14.

Autosomal recessive forms of ARVC, for example Naxos and Carvajal syndromes are recognized, but the majority of cases are caused by autosomal dominant inherited mutations in genes encoding plakophilin 2 (*PKP2*) and other proteins of the cardiomyocyte desmosome. Mutations in transforming growth factor-beta and ryanodine receptor genes may be associated with an ARVC phenotype (Table 1.3) In Naxos disease, there is a cosegregation of cardiac (ARVC), skin (palmoplantar keratosis), and hair (woolly hair) abnormalities which have been mapped on chromosome 17 (locus 17q21)(Rampazzo, Nava et al. 1994).

Table 1.3: Genes associated with ARVC (Haugaa, Haland et al. 2016)

Gene	Proportion (%) of inherited cardiomyopathy caused by pathogenic variants of genes
Desmoplakin (*DSP*) (Andreasen, Nielsen et al. 2013)	2-12
Desmoglein 2 (*DSG2)* (Ackerman, Priori et al. 2011)	5-10
Plakophilin 2 (*PKP2)* (Ackerman, Priori et al. 2011)	25-40
Ryanodine receptor 2 (*RYR2)* (Andreasen, Nielsen et al. 2013)	Unknown
Plakoglobin (*JUP)* (Andreasen, Nielsen et al. 2013)	Unknown
Transmembrane protein 43 (*TMEM43)* (Andreasen, Nielsen et al. 2013)	Unknown
Transforming growth factor-beta (*TGF- ß*) (Ackerman, Priori et al. 2011)	Unknown

Endurance exercise is a known risk factor of ARVC; in fact, 28% of patients in a South African ARVC cohort were professional endurance athletes at some point in their lives (Hendricks,

Watkins et al. 2010). Previous human and animal studies have demonstrated that, if an individual is genetically predisposed to ARVC, endurance exercise is an important environmental factor that leads to ARVC development (Heidbuchel, Hoogsteen et al. 2003, Kirchhof, Fabritz et al. 2006). Because of this, it is important that endurance athletes are screened for heart disease, before and during their sporting careers.

Initial observations suggest that gap junction function is altered by the mechanical defect of the desmosomes. Evidence of myocyte loss or clinical evidence of right ventricular dysfunction may present before electrocardiographic (ECG) changes and arrhythmias (Kaplan, Gard et al. 2004). It has been proposed that similar clinical phenotypes occur that are based on disruption of a 'final common pathway' by mutations in genes encoding proteins in the defined desmosomal pathway (Vatta, Marcus et al. 2007). Examination of the pathogenesis in relation to arrhythmogenesis and disease progression is facilitated by identifying the genetic basis of ARVC (Tsatsopoulou, Protonotarios et al. 2006).

Seven genes have been identified in association with ARVC: plakoglobin (*JUP*) (McKoy, Protonotarios et al. 2000), desmoplakin (*DSP*) (Norgett, Hatsell et al. 2000), plakophilin-2 (*PKP2*) (Gerull, Heuser et al. 2004), desmoglein-2 (*DSG2*) (Awad, Dalal et al. 2006, Pilichou, Nava et al. 2006), desmocollin-2 (*DSC2*) (Syrris, Ward et al. 2006), transforming growth factor beta-3 (*TGFß3*) (Beffagna, Occhi et al. 2005) and transmembrane protein 43 (*TMEM43*) (Merner, Hodgkinson et al. 2008). In patients with an arrhythmic presentation, mutations in ryanodine receptor 2 (*RYR2*) have been reported in ARVC in the absence of significant electrocardiographic or structural abnormalities. At present, catecholaminergic polymorphic ventricular tachycardia is considered a disorder distinct from ARVC (Awad, Calkins et al. 2008).

McKoy and colleagues identified the first disease-causing gene, *JUP*, in patients with Naxos disease. The gene encodes the desmosomal protein plakoglobin which is a major constituent of the cell adhesion junction. Its discovery suggested that ARVC is a cell-to-cell junction disease and stimulated the research of other related genes (McKoy, Protonotarios et al. 2000).

DSP was the first desmosomal protein gene to be associated with the more common autosomal dominant form of ARVC by Rampazzo et al (Rampazzo, Nava et al. 2002). Gerull et al (2004)

identified 25 different mutations in the gene encoding *PKP2*. Recently, *DSG2* mutations have been found in 10% of ARVC unrelated probands (Pilichou, Nava et al. 2006).

Autosomal dominant ARVC has been linked to other genes unrelated to cell adhesion complex, such as the gene *RyR2*and *TGFβ3* the latter of which is responsible for regulation of the production of extracellular matrix components and modulates expression of genes encoding desmosomal proteins (Corrado and Thiene 2006).

It is still unclear how the mutations of *PKP2* and more broadly of desmosomal protein genes cause disease. It has been hypothesized that the lack of normal protein or the incorporation of mutant protein into cardiac desmosomes may invoke myocyte detachment at the intercalated discs, particularly under conditions of mechanical stress, like that occurring during competitive sports activity (McKoy, Protonotarios et al. 2000). Consequently, there is a progressive myocyte death with subsequent repair by fibrofatty replacement. Life-threatening ventricular arrhythmias may occur either during the accelerated phase of myocyte death as abrupt ventricular fibrillation or later in the form of scar-related macro-reentrant ventricular tachycardia. To date, it is still unclear why myocytes are replaced by fibro-fatty tissue, or why the right ventricle is most severely affected. However, It is clear from molecular genetic studies in ARVC families that the genes implicated so far are not the only ones responsible for this cardiac phenotype, and it is possible that with identification of more ARVC-causing genes, these two features of the disease may become more readily understood (Moolman-Smook, Mayosi et al. 2003).

Genetic testing is designated for symptomatic patients with ARVC and family members of a patient with a positive mutation. Relatives with more than one genetic variant are at a higher risk of developing clinical disease, potentially an important determinant of the phenotypic heterogeneity seen within families with ARVC (Quarta, Muir et al. 2011). At present, genetic testing is available for seven types of abnormalities (*DSC2, DSG2, DSP, JUP, PKP2, RYR2* and *TMEM43*) (URL 7).

Because ARVC runs in families, there is an urgent need to screen the first-degree relatives of affected individuals. Death is potentially preventable if there is early diagnosis of the condition

and high-risk individuals are promptly referred for consideration for ICD implantation. Genetic testing yields an answer in one out of four cases and is available in South Africa through the ARVC Registry at the University of Cape Town (Hendricks, Watkins et al. 2010). Desmosomal gene mutation is seen in 40%-50% cases and the clinical application of genetic testing is limited owing to multiple gene mutations and variable penetrance, requiring long-term follow-up (Sen-Chowdhry, Syrris et al. 2007, Cowan, Morales et al. 2008) .

The ARVC registry of South Africa report has made several key observations about the genetic causes of ARVC. Firstly, 30% of participants in the study had a family member who was also affected. For the purposes of the *HeartRhythm* report, genetic analysis was focused on the most common causative gene, *PKP2*. Of the DNA analyzed, 25% had a *PKP2* gene defect that caused disease (Hendricks, Watkins et al. 2010); one mutation in particular, *PKP2* c.1162C>T, was present in four South African probands of European origin which suggested a founder effect (Watkins et al., 2009).

1.3.3 HCM

HCM is recognized as an autosomal dominant disorder that is caused by more than 1400 different mutations in at least eight different genes that code for sarcomeric proteins (Adalsteinsdottir, Teekakirikul et al. 2014) (Table 1.4).

Table 1.4: Genes that cause HCM

Gene	Proportion (%) of inherited cardiomyopathy caused by pathogenic variants of genes
Myosin-7 (*MYH7*) (Marian and Roberts 2001)	35-50
Myosin-binding protein C, cardiac type (*MYBPC3*) (Marian and Roberts 2001)	20-25
Troponin T, cardiac muscle (*TNNT2*) (Richard, Charron et al. 2003)	4
Troponin I, cardiac muscle (*TNNI3*) (Marian and Roberts 2001)	20
Tropomyosin alpha-1 chain (*TPM1*) (Marian and Roberts 2001)	5
Myosin light chain 3 (*MYL3*) (Marian and Roberts 2001)	5
Myosin-6 (*MYH6*) (Jiang, Wakimoto et al. 2013)	Unknown

Telethonin (*TCAP*) (Bos, Poley et al. 2006)	Unknown
Titin (*TTN*) (Bos, Poley et al. 2006)	Less than 5
Myozenin-2(*MYOZ2)* (Ruggiero, Chen et al. 2013)	Unknown

Earlier studies suggesting that only 50% of HCM was familial did not consider incomplete penetrance of the disease mutations, and that individuals may carry a causal mutation with mild or intermittently indiscernible HCM manifestations. With such carriers, their offspring are however still at a 50% risk of inheriting the mutation and may well manifest the disease (Mayosi 2005). The large families with multiple individuals clearly affected with uncomplicated HCM, as described in the early to mid-1900s (Teare 1958, Pare, Fraser et al. 1961) made this cardiomyopathy most amenable to molecular genetic analysis. Over the last decade, it has been demonstrated that HCM is a 'sarcomeropathy' (Thierfelder, Watkins et al. 1994). Autosomal disorders that present in the young include Noonan and LEOPARD syndromes (dominant) and Friedreich's ataxia (recessive) (Elliott and McKenna 2004).

Although extreme hypertrophy is still a predictor of poor prognosis (Spirito, Bellone et al. 2000), hypertrophy and the risk of sudden cardiac death (SCD) are unrelated features (Moolman, Corfield et al. 1997, Moolman-Smook, De Lange et al. 2000). In fact, it was found that defects in some of these sarcomeric protein-encoding genes, such as *TNNT2*, often cause minimal hypertrophy and a high risk of sudden death. Interestingly, these sarcomeric encoding genes have been associated with early SCD (Watkins, McKenna et al. 1995, Moolman, Corfield et al. 1997, Varnava, Baboonian et al. 1999, Moolman-Smook, De Lange et al. 2000). In a South African study, it was found that this was especially true in young male carriers of the *TNNT2*, c.8772C>T (R92W) mutation (Moolman, Corfield et al. 1997, Moolman-Smook, De Lange et al. 2000).

In addition, there have been indications that some of the morphological variants of hypertrophy are associated with specific genes or mutations. For instance, the *TNNT2*, c.8772C>T mutation has been associated, in Japanese patients, with the DCM-like variant with

early cardiac decompensation and progression from hypertrophy to dilation (Fujino, Shimizu et al. 2001).

Studies of the functional effects of these mutations indicated that the encoded faulty proteins interfere with the function and the structure of the sarcomere, and may directly give rise to widespread myofibrillar and myocytic disarray, characteristic of HCM (Marian, Yu et al. , Varnava, Elliott et al. 2001). These functional studies also revealed that most HCM-causing defects result in abnormal calcium (Ca^{2+}) sensitivity, (Redwood, Moolman-Smook et al. 1999) supporting the earlier observation of altered Ca^{2+} handling by HCM hearts (Paulus, Goethals et al. 1992, Schotten, Voss et al. 1999).

In South Africa, however, there are three recurring or founder mutations that have been found in 45% of genotyped patients with European and mixed ancestry (Cantlay, Shokrollahi et al. 1999). Consequently, South African patients with HCM referred for molecular diagnosis are initially screened for the three founder mutations (*MYH7* c.1208G>T, *MYH7* c.2389G>A, and *TNNT2* c.274C>T). Only in the absence of these founder mutations is extensive screening done (Moolman-Smook, Mayosi et al. 2003). Experience elsewhere in the world has elucidated numerous mutations in the sarcomeric protein genes, such that many families have a conserved mutation that is unique to them (Moolman-Smook, De Lange et al. 1999).

Our research group published a paper in 2016 describing the genetic features and outcome of HCM patients in South Africa whose clinical presentation, echocardiographic and electrocardiographic findings were obtained from the Cardiac Clinic at Groote Schuur Hospital; clinical presentation was similar to that reported in other studies (Ntusi et al., 2016). Targeted re-sequencing was performed on 43 HCM patients and included 15 genes known to cause HCM: cardiac myosin binding protein C (*MYBPC3*), cardiac β-myosin heavy chain (*MYH7*), *TNNT2*, cardiac troponin I (*TNNI3*), regulatory light chain of myosin (*MYL2*), essential light chain of myosin (*MYL3*), tropomyosin 1 (*TPM1*), phospholamban (*PLN*), α-actin (*ACTC1*), cysteine and glycine-rich protein 3 (*CSRP3*), AMP-activated protein kinase (*PRKAG2*), α-galactosidase (*GLA*), four-and-a-half LIM domains 1 (*FHL1*), lamin A/C (*LMNA*) and lysosome-associated membrane protein 2 (*LAMP2*). Ten of 35 index cases had disease-causing mutations in only *MYH7* (six cases

or 60%) and *MYBPC3* (four cases or 40%) but not in the other 13 genes screened. The genetic screening gave a 29% yield of causal genetic mutations in South African HCM cases (Ntusi et al., 2016).

Up until 2000, any causative connection between phenocopies of HCM, which do not feature sarcomeric disruption, and classic HCM remained elusive, when mutations in the 5^{l}-activated AMP protein kinase (AMPK) gene were found in individuals featuring HCM and Wolf-Parkinson-White syndrome (Blair, Redwood et al. 2001). The AMPK enzyme acts as the fuel gauge of the myocyte, sensing when adenosine triphosphate (ATP) levels are too low, and activating molecular pathways that lead to increased energy production (Winder and Hardie 1999, Hardie and Hawley 2001). This has led to the proposal that the common basic aetiological principle of cardiac hypertrophies, whether in HCM or in HCM-phenocopies, is related to an imbalance in energy supply and demand (Blair, Redwood et al. 2001). Whatever the primary cause of the energy inequality, prolonged decreased ATP levels will impede Ca^{2+} re-uptake from the cytoplasm into the sarcoplasmic reticulum by the Ca^{2+} ATPase which lead to Ca^{2+} related activation of hypertrophic and arrhythmic pathways(Spindler, Saupe et al. 1998, Somura, Izawa et al. 2001).

1.3.4 RCM

An autosomal dominant mode of inheritance is a hallmark of familial RCM, which in some families is caused by mutations in the *TNNI3*. In other families, RCM is caused by mutations in the desmin (*DES*) gene, normally associated with skeletal myopathy, which is linked to conduction defects. In rare cases, autosomal recessive inheritance can be seen in familial disease such as haemochromatosis caused by mutations in the Haemochromatosis (*HFE*) gene, or with X-linked inheritance (such as Anderson–Fabry disease) (Elliott, Andersson et al. 2008). *MYH7, TNNT2* and *TNNI3* mutations have been associated with RCM, HCM and DCM. The molecular basis of the differences between RCM and HCM associated mutations was that the RCM associated mutations showed much greater Ca^{2+} sensitization than the HCM associated mutations, as demonstrated for *TNNT2* (Pinto, Parvatiyar et al. 2008) and *TNNI3* (Yumoto, Lu et

al. 2005) mutations. In accordance with these findings, it was reported that restrictive phenotype (RCM-like HCM) was uncommon in HCM and may represent a poor prognosis form of HCM with severe diastolic dysfunction (Kubo, Gimeno et al. 2007). Alternatively, a gene dose effect could explain the difference between RCM-associated mutations and DCM-associated mutations. For example, a mutation in *TNNI3* was observed in association with DCM when homozygous (Mogensen, Kubo et al. 2003), but was found as a heterozygous mutation in RCM patients (Murphy, Mogensen et al. 2004).

1.4 Work leading to this thesis

In a study conducted by Dr Mzwandile Mbele through the Cardiovascular Genetics Laboratory at the Hatter Institute For Cardiovascular Research at the University of Cape Town, an ARVC proband (ACM 8.3) was screened for mutations in the desmosomal genes *PKP2, DSG2, DSP, DSC2 and JUP* (Mbele,2014), but no mutations were found in any of these genes.

This family was subsequently prioritized for whole exome sequencing by Dr Mbele as he set out to identify the genetic cause of the disease within this family. The whole exome sequencing experiment yielded a final candidate gene list which contained 13 variants but only two variants were reported to possibly be causative of ARVC. These variants occurred in the genes parvin alpha (*PARVA*) and high mobility group-box containing 3 (*HMGXB3*), neither of which had been reported in association with ARVC before. The *PARVA* c.392A>T (p.D131V) mutation was considered to be a more probable candidate gene for ARVC based on the role of the PARVA protein as a cell adhesion molecule, rare *PARVA* mutations observed in the general population (1 in 12227 population controls), evolutionary conservation of the amino acid change and predicted deleterious effect on protein function. The *HMGXB*3 variant was purported to play a modifying role based on the biological significance of *PARVA* as a plausible ARVC gene, thus it was not pursued further.

1.5. Hypothesis

A previous whole exome sequencing project completed in the United Kingdom (UK) had identified *PARVA* as a disease causing gene in a family with ARVC. We hypothesize that *PARVA* might be a new gene not only for ARVC but for other cardiomyopathies as well.

1.6. Aims

The aims of the study were to:

1. Validate the whole exome sequencing results obtained in the UK on the ACM8 family.
2. Screen for *PARVA* mutations in a large cohort of 180 probands diagnosed with ARVC, DCM, HCM and RCM.

CHAPTER 2: CANDIDATE GENE ANALYSIS OF *PARVA*

2.1. Introduction

In 2012, at the University of Newcastle upon Tyne in the United Kingdom (UK), Dr Mzwandile Mbele performed whole exome sequencing on two severely affected siblings on a South African family with ARVC, referred to as ACM 8. The proband, numbered ACM 8.3 and her younger brother ACM 8.4 (Figure 2.1) were diagnosed with ARVC at the ages of 10 years and 12 years, respectively. Their parents, ACM 8.2 (mother) and ACM 8.1 (father) as well as older brother (ACM 8.9) did not meet the diagnostic criteria for ARVC at the time. This family was prioritized for whole exome sequencing as no causative genetic mutation had been found on routine genetic screening of the five desmosomal genes known to cause ARVC (i.e. *PKP2*, *DSG2*, *DSP*, *DSC2* and *JUP*) (Watkins, Hendricks et al. 2009).

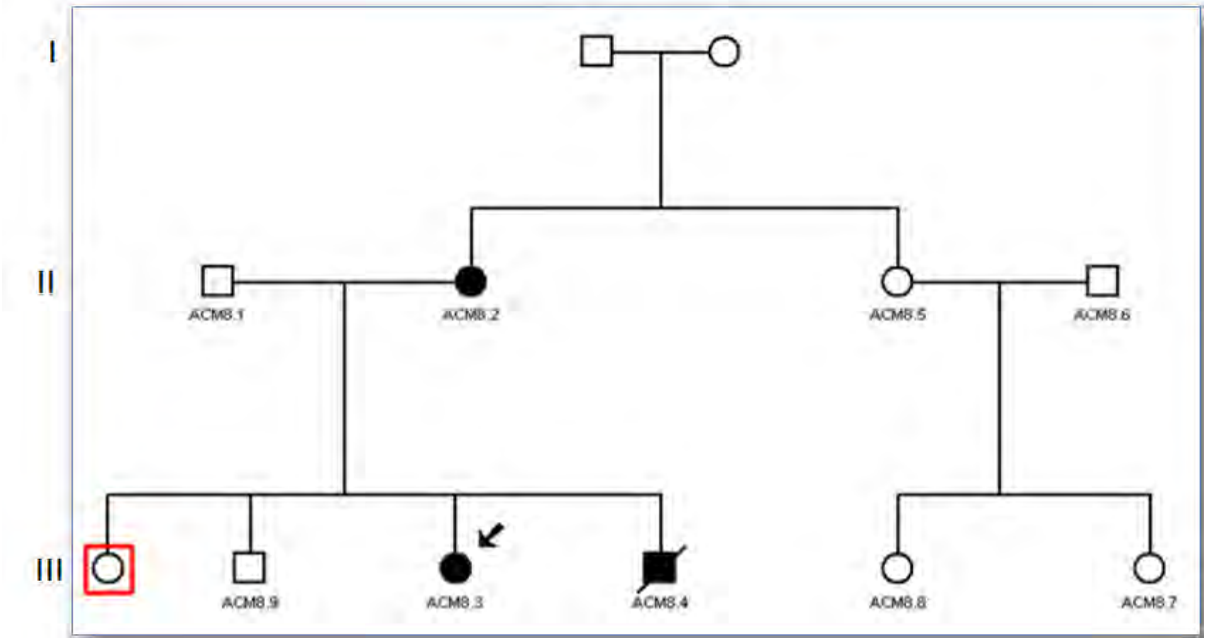


Figure 2.1: Pedigree of the ACM 8 family with ARVC. Circles represent females. Squares represent males. Clear circles and squares represent unaffected family members. Black circles and squares represent affected family members. Crossed symbols represent deceased members of the family. The circle in red represents a family member whose blood was not available for DNA analysis.

In his whole exome sequencing experiment, Mbele used the Illumina NGS platform and after the filtering processes, found that 13 variants (five homozygotes and eight heterozygous) (Table 2.1) were shared by the two affected individuals, ACM 8.3 and ACM 8.4. These variants

occurred in the genes *PARVA*, ankyrin 3 Node Of Ranvier (*ANK3*), anaplastic lymphoma receptor tyrosine kinase (*ALK*), deleted in lung and esophageal cancer 1 (*DLEC1*), WW domain containing adapter with coiled-coil (*WAC*), *HMGXB3*, neuralized homolog (Drosophila) (*NEURL*) and trichorhinophalangeal syndrome I (*TRPS 1*) while the homozygous variants were in the genes ubiquitin protein ligase E3 component N-recognin 4 (*UBR4*), SMAD family 5 (*SMAD5*), glutamate receptor ionotropic AMP3 (*GRIA3*) and zinc finger protein 141 (*ZNF141*). Of the five homozygous variants that were sequenced, four of the variants (*ZNF141* g.331694-331695insC, *SMAD5* g.135513085-135513086insC *GRIA3* g.6586096-65866097insG and *GRIA3* 122336601-122336602insG) were present in the unaffected sibling (ACM 8.9), and therefore did not segregate with disease, whereas the fifth variant (*UBR4* c.15009_229-15009_228 insG) was detected at a frequency of 52% in the white population controls of South Africa thereby relegating its status to that of a single nucleotide polymorphism (SNP). All homozygous variants were also reported in the dbSNP database. Of the eight heterozygous variants, six were present in the unaffected sibling (and therefore did not segregate with disease). Of the remaining two variants, (*HMGBX3* c.607G>A and *PARVA* c.392A>T) the rare *PARVA* c.392A>T variant (frequency 0.007% as per Exome Variant Server (EVS)) was considered the most plausible candidate for ARVC in this family while the *HMGXB3* c.607G>A variant was suggested to play a modifying role.

Table 2.1: Variants detected in two affected family members

Gene	Gene function	DNA change
Heterozygous variants		
Parvin, Alpha (*PARVA*)	Plays a role in the organization of the actin cytoskeleton; and cell adhesion	c.392A>T
Ankyrin 3, Node Of Ranvier (*ANK3*)	Maintenance of ion channels and cell adhesion	c.2978C>T
Anaplastic lymphoma receptor tyrosine kinase (*ALK*)	Plays an important role in the genesis and differentiation of the nervous system	c.4947G>A
Deleted in lung and esophageal cancer 1 (*DLEC1*)	Act as a tumor suppressor by inhibiting cell proliferation	c.262C>G
WW domain containing Adapter with coiled-coil (*WAC*)	Regulates the cell-cycle checkpoint Activation in response to DNA damage	c.1465A>G
HMG box domain containing 3 (*HMGXB3*)	Plays a role in DNA replication and repair	c.607G>A
Neuralized homolog (Drosophila) (*NEURL*)	May function as an E3 ubiquitin-protein ligase to activate monoubiquitination of JAG1	c.563T>C
Trichorhinophalangeal syndrome I *(TRPS 1)*	Transcriptional repressor	c.-9_-8insT
Homozygous variants		
Ubiquitin protein ligase E3 component N-recognin 4*(UBR4)*	It regulates integrin-mediated signalling.	c.15009_229-15009_228inG
SMAD family 5 (*SMAD5)*	It functions as a transcriptional modulator activated by bone morphogenetic protein.	c.1314_1316insC
Glutamate receptor ionotropic AMP3*(GRIA3)*	It functions as a glutamate receptor.	c.382_383insG
Zinc finger protein 141 (ZNF141)	It is involved in transcriptional regulation as a repressor	g.331694-331695insC

Mbele supported his hypothesis by describing the biological plausibility of the role of *PARVA* in cardiomyopathy as a cell adhesion protein, the rarity of *PARVA* mutations in the general

population, evolutionary conservation of the amino acid change and the predicted deleterious effect on protein function.

Mbele then proceeded to screen 66 additional ARVC patients for mutations in *PARVA* and reported a pathogenic non-synonymous variant, c.523A>G (p.K156E), in one ARVC case and a synonymous variant of unknown significance, c.597T>G (p.T199) in another ARVC case.

The *PARVA* gene encodes a PARVA protein which localizes to nascent focal adhesion sites which are sites of interaction between the cell and extracellular matrix (Nikolopoulos and Turner 2000). The PARVA protein forms a ternary complex (PIP) through interactions with integrin linked kinase (ILK) and Particularly interesting new cysteine-histidine-rich protein (PINCH).This complex has been implicated in the control of signalling pathways through phosphorylation of downstream targets, mostly protein kinase B (PKB) and glycogensynthase kinase3 (GSK3). The complex also binds to upstream effectors of the Jun Nterminal kinase (JNK) signalling pathway and regulators of small molecular weight guanine triphosphatases (GTPases) for efficient cell-extracellular matrix interactions for cell signalling (Legate, Montanez et al. 2006). PARVA also binds to a-actinin and Rho GTPase activating protein 31(CdGAP), which clarifies the prominent functions of PARVA in integrin-mediated cell adhesion and actin-dependent processes such as cell shape regulation and cell migration (Sopko, Qin et al. 2011)

The PARVA protein encoded by the *PARVA* gene is 42kDa composed of 13 exons and the most distinct feature is the presence of two calponin homology (CH) domains in their C-terminal regions, named CH1 and CH2 (Sepulveda and Wu 2006) (Figure. 2.2).

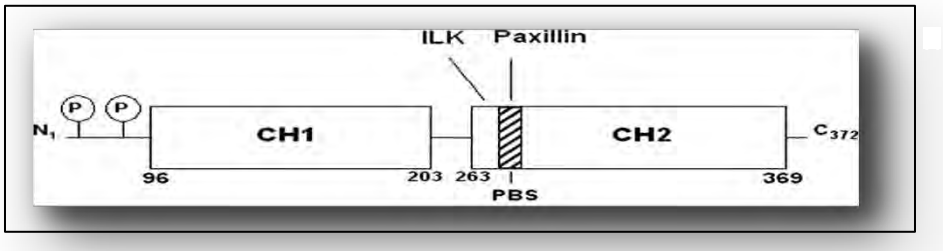


Figure 2.2: Structure of PARVA protein (Sepulveda & Wu, 2006)

Each CH domain is an area of actin-binding and comprises approximately 100 amino acids, separated by a 60-amino acid linker. The N-terminal with 95 amino acids is a major site of phosphorylation which confers conformational change to the whole protein and it has nuclear localization signals as well as src-homology 3 domain (SH3) binding sites of the consensus PXXP, where X is any amino acid but cysteine (Olski, Noegel et al. 2001). The CH2 domain contains paxillin binding site (PBS) that interacts with an integrin-linked kinase.
For this project we continued to explore the role of *PARVA* as a candidate not only for ARVC but for other cardiomyopathy cohorts (DCM, HCM and RCM) as well.

2.2. Aims

The aims of this study were:

1.) To screen for the *PARVA* c.392A>T variant in all members of the ACM8 family to determine segregation with disease.

2.) To screen the cardiomyopathy cohorts (ARVC, DCM, HCM and RCM) consisting of 180 probands for pathogenic *PARVA* variants.

2.3. Methods

2.3.1. ACM 8 Family

All available family members of ACM 8 (ACM 8.1, 8.2, 8.3, 8.4, 8.5, 8.6, 8.7,8.8 and 8.9) were enrolled in the ARVC Registry of South Africa which was established by the Working Group on Registries of the Cardiac Arrhythmia Society of Southern Africa (CASSA), as reported previously (Latib, Michaels et al. 2004). The registry is approved by the Human Research Ethics Committee of the Faculty of Health Sciences of the University of Cape Town (Appendix 1).

The clinical phenotyping of this ACM8 family indicated that both affected children developed severe, early-onset ARVC with the proband presenting at 12 years and requiring a heart transplant at age 23 years. Unfortunately, the proband's affected younger brother died at age

14 years. The proband's mother, ACM8.2, had a later age of onset and displayed mild symptoms of ARVC. We have not phenotyped the father as he has not yet responded to invitations to visit Groote Schuur Hospital.

2.3.2 Cardiomyopathy Cohort

2.3.2.1 ARVC

A diagnostic panel consisting of a group of cardiologists determined, by consensus, whether or not the referred cases met the diagnostic criteria for ARVC set by the Task Force of the Working Group on Myocardial and Pericardial Diseases of the European Society of Cardiology and of the Scientific Council on Cardiomyopathies of the International Society and Federation of Cardiology (McKenna, Thiene et al. 1994). The participants were classified as having definite ARVC (if Task Force criteria were met), probable or possible ARVC (if some criteria were met and no alternative diagnosis was found), or not having ARVC (if there was no evidence of ARVC and/or an alternative diagnosis was present). The diagnosis of ARVC in first-degree relatives of affected individuals was made based on the modified criteria of Hamid and colleagues (Hamid, Norman et al. 2002). For this investigation, 74 ARVC samples were screened.

2.3.2.2 DCM

DCM diagnosis was made on the basis of the definition of the European Society of Cardiology (Elliott *et al*. 2008). For this investigation, 68 DCM samples were screened.

2.3.2.3 HCM

HCM diagnosis was made on the basis of the definition of the European Society of Cardiology (Elliott *et al*. 2008). For this investigation, 28 HCM samples were screened.

2.3.2.4 RCM

RCM diagnosis was made on the basis of the definition of the European Society of Cardiology (Elliott *et al*. 2008). For this investigation, 10 RCM samples were screened.

2.3.3 Normal controls

Controls were chosen from healthy, anonymous blood donors with no history of heart disease from the Western Province Blood Transfusion Service (WPBTS) provided blood samples for DNA isolation. The 232 anonymous blood donors were from 99 people of mixed ancestry, 62 black Africans, 11 people of Indian origin and 60 white South Africans. Informed consent was obtained from all participants (Appendix 3).

2.3.4. DNA extraction

Peripheral blood samples were collected from participants in 5 ml ethylenediaminetetraacetic acid (EDTA) tubes. The genomic DNA (gDNA) was extracted from peripheral blood using the PureGene™ DNA isolation kit (Gentra system, USA) and was extracted according to the manufacturer's instructions (Appendix 4). The gDNA samples were given disease code numbers (ACM to refer to the disease ARVC) to anonymize them and were filed in the database of the Cardiovascular Genetics Laboratory at the Hatter Institute for Cardiovascular Research in Africa, Department of Medicine, University of Cape Town. Extracted DNA samples were stored at -80°C for long term storage.

2.3.5. DNA quality control

2.3.5.1 Nanodrop quantification

All the DNA samples were mixed prior to measurement and were quantified at an absorbance of 260 nm using a NanoDrop ND-2000 spectrophotometer (ThermoScientific, UK). Each sample was measured twice for confirmation of its concentration, and the average of the two measurements was taken as the concentration of each sample. All primary DNA extractions

were also validated using the Qubit™ Fluorometer (Invitrogen, USA) which is considered an accurate determinant of nucleic acid concentrations.

2.3.5.2 Gels

The DNA samples were electrophoresed on an agarose gel to confirm the presence of good quality DNA.

Agarose is composed of repeated agarobiose (L- and D-galactose) subunits whose polymers form a non-covalent association and to give a network of bundles with pore sizes that determine a gel's molecular sieving properties. For the DNA separation in agarose gel electrophoresis, the DNA is loaded into pre-cast wells in the gel and a current applied. The DNA phosphate backbone is negatively charged, thus when placed in an electric field, DNA fragments migrate to the positively charged anode. Because DNA has a uniform mass/charge ratio DNA molecules are separated by size within the agarose gel in a pattern such that its molecular weight is inversely proportional to the distance it traveled (Lee, Costumbrado et al. 2012).

The migrated DNA in the gel was stained by GelRed^{TM} (Biotium, US), a fluorescent nucleic acid stain, for visualization of the DNA. GelRed^{TM} intercalates between the base pairs of the DNA strands and is visualized under ultraviolet (UV) light. The loading of the DNA and the monitoring of electrophoretic progression were assisted with the electrophoresis buffer, 3µl of 1X loading dye with a DNA ladder (New England Biolabs^{R} 100bp DNA ladder) (Appendix 2). The DNA ladder was run alongside extracted samples to give the sizes of the samples electrophoresed through a 1.5% agarose gel in 1X TBE buffer at 120 V for 90 minutes.

All gDNA samples were divided into long term stock solutions which were stored in a -80°C freezer. For experimental purposes, working stocks of DNA were standardized to 50ng/µl and stored at 4°C.

2.3.6 Polymerase Chain Reaction

DNA fragments of interest are amplified by means of the polymerase chain reaction (PCR) prior to HRM in the presence of a third generation intercalating dye called Evagreen™ (Biotium, US).

2.3.6.1. The principle of PCR

PCR is a molecular technique that makes it possible to amplify specific DNA regions. Under specific temperature cycling conditions, primers, which are two complementary oligonucleotide sequences that flank the DNA segment of interest, direct the specific amplification of the target DNA fragment making use of Taq polymerase and deoxynucleotide triphosphates (dNTPs). With each cycle, the concentration of target DNA doubles resulting in generation of multiple target sequence copies to be used in downstream applications (Garrett et al. 2005).

Sample DNA was amplified using real-time PCR in this project which makes it possible to amplify and quantify a DNA sequence simultaneously. This PCR was done in the presence of an EvaGreen™ (Biotium) dye which interacts and intercalates with only double stranded DNA thus fluorescing when bound. A shift in the fluorescence of the solution occurs due to an increase in the amount of DNA in the solution as PCR progresses and this fluorescence can be used to measure DNA concentration at different times during the reaction.

2.3.6.2. Gene annotation and primer design

Primers were designed to amplify all coding regions of *PARVA*. The *PARVA* gene was annotated using the AnnotV9 annotation programme using *PARVA* genetic sequence which was downloaded from Ensembl, GRCh38 (Yates, Akanni et al. 2016) (http://www.ensembl.org, accession number ENSG00000197702, transcript ENST00000334956) (Rebello, 2006; unpublished methodology). Primer sequences were chosen manually by referring to the annotated variants gene sequence, and analyzing the chosen sequences using IDT OligoAnanlyzer (http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/Default.aspx) and NCBI BLAST (Boratyn, Camacho et al. 2013) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) web-based tools for the determination of the optimal primer sequences for the gene amplification.

The primers were required to have:-

(1) A length of between 18 and 25 base pairs

(2) Good sequence diversity with minimal sequence repeats and secondary structure

(3) A melting temperature of between 50°C and 60°C

(4) A GC content of between 45% and 65%, and

(5) A G or C clamp at the 3' end.

An increase in amplicon size in HRM leads to a subsequent decrease in resolution of both the amplicon being screened and thus any mutation that is to be detected. Because of this, the amplicon sequence lengths were restricted to 250-300 bp in length. Using these criteria primer sets were designed for the combined PCR and HRM amplification and analysis of the *PARVA* exons and synthesized by Integrated DNA Technologies (IDT) (IDT, US) (Table 2.2).

The primer pairs used for the HRM analysis of the *PARVA* exons are as shown in Table 2.2 below.

Table 2.2: Primer pairs used for HRM analysis of *PARVA* exons

Primer	PCR products size (bps)	Sequence (5'-3')	Annealing Temperature (°C)
1-1F	246	AAG GGA AAG GCG AGC GTG	55
1-1R		AAT CAT CTT TCT TGC GGG ACG GG	
1-2F	410	AAA GCC GCA GCC TCA GTC	55
1-2R		CTC AGT GTG GGG TGC GCG TCG	
2F	206	TGTTACAGAGCTCCAACCCC	55
2R		GCTGGCATGGGAAAGCTTAA	
3F	327	CCA TCA ACC TGC CCC TCA G	55
3R		GCT GCC ATG ACT TTT TAC CTT CTC	
4F	205	GCTGGTCTGTAAGAAAATACCCC	55
4R		CACACATCCATTGCAGCACT	
5F	285	GCATGCAGTAGGGTTGTCTG	55
5R		ATGGATGGCTTCTTGGGGAA	
6F	248	TGCCTAGAACCTGGAAGAAGA	55
6R		GACCTACTCGACGCACCATA	
7F	187	CAGGAATGCTCATCAGTGTTTC	55
7R		TCCCTCCTGCAATCCCAATA	
8F	194	GCA GTG ATG GAG TGT CCT TTC	55
8R		TGA GAG CCT GTT CCT CCC AG	
9F	241	GAGGCTTCCCATCGGTAGTT	55
9R		TCCCAATGTTCTGTTGCATCTC	
10F	327	TTT CAC CCT CAC CCT TGC CCT C	55
10R		TCC ACA GAG ACT CAG TCC ACC	
11F	227	GGTAGGCTTCAGGTGGCATA	55
11R		TCCCTTGATGTCCCTTCTCC	
12F	244	TGCTGGGCTCCTTCACTTC	55
12R		TGCTTCCCTCTGCTACAAAATC	
13F	258	GAA GGG AGG GGC AGT GTA TG	55
13R		CAC AGT GCA AGA GAC AGG	

2.3.7. HRM analysis

HRM analysis is used to detect any potential mutations in samples analyzed in comparison to control samples.

2.3.7.1 The principle of HRM Analysis

Detection of mutation by HRM analysis is a two-step process involving amplification of the target DNA sequence using PCR and analysis of the produced amplicon by HRM. With the RotorGene 6000 (Corbett Life Sciences, Australia), these two steps can be combined into a single procedure where DNA samples are amplified and then immediately subjected to HRM.

HRM analysis is based on the dissociation behavior of double stranded DNA (dsDNA) dsDNA due to increasing temperature. Melting of dsDNA depends on its GC content and overall distribution of bases.

In HRM analysis, investigation of the melting curve can distinguish between several PCR amplicons with subtle changes in sequence and length, down to the single nucleotide level (Figure 2.3).

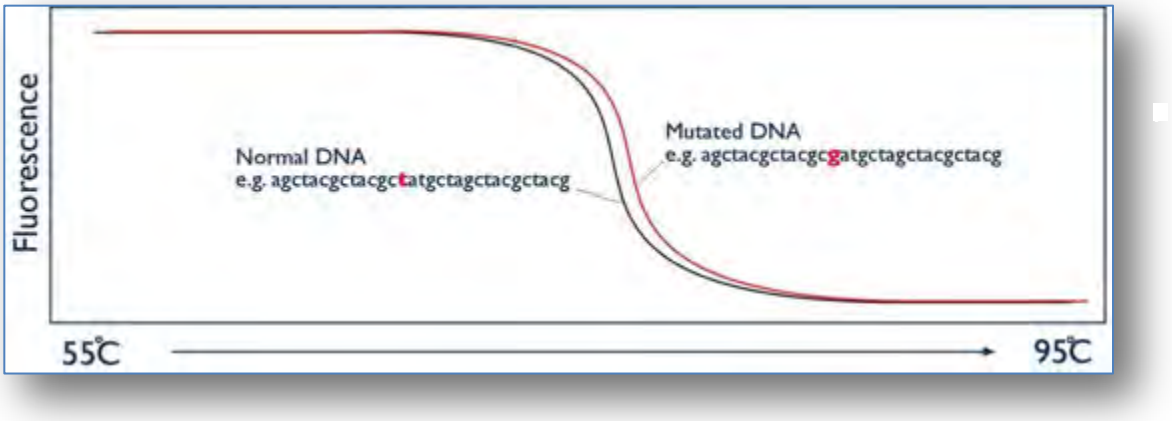


Figure 2.3: Melt curve showing the high efficiency of HRM to single base nucleotide

During HRM, the PCR amplicon is exposed to increasing temperature. The fluorescence of third-generation intercalating dyes (e.g., EvaGreen™) is measured continuously and is plotted against

increasing temperature. At lower temperatures, all DNA is double-stranded thus fluorescence is high. The DNA will start to disassociate into two single strands as the temperature increases, resulting in DNA melting. At this point, the dye is released and the fluorescence will decrease. The melting temperature (Tm) of the DNA sample under analysis is the point of the melt phase where the rate of change in fluorescence is greatest. When the DNA is completely melted, only some background fluorescence will be detected.

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

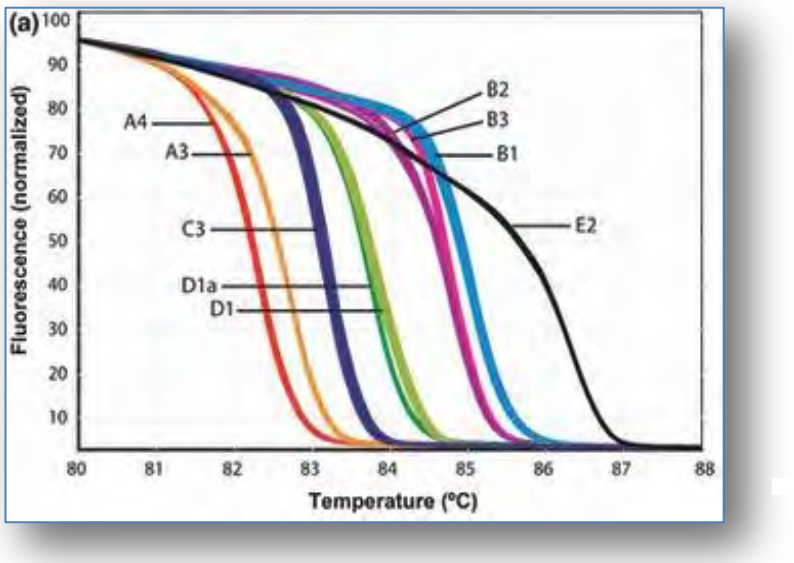


Figure 2.4: Melt curve profiles of different PCR amplicons each with differential melt curves (URL15)

2.3.7.2 HRM Protocol

The reagents used for the combined real-time PCR and HRM of the variant amplicons are shown in Table 2.3 and the reaction conditions are shown in Table 2.4. Scientific Specialities Inc Kits (ssibio, US) were used for HRM analysis.

Table 2.3: Reagents and concentrations in PCR and HRM of the variants exons

REAGENT (STOCK CONCENTRATION)	FINAL CONCENTRATION (µM) OR VOLUME IN SOLUTION QUANTITY (µl)
Forward primer (20µM)	0.4
Reverse primer (20µM)	0.4
dNTPS (20uM)	0.8
GoTaq Polymerase (5U/µl)	0.5U
$MgCl_2$ (25mM)	3
Evagreen (20x)	0.5
DNA	50ng
dH_2O	13.4
Total	25

Table 2.4: Optimized temperature cycling conditions for the PCR and HRM of the variant exons

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

2.3.7.3 Purification of HRM products

After HRM analysis, the samples identified for sequence analysis were purified using the Thermosensitive Alkaline Phosphatase (Applied Biosystems, US) and Exonuclease I (Applied Biosystems) enzymes. The reagents used are shown in Table 2.5. Reactions were incubated at 37°C for 1 hour, after which a 75°C deactivation step was conducted for 15 mins in a MultiGene Gradient thermal cycler (Labnet International).

Table 2.5: Cleanup protocol for variant HRM products

REAGENT	FINAL CONCENTRATION (µM) OR VOLUME IN SOLUTION QUANTITY (µl)
Exonuclease I (20U/µL)	2U
Thermosensitive Alkaline Phosphatase (1U/uL)	1U
Distilled/deionized water	13.9
HRM product	5
Final reaction volume	20

2.3.8 Sanger sequencing

Sanger sequencing was performed on all samples that showed a variation in the melt profile compared to the controls.

2.3.8.1. The principle of Sanger sequencing

Sequencing makes it possible to elucidate the alignment of base pairs in a DNA molecule. To carry out the procedure, the DNA of interest is mixed with DNA polymerase, deoxynucleotide triphosphates (dNTPs), four labelled dideoxynucleotide triphosphates (ddNTPs) and a sequence-specific oligonucleotide primer. DdNTPs and dNTPs are incorporated into the growing strand in a similar manner to each other, but the incorporation of the ddNTPs causes strand termination. For the sequencing reaction, temperature cycling similar to PCR is employed (Hartwell et al. 2004).

The sequencing reaction will generate a complete series of small fragments, each ending with a fluorescently labelled ddNTP. The small fragments can be separated by capillary electrophoresis and detection of the different ddNTPs will allow the DNA sequence of this strand to be determined (Luckey and Smith 1993).

2.3.8.2 Sequencing of variant exons

Sequencing of the HRM products was performed using the BigDye® Terminator v3.1 Sequencing Kit (Applied Biosystems). The reagents used for the direct sequencing reaction of the variants are described in Table 2.6. The temperature cycling conditions were conducted in a MultiGene Gradient thermal cycler (Labnet International) and the sequencing reactions are as described in Table 2.7.

Table 2.6: PCR protocol for the sequencing reactions of variant amplicons

REAGENT	FINAL CONCENTRATION (µM) OR VOLUME IN SOLUTION
Forward Primer (1 µM)	0.1
BigDye ® Terminator v3.1. Ready Reaction Mix (Applied Biosystems)	2 µl
5 X Sequencing Buffer (Applied Biosystems)	1
Distilled/deionized water	9 µl
HRM product	3 µl
FINAL REACTION VOLUME	20 µl

Table 2.7: Optimised cycling conditions for the sequencing reactions of variant amplicons

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

2.3.8.3 Capillary electrophoresis

Analysis of the sequencing products was done using capillary electrophoresis at the DNA Sequencing Unit (Department of Genetics, Stellenbosch University), by means of an ABI PRISM® 3130xl Genetic Analyzer (Applied Biosystems) or an ABI PRISM® 3730xl Genetic Analyzer (Applied Biosystems).

2.3.9 Bioinformatic Tools/ Public Genome Browsers/Databases

The pathogenicity of variants was determined using the three gene prediction tools MutationTaster (Schwarz, Cooper et al. 2014) (http://www.mutationtaster.org/), SIFT (Nelson, Stucker et al. 2016) (http://sift.jcvi.org/) and PolyPhen-2 (Adzhubei, Schmidt et al. 2010) (http://genetics.bwh.harvard.edu/pph2/) (Appendix 5).

2.4. Results

2.4.1. ACM 8 family screen

The *PARVA* c.392A>T variant, identified by Mbele in 2014, was screened in all ACM8 family members in order to confirm its segregation within all members of the ACM8 family. HRM variant analysis of the c.392A>T variant in the family indicated a deviation in HRM profiles between the wildtype and ACM 8.2, 8.3, 8.4 (Figure 2.5 A and B). Sequencing results for the family were compared to a control (Figure 2.5 C and D).

A B

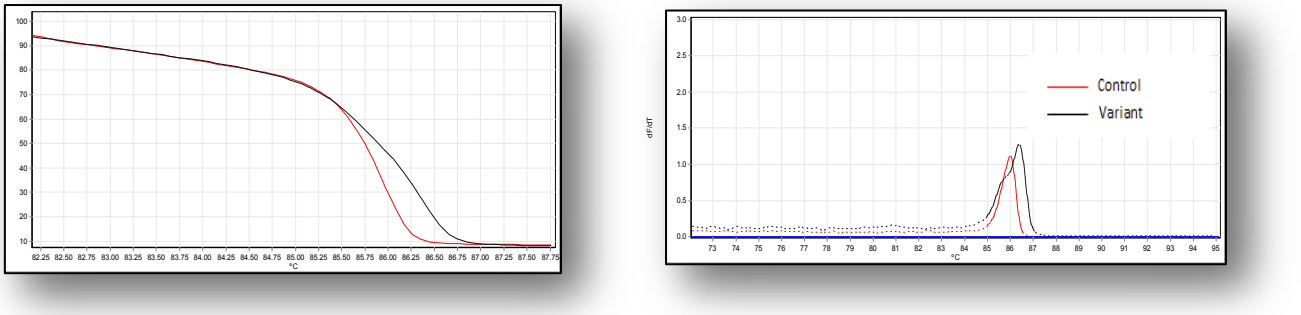


Figure 2.5: Validation of the *PARVA* c.392A>T Exon 3 variant. The black arrow (C-D) points to the region of interest. (A) and (B) HRM graph showing the exon 3 melting curve and peak graphs of sample ACM8.3 (black line) against a negative control (red line);

C D

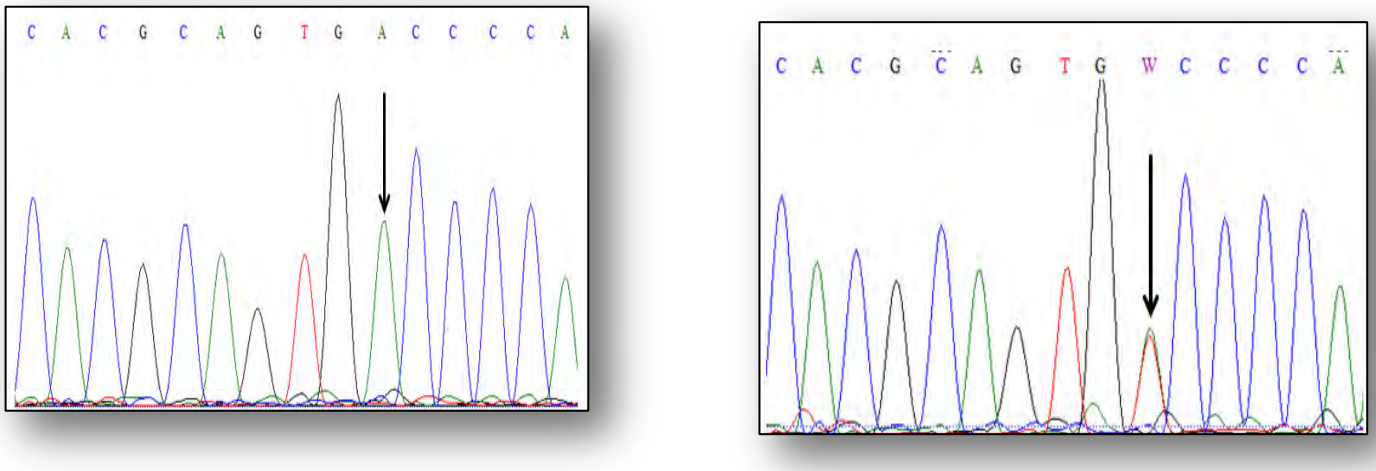


Figure 2.5: Validation of the *PARVA* c.392A>T Exon 3 variant. (C) Sequencing electropherogram showing the *PARVA* sequence in a control sample in the region of interest; (D) Sequencing electropherogram showing the c.392A>T sequence change in ACM 8.3

The c.392A>T variant was only found in the confirmed affected family members, ACM8.2, 8.3 and 8.4 and its segregation pattern is as shown in the pedigree chart below (Figure 2.6).

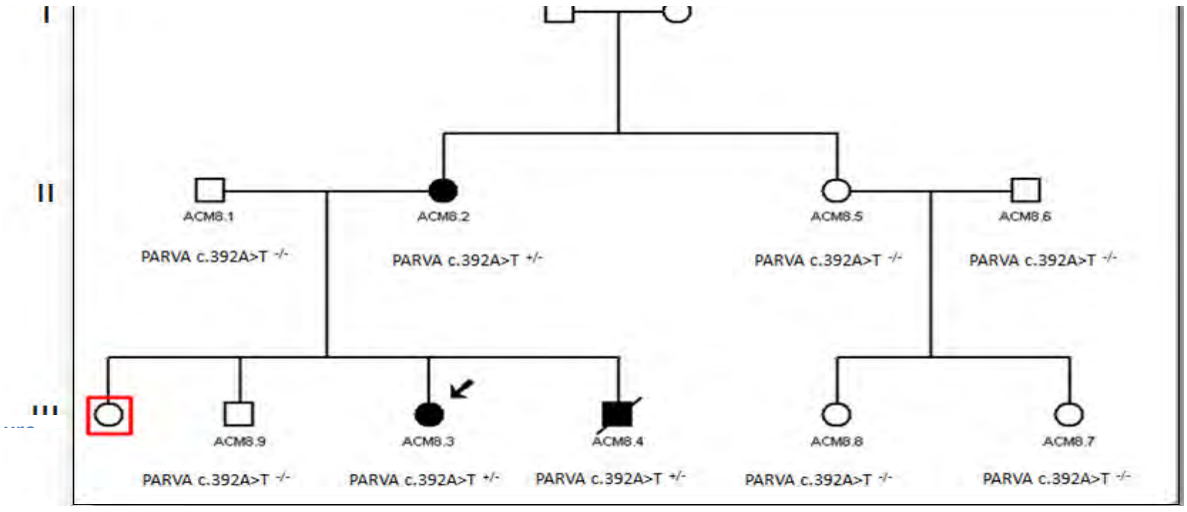


Figure 2.6: Pedigree chart for ACM 8 family showing the segregation of the *PARVA* c.392A>T

2.4.2. Cardiomyopathy cohort screen

Through HRM analysis and Sanger sequencing we screened *PARVA* for pathogenic variations in our cardiomyopathy cohorts (ARVC, DCM, PCM and RCM) consisting of 180 probands. We observed three reported single nucleotide polymorphisms (SNPs) (Table 2.8).

Table 2.8: Results of PARVA screening in a cardiomyopathy cohort

Variant	Location	Frequency in Cardiomyopathy cohort (%)	Frequency in 1000 Genomes population (%)	rs number
c.726C>T p.R242	Exonic	1.5	14	11547363
c.777+10G>A	Intronic	0.9	3	111731473
c.777+26C>G	Intronic	0.9	2	74401006

2.4.3 Variant functional analysis

Of the variants analyzed, only *PARVA* c.392A>T was predicted to be deleterious by all of the three prediction tools as shown in Table 2.9 below. The other variants were synonymous and intronic and were therefore unlikely to affect protein function.

Table 2.9: Variants identified in PARVA

Variant	Location	MutationTaster	Sift	Polyphen
c.392A>T p.D131V	Exonic	Disease causing (0.999)	Damaging (0.01)	Probably damaging (0.990)
c.726C>T p.R242	Exonic	Polymorphism (1.246)	Tolerated (1)	N/A
c.777+10G>A	Intronic	N/A	N/A	N/A
c.777+26C>G	Intronic	N/A	N/A	N/A

2.4.4 Population screening

We screened the *PARVA* c.392A>T variant in order to determine the frequency in the South African background population (Table 2.10). The *PARVA* c.392A>T variant was previously investigated in our research group and reported to be absent in the South African control population (Mbele, 2014) as well as the 1000 Genomes Project database. It was however found in Exome Aggregation Consortium browser (ExAC) (Lek, Karczewski et al. 2016) and EVS but was considered to be rare according to the frequencies observed in the databases mentioned.

Table 2.10: Population frequencies of the PARVA c.392A>T variant

Locus	Gene	Coding	Amino Acid change	Frequency in SA population (%)	Frequency in EXAC (%)	Frequency in 1000 Genomes (%)	Frequency in EVS (%)
chr11:12495505	*PARVA*	c.392A>T	p.Asp131Val	0/60 (0%)	1/32962 (0.03%)	0 (0%)	1/12227 (0.08%)

2.5. Discussion

At the outset of this investigation we wanted to establish the causal role of *PARVA* in ARVC based on previous experiments done by Dr.Mbele that identified it as a possible candidate gene for ARVC.

PARVA has many functional roles that make it a biologically plausible candidate not just for ARVC but for cardiomyopathy as well. This gene (1) forms an integral component of cell adhesion complexes (Nikolopoulos and Turner 2000), (2) has a role in reorganization of the actin cytoskeleton, sarcomere organization, smooth muscle cell contraction, etc. (Wang, Fukuda et al. 2008) and (3) is highly expressed in the heart (Sopko, Qin et al. 2011) .

Experiments in mice have shown that mice deficient of *Parva* express cardiovascular defects in the embryonic stage which includes abnormal vascular beds with dilated blood vessels and pericardial effusion combined with whole body edema and severe bleeding due to vessel rupture (Montanez, Wickstrom et al. 2009). Sopko et.al utilized failing DCM human heart tissues, and found a 2.27-fold enhanced expression of PINCH, while the expression of PARVA and ILK were similarly increased 4- and 10.5-fold, respectively (Sopko, Qin et al. 2011). These findings led them to purport the improvement of cardiac function by enhancing PIP mediated Protein kinase B (Akt) activation (Sopko, Qin et al. 2011). In human cardiac myocyte the ternary complex is localized in the costameres which overlap the Z-line of cardiac myocytes and play a role in myocyte hypertrophy through sarcomere assembly (Chen, Huang et al. 2005).

For my project we continued to screen the ACM 8 family for the *PARVA* c.392A>T variant to determine the segregation with disease and found the variant to occur in both severely affected siblings (ACM 8.3 and ACM 8.4) as well as the mildly affected mother (ACM 8.2). The differences in phenotype amongst the ACM 8 family members may be due to the incomplete penetrance of ARVC or a possible genetic modifier such as *HMGXB3*. Incomplete penetrance arises in the case where individuals harbouring identical mutations show differing phenotypes, and such variation can sometimes be due to different genetic backgrounds or even environmental conditions (Raj, Rifkin et al. 2010). This phenotypic variability in ACM 8 cannot be explained by the *PARVA* c.392A>T variant and called into question the causative role of

PARVA within this family. We also screened our cardiomyopathy cohorts consisting of 180 probands (ACM, DCM, DCM and RCM) for *PARVA* mutations but only common SNPs were identified.

In the initial study in 2012, Dr Mbele had screened 66 additional ARVC patients for mutations in *PARVA* and reported a novel pathogenic non-synonymous c.523A>G variant in one ARVC case (ACM 30.1) and a synonymous variant of unknown significance, c.597T>G in another ARVC case (ACM 32.1). Further investigation of these patients by the cardiomyopathy panel revealed that both probands do not meet the criteria for ARVC even though they had initially been diagnosed with ARVC. The panel found that ACM 30.1, initially classified as having ARVC based on the presence of an arrhythmia with a left bundle branch block morphology and a superior axis, has no other criteria for ARVC; he has a hypertensive heart disease and previous myocardial infarction (S Kraus, personal communication). They also found that ACM 32.1 has an unclassified cardiomyopathy affecting the left ventricle with ventricular tachycardia (with right bundle branch block morphology) but does not fulfil any criteria for ARVC according to Task Force Criteria. These findings by the panel on the phenotype of these two probands now casts doubt on the role of *PARVA* in cardiomyopathy as both probands do not have a definitive diagnosis aside from the mentioned ventricular tachycardia.

In summary, Dr Mbele had identified three *PARVA* variants (c.392A>T; c.523A>G; c.597T>G) in patients with ARVC but we consider these variants to be disease-associated rather than definitely pathogenic for several reasons. Firstly, the *PARVA* c.392A>T variant cannot explain the variation in phenotype that is seen in the ACM 8 family. Secondly, the other two variants were classified as variants of unknown significance as the patients' diagnosis was unclear and we were ambiguous of the role that *PARVA* played in these two patients. We could not comment on familial segregation as the families were not available for phenotyping. We classify these three variants as ARVC-associated but not definitely pathogenic.

In this small study, our observations do not provide strong supporting evidence for *PARVA* as a candidate gene for ARVC or any other cardiomyopathy.

Chapter 3: WHOLE EXOME SEQUENCING OF ACM 8

3.1. Introduction

We elected to validate Dr.Mbele's findings in the South African setting by using the Ion Torrent platform and compare the results to the Illumina platform that had been used by Mbele in the UK. The successful validation of his results would achieve several goals: (1) it would verify that the next generation sequencing (NGS) techniques had been appropriately set up in our laboratory (2) that we had acquired the skills necessary for data analysis of NGS data (3) reproducing the data from the UK whole exome sequencing experiment would allow us to examine the data for other possible causative variants.

Whole genome sequencing and whole exome sequencing (Berglund, Kiialainen et al. 2011) form the new wave of next generation sequencing (NGS) technologies that have enabled the identification of novel genes responsible for disease phenotypes (Mardis 2013). We explored whole exome sequencing technologies as an improvement to the laborious yet fundamental Sanger sequencing method of DNA sequencing (Sanger, Nicklen et al. 1977); this has enabled the accelerated study of DNA on a genome-wide scale thus improving the nature of biological inquiry (Mardis 2013).. NGS has therefore become an invaluable tool in the elucidation of the genetics of cardiomyopathy, where much of the genetic influence on the disease remains to be discovered.

3.2 Aim

The aim of this investigation was to validate the whole exome sequencing findings carried out on an Illumina platform in the UK on an Ion Torrent platform in South Africa.

3.3 Methods

3.3.1 Sample library preparation

The human genomic DNA (gDNA) samples of ACM 8.3 and ACM 8.4 were submitted to the Centre for Proteomic and Genomic Research (CPGR) for whole exome sequencing on the Ion Proton sequencer.

These samples were first analyzed on the NanoDrop ND1000 (Thermo Fisher Scientific) to assess sample purity and quality, the Qubit®2.0 Fluorometer for absolute quantification and finally on a 1% agarose gel to assess integrity of the high molecular weight gDNA.

A total of 100 ng of each sample was used as template in the exome amplification PCR protocol using the Ion AmpliSeq™ Exome RDY Kit (Life Technologies, US). The amplicons were treated with the FuPa Reagent (Life Technologies) to partially digest the primers and phosphorylate the 3'-ends thus ensuring efficient ligation reactions downstream. Amplification products were ligated to Ion Xpress™ Barcode Adapters (Life Technologies) and purified using Agencourt Ampure XP beads (Beckmann Coulter, US). Amplification products with ligated adaptors were then amplified through a limited cycle PCR, and purified with Agencourt Ampure XP beads (Beckmann Coulter). The barcoded amplification products (sequencing libraries) were then quantified by quantitative PCR (qPCR) using the KAPA Library Quantification kit for Ion Torrent (KAPA BioSystems, SA) and fragment size distribution was determined on the BioAnalyzer instrument using the BioAnalyzer High Sensitivity DNA Assay Kit (Agilent, US).

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

3.3.2 Library quality control

Absolute library concentration and fragment size distribution are key quality control (QC) results post library preparation and prior to sequencing. The library concentrations were determined by qPCR. The fragment size distribution analysis was performed on the Bioanalyzer

(Agilent) using the BioAnalyzer High sensitivity DNA Assay kit (Agilent). The Ion AmpliSeq™ Exome RDY library (Ion Torrent) preparation method ideally yields library fragments with a size distributed between 200 and 350 bp.

3.3.3 Data Analysis

A summary of the sequencing data output including the number of reads and quality of data was produced per sequencing run.

The two samples were multiplexed per chip and the QC metrics were generated using Torrent Suite v4.4.3. Overall, the number of reads per sample ranged from ~20 – 55 million. Results of the secondary on-board analysis on Torrent Suite shows a categorization of total data per run following alignment recalibration against the GRCh37.p5 reference from Phred 17, 20 and above (AQ17, AQ20 and "Perfect" corresponds to corresponds to <1% error rate).
Overall, alignment against the GRCh37.p5 (hg19) reference was at least 96% indicating a maximum of 4% unaligned data down to AQ17; thus a minimum of 98% mean raw accuracy

Raw data from the CPGR were uploaded to IonReporter™. The data were initially processed on IonReporter™ before being downloaded and filtered.

3.3.4 Variant filtering

A list of 189 known cardiomyopathy-associated genes, generated from the NGS panel (n=48 genes (GENETests™ ©2016)), PAN cardiomyopathy panel (n=103 genes (Blueprint Genetics © 2016)) and elife (n=159 genes (Zou, Tran et al. 2015)), was used for initial filtering of the whole exome sequencing data. For variants occurring in the gene panel, we looked for variants occurring in both affected siblings and applied the following filters: (1) variants with a minor allele frequency (MAF) of $\leq$1% according to IonReporter™ (2) variants occurring within exons or spanning intron/exon boundaries (3) all variants that occur as missense, nonsense, stop/loss

and insertions and deletions. All variants were validated using Sanger sequencing and in appropriate population controls to determine the frequency in the population.

3.3.5 Bioinformatic Tools/ Public Genome Browsers/Databases

The pathogenicity of all non-synonymous rare variants was determined as described previously (Section *2.3.9*). Only variants predicted to be disease causing by at least two out of the three prediction tools were considered for further analysis. To check for the expression levels of these gene transcripts in the heart, a human gene online database, GeneCards (http://www.genecards.org/) (Stelzer, Plaschkes et al. 2016), was used.

3.3.6. Validation of potentially disease-causing variants

The validation of shortlisted variants was done using PCR and HRM analysis (Sections *2.3.6* and *2.3.7*) followed by Sanger sequencing (Section *2.3.8*) using the primers described in Table 3.1. All available ACM 8 family members were screened for the shortlisted variants in order to determine segregation of these variants with disease.

Table 3.1: Primer set for HRM analysis of c.1162C>T variant in ACM 8 family

Primer	PCR products size (bps)	Sequence (5'-3')	Annealing Temperature (^{o}C)	Variant
4F	291	AGT ATT CGC TGA GTC GTC TCT	55	c.1162C>T
4R		GCA AAG TCA CCA TAA TAG AAG		

3.4. Results

3.4.1. Variant filtering for the known cardiac genes

Using the Ion AmpliSeq™ Exome RDY Kit, a total of 30,500 shared variants were identified between ACM 8.3 and 8.4 at an average sequencing depth of 125X.

We looked for variants occurring in both affected siblings and applied the following filters: (1) variants with a minor allele frequency (MAF) of ≤1% according to IonReporter™, (2) variants occurring within exons or spanning intron/exon boundaries, (3) all variants that occur as missense, nonsense, stop/loss and insertions and deletions. All variants were validated using Sanger sequencing and in appropriate population controls to determine the frequency in the population (as per "Methods" section *3.3.4*).

Filtering for variants appearing in the cardiac gene panel resulted in us mining 232 common variants (Figure 3.1). These variants were subjected to further filtering which allowed us to decrease the total number of variants to 17: nine mutations were homozygous (Table 3.2) and eight were heterozygous variants (Table 3.3). One variant however, *PKP2* c.1162C>T, stood clearly above the rest as this variant has previously been reported to be associated with ARVC by our group (Watkins, Hendricks et al. 2009).

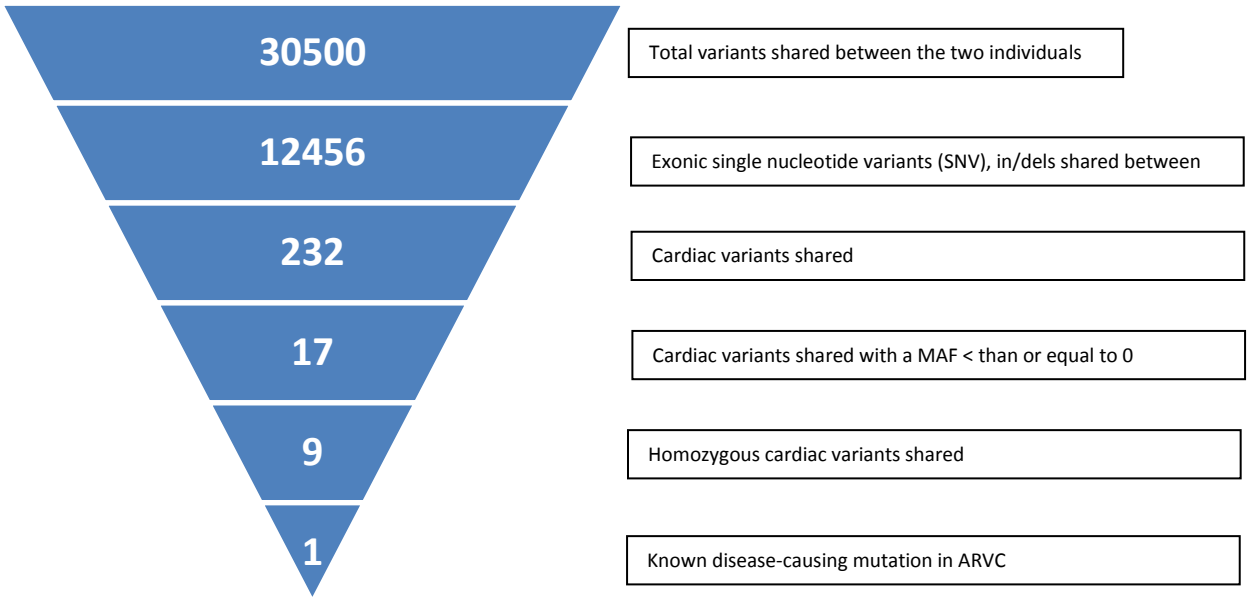


Figure 3.1: Filtering of whole exome sequencing data by using the known cardiac genes as a filter.

Table 3.2: Results of variant filtering in the homozygous cardiac gene variants

Locus	Genes	Coding	Amino Acid Change	Mutationtaster	Sift	Polyphen	Expression in the heart (Log_{10} ppm)
chr12:33021869	*PKP2* (Akinrinade, Ollila et al. 2015)	c.1162C>T	p.Arg388Trp	Polymorphism (0.999)	Damaging 0	Probably damaging (0.999)	251
chr2:179444768	*TTN* (Akinrinade, Ollila et al. 2015)	c.62323G>C	p.Ala20775Pro	Polymorphism (0.999)	Not predicted	Benign (0.000)	1056
chr1:228444565	*OBSCN* (Zou, Tran et al. 2015)	c.4799T>A	p.Val1600Asp	Polymorphism (0.999)	Tolerated 1		20
chr10:29822032	*SVIL* (Zou, Tran et al. 2015)	c.1264G>A	p.Val422Ile	Polymorphism (1)	Not scored	Benign (0.000)	3
chr6:152540278	*SYNE* (Akinrinade, Ollila et al. 2015)	c.21904T>G	p.Phe7302Val	Polymorphism (0.999)	Tolerated 0.55	Benign (0.000)	0.7
chr10:112572458	*RBM20* (Akinrinade, Ollila et al. 2015)	c.2303G>C	p.Trp768Ser	Polymorphism (0.999)	Tolerated 0.85	Benign (0.000)	0.6
chr12:2791205	*CACNA1C* (Zou, Tran et al. 2015)	c.5678A>G	p.Lys1893Arg	Polymorphism (0.999)	Tolerated 0.88	Benign (0.000)	0.54
chr7:91714911	*AKAP9* (Zou, Tran et al. 2015)	c.8935C>T	p.Pro2979Ser	Polymorphism (0.999)	Tolerated 1	Benign (0.000)	0.09
chr3:38739574	*SCN10A* (Zou, Tran et al. 2015)	c.5137A>G	p.Met1713Val	Polymorphism (0.999)	Tolerated 1	Benign (0.000)	0

Table 3.3: Results of variant filtering in the heterozygous cardiac gene variants

Locus	Genes	Coding	Amino Acid Change	Mutationtaster	Sift	Polyphen	Expression in the heart (Log_{10} ppm)
chr10:29821758	*SVIL* (Deo, Musso et al. 2014)	c.1538G>A	p.Ser513Asn	Polymorphism (0.73)	Not scored	Benign (0.004)	3
chr1:144854581	*PDE4DIP* (Zou, Tran et al. 2015)	c.6889A>G	p.Thr2297Ala	Polymorphism 0.99	Tolerated 0.74	Benign 0.000	2
chr1:144871755	*PDE4DIP* (Zou, Tran et al. 2015)	c.5207T>A	p.Val1736Glu	Polymorphism 0.99	Tolerated 0.32	Benign 0.361	2
chr1:144871782	*PDE4DIP* (Zou, Tran et al. 2015)	c.5180T>C	p.Leu1727Pro	Disease causing 0.83	Tolerated 1	Benign 0.000	2
chr1:144874815	*PDE4DIP* (Zou, Tran et al. 2015)	c.4793A>G	p.His1598Arg	Polymorphism 0.99	Tolerated 0.15	Benign 0.001	2
chr1:144879054	*PDE4DIP* (Zou, Tran et al. 2015)	c.4396A>G	p.Arg1466Gly	Polymorphism 0.99	Tolerated 0.22	Possibly damaging (0.835)	2
chr15:99646108	*SYNM* (Zou, Tran et al. 2015)	c.703A>G	p.Arg235Gly	Polymorphism 0.99	Not scored	Benign 0.001	34
chr22:19929263	*COMT* (Zou, Tran et al. 2015)	c.64G>A	p.Val22Met	Polymorphism 0.99	Tolerated 0.24	Benign 0.261	21

3.4.2 Identification of the *PKP2* c.1162C>T founder mutation

Through the abovementioned filtering criteria we identified the homozygous *PKP2* c.1162C>T mutation in the two affected individuals, ACM 8.3 and ACM 8.4 (Table 3.2). This mutation was verified by HRM analysis (Figure 3.2 A and B) and Sanger sequencing (Figure 3.2 C-E).

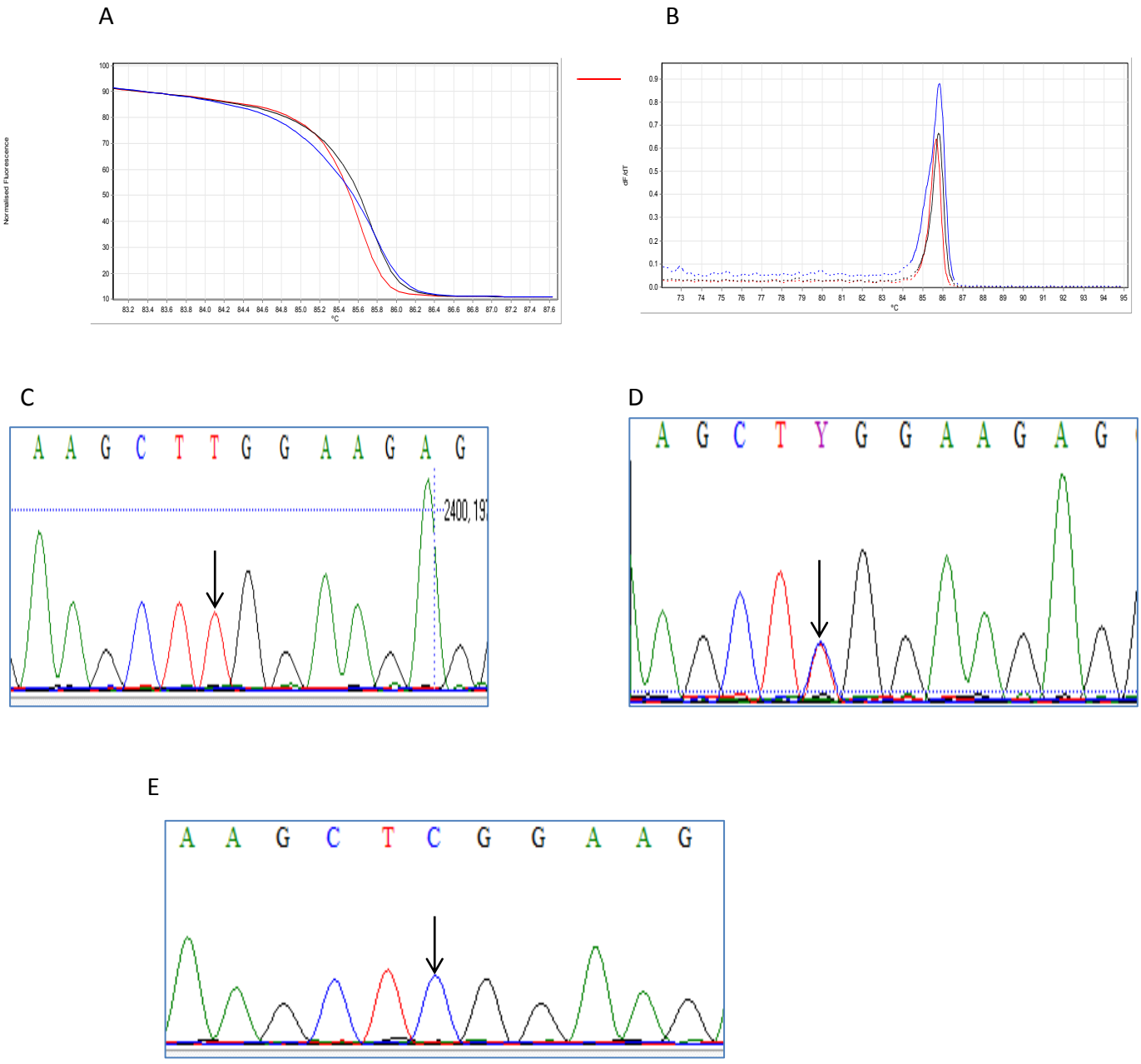


Figure 3.2: Identification of the *PKP2* c.1162 C>T mutation in exon 4 of the affected siblings of the ACM 8 family. HRM showing the exon 4 derivative melt profiles of sample (A) the proband (ACM8.3) and a negative control and (B) the sibling (ACM 8.4) and a negative control; Sequencing electropherogram of the (C) the proband (ACM 8.3), (D) the sibling (ACM 8.4) and (E) the control

3.4.3 Segregation of the *PKP2* c.1162C>T founder mutation

The founder mutation was further screened by HRM analysis and Sanger sequencing in seven other family members, ACM 8.1, 8.2, 8.5, 8.6, 8.7, 8.8 and 8.9 (Figures 3.3 and 3.4). The father (ACM 8.1) and the mother (ACM 8.2) were confirmed to carry the *PKP2* c.1162C>T mutation respectively (Figure 3.3). These results confirm that the siblings (ACM 8.3 and ACM 8.4) had inherited a mutant copy from each parent, resulting in the homozygous genotype (Figure 3.5).

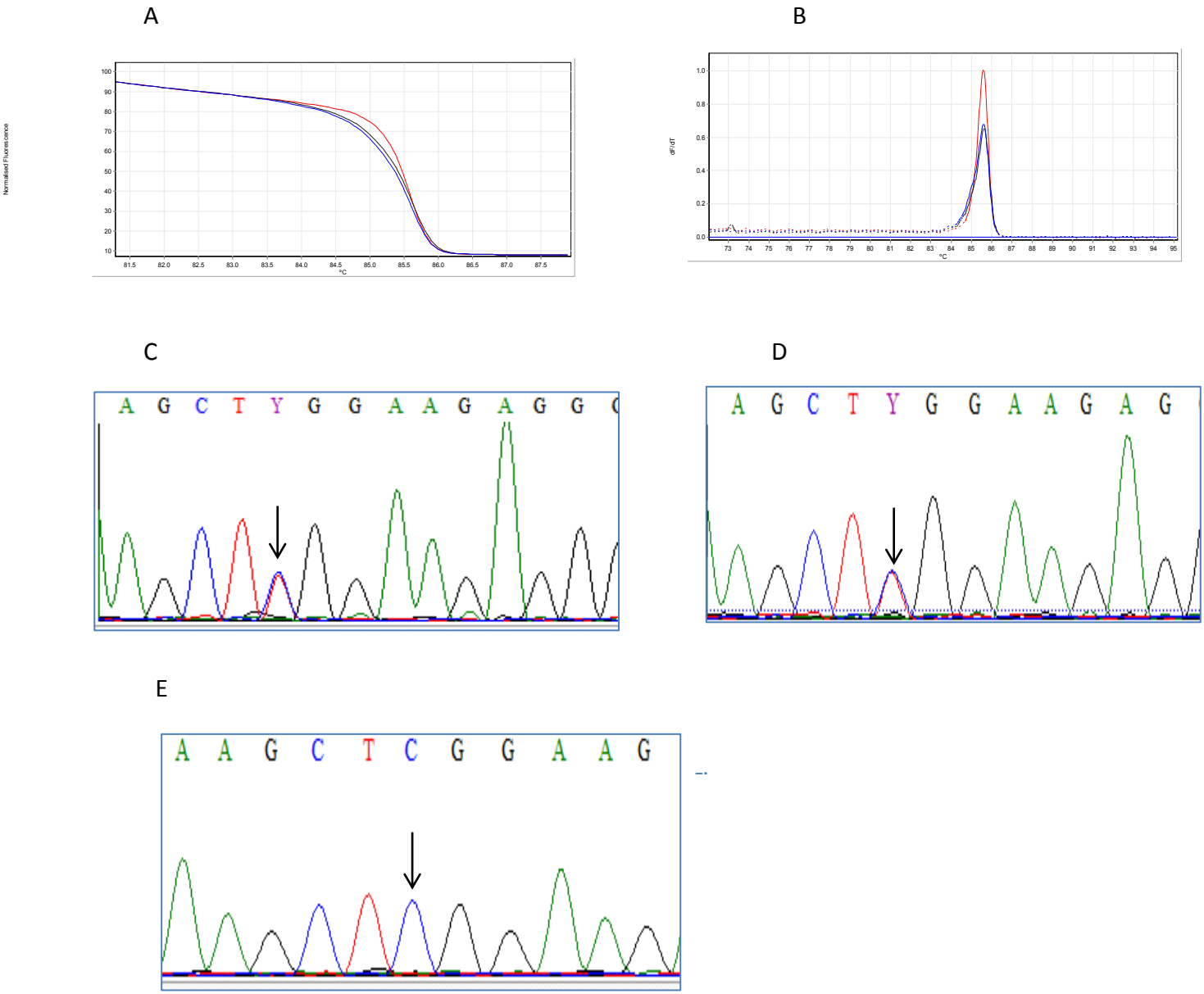


Figure 3.3: Identification of the *PKP2* c.1162 C>T mutation in exon 4 of the parents of the ACM 8 family. HRM showing the exon 4 derivative melt profiles of sample (A) the father (ACM 8.1) and a negative control and (B) the mother (ACM 8.2) and a negative control; Sequencing electropherogram of the (C) the father (ACM8.1), (D) the mother (ACM 8.2) and (E) the control

Samples ACM 8.5, 8.6, 8.7 and 8.8 displayed similar HRM profiles to that of the controls; Sanger sequencing confirmed that these family members were negative for this variant.

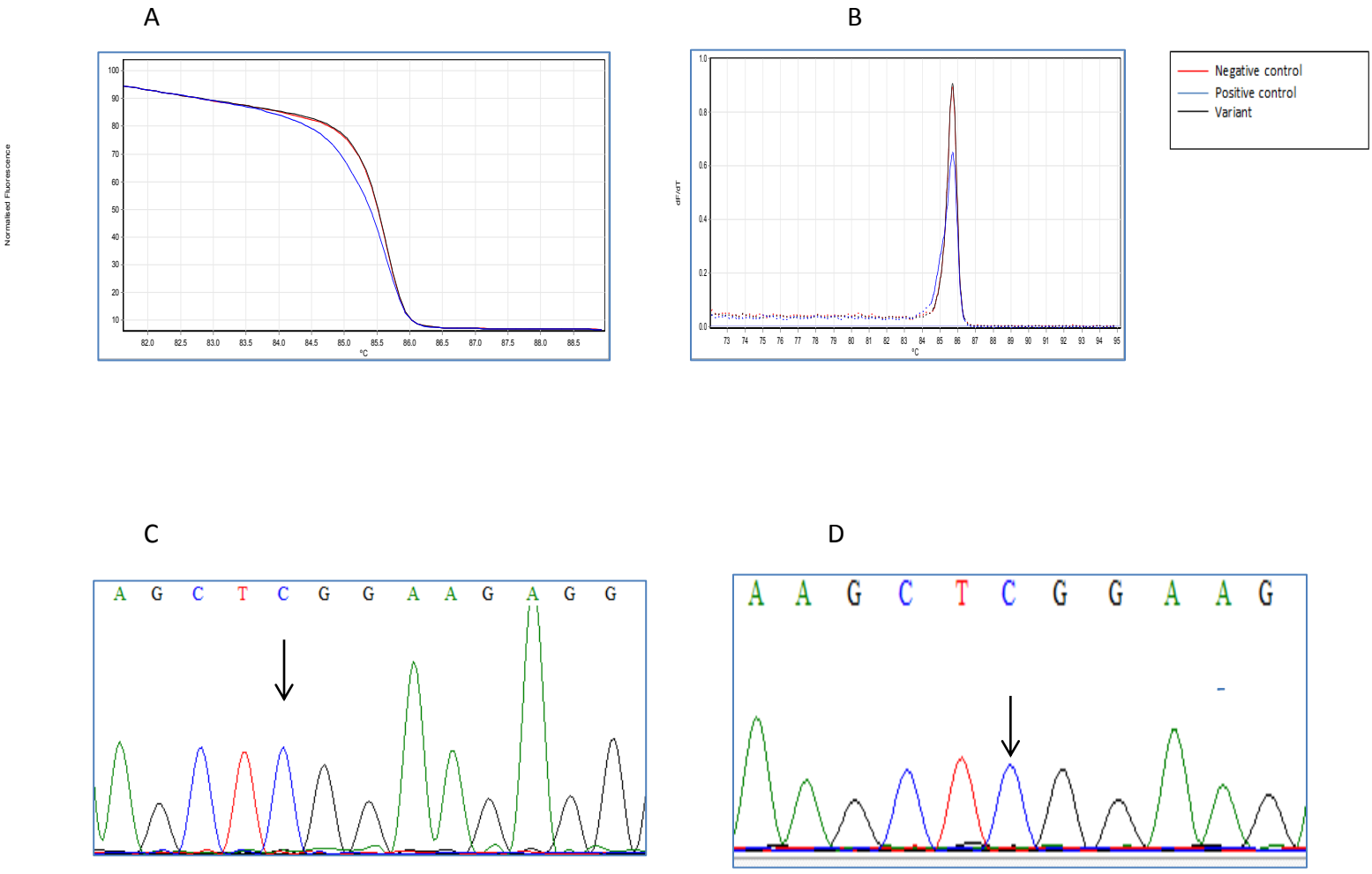


Figure 3.4: Identification of the c.1162 C>T *PKP2* Exon 4 variant in the ACM 8 family members
HRM showing the exon 4 derivative melt profiles of samples (A and B) ACM8.5 and a negative control; Sequencing electropherogram of samples (C) ACM8.5 and (D) the negative control

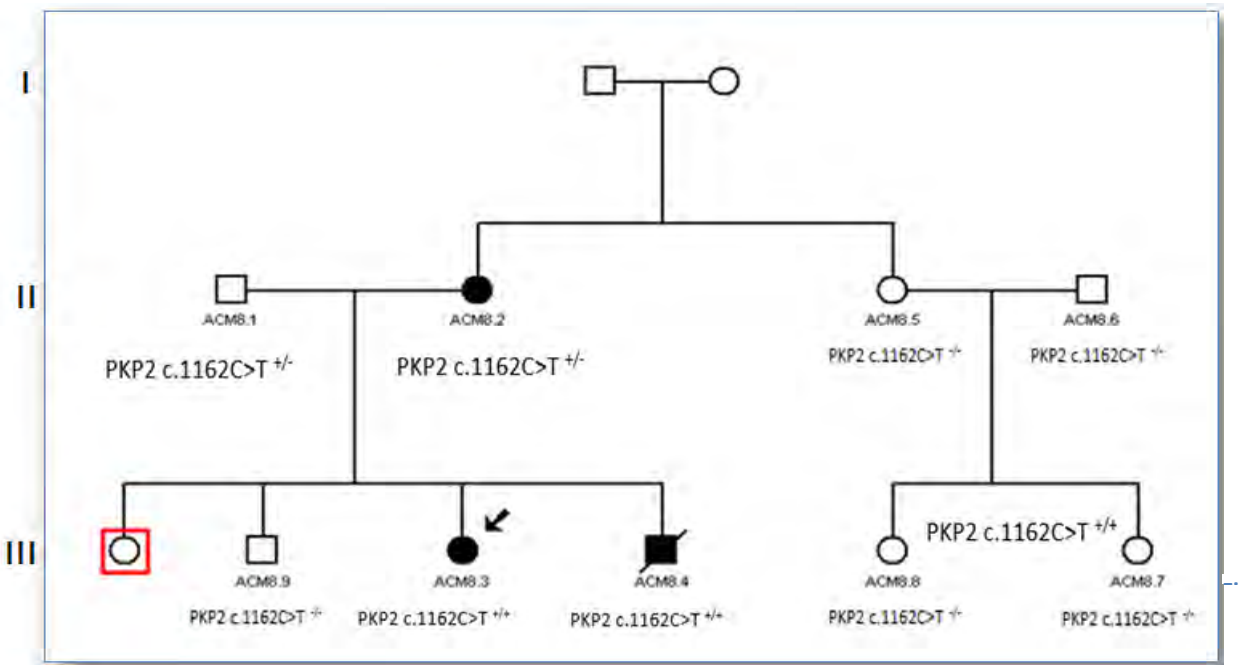


Figure 3.5: Pedigree of the ACM 8 family with ARVC. Circles represent females. Squares represent males. Clear circles and squares represent unaffected family members. Black circles and squares represent affected family members. Crossed symbols represent deceased members of the family. The circle in red represents a family member whose blood was not available for DNA analysis.

3.5 Discussion

This investigation sought to validate the results of prior whole exome sequencing conducted on this family, in which a variant in *PARVA* was reported as the possible causative mutation (Mbele, 2014). Instead, our validation experiment identified variant known disease-causing mutation, *PKP2* c.1162C>T, that was inherited in a homozygous state by the two severely affected siblings with ARVC.

PKP2 encodes a protein which forms part of the cellular complex structure called the desmosome which is involved in intercellular adhesion and signaling in the epithelia and cardiac muscle (Green and Gaudry 2000) (Figure 3.6). *PKP2* is also found in the nucleus, although its function in the nucleus is unknown (Schmidt and Jager 2005).

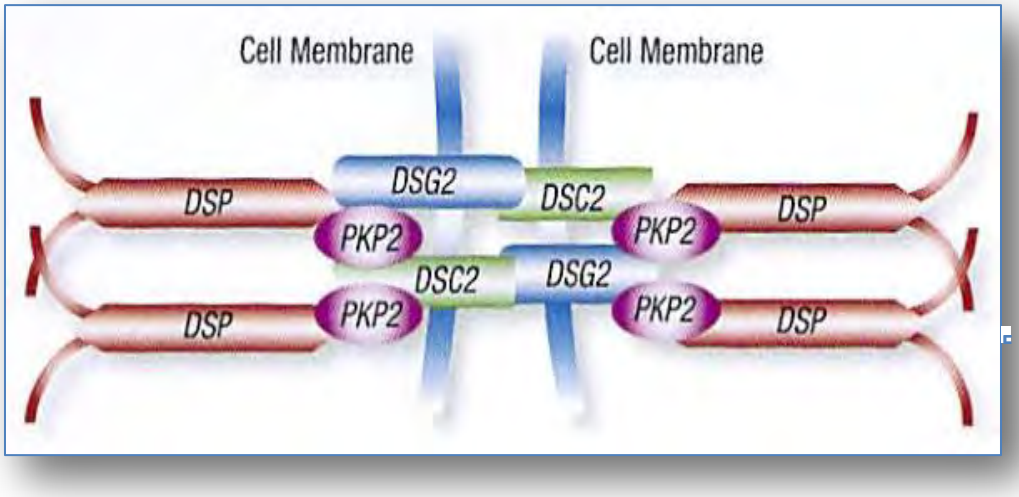


Figure 3.6: An illustration of the cellular desmosomal structure with *PKP2* shown amidst the other desmosomal proteins. (URL16)

PKP2 mutations in 32 of 120 probands (27%) with ARVC of western European descent pointed to an important role of this gene in the pathogenesis of this disease (Gerull, Heuser et al. 2004). Mutations in *PKP2* have been shown to cause ARVC (Gerull, Heuser et al. 2004, van der Zwaag, Cox et al. 2010, Zhang, Tavora et al. 2012, Li Mura, Bauce et al. 2013, Zhou, Chen et al. 2015) and it is estimated that these mutations account for 70% of all mutations associated with ARVC

(van Tintelen, Entius et al. 2006). The mutations generally appear to disrupt the desmosomal assembly and stability and even though the particular pathogenesis of *PKP2* mutations in ARVC is hypothetical, it is speculated that disruption of cardiomyocytes in response to mechanical stretch or stress occurs, particularly in the triangle of dysplasia (right ventricular outflow tract, inferobasal area, and apex) because of impairment of cell-cell contact (Hall, Li et al. 2009). The c.1162C>T missense mutation causes the amino acid encoded to change from a hydrophilic arginine (R) to a hydrophobic tryptophan (W) at protein position 388, a change which may lead to instability of the whole protein because of the PKP2 protein unfolding (Wimley and White 1996). This mutation is known to cause disease when a single mutant allele is present. It is noteworthy that the majority of white South Africans carrying this mutation are of Dutch descent (Watkins, Hendricks et al. 2009).

When we correlate the genotype with the phenotype in this family we find that the individual with the heterozygous form of the *PKP2* c.1162C>T mutation (the mother) displays a milder phenotype and later age of onset compare to the children who were homozygous for the *PKP2* c.1162C>T mutation with a more severe phenotype with a much earlier age of onset. We have previously observed this gene-dose effect in the context of compound heterozygous mutations in *PKP2* (Watkins, Hendricks et al. 2009). The *PKP2* mutation identified in ACM 8 occurred in exon 4 (Figure 3.7) of *PKP2*

Figure 3.7: Figure showing the PKP2 mutation at exon 4 (URL12)

Even though the original intention of this project was to validate the findings of a previous whole exome sequencing project, I found the true disease causing gene within this family.

We are aware that the original candidate gene screen on *PKP2* in 2009 should have detected the c.1162C>T variant and will discuss how this variant could have been missed. Firstly, during the candidate gene screen carried out in 2009, only the student was responsible for analyzing hundreds of electropherograms at a time. The student had been trained and was considered to be experienced in reading the electropherograms- that practice was long since stopped. Secondly, the phenotype indicated that only the mother was affected, this led to a heterozygous autosomal dominant inheritance pattern suggested for this family where both affected siblings inherited the disease from the mother. This could also be the reason why the variant was missed during the UK whole exome sequencing data analysis; they were looking for a heterozygous variant and not a homozygous variant. We were not able to explore the Newcastle results as the raw data remained in Newcastle. We will ask our Newcastle collaborators to confirm the presence of this variant in their Illumina data.

It has to be noted that although both the Illumina and Ion Torrent platforms are NGS techniques, they operate on different principles, which could lead to a difference in performance(Salipante, Kawashima et al. 2014). In the Illumina platform DNA fragments are prepared for sequencing by bridge PCR which simultaneously amplifies single DNA molecules and covalently links amplicons to a solid substrate clusters. In contrast, Ion Torrent sequencing initially prepares templates by using emulsion PCR (Shendure and Ji 2008) whereby PCR reagents, primer-coated particles, and a low concentration of template fragments are combined with oil and emulsified to form picoliter-scale microreactions to achieve clonal amplification of single DNA molecules on the surfaces of individual particles(Shendure and Ji 2008). In terms of data analysis Illumina favours resources that enable full analysis of their data at the location of data generation`, whereas end users employing Ion Torrent platform are able to access the data and share across working groups and used for large-scale customized analysis at sites other than where the data was generated (Glenn 2011). This variance between different platforms can also be affected by knowledge and experience of the bioinformatician analyzing the data.

In light of the homozygous mutation being missed above, we have since put laboratory practices in place that should avoid this scenario from occurring again: (1) All electropherograms are reviewed by a minimum of two experienced staff members and/or postgraduate students (2) Due to the incomplete penetrance of ARVC and to avoid bias we have elected to filter all whole exome sequencing data for ARVC into autosomal dominant and autosomal recessive data streams regardless of the suggested inheritance pattern; this allows us to perform a more robust analysis.

The future of genomic sequencing is very bright with the ability of NGS to integrate a diverse set of data analysis techniques into a single platform being a revolutionary development in the clinical field (Gullapalli, Desai et al. 2012). Nevertheless, it is still a young field and not much is known about it such as how to handle the large data sets being generated and the whole scope of their application and thus a concrete outline on the pitfalls of clinical NGS platform is a future prospect.

Chapter 4: CONCLUSION

A previous whole exome sequencing project completed in the United Kingdom (UK) in 2012 had identified *PARVA* as a candidate gene for cardiomyopathy in the ACM 8 family. We hypothesized that *PARVA* may harbor novel mutations that cause ARVC and other cardiomyopathies. With that in mind we set out to screen for *PARVA* mutations in all members of the ACM 8 family as well as in our cardiomyopathy cohorts (ARVC, DCM, HCM and RCM). We also wanted to validate the whole exome sequencing results obtained in the UK using the Ion Proton platform.

Our screening results indicated that the *PARVA* c.392A>T variant was detected in three family members, two siblings (ACM 8.3 and ACM 8.4) who were severely affected with ARVC at an early age and the mother (ACM 8.2) developed mild ARVC at a later age. The phenotypic variability seen within this family cannot be explained by the *PARVA* c.392A>T variant alone and called into question the causative role of *PARVA* within this family. In addition to the c.392A>T variant, Dr Mbele also found the variants *PARVA* c.523A>G and c.597T>G in two ARVC probands but these probands were later found not to meet the Task Force criteria for ARVC. This casts doubt on the role of *PARVA* in cardiomyopathy. In addition, during our screening, we found no pathogenic mutations in any of the 180 cardiomyopathy probands.

The whole exome sequencing investigation sought to validate the results of prior whole exome sequencing conducted on this family, in which a variant in *PARVA* was reported as the possible causative mutation (Mbele, 2014). Instead, our validation experiment identified the known disease-causing mutation, *PKP2* c.1162C>T, that was inherited in a homozygous state by the two severely affected siblings with ARVC. This mutation was reported by Watkins as a founder mutation of ARVC amongst South Africans of European descent. It is noteworthy that the family in question is of Dutch descent, thus furthering the effect of this variant in causing ARVC and not the initially purported *PARVA* c.392A>T mutation.

Although we found no evidence to support the causal role of *PARVA* in cardiomyopathy, this study was successful in identifying the causal mutation of the ARVC phenotype in this family.

REFERENCES

Literature

URL references

1. Human heart anatomy (2016) http://www.livescience.com/34655-human-heart.html (Accessed on 2016-02-14)

2. Heart structure with function (2016) http://www.newhealthadvisor.com/Heart-Structure-and-Function.html (Accessed on 2016-01-19)

3. Layers of the heart wall (2015) https://www.boundless.com/physiology/textbooks/boundless-anatomy-and-physiology-textbook/cardiovascular-system-the-heart-18/the-heart-172/layers-of-the-heart-walls-864-636/ (Accessed on 2016-03-23)

4. Overview of hypertrophic cardiomyopathy (2016) (https://www.mayoclinic.org.hypertrophic-cardiomyopathy) (Accessed on 2016-03-27)

5. Blausen dilated cardiomyopathy (2015) (https://upload.wikimedia.org/wikipedia/commons/7/75/Blausen_0165_Cardiomyopathy_Dilated.png) (Accessed on 2015-11-29)

6. Cardiomyopathy (2016) (https://www.emedicine.medscape.com) (Accessed on 2016-04-08)

7. Cardiomyopathy panel (2016) (https://www.genetests.org/) (Accessed 2016-05-12)

8. MutationTaster (2014) (http://www.mutationtaster.org/) (Accessed on 2016-06-01)

9. SIFT (http://www.jcvi.org/cms/about/overview/) (Accessed on 2016-06-01)

10. PolyPhen-2 prediction of functional effects of human nsSNPs (http://genetics.bwh.harvard.edu/pph2/) (Accessed on 2016-06-01)

11. DNA melting curve (2016) https://upload /DNA_melting_scematic_curve_2 (Accessed on 2016-03-03)

12. Results from South African registry (2016) https://www.researchgate.net _Results-from-the-South-African-registry-Schematic-representation-of-the-PKP2-mutations (Accessed on 2016-06-23)

13. Hypertrophic cardiomyopathy (2016) https://www.wattpad.com/story/72867711-hypertrophic-cardiomyopathy-therapeutics-market (Accessed on 2016-04-19)

14. Restrictive cardiomyopathy (2016) https://www.drugs.com/cg/restrictive-cardiomyopathy.html (Accessed on 2016-04-21)

15. High resolution melting profiles (2016) https://www.researchgate.net/High-resolution-melting-profiles (Accessed on 2016-03-12)

16. Familion tests http://www.lippman.org/ACC/familiontests.html (Accessed on 2016-06-27)

17. GeneCards (2016) http://www.genecards.org/ (Accessed on 2016-06-06)

18. Ensembl (2016) http://www.ensembl.org, (Accessed on 2016-04-14)

Automated references

Abegaz, B. (1990). "The impact of echocardiography in the diagnosis of hypertrophic cardiomyopathy." East Afr Med J **67**(8): 556-567.

Ackerman, M. J., S. G. Priori, S. Willems, C. Berul, R. Brugada, H. Calkins, A. J. Camm, P. T. Ellinor, M. Gollob, R. Hamilton, R. E. Hershberger, D. P. Judge, H. Le Marec, W. J. McKenna, E. Schulze-Bahr, C. Semsarian, J. A. Towbin, H. Watkins, A. Wilde, C. Wolpert and D. P. Zipes (2011). "HRS/EHRA expert consensus statement on the state of genetic testing for the channelopathies and cardiomyopathies this document was developed as a partnership between the Heart Rhythm Society (HRS) and the European Heart Rhythm Association (EHRA)." Heart Rhythm **8**(8): 1308-1339.

Adalsteinsdottir, B., P. Teekakirikul, B. J. Maron, M. A. Burke, D. F. Gudbjartsson, H. Holm, K. Stefansson, S. R. DePalma, E. Mazaika, B. McDonough, R. Danielsen, J. G. Seidman, C. E. Seidman and G. T. Gunnarsson (2014). "Nationwide study on hypertrophic cardiomyopathy in Iceland: evidence of a MYBPC3 founder mutation." Circulation **130**(14): 1158-1167.

Adzhubei, I., D. M. Jordan and S. R. Sunyaev (2013). "Predicting functional effect of human missense mutations using PolyPhen-2." Curr Protoc Hum Genet **Chapter 7**: Unit7 20.

Adzhubei, I. A., S. Schmidt, L. Peshkin, V. E. Ramensky, A. Gerasimova, P. Bork, A. S. Kondrashov and S. R. Sunyaev (2010). "A method and server for predicting damaging missense mutations." Nat Methods **7**(4): 248-249.

Akinkugbe, O. O., G. D. Nicholson and J. K. Cruickshank (1991). "Heart disease in blacks of Africa and the Caribbean." Cardiovasc Clin **21**(3): 377-391.

Akinrinade, O., L. Ollila, S. Vattulainen, J. Tallila, M. Gentile, P. Salmenpera, H. Koillinen, M. Kaartinen, M. S. Nieminen, S. Myllykangas, T. P. Alastalo, J. W. Koskenvuo and T. Helio (2015). "Genetics and genotype-phenotype correlations in Finnish patients with dilated cardiomyopathy." Eur Heart J **36**(34): 2327-2337.

Amoah, A. G. and C. Kallen (2000). "Aetiology of heart failure as seen from a National Cardiac Referral Centre in Africa." Cardiology **93**(1-2): 11-18.

Andreasen, C., J. B. Nielsen, L. Refsgaard, A. G. Holst, A. H. Christensen, L. Andreasen, A. Sajadieh, S. Haunso, J. H. Svendsen and M. S. Olesen (2013). "New population-based exome data are questioning the pathogenicity of previously cardiomyopathy-associated genetic variants." Eur J Hum Genet **21**(9): 918-928.

Arbustini, E., N. Narula, L. Tavazzi, A. Serio, M. Grasso, V. Favalli, R. Bellazzi, J. A. Tajik, R. O. Bonow, V. Fuster and J. Narula (2014). "The MOGE(S) classification of cardiomyopathy for clinicians." J Am Coll Cardiol **64**(3): 304-318.

Awad, M. M., H. Calkins and D. P. Judge (2008). "Mechanisms of disease: molecular genetics of arrhythmogenic right ventricular dysplasia/cardiomyopathy." Nat Clin Pract Cardiovasc Med **5**(5): 258-267.
Awad, M. M., D. Dalal, E. Cho, N. Amat-Alarcon, C. James, C. Tichnell, A. Tucker, S. D. Russell, D. A. Bluemke, H. C. Dietz, H. Calkins and D. P. Judge (2006). "DSG2 mutations contribute to arrhythmogenic right ventricular dysplasia/cardiomyopathy." Am J Hum Genet **79**(1): 136-142.
Badorff, C., N. Berkely, S. Mehrotra, J. W. Talhouk, R. E. Rhoads and K. U. Knowlton (2000). "Enteroviral protease 2A directly cleaves dystrophin and is inhibited by a dystrophin-based substrate analogue." J Biol Chem **275**(15): 11191-11197.
Batchelder, K. and B. M. Mayosi (2005). "Pentoxifylline for heart failure: a systematic review." S Afr Med J **95**(3): 171-175.
Beck, W. (1978). "Cardiomyopathies in South Africa--a brief survey of the problem and current therapeutic approaches." Postgrad Med J **54**(633): 469-476.
Beffagna, G., G. Occhi, A. Nava, L. Vitiello, A. Ditadi, C. Basso, B. Bauce, G. Carraro, G. Thiene, J. A. Towbin, G. A. Danieli and A. Rampazzo (2005). "Regulatory mutations in transforming growth factor-beta3 gene cause arrhythmogenic right ventricular cardiomyopathy type 1." Cardiovasc Res **65**(2): 366-373.
Berglund, E. C., A. Kiialainen and A. C. Syvanen (2011). "Next-generation sequencing technologies and applications for human genetic history and forensics." Investig Genet **2**: 23.
Bione, S., P. D'Adamo, E. Maestrini, A. K. Gedeon, P. A. Bolhuis and D. Toniolo (1996). "A novel X-linked gene, G4.5. is responsible for Barth syndrome." Nat Genet **12**(4): 385-389.
Blair, E., C. Redwood, H. Ashrafian, M. Oliveira, J. Broxholme, B. Kerr, A. Salmon, I. Ostman-Smith and H. Watkins (2001). "Mutations in the gamma(2) subunit of AMP-activated protein kinase cause familial hypertrophic cardiomyopathy: evidence for the central role of energy compromise in disease pathogenesis." Hum Mol Genet **10**(11): 1215-1220.
Boratyn, G. M., C. Camacho, P. S. Cooper, G. Coulouris, A. Fong, N. Ma, T. L. Madden, W. T. Matten, S. D. McGinnis, Y. Merezhuk, Y. Raytselis, E. W. Sayers, T. Tao, J. Ye and I. Zaretskaya (2013). "BLAST: a more efficient report with usability improvements." Nucleic Acids Res **41**(Web Server issue): W29-33.
Bos, J. M., R. N. Poley, M. Ny, D. J. Tester, X. Xu, M. Vatta, J. A. Towbin, B. J. Gersh, S. R. Ommen and M. J. Ackerman (2006). "Genotype-phenotype relationships involving hypertrophic cardiomyopathy-associated mutations in titin, muscle LIM protein, and telethonin." Mol Genet Metab **88**(1): 78-85.
Bradlow, B. A., M. M. Zion and S. J. Fleishman (1964). "Heart Disease in Africa, with Particular Reference to Southern Africa." Am J Cardiol **13**: 650-669.
Cantlay, A. M., K. Shokrollahi, J. T. Allen, P. W. Lunt, R. A. Newbury-Ecob and C. G. Steward (1999). "Genetic analysis of the G4.5 gene in families with suspected Barth syndrome." J Pediatr **135**(3): 311-315.
Carniel, E., M. R. Taylor, G. Sinagra, A. Di Lenarda, L. Ku, P. R. Fain, M. M. Boucek, J. Cavanaugh, S. Miocic, D. Slavov, S. L. Graw, J. Feiger, X. Z. Zhu, D. Dao, D. A. Ferguson, M. R. Bristow and L. Mestroni (2005). "Alpha-myosin heavy chain: a sarcomeric gene associated with dilated and hypertrophic phenotypes of cardiomyopathy." Circulation **112**(1): 54-59.
Chen, H., X. N. Huang, W. Yan, K. Chen, L. Guo, L. Tummalapali, S. Dedhar, R. St-Arnaud, C. Wu and J. L. Sepulveda (2005). "Role of the integrin-linked kinase/PINCH1/alpha-parvin complex in cardiac myocyte hypertrophy." Lab Invest **85**(11): 1342-1356.
Cirino, A. L. and C. Ho (1993). Hypertrophic Cardiomyopathy Overview. GeneReviews(R). R. A. Pagon, M. P. Adam, H. H. Ardinger et al. Seattle (WA).
Codd, M. B., D. D. Sugrue, B. J. Gersh and L. J. Melton, 3rd (1989). "Epidemiology of idiopathic dilated and hypertrophic cardiomyopathy. A population-based study in Olmsted County, Minnesota, 1975-1984." Circulation **80**(3): 564-572.

Corrado, D., C. Basso, M. Schiavon and G. Thiene (1998). "Screening for hypertrophic cardiomyopathy in young athletes." N Engl J Med **339**(6): 364-369.
Corrado, D. and G. Thiene (2006). "Arrhythmogenic right ventricular cardiomyopathy/dysplasia: clinical impact of molecular genetic studies." Circulation **113**(13): 1634-1637.
Cosnett, J. E. (1962). "Heart disease in the Zulu: especially cardiomyopathy and cardiac infarction." Br Heart J **24**: 76-82.
Cosnett, J. E. and D. J. Pudifin (1964). "Embolic Complications of Cardiomyopathy." Br Heart J **26**: 544-548.
Cowan, J., A. Morales, J. Dagua and R. E. Hershberger (2008). "Genetic testing and genetic counseling in cardiovascular genetic medicine: overview and preliminary recommendations." Congest Heart Fail **14**(2): 97-105.
Daehmlow, S., J. Erdmann, T. Knueppel, C. Gille, C. Froemmel, M. Hummel, R. Hetzer and V. Regitz-Zagrosek (2002). "Novel mutations in sarcomeric protein genes in dilated cardiomyopathy." Biochem Biophys Res Commun **298**(1): 116-120.
Deo, R. C., G. Musso, M. Tasan, P. Tang, A. Poon, C. Yuan, J. F. Felix, R. S. Vasan, R. Beroukhim, T. De Marco, P. Y. Kwok, C. A. MacRae and F. P. Roth (2014). "Prioritizing causal disease genes using unbiased genomic features." Genome Biol **15**(12): 534.
Duboscq-Bidot, L., P. Charron, V. Ruppert, L. Fauchier, A. Richter, L. Tavazzi, E. Arbustini, T. Wichter, B. Maisch, M. Komajda, R. Isnard, E. Villard and E. H. F. Network (2009). "Mutations in the ANKRD1 gene encoding CARP are responsible for human dilated cardiomyopathy." Eur Heart J **30**(17): 2128-2136.
Elliott, P. (2000). "Cardiomyopathy. Diagnosis and management of dilated cardiomyopathy." Heart **84**(1): 106-112.
Elliott, P., B. Andersson, E. Arbustini, Z. Bilinska, F. Cecchi, P. Charron, O. Dubourg, U. Kuhl, B. Maisch, W. J. McKenna, L. Monserrat, S. Pankuweit, C. Rapezzi, P. Seferovic, L. Tavazzi and A. Keren (2008). "Classification of the cardiomyopathies: a position statement from the European Society Of Cardiology Working Group on Myocardial and Pericardial Diseases." Eur Heart J **29**(2): 270-276.
Elliott, P. and W. J. McKenna (2004). "Hypertrophic cardiomyopathy." Lancet **363**(9424): 1881-1891.
Fatkin, D., C. MacRae, T. Sasaki, M. R. Wolff, M. Porcu, M. Frenneaux, J. Atherton, H. J. Vidaillet, Jr., S. Spudich, U. De Girolami, J. G. Seidman, C. Seidman, F. Muntoni, G. Muehle, W. Johnson and B. McDonough (1999). "Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease." N Engl J Med **341**(23): 1715-1724.
Felker, G. M., R. E. Thompson, J. M. Hare, R. H. Hruban, D. E. Clemetson, D. L. Howard, K. L. Baughman and E. K. Kasper (2000). "Underlying causes and long-term survival in patients with initially unexplained cardiomyopathy." N Engl J Med **342**(15): 1077-1084.
Franz, W. M., M. Muller, O. J. Muller, R. Herrmann, T. Rothmann, M. Cremer, R. D. Cohn, T. Voit and H. A. Katus (2000). "Association of nonsense mutation of dystrophin gene with disruption of sarcoglycan complex in X-linked dilated cardiomyopathy." Lancet **355**(9217): 1781-1785.
Fujino, N., M. Shimizu, H. Ino, K. Okeie, M. Yamaguchi, T. Yasuda, H. Kokado and H. Mabuchi (2001). "Cardiac troponin T Arg92Trp mutation and progression from hypertrophic to dilated cardiomyopathy." Clin Cardiol **24**(5): 397-402.
Gerull, B., M. Gramlich, J. Atherton, M. McNabb, K. Trombitas, S. Sasse-Klaassen, J. G. Seidman, C. Seidman, H. Granzier, S. Labeit, M. Frenneaux and L. Thierfelder (2002). "Mutations of TTN, encoding the giant muscle filament titin, cause familial dilated cardiomyopathy." Nat Genet **30**(2): 201-204.
Gerull, B., A. Heuser, T. Wichter, M. Paul, C. T. Basson, D. A. McDermott, B. B. Lerman, S. M. Markowitz, P. T. Ellinor, C. A. MacRae, S. Peters, K. S. Grossmann, J. Drenckhahn, B. Michely, S. Sasse-Klaassen, W. Birchmeier, R. Dietz, G. Breithardt, E. Schulze-Bahr and L. Thierfelder (2004). "Mutations in the desmosomal protein plakophilin-2 are common in arrhythmogenic right ventricular cardiomyopathy." Nat Genet **36**(11): 1162-1164.

Glenn, T. C. (2011). "Field guide to next-generation DNA sequencers." Mol Ecol Resour **11**(5): 759-769.
Grasso, M., M. Diegoli, A. Brega, C. Campana, L. Tavazzi and E. Arbustini (2001). "The mitochondrial DNA mutation T12297C affects a highly conserved nucleotide of tRNA(Leu(CUN)) and is associated with dilated cardiomyopathy." Eur J Hum Genet **9**(4): 311-315.
Green, E. M., H. Wakimoto, R. L. Anderson, M. J. Evanchik, J. M. Gorham, B. C. Harrison, M. Henze, R. Kawas, J. D. Oslob, H. M. Rodriguez, Y. Song, W. Wan, L. A. Leinwand, J. A. Spudich, R. S. McDowell, J. G. Seidman and C. E. Seidman (2016). "A small-molecule inhibitor of sarcomere contractility suppresses hypertrophic cardiomyopathy in mice." Science **351**(6273): 617-621.
Green, K. J. and C. A. Gaudry (2000). "Are desmosomes more than tethers for intermediate filaments?" Nat Rev Mol Cell Biol **1**(3): 208-216.
Grunig, E., J. A. Tasman, H. Kucherer, W. Franz, W. Kubler and H. A. Katus (1998). "Frequency and phenotypes of familial dilated cardiomyopathy." J Am Coll Cardiol **31**(1): 186-194.
Gullapalli, R. R., K. V. Desai, L. Santana-Santos, J. A. Kant and M. J. Becich (2012). "Next generation sequencing in clinical medicine: Challenges and lessons for pathology and biomedical informatics." J Pathol Inform **3**: 40.
Hall, C., S. Li, H. Li, V. Creason and J. K. Wahl, 3rd (2009). "Arrhythmogenic right ventricular cardiomyopathy plakophilin-2 mutations disrupt desmosome assembly and stability." Cell Commun Adhes **16**(1-3): 15-27.
Hamid, M. S., M. Norman, A. Quraishi, S. Firoozi, R. Thaman, J. R. Gimeno, B. Sachdev, E. Rowland, P. M. Elliott and W. J. McKenna (2002). "Prospective evaluation of relatives for familial arrhythmogenic right ventricular cardiomyopathy/dysplasia reveals a need to broaden diagnostic criteria." J Am Coll Cardiol **40**(8): 1445-1450.
Hanson, E. L., P. M. Jakobs, H. Keegan, K. Coates, S. Bousman, N. H. Dienel, M. Litt and R. E. Hershberger (2002). "Cardiac troponin T lysine 210 deletion in a family with dilated cardiomyopathy." J Card Fail **8**(1): 28-32.
Hardie, D. G. and S. A. Hawley (2001). "AMP-activated protein kinase: the energy charge hypothesis revisited." Bioessays **23**(12): 1112-1119.
Haugaa, K. H., T. F. Haland, I. S. Leren, J. Saberniak and T. Edvardsen (2016). "Arrhythmogenic right ventricular cardiomyopathy, clinical manifestations, and diagnosis." Europace **18**(7): 965-972.
Heidbuchel, H., J. Hoogsteen, R. Fagard, L. Vanhees, H. Ector, R. Willems and J. Van Lierde (2003). "High prevalence of right ventricular involvement in endurance athletes with ventricular arrhythmias. Role of an electrophysiologic study in risk stratification." Eur Heart J **24**(16): 1473-1480.
Hendricks, N., D. A. Watkins and B. M. Mayosi (2010). "Lessons from the first report of the Arrhythmogenic Right Ventricular Cardiomyopathy Registry of South Africa." Cardiovasc J Afr **21**(3): 129-130.
Hershberger, R. E. and A. Morales (1993). LMNA-Related Dilated Cardiomyopathy. GeneReviews(R). R. A. Pagon, M. P. Adam, H. H. Ardinger et al. Seattle (WA).
Hulot, J. S., X. Jouven, J. P. Empana, R. Frank and G. Fontaine (2004). "Natural history and risk stratification of arrhythmogenic right ventricular dysplasia/cardiomyopathy." Circulation **110**(14): 1879-1884.
Isaacson, C. (1977). "The changing pattern of heart disease in South African Blacks." S Afr Med J **52**(20): 793-798.
Jiang, J., H. Wakimoto, J. G. Seidman and C. E. Seidman (2013). "Allele-specific silencing of mutant Myh6 transcripts in mice suppresses hypertrophic cardiomyopathy." Science **342**(6154): 111-114.
Kallichurum, S. (1969). "Major aetiological types of heart failure in the Bantu in Durban." S Afr Med J **43**(9): 250-252.

Kamisago, M., S. D. Sharma, S. R. DePalma, S. Solomon, P. Sharma, B. McDonough, L. Smoot, M. P. Mullen, P. K. Woolf, E. D. Wigle, J. G. Seidman and C. E. Seidman (2000). "Mutations in sarcomere protein genes as a cause of dilated cardiomyopathy." N Engl J Med **343**(23): 1688-1696.
Kaplan, S. R., J. J. Gard, N. Protonotarios, A. Tsatsopoulou, C. Spiliopoulou, A. Anastasakis, C. P. Squarcioni, W. J. McKenna, G. Thiene, C. Basso, N. Brousse, G. Fontaine and J. E. Saffitz (2004). "Remodeling of myocyte gap junctions in arrhythmogenic right ventricular cardiomyopathy due to a deletion in plakoglobin (Naxos disease)." Heart Rhythm **1**(1): 3-11.
Khogali, S. S., B. M. Mayosi, J. M. Beattie, W. J. McKenna, H. Watkins and J. Poulton (2001). "A common mitochondrial DNA variant associated with susceptibility to dilated cardiomyopathy in two different populations." Lancet **357**(9264): 1265-1267.
Kimura, A. (2016). "Molecular genetics and pathogenesis of cardiomyopathy." J Hum Genet **61**(1): 41-50.
Kirchhof, P., L. Fabritz, M. Zwiener, H. Witt, M. Schafers, S. Zellerhoff, M. Paul, T. Athai, K. H. Hiller, H. A. Baba, G. Breithardt, P. Ruiz, T. Wichter and B. Levkau (2006). "Age- and training-dependent development of arrhythmogenic right ventricular cardiomyopathy in heterozygous plakoglobin-deficient mice." Circulation **114**(17): 1799-1806.
Kubo, T., J. R. Gimeno, A. Bahl, U. Steffensen, M. Steffensen, E. Osman, R. Thaman, J. Mogensen, P. M. Elliott, Y. Doi and W. J. McKenna (2007). "Prevalence, clinical significance, and genetic basis of hypertrophic cardiomyopathy with restrictive phenotype." J Am Coll Cardiol **49**(25): 2419-2426.
Lee, P. Y., J. Costumbrado, C. Y. Hsu and Y. H. Kim (2012). "Agarose gel electrophoresis for the separation of DNA fragments." J Vis Exp(62).
Legate, K. R., E. Montanez, O. Kudlacek and R. Fassler (2006). "ILK, PINCH and parvin: the tIPP of integrin signalling." Nat Rev Mol Cell Biol **7**(1): 20-31.
Lek, M., K. J. Karczewski, E. V. Minikel, K. E. Samocha, E. Banks, T. Fennell, A. H. O'Donnell-Luria, J. S. Ware, A. J. Hill, B. B. Cummings, T. Tukiainen, D. P. Birnbaum, J. A. Kosmicki, L. E. Duncan, K. Estrada, F. Zhao, J. Zou, E. Pierce-Hoffman, J. Berghout, D. N. Cooper, N. Deflaux, M. DePristo, R. Do, J. Flannick, M. Fromer, L. Gauthier, J. Goldstein, N. Gupta, D. Howrigan, A. Kiezun, M. I. Kurki, A. L. Moonshine, P. Natarajan, L. Orozco, G. M. Peloso, R. Poplin, M. A. Rivas, V. Ruano-Rubio, S. A. Rose, D. M. Ruderfer, K. Shakir, P. D. Stenson, C. Stevens, B. P. Thomas, G. Tiao, M. T. Tusie-Luna, B. Weisburd, H. H. Won, D. Yu, D. M. Altshuler, D. Ardissino, M. Boehnke, J. Danesh, S. Donnelly, R. Elosua, J. C. Florez, S. B. Gabriel, G. Getz, S. J. Glatt, C. M. Hultman, S. Kathiresan, M. Laakso, S. McCarroll, M. I. McCarthy, D. McGovern, R. McPherson, B. M. Neale, A. Palotie, S. M. Purcell, D. Saleheen, J. M. Scharf, P. Sklar, P. F. Sullivan, J. Tuomilehto, M. T. Tsuang, H. C. Watkins, J. G. Wilson, M. J. Daly, D. G. MacArthur and C. Exome Aggregation (2016). "Analysis of protein-coding genetic variation in 60,706 humans." Nature **536**(7616): 285-291.
Lewis, C. M., P. M. Barnard, J. J. Van Der Walt and A. J. Brink (1973). "Haemodynamic basis for pericardiotomy as palliative treatment of idiopathic endomyocardiopathy." Recent Adv Stud Cardiac Struct Metab **2**: 797-814.
Li Mura, I. E., B. Bauce, A. Nava, M. Fanciulli, G. Vazza, E. Mazzotti, I. Rigato, M. De Bortoli, G. Beffagna, A. Lorenzon, M. Calore, E. Dazzo, C. Nobile, M. L. Mostacciuolo, D. Corrado, C. Basso, L. Daliento, G. Thiene and A. Rampazzo (2013). "Identification of a PKP2 gene deletion in a family with arrhythmogenic right ventricular cardiomyopathy." Eur J Hum Genet **21**(11): 1226-1231.
Link, M. S., D. Laidlaw, B. Polonsky, W. Zareba, S. McNitt, K. Gear, F. Marcus and N. A. Estes, 3rd (2014). "Ventricular arrhythmias in the North American multidisciplinary study of ARVC: predictors, characteristics, and treatment." J Am Coll Cardiol **64**(2): 119-125.
Luckey, J. A. and L. M. Smith (1993). "Optimization of electric field strength for DNA sequencing in capillary gel electrophoresis." Anal Chem **65**(20): 2841-2850.
Maharaj, B. and M. G. Hammond (1990). "HLA-A, B, DR, and DQ antigens in black patients with idiopathic dilated cardiomyopathy." Am J Cardiol **65**(20): 1402-1403.

Marcus, F. I., W. J. McKenna, D. Sherrill, C. Basso, B. Bauce, D. A. Bluemke, H. Calkins, D. Corrado, M. G. Cox, J. P. Daubert, G. Fontaine, K. Gear, R. Hauer, A. Nava, M. H. Picard, N. Protonotarios, J. E. Saffitz, D. M. Sanborn, J. S. Steinberg, H. Tandri, G. Thiene, J. A. Towbin, A. Tsatsopoulou, T. Wichter and W. Zareba (2010). "Diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia: proposed modification of the Task Force Criteria." Eur Heart J **31**(7): 806-814.
Mardis, E. R. (2013). "Next-generation sequencing platforms." Annu Rev Anal Chem (Palo Alto Calif) **6**: 287-303.
Marian, A. J. and R. Roberts (2001). "The molecular genetic basis for hypertrophic cardiomyopathy." J Mol Cell Cardiol **33**(4): 655-670.
Marian, A. J., Q. T. Yu, D. L. Mann, F. L. Graham and R. Roberts (1995). "Expression of a mutation causing hypertrophic cardiomyopathy disrupts sarcomere assembly in adult feline cardiac myocytes." Circ Res **77**(1): 98-106.
Marin-Garcia, J., O. Zoubenko and M. J. Goldenthal (2002). "Mutations in the cardiac mitochondrial DNA control region associated with cardiomyopathy and aging." J Card Fail **8**(2): 93-100.
Mayosi, B. M. (2007). "Contemporary trends in the epidemiology and management of cardiomyopathy and pericarditis in sub-Saharan Africa." Heart **93**(10): 1176-1183.
Mayosi, B. M., S. Khogali, B. Zhang and H. Watkins (1999). "Cardiac and skeletal actin gene mutations are not a common cause of dilated cardiomyopathy." J Med Genet **36**(10): 796-797.
McGlashan, N. D. (1988). "Southern African cardiomyopathy in the Republic of South Africa, 1978-1980." Afr J Med Med Sci **17**(1): 33-46.
McKenna, W. J., G. Thiene, A. Nava, F. Fontaliran, C. Blomstrom-Lundqvist, G. Fontaine and F. Camerini (1994). "Diagnosis of arrhythmogenic right ventricular dysplasia/cardiomyopathy. Task Force of the Working Group Myocardial and Pericardial Disease of the European Society of Cardiology and of the Scientific Council on Cardiomyopathies of the International Society and Federation of Cardiology." Br Heart J **71**(3): 215-218.
McKoy, G., N. Protonotarios, A. Crosby, A. Tsatsopoulou, A. Anastasakis, A. Coonar, M. Norman, C. Baboonian, S. Jeffery and W. J. McKenna (2000). "Identification of a deletion in plakoglobin in arrhythmogenic right ventricular cardiomyopathy with palmoplantar keratoderma and woolly hair (Naxos disease)." Lancet **355**(9221): 2119-2124.
McNair, W. P., L. Ku, M. R. Taylor, P. R. Fain, D. Dao, E. Wolfel, L. Mestroni and G. Familial Cardiomyopathy Registry Research (2004). "SCN5A mutation associated with dilated cardiomyopathy, conduction disorder, and arrhythmia." Circulation **110**(15): 2163-2167.
McNally, E. M., J. R. Golbus and M. J. Puckelwartz (2013). "Genetic mutations and mechanisms in dilated cardiomyopathy." J Clin Invest **123**(1): 19-26.
Merner, N. D., K. A. Hodgkinson, A. F. Haywood, S. Connors, V. M. French, J. D. Drenckhahn, C. Kupprion, K. Ramadanova, L. Thierfelder, W. McKenna, B. Gallagher, L. Morris-Larkin, A. S. Bassett, P. S. Parfrey and T. L. Young (2008). "Arrhythmogenic right ventricular cardiomyopathy type 5 is a fully penetrant, lethal arrhythmic disorder caused by a missense mutation in the TMEM43 gene." Am J Hum Genet **82**(4): 809-821.
Mogensen, J., T. Kubo, M. Duque, W. Uribe, A. Shaw, R. Murphy, J. R. Gimeno, P. Elliott and W. J. McKenna (2003). "Idiopathic restrictive cardiomyopathy is part of the clinical expression of cardiac troponin I mutations." J Clin Invest **111**(2): 209-216.
Montanez, E., S. A. Wickstrom, J. Altstatter, H. Chu and R. Fassler (2009). "Alpha-parvin controls vascular mural cell recruitment to vessel wall by regulating RhoA/ROCK signalling." EMBO J **28**(20): 3132-3144.
Moolman-Smook, J. C., J. De Lange, A. Brink and A. Corfield (2000). "Hypertrophic cardiomyopathy repealing tenets in South Africa." Cardiovasc J S Afr **11**(4): 202-209.

Moolman-Smook, J. C., W. J. De Lange, E. C. Bruwer, P. A. Brink and V. A. Corfield (1999). "The origins of hypertrophic cardiomyopathy-causing mutations in two South African subpopulations: a unique profile of both independent and founder events." Am J Hum Genet **65**(5): 1308-1320.
Moolman-Smook, J. C., B. M. Mayosi, P. A. Brink and V. A. Corfield (2003). "Molecular genetics of cardiomyopathy: changing times, shifting paradigms." Cardiovasc J S Afr **14**(3): 145-155.
Moolman, J. C., V. A. Corfield, B. Posen, K. Ngumbela, C. Seidman, P. A. Brink and H. Watkins (1997). "Sudden death due to troponin T mutations." J Am Coll Cardiol **29**(3): 549-555.
Murphy, R. T., J. Mogensen, A. Shaw, T. Kubo, S. Hughes and W. J. McKenna (2004). "Novel mutation in cardiac troponin I in recessive idiopathic dilated cardiomyopathy." Lancet **363**(9406): 371-372.
Nelson, M. I., K. M. Stucker, S. A. Schobel, N. S. Trovao, S. R. Das, V. G. Dugan, S. W. Nelson, S. Sreevatsan, M. L. Killian, J. M. Nolting, D. E. Wentworth and A. S. Bowman (2016). "Introduction, evolution, and dissemination of influenza A viruses in exhibition swine, USA, 2009-2013." J Virol.
Nikolopoulos, S. N. and C. E. Turner (2000). "Actopaxin, a new focal adhesion protein that binds paxillin LD motifs and actin and regulates cell adhesion." J Cell Biol **151**(7): 1435-1448.
Norgett, E. E., S. J. Hatsell, L. Carvajal-Huerta, J. C. Cabezas, J. Common, P. E. Purkis, N. Whittock, I. M. Leigh, H. P. Stevens and D. P. Kelsell (2000). "Recessive mutation in desmoplakin disrupts desmoplakin-intermediate filament interactions and causes dilated cardiomyopathy, woolly hair and keratoderma." Hum Mol Genet **9**(18): 2761-2766.
Ntusi, N. B., M. Badri, F. Gumedze, A. Wonkam and B. M. Mayosi (2011). "Clinical characteristics and outcomes of familial and idiopathic dilated cardiomyopathy in Cape Town: a comparative study of 120 cases followed up over 14 years." S Afr Med J **101**(6): 399-404.
Olski, T. M., A. A. Noegel and E. Korenbaum (2001). "Parvin, a 42 kDa focal adhesion protein, related to the alpha-actinin superfamily." J Cell Sci **114**(Pt 3): 525-538.
Olson, T. M., S. Illenberger, N. Y. Kishimoto, S. Huttelmaier, M. T. Keating and B. M. Jockusch (2002). "Metavinculin mutations alter actin interaction in dilated cardiomyopathy." Circulation **105**(4): 431-437.
Olson, T. M., V. V. Michels, S. N. Thibodeau, Y. S. Tai and M. T. Keating (1998). "Actin mutations in dilated cardiomyopathy, a heritable form of heart failure." Science **280**(5364): 750-752.
Pare, J. A., R. G. Fraser, W. J. Pirozynski, J. A. Shanks and D. Stubington (1961). "Hereditary cardiovascular dysplasia. A form of familial cardiomyopathy." Am J Med **31**: 37-62.
Paulus, W. J., M. A. Goethals and S. U. Sys (1992). "Failure of myocardial inactivation: a clinical assessment in the hypertrophied heart." Basic Res Cardiol **87 Suppl 2**: 145-161.
Pilichou, K., A. Nava, C. Basso, G. Beffagna, B. Bauce, A. Lorenzon, G. Frigo, A. Vettori, M. Valente, J. Towbin, G. Thiene, G. A. Danieli and A. Rampazzo (2006). "Mutations in desmoglein-2 gene are associated with arrhythmogenic right ventricular cardiomyopathy." Circulation **113**(9): 1171-1179.
Pinto, J. R., M. S. Parvatiyar, M. A. Jones, J. Liang and J. D. Potter (2008). "A troponin T mutation that causes infantile restrictive cardiomyopathy increases Ca2+ sensitivity of force development and impairs the inhibitory properties of troponin." J Biol Chem **283**(4): 2156-2166.
Powell, S. J. and R. Wright (1965). "Cardiomyopathy in Durban." S Afr Med J **39**(42): 1062-1066.
Quarta, G., A. Muir, A. Pantazis, P. Syrris, K. Gehmlich, P. Garcia-Pavia, D. Ward, S. Sen-Chowdhry, P. M. Elliott and W. J. McKenna (2011). "Familial evaluation in arrhythmogenic right ventricular cardiomyopathy: impact of genetics and revised task force criteria." Circulation **123**(23): 2701-2709.
Raj, A., S. A. Rifkin, E. Andersen and A. van Oudenaarden (2010). "Variability in gene expression underlies incomplete penetrance." Nature **463**(7283): 913-918.
Rakar, S., G. Sinagra, A. Di Lenarda, A. Poletti, R. Bussani, F. Silvestri and F. Camerini (1997). "Epidemiology of dilated cardiomyopathy. A prospective post-mortem study of 5252 necropsies. The Heart Muscle Disease Study Group." Eur Heart J **18**(1): 117-123.

Rampazzo, A., A. Nava, G. A. Danieli, G. Buja, L. Daliento, G. Fasoli, R. Scognamiglio, D. Corrado and G. Thiene (1994). "The gene for arrhythmogenic right ventricular cardiomyopathy maps to chromosome 14q23-q24." Hum Mol Genet **3**(6): 959-962.
Rampazzo, A., A. Nava, S. Malacrida, G. Beffagna, B. Bauce, V. Rossi, R. Zimbello, B. Simionati, C. Basso, G. Thiene, J. A. Towbin and G. A. Danieli (2002). "Mutation in human desmoplakin domain binding to plakoglobin causes a dominant form of arrhythmogenic right ventricular cardiomyopathy." Am J Hum Genet **71**(5): 1200-1206.
Rathore, S. S., J. P. Curtis, Y. Wang, M. R. Bristow and H. M. Krumholz (2003). "Association of serum digoxin concentration and outcomes in patients with heart failure." JAMA **289**(7): 871-878.
Redwood, C. S., J. C. Moolman-Smook and H. Watkins (1999). "Properties of mutant contractile proteins that cause hypertrophic cardiomyopathy." Cardiovasc Res **44**(1): 20-36.
Richard, P., P. Charron, L. Carrier, C. Ledeuil, T. Cheav, C. Pichereau, A. Benaiche, R. Isnard, O. Dubourg, M. Burban, J. P. Gueffet, A. Millaire, M. Desnos, K. Schwartz, B. Hainque, M. Komajda and E. H. F. Project (2003). "Hypertrophic cardiomyopathy: distribution of disease genes, spectrum of mutations, and implications for a molecular diagnosis strategy." Circulation **107**(17): 2227-2232.
Richardson, P., W. McKenna, M. Bristow, B. Maisch, B. Mautner, J. O'Connell, E. Olsen, G. Thiene, J. Goodwin, I. Gyarfas, I. Martin and P. Nordet (1996). "Report of the 1995 World Health Organization/International Society and Federation of Cardiology Task Force on the Definition and Classification of cardiomyopathies." Circulation **93**(5): 841-842.
Rodriguez-Perez, J. M., J. M. Fragoso, E. Alvarez-Leon, N. Martinez-Rodriguez, G. J. Gallardo, S. Ines-Real, J. Granados, P. A. Reyes and G. Vargas-Alarcon (2007). "MHC class II genes in Mexican patients with idiopathic dilated cardiomyopathy." Exp Mol Pathol **82**(1): 49-52.
Ruggiero, A., S. N. Chen, R. Lombardi, G. Rodriguez and A. J. Marian (2013). "Pathogenesis of hypertrophic cardiomyopathy caused by myozenin 2 mutations is independent of calcineurin activity." Cardiovasc Res **97**(1): 44-54.
Salipante, S. J., T. Kawashima, C. Rosenthal, D. R. Hoogestraat, L. A. Cummings, D. J. Sengupta, T. T. Harkins, B. T. Cookson and N. G. Hoffman (2014). "Performance comparison of Illumina and ion torrent next-generation sequencing platforms for 16S rRNA-based bacterial community profiling." Appl Environ Microbiol **80**(24): 7583-7591.
Sanger, F., S. Nicklen and A. R. Coulson (1977). "DNA sequencing with chain-terminating inhibitors." Proc Natl Acad Sci U S A **74**(12): 5463-5467.
Schmidt, A. and S. Jager (2005). "Plakophilins--hard work in the desmosome, recreation in the nucleus?" Eur J Cell Biol **84**(2-3): 189-204.
Schotten, U., S. Voss, T. B. Wiederin, M. Voss, F. Schoendube, P. Hanrath and C. Schumacher (1999). "Altered force-frequency relation in hypertrophic obstructive cardiomyopathy." Basic Res Cardiol **94**(2): 120-127.
Schwarz, J. M., D. N. Cooper, M. Schuelke and D. Seelow (2014). "MutationTaster2: mutation prediction for the deep-sequencing age." Nat Methods **11**(4): 361-362.
Seftel, H. C. (1973). "Cardiomyopathies in Johannesburg Bantu." S Afr Med J **47**(8): 321-324.
Sen-Chowdhry, S., P. Syrris and W. J. McKenna (2007). "Role of genetic analysis in the management of patients with arrhythmogenic right ventricular dysplasia/cardiomyopathy." J Am Coll Cardiol **50**(19): 1813-1821.
Sepulveda, J. L. and C. Wu (2006). "The parvins." Cell Mol Life Sci **63**(1): 25-35.
Shendure, J. and H. Ji (2008). "Next-generation DNA sequencing." Nat Biotechnol **26**(10): 1135-1145.
Shoeman, R. L., C. Kesselmier, E. Mothes, B. Honer and P. Traub (1991). "Non-viral cellular substrates for human immunodeficiency virus type 1 protease." FEBS Lett **278**(2): 199-203.
Sliwa, K., A. Damasceno and B. M. Mayosi (2005). "Epidemiology and etiology of cardiomyopathy in Africa." Circulation **112**(23): 3577-3583.

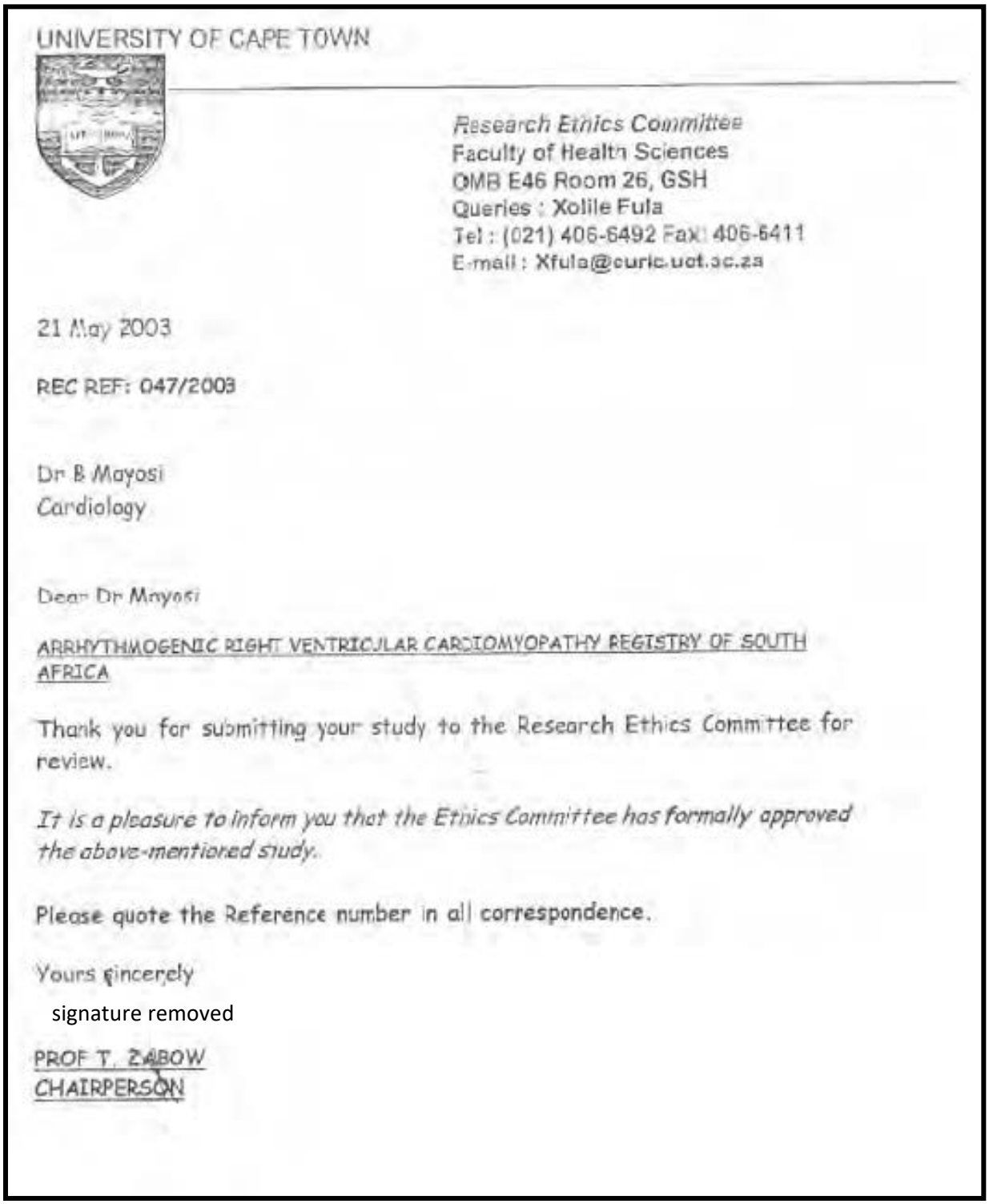
Sliwa, K., D. Wilkinson, C. Hansen, L. Ntyintyane, K. Tibazarwa, A. Becker and S. Stewart (2008). "Spectrum of heart disease and risk factors in a black urban population in South Africa (the Heart of Soweto Study): a cohort study." Lancet **371**(9616): 915-922.
Sliwa, K., A. Woodiwiss, E. Libhaber, F. Zhanje, C. Libhaber, R. Motara and R. Essop (2004). "C-reactive protein predicts response to pentoxifylline in patients with idiopathic dilated cardiomyopathy." Eur J Heart Fail **6**(6): 731-734.
Somura, F., H. Izawa, M. Iwase, Y. Takeichi, R. Ishiki, T. Nishizawa, A. Noda, K. Nagata, Y. Yamada and M. Yokota (2001). "Reduced myocardial sarcoplasmic reticulum Ca(2+)-ATPase mRNA expression and biphasic force-frequency relations in patients with hypertrophic cardiomyopathy." Circulation **104**(6): 658-663.
Sopko, N., Y. Qin, A. Finan, A. Dadabayev, S. Chigurupati, J. Qin, M. S. Penn and S. Gupta (2011). "Significance of thymosin beta4 and implication of PINCH-1-ILK-alpha-parvin (PIP) complex in human dilated cardiomyopathy." PLoS One **6**(5): e20184.
Spindler, M., K. W. Saupe, M. E. Christe, H. L. Sweeney, C. E. Seidman, J. G. Seidman and J. S. Ingwall (1998). "Diastolic dysfunction and altered energetics in the alphaMHC403/+ mouse model of familial hypertrophic cardiomyopathy." J Clin Invest **101**(8): 1775-1783.
Spirito, P., P. Bellone, K. M. Harris, P. Bernabo, P. Bruzzi and B. J. Maron (2000). "Magnitude of left ventricular hypertrophy and risk of sudden death in hypertrophic cardiomyopathy." N Engl J Med **342**(24): 1778-1785.
Steenekamp, J. H., I. W. Simson and W. Theron (1992). "Cardiovascular causes of death at Tshepong Hospital in 1 year, 1989-1990. A necropsy study." S Afr Med J **81**(3): 142-146.
Stein, H., M. H. Shnier, S. Wayburne and C. Isaacson (1964). "Cardiomyopathy in African Children." Arch Dis Child **39**: 610-617.
Stelzer, G., I. Plaschkes, D. Oz-Levi, A. Alkelai, T. Olender, S. Zimmerman, M. Twik, F. Belinky, S. Fishilevich, R. Nudel, Y. Guan-Golan, D. Warshawsky, D. Dahary, A. Kohn, Y. Mazor, S. Kaplan, T. Iny Stein, H. N. Baris, N. Rappaport, M. Safran and D. Lancet (2016). "VarElect: the phenotype-based variation prioritizer of the GeneCards Suite." BMC Genomics **17 Suppl 2**: 444.
Syrris, P., D. Ward, A. Evans, A. Asimaki, E. Gandjbakhch, S. Sen-Chowdhry and W. J. McKenna (2006). "Arrhythmogenic right ventricular dysplasia/cardiomyopathy associated with mutations in the desmosomal gene desmocollin-2." Am J Hum Genet **79**(5): 978-984.
Tabib, A., R. Loire, L. Chalabreysse, D. Meyronnet, A. Miras, D. Malicier, F. Thivolet, P. Chevalier and P. Bouvagnet (2003). "Circumstances of death and gross and microscopic observations in a series of 200 cases of sudden death associated with arrhythmogenic right ventricular cardiomyopathy and/or dysplasia." Circulation **108**(24): 3000-3005.
Tashiro, A., T. Masuda and I. Segawa (1990). "Morphometric comparison of mitochondria and myofibrils of cardiomyocytes between hypertrophic and dilated cardiomyopathies." Virchows Arch A Pathol Anat Histopathol **416**(6): 473-478.
Teare, D. (1958). "Asymmetrical hypertrophy of the heart in young adults." Br Heart J **20**(1): 1-8.
Thierfelder, L., H. Watkins, C. MacRae, R. Lamas, W. McKenna, H. P. Vosberg, J. G. Seidman and C. E. Seidman (1994). "Alpha-tropomyosin and cardiac troponin T mutations cause familial hypertrophic cardiomyopathy: a disease of the sarcomere." Cell **77**(5): 701-712.
Towbin, J. A. (1998). "The role of cytoskeletal proteins in cardiomyopathies." Curr Opin Cell Biol **10**(1): 131-139.
Towbin, J. A. and N. E. Bowles (2002). "The failing heart." Nature **415**(6868): 227-233.
Towbin, J. A., J. F. Hejtmancik, P. Brink, B. Gelb, X. M. Zhu, J. S. Chamberlain, E. R. McCabe and M. Swift (1993). "X-linked dilated cardiomyopathy. Molecular genetic evidence of linkage to the Duchenne muscular dystrophy (dystrophin) gene at the Xp21 locus." Circulation **87**(6): 1854-1865.

Tsatsopoulou, A. A., N. I. Protonotarios and W. J. McKenna (2006). "Arrhythmogenic right ventricular dysplasia, a cell adhesion cardiomyopathy: insights into disease pathogenesis from preliminary genotype--phenotype assessment." Heart **92**(12): 1720-1723.
van der Zwaag, P. A., M. G. Cox, C. van der Werf, A. C. Wiesfeld, J. D. Jongbloed, D. Dooijes, H. Bikker, R. Jongbloed, A. J. Suurmeijer, M. P. van den Berg, R. M. Hofstra, R. N. Hauer, A. A. Wilde and J. P. van Tintelen (2010). "Recurrent and founder mutations in the Netherlands : Plakophilin-2 p.Arg79X mutation causing arrhythmogenic right ventricular cardiomyopathy/dysplasia." Neth Heart J **18**(12): 583-591.
van Tintelen, J. P., M. M. Entius, Z. A. Bhuiyan, R. Jongbloed, A. C. Wiesfeld, A. A. Wilde, J. van der Smagt, L. G. Boven, M. M. Mannens, I. M. van Langen, R. M. Hofstra, L. C. Otterspoor, P. A. Doevendans, L. M. Rodriguez, I. C. van Gelder and R. N. Hauer (2006). "Plakophilin-2 mutations are the major determinant of familial arrhythmogenic right ventricular dysplasia/cardiomyopathy." Circulation **113**(13): 1650-1658.
Varnava, A., C. Baboonian, F. Davison, L. de Cruz, P. M. Elliott, M. J. Davies and W. J. McKenna (1999). "A new mutation of the cardiac troponin T gene causing familial hypertrophic cardiomyopathy without left ventricular hypertrophy." Heart **82**(5): 621-624.
Varnava, A. M., P. M. Elliott, C. Baboonian, F. Davison, M. J. Davies and W. J. McKenna (2001). "Hypertrophic cardiomyopathy: histopathological features of sudden death in cardiac troponin T disease." Circulation **104**(12): 1380-1384.
Vatta, M., F. Marcus and J. A. Towbin (2007). "Arrhythmogenic right ventricular cardiomyopathy: a 'final common pathway' that defines clinical phenotype." Eur Heart J **28**(5): 529-530.
Wang, X., K. Fukuda, I. J. Byeon, A. Velyvis, C. Wu, A. Gronenborn and J. Qin (2008). "The structure of alpha-parvin CH2-paxillin LD1 complex reveals a novel modular recognition for focal adhesion assembly." J Biol Chem **283**(30): 21113-21119.
Watkins, D. A., N. Hendricks, G. Shaboodien, M. Mbele, M. Parker, B. Z. Vezi, A. Latib, A. Chin, F. Little, M. Badri, J. C. Moolman-Smook, A. Okreglicki, B. M. Mayosi and A. R. o. t. C. A. S. o. S. Africa (2009). "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 **6**(11 Suppl): S10-17.
Watkins, D. A. and B. M. Mayosi (2009). "The contribution of South Africans to the subject of dilated cardiomyopathy - with reference to : cardiovascular collagenosis with parietal endocardial thrombosis : a clinicopathologic study of forty cases." Cardiovasc J Afr **20**(1): 11-16.
Watkins, H., W. J. McKenna, L. Thierfelder, H. J. Suk, R. Anan, A. O'Donoghue, P. Spirito, A. Matsumori, C. S. Moravec, J. G. Seidman and et al. (1995). "Mutations in the genes for cardiac troponin T and alpha-tropomyosin in hypertrophic cardiomyopathy." N Engl J Med **332**(16): 1058-1064.
Wichter, T., M. Borggrefe, W. Haverkamp, X. Chen and G. Breithardt (1992). "Efficacy of antiarrhythmic drugs in patients with arrhythmogenic right ventricular disease. Results in patients with inducible and noninducible ventricular tachycardia." Circulation **86**(1): 29-37.
Wimley, W. C. and S. H. White (1996). "Experimentally determined hydrophobicity scale for proteins at membrane interfaces." Nat Struct Biol **3**(10): 842-848.
Winder, W. W. and D. G. Hardie (1999). "AMP-activated protein kinase, a metabolic master switch: possible roles in type 2 diabetes." Am J Physiol **277**(1 Pt 1): E1-10.
Yates, A., W. Akanni, M. R. Amode, D. Barrell, K. Billis, D. Carvalho-Silva, C. Cummins, P. Clapham, S. Fitzgerald, L. Gil, C. G. Giron, L. Gordon, T. Hourlier, S. E. Hunt, S. H. Janacek, N. Johnson, T. Juettemann, S. Keenan, I. Lavidas, F. J. Martin, T. Maurel, W. McLaren, D. N. Murphy, R. Nag, M. Nuhn, A. Parker, M. Patricio, M. Pignatelli, M. Rahtz, H. S. Riat, D. Sheppard, K. Taylor, A. Thormann, A. Vullo, S. P. Wilder, A. Zadissa, E. Birney, J. Harrow, M. Muffato, E. Perry, M. Ruffier, G. Spudich, S. J. Trevanion, F. Cunningham, B. L. Aken, D. R. Zerbino and P. Flicek (2016). "Ensembl 2016." Nucleic Acids Res **44**(D1): D710-716.
Yumoto, F., Q. W. Lu, S. Morimoto, H. Tanaka, N. Kono, K. Nagata, T. Ojima, F. Takahashi-Yanaga, Y. Miwa, T. Sasaguri, K. Nishita, M. Tanokura and I. Ohtsuki (2005). "Drastic Ca2+ sensitization of

myofilament associated with a small structural change in troponin I in inherited restrictive cardiomyopathy." Biochem Biophys Res Commun **338**(3): 1519-1526.
Zhang, M., F. Tavora, J. B. Oliveira, L. Li, M. Franco, D. Fowler, Z. Zhao and A. Burke (2012). "PKP2 mutations in sudden death from arrhythmogenic right ventricular cardiomyopathy (ARVC) and sudden unexpected death with negative autopsy (SUDNA)." Circ J **76**(1): 189-194.
Zhou, X., M. Chen, H. Song, B. Wang, H. Chen, J. Wang, W. Wang, S. Feng, F. Zhang, W. Ju, M. Li, K. Gu, K. Cao, D. W. Wang and B. Yang (2015). "Comprehensive analysis of desmosomal gene mutations in Han Chinese patients with arrhythmogenic right ventricular cardiomyopathy." Eur J Med Genet **58**(4): 258-265.
Zou, J., D. Tran, M. Baalbaki, L. F. Tang, A. Poon, A. Pelonero, E. W. Titus, C. Yuan, C. Shi, S. Patchava, E. Halper, J. Garg, I. Movsesyan, C. Yin, R. Wu, L. D. Wilsbacher, J. Liu, R. L. Hager, S. R. Coughlin, M. Jinek, C. R. Pullinger, J. P. Kane, D. O. Hart, P. Y. Kwok and R. C. Deo (2015). "An internal promoter underlies the difference in disease severity between N- and C-terminal truncation mutations of Titin in zebrafish." Elife **4**: e09406.

APPENDICES

Appendix 1: Approval letter from Ethics committee



UNIVERSITY OF CAPE TOWN

Research Ethics Committee
Faculty of Health Sciences
OMB E46 Room 26, GSH
Queries : Xolile Fula
Tel : (021) 406-6492 Fax: 406-6411
E-mail : Xfula@curic.uct.ac.za

21 May 2003

REC REF: 047/2003

Dr B Mayosi
Cardiology

Dear Dr Mayosi

ARRHYTHMOGENIC RIGHT VENTRICULAR CARDIOMYOPATHY REGISTRY OF SOUTH AFRICA

Thank you for submitting your study to the Research Ethics Committee for review.

It is a pleasure to inform you that the Ethics Committee has formally approved the above-mentioned study.

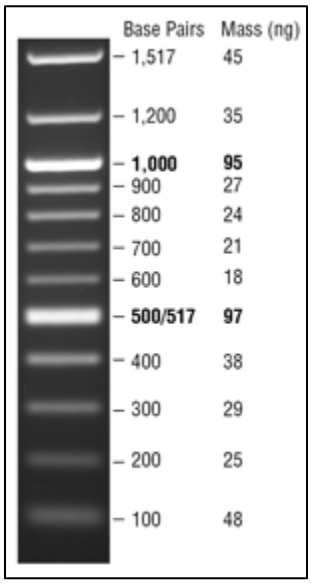
Please quote the Reference number in all correspondence.

Yours sincerely

signature removed

PROF T. ZABOW
CHAIRPERSON

Appendix 2: New England Biolabs^{R} 100bp DNA ladder



Appendix 3: ARVC study patient consent form

UNIVERSITY OF CAPE TOWN

Arrhythmogenic Right Ventricular Cardiomyopathy Registry of South Africa
Department of Medicine
Cardiology Research Office
Room H46/47 OMB
Groote Schuur Hospital
OBSERVATORY. 7925
Tel: +27-21-447 2777
Fax: +27-21-447 2765 **Email: veronica.francis@uct.ac.za**

Date___/___/_____
DD MM YEAR

Informed Consent Form

I agree to participate in the study of Arrhythmogenic Right Ventricular Cardiomyopathy (ARVC), a heart muscle disease (cardiomyopathy) that predominantly affects the right ventricle and may lead to ventricular arrhythmias and heart failure. I understand that I will be interviewed about my medical history, family history and medications, and that I may undergo further cardiac investigations. However, I will only undergo those investigations necessary to confirm or exclude the diagnosis of ARVC. In addition, I will have a blood sample drawn consisting of 5 tubes for a total of 25 ml of blood. This blood will be used to test for & study the genetic factors that cause ARVC. After my initial consultation, I agree to be followed up annually to assess if there have been any changes in my medical history and medications. As ARVC is a familial disease in 30-50% of people, I agree to have my first-degree relatives contacted for clinical screening & participation in this study. However, each of them will be counselled about ARVC and will also have to sign a consent form before they will be included in this study.

I understand that my participation in this study is entirely voluntary. All information gathered in this study is strictly confidential, and will only be used for research relating to the study of ARVC. This information (including genetic material) will not be used to generate a profit. As well, genetic material will not be used for the purpose of gene alteration, and, prior to blood sampling, I will sign a separate DNA consent form that governs the use of genetic material under the rules of the University of Cape Town Research Ethics Committee. I will not be identified in any published report. I am free to refuse to participate or withdraw from the study at any time, without jeopardising my future care. If I have any questions, I may contact _____at_____.

I agree to participate in the study and I have been given a copy of this form.

Subject name	Subject signature	Date
Witness name	Witness signature	Date
Investigator name	Investigator signature	Date

ARVCptcons.doc Feb2009.vers1.0

Appendix 4: Puregene DNA extraction protocol

Protocol: DNA Purification from Buffy Coat Using the Gentra Puregene Blood Kit

This protocol is for purification of genomic DNA from buffy coat prepared from 3 ml whole blood using the Gentra Puregene Blood Kit.

Buffy Coat

Things to do before starting

- Preheat water bath to 65°C for use in step 19 of the procedure.
- Buffy coat is a leukocyte-enriched fraction of whole blood. Preparing a buffy coat fraction from whole blood is simple and yields approximately 5–10 times more DNA than an equivalent volume of whole blood. Prepare buffy coat by centrifuging whole blood at 2500 x *g* for 10 minutes at room temperature (15–25°C). After centrifugation, three different fractions are distinguishable: the upper clear layer is plasma; the intermediate layer is buffy coat, containing concentrated leukocytes; and the bottom layer contains concentrated erythrocytes.
- Frozen buffy coat should be thawed quickly in a 37°C water bath with mild agitation and stored on ice before beginning the procedure.
- Optional: Preheat water bath to 37°C for use in step 8 of the procedure.

Procedure

1. **If the buffy coat preparation contains red blood cells, continue with step 2. Otherwise, pipet 3 ml Cell Lysis Solution into a 15 ml centrifuge tube, add 150–250 µl sample, and continue with step 8.**
2. **Dispense 3 volumes RBC Lysis Solution into a 15 ml centrifuge tube (e.g., if processing 250 µl buffy coat, dispense 750 µl RBC Lysis Solution). Add 150–250 µl buffy coat preparation.**
3. **Invert to mix, and incubate for 10 min at room temperature (15–25°C). Invert again at least once during the incubation.**
4. **Centrifuge for 5 min at 2000 x *g*.**
5. **Carefully discard the supernatant by pipetting or pouring, leaving approximately 100–200 µl of the residual liquid and the pellet.**
6. **Vortex the tube vigorously to resuspend the pellet in the residual liquid.**
   Vortexing greatly facilitates cell lysis in the next step.
   The pellet should be completely dispersed after vortexing.

Buffy Coat

7. **Add 3 ml Cell Lysis Solution, and pipet up and down or vortex vigorously to lyse the cells.**
   Usually no incubation is required; however, if cell clumps are visible, incubate at 37°C until the solution is homogeneous.
   Samples are stable in Cell Lysis Solution for at least 2 years at room temperature.
8. **Optional: If RNA-free DNA is required, add 15 µl RNase A Solution, and mix by inverting 25 times. Incubate for 15 min at 37°C. Then incubate for 3 min on ice to quickly cool the sample.**
9. **Add 1 ml Protein Precipitation Solution, and vortex vigorously for 20 s at high speed.**
10. **Centrifuge for 5 min at 2000 x *g*.**
    The precipitated proteins should form a tight, white or brown pellet. If the protein pellet is not tight, incubate on ice for 5 min and repeat the centrifugation.
11. **Pipet 3 ml isopropanol into a clean 15 ml centrifuge tube and add the supernatant from the previous step by pouring carefully.**
    Be sure the protein pellet is not dislodged during pouring.
12. **Mix by inverting gently 50 times.**
13. **Centrifuge for 3 min at 2000 x *g*.**
    The DNA will be visible as a small white pellet.
14. **Carefully discard the supernatant, and drain the tube by inverting on a clean piece of absorbent paper, taking care that the pellet remains in the tube.**
15. **Add 3 ml of 70% ethanol and invert several times to wash the DNA pellet.**
16. **Centrifuge for 1 min at 2000 x *g*.**
17. **Carefully discard the supernatant. Drain the tube on a clean piece of absorbent paper, taking care that the pellet remains in the tube. Allow to air dry for 5–10 min.**
    The pellet might be loose and easily dislodged. Avoid over-drying the DNA pellet, as the DNA will be difficult to dissolve.
18. **Add 300 µl DNA Hydration Solution and vortex for 5 s at medium speed to mix.**
19. **Incubate at 65°C for 1h to dissolve the DNA.**
20. **Incubate at room temperature overnight with gentle shaking. Ensure tube cap is tightly closed to avoid leakage. Samples can then be centrifuged briefly and transferred to a storage tube.**

Buffy Coat

Appendix 5: Bioinformatics tools methodology

MutationTaster employs a Bayes classifier to eventually predict the disease potential of an alteration based on the input which is the name of the gene and the particular base alteration. The Bayes classifier calculates probabilities for the alteration to be either a disease mutation or a harmless polymorphism. The output is a probability value which represents the probability of the prediction whereby a value close to 1 indicates a high 'security' of the prediction

SIFT prediction tool predicts the effect of a nonsynonymous coding single base variant and the input includes variant position and change in form of chromosomal coordinates. The coordinates are aligned against the Homo sapiens GRCh37 Ensembl63 assembly and the output is a prediction value which can either be damaging or tolerated with a SIFT score based on the effect of the variant on protein effect.

PolyPhen-2 employs structural and comparative evolutionary considerations to predict the possible impact of amino acid substitutions on the stability and function of human proteins. It performs functional annotation of single-nucleotide polymorphisms (SNPs), maps coding SNPs to gene transcripts, extracts protein sequence annotations and structural attributes, and builds conservation profiles all made possible due to a high-quality multiple protein sequence alignment pipeline and a prediction method employing machine-learning classification. It then estimates the probability of the missense mutation being damaging based on a combination of all these properties. The prediction outcome can be one of benign, possibly damaging or probably damaging. A “Score” output accompanying the outcome is the probability of the substitution being damaging; “sensitivity” and “specificity” correspond to prediction confidence (Adzhubei, Jordan et al. 2013).