A Candidate Gene Analysis of Arrhythmogenic Right Ventricular Cardiomyopathy (ARVC)

by

Janine du Preez

DPRJAN003

SUBMITTED TO THE UNIVERSITY OF CAPE TOWN

In fulfilment of the requirements for the degree

MSc (Med) in Human Genetics

Faculty of Health Sciences

UNIVERSITY OF CAPE TOWN

15 March 2006

Supervisor: Prof. Bongani M. Mayosi, Department of Medicine, University of Cape Town

DECLARATION

I, Janine du Preez, hereby declare that the work on which this dissertation 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: .....

Date: 15 March 2006

ACKNOWLEDGEMENTS

I am grateful to the **patients and family members** who participated in this study.

I would like express my gratitude to my supervisor, **Professor Bongani Mayosi,** for his guidance, insight and enthusiasm throughout this project. His input in the writing of this dissertation has been invaluable.

To all at the **Division of Human Genetics**, you make the Division what it is. A big THANK-YOU to **George Rebello**, **Lisa Roberts**, **Alvera Vorster** and **Janine Scholefield** for all their help and advice.

Special thanks go to **Assoc. Prof. Jacquie Greenberg** for ensuring that I keep "head above water".

I am grateful to the **Department of Medicine**, the **University of Cape Town**, **Ernst and Ethel Eriksen Trust, Guy Elliott** and **Shell** for their financial assistance.

I thank **my parents** for their support and encouragement throughout my studies.

To my fiancé, **Rory**, for all his support, encouragement and understanding.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

°C = degrees Celsius

AD = Autosomal Dominant

ALVC = Arrhythmogenic Left Ventricular Cardiomyopathy

APC = Anterior Polar Cataract

APS = ammonium persulphate

AR = Autosomal Recessive

ARVC = Arrhythmogenic Right Ventricular Cardiomyopathy

AV = Atrioventricular

BLAST = Basic Local Alignment Search Tool

bp = Base pairs

Ca^{2+} = Calcium

CAD = Coronary artery disease

CASSA = Cardiac Arrhythmia Society of South Africa

CBX5 = Chromobox protein homolog 5

cm = Centimetre

CPVT = Catecholaminergic Polymorphic Ventricular Tachycardia

CVD = Cardiovascular disease

DCM = Dilated cardiomyopathy

dH_20 = Distilled water

DHPLC = Denaturing high-performance liquid chromatography

DM = Dense midline

DNA = Deoxyribonucleic acid

dNTPs = Deoxytrinucleotide phosphate

DSC = Desmocollin

DSG2 = Desmoglein-2

DSP = Desmoplakin

E = Exon

EC coupling = Excitation-contraction coupling

ECG = Electrocardiogram

F = Forward primer

H_20 = Water

HCl = Hydrochloric acid

HCM = Hypertrophic cardiomyopathy

HGVS = Human Genome Variation Society

HP1 = Heterochromatin protein 1

I = Intron

ICD = Implantable cardioverter defibrillator

IDP = Inner dense plaque

IFs = Intermediate filaments

ISFC = International Society and Federation of Cardiology

JUP = Plakoglobin

kcal/mol = Kilocalories per mole

KCl = Potassium chloride

Lamr1 = Laminin receptor 1 *Mus musculus*

LAMR1 = Laminin receptor 1 *Homo sapiens*

LV = Left Ventricle

LVNC = Left Ventricular Non-Compaction

$MgCl_2$ = Magnesium chloride

MIM = Mendelian Inheritance in Man

µg = Microgram

µl	= Microlitre
µM	= Micromolar
ml	= Milliliter
mM	= Millimolar
MRC	= Medical Research Council
MRI	= Magnetic resonance image
NCBI	= National Center for Biotechnology Information
ODP	= Outer dense plaque
PAGE	= Polyacrylamide gel electrophoresis
PCR	= Polymerase chain reaction
PKP2	= Plakophilin-2
PM	= Plasma membrane
R	= Reverse primer
RCM	= Restrictive cardiomyopathy
RV	= Right ventricle
RyR2	= Cardiac ryanodine receptor
SV	= Sequence variants
Ta	= Annealing temperature
TBE buffer	= Tris Borate EDTA buffer
TEMED	= N'N'N'N'-tetramethylethylenediamine
TGFβ3	= Transforming growth factor-β3
UCT	= University of Cape Town
UTR	= Untranslated region
V	= Volts
WHO	= World Health Organisation
w/v	= Weight per volume

ABSTRACT

Background: Heart failure is a major public health problem throughout the world. In South Africa 17% of mortality is attributed to cardiovascular disease (CVD). Heart failure may be either ischemic or non-ischemic in origin. A significant proportion of non-ischemic heart failure is due to cardiomyopathy. There are currently five types of cardiomyopathy recognised, of which arrhythmogenic right ventricular cardiomyopathy (ARVC) is one. ARVC is familial in 30 to 50% of cases and it is inherited in an autosomal dominant or an autosomal recessive manner. Twelve chromosomal loci have been linked to ARVC and six genes have been identified. In 2004 Asano and colleagues reported a mouse model of ARVC that established *LAMR1* and *CBX5* as candidate genes for the human form of ARVC.

Methods: *LAMR1* and *CBX5* were screened for mutations in a cohort of 22 unrelated index patients, from the ARVC Registry of South Africa. The patients are from different ethnic groups of which Caucasian is the majority. *LAMR1* was screened by sequencing and *CBX5* was screened by DHPLC on the WAVE System. The sequence variants identified in *LAMR1* were subjected to haplotype analysis.

Results: A total of 17 sequence variants (SVs) were identified in *LAMR1* and were named SV 1-17. Of these variants 10 (i.e., SV 1, 2, 5, 6, 8, 9, 10, 14, 16 and 17) were known as SNPs on the NCBI SNP database. None of the seven novel variants (i.e., SV 3, 4, 7, 11, 12, 13 and 15) were predicted to have a pathogenic effect on protein structure or function. SV 3 is a single base substitution in the promoter region of *LAMR1*, 140 bp upstream to exon 1. SV 4 was found to be a common polymorphism with a carrier frequency of 4.6% in the background population. SV 7, 13 and 15 are intronic single base substitutions and SV 11 is an exonic single base substitution that does not change the amino acid. SV 12 was homozygous in one patient with apparent recessive ARVC but heterozygous in an unaffected and affected family member. SV 12 was therefore thought not to be disease causing. Haplotype analysis of *LAMR1* reveals that this gene contains either one or two haplotype blocks depending on the block definition used. Two haplotypes account

for nearly half of all the chromosomes. The haploytype analysis is however limited as it was calculated based on a small group of patients from different ethnic origins.

In *CBX5* a known SNP was identified in exon 2. Aberrant chromatographic profiles were identified in exons 3, 4 and 5. No sequence variant was identified in exon 4 upon sequencing. The cause of the aberrant profiles in exons 3 and 5 was not identified as sequencing with either the forward or reverse primer did not produce good quality sequence. These sequence variants however do not appear to be disease causing. The SV in *CBX5* exon 3 appears to be a common polymorphism and the SV in exon 5 was not identified in the patient's affected sibling.

Conclusion: No conclusive evidence of disease-causing mutations for ARVC was found in the *LAMR1* gene. Functional studies on the sequence variants are however required to conclude if *LAMR1* is a disease gene of ARVC. Preliminary results for the *CBX5* mutation screen did not reveal a disease causing mutation in the South African patients.

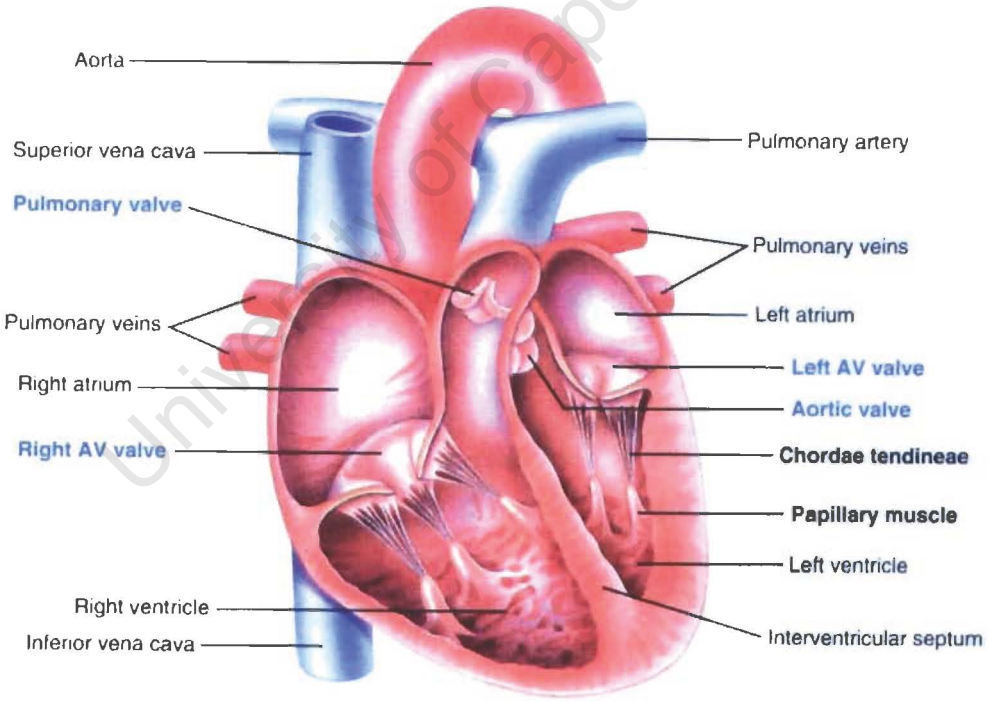
Chapter 1:

Introduction

1 INTRODUCTION

1.1 Structure and Function of the Heart

The heart is the first functional organ in the developing embryo and it starts beating within three weeks after conception (Gilbert 2000; Sherwood 2004). The heart is about the size of a clenched fist and it is located in the thoracic cavity between the sternum and the vertebrae. It is divided into a right and a left half by the atrioventricular septum. Each half is further divided into an upper and lower chamber called the atrium and ventricle respectively (Sherwood 2004). An atrioventricular (AV) valve is positioned between each atrium and ventricle and a semilunar valve is located between each ventricle and the artery leading out from it (Starr and Taggart 1998). The heart valves help to maintain a one-way flow of blood (Van De Graaff 2002). The structure of the heart is illustrated in Figure 1-1.

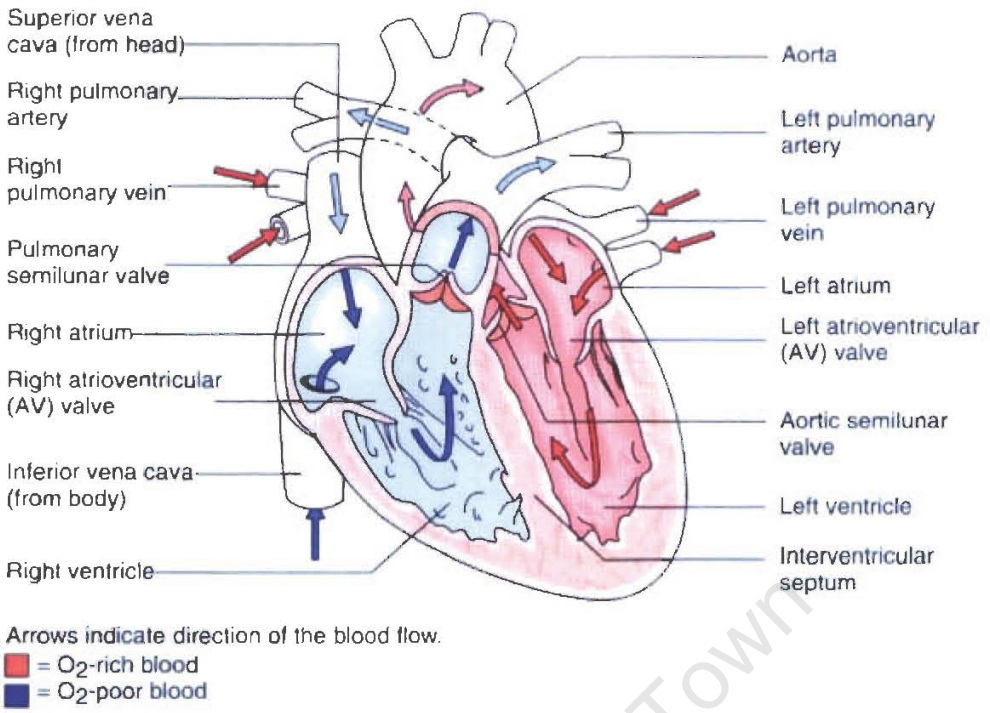


(Sherwood 2004)

Figure 1-1: Longitudinal section of the heart, depicting the four heart chambers.

The heart beats an estimated 42 million times a year (Van De Graaff 2002). Each time the heart beats, its four chambers go through phases of contraction (systole) and relaxation (diastole). When the heart is relaxed, the atria fill with blood. The increase in fluid pressure in the atria forces the AV valves open. Blood flows into the ventricles, which completely fill when the atria contract. As the filled ventricles start to contract, the rising fluid pressure forces the AV valves shut, preventing the backflow of blood into the atria. The pressure in the ventricles rises above that in the great arteries, forcing the semilunar valves open and the blood leaves the heart. The ventricles then relax, the semilunar valves close, and the atria fill with blood again to repeat the cycle (Starr and Taggart 1998).

Although the left and the right sides of the heart contract and relax together, the two sides function as two separate pumps. Deoxygenated blood, returning from systemic circulation, enters the right atrium via the vena cavae. This O_2-poor blood flows from the right atrium into the right ventricle, which pumps it out through the pulmonary artery to the lungs. Within the lungs, the blood becomes oxygenated and it is returned to the heart via the pulmonary veins. This O_2-rich blood enters the left atrium and flows into the left ventricle, which pumps it out through the aorta. Major arteries branch from the aorta to supply the various tissues of the body (Sherwood 2004). The blood flow through the heart is illustrated in Figure 1-2.



(Sherwood 2004)

Figure 1-2: Longitudinal section of the heart, depicting blood flow through the heart.

1.2 Cardiomyopathies

Heart failure is defined as a clinical syndrome of effort intolerance that is associated with fluid retention due to cardiac dysfunction. Heart failure is a major public health problem throughout the world (Seidman and Seidman 2001). In Europe, 50 million of the population of 1,000 million suffer from heart failure, and in the USA, 4.9 million patients are treated for heart failure each year (Moolman-Smook et al. 2003). In South Africa, similar statistics are not available however the Medical Research Council (MRC) reports that deaths from all forms of cardiovascular disease (CVD) account for 17% of total mortality (Bradshaw et al. 2003).

Heart failure may be either ischemic or non-ischemic in origin. The cause of a significant proportion of non-ischemic heart failure is cardiomyopathy (Moolman-Smook et al. 2003). In 1980 a Task Force, set up by the World Health Organisation

(WHO) and the International Society and Federation of Cardiology (ISFC), defined cardiomyopathy as a "heart muscle disease of unknown cause" (Thiene et al. 2004). In 1995 the Task Force redefined cardiomyopathy as a "disease of the myocardium associated with cardiac dysfunction" to reflect the new genetic discoveries (Thiene et al. 2005). There are currently five types of cardiomyopathy namely: dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM), restrictive cardiomyopathy (RCM), arrhythmogenic right ventricular cardiomyopathy (ARVC) and unclassified cardiomyopathy.

1.2.1 Dilated Cardiomyopathy (DCM)

DCM is characterised by dilatation and impaired contraction of the left ventricle or both ventricles. It may be idiopathic, familial/genetic, viral and/or immune, alcoholic/toxic or associated with recognised cardiovascular disease (Richardson et al. 1996). DCM may also be associated with other organ or muscle abnormalities. Familial DCM may be inherited via autosomal dominant, autosomal recessive, X-linked or mitochondrial transmission (Priori et al. 1999). The disease genes identified encode cytoskeletal proteins related to force transmission e.g. dystrophin (Thiene et al. 2005).

1.2.2 Hypertrophic Cardiomyopathy (HCM)

HCM is characterised clinically by unexplained cardiac hypertrophy. Hypertrophy is usually asymmetric and involves the interventricular septum. Familial disease with autosomal dominant inheritance predominates (Richardson et al. 1996). Most of the mutations have been found in genes encoding sarcomeric proteins e.g. β-myosin heavy chain, thus impairing force production (Thiene et al. 2005).

1.2.3 Restrictive Cardiomyopathy (RCM)

RCM is characterised by restrictive filling and reduced diastolic volume of either or both ventricles. Systolic function and wall thickness are normal or near-normal

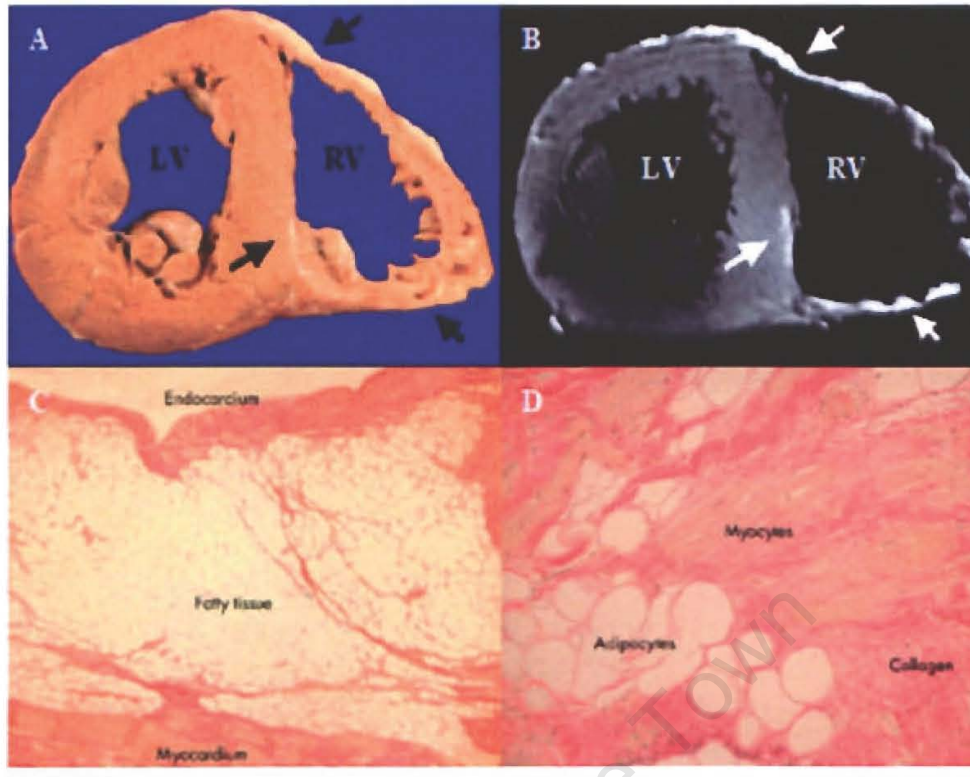
(Richardson et al. 1996). RCM may be primary or secondary. Primary disease includes idiopathic RCM, which may be characterised by skeletal myopathy and autosomal dominant inheritance (Hughes and McKenna 2005). Idiopathic RCM appears to be a sarcomeric disease, like HCM, due to mutation of the cardiac troponin I (Thiene et al. 2005).

1.2.4 Arrhythmogenic Right Ventricular Cardiomyopathy (ARVC)

Two pathological forms of ARVC have been reported, a fatty and fibrofatty type. The **fatty variant** is found only in the right ventricle. The myocardium is partially or almost completely replaced with fat but there is no thinning of the right ventricular wall. In the **fibrofatty variant**, the myocardium is replaced with fat and fibrous tissue which leads to a thinning of the right ventricular wall. This form is associated with inflammatory infiltrates (Thiene et al. 1991; Wolfe and Corwin 2005). ARVC is characterised by progressive fatty or fibrofatty replacement of right ventricular myocardium (Richardson et al. 1996). Fatty or fibrofatty infiltration usually affects the right ventricular outflow tract, the apex and the inflow tract (subtricuspid area) which is known as the "triangle of dysplasia". The septum is usually spared (Paul et al. 2003).

Most ARVC patients are male with a male:female ratio of 2.7:1. The male predominance is unexplained (Wolfe and Corwin 2005). Figure 1-3 shows the macroscopic structure, magnetic resonance image (MRI) and histopathology of a heart with ARVC.

It is thought that ARVC may be a milder form of Uhl's anomaly. This disease is characterised by partial or complete absence of the right ventricular myocardium and it presents in infancy as heart failure. The pathogenesis of Uhl's anomaly has not been determined (Dokuparti et al. 2005).



(John et al. 2004)

Figure 1-3: A, A short axis slice of a heart with ARVC. Yellow streaks of fat are indicated with arrows. **B**, A MRI of the same slice, which shows an increased signal intensity corresponding to the yellow streaks seen in *A*. **C&D**, A histopathology specimen taken from the RV showing large numbers of fat cells. LV = Left Ventricle, RV = Right Ventricle

1.2.5 Unclassified Cardiomyopathy

Cases that do not fit into any of the above mentioned groups, e.g., isolated left ventricular non-compaction (LVNC), fall into the group unclassified cardiomyopathy (Richardson et al. 1996; Hughes and McKenna 2005).

This literature review will, from here on, focus on ARVC as this is the disease under investigation

1.3 Historical Background of ARVC

In 1961, Dalla-Volta and colleagues described an "auricularization of the right ventricular curve" in a patient with "primitive endocardial fibrosis" (Dalla Volta et al. 1961; Paul et al. 2003). In 1977, Fontaine and colleagues consolidated these findings and called it "right ventricular dysplasia" (i.e., an abnormality of development) (Paul et al. 2003; Michael et al. 2004). Once the cause of this disease was identified as progressive myocardial loss, the term "cardiomyopathy" was considered more appropriate than "dysplasia" (Thiene et al. 1991). In 1995, ARVC was added to list of cardiomyopathies by the WHO/ISFC Task Force (Richardson et al. 1996). The first case series of ARVC was reported for the first time in South Africa in 2000 by Munclinger and colleagues (Munclinger et al. 2000).

1.4 Epidemiology of ARVC

The exact prevalence of ARVC is unknown due to incomplete penetrance of the disease and under diagnosis by clinicians and pathologists (Paul et al. 2003; Wolfe and Corwin 2005).

In Italy, preliminary studies showed ARVC to account for 20% of all sudden deaths in individuals under the age of 35 years and 27% of sudden deaths in young athletes (Thiene et al. 1988; Corrado et al. 1990). In a 21-year prospective cohort study, ARVC and coronary artery disease (CAD) were associated with the greatest risk of sudden death in Italian athletes (Corrado et al. 2003). In the USA, ARVC accounts for only 2% of sudden death in athletes while HCM is the most common cause of sudden death in athletes causing 30% of cases (Maron et al. 1996).

The regional variation in ARVC prevalence may be due to two reasons. First, the genetic predisposition, dietary pattern and environmental conditions of Italians are different from those of Americans (Williams and Chen 2003). Second, in the 21-year Italian study, the examination of all hearts was performed by the same team of experienced cardiovascular pathologists (Corrado et al. 2003).

The prevalence of ARVC was determined as 1 in 1000 inhabitants served by the Academic Teaching Hospital in Quedlinburg, Germany (Peters et al. 2004). It was a retrospective study over a 5-year time period which revealed a much higher prevalence than that of other centres.

The scarcity of reports on ARVC in Africa is thought to be due to the lack of sophisticated cardiac electrophysiology facilities and expertise required for the diagnosis of the disease (Sliwa et al. 2005). An initial report from the ARVC Registry of South Africa suggests that ARVC occurs in all segments of the population and that its clinical features and prognosis are similar to international findings (Latib et al. 2004; Sliwa et al. 2005).

1.5 Diagnosis of ARVC

In 1994, an "International Task Force for ARVC" proposed standardised diagnostic criteria (McKenna et al. 1994). The diagnosis is based on major and minor criteria that encompass clinical, structural and functional features (Hodgkinson et al. 2005). A list of the diagnostic criteria is given in Table 1-1. A patient is diagnosed with ARVC if they have two major criteria, one major plus two minor criteria or four minor criteria from different diagnostic categories (McKenna et al. 1994).

Although the International Task Force criteria facilitate clinical diagnosis, they lack sensitivity for early disease when clinical findings are often subtle (Syrris et al. 2006). The experience of hypertrophic and dilated cardiomyopathy has been that although many relatives may have phenotypic abnormalities that are "nondiagnostic", they are nevertheless indicative of disease (Hamid et al. 2002). In a study of 67 families Hamid and colleagues found familial ARVC present in 28% of index patients. When all cardiovascular abnormalities are considered, 48% had evidence for familial ARVC. This lead Hamid and colleagues to propose broader diagnostic criteria for familial ARVC (Hamid et al. 2002).

In 2006, Syrris and colleagues screened 100 white European ARVC patients from the United Kingdom for mutations in plakophilin-2 (*PKP2*) (Syrris et al. 2006). Mutations were identified in 11 patients and their families were investigated, both clinically and genetically, to identify genotype-phenotype correlations. It was however found that the clinical expression of *PKP2* mutations is heterogeneous even among first-degree relatives. A retrospective evaluation of the genotyped family members revealed that strict adherence to the International Task Force diagnostic criteria would have led to the false-negative diagnosis of probands and relatives that are in fact mutation-positive but exhibit incomplete disease expression. Use of the modified diagnostic criteria, proposed by Hamid and colleagues, did expand the diagnostic yield but it also led to false-positives, i.e., several family members assumed to be clinically affected were found to be mutation-negative. This study highlights the need for a more accurate set of diagnostic criteria. The identification of all ARVC disease genes would aid diagnosis of asymptomatic mutation carriers at risk of developing the disease (Syrris et al. 2006).

Table 1-1: Criteria for diagnosis of ARVC

	Major criteria	Minor criteria
1. Global and/or regional dysfunction and structural alterations	Severe dilatation and reduction of right ventricular ejection fraction with no (or only mild) LV impairment Localised right ventricular aneurysms (akinetic or dyskinetic areas with diastolic bulging) Severe segmental dilation of the right ventricle	Mild global right ventricular dilatation and/or ejection fraction reduction with normal left ventricle Mild segmental dilatation of the right ventricle Regional right ventricular hypokinesia
2. Tissue characterisation of walls	Fibrofatty replacement of myocardium on endomyocardial biopsy	
3. Repolarisation abnormalities		Inverted T waves in right precordial leads (V2 and V3) (people aged more than 12 yr; in absence of right bundle branch block)
4. Depolarisation/conduction abnormalities	Epsilon waves or localised prolongation (>110 ms) of the QRS complex in right precordial leads (V1-V3)	Late potentials (signal averaged ECG)
5. Arrhythmias		Left bundle branch block type ventricular tachycardia (sustained and non-sustained) (ECG, Holter, exercise testing) Frequent ventricular extrasystoles (more than 1000/24h) (Holter)
6. Family history	Familial disease confirmed at necropsy or surgery	Familial history of premature sudden death (<35yr) due to suspected right ventricular dysplasia Familial history (clinical diagnosis based on present criteria)

(McKenna et al. 1994)

* Detected by echocardiography, angiography, magnetic resonance imaging or radionuclide scintigraphy. ECG, electrocardiogram; LV, left ventricle

1.6 Treatment of ARVC

Treatment of ARVC is focused on controlling ventricular arrhythmias and preventing sudden death (Frias 2005).

Patients with well-tolerated, non-life-threatening ventricular arrhythmias are treated with anti-arrhythmic drugs. Beta-blockers alone or in combination with sotalol or amiodarone may be used (Campbell 2005). ARVC patients should avoid class Ia antiarrhythmic agents such as propafenone which has been associated with incessant ventricular tachycardia in these patients (Michael et al. 2004).

Patients with life-threatening ventricular arrhythmias are best treated with an implantable cardioverter defibrillator (ICD). A study by Hodgkinson et al. demonstrated that ICD therapy improved survival of males. The effectiveness of ICD therapy in preventing sudden cardiac death in women was however not demonstrated (Hodgkinson et al. 2005). This supports the clinical observation of a milder disease expression and better long-term prognosis of ARVC in women (Wichter and Breithardt 2005).

In patients with severe right ventricular or biventricular failure, treatment consists of pharmacologic therapy for heart failure and heart transplantation (Michael et al. 2004; Frias 2005).

1.7 Genetics of ARVC

ARVC is familial in 30 to 50% of cases and it is inherited in an autosomal dominant or an autosomal recessive manner (Hamid et al. 2002; Michael et al. 2004). Twelve chromosomal loci have been linked to ARVC, and six genes have been identified, namely plakoglobin (*JUP*), cardiac ryanodine receptor (*RyR2*), desmoplakin (*DSP*), plakophilin-2 (*PKP2*), transforming growth factor-β3 (*TGFβ3*) and desmoglein-2 (*DSG2*). The chromosomal loci and disease genes are given in Table 1-2.

Table 1-2: Chromosomal loci and genes involved in ARVC

Type of ARVC	Reference	MIM	Chromosome	Mode of inheritance	Gene	Reference
ARVC-1	(Rampazzo et al. 1994)	107970	14q23-q24	AD	Transforming Growth Factor-β3 (*TGFβ3*)	(Beffagna et al. 2005)
ARVC-2	(Rampazzo et al. 1995)	600996	1q42-q43	AD	Cardiac Ryanodine Receptor (*RyR2*)	(Tiso et al. 2001)
ARVC-3	(Severini et al. 1996)	602086	14q12-q22	AD	unknown	-
ARVC-4	(Rampazzo et al. 1997)	602087	2q32.1-q32.3	AD	unknown	-
ARVC-5	(Ahmad et al. 1998)	604400	3p23	AD	unknown	-
ARVC-6	(Li et al. 2000) (Matolweni et al.)	604401	10p12-p14	AD	unknown	-
ARVC-7	(Melberg et al. 1999)	609160	10q22.3	AD	unknown	-
ARVC-8	(Rampazzo et al. 2002)	607450	6p24	AD/AR	Desmoplakin (*DSP*)	(Rampazzo et al. 2002) (Alcalai et al. 2003)
ARVC-9	(Gerull et al. 2004)	609040	12p11	AD	Plakophilin-2 (*PKP2*)	(Gerull et al. 2004)
ARVC-10	(Pilichou et al. 2006)	nm	18q12.1	AD	Desmoglein-2 (*DSG2*)	(Pilichou et al. 2006)
Naxos	(Coonar et al. 1998)	601214	17q21	AR	Plakoglobin (*JUP*)	(McKoy et al. 2000)
ARVC/APC	(Frances et al. 1997)	115650	14q24-q terminal	AR	unknown	-

(Paul et al. 2003)

Note: The types of ARVC (i.e. ARVC 1-10) are classified based on the different loci mapped for the condition (Dokuparti et al. 2005).

APC Anterior Polar Cataract; *MIM*^{TM} Mendelian Inheritance in Man (http://www.ncbi.nlm.nih.gov/Omim/); *AD* Autosomal Dominant; *AR* Autosomal Recessive, *nm* not mentioned

Four out of the six ARVC genes, namely plakoglobin, desmoplakin, plakophilin-2 and desmoglein-2, encode cell junction proteins that are found in the desmosome. It has therefore been suggested that ARVC be renamed as a cell junction cardiomyopathy (Thiene et al. 2005).

1.7.1 Structure and Function of the Desmosome

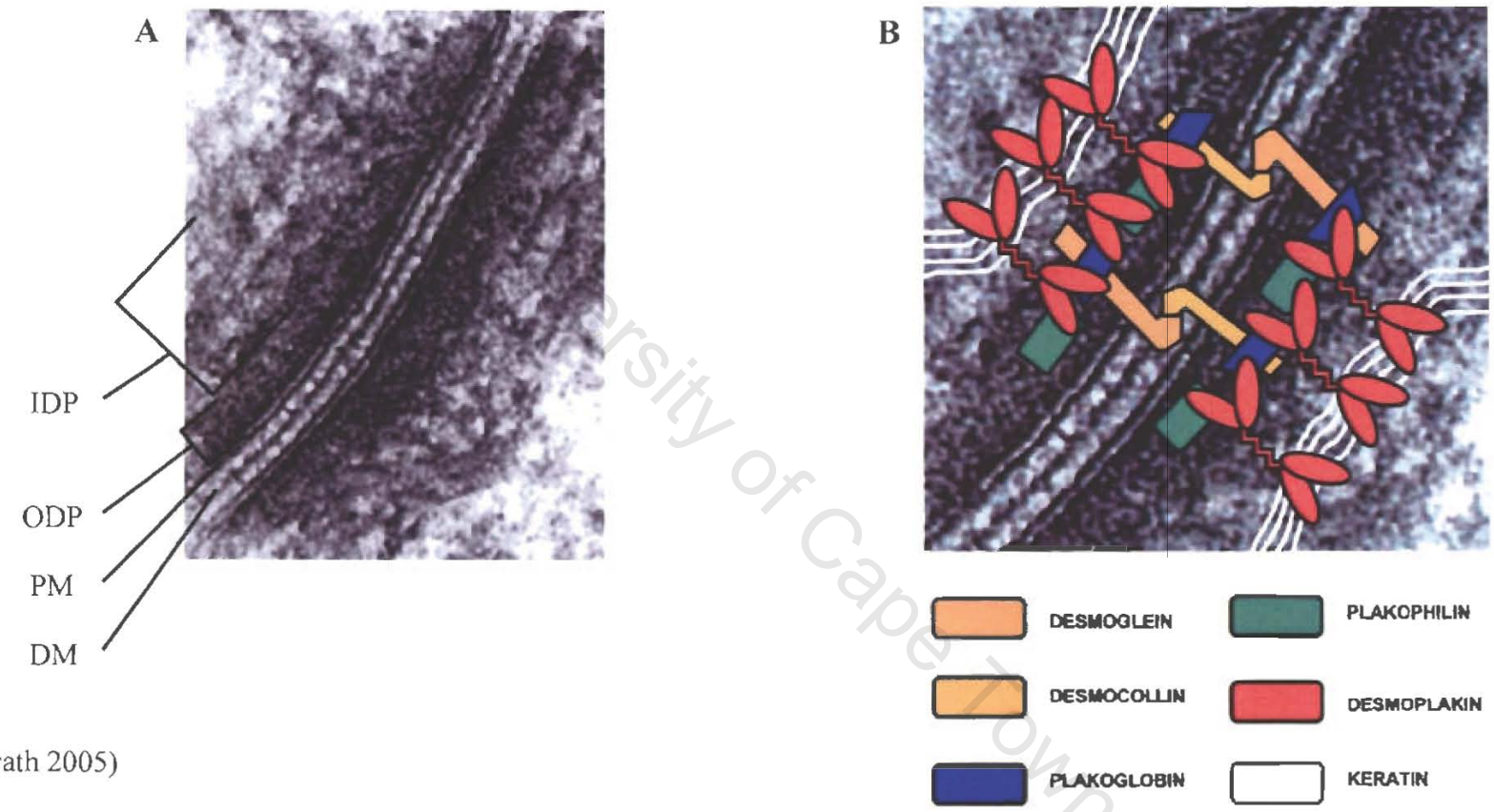
Desmosomes belong to a class of intercellular junctions called "anchoring junctions" (Yin and Green 2004). They are found primarily in epithelial tissues but they are also present in the meninges, the dendritic reticulum cells of lymph node follicles and the myocardium. Desmosomes provide mechanical integrity to these tissues by anchoring intermediate filaments (IFs) to sites of strong adhesion (McGrath 2005).

Desmosomes contain proteins from three distinct gene families, namely:

1. **cadherins:** desmogleins (Dsg1-4) and desmocollins (Dsc1-3)
2. **armadillo proteins:** plakoglobin, plakophilins (PKP1-3) and p0071 (also known as PKP4)
3. **plakins:** desmoplakins I and II, plectin, envoplakin, periplakin and epiplakin (Yin and Green 2004; McGrath 2005).

Figure 1-4 depicts the ultrastructure of the desmosome and the interaction of the major desmosomal proteins.

It has been hypothesised that mutations in plakoglobin, desmoplakin, plakophilin-2 and desmoglein-2 cause ARVC by disrupting the desmosomes which results in cell death and fibrofatty replacement (McKoy et al. 2000).



(McGrath 2005)

Figure 1-4: Desmosome ultrastructure and molecular model in human skin.

A, An electron micrograph of a desmosome in human skin, showing the regions of the junction, namely: the inner dense plaque (IDP), the outer dense plaque (ODP), the plasma membrane (PM) and the dense midline (DM). **B**, A molecular model which depicts the relative organisation of the major desmosomal proteins.

1.7.2 Disease Genes

1.7.2.1 Plakoglobin (*JUP*)

A homozygous 2 base pair deletion in plakoglobin causes Naxos disease (McKoy et al. 2000). This disease was first described in 1986, by Protonotarios and colleagues, in families originating from the Greek island of Naxos (Protonotarios et al. 1986). It is an autosomal recessive form of ARVC associated with thickening of the palms and soles (palmoplantar keratoderma) and woolly hair (Protonotarios et al. 2002). The skin and hair phenotype is fully expressed from early infancy, while the cardiac abnormalities are 100% penetrant by adolescence (Protonotarios et al. 2001).

1.7.2.2 Desmoplakin (*DSP*)

Desmoplakin has been implicated in a number of cardiac conditions to date.

The first recessive mutation to be described caused heart disease, palmoplantar keratoderma and woolly hair (Norgett et al. 2000). This cardiocutaneous syndrome is named Carvajal syndrome after Dr. Luis Carvajal-Huerta who first described it in 1998 (Carvajal-Huerta 1998). Clinical examination diagnosed the heart disease as dilated cardiomyopathy but analysis of the cardiac pathology has revealed a cardiomyopathy with unique pathologic features (Kaplan et al. 2004). According to Norman et al., the cardiac disease is typical of ARVC with predominant left ventricle involvement (Norman et al. 2005).

A heterozygous and a homozygous missense mutation, in different exons of the *DSP* gene, were reported to cause autosomal dominant and autosomal recessive ARVC, respectively (Rampazzo et al. 2002; Alcalai et al. 2003). The recessive ARVC family reported by Alcalai is characterised by a pemphigus-like skin disorder and woolly hair (Alcalai et al. 2003). In a recent study by Pilichou and colleagues, mutations in *DSP* accounted for 16% of ARVC (Pilichou et al. 2006).

Recently, a heterozygous single base pair insertion in *DSP* was identified and it is reported to cause autosomal dominant Arrhythmogenic Left Ventricular Cardiomyopathy (ALVC) (Norman et al. 2005).

1.7.2.3 Plakophilin-2 (*PKP2*)

Thirty different heterozygous mutations in *PKP2* have been reported to cause 11-27% of cases of autosomal dominant ARVC (Gerull et al. 2004; Syrris et al. 2006). These mutations cover a range of mutation types, from missense and nonsense mutations to deletions, insertions, deletion/insertions and splice site mutations.

1.7.2.4 Desmoglein-2 (*DSG2*)

Nine heterozygous mutations in *DSG2* were detected in 8 probands. Of these mutations 5 were missense and there was a single insertion, deletion, nonsense and splice site mutation. *DSG2* has been reported to account for 10% of ARVC (Pilichou et al. 2006).

1.7.2.5 Cardiac Ryanodine Receptor (*RyR2*)

RyR2 is a calcium (Ca^{2+})-release channel found on the sarcoplasmic reticulum (an intracellular $Ca2^{+}$ storage organelle). In the heart, Ca^{2+} regulates muscle contraction and electrical signals that determine the cardiac rhythm. RyR2 is required for excitation-contraction (EC) coupling and maintaining intracellular calcium homeostasis (Marks 2000; Tiso et al. 2001; Chelu et al. 2004).

Missense mutations in *RyR2* cause ARVC-2 and catecholaminergic polymorphic ventricular tachycardia (CPVT) (Priori et al. 2001; Tiso et al. 2001). CPVT is characterised by stress-induced, bidirectional ventricular tachycardia in the absence of structural heart disease and seven mutations have been identified (Laitinen et al. 2001; Priori et al. 2001).

Four mutations were identified in four families with ARVC-2. ARVC-2 differs clinically from the other types of ARVC due to the presence of effort-induced ventricular arrhythmias, high penetrance and 1:1 male:female ratio among affected patients (Tiso et al. 2001).

It has been hypothesised that mutations in *RyR2* cause ARVC-2 by altering the ability of the channel to remain closed which leads to Ca^{2+} leaking from the channel. The leaking of Ca^{2+} will in turn lead to arrhythmias and an imbalance of intra-cellular Ca^{2+} can trigger apoptosis (Chelu et al. 2004; Dokuparti et al. 2005).

1.7.2.6 Transforming Growth Factor-β3 (*TGFβ3*)

In 2005 Beffagna and colleagues reported transforming growth factor-β3 (*TGFβ3*) as the disease gene involved in ARVC-1. TGFβ3 belongs to a group of regulatory cytokines which are important in development and tissue homeostasis. They identified a nucleotide substitution in the 5' untranslated region (UTR) of an ARVC-1 family and in the 3' UTR of an unrelated ARVC patient. Neither nucleotide change was found in 300 control subjects and *in vitro* expression assays showed mutated UTRs to be about 2.5-fold more active than wild-types (Beffagna et al. 2005).

In an editorial, Nattel and Schott mention a number of issues that need to be resolved before TGFβ3 overexpression can be accepted as the molecular basis of ARVC-1. Their first concern is that two families with linkage to ARVC-1 do not contain *TGFβ3* mutations. The *in vivo* overexpression of TGFβ3 needs to be confirmed in the mutation carriers and the biological connection between TGFβ3 overexpression and ARVC needs to be established. TGFβ3 is ubiquitously expressed yet mutations in this gene result in only a cardiac phenotype (Nattel and Schott 2005).

1.8 ARVC Registries

A multidisciplinary and multi-centre study of ARVC has been established in Europe and North America. The European study involves seven countries, namely: Italy, Germany, France, United Kingdom, Greece, Poland and Cyprus (Basso et al. 2004). The North America study is a collaboration of research groups from the University of Arizona, Baylor College of Medicine and the University of Rochester. Patients suspected of having ARVC are screened in enrolling medical centres, geographically distributed in the United States and Canada (Marcus et al. 2003). The aim of these studies is to investigate the clinical, pathological and genetic features of ARVC.

A national ARVC registry was established in South Africa in 2004 under the direction of the Cardiac Arrhythmia Society of South Africa (CASSA). The co-ordinating centre is in the Cardiac Clinic at Groote Schuur Hospital in Cape Town. A total of 32 patients from 27 families met the standarised diagnostic criteria for ARVC (Latib et al. 2004).

1.9 Aims and Objectives of this Investigation

In 2004 Asano and colleagues reported a mouse model of ARVC caused by mutation of the laminin receptor 1 gene (*Lamr1*) (Asano et al. 2004). Lamr1 is a ribosomal protein localised in the nucleus and involved in apoptosis. An *in vitro* study of cardiomyocytes expressing the mutated Lamr1 showed early cell death accompanied by alteration of the chromatin architecture. It was also found that heterochromatin protein 1 (Hp1) bound to mutant Lamr1. Hp1 is localised at heterochromatin sites, where it mediates gene silencing and it was concluded that the mutant Lamr1 interacts with Hp1 to cause degeneration of cardiomyoctes (MIM: 604478). This rodent model of ARVC thus established *LAMR1* and *CBX5* as candidate genes for the human form of ARVC. (The official name and symbol for Hp1 in *homo sapiens* is chromobox protein homolog 5 (CBX5) and hereafter will be referred to as such).

The aim of this study was therefore to perform mutation screening of *LAMR1* and *CBX5* in the patients enrolled in the South African ARVC registry. There were a number of factors underlying the plausibility of *LAMR1* and/or CBX5 as ARVC disease genes. First, the heart disease found in the mice is characteristic of human ARVC. In the mice the cardiomyocyte degeneration proceeded from the outer to the inner part of the ventricular wall as in human ARVC. Second, *LAMR1* maps to a known ARVC locus, namely ARVC-5. Third, apoptosis in ARVC has been shown in the autopsy specimens of the right ventricular myocardium (Dokuparti et al. 2005).

Identification of all ARVC disease genes will aid elucidation of the pathogenic pathway leading to the condition and assist with the clinical diagnosis of ARVC (Dokuparti et al. 2005).

Chapter 2:

LAMR1 methods

2 METHODS: MUTATION SCREENING OF THE *LAMR1* GENE

2.1 Patients

A total of 22 unrelated index patients, from the ARVC Registry of South Africa, provided informed consent to participate in genetic studies (Appendix 1) and contributed 8ml of blood for DNA extraction. The mean age at diagnosis was 33 years (range 14-53 years) and 77% were male. The demographics of each patient are given in Table 2-1. The majority of patients are of Caucasian origin. The blood specimens and consent forms were sent to the molecular genetics laboratory of the Division of Human Genetics, University of Cape Town (UCT). Patient information was captured onto the Human Genetics Laboratory Database and a laboratory reference code was assigned to each specimen to ensure confidentiality. DNA was extracted from lymphocytes using the PUREGENE™ DNA Isolation Kit (*Gentra Systems*, Adcock Ingram Scientific). The DNA was stored as two stock samples, one sample at -20°C and the other sample at -70°C. DNA samples from Caucasian individuals, with no history of heart disease, were provided by the Cape Heart Centre, UCT and served as controls.

Table 2-1: Demographics of the ARVC patients

Patient	Age at diagnosis (in years)	Sex	Ethnic origin
ACM 1.2 KEV	20	Male	Caucasian
ACM 2.4 MRC	25	Male	Caucasian
ACM 3.1 CHE	39	Female	Caucasian
ACM 4.1 ROS	26	Female	Mixed Ancestry
ACM 5.1 JOH	44	Female	Caucasian
ACM 6.1 HAV	52	Male	Mixed Ancestry
ACM 8.3 MIC	14	Female	Caucasian
ACM 9.1 POU	27	Male	Caucasian
ACM 10.1 HIL	39	Male	Mixed Ancestry
ACM 11.2 JON	15	Male	Caucasian
ACM 12.1 HER	50	Male	Caucasian
ACM 13.1 JOH	18	Male	Caucasian
ACM 14.1 DER	43	Male	Caucasian
ACM 15.1 VIG	27	Male	Caucasian
ACM 16.1 HEN	35	Male	Mixed ancestry
ACM 17.1 AKH	22	Male	Black
ACM 18.1 TIA	21	Male	Caucasian
ACM 19.2 CHR	19	Female	Caucasian
ACM 20.1 ENA	53	Female	Caucasian
ACM 21.1 COL	51	Male	Caucasian
ACM 22.1 AND	38	Male	Mixed Ancestry
ACM 23.1 GAM	40	Male	Indian

2.2 Polymerase Chain Reaction (PCR)

2.2.1 Primer Design

Primer pairs were designed for the seven exons of the *LAMR1* gene. The sequence of these primers is given in Table 2-2. An annotation program **ANNOTV9** (designed by Dr George Rebello, Division of Human Genetics, UCT) was used to designate the exons and reported single nucleotide polymorphisms (SNPs) in the genome sequence. **Primer 3**, a program freely available on the internet, was used to design the primers. Primers were designed to extend at least 60bp into the intron on either side of the exon in order to identify any variants that may affect splicing. The primers were ordered from Molecular and Cell Biology, UCT. The primer pair for exon 4 had to be redesigned as primer LAMR1_4F was designed over a region that contained two SNPs. To avoid designing primers over sequence with reported SNPs the "< >" signs can be used in Primer 3 to demarcate sequence not suitable for primers. The new primer pair was named Lamr1_4bF & R and was ordered from Inqaba Biotechnical Industries [Pty] Ltd.

Table 2-2: *LAMR1* primer sequences

Exon	Name of primer	Primer sequence (5' – 3')	Amplicon size (in bp)
1	LAMR1_1F	taacttgagttccgcgctct	599
	LAMR1_1R	aggatgttagcccgctttct	
2	LAMR1_2F	gccgtccacactatttcctc	597
	LAMR1_2R	gttcatgacccaacccaagt	
3	LAMR1_3F	ctgagtgtcggaagtgtgct	459
	LAMR1_3R	tgtggatctcctgacctcgt	
4	LAMR1_4F	tgtgcctgctttacatggag	462
	LAMR1_4R	ggccagtcagtagcctcaac	
	LAMR1_4bF	tgcaggaaatctctggagaag	573
	LAMR1_4bR	cccttgtggtgttagcgaaa	
5 and 6	LAMR1_5+6F	tgctccctgttacaattgctt	677
	LAMR1_5+6R	cagatgctgtagcacactgga	
7	LAMR1_7F	cctctgtgcctattcagcaa	570
	LAMR1_7R	aggaaggaagctcagggaaa	

F = forward primer; R = reverse primer

2.2.2 PCR Optimisation

PCR was optimised for each amplicon using a PCR optimisation kit (*Roche Diagnostics South Africa*). The kit consists of 16 PCR buffers (buffer A-P) (10 x conc.) each containing 100mM Tris-HCl and 500mM KCl. The buffers differ in $MgCl_2$ concentration and pH-value as indicated in Table 2-3.

Table 2-3: PCR optimisation buffers

pH	$MgCl_2$ concentration [mM]			
	1.0	1.5	2.0	2.5
8.3	A	B	C	D
8.6	E	F	G	H
8.9	I	J	K	L
9.2	M	N	O	P

PCR was performed in 0.2ml tubes (*ABgene*, Southern Cross Biotechnology) in a total of 25µl containing 100ng of DNA, 0.8µM of each primer, 0.2mM of each of the 4 dNTPs (*Invitrogen*, Laboratory Specialist Services or *Bioline*, Celtic Molecular Diagnostics), 1x PCR optimisation buffer (*Roche*), 0.5 Unit of *Taq* polymerase (*Invitrogen*, Laboratory Specialist Services or *Promega* GoTaq^R, Whitehead Scientific), made up with SABAX H_2O (*Adcock Ingram Critical Care*). A total of 16 tubes were set up, each tube containing the same DNA sample but a different PCR optimisation buffer. PCR cycling was carried out on a GeneAmp^R PCR System 9700 (*PE Applied Biosystems)* at the cycling conditions given in Table 2-4. The melting temperature (Tm) was calculated for each primer using the formula:

$$\mathbf{Tm = 4(G + C) + 2(A + T)}$$

The annealing temperature was then calculated by subtracting 5°C from the lowest melting temperature for the primer pair. The calculated annealing temperature for each exon is given in Table 2-5.

Table 2-4: PCR cycling conditions for *LAMR1* optimisation

Step	Temperature	Time	Number of Cycles
Initial Denaturation	94°C	4 minutes	1 cycle
Denaturation	94°C	30 seconds	
Annealing	Ta*	30 seconds	30 cycles
Extension	72°C	30 seconds	
Final Extension	72°C	7 minutes	1 cycle

* Ta = annealing temperature and is given in Table 2-5.

Table 2-5: Annealing temperatures calculated for each *LAMR1* exon

Exon number	Annealing temperature
1	55°C
2	55°C
3	57°C
4	55°C
4b	55°C
5+6	55°C
7	55°C

PCR products were electrophoresed through a 1% agarose gel containing ethidium bromide (Appendix 2) to determine which buffer produced the best amplification. The PCR products were mixed with agarose loading dye (Appendix 3) and electrophoresed alongside either 0.5µg 1kb Plus DNA Ladder (*Invitrogen*, Lab Specialist Services) or 0.5µg GeneRuler^{TM} 100bp DNA Ladder Plus (*Fermentas*, Inqaba) (Appendix 4). The buffer chosen for PCR of each exon in patient DNA is given in Table 2-6.

Table 2-6: PCR optimisation buffer for PCR of each *LAMR1* exon

Exon	PCR optimisation buffer
1	I
2	H
3	B
4	B
4b	D
5+6	D
7	O

The PCR optimisation worked well for exons 2, 4, 4b and 5+6. The amplification for exons 1 and 7 was poor. The number of cycles for exon 1 was increased from 30 to 35 cycles and the annealing temperature for exon 7 was decreased from 55°C to 50°C. No product was obtained for PCR optimisation of exon 3 so the cycling conditions were changed to that of a touchdown PCR where for 10 cycles the annealing temperature starts high and then decreases by 0.5°C each cycle. The cycling conditions for each exon are given in Table 2-7.

Table 2-7: PCR cycling conditions for each *LAMR1* exon

A. Exon 1

Step	Temperature	Time	Number of Cycles
Initial Denaturation	94°C	4 minutes	1 cycle
Denaturation	94°C	30 seconds	
Annealing	55°C	30 seconds	35 cycles
Extension	72°C	30 seconds	
Final Extension	72°C	7 minutes	1 cycle

B. Exon 2, 4, 4b and 5+6

Step	Temperature	Time	Number of Cycles
Initial Denaturation	94°C	4 minutes	1 cycle
Denaturation	94°C	30 seconds	
Annealing	55°C	30 seconds	30 cycles
Extension	72°C	30 seconds	
Final Extension	72°C	7 minutes	1 cycle

C. Exon 3

Step	Temperature	Time	Number of Cycles
Initial Denaturation	95°C	3 minutes	1 cycle
Denaturation	94°C	15 seconds	
Annealing	62°C-57°C	15 seconds	10 cycles
Extension	72°C	30 seconds	
Denaturation	89°C	15 seconds	
Annealing	57°C	15 seconds	20 cycles
Extension	72°C	30 seconds	
Final Extension	72°C	5 minutes	1 cycle

D. Exon 7

Step	Temperature	Time	Number of Cycles
Initial Denaturation	94°C	4 minutes	1 cycle
Denaturation	94°C	30 seconds	
Annealing	50°C	30 seconds	30 cycles
Extension	72°C	30 seconds	
Final Extension	72°C	7 minutes	1 cycle

2.2.3 Amplification of Patient DNA

PCR of the patient DNA was performed using the same reagents at the same concentrations as detailed above. The buffer used is indicated in Table 2-6 (page 40). A negative control, where H_2O replaced DNA, was included with every PCR set up to rule out contamination. PCR cycling was carried out on a GeneAmp^R PCR System 9700 (*PE Applied Biosystems*) at the cycling conditions given in Table 2-7.

To confirm amplification, PCR products were electrophoresed through a 1% agarose gel containing ethidium bromide (Appendix 2). The PCR products were mixed with agarose loading dye (Appendix 3) and electrophoresed alongside either 0.5µg 1kb Plus DNA ladder (*Invitrogen*, Laboratory Specialist Services) or 0.5µg GeneRuler^{TM} 100bp DNA Ladder Plus (*Fermentas*, Inqaba) (Appendix 4).

If the amplification of a DNA sample was poor, the concentration of the DNA added to the PCR reaction was increased. If this did not improve amplification, then the annealing time, extension time and number of cycles were increased in the cycling conditions. These cycling conditions are given in Table 2-8.

Table 2-8: PCR cycling conditions for increased PCR product

Step	Temperature	Time	Number of Cycles
Initial Denaturation	95°C	2 minutes	1 cycle
Denaturation	95°C	30 seconds	
Annealing	Ta*	1 minute	35 cycles
Extension	72°C	1 minute	
Final Extension	72°C	5 minutes	1 cycle

* Ta = annealing temperature and is the same as indicated in Table 2-7.

Amplification of patient DNA for exon 4 yielded two PCR products of slightly different sizes. The cycling conditions were therefore changed to that of a touchdown PCR to eliminate non specificity. These cycling conditions are given in Table 2-9.

Table 2-9: Touchdown PCR cycling conditions for *LAMR1* exon 4

Step	Temperature	Time	Number of Cycles
Initial Denaturation	95°C	3 minutes	1 cycle
Denaturation	94°C	15 seconds	
Annealing	60°C-55°C	15 seconds	10 cycles
Extension	72°C	30 seconds	
Denaturation	89°C	15 seconds	
Annealing	55°C	15 seconds	20 cycles
Extension	72°C	30 seconds	
Final Extension	72°C	5 minutes	1 cycle

2.3 Sequencing

Sequencing was performed in both directions to screen for mutations in the exons. All 25µl of each of the PCR products were electrophoresed through a 1% agarose gel containing ethidium bromide (Appendix 2). The DNA fragments were excised from the agarose gel using a clean, sharp scalpel and placed in 1.5µl tubes (*Eppendorf*, Merck). This step ensures that excess primers and non-specific amplification products are removed, which may affect the sequencing quality. An empty 1.5ml tube was used to tare the scale and the DNA fragments were weighed. The DNA was extracted and purified from the agarose gel using the QIAquick Gel Extraction Kit (*QIAGEN*, Southern Cross Biotechnology) or the Wizard^{R} SV Gel and PCR Clean-Up System (*Promega*, Whitehead Scientific) in accordance with the

manufacturer's instructions. DNA was eluted with 30µl and 50µl SABAX H_2O (*Adcock Ingram Critical Care*) respectively.

Cycle sequencing was performed in 0.2ml tubes (*ABgene*, Southern Cross Biotechnology) in a total of 20µl containing ~100ng PCR product, 3.2µM forward or reverse primer (described in Table 2-2 page 37), 2µl of the BigDye^R Terminator v3.1 Cycle Sequencing mix (*Applied Biosystems*), 2µl BigDye^R Terminator v3.1 5x Sequencing Buffer (*Applied Biosystems*), made up with SABAX H_2O (*Adcock Ingram Critical Care*). The thermal cycling was carried out on a GeneAmp^R PCR System 9700 (*PE Applied Biosystems*) and the cycling conditions are given in Table 2-10.

Table 2-10: Sequencing cycling conditions

Temperature	Time	Number of Cycles
96°C	1 minute	1 cycle
95°C	45 seconds	
Ta*	4 minutes	25 cycles

* Ta (annealing temperature) for exons 1 to 5+6 was 55°C and for exon 7 was 50°C.

Excess di-deoxynucleotide terminators were removed from completed DNA sequencing reactions using Centri-sep columns (*Princeton Separations*, Whitehead Scientific), in accordance with the manufacturer's instructions. If the columns were used previously, 7% Sephadex^{TM} G-50 Fine (*Amersham Biosciences*, Separation Scientific) was prepared and 900µl was pipetted into each column. The columns were left to stand for at least 30 minutes at room temperature before following the manufacturer's protocol.

Sequencing was performed on the ABI Prism^{R} 3100 Genetic Analyzer (*Applied Biosystems*) and analysed using the Sequence Analysis version 3.7 software (*Applied Biosystems*). Due to time restrictions some samples were sequenced by Mshengu Tshabalala, in the Division of Human Genetics, UCT and Genecare, Christiaan Barnard Memorial Hospital, Cape Town.

The electropherograms were analysed by eye to identify heterozygous variants. To identify homozygous variants, the patients' sequences were aligned to the wild type sequence. BioEdit, a program downloaded from the internet, was used to align the sequences and the wild type sequence was obtained from the National Center for Biotechnology Information (NCBI) database (NT_022517.16).

2.4 Polyacrylamide Gel Electrophoresis (PAGE) Analysis

A 20bp deletion was identified in the promoter region. The background population was screened for this variant to determine if it is a novel variant associated with ARVC or a polymorphism. PCR of amplicon 1 was carried out, according to the conditions described above, and polyacrylamide gel electrophoresis (PAGE) was used to visualise the genotype of each PCR product.

Polyacrylamide gel was set between two glass plates (11.8cm x 22cm). One of the plates contained blocks of Dymo tape (*Waltons*) for well formation which is referred to as the well plate. The other plate is referred to as the back plate. Both glass plates were cleaned thoroughly with ethanol. The well plate was cleaned with acetone (to remove the ethanol) and Gel Slick^{R} Solution (*BioWhittaker Molecular Applications*, Adcock Ingram Scientific) was applied to prevent the gel from adhering to this plate. A mixture of plate glue (*SIGMA*) and 10% acetic acid was wiped vigorously over the back plate so that the gel would adhere to this plate. Excess plate glue was removed with ethanol. The plates were clamped together (cleaned sides facing inwards) with

a 0.5cm wide spacer in between. The plates were set upright with the wells at the bottom. Figure 2-1 illustrates the set up of the plates.

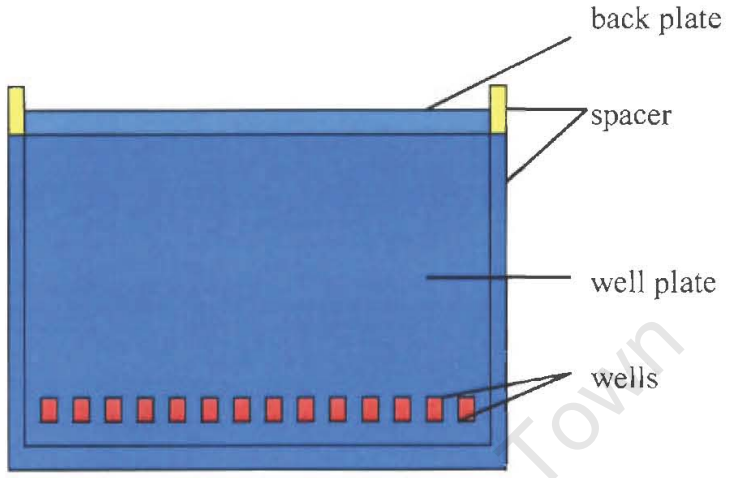


Figure 2-1: Set up of plates for polyacrylamide gel

A polyacrylamide gel was prepared by mixing 17ml 8% polyacrylamide solution (Appendix 2) with 200µl 10% ammonium persulphate (APS) and 20µl N'N'N'N'-tetramethylethylenediamine (TEMED) (*BDH Laboratory Supplies*). APS and TEMED cause polymerisation of the polyacrylamide solution, therefore this mixture was quickly poured between the plates using a 10 ml syringe. The gel was allowed to polymerise for approximately an hour before use. The clamps were removed, the plates separated and the back plate (containing the gel) was put into position on the "Multiphor II Electrophoresis System" (*Amersham Pharmacia Biotechnology*, Separation Scientific). Six filter paper strips (5cm x 28cm) were soaked in 1x TBE (Appendix 3). Three strips were placed on both the top and bottom of the gel. A volume of 3µl PCR product was mixed with 3µl agarose loading dye (Appendix 3) of which 3µl was loaded into the well. The PCR products were electrophoresed through the gel at approximately 355V for two hours alongside a 1kb Plus DNA ladder (*Invitrogen*, Lab Specialist Services). The gel was stained in silver staining

solution 1 (Appendix 3) for 10 minutes. The gel was then washed in dH_20 to remove any excess silver nitrate before being counter stained in silver staining solution 2 (Appendix 3) for 5-10 minutes.

2.5 Haplotyping

Haplotype analysis was performed by Dr Adebowale Adeyemo at the National Human Genome Center, Howard University, Washington DC, USA. Haplotype blocks were defined using two methods: the confidence interval method and the solid spine of LD method, both implemented in the *Haploview* program (Barrett et al. 2005). Haplotype construction was done using the PHASE algorithm (Stephens et al. 2001; Stephens and Donnelly 2003; Stephens and Scheet 2005). This is a Bayesian approach that uses coalescent-based methods to improve accuracy of haplotype estimates from unrelated individuals sampled from a population. *PHASE* version 2.1.1 was used for the analysis.

Chapter 3:

LAMR1 results

3 RESULTS: MUTATION SCREENING OF THE *LAMR1* GENE

3.1 Sequencing

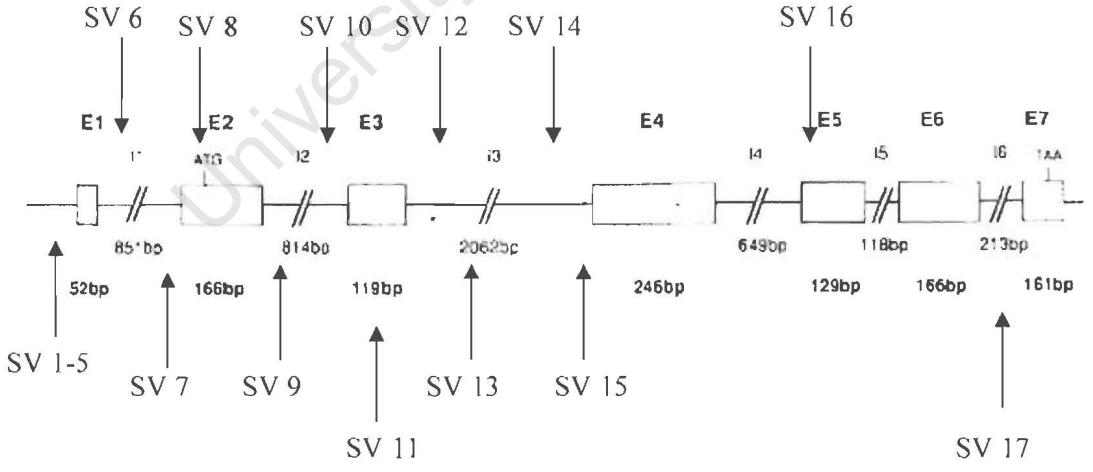
Overall, a total of 17 different sequence variants (SV) were identified in the *LAMR1* gene and they have been named SV 1-17. Fifteen of the SVs are single base substitutions. SV 4 is a 20bp deletion in the promoter region of *LAMR1*, 47bp upstream to exon 1 and SV 12 is a TA insertion in intron 3. A summary of the SVs is given in Table 3-1 and the location of each SV is indicated in Figure 3-1. The SVs have been annotated in the genomic sequence in Appendix 5 along with the nature of the change and the SNP accession number (if it is known).

Of the variants found, 10 (i.e., SV 1, 2, 5, 6, 8, 9, 10, 14, 16 and 17) have been reported as single nucleotide polymorphisms (SNPs) on the NCBI SNP database. The sequence variants found in each patient are indicated in Table 3-2 except for SV 14 and 17. SV 14 was excluded from the Table as it was only identified in one patient. The sequences of the other patients did not cover this SV. SV 17 was excluded from the Table as the genotype of this variant could not be accurately determined in six patients. The genotype for SV 13 was not determined in half the patients. When sequencing with the forward primer, patients heterozygous for the TA insert (SV 12) resulted in the rest of the sequence becoming unreadable. When sequencing with the reverse primer, SV 13 is situated too close to the reverse primer to be accurately genotyped.

Table 3-1: A summary of the sequence variants identified in the *LAMR1* gene

SV number	Nature of change	SNP accession number	Physical location
SV 1	C>G	11717439	2847
SV 2	G>A	3772147	2859
SV 3	C>A	-	2861
SV 4	20bp deletion	-	2934_2953
SV 5	C>T	3772146	3000
SV 6	T>C	3772144	3220
SV 7	G>A	-	3852
SV 8	T>C	1803893	3933
SV 9	A>G	3772141	4237
SV 10	C>T	3772138	4861
SV 11	C>T	-	4993
SV 12	TA insertion	-	5075_5076
SV 13	G>A	-	5158
SV 14	A>G	2077798	6902
SV 15	A>C	-	7005
SV 16	G>A	2269350	7956
SV 17	C>T	2723	8406

Note: SVs are numbered in base pairs relative to the genomic sequence in Appendix 5. SVs recorded in the NCBI SNP database are highlighted in red.



(Jackers et al. 1996)

Figure 3-1: Schematic diagram of the *LAMR1* gene illustrating the location of the 17 sequence variants. E = exon, I = intron

Table 3-2: The *LAMR1* sequence variants found in each patient

	SV 1	SV 2	SV 3	SV 4	SV 5	SV 6	SV 7	SV 8	SV 9	SV 10	SV 11	SV 12	SV 13	SV 15	SV 16
ACM 1.2 KEV	1 1	1 2	1 1	1 1	2 2	2 2	1 1	2 2	2 2	2 2	1 1	2 2	2 2	1 1	2 2
ACM 2.4 MRC	1 1	1 2	1 1	1 1	1 2	1 2	1 1	1 2	1 2	1 2	1 1	1 2		1 1	1 2
ACM 3.1 CHE	1 2	1 1	1 2	1 1	1 1	1 1	1 1	2 2	1 1	1 2	1 1	1 2		1 1	1 1
ACM 4.1 ROS	1 1	1 1	1 1	1 1	1 2	1 2	1 1	1 2	1 2	1 2	1 1	1 2		1 1	1 2
ACM 5.1 JOH	1 2	1 2	1 2	1 1	1 2	1 2	1 1	2 2	1 2	1 2	1 1	1 2		1 1	1 2
ACM 6.1 HAV	1 2	1 2	1 2	1 1	1 2	1 2	1 1	2 2	1 2	1 2	1 1	1 2		1 1	1 2
ACM 8.3 MIC	1 2	1 1	1 2	1 1	1 1	1 1	1 1	1 2	1 1	1 1	1 1	1 1	1 1	1 1	1 1
ACM 9.1 POU	1 1	1 2	1 1	1 2	1 2	1 2	1 1	1 2	1 2	1 2	1 1	1 2		1 1	1 2
ACM 10.1 HIL	1 1	1 1	1 1	1 1	1 2	1 2	1 1	1 2	1 2	1 2	1 1	1 2		1 1	1 2
ACM 11.2 JON	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1
ACM 12.1 HER	1 1	1 1	1 1	1 1	1 2	1 2	1 1	1 2	1 2	1 2	1 1	1 2		1 1	1 2
ACM 13.1 JOH	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1
ACM 14.1 DER	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 2	1 1	1 1	1 2	1 1	1 1	1 1	1 1
ACM 15.1 VIG	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 2	1 2	1 2	1 1	1 2		1 1	1 2
ACM 16.1 HEN	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1
ACM 17.1 AKH	1 1	1 1	1 1	1 1	1 1	1 1	1 2	1 2	1 1	1 1	1 1	1 1	1 1	1 1	1 1
ACM 18.1 TIA	1 1	1 2	1 1	1 1	1 2	1 2	1 1	1 2	1 2	1 2	1 1	1 2		1 1	1 2
ACM 19.2 CHR	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1
ACM 20.1 ENA	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1
ACM 21.1 COL	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1
ACM 22.1 AND	1 1	1 2	1 1	1 1	1 2	1 2	1 1	1 2	1 2	1 2	1 1	1 2		1 2	1 2
ACM 23.1 GAM	1 1	2 2	1 1	1 1	2 2	2 2	1 1	2 2	2 2	2 2	1 1	2 2	2 2	1 1	2 2

1 = the wild type sequence (according to NBCI accession no. NT_022517.16), 2 = the sequence variant

yellow = homozygous for the wild type allele, green = heterozygous, blue = homozygous for the sequence variant

The seven novel variants are SV 3, 4, 7, 11, 12, 13 and 15. Four of these variants, namely SV 4, 7, 11 and 15 were identified only once and each time in a different patient. SV 3 is a single base substition in the promoter region of *LAMR1*, 140 bp upstream to exon 1. SV 4 (20bp deletion) was screened in 3 family members and 65 ethnically matched population controls to determine if it is a novel variant associated with ARVC or a polymorphism. The result of this screen is given in the PAGE analysis section (page 54). SV 7, 13 and 15 are intronic sequence variants and SV 11 is an exonic sequence variant. This single base substitution (C>T) is at the third "wobble" position of the codon. The codon TCC changes to TCT but the amino acid remains the same, namely Serine. SV 12 (TA insertion) was further investigated in patient ACM 1.2 KEV with apparent recessive ARVC who was homozygous for the TA insertion. The pedigree of this patient is illustrated in Figure 3-2 and SV 12 is illustrated in Figure 3-3. The unaffected mother and affected brother were sequenced for amplicon 3 and both were heterozygous for SV 12, suggesting that this change is unlikely to be disease causing.

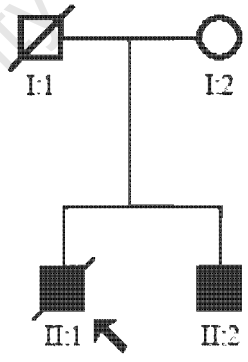


Figure 3-2: Pedigree of family 1

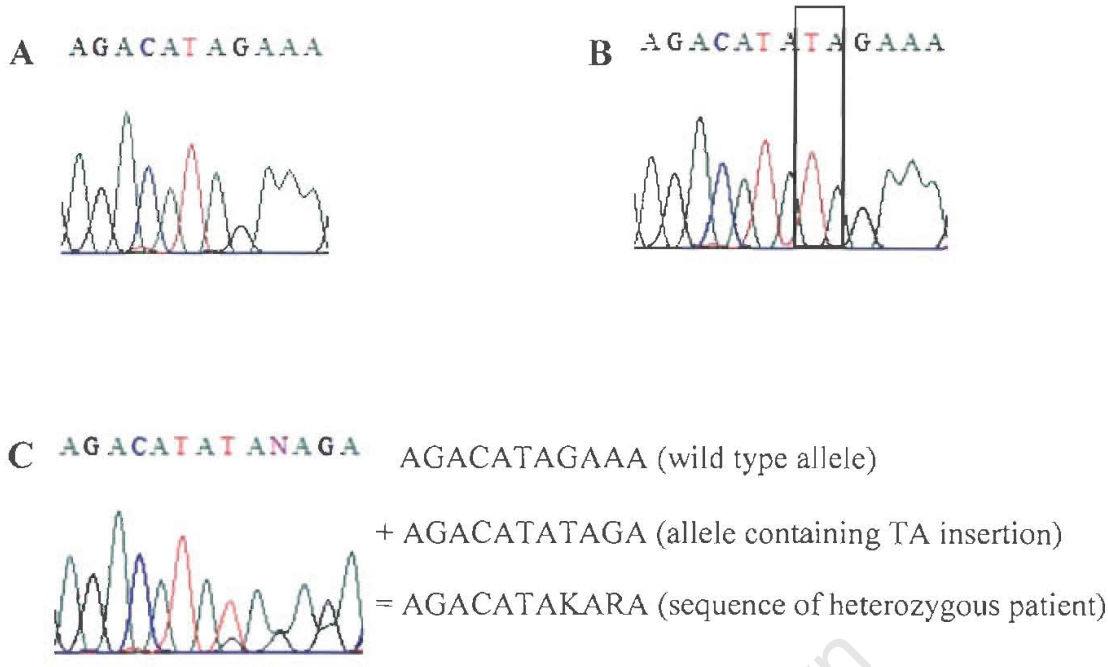


Figure 3-3: A, A patient with wild type sequence (according to NBCI accession no. (NT_022517.16). **B**, A patient homozygous for the TA insertion. **C**, A patient heterozygous for the TA insertion. The derivation of the sequence for this patient is given alongside.

To determine the effect of SV 3, 7, 11, 13 and 15 on splicing, Splice Site Prediction by Neural Network was used which is available on the Berkeley Drosophila Genome Project website (www.fruitfly.org/seq_tools/splice.html). The splice sites could not be determined for the sequence surrounding SV 3 and 15. SV 7, 11 and 13 were predicted not to interfere with splicing.

3.2 PAGE Analysis

The SV 4 (20bp deletion) found in patient ACM 9.1 POU was also detected in 2 unaffected family members and in 3/65 ethnically matched population controls (i.e. Caucasian South Africans), suggesting that it is common polymorphism with a carrier frequency of 4.6%. The three controls, with the variant profile on PAGE, were sequenced to confirm the presence of the 20bp deletion. The pedigree of patient ACM 9.1 POU is illustrated in Figure 3-4 along with the family genotyping results.

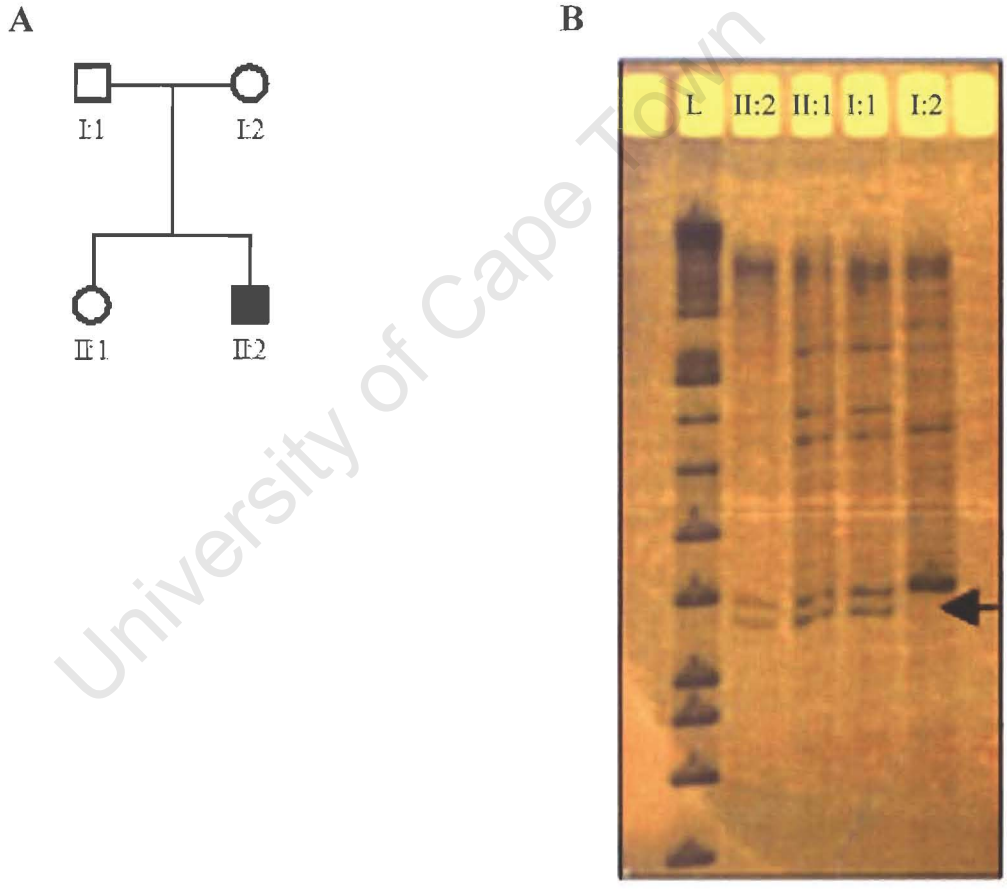


Figure 3-4: A, Pedigree of family 9. **B**, Genotype results for SV 4 in family 9. Pedigree numbers are indicated above each lane. The DNA fragments with the 20bp deletion are indicated with an arrow. L = 1 Kb+ DNA Ladder

3.3 Haplotyping

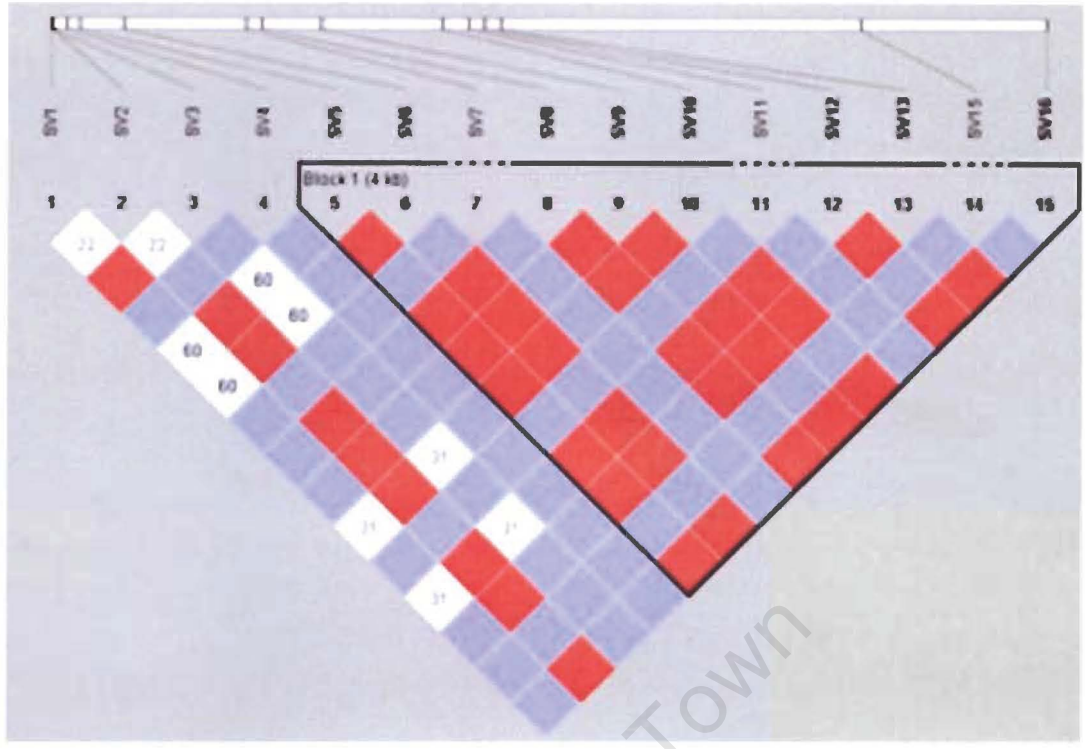
The haplotype block definitions by the two methods, namely confidence interval and solid spine of LD, are shown in Figure 3-5 on D'/LOD plots. There is clearly a block running from SV 5 to SV 16. Another block may also be formed by SV 1-SV 4, depending on the block definition used.

The haplotypes with estimated population frequency of at least 0.01 (or 1%) are shown in Table 3-3 below. The two haplotypes in bold font below account for nearly half of all chromosomes. The haplotypes are however from a group of patients with ARVC and may not reflect the general population haplotype frequency.

Table 3-3: *LAMR1* haplotypes with an estimated population frequency of at least 1% in ARVC patients

Haplotype	Freq	SE
111122122212212	0.072582	0.012482
111111112211111	0.014795	0.010833
111111111211211	0.014037	0.011045
111111111211111	0.016622	0.016681
111111111222121	0.034996	0.015350
111111111111211	0.010362	0.013977
111111111111111	**0.386538**	**0.028110**
111111221222121	0.037263	0.023463
111211111111111	0.031793	0.018631
121122122212212	**0.101894**	**0.016953**
121111111111111	0.073574	0.010807
212111121111111	0.080543	0.013341

A



B

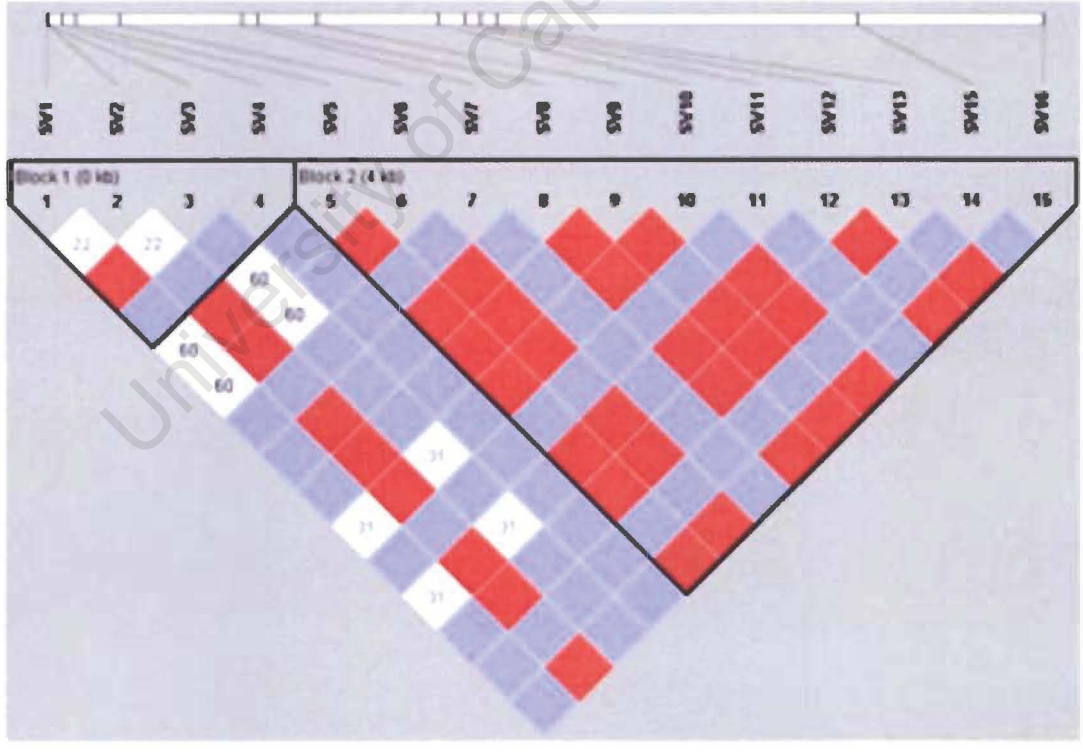


Figure 3-5: Haploblock type definition by the **A**, confidence interval method and **B**, solid spine of LD method. Strong linkage disequilibrium (LD) between markers is indicated by a red block, inconclusive LD by white block, and strong recombination by blue.

Chapter 4:

CBX5 methods

4 METHODS: MUTATION SCREENING OF THE *CBX5* GENE

4.1 Patients

The patient cohort described in "Methods: Mutation Screening of the *LAMR1* Gene" was screened for mutations in the *CBX5* gene. Controls were used in the dHPLC screen. They were of Caucasian and black African descent and were archived in the Division of Human Genetics, UCT.

4.2 PCR

4.2.1 Primer Design

Primer pairs were designed for exons 2 to 5 of *CBX5*. The sequence of these primers is given in Table 4-1. Exon 1 was excluded from this mutation screen as it is not translated. The disease model, described by Asano and colleagues, was due to Hp1 binding specifically to mutant Lamr1 (Asano *et al.*, 2004). The hypothesis was therefore to screen exons encoding CBX5 protein for mutations that might lead to a greater affinity of CBX5 for LAMR1, and so lead to ARVC. ANNOTV9 and Primer 3 were used to design primers. OligoAnalyzer 3.0 and Basic Local Alignment Search Tool (BLAST) were used to evaluate the primers.

OligoAnalyzer 3.0, a program freely available on the internet (http://scitools.idtdna.com/scitools/Applications/OligoAnalyzer/Default.aspx), was used to identify possible:

1. **Hairpin structures:** annealing of the primer with itself
2. **Homo-dimers**: annealing of primers of the same type
3. **Hetero-dimers**: annealing within primer pairs

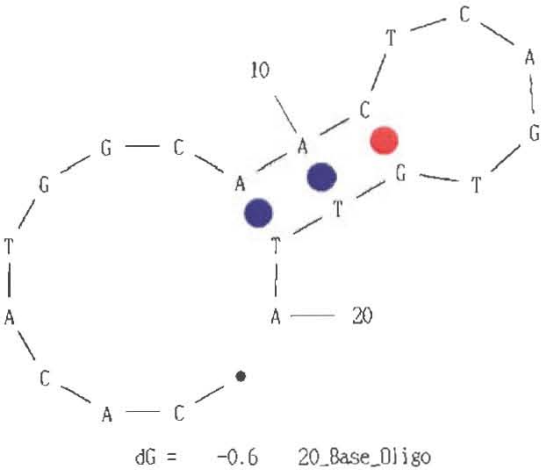
OligoAnalyzer 3.0 produces a graphical representation and a delta g value for each possible secondary structure of a primer. An example of each secondary structure is given in Figure 4-1. If the delta g value was -9kcal/mol or more negative, the primer was redesigned (OligoAnalyzer 3.0 technical support, personal communication).

Table 4-1: *CBX5* primer sequences

Exon	Name of primer	Primer sequence	Amplicon size (in bp)
2	CBX5_2F	cacatggcaactcagtgtta	402
	CBX5_2R	ctcttaggaagagtgcaggag	
3	CBX5_3F	gtccgaatgtcctctttgaa	449
	CBX5_3R	ggagagggagctgtaatcaa	
4	CBX5_4F	ttagggttcaccaccagtct	450
	CBX5_4R	ggtttgggaatatgcagtct	
5	CBX5_5F	aggtgaggaggaaatcaagtc	484
	CBX5_5R	ctcagtatgtaaatgcctggtc	

BLAST, which is available on the NCBI website, was used to search for the primer in the human genome. If the 3' end of the primer was found anywhere else in the genome, the primer was redesigned. The primers were ordered from Inqaba Biotechnical Industries [Pty] Ltd.

A



B

HOMO-DIMER ANALYSIS

Dimer Sequence
5'- CACATGGCAACTCAGTGTTA -3'

Maximum Delta G -34.4 kcal/mole

Delta G -5.38 kcal/mole
Base Pairs 4

```
5'              CACATGGCAACTCAGTGTTA
                    ||||
3' ATTGTGACTCAACGGTACAC
```

C

Primary Sequence
5'- CACATGGCAACTCAGTGTTA -3'

Secondary Sequence
5'- CTCTTAGGAAGAGTGCAGGAG -3'

Maximum Delta G -36.87 kcal/mole

Delta G -5.09 kcal/mole
Base Pairs 3

```
5' CACATGGCAACTCAGTGTTA
        |||
3'  GAGGACGTGAGAAGGATTCTC
```


Figure 4-1: Secondary structure prediction by OligoAnalyzer 3.0. **A**, Hairpin structure **B**, Homo-Dimer **C**, Hetero-Dimer.

4.2.2 PCR Optimisation

PCR was optimised for each amplicon using the gradient function on the Px2 Thermal Cycler (*Thermo Electron Corporation*). PCR was performed in 0.2ml tubes (*ABgene*, Southern Cross Biotechnology) in a total of 25µl containing 100ng of DNA, 0.4µM of each primer, 0.2mM of each of the 4 dNTPs (*Bioline*, Celtic Molecular Diagnostics), 1x PCR Buffer (*Promega*, Whitehead Scientific; pH 8.5., 1.5mM $MgCl_2$), 0.5 unit of *Taq* polymerase (*Promega* GoTaq^R, Whitehead Scientific), made up with SABAX H_2O (*Adcock Ingram Critical Care*). A total of 12 tubes were set up each with the same DNA sample but a different annealing temperature. The PCR cycling conditions are given in Table 4-2 and the annealing temperatures are in Table 4-3.

Table 4-2: PCR cycling conditions for *CBX5* optimisation

Step	Temperature	Time	Number of Cycles
Initial Denaturation	95°C	5 minutes	1 cycle
Denaturation	95°C	30 seconds	
Annealing	50-60°C*	30 seconds	25 cycles
Extension	72°C	30 seconds	
Final Extension	72°C	7 minutes	1 cycle

* Each of the 12 tubes had a different annealing temperature ranging from 50-60°C. The exact annealing temperature for each tube is given in Table 4-3.

Table 4-3: Annealing temperatures for *CBX5* optimisation

Tube number	Annealing temperature (in ^{o}C)
1	50.0
2	50.3
3	50.9
4	51.8
5	52.9
6	54.2
7	55.5
8	56.9
9	58.4
10	59.3
11	59.8
12	60.1

The PCR products were electrophoresed through a 1.5% agarose gel (Appendix 2) to determine which annealing temperature produced the best amplification. A volume of 5µl of PCR product was mixed with 2.5µl agarose loading dye containing SYBR^{R} Gold nucleic acid gel stain (*Molecular Probes*) (Appendix 3) and electrophoresed alongside 0.5µg GeneRuler^{TM} 100bp DNA Ladder Plus (*Fermentas,* Inqaba) (Appendix 4). The annealing temperature chosen for each exon is given in Table 4-4. If multiple bands looked equally bright, the annealing temperature was calculated (as described in "Methods: Mutation Screening of the *LAMR1* Gene") to aid decision-making.

Table 4-4: Annealing temperature for PCR of each *CBX5* exon

Exon number	Annealing temperature (in ^{o}C)
2	53
3	54
4	53
5	52

Note: The annealing temperature that produced the best amplification was rounded off to the nearest whole number.

4.2.3 Amplification of Patient DNA

PCR of the patient DNA was performed using the same reagents at the same concentrations as detailed above. A negative control, where H_2O replaced DNA, was included with every PCR set up to rule out contamination. PCR cycling was carried out on a Px2 Thermal Cycler (*Thermo Electron Corporation*) at the cycling conditions given in Table 4-2 and the annealing temperature in Table 4-4. The cycling number was however changed from 25 cycles to 30 cycles to increase the amount of PCR product.

To confirm amplification, the PCR products were electrophoresed through a 1.5% agarose gel (Appendix 2). A volume of 5µl of PCR product was mixed with 2.5µl agarose loading dye containing SYBR^{R} Gold nucleic acid gel stain (*Molecular Probes*) (Appendix 3) and electrophoresed alongside 0.5µg GeneRuler^{TM} 100bp DNA Ladder Plus (*Fermentas*, Inqaba) (Appendix 4).

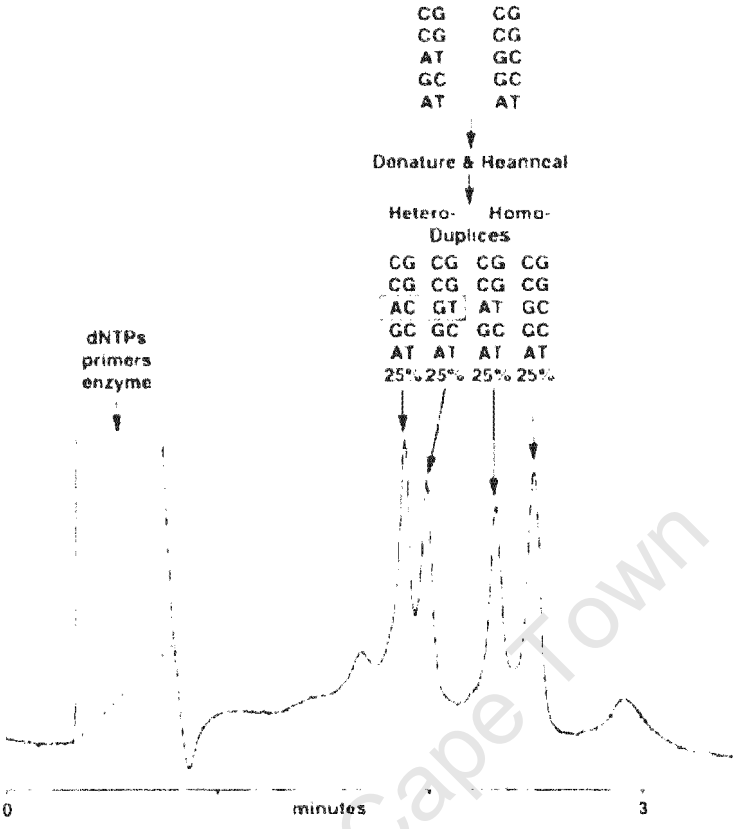
4.3 Denaturing High Performance Liquid Chromatography (DHPLC)

DHPLC, performed by the WAVE^{R} System (*Transgenomic*), was used to screen for mutations in the exons. This method was initially developed in 1995 by Oefner and Underhill (Oefner and Underhill 1995).

4.3.1 Principle of DHPLC

DHPLC, under partial denaturing conditions, can detect single base substitutions, small insertions and deletions (Xiao and Oefner 2001). Prior to DHPLC analysis, PCR products are denatured at 95°C for 5 minutes and re-annealed over 45 minutes by gradual cooling to 25°C (WAVE^{R} manual). If one of the chromosomes contains a sequence variant, the PCR products will re-anneal to form the original homoduplexes and the sense and anti-sense stands from different PCR products will form heteroduplexes (Xiao and Oefner 2001).

The PCR products are loaded on the WAVE machine which injects the sample onto the separation cartridge. The cartridge is filled with alkylated nonporous polystyrene-divinylbenzene (PS-DVB) beads which are electrostatically neutral and hydrophobic. To aid adsorption of nucleic acids to the beads, triethylammonium acetate (TEAA) is added. This reagent is positively charged and hydrophobic therefore it interacts with nucleic acids (which are negatively charged) and the hydrophobic beads. Increasing concentrations of acetronitrile flows across the cartridge. The hydrophobic interaction between the cartridge and DNA/TEAA is broken and the PCR product elutes (WAVE^{R} manual). Under partial denaturing conditions, heteroduplexes denature more extensively at the analysis temperature and therefore elute earlier than the homoduplexes that undergo less pronounced denaturation. The ideal chromatographic profile is illustrated in Figure 4-2. The order of elution of the four species is primarily determined by neighbouring stacking interactions (Xiao and Oefner 2001).



(Xiao and Oefner 2001)

Figure 4-2: An ideal chromatographic profile for a mismatch.

4.3.2 Design of WAVE methods

A WAVE method consists of an analysis temperature and DHPLC conditions. WAVEMAKER™ Software Version 4.1 was used to identify the melting domains in an amplicon and a WAVE method was designed for each domain. The analysis temperature for a domain was chosen where the helical fraction (i.e., double strandedness) of the domain is between 50 - 90%. The DHPLC conditions "rapid DNA for mutation detection" was selected where the acetronitrile flow rate is 1.5ml/min and the slope of the acetronitrile gradient is 5.0. These initial conditions were chosen as it runs samples in the shortest time possible. WAVEMAKER™ Software predicts when the sample will elute based on sequence composition and the WAVE method. In order to elute the sample between the injection peak and wash

peak the gradient duration was increased or a time shift was added. The WAVE methods were tested on two control samples. A volume of 7µl of PCR product was injected onto the cartridge for each WAVE method. Suitable chromatographic profiles were obtained for exons 4 and 5. The WAVE methods were redesigned for exons 2 and 3. If the elution peak was of insufficient magnitude the injection volume was increased. If the sample eluted to close the injection peak or the wash peak a time shift was added and either the gradient duration was increased or the sample was run on "slow DNA for mutation detection". The DHPLC conditions "slow DNA for mutation detection" has an acetronitrile flow rate is 0.9ml/min and the slope of the acetronitrile gradient is 2.0. In exon 2, two out of the three WAVE methods were redesigned but this did not improve the chromatographic profiles. This exon was therefore screened by sequencing.

4.3.3 Mutation Screening of Patients

In order to identify recessive mutations, patient PCR product was mixed with control PCR product in an equimolar ratio. Heteroduplex formation was carried out on a Hybaid TouchDown Thermal Cycler. The injection volume of each mixed PCR product was 7µl for each WAVE method except for two methods for exon 3 which had 14µl injected.

4.4 Sequencing

Exon 2 samples and those samples that showed aberrant chromatographic profiles were sequenced as in "Methods: Mutation Screening of *LAMR1* Gene" with a few deviations.

The purified PCR products were quantified using the UVIpro Gold Gel Documentation System (*UVItec*, Whitehead Scientific). This was performed to accurately determine the volume of PCR product to add to the cycle sequencing reaction. A volume of 4µl of PCR product was mixed with 10µl agarose loading dye and 1µl agarose loading dye containing SYBR^{R} Gold nucleic acid gel stain

(*Molecular Probes*) (Appendix 3). The PCR products were electrophoresed through a 1% agarose gel alongside 1µl (0.5µg) GeneRuler^{TM} 100bp DNA Ladder Plus (*Fermentas,* Inqaba) (Appendix 4) mixed with 7µl SABAX H_2O (*Adcock Ingram*) and 2µl agarose loading dye containing Syber Gold. The Gel Documentation System quantified the purified PCR product by comparing it to the known concentration of the DNA fragments of the ladder.

Cycle sequencing was performed with 4µl of the BigDye^{R} Terminator v3.1 Cycle Sequencing mix (*Applied Biosystems*) and the thermal cycling was carried out on a Hybaid TouchDown Thermal Cycler. Due to time restrictions some samples were sequenced by Genecare, Christiaan Barnard Memorial Hospital, Cape Town. The electropherograms were analysed, as mentioned previously, and the wild type sequence was obtained from the NCBI database (NC_000012.9).

Chapter 5:

CBX5 results

5 RESULTS: MUTATION SCREENING OF THE *CBX5* GENE

5.1 DHPLC

5.1.1 Exon 3

Exon 3 was analysed with three WAVE methods. The chromatographic profile observed at each WAVE method fell into one of three groups namely:

1. wild type chromatographic profiles, i.e., profiles that resembled the control sample (Figure 5-1A)
2. aberrant chromatographic profiles indicative of a sequence variant (Figure 5-1B)
3. spurious chromatographic profiles that could not be definitively classified as wild type or aberrant (Figure 5-1C)

One patient from each group was chosen for sequencing. The results of the screen are summarised in Table 5-1 and an example of each profile is given in Figure 5-1.

Table 5-1: Summary of the *CBX5* exon 3 DHPLC screen

Patients	Methods (in ^{o}C)		
	55.7	**54.7**	**52.7**
ACM 1.2 KEV*	2	2	2
ACM 2.4 MRC	2	2	2
ACM 3.1 CHE	1	1	1
ACM 4.1 ROS	2	2	2
ACM 5.1 JOH*	1	1	1
ACM 6.1 HAV	1	1	1
ACM 8.3 MIC*	3	3	3
ACM 9.1 POU	3	3	3
ACM 10.1 HIL	1	1	1
ACM 11.2 JON	1	1	1
ACM 12.1 HER	3	3	3
ACM 13.1 JOH	2	2	2
ACM 14.1 DER	2	2	2
ACM 15.1 VIG	2	2	2
ACM 16.1 HEN	2	2	2
ACM 17.1 AKH	2	2	2
ACM 18.1 TIA	2	2	2
ACM 19.2 CHR	2	2	2
ACM 20.1 ENA	2	2	2
ACM 21.1 COL	2	2	2
ACM 22.1 AND	3	3	3
ACM 23.1 GAM	2	2	2

Note: 1 = wild type profile, 2 = aberrant profile, 3 = spurious profile
Patients with the same chromatographic profile at all three WAVE methods are highlighted in the same colour. Patients chosen for sequencing are indicated with a *

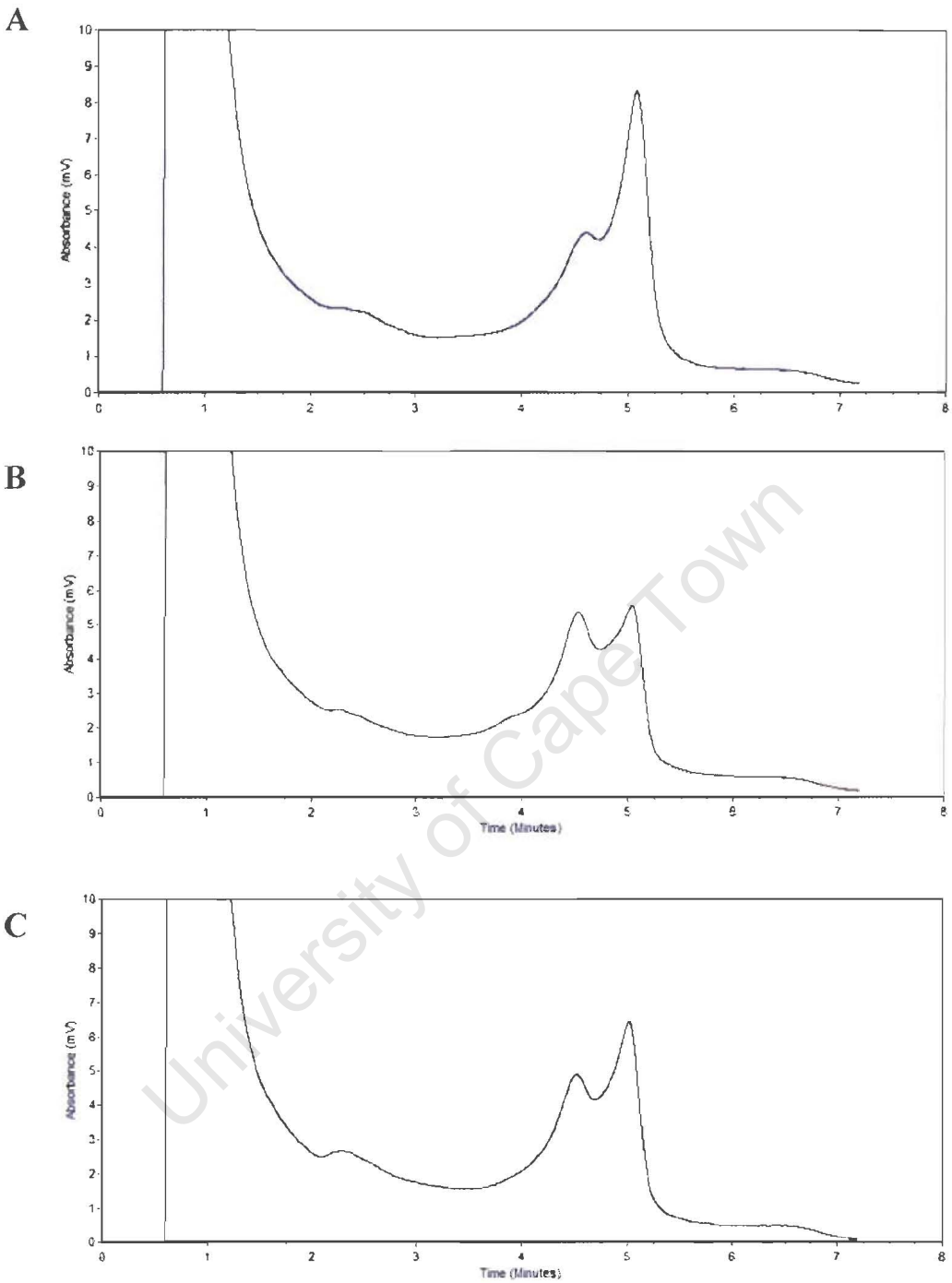


Figure 5-1: Chromatograms of exon 3 @ 55.7 °C indicating **A**, wild type profile **B**, aberrant profile and **C**, spurious profile.

5.1.2 Exon 4

Minor differences in chromatographic profile were observed in two out of four WAVE methods in patient ACM 3.1 CHE. These chromatograms are shown in Figure 5-2.

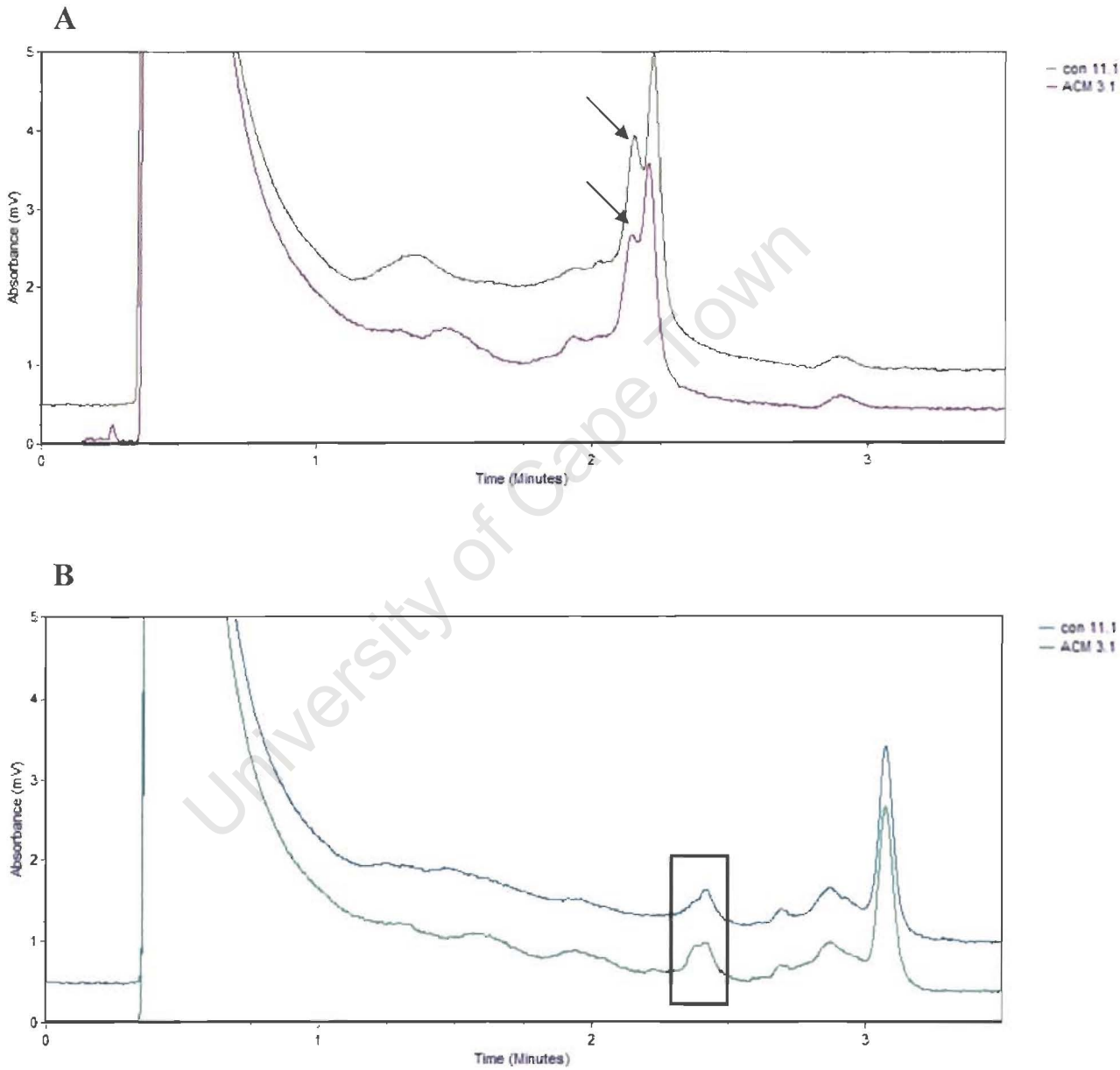


Figure 5-2: Chromatograms of exon 4, patient ACM 3.1 CHE and control (con 11.1) at **A**, 57.5°C and **B**, 55.2°C. Minor differences between control and patient profile are indicated.

5.1.3 Exon 5

Differences in chromatographic profile were observed at all three WAVE methods in patient ACM 19.2 CHR. These chromatograms are given in Figure 5-3. DNA for this patient's affected sibling (ACM 19.1 TRE) was archived in the Division of Human Genetics, UCT and was subsequently screened for this exon. The sibling did not have the same chromatographic profile as the patient and the polymorphism was therefore concluded not to be disease causing. The patient and sibling chromatograms are given in Figure 5-4.

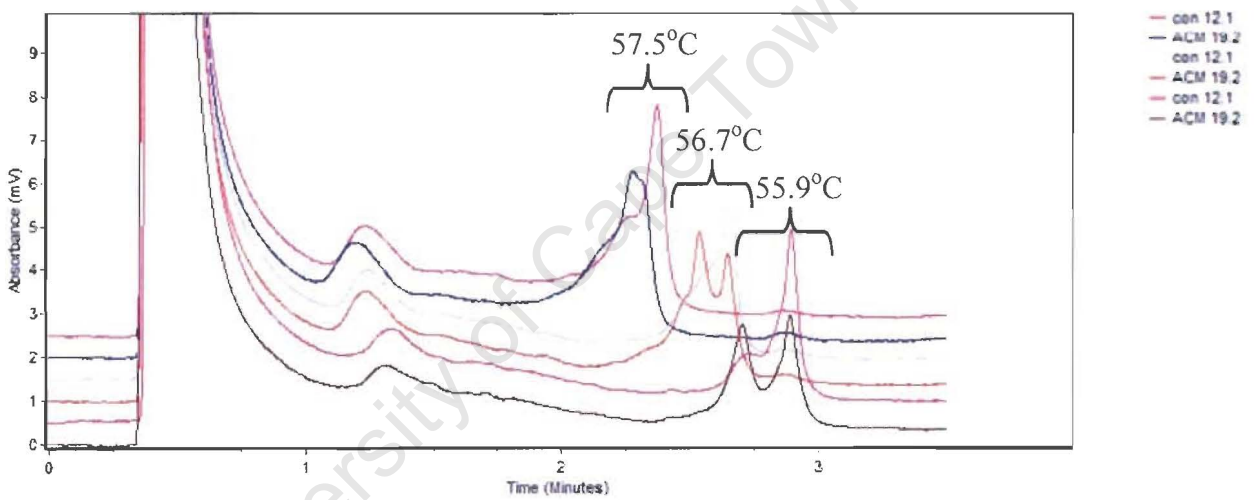


Figure 5-3: Chromatograms of exon 5, patient ACM 19.2 CHR and control (con 12.1) at all three WAVE methods

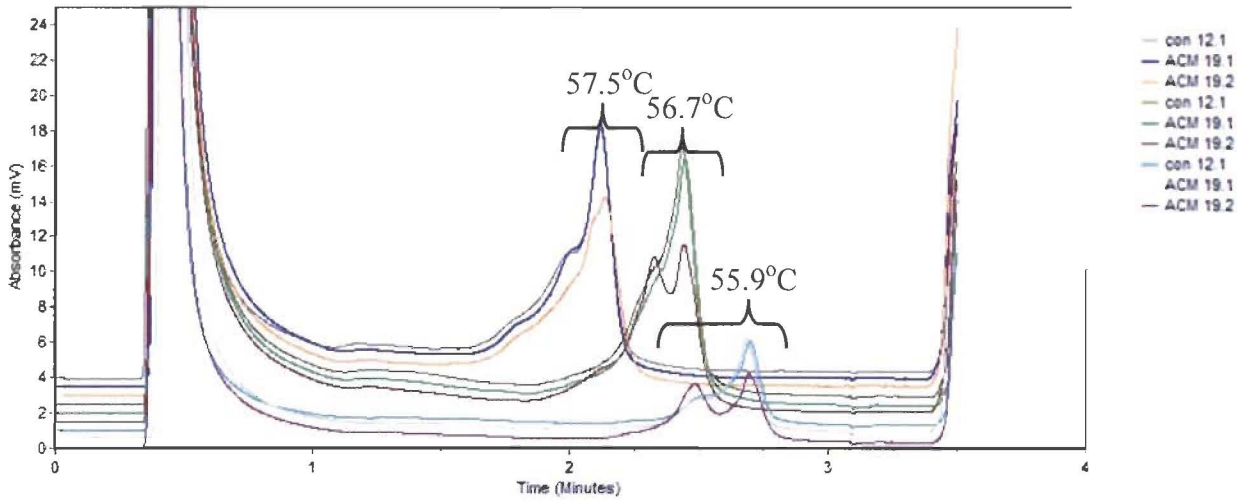


Figure 5-4: Chromatograms of exon 5, patient ACM 19.2 CHR, sibling ACM 19.1 TRE and control (con 12.1) at all three WAVE methods

5.2 DNA Sequencing

5.2.1 Exon 2

One sequence variant, a T to C substitution, was identified in exon 2. It is reported as a SNP on the NCBI SNP database (accession no. 1140681). Seven patients were homozygous wild type (according to NC_000012.9) (i.e. T/T), ten patients were heterozygous for the variant (i.e. T/C or Y) and five patients were homozygous for the variant (i.e. C/C). The status of each patient for the SNP is shown in Figure 5-5. Sequencing with the forward primer did not produce good quality sequence.

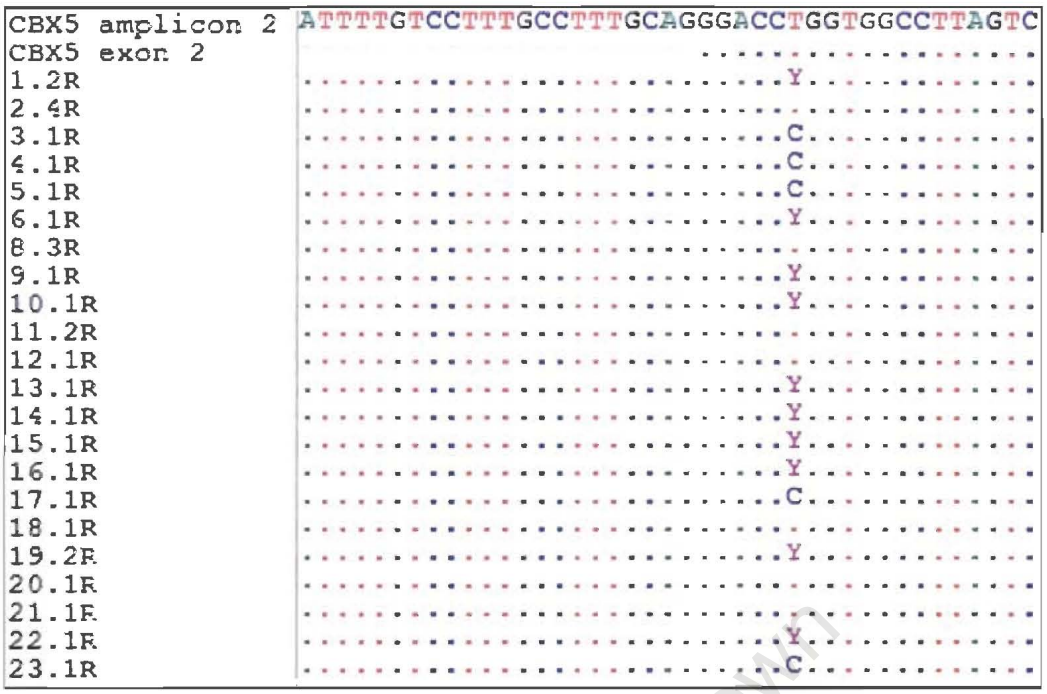


Figure 5-5: Part of the alignment between the reverse sequences and the wild type sequence which shows the status of each patient for the SNP.

5.2.2 Exon 3

The sequence variants were not identified in exon 3 as sequencing with either the forward or the reverse primer produced sequence of poor quality.

5.2.3 Exon 4

No sequence variant was identified in patient ACM 3.1 CHE upon sequencing. The minor differences in the chromatographic profiles were therefore not indicative of a variant.

5.2.4 Exon 5

The sequence variant in ACM 19.2 CHR was not identified as sequencing with either the forward or the reverse primer produced sequence of poor quality.

Chapter 6:

Discussion

6 DISCUSSION

6.1 Main Findings

The aim of this study was to screen *LAMR1* and *CBX5* genes for mutations that cause ARVC in a cohort of South African patients. *Lamr1* was reported to cause ARVC in a mouse model and was a biologically plausible candidate gene. *CBX5* was screened as the mouse homolog, Hp1, was reported to bind to the mutant Lamr1.

The candidate genes identified by the mouse model were considered important as phenotypic homologies with other organisms have been useful in identifying human disease genes. A successful example of this approach is when phenotypic resemblances led to the identification of *MSH2* and *MLH1* (human homologs of yeast mismatch repair genes) as disease genes in hereditary colon cancer (Strachan and Read 1999).

A definitive method of mutation detection was used for the *LAMR1* gene because of the relatively small number of patients (i.e., 12) that was available at the beginning of the study, and the urgency to replicate the findings of Asano et al. in man. Furthermore, it was felt that a re-sequencing study of the *LAMR1* gene would provide useful information about the structure of variation at this locus. When disease causing mutations were not found in the *LAMR1* gene, mutation screening of the *CBX5* gene was conducted by DHPLC which is cost-effective, rapid and nearly 100% of sequence variations can be detected (Oefner and Underhill 1999; Xiao and Oefner 2001).

6.1.1 LAMR1

A total of 17 sequence variants were identified in *LAMR1* and they were named SV 1-17. Of these variants 10 (i.e., SV 1, 2, 5, 6, 8, 9, 10, 14, 16 and 17) were reported as SNPs on the NCBI SNP database. None of the seven novel variants (i.e., SV 3, 4, 7, 11, 12, 13 and 15) were predicted to have a pathogenic effect on protein structure or function, although this remains to be confirmed in functional studies for some of the variants.

SV 3 is a single base substitution in the promoter region of *LAMR1*, 140bp upstream to exon 1. The possibility that this variant, which occurred in 4 patients, affects transcription has not been excluded in this study. Four variants, namely SV 4, 7, 11 and 15 were identified only once and each time in a different patient. SV 4 was found to be a common polymorphism with a carrier frequency of 4.6% in the background population. SV 7, 13 and 15 are intronic sequence variants and SV 11 is an exonic variant that does not change the amino acid. SV 12 was homozygous in one patient with apparent recessive ARVC but heterozygous in unaffected and affected family members. SV 12 was therefore thought not to be disease causing.

To determine the effect of SV 3, 7, 11, 13 and 15 on splicing, Splice Site Prediction by Neural Network was used which is available on the Berkeley Drosophila Genome Project website (www.fruitfly.org/seq_tools/splice.html). The splice sites could not be determined for the sequence surrounding SV 3 and 15. SV 7, 11 and 13 were predicted not to interfere with splicing. .

For SV 4 (20bp deletion) and SV 12 (2bp insertion) the exact bases deleted or inserted could not be determined. According to the Human Genome Variation Society (HGVS) "the most 3' position possible is arbitrarily assigned to have been changed". The 20bp deletion, [2921]tctccggctc cg**ccettctt ccgctcgact ttc**tttgcca to tctccggctc cgctttgcca, is therefore indicated in Appendix 5 as **2934_2953del20** and not as 2933_2952del20. The 2bp insertion, [5071]acatagaaag to acat**a**tagaaag, is indicated in Appendix 5 as **5075_5076insTA** and not as 5074_5075insAT.

There are two possible explanations why disease causing mutations were not found in *LAMR1*. First, this gene, like *TGFβ3,* may be a rare cause of ARVC and disease causing mutations were therefore not identified due to the small number of patients screened. Pilichou and colleagues reported that in a cohort of 80 unrelated ARVC probands, two probands carried a pathogenic or causative *TGFβ3* mutation (Pilichou et al. 2006). Second, the molecular mechanism by which *LAMR1* causes disease in the mouse may not be applicable to man.

Structure of variation in *LAMR1*

Single base substitutions were the most frequent form of sequence variation with a frequency of 88% (15 of 17 variants) compared to 12% (2 of 17 variants) observed for insertion/deletion variants. This frequency of sequence variation is consistent with other studies of human DNA sequence variation (Kwok et al. 1996; Nickerson et al. 1998) .

The average distance between adjacent SNPs in the human genome has been reported as 1871bp (Hinds et al. 2005). The single base substitutions in the *LAMR1* gene are however at much closer intervals with the first five substitutions found clustered in amplicon 1 (599 bp). There were five reported SNPs, two SNPs in amplicon 1 and three SNPs in amplicon 7 (570 bp), that were covered by sequencing but were not identified in any of the ARVC patients.

Haplotypes are a combination of alleles at different markers along the same chromosome that tend to be inherited as a unit. Haplotypes are commonly used to localise a disease-conferring gene or locus by genetic association or linkage studies. SNPs are the choice of DNA-based markers for genetic case-control association studies due to their abundance in the human genome and their potential to be genotyped in a large-scale, automated fashion (Crawford and Nickerson 2005). In candidate gene association studies, haplotypes of linked SNPs make a locus more polymorphic and thus more informative. This is illustrated in a haplotype analysis of the aldosterone synthase gene and heart size (Mayosi et al. 2003).

Haplotype analysis of the *LAMR1* gene reveals that this gene contains either one or two haplotype blocks depending on the block definition used. Haplotype blocks are regions of low haplotype diversity due to high linkage disequilibrium. The boundaries of the blocks could have been generated by hotspots of recombination or by other forces such as population history (i.e., founder effect, drift and selection) (Crawford and Nickerson 2005). Two haplotypes have been found to account for nearly half of all chromosomes. The haplotypes are however from a group of patients with ARVC and may not reflect the general population haplotype frequency.

6.1.2 CBX5

In *CBX5* a sequence variant was identified in exon 2 that has been reported as an SNP on the NCBI SNP database (accession no. 1140681). In exon 3 the chromatographic profiles of the patients fell into one of three groups, i.e., wild type, aberrant or spurious. The cause of the aberrant and spurious profiles was not identified as sequencing with either the forward or reverse primer did not produce sequence of good quality. It is unlikely that this sequence variant(s) is disease causing as it appears to be a common polymorphism with 13 patients with an aberrant profile and 4 patients with a spurious profile. No sequence variant was identified in exon 4 upon sequencing. Aberrant chromatographic profiles were observed for one patient in exon 5 at all three WAVE methods. The sequence variant was not identified because neither the forward nor the reverse primer produced readable sequence. It is unlikely that this sequence variant is disease causing as it was not found in this patient's affected sibling. Preliminary results for the *CBX5* mutation screen indicate that this gene may not be a disease gene in South African patients.

6.2 Troubleshooting

6.2.1 Primer Design

To avoid designing primers over sequence with reported SNPs (as was the case for primer LAMR1_4F) the "< >" signs were used in Primer 3 to demarcate *CBX5* sequence not suitable for primers.

OligoAnalyzer 3.0 and BLAST were used to evaluate the primers designed for *CBX5*. These two programs greatly improved primer design as is evident in the comparison between *CBX5* and *LAMR1* primers. The *CBX5* PCRs were optimised in one step and no difficulty was experienced in amplifying patient DNA samples. Some of the *LAMR1* PCRs were optimised in one step but other PCRs required many changes to the PCR cycling conditions before they were considered optimised. When amplifying the patient DNA samples not all samples would amplify in the first PCR. The DNA concentration was increased and cycling conditions with increased annealing time, extension time and number of cycles was employed to amplify problematic samples.

6.2.2 DHPLC

Exon 2 was not screened by DHPLC as two of the three WAVE methods, namely the WAVE method at 58.5°C and 56.5°C, did not produce suitable chromatographic profiles for analysis. The chromatograms of exon 2 are shown in Figure 6-1. This exon was therefore screened by sequencing.

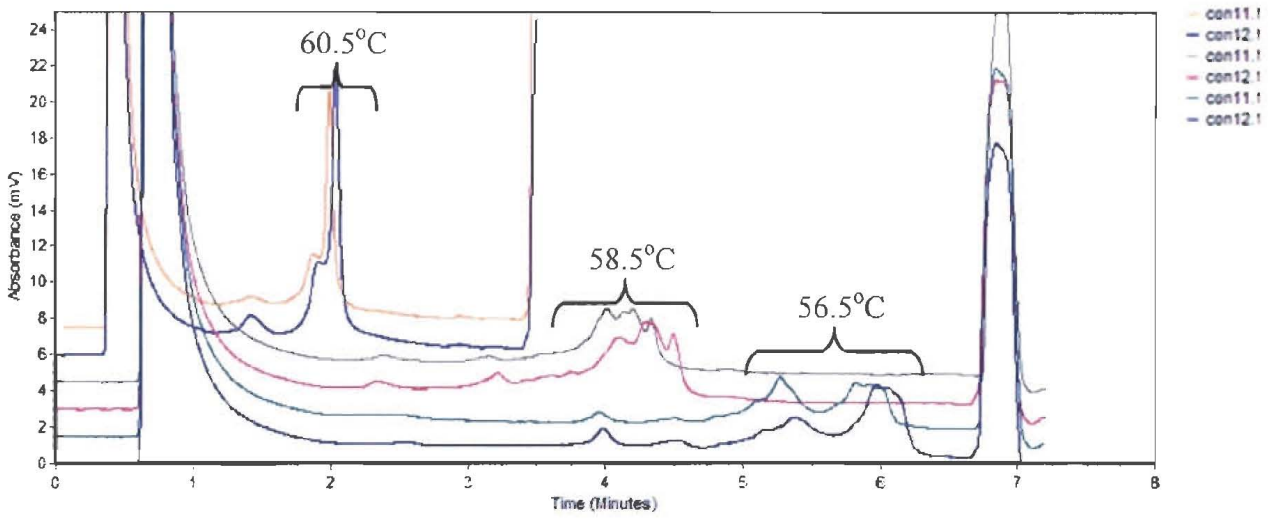


Figure 6-1: Chromatograms of exon 2, control samples con 11.1 and 12.1

6.2.3 Sequencing

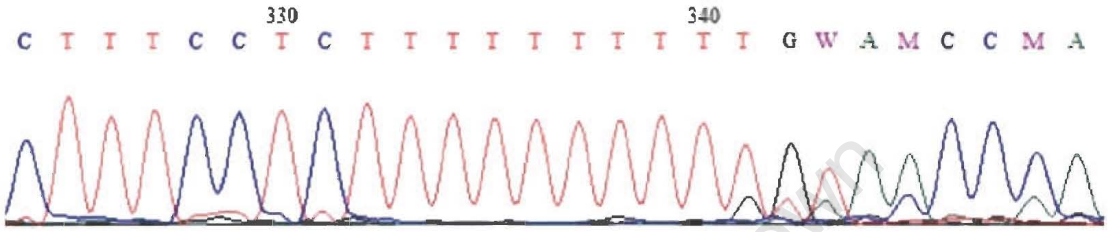
Difficulties were experienced in sequencing *LAMR1* exons 4, 5 and 6 as well as *CBX5* exons 2, 3 and 5.

Analysis of the sequence obtained for *LAMR1* exon 4 revealed that the wrong genomic sequence had been sequenced. Inspection of the exon 4 primers revealed that the forward primer was designed over a region that contained two SNPs. One of the SNPs was situated 2bp from the 3' end. The exon 4 primers were subsequently redesigned and named LAMR1_4bF & R.

Exons 5 and 6 were amplified in one amplicon. The sequence obtained from the 5+6 amplicon was of good quality until after a stretch of 10 "T" bases. This is illustrated in Figure 6-2. According to an online guide to sequencing by Hill and Helps (Chapter 8, Online resources), Taq DNA polymerase is not able to efficiently read through homopolymeric stretches of bases due to its lack of processivity. When Taq DNA polymerase reaches a homopolymeric stretch of bases one of two things can occur. The polymerase can either prematurely drop off the DNA, which results in

premature loss of signal strength, or it can "slip", which results in high background after the homopolymeric region. The latter appears to be the case in sequencing of amplicon 5+6. This did not however pose much of a problem as the 10 "T" bases were situated in the intron between exons 5 and 6. Exon 5 was therefore sequenced with the forward primer and exon 6 was sequenced with the reverse primer.

A



B

Figure 6-2: Sequence obtained for one patient where sequencing was performed with **A**, LAMR1_5+6F and with **B**, LAMR1_5+6R. The same stretch of sequence is shown; the reverse sequence has been reverse complemented by BioEdit.

Poor sequence quality was obtained when sequencing was performed with primers LAMR1_4bR, CBX5_2F, CBX5_3F&R and CBX5_5F&R. The sequence obtained for *CBX5* exons 2 and 3 is shown in Figure 6-3 and 6-4 respectively. These primers produced sequence of similar poor quality suggesting a common problem.

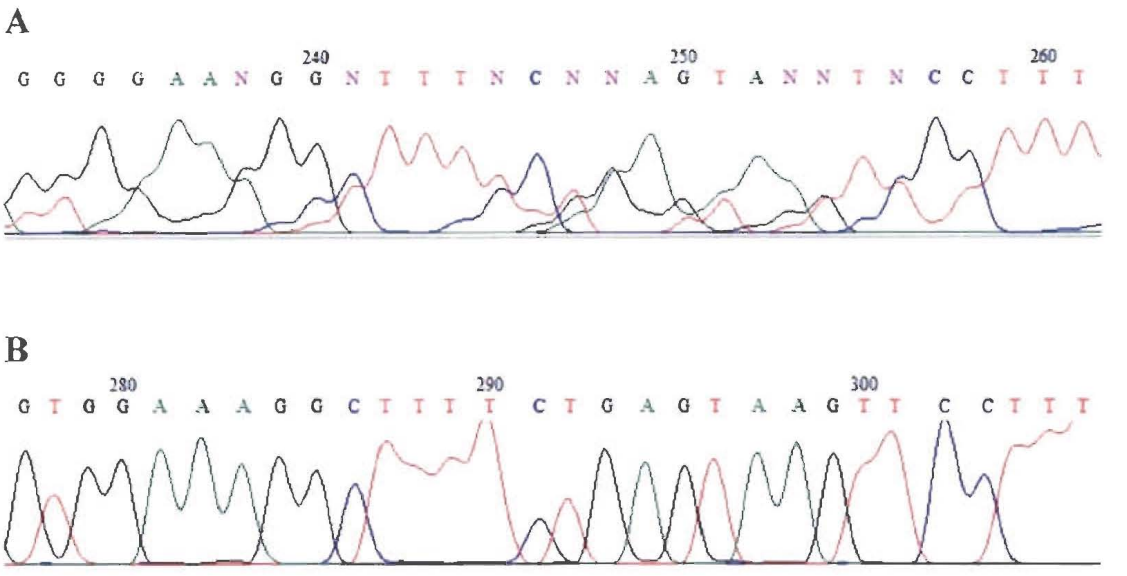


Figure 6-3: Sequence obtained for one patient where sequencing was performed with **A**, CBX5_2F and with **B**, CBX5_2R. The same stretch of sequence is shown; the reverse sequence has been reverse complemented by BioEdit.

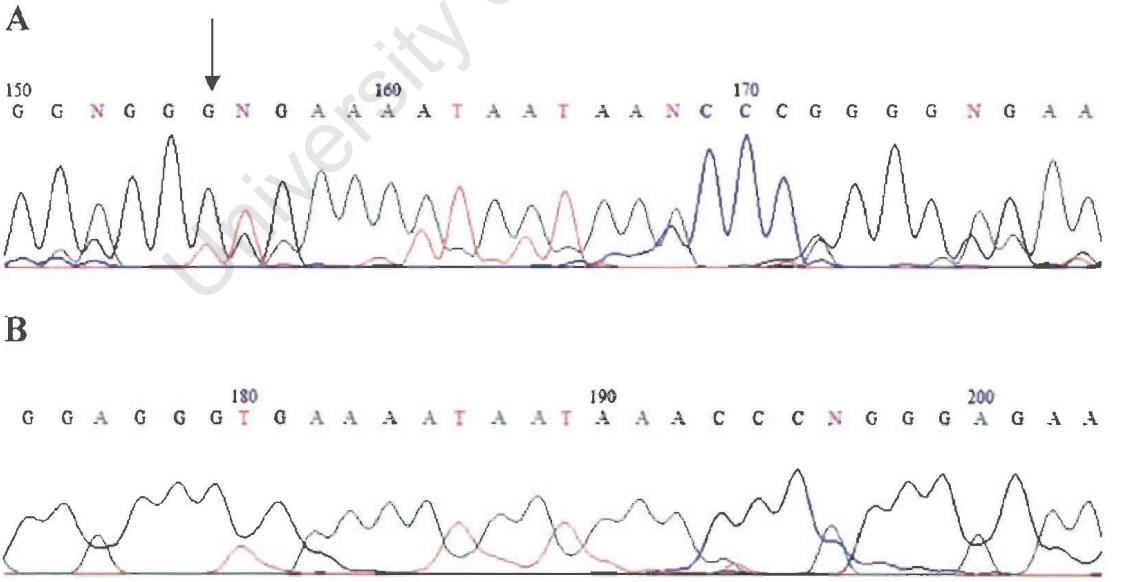


Figure 6-4: Sequence obtained for one patient where sequencing was performed with **A**, CBX5_3F and with **B**, CBX5_3R. The same stretch of sequence is shown; the reverse sequence has been reverse complemented by BioEdit.

On closer inspection of the sequences one observes a small peak before each correct peak, e.g., in Figure 6-4 at position 155 one observes a black peak indicating a G (which is the expected letter at this position) underneath which is a small red peak. This small red peak corresponds to the next letter at position 156 which is a red peak indicating a T (which is the expected letter at this position). The same is observed in the reverse sequence except that the small peak occurs after each correct peak because the sequence has been reverse complemented.

An online guide to sequencing by Hill and Helps describes this phenomenon and attributes it to poor quality primer. Primers are synthesised from the 3' to 5' end and primers that are the result of poor synthesis will have common 3' ends and different 5' ends. According to Hill and Helps these primers produce sequences containing a background "pre-read" where a small peak will be seen before each correct peak.

The purchase of the UVIpro Gold Gel Documentation System (*UVItec*, Whitehead Scientific) meant that the purified PCR product could be accurately quantified to determine the correct amount to add to the cycle sequencing reaction. This improved the sequence quality as previously the amount of purified PCR product added was estimated from the intensity of the PCR product on the gel photo.

6.3 Conclusion

This is the first re-sequencing study of the *LAMR1* gene in man to identify sequence variants that may cause ARVC. Although variants that could unequivocally be considered to cause ARVC were not detected, 7 new sequence variants have been discovered and a haplotype structure for the *LAMR1* gene has been proposed. This information will be invaluable in future gene association studies of *LAMR1*. Preliminary results for the *CBX5* mutation screen indicate that this gene may not be a disease gene in South African patients.

6.4 Future Work

The primers need to be re-synthesised for *CBX5* exons 3 and 5 to identify the sequence variants in these exons. If the sequence variant in *CBX5* exon 3 is found to be different in the groups with aberrant and spurious profiles then the variants need to be confirmed in each patient. The variants can be identified either by restriction enzyme digestion or by DHPLC. In the latter method each patient PCR product would be mixed with the PCR product of the patient sequenced in their group before the sample undergoes heteropduplex formation and is loaded on the WAVE^{R} System. If the patients have been correctly classified then all the chromatograms will look the same.

Further investigation of *LAMR1* sequence variants SV 3 and 15 is needed as intronic variants have been reported to affect splicing of the gene transcript and cause disease. A splicing functional assay or an *in vitro* splicing assay can be performed to determine if the variant affects splicing (Pagani and Baralle 2004). The two *LAMR1* related genes identified by Asano and colleagues also need to been screened.

All seven novel sequence variants identified in this study will be submitted to the SNP database.

Further research needs to be carried out to identify the disease causing genes in ARVC. Genes encoding desmosomal proteins would be a good place to start as these genes are the major cause of ARVC to date. Desmocollin-2 is expressed ubiquitously in desmosomal tissues and would be a good candidate gene to screen (Garrod et al. 2002).

Chapter 7:

References

7 REFERENCES

Ahmad F, Li D, Karibe A, Gonzalez O, Tapscott T, Hill R, Weilbaecher D, Blackie P, Furey M, Gardner M, Bachinski LL, Roberts R (1998) Localization of a gene responsible for arrhythmogenic right ventricular dysplasia to chromosome 3p23. Circulation 98:2791-2795

Alcalai R, Metzger S, Rosenheck S, Meiner V, Chajek-Shaul T (2003) A recessive mutation in desmoplakin causes arrhythmogenic right ventricular dysplasia, skin disorder, and woolly hair. J Am Coll Cardiol 42:319-327

Asano Y, Takashima S, Asakura M, Shintani Y, Liao Y, Minamino T, Asanuma H, Sanada S, Kim J, Ogai A, Fukushima T, Oikawa Y, Okazaki Y, Kaneda Y, Sato M, Miyazaki J, Kitamura S, Tomoike H, Kitakaze M, Hori M (2004) Lamr1 functional retroposon causes right ventricular dysplasia in mice. Nat Genet 36:123-130

Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263-265

Basso C, Wichter T, Danieli GA, Corrado D, Czarnowska E, Fontaine G, McKenna WJ, Nava A, Protonotarios N, Antoniades L, Wlodarska K, D'Alessi F, Thiene G (2004) Arrhythmogenic right ventricular cardiomyopathy: clinical registry and database, evaluation of therapies, pathology registry, DNA banking. Eur Heart J 25:531-534

Beffagna G, Occhi G, Nava A, Vitiello L, Ditadi A, Basso C, Bauce B, Carraro G, Thiene G, Towbin JA, Danieli GA, Rampazzo A (2005) Regulatory mutations in transforming growth factor-beta3 gene cause arrhythmogenic right ventricular cardiomyopathy type 1. Cardiovasc Res 65:366-373

Bradshaw D, Groenewald P, Laubscher R, Nannan N (2003) Initial estimates from the South African National Burden of Disease Study, 2000 MRC Policy Brief No1. MRC, Cape Town

Campbell RM (2005) The treatment of cardiac causes of sudden death, syncope, and seizure. Semin Pediatr Neurol 12:59-66

Carvajal-Huerta L (1998) Epidermolytic palmoplantar keratoderma with woolly hair and dilated cardiomyopathy. J Am Acad Dermatol 39:418-421

Chelu MG, Danila CI, Gilman CP, Hamilton SL (2004) Regulation of ryanodine receptors by FK506 binding proteins. Trends Cardiovasc Med 14:227-234

Coonar AS, Protonotarios N, Tsatsopoulou A, Needham EW, Houlston RS, Cliff S, Otter MI, Murday VA, Mattu RK, McKenna WJ (1998) Gene for arrhythmogenic right ventricular cardiomyopathy with diffuse nonepidermolytic palmoplantar keratoderma and woolly hair (Naxos disease) maps to 17q21. Circulation 97:2049-2058

Corrado D, Basso C, Rizzoli G, Schiavon M, Thiene G (2003) Does sports activity enhance the risk of sudden death in adolescents and young adults? J Am Coll Cardiol 42:1959-1963

Corrado D, Thiene G, Nava A, Rossi L, Pennelli N (1990) Sudden death in young competitive athletes: clinicopathologic correlations in 22 cases. Am J Med 89:588-596

Crawford DC, Nickerson DA (2005) Definition and clinical importance of haplotypes. Annu Rev Med 56:303-320

Dalla Volta S, Battaglia G, Zerbini E (1961) "Auricularization' of right ventricular pressure curve. Am Heart J 61:25-33

Dokuparti MV, Pamuru PR, Thakkar B, Tanjore RR, Nallari P (2005) Etiopathogenesis of arrhythmogenic right ventricular cardiomyopathy. J Hum Genet 50:375-381

Frances R, Rodriguez Benitez AM, Cohen DR (1997) Arrhythmogenic right ventricular dysplasia and anterior polar cataract. Am J Med Genet 73:125-126

Frias PA (2005) Hypertrophic cardiomyopathy and arrhythmogenic right ventricular dysplasia in young patients. Semin Pediatr Neurol 12:25-31

Garrod DR, Merritt AJ, Nie Z (2002) Desmosomal cadherins. Curr Opin Cell Biol 14:537-545

Gerull B, Heuser A, Wichter T, Paul M, Basson CT, McDermott DA, Lerman BB, Markowitz SM, Ellinor PT, MacRae CA, Peters S, Grossmann KS, Drenckhahn J, Michely B, Sasse-Klaassen S, Birchmeier W, Dietz R, Breithardt G, Schulze-Bahr E, Thierfelder L (2004) Mutations in the desmosomal protein plakophilin-2 are common in arrhythmogenic right ventricular cardiomyopathy. Nat Genet 36:1162-1164

Gilbert SF (2000) Chapter 15: Lateral plate mesoderm and endoderm. In: Developmental Biology. Sinauer Associates, Inc., Massachusetts, pp 471 & 472

Hamid MS, Norman M, Quraishi A, Firoozi S, Thaman R, Gimeno JR, Sachdev B, Rowland E, Elliott PM, McKenna WJ (2002) Prospective evaluation of relatives for familial arrhythmogenic right ventricular cardiomyopathy/dysplasia reveals a need to broaden diagnostic criteria. J Am Coll Cardiol 40:1445-1450

Hinds DA, Stuve LL, Nilsen GB, Halperin E, Eskin E, Ballinger DG, Frazer KA, Cox DR (2005) Whole-genome patterns of common DNA variation in three human populations. Science 307:1072-1079

Hodgkinson KA, Parfrey PS, Bassett AS, Kupprion C, Drenckhahn J, Norman MW, Thierfelder L, Stuckless SN, Dicks EL, McKenna WJ, Connors SP (2005) The impact of implantable cardioverter-defibrillator therapy on survival in autosomal-dominant arrhythmogenic right ventricular cardiomyopathy (ARVD5). J Am Coll Cardiol 45:400-408

Hughes SE, McKenna WJ (2005) New insights into the pathology of inherited cardiomyopathy. Heart 91:257-264

Jackers P, Minoletti F, Belotti D, Clausse N, Sozzi G, Sobel ME, Castronovo V (1996) Isolation from a multigene family of the active human gene of the metastasis-associated multifunctional protein 37LRP/p40 at chromosome 3p21.3. Oncogene 13:495-503

John AS, Mohiaddin RH, Sheppard MN (2004) Arrhythmogenic right ventricular cardiomyopathy. Heart 90:1102

Kaplan SR, Gard JJ, Carvajal-Huerta L, Ruiz-Cabezas JC, Thiene G, Saffitz JE (2004) Structural and molecular pathology of the heart in Carvajal syndrome. Cardiovasc Pathol 13:26-32

Kwok PY, Deng Q, Zakeri H, Taylor SL, Nickerson DA (1996) Increasing the information content of STS-based genome maps: identifying polymorphisms in mapped STSs. Genomics 31:123-126

Laitinen PJ, Brown KM, Piippo K, Swan H, Devaney JM, Brahmbhatt B, Donarum EA, Marino M, Tiso N, Viitasalo M, Toivonen L, Stephan DA, Kontula K (2001) Mutations of the cardiac ryanodine receptor (RyR2) gene in familial polymorphic ventricular tachycardia. Circulation 103:485-490

Latib MA, Michael KA, Mayosi BM (2004) Initial Report of the Arrhythmogenic Right Ventricular Cardiomyopathy Registry of South Africa: 2004 [Abstract]. Cardiovascular Journal of South Africa 15:237 & 238

Li D, Ahmad F, Gardner MJ, Weilbaecher D, Hill R, Karibe A, Gonzalez O, Tapscott T, Sharratt GP, Bachinski LL, Roberts R (2000) The locus of a novel gene responsible for arrhythmogenic right-ventricular dysplasia characterized by early onset and high penetrance maps to chromosome 10p12-p14. Am J Hum Genet 66:148-156

Marcus F, Towbin JA, Zareba W, Moss A, Calkins H, Brown M, Gear K (2003) Arrhythmogenic right ventricular dysplasia/cardiomyopathy (ARVD/C): a multidisciplinary study: design and protocol. Circulation 107:2975-2978

Marks AR (2000) Cardiac intracellular calcium release channels: role in heart failure. Circ Res 87:8-11

Maron BJ, Shirani J, Poliac LC, Mathenge R, Roberts WC, Mueller FO (1996) Sudden death in young competitive athletes. Clinical, demographic, and pathological profiles. Jama 276:199-204

Matolweni LO, Bardien S, Rebello G, Oppon E, Munclinger MJ, Ramesar R, Mayosi BM Arrhythmogenic right ventricular dysplasia/cardiomyopathy type 6 (ARVC6): corroboration of locus assignment, narrowing of the critical region and mutation screening of three candidate genes. (submitted)

Mayosi BM, Keavney B, Watkins H, Farrall M (2003) Measured haplotype analysis of the aldosterone synthase gene and heart size. Eur J Hum Genet 11:395-401

McGrath JA (2005) Inherited disorders of desmosomes. Australas J Dermatol 46:221-229

McKenna WJ, Thiene G, Nava A, Fontaliran F, Blomstrom-Lundqvist C, Fontaine G, Camerini F (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 71:215-218

McKoy G, Protonotarios N, Crosby A, Tsatsopoulou A, Anastasakis A, Coonar A, Norman M, Baboonian C, Jeffery S, McKenna WJ (2000) Identification of a deletion in plakoglobin in arrhythmogenic right ventricular cardiomyopathy with palmoplantar keratoderma and woolly hair (Naxos disease). Lancet 355:2119-2124

Melberg A, Oldfors A, Blomstrom-Lundqvist C, Stalberg E, Carlsson B, Larrson E, Lidell C, Eeg-Olofsson KE, Wikstrom G, Henriksson G, Dahl N (1999) Autosomal dominant myofibrillar myopathy with arrhythmogenic right ventricular cardiomyopathy linked to chromosome 10q. Ann Neurol 46:684-692

Michael KA, Latib MA, Matolweni LO, Mayosi BM (2004) Arrhythmogenic right ventricular cardiomyopathy: pathogenesis, diagnosis and management. SA Heart 1:13-22

Moolman-Smook JC, Mayosi BM, Brink PA, Corfield VA (2003) Molecular genetics of cardiomyopathy: changing times, shifting paradigms. Cardiovasc J S Afr 14:145-155

Munclinger MJ, Patel JJ, Mitha AS (2000) Follow-up of patients with arrhythmogenic right ventricular cardiomyopathy dysplasia. S Afr Med J 90:61-68

Nattel S, Schott JJ (2005) Arrhythmogenic right ventricular dysplasia type 1 and mutations in transforming growth factor beta3 gene regulatory regions: a breakthrough? Cardiovasc Res 65:302-304

Nickerson DA, Taylor SL, Weiss KM, Clark AG, Hutchinson RG, Stengard J, Salomaa V, Vartiainen E, Boerwinkle E, Sing CF (1998) DNA sequence diversity in a 9.7-kb region of the human lipoprotein lipase gene. Nat Genet 19:233-240

Norgett EE, Hatsell SJ, Carvajal-Huerta L, Cabezas JC, Common J, Purkis PE, Whittock N, Leigh IM, Stevens HP, Kelsell DP (2000) Recessive mutation in desmoplakin disrupts desmoplakin-intermediate filament interactions and causes dilated cardiomyopathy, woolly hair and keratoderma. Hum Mol Genet 9:2761-2766

Norman M, Simpson M, Mogensen J, Shaw A, Hughes S, Syrris P, Sen-Chowdhry S, Rowland E, Crosby A, McKenna WJ (2005) Novel mutation in desmoplakin causes arrhythmogenic left ventricular cardiomyopathy. Circulation 112:636-642

Oefner PJ, Underhill PA (1995) Comparative DNA sequencing by denaturing high-performance liquid chromatography (DHPLC). American Journal of Human Genetics 57:A266

Oefner PJ, Underhill PA (1999) DNA mutation detection using denaturing high-performance liquid chromatography (DHPLC). In: Current Protocols in Human Genetics. Wiley-Interscience, New York, pp 1 - 12

Pagani F, Baralle FE (2004) Genomic variants in exons and introns: identifying the splicing spoilers. Nat Rev Genet 5:389-396

Paul M, Schulze-Bahr E, Breithardt G, Wichter T (2003) Genetics of arrhythmogenic right ventricular cardiomyopathy--status quo and future perspectives. Z Kardiol 92:128-136

Peters S, Trummel M, Meyners W (2004) Prevalence of right ventricular dysplasia-cardiomyopathy in a non-referral hospital. Int J Cardiol 97:499-501

Pilichou K, Nava A, Basso C, Beffagna G, Bauce B, Lorenzon A, Frigo G, Vettori A, Valente M, Towbin J, Thiene G, Danieli GA, Rampazzo A (2006) Mutations in Desmoglein-2 Gene Are Associated With Arrhythmogenic Right Ventricular Cardiomyopathy. Circulation

Priori SG, Barhanin J, Hauer RN, Haverkamp W, Jongsma HJ, Kleber AG, McKenna WJ, Roden DM, Rudy Y, Schwartz K, Schwartz PJ, Towbin JA, Wilde AM (1999) Genetic and molecular basis of cardiac arrhythmias: impact on clinical management parts I and II. Circulation 99:518-528

Priori SG, Napolitano C, Tiso N, Memmi M, Vignati G, Bloise R, Sorrentino V, Danieli GA (2001) Mutations in the cardiac ryanodine receptor gene (hRyR2) underlie catecholaminergic polymorphic ventricular tachycardia. Circulation 103:196-200

Protonotarios N, Tsatsopoulou A, Anastasakis A, Sevdalis E, McKoy G, Stratos K, Gatzoulis K, Tentolouris K, Spiliopoulou C, Panagiotakos D, McKenna W, Toutouzas P (2001) Genotype-phenotype assessment in autosomal recessive arrhythmogenic right ventricular cardiomyopathy (Naxos disease) caused by a deletion in plakoglobin. J Am Coll Cardiol 38:1477-1484

Protonotarios N, Tsatsopoulou A, Patsourakos P, Alexopoulos D, Gezerlis P, Simitsis S, Scampardonis G (1986) Cardiac abnormalities in familial palmoplantar keratosis. Br Heart J 56:321-326

Protonotarios NI, Tsatsopoulou AA, Gatzoulis KA (2002) Arrhythmogenic right ventricular cardiomyopathy caused by a deletion in plakoglobin (Naxos disease). Card Electrophysiol Rev 6:72-80

Rampazzo A, Nava A, Danieli GA, Buja G, Daliento L, Fasoli G, Scognamiglio R, Corrado D, Thiene G (1994) The gene for arrhythmogenic right ventricular cardiomyopathy maps to chromosome 14q23-q24. Hum Mol Genet 3:959-962

Rampazzo A, Nava A, Erne P, Eberhard M, Vian E, Slomp P, Tiso N, Thiene G, Danieli GA (1995) A new locus for arrhythmogenic right ventricular cardiomyopathy (ARVD2) maps to chromosome 1q42-q43. Hum Mol Genet 4:2151-2154

Rampazzo A, Nava A, Malacrida S, Beffagna G, Bauce B, Rossi V, Zimbello R, Simionati B, Basso C, Thiene G, Towbin JA, Danieli GA (2002) Mutation in human desmoplakin domain binding to plakoglobin causes a dominant form of arrhythmogenic right ventricular cardiomyopathy. Am J Hum Genet 71:1200-1206

Rampazzo A, Nava A, Miorin M, Fonderico P, Pope B, Tiso N, Livolsi B, Zimbello R, Thiene G, Danieli GA (1997) ARVD4, a new locus for arrhythmogenic right ventricular cardiomyopathy, maps to chromosome 2 long arm. Genomics 45:259-263

Richardson P, McKenna W, Bristow M, Maisch B, Mautner B, O'Connell J, Olsen E, Thiene G, Goodwin J, Gyarfas I, Martin I, Nordet P (1996) Report of the 1995 World Health Organization/International Society and Federation of Cardiology Task Force on the Definition and Classification of cardiomyopathies. Circulation 93:841-842

Seidman JG, Seidman C (2001) The genetic basis for cardiomyopathy: from mutation identification to mechanistic paradigms. Cell 104:557-567

Severini GM, Krajinovic M, Pinamonti B, Sinagra G, Fioretti P, Brunazzi MC, Falaschi A, Camerini F, Giacca M, Mestroni L (1996) A new locus for arrhythmogenic right ventricular dysplasia on the long arm of chromosome 14. Genomics 31:193-200

Sherwood L (2004) Chapter 9: Cardiac Physiology. In: Human Physiology. Brooks/Cole, California, pp 303-307

Sliwa K, Damasceno A, Mayosi BM (2005) Epidemiology and etiology of cardiomyopathy in Africa. Circulation 112:3577-3583

Starr C, Taggart R (1998) Chapter 39: Circulation. In: Biology: The Unity and Diversity of Life. Wadsworth Publishing Company, California, pp 658 & 659

Stephens M, Donnelly P (2003) A comparison of bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet 73:1162-1169

Stephens M, Scheet P (2005) Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. Am J Hum Genet 76:449-462

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68:978-989

Strachan T, Read AP (1999) Chapter 15: Identifying human disease genes. In: Human Molecular Genetics 2. John Wiley & Sons Inc., New York, pp 367-372

Syrris P, Ward D, Asimaki A, Sen-Chowdhry S, Ebrahim HY, Evans A, Hitomi N, Norman M, Pantazis A, Shaw AL, Elliott PM, McKenna WJ (2006) Clinical

expression of plakophilin-2 mutations in familial arrhythmogenic right ventricular cardiomyopathy. Circulation 113:356-364
Thiene G, Basso C, Calabrese F, Angelini A, Valente M (2005) Twenty years of progress and beckoning frontiers in cardiovascular pathology: cardiomyopathies. Cardiovasc Pathol 14:165-169
Thiene G, Corrado D, Basso C (2004) Cardiomyopathies: is it time for a molecular classification? Eur Heart J 25:1772-1775
Thiene G, Corrado D, Nava A, Rossi L, Poletti A, Boffa GM, Daliento L, Pennelli N (1991) Right ventricular cardiomyopathy: is there evidence of an inflammatory aetiology? Eur Heart J 12 Suppl D:22-25
Thiene G, Nava A, Corrado D, Rossi L, Pennelli N (1988) Right ventricular cardiomyopathy and sudden death in young people. N Engl J Med 318:129-133
Tiso N, Stephan DA, Nava A, Bagattin A, Devaney JM, Stanchi F, Larderet G, Brahmbhatt B, Brown K, Bauce B, Muriago M, Basso C, Thiene G, Danieli GA, Rampazzo A (2001) Identification of mutations in the cardiac ryanodine receptor gene in families affected with arrhythmogenic right ventricular cardiomyopathy type 2 (ARVD2). Hum Mol Genet 10:189-194
Van De Graaff KM (2002) Chapter 16: Circulatory System. In: Human Anatomy. McGraw-Hill, New York, pp 545-550
Wichter T, Breithardt G (2005) Implantable cardioverter-defibrillator therapy in arrhythmogenic right ventricular cardiomyopathy: a role for genotyping in decision-making? J Am Coll Cardiol 45:409-411
Williams RG, Chen AY (2003) Identifying athletes at risk for sudden death. J Am Coll Cardiol 42:1964-1966
Wolfe BA, Corwin EJ (2005) ARVD: a silent killer. Nurse Pract 30:16-18, 23-17; quiz 27-19
Xiao W, Oefner PJ (2001) Denaturing high-performance liquid chromatography: A review. Hum Mutat 17:439-474
Yin T, Green KJ (2004) Regulation of desmosome assembly and adhesion. Semin Cell Dev Biol 15:665-677

Chapter 8:

Electronic resources

8 ELECTRONIC RESOURCES

BLAST search on NCBI

http://www.ncbi.nlm.nih.gov/blast/

Hill AJM, Helps, NR (2000) A Guide to Automated DNA Sequencing v1.2.

http://www.dnaseq.co.uk/Downloads/sequencing_guide.pdf

Human Genome Variation Society (HGVS) mutation nomenclature rules

http://www.HGVS.org/mutnomen/

OligoAnalyzer

http://scitools.idtdna.com/scitools/Applications/Oligoanalyzer/Default.aspx

Primer 3

http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi

Splice Site Prediction by Neural Network available on the Berkeley Drosophila Genome Project website

www.fruitfly.org/seq_tools/splice.html

Chapter 9:

Appendices

REQUEST FOR MOLECULAR STUDIES (DNA)

Molecular Laboratory
Division of Human Genetics
Suite 3.21, IIDMM,
UCT Medical School, Observatory 7925

Tel: (021) 406 6425 Fax: (021) 448-0906

Blood should be drawn in 2 plastic EDTA Tubes
(Purple top) – 10ml each using a yellow barrel
Each tube should be inverted to mix and should be
clearly labelled with the patient's name and DOB
Keep blood in fridge at 4°C until able to send to laboratory

Please DO NOT send specimens on ice or frozen.

Please fill in all the information requested:

Surname: _____ First Name(s): _____

New Family: Yes ☐ No ☐ (If no, please fill in family name) Family name: _____

Medical Aid: _____ Medical Aid No: _____

Sex: M ☐ F ☐ Date of Birth: Year: _____ Month: _____ Day: _____

Number of children: _____

Ethnic Origin: (please indicate ancestry of both your mother and father) _____

Contact Address: _____ Town: _____ Fax: Tel: _____

Referring Doctor/Sister _____ Town: _____ Fax: Tel: _____

Hospital or Address: _____ Town: _____ Fax: Tel: _____

Reason for Referral (Clinical diagnosis):

Affected ☐ At Risk ☐ Carrier ☐ Spouse ☐ Query ☐ Unaffected ☐

Arrhythmogenic Right Ventricular Cardiomyopathy ☐

Additional disorders (apparent or previously treated): _____

Additional family history _____

Clinical Details:

Physical disability ☐ Mental retardation ☐ Deafness ☐ Impaired vision ☐ Night blindness ☐

Other: _____

Have samples from this patient been sent to a DNA lab before? (DELETE WHERE NOT APPLICABLE) YES / NO / Don't Know

If Yes, where: _____

For Laboratory use only:

DNA number: _____ *Vol Blood:* _____ *(ml) Other:* _____

Date Received: Year: _____ *Month* _____ *Day:* _____ *Computer Index No* _____

CONSENT FOR DNA ANALYSIS AND STORAGE

1 I, _____, request that an attempt be made using genetic material to assess the probability that I might have inherited a disease-causing mutation in the gene for arrhythmogenic right ventricular cardiomyopathy

2 I understand that the genetic material for analysis is to be obtained from: blood cells/other (specify) (DELETE WHERE NOT APPLICABLE).

3 I request that **no** portion of the sample be stored for later use. ☐ (MARK IF APPLICABLE)

Or

I request that a portion of the sample be stored indefinitely for (DELETE WHERE NOT APPLICABLE):

(a) possible re-analysis
(b) analysis for the benefit of members of my immediate family
(c) research purposes, subject to the approval of the University of Cape Town Research Ethics Committee, provided that any information from such research will remain confidential

4 The results of the analysis carried out on this sample of stored biological material will be made known to me, via my doctor, in accordance with the relevant protocol, if and when available.
In addition, I authorise that they may be made known to: (DELETE WHERE NOT APPLICABLE):
other doctors involved in my care
the following family members: _____
other: _____

5 I authorise / do not authorise my doctor(s) (DELETE WHERE NOT APPLICABLE) to provide relevant clinical details to the Division of Human Genetics, UCT.

6 I have been informed that:

(a) there are risks and benefits associated with genetic analysis and storage of biological material and these have been explained to me
(b) the analysis procedure is specific to the genetic condition mentioned above and cannot determine the complete genetic makeup of an individual.
(c) the genetics laboratory is under an obligation to respect medical confidentiality
(d) genetic analysis may not be informative for some families or family members
(e) even under the best conditions, current technology of this type is not perfect and could lead to incorrect results
(f) where biological material is used for research purposes, there may be no direct benefit to me

7 I understand that I may withdraw my consent for any aspect of the above at any time without this affecting my future medical care

8 **ALL OF THE ABOVE HAS BEEN EXPLAINED TO ME IN A LANGUAGE THAT I UNDERSTAND AND MY QUESTIONS ANSWERED BY:**

_____ DATE: _____

Patient signature _____ **Witnessed consent** _____

NOTE - PLEASE INSERT A FAMILY PEDIGREE DRAWING ON THE REVERSE OF THIS FORM

Appendix 2: Gel Formulas

1% (w/v) Agarose gel with ethidium bromide (100ml)

1g Agarose

100ml 1X TBE (see Appendix 3)

6µl Ethidium bromide (10mg/ml)

1.5% (w/v) Agarose gel (100ml)

1.5g Agarose

100ml 1X TBE (see Appendix 3)

8% PAGE Gel

200ml 40% Acrylagel

100ml 10X TBE

422g Urea

Make up to 1L with dH_2O

Microwave for 5 minutes to dissolve urea

Appendix 3: Solutions and Buffers

10X Tris Borate EDTA (TBE) buffer (2L)

216g Tris

110g Boric acid

14.8g EDTA

Make up to 2L with dH_2O

Agarose Loading Dye

0.125g (0.25%) Bromophenol blue

20g (40%) Sucrose

Make up to 50ml with dH_2O and ensure the pH is 8 or more basic

Agarose Loading Dye with SYBR^R Gold nucleic acid gel stain

Make a 1:500 dilution of SYBR^R Gold in agarose loading dye (i.e., 2µl:1ml)

Silver staining solutions

Solution 1:

1g Silver Nitrate

Make up to 1L with dH_2O

Solution 2:

15g NaOH

800ml dH_2O

Mix on magnetic stirrer

Add 10ml formaldehyde

Make up to 1L

Appendix 4: DNA Ladders

Figure 9-1: 1 Kb Plus DNA Ladder. A volume of 0.7μg of DNA ladder was electrophoresed on a 0.9% agarose gel.

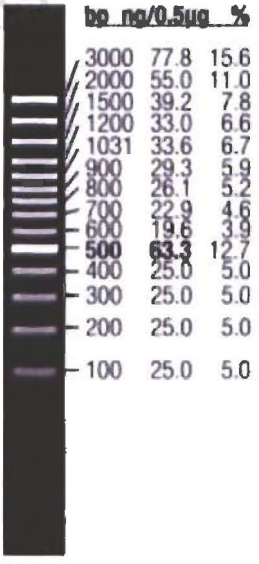


Figure 9-2: GeneRuler^{TM} 100bp DNA Ladder Plus. A volume of 0.5μg of DNA ladder was electrophoresed on a 1.7% agarose gel.

Appendix 5: *LAMR1* annotated sequence

The *LAMR1* exons and sequence variants have been annotated in the genome sequence. A "^" sign precedes the sequence variant. In the following line the position of this sequence variant (in this annotation), the nature of the sequence variant and the SNP accession number (if it is known) is given. The sequence variants identified in this study are highlighted and labelled SV 1-17.

```
Annotv9 - George Rebello, May 2004

Summary of Exons in sequence:

LAMR1-Exon1(3001-3051 -> 51bp)
LAMR1-Exon2(3908-4073 -> 166bp)
LAMR1-Exon3(4893-5011 -> 119bp)
LAMR1-Exon4(7041-7286 -> 246bp)
LAMR1-Exon5(7936-8064 -> 129bp)
LAMR1-Exon6(8183-8348 -> 166bp)
LAMR1-Exon7(8561-8721 -> 161bp)
Total size for LAMR1 = 1038 bp

Total of 72 SNPs in annotation

Total of 3 STSs in annotation


RNA splicing branch sites are searched for in the 60 bases before the
splice acceptor site (3' of intron), only 4 patterns are searched for:
         "TCCTRAY(found in U12 introns)", "YNYYRAY", "YRYRAY", and "CTRAY"

Note! IUB codes used for DNA in Branch site descriptors:
        A = Adenosine, C = Cytosine, G = Guanosine, T = Thymidine,
        R = A or G (puRine), Y = C or T (pYrimidine), K = G or T (Keto),
        M = A or C (aMino), S = G or C (Strong), W = A or T (Weak),
        N = A or C or G or T (aNy base).

Annotation based on:
LOCUS       NM_002295               1039 bp    mRNA    linear   PRI 20-MAY-2004
DEFINITION  Homo sapiens laminin receptor 1 (ribosomal protein SA, 67kDa)
            (LAMR1), mRNA.


Annotated sequence file:

LOCUS       NT_022517              11721 bp    DNA     linear   CON 19-FEB-2004
DEFINITION  Homo sapiens chromosome 3 genomic contig.
ACCESSION   NT_022517 REGION: 39370797..39382517
VERSION     NT_022517.16  GI:37550163
```

```
ORIGIN
      1 aatactctag agtctaggat cgtggtttrt aaactcaata tggctgcaca ttggaatcat
     61 ctggaaaatg tt^ggggctca gaaaatgata ccccctgactg aataaggcac tttggcatgc
        (73)(a/g)(dbSNP:13062836)
    121 tgcatgcttt gaatttaaaga aaatkgaaag gcttctgaaa tagcgtcaga accaaggttt
    181 ctttctgacc ttccccctaac cccatctctc aatccccttt tcccaaagca tccggagggg
    241 ctttgtctga agttcccctta tctgaataag gtaggtkcct ccagaataaa atacaatagt
    301 cttaaaactt cctcccctaag aatctcatca ggtaaccagg aaatattaac cactgggaag
    361 agaaaagact aaaagtcatc a^ccatgccca gacagactct tcatctattt ttctgagggc
        (382)(a/c)(dbSNP:13067212)
    421 agcttttgag agattttcca ggagacatta tttgcctaat aaggcaacct ttggtcacag
    481 tgaagttcct ccccttaccc tttgctgctg tttccc^ccag agctcacagg aactttgtct
        (517)(c/t)(dbSNP:6806261)
    541 caggcaaatt ttctgtttct tgaactcatt cctat^ccttc caaaaatcac ttactagccc
        (576)(c/t)(dbSNP:6806372)
    601 tctaaattgt ttaccatccc atctcccttc cctgtgaagt agatgctatt taagcttcag
    661 ccat^ctggcc cttctttgag tcttatattt c^gtgtggctc ctgtgcataa gcacattaat
        (665)(a/c)(dbSNP:6806483), (692)(a/g)(dbSNP:6781576)
    721 aaatgtgtat gttttgg^cctg gtgtggtggc tcacgcctgt aatcccagca ctttgggagg
        (737)(c/t)(dbSNP:13067935)
    781 ccgaggtggg cagataacct gaggtcagga gttcaagtcc agcctgacca acatggagca
    841 accccatttc tactaaaaat acaaaattag ccgggtgtgg tggcgcatgc ctataatccc
    901 agctactcga gaggctgagg caggagaatt gcttgaaccc gagagg^tgga ggttgcgttg
        (947)(c/t)(dbSNP:2133578)
    961 agccgagatt gagccactgc actccagcct gggcaacaag agcaaaactc cgtctcaaaa
   1021 aaaaaaagtg tatgtttctt ctcctgtt^aa tttgrctatt gtcagttcat tttagcagtg
        (1049)(a/g)(dbSNP:9817347)
   1081 aac^actcaga gggtggaggg gaagccccta caagagcttt aaaagctact gatgcttggg
        (1084)(a/g)(dbSNP:2133579)
   1141 tctcactcag agattctgat atattcagcc caggtgtaag attttaaaag ctgcccaggt
   1201 gatcctaatg ccaaggctga ggcccctgga ggaaccctgc agagctgcac tcactactca
   1261 gcaat^cccaa tccacacgtg aagggttaag caggagtaca tttgggtgca gggggcaggg
        (1266)(c/t)(dbSNP:2173603)
   1321 catkgttttg ggactcaggg gcctctgagg aggcttgtgg cctgagtttc tgccaacaaa
   1381 ctgaatgcct cctcttccca taacaaagcg tagaagcggg aggcatttgc attrgccgca
   1441 aggggctgtg actttctctga atttgtccct caagccccag gtgcttcagc cttggtgggt
   1501 gatgaagcag agcagcctgg gggtgacctg tgctagggtc caagggtttc caagctcctc
   1561 agaaaatcat actggcttca gtggagctgc agagccactt gcaggagaac ctggatgtgc
   1621 tgtcctttgg aagagtggcc ataagggtct acgagttcat tttagggtgg gctcagaaat
   1681 cagatkgact cagcaggccc cagagaggaa cctgcacttt ttttttttct gacggagtct
   1741 tactctgtcg cccaggctgg agtgcaagtg gcgcaatctc cactcac^tac aacctctgcc
        (1788)(c/t)(dbSNP:13096784)
   1801 tcctgggttc aagcaattct cccgcctcag cctcccgagt agctgggatt acaggcatgc
   1861 c^tggctaatt tttgaatttt tagtagatac cgggtttcgc catgttggcc aggctggtct
        (1862)(c/t)(dbSNP:13096961)
   1921 cttgaactcc tggcctcaag ggatccgccc gcgtcagcct ccgaaagtgc tgggattaca
   1981 gccaccgcgc ccctcaattt aactcctatt ttttgttttg ttttttgg^ag acagaatttc
        (2029)(a/g)(dbSNP:6776627)
   2041 gctctattgc ccaggctgga gtgcagcggt gcaatcacat tcactgcag^a cttgacctct
        (2090)(a/c)(dbSNP:4676494)
   2101 aggctcaagc tatcctcctg cctcagcctc ctgagtagct gagactacag ggacgcacca
   2161 tcatgattgg ttaatttttg tatttttttgt ggagatgggg gtctccccat gttttcccagg
   2221 ctagtcttga acttctgagc tcaggcaatc cacccacctc ggcctcccaa agtgctggga
   2281 ttacaggcac aagccaccg^c gcccggcaaa cttatctcct actgatggat a^ttgcaaaat
        (2300)(c/g)(dbSNP:7611864), (2332)(c/t)(dbSNP:7625535)
   2341 ttttcttcta c^caaaaaaaa a^aaaaaaaaa aagatagata taatcca^ta^c catctgttca
        (2352)(a/-)(dbSNP:11409391), (2362)(a/-)(dbSNP:11394743),
(2388)(c/t)(dbSNP:7625610), (2390)(c/t)(dbSNP:11129828)
   2401 atcttgtcct ttttaacttt ttcagggaaa cttccccccag gtgatagatg gatagataca
   2461 tgagatagа^c ataacgcata cgatcagttc aatctt^atct tttttaactt tttcagagaa
        (2470)(c/t)(dbSNP:7612064), (2497)(a/g)(dbSNP:7623381)
```

2521 aacttccctt aagtgaacat ttaaatctga attacgtcct gttaaactgt tctccaggaa
2581 aataaaataa aataaatctt caagtttttg tttacctaac aatttgttgt gtcgaacaaa
2641 ccttcctact tttcaggtaa caaaatggca gcttaggcta gaaagccgct catattcgca
2701 ggtacaaggg ctgggtaaga acgccccgcc tggctgacta acttgagttc cgcgctctgg
2761 acaggaatta tgcacagggc gtcgctgtgg cactagaaac cccaaagtca caagcgcccc
2821 agatccgacc aggatgccgc tac^cgg^ctac agccca^ga^gg ^ccgctcctg cggcgcaagc
(2844)(c/g)(dbSNP:11925942), SV 1 (2847)(c/g)(dbSNP:11717439),
(2857)(a/g)(dbSNP:11926001), SV 2 (2859)(a/g)(dbSNP:3772147),
SV 3 (2861) (c/a)
2881 ccgccttcct gaagaaaaag cccagtcccg gcagcgttct tctccggctc cgc^ccttctt
2941 ccgctcgact ttctttgcca ttgg^ctgaca acggagtaca taaggacgtc atttcctgc^c
SV 4 (2934_2953) (del20), (2965)(branch site - CTRAY),
SV 5 (3000)(c/t)(dbSNP:3772146)

LAMR1-Exon1(3001-3051 -> 51bp)
3001 GCCTGTCTTT TCCGTGCTAC CTGCAGAGGG GTCCATACGG CGTTGTTCTG Ggtgagttcc
3061 gtgtagcgtc cctggcgcct tccagggcta gaaaaatgag cttttcctgc tcaaatgaag
3121 ggtgagaaga ctagtgatga aagccggtca gactggatct gtctcccgcc cggcgcgccc
3181 caccttaggc ctgcggcccg cacgtggcca ggct^cgggc^t ggcgggttcc cagagtgccc
(3215)(c/t)(dbSNP:3772145), SV 6 (3220)(c/t)(dbSNP:3772144)
3241 cgggagcggg tggaggccgc cctccagcgg aggctccgag ctggggttcg gaccaggccg
3301 cgggtgggcg ggagtgcaga aagcgggcta acatcctgtg ttgctatccc ttcggagtcc
3361 cacacggcgg tgagtc^tagg cccaggcgct gatttacacc agctacttgg gctgg^tcggg
(3377)(t/g)(dbSNP:3772143), (3416)(c/t)(dbSNP:11706951)
3421 ttttcccttc gcgccgtgcg ggtcaggagt taaggttctc gggtttttag acaaacaagt
3481 ggtgacagca cagcgaagta attccaaagc atccgcctac aatctgc^ttg aaaatgtctg
(3528)(a/t)(dbSNP:3772142)
3541 aaaacaattc atgccttttt tgcctttagt ttgcatattc caaacatggc tgctctttt^g
(3600)(a/g)(dbSNP:1472508)
3601 tatctagttg ttaacttggc gcatccacaa cttttcctta attcctatct tgagaagtgt
3661 tgaatttcca ttcgctaatt tcgtgtagtt ttattactcg gttactctgc cgtccacact
3721 atttcctcag taagatgtgc gctgttccgt aatacacgac atgtatgggt taactttctg
3781 tttacccttc actacactgt aagctcaatg cctgacacta tagcagatgg agtttttggt
3841 tgcttttaag g^gtgtgccct acttaactca atggaatgaa aagaaatagg ttgctctctt
SV 7 (3852) (g/a)

LAMR1-Exon2(3908-4073 -> 166bp)
3901 atttcagATT CCCGTCGTAA CTTAAAGGGA AA^TTTTCACA ^ATGTCCGGAG CCCTTGATGT
SV 8 (3933)(c/t)(dbSNP:1803893), start
3961 CCTGCAAATG AAGGAGGAGG ATGTCCTTAA GTTCCTTGCA GCAGGAACCC ACTTAGGTGG
4021 CACCAATCTT GACTTCCAGA TGGAACAGTA CATCTATAAA AGGAAAAGTG ATGgttagtc
4081 attgctttaa ttttttgtta ctccagctgt aagtacaaat tttgagcttg ctattctcgt
4141 ggttagttct gggtaatttc tttctatctt ccttaaatga agccagaccc ctacgttgaa
4201 aacatacttt aaataaatgt tcttgttatt gagaga^aaag atttctgctc caggggagga
SV 9 (4237)(a/g)(dbSNP:3772141)
4261 atatgtcagt tgtctgagtt acgggacttg ggttgggtca tgaacatgag aaagttcatt
4321 ggctgagttt ctaatagctc gcatagtgcc ttcaacttca taggtactta cacaggctat
4381 tgaactgaat tttgagtgga aagtggggta agaggtggga atgaacaata tggtttggaa
4441 gaaccagagg atggaaggca ttaagttggc taaggcactt acttatttta gagataaagc
4501 taccaaggac ttgggagtga catgctgtaa aggagtttta ggatgaatct tgcatcagtg
4561 cagat^tataa aagggaaaga gtggcagaaa gccaggaagt ggatt^attaa gaaccaatga
(4566)(t/g)(dbSNP:3772140), (4606)(a/c)(dbSNP:3772139)
4621 tggaataagt acagggaaag agaaaggaaa gatggattta gccaactgaa tctaggctgc
4681 acactattaa agctcagggt ggaggccagt cttggctcat gaacttctga gtgtcggaag
4741 tgtgctatat caatggcagg attttcgcta acaccagtag agcttgcctc tatgactgga
4801 gtttggtagt actcgctgcc acatagagta aacttgagaa gaatgtttgc acagccaggt

LAMR1-Exon3(4893-5011 -> 119bp)
4861 ^caagtgttac aaatccttct gccctcactt agGCATCTAT ATCATAAATC TCAAGAGGAC
SV 10 (4861)(c/t)(dbSNP:3772138)
4921 CTGGGAGAAG CTTCTGCTGG CAGCTCGTGC AATTGTTGCC ATTGAAAACC CTGCTGATGT
4981 CAGTGTTATA TC^CTCCAGGA ATACTGGCCA Ggtttgtgga acagtggtta gtttttatat
SV 11 (4993) (c/t)
5041 tatagaaata aagcttacag acattgtgag acata^gaaag acataagaaa acagaaataa

SV 12 (5075_5076) (insTA)
5101 ctcaggccgg gcgcggtggc tacgcctgta atcccagcac tttgggaggt ggagtcc^ggc
SV 13 (5158) (g/a)
5161 gtattacgag gtcaggagat ccacaccacg gtgaagcccc atctctacta aaaatacaaa
5221 aaattagttg ggcgtggggg cgggtgcctg tagtcccagc tactcaggag gccaa^agtgg
(5276)(a/g)(dbSNP:7632488)
5281 agaatggcgt gaacccggtg ggcggagctt gcagtgagct gagatgcgcc actgtactcc
5341 agcctgggtg acagagccga gactctgtct caaaaaaaaa aaaaaagaaa atagaaaact
5401 caaactccac tgttaagaga tcttaagttg gactgagcca tagaaatcgc ccttatgacc
5461 tatgaccttc ctttttttt ttttttttt ttttccttca agacagtctc tcgctctgtt
5521 gcctgttgtc catcctggac tgcagtggtg tgatatcggc tcattgcagc ctttgcctcc
5581 taagctcaag ggattctcct gcctcatcct cccgagtagc tggaattaca ggcatgtgcc
5641 accatgcctg actgattttt gtatttttag tagagacatg gttttaccat ttggccaggc
5701 tggtcttaaa cgcctgatct caagtgatcc tcccacctcg gcctcccaaa gtgttgggat
5761 tacaggcgta agccactgcg cccagcctgt aagacctttt caacatacat gctgggcatc
5821 tggcagtatg ttgaacttta taatttcatt ttgtgtagac ttgtgccact aaatatctag
5881 aaattctatt aaagtggctt atcacatttt cccatagtca aagtttttac cttaaaacaa
5941 tgaaatactg ggacagaaat tttttaagaa tcttagccag tcatgacggc atgcacccat
6001 agttcccgtt aattgggagt ctgagccagg agatccctca agtccaggag ctgggggctg
6061 cagtaagtta tgattttgcc actgcactcc aagcctgggg cacagagtga gatcctgttt
6121 cttaaaa^aag ccagtcatgg gtggctcatg cttttgtaat atggggtact ttgtgaggct
(6128)(a/c)(dbSNP:11921564)
6181 gaggctggtg gcttgcttga gcccagaagt tcgagaccag ggtgggcaac gtggcaaaac
6241 tccgtctcta ccaaaattaa caaaaacgag ctgggtgtgg tgttggcgtg cctgtagtct
6301 cagctgaggc t^aaggtggg^a agattgctag aacacaggag gttgaggcag cagggagcca
(6312)(a/g)(dbSNP:7619128), (6320)(a/g)(dbSNP:7619131)
6361 agactgtcca ctgcatccca gcctgagtga cagagtgagg ccttgtctca aagagaacct
6421 ctgccaactt gaatgaaaaa taccttaatt tgtattttta agaaatgact agttagtaaa
6481 acttgaatct tttttgccac agctcta^ttt taataatttt ^tctgttgaga aaactagtag
(6508)(a/t)(dbSNP:7621701), (6521)(t/-)(dbSNP:11431351)
6541 atttcttgaa ttctaagagt tctaacaggg attttattca tcttaactat tttgacagat
6601 ttaaaaatta tgtagtaaga ctttatttca atgttgacag gacgggttcc catgtgtgtt
6661 tttgcagagg gttattcctg aaactgactt gttactagga ccacatagtt tggtacacag
6721 tagaaaaaag gcagttagtt attgctgtag aatgaactga gtgacagttc ttgcttgctt
6781 gagaaaaata tactgagcat ctagattggg ccaggtagtg ttgctaggtg ctc^gggatat
(6834)(a/g)(dbSNP:1318417)
6841 atagtagaaa aacaagcctg tctttttttaa tgtatgcagg aaatctctgg agaagtaaga
6901 g^aggttaagg aaagctgg^gt atgt^gcctgc tttacatgga gtagtagtga ttaagttggc
SV 14 (6902)(a/g)(dbSNP:2077798), (6919)(a/g)(dbSNP:2077797),
(6925)(t/g)(dbSNP:2077799)
6961 cagtgcccag aagtgcttac tgggtgccag gattgttgct gtgg^aatatc gagtac^cact
SV 15 (7005) (a/c), (7017)(branch site - YNYYRAY)
LAMR1-Exon4(7041-7286 -> 246bp)
7021 aacttttaaa ttcttcaaag AGGGCTGTGC TGAAGTTTGC TGCTGCCACT GGAGCCACTC
7081 CAATTGCTGG CCGCTTCACT CCTGGAACCT TCACTAACCA GATCCAGGCA GCCTTCCGGG
7141 AGCCACGGCT TCTTGTGGTT ACTGACCCCA GGGCTGACCA CCAGCCTCTC ACGGAGGCAT
7201 CTTATGTTAA CCTACCTACC ATTGCGCTGT GTAACACAGA TTCTCCTCTG CGCTATGTGG
7261 ACATTGCCAT CCCATGCAAC AACAAGgtaa tgattttagg atctagagtt tgtgaatgcg
7321 tgctctagaa aaaacattcc tgtgcacatt gttagagctt ggagttgagg ctactgactg
7381 gccgatgaac tcgcaagtgt aggtagtgtg ctacatgagg ggcaagtttt cgctaacacc
7441 acaagggtct ctggcccaat gagtggagtt tgatagtaat tcttgctaca agtataacat
7501 tactgcatga cagctttgtg gagaaatgaa aacatttgga aaatagtgtg ttcttctgcc
7561 tttgtcc^atg tttcttgctc aggcctcagg cacttggcct ttgttttcac accaacttga
(7568)(a/g)(dbSNP:1392192)
7621 tgggttctac tataagatgc taaaaaggtg ggttgtgtgt ggttcagatt ggttgatttg
7681 taaccctagt tgtgcataag aattgcccag atatttttt taaatgccag tgtccaggca
7741 tttttttttc aatccctatg attccagtgt gcagacaagg ttgaaaatcc ctatttataa
7801 tgctccctgt tacaattgct tttcaactaa gagatgtctg tacttttggt ac^ctagatga
(7853)(c/t)(dbSNP:2269349)
7861 tgtaagagcc aggaaggtgc ttgctgtttg ggtttgacca agtgtcact^t tttaataatc
(7910)(branch site - YNYYRAY)

```
LAMR1-Exon5(7936-8064 -> 129bp)
  7921 tgccactctt ggcagGGAGC TCACTCAGTG GGTTT^GATGT GGTGGATGCT GGCTCGGGAA
       SV 16(7956)(a/g)(dbSNP:2269350)
  7981 GTTCTGCGCA TGCGTGGCAC CATTTCCCGT GAACACCCAT GGGAGGTCAT GCCTGATCTG
  8041 TACTTCTACA GAGATCCTGA AGAGgtaagc ttctccaaag gcttgtggtt acataagcaa
  8101 attggacgac ttggactgtg cttctaggaa gcaaaacttg tcagtccctg taagtctttc
LAMR1-Exon6(8183-8348 -> 166bp)
  8161 ctctttttt tt^tgtaaccc agATTGAAAA AGAAGAGCAG GCTGCTGCTG AGAAGGCAGT
       (8173)(branch site - YRYRAY)
  8221 GACCAA^GG^AG GAATTTCAGG GTGAATGGAC TGCTCCCGCT CCTGAGTTCA CTGCTACTCA
       (8227-8498)(STS:RP_SA_1), (8229-8474)(STS:G59752)
  8281 GCCTGAGGTT GCAGACTGGT CTGAAGGTGT ACAGGTGCCC TCTGTGCCTA TTCAGCAATT
  8341 CCCTACTGgt atgtatcagg atagaggtga atcaagctga tattttgca^a cttctcagtt
       (8390)(a/g)(dbSNP:2276714)
  8401 ttatt^ctaac tttaatgatc tctgtgactt ttatactagc tttaagaggt tttcattcca
       SV 17 (8406)(c/t)(dbSNP:2723)
  8461 gtgtgctaca gcatctgata gactgctgtt gg^gagtgggg taaggaaaaa tactacattg
       (8493)(t/g)(dbSNP:7635671)
LAMR1-Exon7(8561-8721 -> 161bp)
  8521 aggacagagc tgatggcttt ttttggtatt c^tcttaacag AAGACTGGAG ^C^GCTCAGCCT
       (8552)(branch site - YNYYRAY), (8571)(c/t)(dbSNP:2724),
(8572)(a/g)(dbSNP:2725)
  8581 GCCACGGAAG ACTGGTCTGC AGCTCCCACT GCTCAGGCCA CTGA^ATGGGT AGGAGCAACC
       (8625-8788)(STS:RH92628)
  8641 ACTGACTGGT CT^TAAGCTGT TCTTGCATAG GCTCTTAAGC AGCATGGAAA AATGGTTGAT
       stop
  8701 GGAAAATAAA CATCAGTTTC Taaaagttgt cttcatttag tttgcttttt actccagatc
  8761 agaatacctg ggattgcata tcaaagcata ataataaata catgtctcga catgagttgt
  8821 acttctaaag cccactgtag atagtgtata ttgcttttca cttcagaatt tccctgagct
  8881 tccttcctgt gtggagaaca tgtgtgctgt gaaaatagat gtaagtggtt acacatactt
  8941 agttggtaac ctcaggaata cttttcc^gag gccctgatga tcttgctttt attaccattt
       (8968)(c/g)(dbSNP:7638276)
  9001 gcgtaggtaa gttctataga gcttaggttt ttatttttga gagtcttcct gtcacttagg
  9061 ttaactattc ctggaagccc tggagaggcg atggtccttc cccacattga gagacattaa
  9121 ttgccccctct tgtcagcttg cctacaggtt ctgggagagg agaggtatgt atgtatttaa
  9181 aaatgtgatg ttctccatat ttgtctttgc cggggagggg gttggggaga cagtctttcc
  9241 ctgttgccca ggctggagta cagtggcgtg atctcagttc actgcaacct ctgcctcctg
  9301 ggttcaagcg attcttgtgc ctcaggctcc ttcgtagctg ggatttcagg tgtgcactgc
  9361 cacacctagc taatttttgt atttgtagta gagatggttc tgccaggctg gtctccaact
  9421 cctggcctca tgtgacccgc cccaaaatgc tgggattaca ggcctgagcc tgaccatatt
  9481 tgtttttttat gtattaatgc tggtggcttc atcctaaacg gatttcaaga ttttcattgc
  9541 ctattaaaaa aatactcgtg gctgtcatgt tggctcaccc atgatcccag cactttggga
  9601 ggcagaggca aagcgattgc ttgagctcag gagttcagga ccagcctggg caa^tgtggta
       (9654)(c/t)(dbSNP:13070037)
  9661 aaaccccatc tctctccaaa agatacaaaa aattggcagg gtgtgatggc atactcctgt
  9721 agtacctcag ctacacagga gactgaggtg ggagaatcac ctgaggccag gagttggagg
  9781 ctgcagtgag ctgtgattga gcccctgtat actctagcct gggtgacgaa gtgaggccct
  9841 gtctcaaaaa gattatagag acatttccat agattggatt gaatgcagat tttgagatgg
  9901 cctgaactga tttacatgga ttcatacatt ctcttggaaa atttacaaaa tacatgtaca
  9961 tgtgaccaca ttcctagggc tccttccggt ggtgctgagg caatgggaac actgca^gcct
       (10017)(a/g)(dbSNP:7641291)
 10021 caagcattgt cttcatggga agcctgctta tggctgtaag cgaggctggg aaaatctgcc
 10081 ctataatgtc tccaagttga atggcattca ^gtgagattta tggatcctga agttgcaagt
       (10111)(a/g)(dbSNP:12638378)
 10141 ggttagcaa^c actggtagag attgctggaa tcaggatggt gggaagcagg gacagtttac
       (10150)(c/t)(dbSNP:12638341)
 10201 cagaaggcta ctgg^tcaaca ggggagttgg agaagcattg cgaggtggtg tgagcctgga
       (10215)(c/t)(dbSNP:7633878)
 10261 gtaaggtata ctaagctcat ttatgagcgg gttctgggtg gagtctgttc ag^cgaacaga
       (10313)(c/t)(dbSNP:13095720)
 10321 gaaccttggc atttctggac acttgaagat actgtggaaa acatccgttg tgactatcca
```

```
 10381 gcaacctcca ggattacccc atcttac^tgg ggtcccatgc aaacccagcc tctccataga
       (10408)(c/t)(dbSNP:7634078)
 10441 cttgtgtcct agtgccaagc acctccaatc aggagacctt ggtagcattc tcaaagcagt
 10501 tccagaaagt ctctttttct agattctgag gttagggagg aggtagtctc ccttggcagg
 10561 gttcattcat gttcagctaa gtggaagcca gcaaatacta gaggggattt ctctcagacc
 10621 cactcagctg gggttagcct tcccctg^acag gcctiaacct ttttataacc ttgtatatct
       (10647)(a/g)(dbSNP:13096175)
 10681 gaaaattagga tcctctcccc aactccactc ccctggccct gtaccatatc accaactgct
 10741 agtggattct ggctcaggtt tgggatcacc atgctctcat gggacccagc agctgcccat
 10801 ggccacagag tcagtgtccc caagctgtca ccatctggaa tagccatctt aaactagcca
 10861 gttcaaacag ttttctcttg cctggtggtg cccacctgat gtagggcagc tcataagcac
 10921 tgctgaatca gtgacaatct gtgccccctg ggtgttcagc agcctgctac tagtaactta
 10981 ccccccagtgg tgtttgcaga attgagtctt cctgactcaa ttcctagc^at gaggccattt
       (11029)(a/g)(dbSNP:12491038)
 11041 ttaaggtgag ggagaaagccg tt^gcctcctt cactgctagg cactgctctt gctttggccc
       (11063)(c/g)(dbSNP:11720637)
 11101 tgtgagccgc ctcccttctt gccctggcac atctccattt ccttctacac aggaattgca
 11161 tttctctcag taagttattc ^aatttccttg ttcagc^ctga tcttacaaat gtagcttgct
       (11181)(a/t)(dbSNP:12491062), (11197)(c/g)(dbSNP:12497517)
 11221 ctaaatattt tacttgttca caacttcaag ctcttcatac acttttgggc agtttgtctt
 11281 ctatgcttaa tagaaactga atcccacaca gaactctagt tttcctctac tttcctaagt
 11341 agcttaggga ctcctttctt aggactgtgt gcaataaaca catttttct agtacattcc
 11401 tagtatgagg gcaaatctga gcatcctcat tggaatcagc ctatcacaaa agctggcact
 11461 tggtaaatgc tagtaagtac tgagtgaaaa gaaaaagtta agaataaatc tgagggccag
 11521 gtggctcatg cctgtaatct cagcactttg ggagaccgag gtgggcgggt cacctgagga
 11581 caggagtttg agaccaacct ggccaaaatg gagaaacccc atctctacta aaaatacaaa
 11641 a^aattagcca gg^cctggtgg tgtgtgcctg taatcccagc tactcaggaa gctgaggcag
       (11642)(a/c)(dbSNP:11129829), (11653)(a/c)(dbSNP:13059668)
 11701 gagaatcgct taaatctggg a
//
```