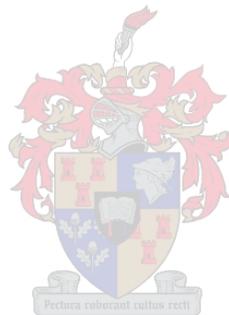


A candidate and novel gene search to identify the PFHBII-causative gene

Pedro Fernandez



Dissertation presented for approval for the degree Doctor of Philosophy at the University of
Stellenbosch

Promoter: Prof Valerie A Corfield

December 2004

“Declaration”

I, the undersigned, hereby declare that the work contained in this dissertation is my own original work and that I have not previously in its entirety or in part submitted it at a university for a degree

Signature:

Date:

Abstract

Heart failure due to cardiomyopathy or cardiac conduction disease is a major cause of mortality and morbidity in both developed and developing countries. Although defined as separate clinical entities, inherited forms of cardiomyopathies and cardiac conduction disorders have been identified that present with overlapping clinical features and/or have common molecular aetiologies.

The objective of the present study was to identify the molecular cause of progressive familial heart block type II (PFHBII), an inherited cardiac conduction disorder that segregates in a South African Caucasian Afrikaner family (Brink and Torrington, 1977). The availability of family data tracing the segregation of PFHBII meant that linkage analysis could be employed to identify the chromosomal location of the disease-causative gene. Human Genome Project (HGP) databases have provided additional resources to facilitate the identification of positional candidate genes.

Clinical examinations were performed on individuals of the PFHBII-affected family, and, where available, clinical records of subjects examined in a previous study by Brink and Torrington (1977) were re-assessed. Retrospective data suggested redefining the classification of PFHBII. Subsequently, linkage analysis was used to test described dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM) and cardiac conduction-causative loci on chromosomes 1, 2, 3, 6, 7, 9, 11, 14, 15 and 19 for their involvement in the development of PFHBII. Once a locus was mapped, bioinformatics tools were applied to identify and prioritise positional candidate genes for mutation screening.

The retrospective and prospective clinical study redefined PFHBII as a cardiac conduction and DCM-associated disorder and simultaneously allowed more family members to be traced.

Fortuitously, candidate loci linkage analysis mapped the PFHBII locus to chromosome 1q32, to a region that overlapped a previously described DCM-associated disorder (CMD1D), by the generation of a maximum pairwise lod score of 3.13 at *DIS3753* (theta [θ]=0.0) and a maximum multipoint lod score of 3.7 between *DIS3753* and *DIS414*. However, genetic fine mapping and haplotype analysis placed the PFHBII-causative locus distal to the CMD1D locus, within a 3.9 centimorgan (cM) interval on chromosome 1q32.2-q32.3, telomeric of *DIS70* and centromeric of *DIS505*. Bioinformatics analyses prioritised seven candidate genes for mutation analysis, namely, a gene encoding a potassium channel (*KCNH1*), an extracellular matrix protein (*LAMB3*), a protein phosphatase (*PPP2R5A*), an adapter protein that interacts with a cytoskeletal protein (*T3JAM*), a putative acyltransferase (*KIAA0205*) and two genes encoding proteins possibly involved in energy homeostasis (*RAMP* and *VWS59*). The PFHBII-causative mutation was not identified, although single sequence variations were identified in four of the seven candidate genes that were screened.

Although the molecular aetiology was not established, the present study defined the underlying involvement of DCM in the pathogenesis of PFHBII. The new clinical classification of PFHBII has been published (Fernandez et al., 2004) and should lead to tracing more affected individuals in South Africa or elsewhere. The identification of a novel disease-causative locus may point toward the future identification of a new DCM-associated aetiology, which, in turn, might provide insights towards understanding the associated molecular pathophysiologies of heart failure.

Opsomming

Hartversaking as gevolg van kardiomiopatie of kardiaal geleidingsiekte is 'n hoof-oorsaak van mortaliteit and morbiditeit in beide ontwikkelde en ontwikkelende lande. Alhoewel gedefinieer as verskillende kliniese entiteite is oorerflike vorms van kardiomiopatie en kardiaal geleidingsstoornisse geïdentifiseer met oorvleuelende kliniese eienskappe en/of molukulêre oorsake.

Die doelwit van hierdie studie was om die molukulêre oorsaak van progressiewe familiële hartblok tipe II (PFHBII), 'n oorerflike kardiaal geleidingsstoornis, wat in 'n Suid-Afrikaanse Kaukasiër familie segregeer (Brink en Torrington, 1977), te identifiseer. Die beskikbaarheid van familie data, beteken dat koppelingsanalise gebruik kan word om die chromosomale posisie van die siekte-veroorsakende geen te identifiseer. Menslike Genoom Projek (MGP) databanke het addisionele hulpbronne beskikbaar gestel om die identifikasie van posisionele kandidaat gene te vergemaklik.

Kliniese ondersoek is uitgevoer op PFHBII-geaffekteerde familieledes, en waar beskikbaar is kliniese rekords van persone, wat in 'n vorige studie deur Brink en Torrington (1977) geassesseer was, herontleed. Retrospektiewe data-analise het die kliniese herdefinisie van PFHBII voorgestel. Daarna is koppelingsanalise gebruik om dilateerde kardiomiopatie (DKM), hipertrofiese kardiomiopatie (HKM) en kardiaal geleidingsiekte-veroorsakende loki op chromosoom 1, 2, 3, 6, 7, 9, 11, 14, 15 en 19 te ondersoek vir hul moontlike bydrae tot die ontwikkeling van PFHBII. Toe die lokus gekarteer was, is bioinformatiese ondersoek gebruik om posisionele kandidaat gene te identifiseer en prioritiseer vir mutasie analise.

Die retrospektiewe en prospektiewe kliniese ondersoek het PFHBII herdefinieer as 'n geleidingsstoornis en DKM-verbode siekte, en terselfde tyd het dit gelei tot die opsporing

van nog familieleden. Toevallig het kandidaat loki-analise die PFHBII lokus op chromosoom 1q32 gekarteer, na 'n gebied wat met 'n voorheen-beskyfde DKM-verbode stoornis (CMD1D) ooreenstem, met die opwekking van 'n maksimum paargewyse lod-getal van 3.13 by *DIS3753* ($\theta = 0.0$) en 'n maksimum multipunt lod-getal van 3.7 tussen *DIS3753* en *DIS414*. Genetiese fynkartering en haplotipe-analise het die PFHBII-veroorsakende lokus afwaards van die CMD1D lokus geplaas, in 'n 3.9 centimorgan (cM) gebied op chromosoom 1q32.2-q32.3, telomeries van *DIS70* en sentromeries van *DIS505*. Bioinformatiese analise het daartoe gelei dat sewe kandidaat gene vir mutasie analise geprioritiseer is, naamlik, gene wat onderskeidelik 'n kalium kanaal (*KCNH1*), 'n ekstrasellulêre matriksproteïen (*LAMB3*), 'n proteïen fosfatase (*PPP2R5A*), 'n aansluitende proteïen wat met 'n sitoskilet proteïen bind (*T3JAM*), 'n asieltransferase (*KIAA0205*) en twee gene moontlik betrokke in energie homeostase (*RAMP* en *VWS59*) encodeer. Die PFHBII-veroorsakende geen is nie geïdentifiseer nie, alhoewel enkele volgorde-wisselings geïdentifiseer is in vier van die sewe geanaliseerde kandidaat gene.

Alhoewel die molekulêre oorsaak van die siekte nie vasgestel is nie, het die huidige studie die onderliggende betrokkenheid van DKM in die pathogenese van PFHBII gedefinieer. Die nuwe kliniese klassifikasie van PFHBII is gepubliseer (Fernandez et al., 2004) en sal lei tot die identifisering van nog geïmpakteerde persone in Suid Afrika of in ander lande. Die identifisering van 'n nuwe siekte-verbode lokus mag lei tot die toekomstige identifisering van 'n nuwe DKM-verbode genetiese oorsaak wat, opsig self, dalk insig kan gee in die molekulêre patofisiologie van hartversaking.

To my mother and late father.

You instilled in me the yearning to acquire knowledge.

A mind without instruction can no more bear fruit than can a field, however fertile, without cultivation.

Cicero

Acknowledgements

This dissertation would not be possible without the unselfish support and guidance from the following individuals:

Prof. Valerie Corfield, my promoter and mentor, you have trained me to always strive for excellence

Prof. Paul Brink, thank you for your invaluable clinical and statistical commentary

Dr. Johanna (Hanlie) Moolman-Smook, for always giving me words of encouragement

Althea Goosen, thank you for gathering the genealogical and clinical data

Past and present members of lab F445 who have helped me to maintain my sanity during the course of this study

Luigi and Sabrina, thank you for supporting me

Mario, we are genetic experiments in our own right, thank you for being academically competitive – you motivated me by setting such high standards

Last, but not least, to my incredible wife Mindy, thank you for your unyielding patience and understanding – you were my pillar of strength when I had doubts.

TABLE OF CONTENTS

Declaration	i
Abstract	ii
Opsomming	iv
Dedication	vi
Acknowledgements	vii
Index	viii
Abbreviations and acronyms	ix
Glossary	xi
List of figures	xvi
List of tables	xviii
Chapter 1: Introduction	1
Chapter 2: Materials and methods	53
Chapter 3: Results	85
Chapter 4: Discussion	116
Appendix A	143
Appendix B	149
Appendix C	151
Appendix D	153
Appendix E	159
Appendix F	160
Appendix G	164
References	165
Publications	205

Abbreviations and Acronyms

AV	atrioventricular
bpm	beats per minute
BrS	Brugada syndrome
BTHS	Barth syndrome
CCD	cardiac conduction defect
CHB	complete heart block
cM	centimorgan
°C	degrees celsius
DCM	dilated cardiomyopathy
EBV	Epstein-Barr Virus
ECG	electrocardiograph/electrocardiogram
EF	ejection fraction
HCM	hypertrophic cardiomyopathy
HGP	Human Genome Project
ICCD	isolated cardiac conduction disorder
IVF	idiopathic ventricular fibrillation
LAHB	left anterior hemiblock
LBBB	left bundle branch block
Lod	log of the odds ratio
LPHB	left posterior hemiblock
LQTS	long QT syndrome
LVEDD	left ventricular end diastolic diameter
MIM	Mendelian Inheritance in Man
PCCD	progressive cardiac conduction disease
PCR	polymerase Chain Reaction

PFHBI	progressive familial heart block type I
PFHBII	progressive familial heart block type II
RBBB	right bundle branch block
rpm	revolutions per minute
SB	sinus bradycardia
SSCP	single stranded conformation polymorphism
STACK	Sequence Tag Alignment and Consensus Knowledgebase
UTR	untranslated region
α	alpha
β	beta
δ	delta
γ	gamma
θ	theta
Ca^{2+}	calcium ion
K^{+}	potassium ion
M	molar
mg	milligram
ml	millilitre
mM	millimolar
Na^{+}	sodium ion
ng	nanogram
ρmol	picomole
μg	microgram
μl	microlitre
μM	micromolar
nm	nanometre

Glossary

ab initio - software that attempts to predict genes from sequence data without using prior knowledge about similarities to other genes.

Allele - alternative form of a genetic locus; a single allele for each locus is inherited from each parent (one or two more forms of a particular gene or marker).

Autosomal dominant - a gene (or disease allele) on one of the non-sex chromosomes that is expressed, even if only one copy is present. The chance of passing the gene (or disease allele) to offspring is 50% for each pregnancy.

Autosomal recessive - a trait (or a disease) that is produced only when two copies of a gene (or disease allele) are present.

Bioinformatics - the science of managing and analysing biological data using advanced computing techniques.

Blast (Basic Local Alignment Search Tool) - a fast heuristic search tool developed by Altschul, Gish, Miller, Myers and Lipman at the National Centre for Biotechnology Information (NCBI) that allows sequence search queries against the Genbank® database via the worldwide web (WWW). BLAST is able to detect relationships among sequences that share only isolated regions of similarity.

BLOCKS - an online database of multiply aligned ungapped segments corresponding to the most highly conserved regions of proteins documented in the PROSITE and Interpro database.

Centimorgan (cM) - a unit of genetic distance equivalent to a 1% recombination during meiosis. One centimorgan is equivalent to a physical distance of approximately one megabase (Mb) in a genome.

cDNA - complementary DNA- a piece of DNA copied *in vitro* from messenger ribonucleic acid (mRNA) by a reverse transcriptase enzyme.

Codon – the basic unit of the genetic code, comprising three-nucleotide sequences of mRNA, which specifies an amino acid or translational stop signal.

Contiguous - the result of joining overlapping collections of sequences or clones.

Domain - a discrete portion of a protein with its own function. The combination of domains in a single protein determines its overall function.

Draft sequence - the sequence produced by combining the information from the individual sequenced clones and positioning the sequence along the physical map of chromosomes.

E-value (expectation value) - the number of different alignments with scores equivalent to or better than a score expected to occur in a database search by chance. The lower the E-value, the more significant the score.

Expressed sequence tag (EST) – an expressed sequence tag is obtained by performing a single raw sequence read from the 3' or 5' end of a cDNA clone.

Finished sequence - high-quality, low error, gap-free DNA sequence of the human genome.

Freeze dates - snapshot of the most recent accession for each sequenced clone and ancillary data taken at a particular date.

Genetic map - a genome map in which polymorphic loci are positioned relative to one another on the basis of the frequency with which they recombine during meiosis. The unit of distance is a cM, where 1cM denotes a 1% chance of recombination.

Genotype - the set of genes that an individual carries; refers to the particular pair of alleles (alternative forms of a gene) that a person has at a given locus.

Haploinsufficiency - an individual who is heterozygous for a certain gene mutation or hemizygous at a particular locus, often due to a deletion of the corresponding allele, is clinically affected because a single copy of the normal gene is incapable of providing sufficient protein production as to assure normal function.

Haplotype - a particular combination of alleles or sequence variations that are closely linked or likely to be inherited together on chromosomes.

In silico - the term "*in silico*" has been introduced into life sciences and is similar to the terms "*in vivo*" (in the living system) and "*in vitro*" (in the test tube). It implies the gain of insights by theoretical considerations, simulations and experiments conducted in the silicon-based technology of a computer.

Linkage - the tendency of genes or other DNA sequences at specific loci to be inherited together as a consequence of their physical proximity on a single chromosome.

Lod score - a measure of the likelihood of genetic linkage between loci. A lod score greater than +3 is often taken as evidence of linkage; a score less than -2 is often taken as evidence against linkage, or exclusion.

Megabase (Mb) - unit of length for DNA fragments equal to 1 million nucleotides and roughly equal to 1 cM.

Motif - a region within a group of related protein or DNA sequences that is evolutionarily conserved, presumably due to its functional importance.

mRNA (messenger RNA) - RNA molecules synthesised from a DNA template (transcribed) which then serves as a template for the synthesis of a protein (translation).

NCBI (National Centre for Biotechnology Information) - the National Centre for Biotechnology Information (NCBI) is part of the National Library of Medicine (NLM). Its mission is to develop new information technologies to aid in the understanding of fundamental molecular and genetic processes that control health and disease. NCBI developed and maintains the Entrez Search System and PubMed database.

Pfam - an online database of protein families and multiple sequence alignments covering many common protein domains, created by ELL Sonnhammer, SR Eddy, E Birney, A Bateman and R Durbin.

Phenotype - the physical characteristics of a cell or organism as determined by the genetic constitution; also the specific physical characteristics of a disorder.

Physical map - a map of the locations of identifiable landmarks on DNA (e.g., restriction-enzyme cutting sites, genes). Distance is measured in base pairs. For the human genome, the lowest-resolution physical map is the banding patterns on the 24 different chromosomes; the highest-resolution map is the complete nucleotide sequence of the chromosomes.

PROSITE - an authoritative database of protein families and domains. It consists of biologically significant protein sites, patterns and profiles.

Single nucleotide polymorphism (SNP) – alternative alleles (a single base pair substitution, an insertion or deletion) present at a frequency of at least 1% in a population.

STACK - the Sequence Tag Alignment and Consensus Knowledgebase (STACK) is generated by processing EST and mRNA sequences obtained from Genbank through a pipeline consisting of masking, clustering, alignment and variation analysis steps. STACK database is organised into 15 tissue-based categories and one disease category.

Recombination - the process by which DNA is exchanged between homologous chromosome pairs during egg or sperm formation. Recombination has the effect of making the chromosomes of the offspring distinct from those of the parent.

Transcript map - a map of locations of gene (mRNA, cDNA or EST) sequences that code for a protein.

Transition - a type of nucleotide-pair substitution involving the replacement of a purine with another purine, or of a pyrimidine with another pyrimidine for example GC with AT.

Transversion - a type of nucleotide-pair substitution involving the replacement of a purine with a pyrimidine, or vice versa, for example GC with TA.

UniGene - an online database of non-redundant clustered Genbank and EST sequences for human, mouse, rat, cow, clawed frog and zebrafish. Each cluster contains sequences that

represent a unique gene, as well as related information such as tissue types in which the gene has been expressed and map location.

X-linked – a mutation in a gene on the X chromosome passed through a family, resulting in a specific trait or disease to be seen more commonly in males than females.

List of figures

Figure	Page	
1.1	The four-generation pedigree described by Brink and Torrington (1977) showing the autosomal dominant transmission of PFHBII	7
1.2	Human cytogenetic map indicating mapped DCM-causative loci and genes	12
1.3	The cellular localisation and interaction of the proteins constituents of the cardiomyocyte	19
1.4	Protein components of the sarcoplasmic reticulum (SR) involved in regulating Ca ²⁺ concentrations	22
1.5	Human cytogenetic map indicating mapped HCM-causative loci and genes	27
1.6	Anatomy of the cardiac conduction system	35
1.7	Human cytogenetic map indicating mapped cardiac conduction disease loci and genes	37
2.1	The South African PFHBII-affected kindred	56
2.2	Genetic markers and selected genomic clones that span the PFHBII locus on chromosome 1q32.2-q32.3	71
3.1	Clinical designations of members of the PFHBII-affected family, genotype assignments and haplotype construction with markers at chromosome 1q32.2-q32.3	92, 93
3.2	Representative autoradiograph of a genotype analysis in the family in which PFHBII segregates	94
3.3	Multipoint lod score curve at chromosome 1q32.2-q32.3	98
3.4	Representative agarose gel analysis of mapping an STR marker onto genomic clone-inserts	100

3.5	Integrated genetic, physical and transcript map of the PFHBII locus	101
3.6	Representative PCR-SSCP gel electrophoresis analysis of genes in which no mobility variations were identified	106
3.7	PCR-SSCP analysis of exon 11a of <i>KCNHI</i>	108
3.8	Sequence analysis of exon 11a of <i>KCNHI</i>	109
3.9	Sequence analysis of the 3' UTR of <i>KCNHI</i>	111
3.10	Sequencing analysis of exon 11b of <i>RAMP</i>	113
3.11	Protein sequence alignment analysis of L2DTL	114
4.1	The geographical distribution of members of the South African PFHBII-affected family	121
4.2	Tracing the origins of PFHBII by extensive pedigree analysis	128

List of tables

Table		Page
1.1	Timeline of familial DCM-associated disorders	9
1.2	Timeline of familial isolated cardiac conduction disorders	34
1.3	A representative list of disease-causative genes associated with cardiac disorders, channelopathies or skeletal myopathies that were identified using bioinformatics analysis	47
2.1	Primer sequences and annealing temperatures of selected STR markers used in a candidate loci linkage analysis of members of the PFHBII-affected family	62
2.2	Genomic clones used in the construction of a genetic, physical and transcript map of the PFHBII locus	70
3.1	Clinical characteristics of subjects from the PFHBII-affected family	87
3.2	Comparison of heterozygosity values for STR markers at chromosome 1q32.2-q32.3	94
3.3	Pairwise lod score analysis indicating exclusion of previously mapped causative loci	96
3.4	Pairwise lod scores between PFHBII and markers on chromosome 1q32.2-q32.3	97
3.5	Pairwise lod scores generated by changing the phenotypic status of individual III:11	97
3.6	Diseases caused by defects in genes mapped to within the PFHBII locus	103
3.7	Candidate genes selected for mutation screening	104
3.8	Summary of the sequence variations identified in the present study	115
4.1	Comparison of database entries showing changes to exon numbers of the candidate genes screened at the PFHBII locus	136

1.4.4.1 Progressive cardiac conduction defect	38
1.4.4.2 Cardiac conduction defect (CCD)	39
1.4.5 Mechanisms of cardiac conduction disorders	39
1.4.5.1 Alterations in ion channel electrophysiology	39
1.4.6 A synopsis of familial cardiac conduction disorders	41
1.4.7 Familial cardiomyopathies and cardiac conduction diseases in South Africa: contributions to global knowledge	42
1.4.7.1 Diseases of the myocardium	42
1.4.7.2 Diseases of the cardiac conduction system	43
1.4.8 The implication of establishing common clinical and genetic links between diseases	44
1.5 The HGP: background	44
1.5.1 Identifying disease-causative genes and the HGP	45
1.5.1.1 The pre-HGP era: 1970-1988	45
1.5.1.2 Successful applications of HGP resources	46
1.5.2 Limitations of the HGP: <i>caveat emptor</i>	47
1.5.2.1 Sequence errors	47
1.5.2.2 Gene annotation and prediction accuracy: “real” versus “virtual” genes	48
1.5.3 Current sequence and annotation status of the HGP	49
1.5.3.1 The status of complete human genome	49
1.5.3.2 The status of human chromosome 1	49
1.6 The present study	50

Introduction

I have long felt that biology ought to seem as exciting as a mystery story, for a mystery story is exactly what biology is.

- Richard Dawkins, "The Selfish Gene"

1.1 Preface

In 1950, Paul Wood stated, "Heart failure is a state in which the heart fails to maintain an adequate circulation for the needs of the body despite a satisfactory filling pressure". More than 50 years later, this condition is still recognised as the inability of the heart to function efficiently and is considered to be a major cause of morbidity and mortality (Braunwald, 1980; Eriksson, 1995; Giles, 1997; Towbin and Bowles, 2002). Estimates from the United States of America (USA) indicate that heart failure affects more than 60 million individuals in that country and, per annum, accounts for more than US\$200 million in associated hospitalisation costs (Manolio et al., 1992; Schönberger and Seidman, 2001; Seidman and Seidman, 2001; Towbin and Bowles, 2002). In Europe, it is estimated that more than 50 million individuals are affected by heart failure (Franz et al., 2001). However, in developing countries, the societal and economic cost of heart failure have not been estimated, although it has been proposed that changing lifestyles may cause similar patterns of incidence to that observed in the USA and Europe (Garros et al., 1980; World Health Organisation (WHO) website, <http://www.who.org>).

Heart failure results when the heart is no longer able to pump efficiently because of abnormal cardiac remodelling, usually in the form of hypertrophy of the myocardial wall or dilatation of the cardiac ventricles (Braunwald, 1980). Normally, the heart undergoes compensatory cardiac remodelling in response to coronary artery disease (myocardial infarction and ischaemia), valvular heart disease, alcohol and certain drugs, toxins, arrhythmias or cardiomyopathies (Braunwald, 1980). Occasionally the remodelling may become maladaptive and trigger events

that lead to complete cardiac malfunction (Towbin and Bowles, 2004). In the past, the primary focus of medical research was to identify effective means to prevent heart failure that had developed as a secondary consequence to pathological insults such as inflammation, infection, drugs, alcohol or toxins (Braunwald, 1980). However, in the 1990s, it was realised that the progression to heart failure was mostly associated with large families in which many members presented with primary forms of cardiomyopathy (Maron et al., 1995). Furthermore, inter- and intra-familial clinical variability was often demonstrated, which led to the suggestion that heart failure itself was a primary condition and that the phenotypic variation was a consequence of genetic heterogeneity (Maron et al., 1984; Graber et al., 1986; Solomon et al., 1990; Michels et al., 1992; Grünig et al., 1998).

The relevance of heart failure to the present study

The present study will demonstrate that progressive familial heart block type II (PFHBII) (Mendelian Inheritance in Man [MIM] accession number 140400) is a primary form of cardiomyopathy that is complicated by cardiac conduction defects that shows progression to heart failure. The cause of PFHBII is unknown, but it is hypothesised that the clinical and pathophysiological characteristics of other diseases that are associated with heart failure may provide clues to the likely molecular aetiology of the disorder. Accordingly, this literature review aims to give a clinical overview and to describe the molecular aetiologies that have been identified for two types of primary cardiomyopathy, namely, dilated cardiomyopathy (DCM) and hypertrophic cardiomyopathy (HCM), in addition to providing insight into the clinical and genetic characteristics of inherited cardiac conduction disorders. Common clinical features and molecular aetiologies between different clinical conditions will be reviewed, thereby providing further clues to the possible cause of PFHBII. As initiatives such as the Human Genome Project (HGP) have provided resources that have accelerated the identification of numerous disease-

causative aetiologies, the application of these resources will also be reviewed.

1.2 PFHBII

1.2.1 History and genealogy

In 1977, Brink and Torrington described PFHBII as an adult-onset cardiac conduction disorder of unknown aetiology displaying an *autosomal dominant* pattern of inheritance, which segregates in a South African Caucasian Afrikaner family. Genealogical studies performed by Brink and Torrington in 1977 (Fig. 1.1) had traced the segregation of the disorder in a family descended from a Dutch immigrant who arrived in South Africa in 1713 and who had settled in a region in the Eastern Cape Province seven years later.

For the present study, P. Brink and A. Goosen (personal communication) traced the siblings and progeny of subjects that had previously been assessed, or were interviewed by Brink and Torrington (1977) and Torrington (PhD thesis, 1979), some 25 years ago. These investigations could only identify the disease in the family described by Brink and Torrington (1977), and for this reason, the present study has focused on the same kindred, even though previous studies have supported the notion that the occurrence of PFHBII may be more widespread than merely in one kindred (see section 4.5).

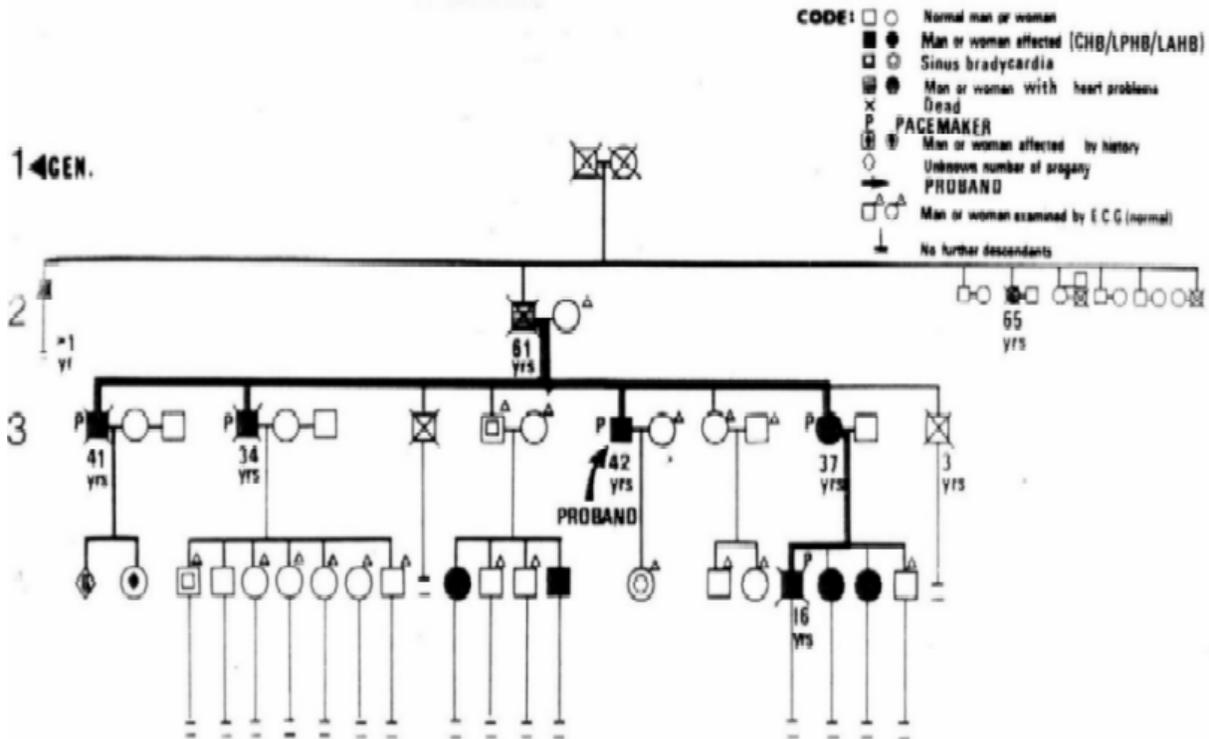


Fig. 1.1

The four-generation pedigree described by Brink and Torrington (1977) showing the autosomal dominant transmission of PFHBII.

The study by Brink and Torrington (1977) had identified 144 members of a family in which PFHBII segregates. Twenty-four of the family members were clinically assessed, of which 13 were identified with cardiac conduction defects.

1.2.2 Clinical characteristics

In their publication, Brink and Torrington (1977) defined the clinical features of PFHBII on electrocardiograph (ECG) analysis as isolated sinus bradycardia (SB), isolated left posterior hemiblock (LPHB) associated with syncopal episodes, Stokes-Adams seizures (light-headedness) or complete heart block (CHB) with narrow QRS complexes. Almost 10 years thereafter, A. Brink suggested that PFHBII was possibly associated with an inherited cardiomyopathy (Van der Merwe et al., 1986), although no data was supplied to substantiate this claim. Later, P. Brink (personal communication) indicated that he had identified individuals in the PFHBII-affected

family that demonstrated progression from conduction defects to DCM. This data promoted the possibility that cardiomyopathy was a primary feature of the disorder.

Presently, the molecular cause of PFHBII has not been established, although the availability of a number of affected individuals in a family has allowed the application of *linkage* analysis and positional cloning or positional candidate gene strategies to search for the disease-causative gene. Since inherited cardiac diseases with clinical features similar to PFHBII have been described globally and in South Africa, it is possible that this disorder has been described elsewhere, although by a different clinical designation. Therefore, the following sections will give an overview of the clinical features and molecular aetiologies of conditions akin to PFHBII, which may be associated with progression to heart failure.

1.3 Familial cardiomyopathies

1.3.1 Familial DCM: history and clinical classification

In 1961, Whitfield documented a family that presented with a “cardiomyopathy” that was, at the time, defined as a disease of the heart muscle for which the cause was unknown. Subsequently, Kariv et al., (1966) identified another family with a cardiomyopathy that was accompanied by Stokes-Adams seizures, while, four years later, Goodwin (1970) identified families that either presented with a congestive type of cardiomyopathy or a clinical form that was associated with hypertrophy (see section 1.3.5). The congestive form of cardiomyopathy that Goodwin (1970) had described was characterised by ventricular dilatation and contractile dysfunction, hence its subsequent classification as DCM (Gardner et al., 1987). Currently, numerous studies have described families with autosomal dominant, *autosomal recessive* and *X-linked* inherited forms of DCM that may occur in isolation or in conjunction with conduction defects, arrhythmias or skeletal myopathies (Table 1.1).

Table 1.1*Timeline of familial DCM-associated disorders*

Authors	Clinical features	Country where described	#Mode of inheritance
Emanuel (1972)	Isolated DCM	England	Recessive
Waxman et al., (1974)	Ventricular dilatation, conduction defects, ventricular tachycardia, atrial fibrillation and mitral insufficiency	Canada	Dominant
Brink et al., (1976)	Restrictive DCM and dysrhythmias	South Africa	Dominant
Ross et al., (1978)	DCM and arrhythmias	North America	Dominant
Sekiguchi et al., (1978)	DCM and conduction defects	Japan	Recessive
Moller et al., (1979)	DCM, septal hypertrophy, arrhythmias and conduction defects	Germany	Dominant
Marcus et al., (1982)	Right ventricular dysplasia, ventricular tachycardias, right-heart failure or asymptomatic cardiomegaly	France	Unknown
Barth et al., (1983)	DCM, neutropenia and mild skeletal myopathy	Netherlands	X-linked
Ibsen et al., (1985)	Right ventricular dilatation, ventricular arrhythmias and conduction defects	Netherlands	Unknown
Krasnow et al., (1985)	DCM, cataracts and hip-spine disease	North America	Recessive
Protonotarios et al., (1986)	Naxos syndrome characterised by right ventricular dysplasia, woolly hair and keratinosis	Greece	Recessive
Berko and Swift (1987)	DCM and mild skeletal myopathy, late progression to heart failure in females	North America	X-linked
Schmidt et al., (1988)	DCM, dysrhythmias, dyspnoea and exercise-induced sinus tachycardia.	North America	X-linked
Graber et al., (1986)	DCM and conduction defects	North America	Dominant
Kass et al., (1994)	DCM and conduction defects	North America	Dominant
Durand et al., (1995b)	DCM and conduction defects	North America	Dominant
Krajinovic et al., (1995)	Isolated DCM	Italy	Dominant
Rampazzo et al., (1995)	Exercised-induced ventricular tachycardias, right ventricular degeneration	Italy	Dominant
Olson and Keating (1996)	DCM, sinus node dysfunction, supraventricular tachyarrhythmia, conduction defects and stroke	North America	Dominant
Messina et al., (1997)	DCM, skeletal myopathy and conduction defects	North America	Dominant
Olson et al., (1998)	Isolated DCM	North America	Dominant
Li et al., (1999)	Isolated DCM	North America	Dominant
Kamisago et al., (2000)	Early onset DCM	North America	Dominant
Schönberger et al., (2000)	DCM and sensorineural hearing loss	North America	Dominant
Norgett et al., (2000)	DCM, woolly hair and keratoderma	Ecuador	Recessive
Olson et al., (2001)	Isolated DCM	North America	Dominant
Daehmlow et al., (2002)	Isolated DCM	Germany	Dominant
Hanson et al., (2002)	DCM and conduction defects	North America	Dominant
Murphy et al., (2004)	Isolated DCM	England	Recessive

= autosomal inheritance pattern, except for X-linked disorders; DCM = dilated cardiomyopathy

1.3.2 Molecular genetics of DCM

At the onset of the 1970s, numerous families with DCM-associated disorders that progressed to heart failure had already been described (Table 1.1), although the prevalence of the condition in those lineages had not been determined. However, in 1978, Sekiguchi and colleagues reported that 31% of patients with DCM in the Japanese population had a family history of heart failure. The subsequent decade-and-a-half produced a number of reports describing families that presented with DCM either in isolation or with accompanying features (Table 1.1). Prior to the report by Sekiguchi et al., (1978), DCM-associated disorders had mainly been portrayed as conditions that occurred as a result of environmental factors or lifestyle, as described by Braunwald in “*A Textbook of Cardiovascular Medicine*” (Braunwald, 1980). Later, Michels and colleagues (1992) initiated a study to establish whether familial forms of DCM were as frequent as conditions caused by environmental factors. The study demonstrated that at least 20% of the individuals with DCM-associated defects had a familial form of the disease (Michels et al., 1992). As numerous reports accumulated describing families with DCM in which the disease-associated loci had been identified (see sections hereafter), Grünig et al., (1998) re-evaluated the previous data on the prevalence of inherited DCM. These authors subsequently suggested that as many as 35% of patients with DCM had a familial form of the disease, of which autosomal dominant inheritance was the most common pattern.

1.3.2.1 Molecular aetiologies of autosomal dominant DCM

1.3.2.1.1 Disease loci and causative genes associated with isolated DCM

Currently, causative mutations associated with isolated forms of DCM have been described in 11 genes, namely, the gene encoding cardiac troponin T (*TNNT2*) mapped on chromosome 1q32 (Durand et al., 1995b; Kamisago et al., 2000; Li et al., 2001) (MIM 601494), titin (*TTN*) on chromosome 2q31 (Siu et al., 1999; Gerull et al., 2002) (MIM 604145), desmin (*DES*) on

chromosome 2q35 (Li et al., 1999) (MIM 604765), delta (δ)-sarcoglycan (*SGCD*) on chromosome 5q33 (Tsubata et al., 2000) (MIM 606685), phospholamban (*PLB*) on chromosome 6q22 (Schmitt et al., 2003) (MIM 172405), metavinculin (*VCL*) on chromosome 10q22-q23 (Bowles et al., 1996; Olson et al., 2002) (MIM 193065), cardiac myosin-binding protein C (*MYBPC3*) on chromosome 11p11 (Daehmlow et al., 2002) (MIM 115197), cardiac muscle LIM protein (*CSRP3*) on chromosome 11p15 (Knöll et al., 2002) (MIM 607482), beta (β)-myosin heavy chain (*MYH7*) on chromosome 14q12 (Kamisago et al., 2000) (MIM 115200), cardiac actin (*ACTC*) on chromosome 15q14 (Olson et al., 1998) (MIM 115200) and alpha (α)-tropomyosin (*TPMI*) on chromosome 15q22 (Olson et al., 2001) (MIM 115196) (Fig. 1.2). Additionally, two isolated DCM disorders have been mapped to loci on chromosome 6q12-q16 (*CMD1K*) (Sylvius et al., 2001) (MIM 605582) and chromosome 9q12-q13 (*CMD1B*) (Krajinovic et al., 1995) (MIM 600884) (Fig. 1.2), although their associated molecular pathophysiology has yet to be identified.

1.3.2.1.2 Disease loci and causative genes associated with DCM and conduction defects

Various studies have described families in which cardiac conduction abnormalities accompany DCM (Sekiguchi et al., 1978; Graber et al., 1986; Kass et al., 1994; Durand et al., 1995b; Hanson et al., 2002) (Table 1.1). Presently, causative mutations associated with disorders that present with DCM and cardiac conduction defects have been identified in two genes, namely the gene encoding lamin A and C (*LMNA*) on chromosome 1p1-1q21 (Kass et al., 1994; Fatkin et al., 1999) (MIM 115200) and *TNNT2* (Hanson et al., 2002) (Fig. 1.2). Two additional loci have been described, for which the molecular aetiology has not been identified. These loci map to chromosome 2q14-q22 (*CMD1H*) (Jung et al., 1999) (MIM 604288) and chromosome 3p21-p25 (*CMD1E*) (Olson and Keating, 1996) (MIM 601154) (Fig. 1.2).

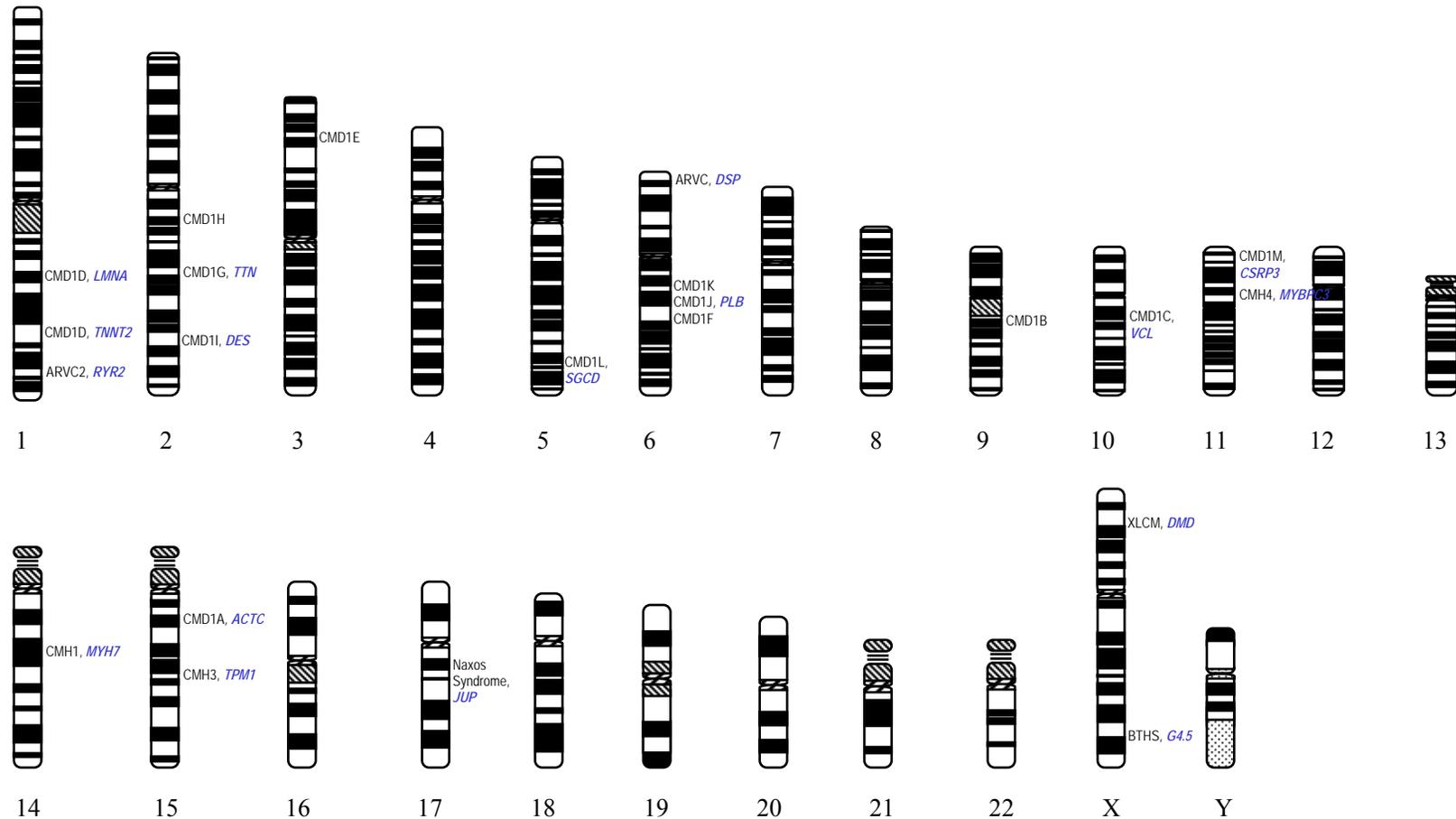


Fig. 1.2

Human cytogenetic map indicating mapped DCM-causative loci and genes.

ACTC = cardiac actin; ARVC = arrhythmogenic right ventricular cardiomyopathy; BTHS = Barth syndrome; CMD = dilated cardiomyopathy locus; CMH = hypertrophic cardiomyopathy locus; CSRP3 = muscle LIM protein; DES = desmin; DMD = dystrophin; DSP = desmoplakin; G4.5 = tafazzin; JUP = plakoglobin; LMNA = lamin A and C; MYBPC3 = cardiac myosin-binding protein C; MYH7 = β -myosin heavy chain; PLB = phospholamban; RYR2 = ryanodine receptor type 2; SGCD = δ -sarcoglycan; TNNT2 = cardiac troponin T; TPM1 = α -tropomyosin; TTN = titin; VCL = metavinculin; XLCM = X-linked DCM. Disease-causative genes are shown in blue.

1.3.2.1.3 Disease loci and causative mutations associated with DCM, skeletal myopathies and conduction defects

Syndromes that are associated with skeletal myopathies and DCM are fairly common, although these have been mainly associated with X-linked conditions (see section 1.3.2.3). Conversely, autosomal dominantly inherited DCM-associated disorders that present with skeletal myopathies are less prevalent. In 1997, Messina et al., described mapping a DCM-causative locus that is associated with skeletal myopathies accompanied by conduction defects to chromosome 6q23 (CMD1F) (MIM 602067) (Fig. 1.2). Currently, there are no other reports of families with autosomal dominantly inherited primary DCM that present with skeletal myopathies and conduction defects. However, families with Kearns-Sayre syndrome (Kearns, 1965; Drachman, 1975) (MIM 530000), a neuromuscular disease caused by deletion mutations in muscle mitochondria, which may result in chronic, progressive or external ophthalmoplegia (eye muscle paralysis), hearing loss, pigment degeneration of the retina, conduction defects and DCM, have been reported.

1.3.2.1.4 Disease loci and causative genes associated with DCM and sensorineural hearing loss

Schönberger et al., (2000) characterised a family with a rare form of DCM that was accompanied by sensorineural hearing loss, for which they mapped the causative locus to a region on chromosome 6q23-q24 (CMD1J) (MIM 605362) (Fig. 1.2). The causative gene for this disorder remains unknown and no additional autosomal dominant DCM-associated disorders with abnormal auditory function have been described, although, sensorineural deafness may be a clinical feature that is associated with Kearns-Sayre syndrome (Kearns, 1965) (see previous section).

1.3.2.1.5 Disease loci associated with cardiac arrhythmias

Arrhythmogenic right ventricular dysplasia (ARVD) (MIM 107970) (reclassified as arrhythmogenic right ventricular cardiomyopathy [ARVC] in 1996 by the WHO) is a disorder characterised by supraventricular or ventricular arrhythmias, conduction defects, right ventricular dysplasia or dilatation leading to ventricular degeneration brought about by fibro-fatty deposits in the ventricular free wall (Marcus et al., 1982). By 1985, a number of families and sporadic cases of individuals with ARVC had been described (Marcus et al., 1982; Ibsen et al., 1985) (Table 1.1), even though no family studies had been undertaken at the time, to elucidate the genetic cause of the condition. About 10 years later, Rampazzo et al., (1995) described a family that presented with an ARVC-like disorder (termed ARVC2) (MIM 600996) that was characterised by exercise-induced polymorphic ventricular tachycardias and fatty-fibrous replacement of the right ventricle, without notable structural changes to the heart. The same authors also reported mapping the causative gene for ARVC2 to a locus on chromosome 1q42-q43 (Rampazzo et al., 1995) (Fig. 1.2). Subsequently, Tiso et al., (2001) described a positional candidate gene analysis that resulted in the identification of four ARVC2-causative mutations in four unrelated families in the gene that encodes the cardiac ryanodine receptor type 2 (*RYR2*) (Fig. 1.2).

1.3.2.2 Molecular aetiologies of autosomal recessive DCM

Autosomal recessive inheritance has been reported in 16% of all described families that present with DCM (Towbin and Bowles, 2001). One of the earliest reports of autosomal recessive isolated DCM was by Emmanuel (1972). Thereafter, families were identified in which autosomal recessive DCM occurred in isolation (Krasnow et al., 1985) or was accompanied by features such as cataracts, joint disease, woolly hair, keratinosis or keratoderma (Protonotarios et al., 1986). The availability of families with recessive forms of DCM, has allowed the application of linkage analyses to identify the causative genes for these disorders.

To this end, Coonar and colleagues (1998) described the results of a linkage study in consanguineous Greek families with Naxos syndrome (Protonotarios et al., 1986) (MIM 601214) (Table 1.1), which placed the causative gene at a locus on chromosome 17q21 (Fig. 1.2). Two years later, McKoy et al., (2000) reported that they had performed positional candidate gene analyses at the Naxos syndrome locus and had identified a homozygous deletion mutation in the plakoglobin protein-encoding gene (*JUP*) (Fig. 1.2). Soon thereafter, Norgett et al., (2000) reported that they had identified a mutation in the gene that encodes desmoplakin (*DSP*) (MIM 605676) (Fig. 1.2), which caused DCM, woolly hair and keratoderma in three families of Ecuadorian descent.

More recently, Murphy et al., (2004) reported that they had identified an autosomal recessive DCM-causative mutation in the gene that encodes cardiac troponin I (*TNNI3*) (MIM 115210). Although only two subjects were identified with the mutation in troponin I, these authors suggested that other recessive disease genes might be identified using techniques specifically designed to identify homozygous sequence variations (Murphy et al., 2004).

1.3.2.3 Molecular aetiologies of X-linked DCM

X-linked cardiomyopathies are common conditions that affect approximately one in 3500 males (Fatkin and Graham, 2002). Barth et al., (1983) described a family with a syndromic form of cardiomyopathy that presented with DCM and mild skeletal myopathy (Table 1.1). Four years later, Berko and Swift (1987) described another family in which teenage male subjects presented with DCM and mild skeletal myopathy and older female subjects presented with late-onset atypical chest pain and progression to heart failure. This condition was subsequently termed X-linked DCM (XLCM) (MIM 302045). The following year, Schmidt et al., (1988) described a

third family with an X-linked form of DCM that presented with DCM, dysrhythmias, dyspnoea and exercise-induced sinus tachycardia. Genetic studies in these families with the syndromic form of DCM and X-linked type resulted in mapping the associated disease loci, as well as identifying the molecular aetiologies of these conditions (see below).

1.3.2.3.1 Barth syndrome

Bolhuis and colleagues (1991) had mapped a condition, which was termed Barth syndrome (BTHS) (MIM 302060) because Barth et al., (1983) had first described the disease, to a locus on chromosome Xq28 (Fig. 1.2). Later, Bione et al., (1996) reported that they had performed a positional candidate gene analysis at the locus on chromosome Xq28 and had identified two BTHS-causative mutations in one family in the tafazzin protein-encoding gene (*G4.5*). Presently, this is the only report of cardiomyopathy-associated mutations in *G4.5*.

1.3.2.3.2 XLCM

Towbin and colleagues (1993) described a linkage study that had been performed in the North American family previously described by Berko and Swift (1987), which resulted in mapping XLCM to a locus on chromosome Xp21.1 (Fig. 1.2). Five years prior to the study by Towbin and colleagues (1993), deletion mutations that cause two forms of muscular dystrophy, namely Duchenne (Hoffman et al., 1987) and Becker muscular dystrophy (Koenig et al., 1988) were identified in the dystrophin protein-encoding gene (*DMD*). Because cardiomyopathy was a presenting feature of Duchenne's and Becker's muscular dystrophy, Muntoni and colleagues (1993) screened *DMD* in an Italian family with XLCM, which resulted in these authors describing a disease-causative deletion mutation in the promoter region of the gene. Later, in 1998, Ferlini and colleagues reported that they had identified another mutation in *DMD* that

activated a cryptic splice site in intron 11 and was associated with XLCM in another Italian family.

1.3.3 Mechanisms of familial DCM

The identification of the molecular aetiologies of various DCM-related disorders have provided clues that aided in elucidating the pathophysiological mechanisms of the disease. Moreover, DCM-causing mutations were identified in genes that encode, among others, structural protein constituents of cardiomyocytes (cardiac muscle cells), hence the formulation of hypotheses that explain the pathological pathways associated with heart failure.

1.3.3.1 Alterations in cardiomyocyte structural proteins

1.3.3.1.1 Structure of the cardiomyocyte

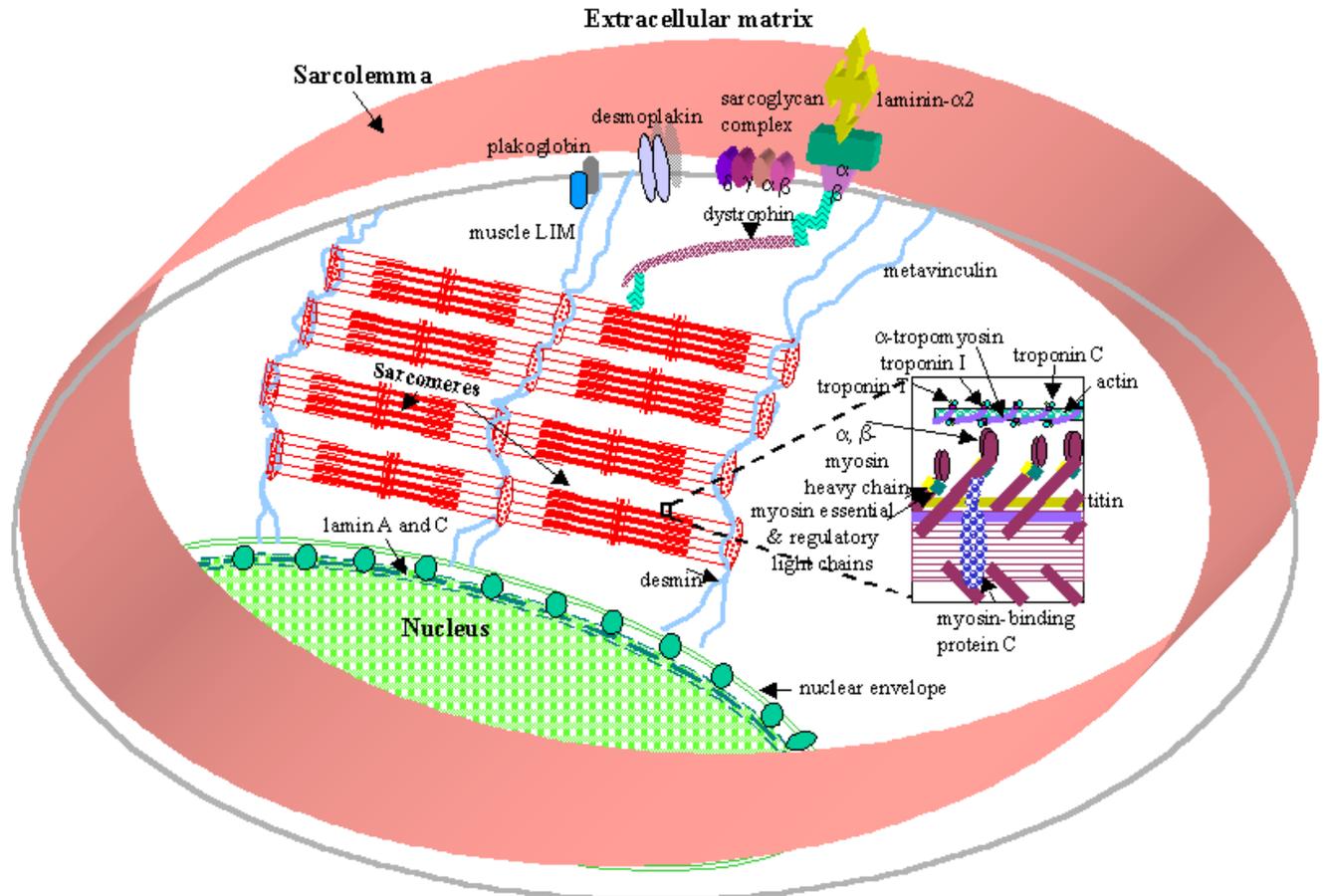
In order to illustrate the mechanisms of DCM that occur as a consequence of alterations in cardiomyocyte structure, the architecture of the muscle cell first has to be understood. Cardiomyocytes consist of cytoskeletal and filamentous proteins that maintain the structural integrity of the cell, and sarcomeric proteins, which constitute the functional unit of muscle contraction (Jaenicke et al., 1990) (Fig. 1.3). A sarcolemmal membrane consisting of anchoring and channel proteins, which allow cell-to-cell contact and the influx/efflux of nutrients and ions, encloses the cardiomyocyte (Fig. 1.3). The nucleus of the cardiomyocyte is embedded within the cytosol of the cell and is surrounded by a nuclear envelope that consists of structural and channel proteins (Fig. 1.3). It is important to remember that this is a rudimentary description of the structure of the cardiomyocyte, as there are a myriad of other proteins that interact with the constituents described above. The present study will focus on the cardiomyocytic components in which mutations have been identified that cause familial forms of cardiomyopathy.

1.3.3.1.2 Structural alterations and defective force transmission

Physiological functioning of the heart requires that the contractile force that is generated by the sarcomere be transmitted via the cytoskeleton to the sarcolemmal membrane and to adjacent myocytes (Fung et al., 1995). In 1993, Muntoni et al., described an XLCM-causative deletion mutation in dystrophin, a cytoskeletal protein that is attached to the sarcolemma (Fig. 1.3). Fortuitously, this finding was later to provide clues to the potential pathological mechanisms of DCM. Thus, when Olson et al., (1998) identified two DCM-causative mutations in cardiac actin that were located in amino acid residues involved in actin-cytoskeletal interactions, correlations were drawn between these defects and that described in dystrophin, which led to the suggestion that DCM occurred as a consequence of alterations in cytoskeletal integrity and reduced force transmission. For this reason, in 1998, Towbin suggested that DCM be classified as a “cytoskeletalopathy”. Subsequent studies describing DCM-causing mutations in the cytoskeletal proteins desmin (Li et al., 1999), metavinculin (Olson et al., 2002) and cardiac muscle LIM protein (Knöll et al., 2002) provided further evidence to support Towbin’s (1998) description of DCM as a cytoskeletalopathy.

At about the same time, four other studies had described DCM-associated mutations in genes that encode the membrane-associated proteins lamin A and C (Fatkin et al., 1999), desmoplakin (Norgett et al., 2000), δ -sarcoglycan (Tsubata et al., 2000) and plakoglobin (McKoy et al., 2001). The lamin A and C proteins stabilise the nuclear envelope (Lin and Worman, 1993, Hutchison, 2002), while desmoplakin, δ -sarcoglycan and plakoglobin are attached to the sarcolemma and form cell-to-cell junctions (Norgett et al., 2000; Tsubata et al., 2000). Although the lamin A and C, desmoplakin, δ -sarcoglycan and plakoglobin proteins are membrane bound, they are attached to other cytoskeletal proteins that, in turn, are linked to the sarcomere (Fig. 1.3). Therefore, it was

suggested that the DCM-causing mutations in these proteins also reduce cardiac force transmission (Fatkin et al., 1999; Tsubata et al., 2000).



Reproduced and adapted from Moolman-Smook et al., (2003)

Fig. 1.3

The cellular localisation and interaction of the proteins constituents of the cardiomyocyte.

Schematic representation of a cross-section through a cardiomyocyte indicating the cellular location and interactions of proteins in which mutations have been identified that cause DCM and HCM (see text).

1.3.3.1.3 Structural alterations and defective force generation

The identification of DCM-causing defects in proteins involved in contractile force transmission set a precedent for screening genes encoding cytoskeletal proteins, although most of these studies were unsuccessful. However, in 2000, Kamisago and colleagues reported mapping a DCM-causative locus to chromosome 14q12.2-q11.2 and concomitantly described mutations in the sarcomeric β -myosin heavy chain protein, which, had previously been implicated in the pathogenesis of HCM (Geisterfer-Lowrance et al., 1990). The same group then performed candidate gene analyses of other sarcomeric protein-encoding genes in which mutations had been shown to cause HCM and subsequently described a DCM-associated mutation in cardiac troponin T (Kamisago et al., 2000)

The finding by Kamisago et al., (2000) was important because mutations in β -myosin heavy chain and troponin T had been shown to cause HCM by altering the stoichiometry of the sarcomeric proteins, thereby causing a reduction in cardiac contractile force (Thierfelder et al., 1994; Watkins et al., 1996). Thus, based on the functional *domains* in which the particular DCM-causing mutations were identified, Kamisago and colleagues (2000) suggested that defects in β -myosin heavy chain and troponin T altered stereospecific myosin-actin interactions and actin-myosin ATPase activity, respectively, which in both instances, resulted in DCM by reducing the overall generation of contractile force.

Additionally, in the same study, Kamisago et al., (2000) proposed that at least 10% of all cases of familial DCM were likely to be attributed to defects in genes that encode sarcomeric proteins. Consequently, other studies described DCM-associated mutations in the genes encoding the sarcomeric proteins α -tropomyosin (Olson et al., 2001), cardiac myosin-binding protein C

(Daehmlow et al., 2002) and titin (Gerull et al., 2002). Incidentally, separate mutations in these sarcomeric proteins had also been implicated as cause of HCM (Thierfelder et al., 1994, Watkins et al., 1995; Moolman-Smook et al., 1998; Satoh et al., 1999). Therefore, it was suggested that reduced contractile force was also a primary cause of DCM (Schönberger and Seidman, 2001; Seidman and Seidman, 2001).

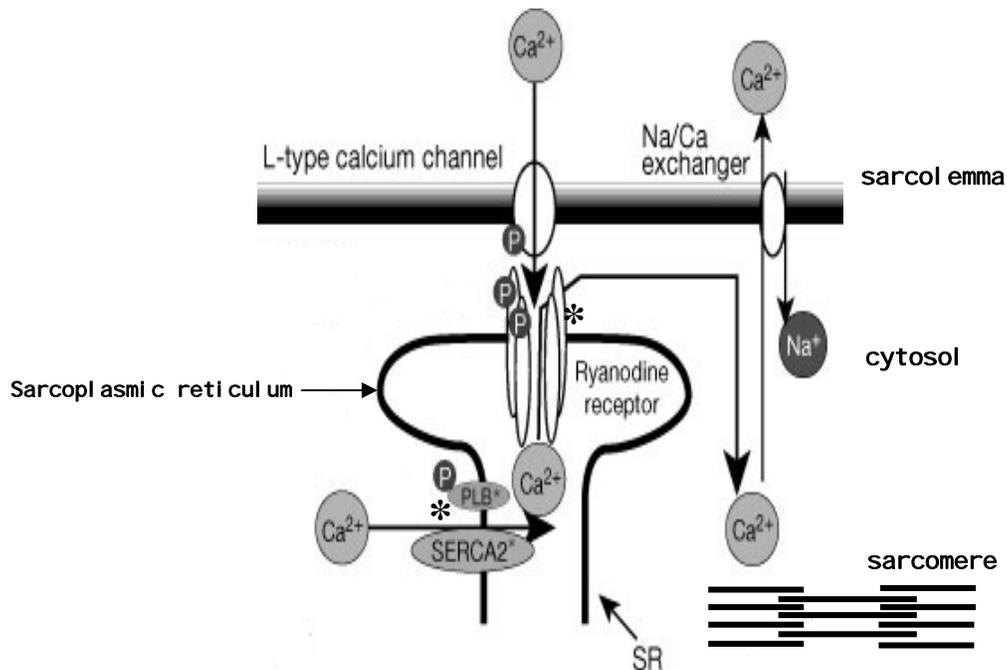
1.3.3.2 Alterations in cardiomyocyte calcium (Ca^{2+}) homeostasis

Given the multitude of proteins that participate in cardiac contractile force generation and force transmission, it would be expected that more DCM-causative mutations would have been identified in genes whose protein products are components of the sarcomere or cytoskeleton. However, to date, no defects have been described in other sarcomeric or cytoskeletal proteins, despite many extensive investigations. The lack of success suggested that additional pathophysiological mechanisms might be involved in the pathogenesis of DCM.

The first indication of other DCM-associated disease mechanisms developed as a consequence of the identification of ARVC2-causative defects in the cardiac ryanodine receptor 2 (Tiso et al., 2001). The precise pathological mechanism that causes ARVC has remained inconclusive, although the identification of mutations in this receptor, as well as the recognition that right cardiac ventricular degeneration was a defining characteristic of ARVC, provided possible clues.

The ryanodine receptor regulates the release of Ca^{2+} from the sarcoplasmic reticulum (SR) (Fig. 1.4) and it has been suggested that mutations in this receptor may increase the release of Ca^{2+} from the SR (Fatkin and Graham, 2002). The consequent elevated cytosolic Ca^{2+} levels are thought to disrupt excitation-contraction coupling and thereby trigger arrhythmias or promote cell

death, the latter being suggested as a feasible explanation for the degeneration of the right ventricle that is frequently observed in ARVC (Fatkin and Graham, 2002).



Reproduced and adapted from Chien (2000)

Fig. 1.4

Protein components of the sarcoplasmic reticulum (SR) involved in regulating Ca^{2+} concentrations.

During cardiac contractile force generation, Ca^{2+} enters into the cardiomyocyte via L-type Ca^{2+} channels and $\text{Na}^+/\text{Ca}^{2+}$ exchangers and triggers the ryanodine receptors to release Ca^{2+} into the cytosol, where it binds to the sarcomeric proteins that are involved in contractile force generation. During cardiac relaxation, the Ca^{2+} -ATPase (SERCA2) pump, which are modulated by the phosphoprotein phospholamban (PLB), remove Ca^{2+} from the cytosol and restores sarcoplasmic reticulum basal levels. An asterisk denotes a protein in which mutations have been shown to cause DCM.

In addition, a missense mutation that causes DCM and rapid progression to heart failure was recently described in phospholamban (Schmitt et al., 2003), a transmembrane protein that regulates the activity of a cardiac Ca^{2+} -ATPase (also called SERCA2), which, in turn, is responsible for the re-uptake of Ca^{2+} from the cytosol (McTiernan et al., 1999) (Fig. 1.4). Interestingly, decreased SERCA2 activity has been associated with progression to heart failure (Struder et al., 1994), although no causative mutations have yet been described in this protein. Therefore, Schmitt et al., (2003) proposed that a mutation in phospholamban caused DCM and acute heart failure because of a disturbance in Ca^{2+} homeostasis, which was a direct result of the constitutive inhibition of SERCA2.

1.3.3.3 Alterations in energy homeostasis

The molecular aetiologies and underlying pathophysiologies of autosomal dominant, recessive and X-linked inherited disorders, in which DCM is a primary feature, have, in fact, only been established in the last decade. However, Ozawa et al., (1990) had previously demonstrated that mitochondrial defects caused sporadic forms of HCM and DCM. In the same year, Tanaka et al., (1990) identified a mutation in a mitochondrial protein-encoding gene (*tRNA-ILE*) that causes fatal infantile cardiomyopathy. The mutation was thought to affect cellular energy generation and metabolism, given the role of mitochondria in the electron transport system and its production, via oxidative phosphorylation, of the energy substrate adenosine triphosphate (ATP) (Tanaka et al., 1990). The five years that followed produced two studies that described mutations in *tRNA-ILE* that encodes a mitochondrial protein (Silvestri et al., 1994; Casali et al., 1995) and which were associated with cardiomyopathy and skeletal myopathy. However, none of the studies implicating mitochondrial defects had established whether the mutations were the cause or the consequence of the cardiomyopathy, given the fact that the mitochondrial genome is prone to spontaneous mutations related to physiological ageing (Melov et al., 1995).

The identification of BTHS-causative mutations in the enzyme tafazzin (Bione et al., 1996) re-established the hypothesis that deficits in energy homeostasis caused DCM-associated disorders. Defects in tafazzin, which belongs to a protein family known as acyltransferases, are thought to cause cardiomyopathy because of the enzyme's inability to catalyse the production of specific glycerophospholipids (Bissler et al., 2002). The glycerophospholipids form part of the inner mitochondrial membrane and play a role in respiratory-chain transport and oxidative phosphorylation required to produce ATP (Neuwald, 1997). The data indicated that mutations in tafazzin cause BTHS as a result of a reduction in the production of energy (Bissler et al., 2002).

1.3.4 A synopsis of familial DCM

Familial DCM is a clinically variable and genetically heterogeneous condition, which may account for up to 35% of seemingly idiopathic forms of DCM. The causative genes for a number of familial DCM disorders have been identified, although, for some, the associated molecular aetiologies have yet to be elucidated. However, the identification of common genetic defects has suggested that the underlying causes of DCM are alterations in cardiomyocyte structure or Ca^{2+} and/or energy homeostasis. In addition, DCM shares common molecular aetiologies with other cardiomyopathies that are associated with progression to heart failure, notably ARVC and HCM (also see sections hereafter). Consequently, the accumulation of knowledge might provide clues for the elucidation of the molecular aetiologies of disorders for which, presently, only disease causative loci have been identified.

1.3.5 Familial HCM: history and clinical classification

The first description of a familial form of a cardiomyopathy was by Teare (1958), who described six families in which young adults presented with asymmetric thickening (hypertrophy) of the myocardial wall. At the time, the cardiac hypertrophy was not considered a disease, but rather a physiological adaptation of the heart toward mechanical stress and reduced cardiac output (Teare, 1958). Subsequently, Nasser et al., (1967) and Maron et al., (1974) identified additional families that presented with asymmetric or obstructive forms of cardiac hypertrophy. Then, in a landmark study by Maron and Epstein (1979), the term “hypertrophic cardiomyopathy” was suggested as the preferential nomenclature of families with asymmetric septal hypertrophy (ASH) – this study pioneered the present classification of different forms of the cardiac condition under one clinical term known as HCM.

Histological and morphological studies have further defined HCM as a complex disease of the myocardium that is characterised by myofibrillar disarray, interstitial fibrosis, left ventricular hypertrophy and/or asymmetric interventricular septal hypertrophy (Maron, 1984; Maron et al., 1995). The underlying clinical feature of HCM, in the absence of hypertension or aortic stenosis, was demonstrated to be an impairment of cardiac relaxation (diastole) (Maron et al., 1995; Maron, 1997). The condition exhibits a variety of associated clinical features that range from mild symptoms such as shortness of breath (dyspnoea) and chest pains (angina), to more severe features, such as atrial or ventricular arrhythmias and sudden death, which may occur with or without the presence of left ventricular or interventricular septal hypertrophy (Maron, 1995; Maron, 1997). Furthermore, a number of studies had mapped disease loci and had identified HCM-causative mutations in sarcomeric proteins (Jarcho et al., 1989; Geisterfer-Lowrance et al., 1990; Tanigawa et al., 1990; Carrier et al., 1993; Watkins et al., 1993; Thierfelder et al., 1994; Watkins et al., 1995). Consequently, the molecular genetic data demonstrated that the clinical

variability of HCM was due to genetic locus heterogeneity (Maron, 1997; Marian and Roberts, 2001).

1.3.6 Molecular genetics of familial HCM

1.3.6.1 Molecular aetiologies of autosomal dominant HCM

Dominant inheritance patterns occur in approximately 50% of familial HCM (Maron et al., 1984). Presently, HCM-causative mutations have been described in 13 genes, namely *TNNT2* on chromosomes 1q32 (Watkins et al., 1993; Thierfelder et al., 1994; Moolman et al., 1997) (MIM 191045), *TTN* on chromosomes 2q31 (Sato et al., 1999) (MIM 188840), the regulatory myosin light chain (*MYL3*) gene on chromosome 3p21 (Poetter et al., 1996) (MIM 160790), the troponin C (*TNNC1*) gene on chromosome 3p14-p21 (Hoffmann et al., 2001) (MIM 191040), the gene encoding the alpha subunit of adenosine monophosphate (AMP)-activated protein kinase (*PRKAG2*) on chromosome 7q36 (MacRae et al., 1995; Blair et al., 2001; Arad et al., 2002) (MIM 194200), *MYBPC3* on chromosome 11p11 (Carrier et al., 1993; Watkins et al., 1995; Niimura et al., 1998; Moolman-Smook et al., 1998) (MIM 115197), *CSRP3* on chromosome 11p15 (Geier et al., 2003) (MIM 600824), the essential myosin light chain (*MYL2*) gene on chromosome 12q23-q24 (Poetter et al., 1996) (MIM 160781), the α -myosin heavy chain (*MYH6*) gene on chromosome 14q12 (Tanigawa et al., 1990; Niimura et al., 2002) (MIM 192600), *MYH7* on chromosome 14q12 (Jarcho et al., 1989; Geisterfer-Lowrance et al., 1990; Moolman et al., 1995) (MIM 192600), *ACTC* on chromosome 15q14 (Mogensen et al., 1999; Olson et al., 2000) (MIM 192600), *TPM1* on chromosome 15q22 (Thierfelder et al., 1993; Thierfelder et al., 1994) (MIM 115196) and *TNNI3* (Kimura et al., 1997) (MIM 191044) (Fig. 1.5).

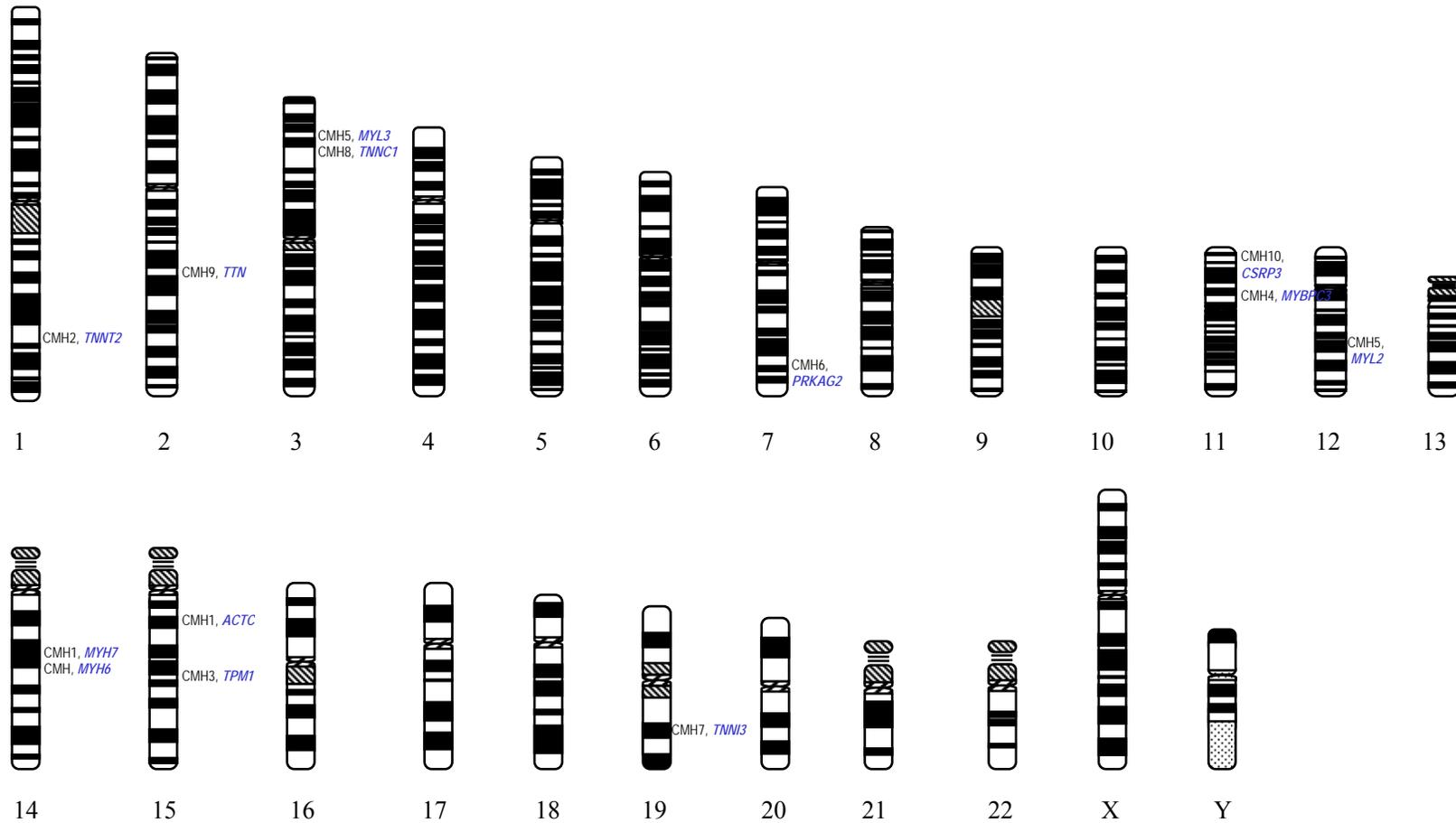


Fig. 1.5

Human cytogenetic map indicating mapped HCM-causative loci and genes.

ACTC = cardiac actin; *CMH* = hypertrophic cardiomyopathy (HCM) locus; *CSRP3* = muscle LIM protein; *MYBPC3* = cardiac myosin-binding protein C; *MYL2* = essential myosin light chain; *MYL3* = regulatory myosin light chain; *MYH6* = α -myosin heavy chain; *MYH7* = β -myosin heavy chain; *PLB* = phospholamban; *PRKAG2* = adenosine monophosphate (AMP)-activated protein kinase; *TNNC1* = cardiac troponin C; *TNNI3* = cardiac troponin I; *TNNT2* = cardiac troponin T; *TPM1* = α -tropomyosin. Disease-causative genes are shown in blue.

1.3.6.2 Molecular mechanisms of familial HCM

By 2002, HCM-associated mutations had been identified in 11 genes whose protein products are components of the sarcomere, specifically β -myosin heavy chain (Geisterfer-Lowrance et al., 1990), α -myosin heavy chain (Tanigawa et al., 1990), cardiac troponin T and α -tropomyosin (Thierfelder et al., 1994), cardiac myosin-binding protein C (Watkins et al., 1995; Niimura et al., 1998; Moolman-Smook et al., 1998); the essential and regulatory myosin light chains (Poetter et al., 1996), troponin I (Kimura et al., 1997), cardiac actin (Mogensen et al., 1999; Olson et al., 2000), titin (Sato et al., 1999) and troponin C (Hoffmann et al., 2001) (Fig. 1.3). Defects in these proteins indicated that HCM was primarily a disease of the sarcomere, hence the term “sarcomeropathy” was coined by Thierfelder et al., in 1994, and, more importantly, it pointed to the common underlying disease mechanism.

1.3.6.2.1 Alterations in force generation

Animal model studies have been performed to determine the physiological effect of HCM-associated mutations and, in most cases, it demonstrated that the mutant proteins are effectively incorporated into the cardiac myofilament, but become “poison peptides” that impaired the normal functioning of the sarcomere (Watkins et al., 1996; Rust et al., 1999). Mutations that alter proteins and cause HCM in humans have been genetically engineered into the genomes of various animal models in order to gain a better understanding of their associated pathophysiology. However, the conclusions drawn from these studies have been inconsistent (Maass and Leinwand, 2000), as will be discussed below.

Tardiff et al., (1998) demonstrated that mutations in troponin T altered α -tropomyosin-actin activity and caused the cardiac myofibres to be in a hypercontractile state; thus increasing force

generation. However, the animal model study by Yang and colleagues (1998) that was designed to determine the effects of truncation mutations in cardiac myosin-binding protein C demonstrated that the altered protein resulted in a leftward shift in the pCa^{2+} -force curve, thereby reducing the power stroke because the myofibres were in a hypocontractile state. An analogous study undertaken thereafter demonstrated that different truncation mutations in cardiac myosin-binding protein C caused a hypercontractile state of the myofibres (Witt et al., 2001), thereby contradicting the previous study by Yang et al., (1998). Additionally, Spindler et al., (1998) demonstrated that the myofibres in mice with the α -myosin heavy chain-403 (α MHC^{403/+}) heterozygous mutation, which was analogous to the arginine-403 glutamine HCM-causative mutation in humans, demonstrated hypocontractility and a reduction in overall contractile work.

However, there was a turning point in 1999, when Redwood and colleagues noted that different HCM-causative mutations each resulted in abnormal Ca^{2+} sensitivities and thus proposed that defective proteins increased contractile velocity and created a hypercontractile state. Redwood et al., (1999) suggested that the same mutant proteins could also reduce the maximum contractile force that was generated, thereby creating an opposing hypocontractile state. A more important theme of this review was the hypothesis that altered cardiac energetics gave rise to the hypertrophy associated with some forms of HCM (Redwood et al., 1999). This statement challenged the dogma at the time, but by design, it encouraged other investigations into the role of cellular energy in HCM.

1.3.6.2.2 Alterations in cellular energy homeostasis: a “Pandora’s box”

The cell energetics hypothesis proposed by Redwood et al., (1999) was most likely based on the research by Spindler et al., (1998) and Sweeney et al., (1998), who had demonstrated in animal

model studies that HCM-causative mutations in the β -myosin heavy chain placed a greater energy demand on cardiac myofibrils. However, Blair et al., (2001) provided the first evidence in a human model of the role of cellular energy homeostasis when they described genetic defects, which were associated with HCM accompanied by Wolff-Parkinson-White (WPW) syndrome, in *PRKAG2*, the gene that encodes the gamma (γ)-2 subunit of AMP-activated protein kinase (AMPK).

The enzyme AMPK serves as a cellular “fuel gauge” by activating glycolysis and fatty acid (FA) uptake when the cellular ATP level falls below threshold concentrations (Hardie and Carling, 1997; Blair et al., 2001). Prior to Blair and colleagues’ (2001) findings, Milan et al., (2000) had demonstrated that mutations in a gene that encodes the γ -3 subunit of AMPK caused a glycogen storage disorder that was associated with skeletal muscle abnormalities in pigs. Although the Milan et al., (2000) report did not describe any cardiac abnormalities, it was important because it demonstrated severe porcine skeletal muscle hypertrophy. Thus, Blair et al., (2001) suggested that the underlying mechanism of cardiac hypertrophy was an inability by the cardiomyocytes to maintain an adequate cellular energy supply, given that cardiac hypertrophy could be a primary feature associated with certain HCM-causative mutations.

The identification of mutations in AMPK contradicted previous findings that solely implicated sarcomeric proteins in the pathogenesis of HCM. Nonetheless, it offered an explanation as to how hyper- or hypo-contractile states of the myofibres resulting from sarcomeric protein mutations could both cause HCM.

The energy homeostasis hypothesis

It has been proposed that, for HCM, a decreased supply of ATP (energy) might inhibit Ca^{2+} re-uptake by SERCA2 (Ashrafian et al., 2003; reviewed by Moolman-Smook et al., 2003). The reduced SERCA2 activity causes an increased availability of cytosolic Ca^{2+} that could trigger hypertrophic pathways (Spindler et al., 1998), or, in extreme cases, provide substrates for arrhythmogenic events resulting in sudden cardiac death (Somura et al., 2001). The hypothesis that alterations in energy homeostasis and changes in Ca^{2+} handling are involved in the pathogenesis of DCM has additionally been suggested (Seidman and Seidman, 2001; Watkins, 2003) (also refer to section 1.3.3.2 and 1.3.3.3). Consequently, unlocking the “Pandora’s box” has unravelled some of the mysteries of the cardiomyopathies, which lead to the suggestion of “a unifying hypothesis” to describe the common pathophysiologies in DCM and HCM (Fatkin and Graham, 2002). This paradigm shift has provided impetus towards the search for other proteins that may play a role in Ca^{2+} or cellular energy homeostasis (Seidman and Seidman, 2002; Watkins, 2003). Additionally, this data is important for the present study, as it opens a new vista of likely candidate genes in which to search for the PFHBII-causative mutation.

1.3.7 A synopsis of familial HCM

Presently, familial HCM-causative mutations have been described in 11 genes encoding protein components of the sarcomere, one gene encoding a cytoskeletal protein (*CSRP3*) and a gene encoding a regulatory enzyme (*PRKAG2*). Functional studies to determine the physiological effects of HCM-causative mutations have indicated that deficits in force generation are an underlying mechanism in the pathogenesis of the condition. In addition, the identification of HCM-associated defects in an enzyme (AMPK) that is involved in maintaining cellular ATP levels has suggested that improper cellular energy homeostasis may be an additional cause of HCM. This hypothesis of improper energy homeostasis ties in with other speculations supporting

the role of Ca^{2+} and provides plausible evidence that common processes might be involved in the pathogenesis of all forms of cardiomyopathy.

1.4 Familial cardiac conduction disorders

The previous sections aimed to draw a clinical and aetiological comparison between PFHBII and diseases associated with cardiomyopathies. Pathologically, isolated (“pure”) cardiac conduction abnormalities are usually the first clinical manifestation of PFHBII, so much so that Brink and Torrington (1977) classified the disorder as a cardiac conduction disease. Based on later investigations by P. Brink (personal communication), as well as the present study, it is evident that PFHBII is a primary cardiomyopathy. However, it is possible that the cause of other pure forms of cardiac conduction disease might well provide insight to the identity of the PFHBII-causative gene. Thus, a review of these forms of cardiac disorders is provided hereafter.

1.4.1 History and clinical features of isolated cardiac conduction diseases

At the beginning of the 20th century, Morquio (1901) provided the first contemporary documentation of a familial cardiac conduction disorder, by describing a family that presented with clinical features of syncope (loss of consciousness) and SB. Almost three decades later, both Aylward (1928) and Aitken (1932) described individuals in two different families who presented with characteristics that suggested congenital forms of cardiac conduction disease. The study by Aitken et al., (1932) was particularly significant because it demonstrated the first use of ECG analysis to diagnose cardiac conduction defects. In the decades that followed, a multitude of studies described individuals in families with clinical features that may be associated with cardiac conduction disease – a summary of the clinical features and mode of inheritance is provided in Table 1.2. The leading characteristics that were identified by ECG analysis were bi-fascicular block (BFB), complete atrioventricular (AV) block (also defined as CHB), left anterior

hemiblock (LAHB), left axis deviation (LAD), left bundle branch block (LBBB) occurring in conjunction with broad QRS complexes, PR interval prolongation (first-degree heart block), right bundle branch block (RBBB), syncope and/or sudden death (Table 1.2).

In the context of the present study, it has to be noted that numerous inherited cardiac conduction disorders have been described in South African families (Combrink et al., 1962; Myburgh and Steenkamp, 1973; Brink and Torrington, 1977; Torrington et al., 1986; Van der Merwe et al., 1986; Brink et al., 1995). Later sections will illustrate the relevance of these investigations and their impact on the global understanding of cardiac conduction diseases.

1.4.2 The cardiac conduction system: structural complexity and disease manifestation

The cardiac conduction system is comprised of specialised connective tissue and structures that resemble undeveloped myofibrils (James and Sherf, 1971; Virágh and Challice 1982; James, 2002). A number of studies have demonstrated that different proteins are expressed in the different regions of the atrial and ventricular conduction systems (Anderson and Ho, 1998; Schram et al., 2001; Pennisi et al., 2002; Myers and Fishman, 2003). Generally, disorders of the cardiac conduction system have a common pathological feature of causing a delay or complete block of the electrical impulse that is generated at the SA node and progresses via connective fibres through the AV node, Bundle of His and the cardiac bundle branches toward the Purkinje fibres of the heart (Fig. 1.6). The particular location of the conduction abnormality has been shown to be significant because it accounts for the clinical and electrophysiological features of that disorder (Waxman et al., 1974). Thus, by identifying the location of the conduction abnormality, insights can be gained to the likely cause of the defect.

Table 1.2*Timeline of familial isolated cardiac conduction disorders*

Authors	Clinical features	Country where described	#Mode of inheritance
Morquio (1901)	Syncope, SB	France	Unknown
Aylward (1928)	Congenital heart block	England	Unknown
Aitken (1932)	Congenital heart block	England	Unknown
Conner et al., (1959)	CHB	North America	Unknown
Combrink et al., (1962)	CHB, RBBB	South Africa	Dominant
Lenègre (1964)	RBBB, LBBB with broad QRS complexes, CHB, syncope and sudden death	France	Dominant
Gazes et al., (1965)	Congenital heart block	North America	Dominant
Wagner and Hall (1967)	CHB	North America	Dominant
Lev (1970)	RBBB, LBBB with broad QRS complexes, CHB, syncope and sudden death	France	Dominant
Sarachek and Leonard (1972)	SB, congenital heart block, adult-onset first-degree heart block and second-degree heart block	North America	Dominant
Myburgh and Steenkamp (1973)	Congenital AV block	South Africa	Unknown
Schaal et al., (1973)	RBBB, LAD and CHB	North America	Dominant
Lynch et al., (1975)	SB and arrhythmias	North America	Dominant
Brink and Torrington (1977)	Type I – RBBB, LAHB, CHB with broad QRS complexes; Type II – SB, LPHB, CHB with narrow QRS complexes	South Africa	Dominant
Stephan (1978)	RBBB and CHB	Lebanon	Dominant
Barak et al., (1987)	First, second or complete AV block, Adam-Stokes attacks and sinus node dysfunction	Israel	Dominant
Schott et al., (1999)	Type I - Adult onset progressive RBBB and AV block; Type II – non-progressive	Netherlands and France	Dominant
Tan et al., (2001)	broad P waves, first-degree heart block, broad QRS complexes and severe SB	Netherlands	Dominant

= autosomal inheritance pattern, except where the type of inheritance is unknown; AV = atrioventricular; CHB = complete heart block; LAD = left axis deviation; LAHB = left anterior hemiblock; RBBB = right bundle branch block; SB = sinus bradycardia

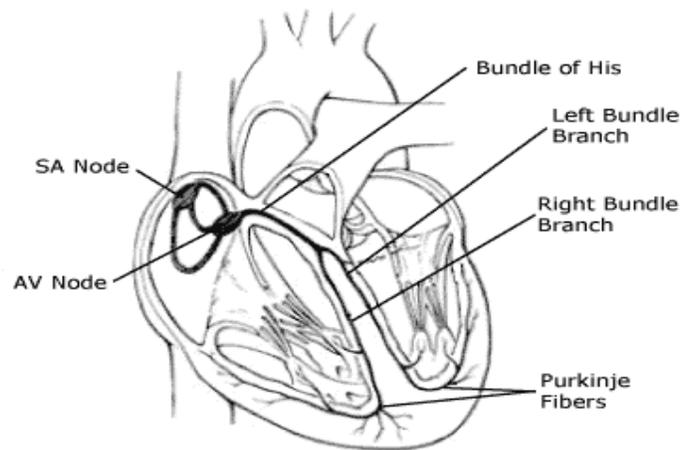


Fig. 1.6

Anatomy of the cardiac conduction system.

AV = atrioventricular; SA = sinoatrial

1.4.3 Molecular genetics of familial isolated cardiac conduction disorders

1.4.3.1 Mapping disease loci for progressive familial heart block type I (PFHBI) and isolated cardiac conduction disease (ICCD)

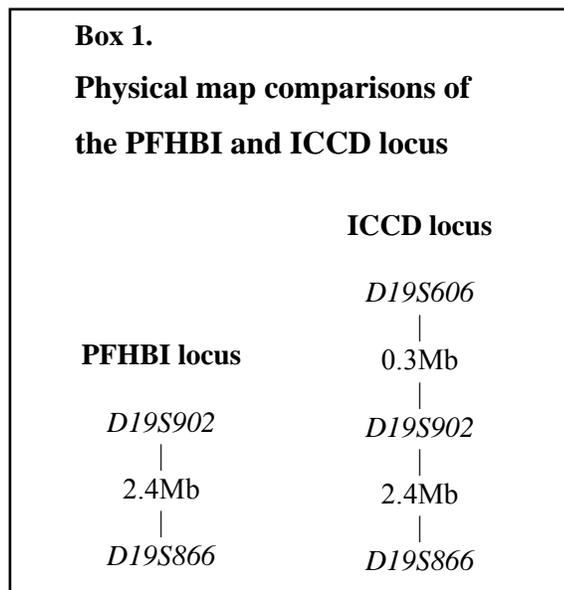
PFHBI

In 1962, Combrink and colleagues first documented a familial cardiac conduction disease in South Africa when they described a Caucasian family that presented with RBBB and CHB. Fifteen years thereafter, Brink and Torrington (1977) described another Caucasian South African family that presented with RBBB, LAHB, BFB or CHB with broad QRS complexes, a disorder they classified as progressive familial heart block type I (PFHBI) (MIM 113900). In addition, Brink and Torrington (1977) noted that the clinical features of PFHBI were similar to the features that Combrink et al., (1962) had previously described in a South African family, and thus suggested that the two families belonged to the same kindred (Brink and Torrington, 1977). Later, in 1986, Torrington and colleagues traced nine additional PFHBI-affected families, although, to date, it is not apparent whether any of these families were related to the family described by Combrink et al., (1962). In the mid-990s, Brink and colleagues (1995) reported the

results of genome-wide linkage analysis, which positioned the PFHBI-causative gene on chromosome 19q13.3 (Fig. 1.7). Subsequent *haplotype* analysis refined the location of the PFHBI-causative gene to within a 2.4 *megabase* (Mb) interval flanked by markers *D19S902* and *D19S866* (Box 1) (Z. Arieff, personal communication).

ICCD

In the same year that Brink and colleagues had described mapping the PFHBI locus, de Meeus et al., (1995) reported mapping the conduction disorder that was previously described by Stephan in 1978 in a Lebanese family, a condition termed isolated cardiac conduction disease (ICCD), to a region on chromosome 19q13.3 (Fig. 1. 7). The ICCD-causative locus spans a 2.7Mb region flanked by markers *D19S606* and *D19S866*, a region that overlaps with the PFHBI locus (Box. 1) (Z. Arieff, personal communication). Although identified in different countries and in families of different ethnicity, because the two diseases map to within a common chromosomal interval, it has been suggested that PFHBI and ICCD are the same disease even though classified by different names (P. Brink, personal communication).



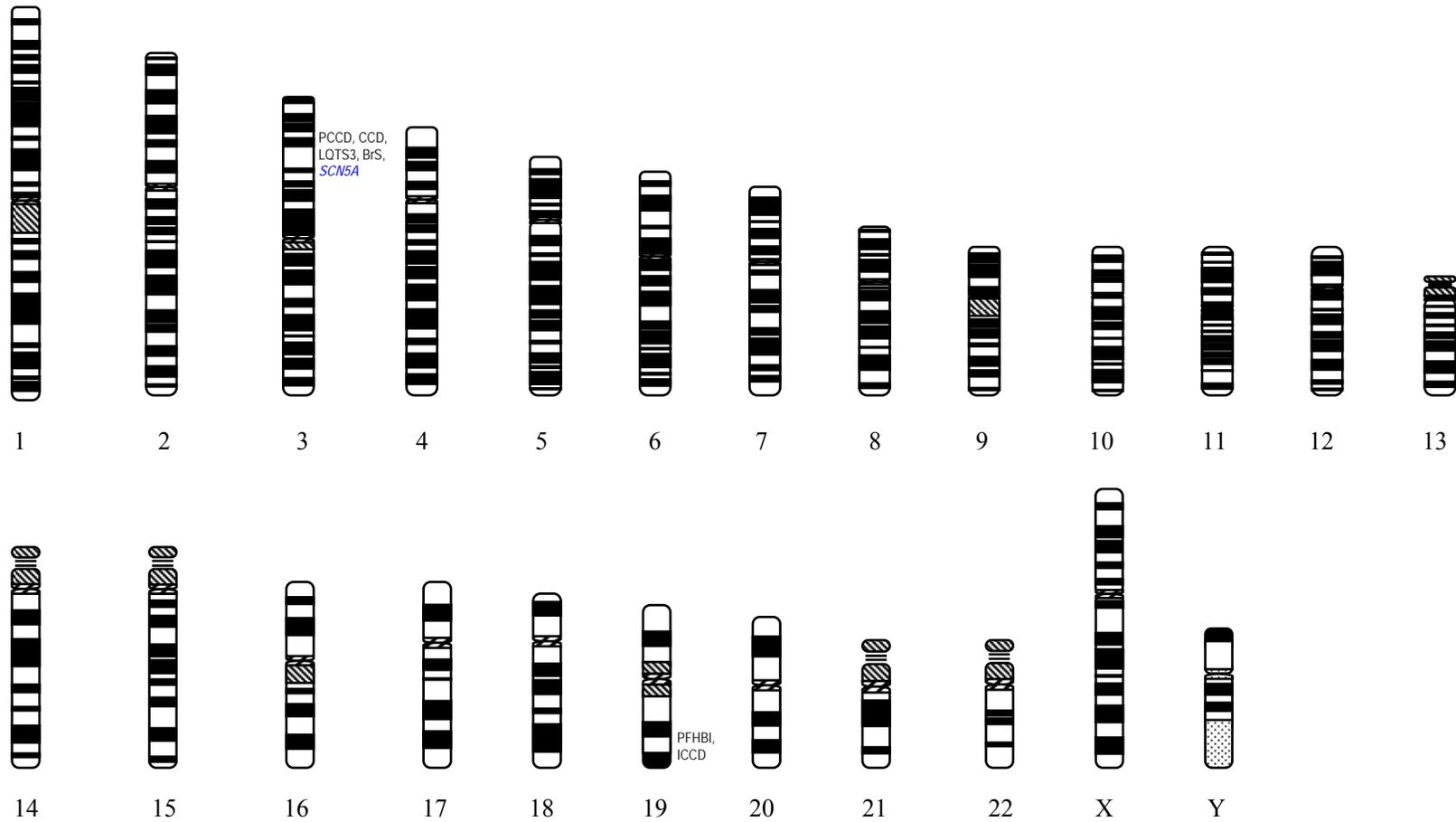


Fig. 1.7

Human cytogenetic map indicating mapped cardiac conduction disease loci and genes.

BrS = Brugada syndrome; *CCD* = cardiac conduction defect; *ICCD* = isolated cardiac conduction disease; *LQTS3* = long QT syndrome type 3; *PCCD* = progressive cardiac conduction disease; *PFHBI* = progressive familial heart block type I; *SCN5A* = cardiac Na^+ channel. Disease-causative genes are shown in blue.

1.4.4 Molecular aetiologies of isolated cardiac conduction disorders

1.4.4.1 Progressive cardiac conduction defect

Lenègre (1964) and Lev et al., (1970) described a number of French families that presented with cardiac conduction disturbances that were characterised by RBBB, LBBB with broad QRS complexes, CHB, syncope and progression to sudden death (Table 1.2). Later, the term Lenègre-Lev disease was used to describe the clinical entities described by Lenègre (1964) and Lev et al., (1970), while, more recently, progressive cardiac conduction defects (PCCD) (Schott et al., 1999) (MIM 600163) is increasingly being used to describe conditions with these conduction abnormalities.

Although PCCD is the most common form of isolated cardiac conduction anomaly, little was known about its underlying molecular aetiology, until 1999, when Schott and colleagues reported that they had performed a linkage study in a French and a Dutch kindred with PCCD and had mapped the causative gene to a locus on chromosome 3p21 (Fig. 1.7). Their subsequent candidate gene mutation analyses at the locus identified a PCCD-causative splice site mutation, which segregated in the French, but not the Dutch, family in a gene encoding a cardiac-expressed sodium (Na^+) channel (*SCN5A*). Further mutation analyses of *SCN5A* identified a unique deletion mutation that segregated exclusively in the affected Dutch family (Schott et al., 1999). The findings by Schott et al., (1999) were interesting because mutations had previously been described in *SCN5A* that caused three types of inherited cardiac arrhythmias known as long QT syndrome (LQTS) type 3 (LQT3) (Wang et al., 1995) (MIM 603830), Brugada syndrome (BrS) (Brugada and Brugada, 1992; Chen et al., 1998; Bezzina et al., 1999; Rook et al., 1999) (MIM 601144) and idiopathic ventricular fibrillation (IVF) (Chen et al., 1998) (MIM 600163). Although LQT3, BrS and IVF are associated with rapid heart rhythms, while PCCD is associated with a slowing of the heart rate, the identification of cardiac conduction- and arrhythmia-

causative mutations in *SCN5A* was significant because it provided clues to the underlying mechanism of diseases of the cardiac conduction system. This will be discussed in section 1.4.5.

1.4.4.2 Cardiac conduction defect (CCD)

Tan et al., (2001) described a family that presented with broad P waves, first-degree heart block, broad QRS complexes and severe SB, a disorder designated cardiac conduction defect (CCD) (MIM 115080). The previous identification of mutations in a cardiac Na⁺ channel by Schott et al., (1999) as cause of PCCD (see previous section) suggested that other cardiac conduction abnormalities might have similar molecular aetiologies. Using this rationale, Tan et al., (2001) performed a positional independent candidate gene analysis, which resulted in the identification of a CCD-causative mutation in *SCN5A* (Tan et al., 2001) (Fig. 1.7). These results aided in identifying the pathological mechanisms of different cardiac conduction system disorders (Tan et al., 2001).

1.4.5 Mechanisms of cardiac conduction disorders

1.4.5.1 Alterations in ion channel electrophysiology

Prior to the reports by Schott et al., (1999) and Tan et al., (2001), the underlying disease mechanisms of familial cardiac conduction disorders had not been established. However, the pathophysiologies of the arrhythmic disorders LQTS types 1-6 (MIM 192500, 152427, 600919, 176261, 603830 and 603796) and BrS (Brugada and Brugada, 1992), which could also be accompanied by cardiac conduction defects such as LAD and RBBB, were well established and served as a model that helped to elucidate the mechanisms that lead to cardiac conduction defects.

It had been demonstrated that genetic defects in various genes that encode two potassium (K^+) channels and their associated subunits (Wang et al., 1996; Curran et al., 1995; Splawski et al., 1997; Tyson et al., 1997; Abbott et al., 1999) and a Na^+ channel (Wang et al., 1995) were the underlying causes of LQTS, while BrS was also caused by mutations in a gene encoding a cardiac Na^+ channel (Chen et al., 1998; Bezzina et al., 1999; Rook et al., 1999). Interestingly, the LQT3- (Wang et al., 1995), BrS- (Chen et al., 1998; Bezzina et al., 1999; Rook et al., 1999) and IVF-causative mutations (Chen et al., 1998) have been identified in *SCN5A*, the same cardiac Na^+ channel in which mutations were shown to cause PCCD and CCD. These common molecular aetiologies suggested that cardiac conduction disorders and arrhythmias shared similar pathophysiologies.

Pathophysiology of CCD

Functional electrophysiological studies indicated that LQTS-causative mutations in *SCN5A* altered cardiac electrical conductance by delaying channel repolarisation, which, in turn, causes a slowing of the overall heart rhythm and eventually results in the prolongation of the QT interval (Wang et al., 1995). Alternately, BrS-causative mutations in *SCN5A* were shown to increase cardiac Na^+ channel repolarisation, causing the heart's electrical activity to become disordered, thus triggering unsynchronised atrial and ventricular contractions (Chen et al., 1998). Consequently, Tan et al., (2001) performed similar functional studies to determine the effect of the CCD-causative mutation on Na^+ channel electrophysiology and demonstrated that it caused a delay in channel repolarisation, but did not alter the overall rhythm of the heart. Consequently, the delayed channel repolarisation caused a slowing of the conduction impulse between the atrial and the ventricular conduction fibres without triggering the arrhythmic events associated with LQTS and BrS (Tan et al., 2001).

Pathophysiology of PCCD

In 2003, Probst et al., described the results of an *in vitro* cell-transfection study to determine the functional consequences of a splice site mutation in *SCN5A* (IVS.22+2 T→C) that had previously been described by Schott et al., (1999) as cause of PCCD. Three fundamental outcomes were derived from this study. Firstly, the study predicted that the mutation caused an in-frame skipping of exon 22, which resulted in a Na⁺ channel in which the voltage-sensitive DIIS4 domain was absent. Secondly, the study showed that PCCD was caused by *haploinsufficiency*, because the mutated copy of the protein was processed correctly within the cardiac cell membrane, but the defective protein did not generate an inward current, which resulted in a reduction in the velocity of the conduction impulse (Probst et al., 2003). Lastly, Lenègre (1963) and Lev et al., (1964) had suggested that natural fibrosis of the cardiac conduction system caused progressive AV block, although fibrosis was observed in elderly individuals who did not have AV block (Davies et al., 1967). Taking this into consideration, Probst et al., (2003) suggested that, for PCCD, the progression to CHB spanning a number of years occurred as a result of normal age-dependent fibrosis of the conduction system, which exacerbated the alteration in the conduction velocity that was caused by the genetic defect in the cardiac Na⁺ channel.

1.4.6 A synopsis of familial isolated cardiac conduction disorders

Cardiac conduction system disorders are clinically variable, although the location of the abnormality has provided insight into the probable causes of these defects. The causative gene(s) for two familial cardiac conduction disorders (PFHBI and ICCD) have been mapped to the same locus on chromosome 19q13.3, while the underlying cause of two other forms of cardiac conduction disorders (PCCD and CCD) have been shown to be genetic defects in a cardiac Na⁺ channel. These genetic defects were shown to cause alterations in channel electrophysiology by slowing the conduction impulse. Genetic defects in *SCN5A* have additionally been associated

with the development of the cardiac arrhythmic disorders, LQTS and BrS, which indicates that common molecular aetiologies may exist for different clinical entities that affect cardiac electrophysiology.

1.4.7 Familial cardiomyopathies and cardiac conduction diseases in South Africa: contributions to global knowledge

1.4.7.1 Diseases of the myocardium

In the USA, DCM is estimated to affect nearly 40 in 100 000 individuals (Durand et al., 1995a), and, furthermore, is considered to be the foremost indication of heart transplantations in older individuals (Seidman and Seidman, 2001; Towbin and Bowles, 2004). However, in South Africa, current statistics on the prevalence of DCM have been difficult to obtain, which probably reflects a lack of data because of the complex nature of the condition, rather than the low prevalence of the disease (Moolman-Smook et al., 2003). One genetic study has been performed in South African populations in which the prevalence of DCM-associated mutations in cardiac and skeletal actin was determined (Mayosi et al., 1999). These authors suggested that the global prevalence of DCM-causative mutations in cardiac and skeletal actin was low in comparison to mutations in other DCM-associated genes. Later, Takai et al., (1999) and Karkkainen et al., (2002) provided more evidence that supported Mayosi and colleagues' original proposal.

On the other hand, HCM has an estimated prevalence of 0.2% (or one in 500) in populations in the USA and Europe and has been shown to be a leading cause of death in young adults in those countries (Maron et al., 1995). As with DCM, the prevalence of HCM has not been established in South Africa, although numerous important HCM-associated findings have been reported from this country (Moolman et al., 1995; Moolman et al., 1997; Moolman-Smook et al., 1998; Moolman-Smook et al., 1999; Moolman-Smook et al., 2002).

Although DCM and HCM are separate clinical entities, myocyte hypertrophy is often observed in patients with DCM, while individuals with HCM have been shown to progress to DCM (Seidman and Seidman, 2000). The novel arginine-92 tryptophan (Arg92Trp) mutation in cardiac troponin T, which was shown to segregate in two South African families (Moolman et al., 1997), is an example of the latter. The mutation causes mild or undetectable hypertrophy, although numerous individuals with this defect have been identified that demonstrate rapid progression to a DCM-like *phenotype* that is associated with a high incidence of sudden cardiac death (J. Moolman-Smook, personal communication).

Other studies in South African families have also contributed significantly to the overall understanding of the molecular mechanisms of cardiomyopathies. The study by Moolman-Smook et al., (1998) described the first missense mutation in the C5 domain of cardiac myosin-binding protein C that causes adult-onset HCM. Subsequent investigations have demonstrated that HCM-causative missense mutations modulate the interactions between three myosin-binding protein C domains (Moolman-Smook et al., 2002). Significantly, Moolman-Smook et al., (2002) proposed a new model for the arrangement of cardiac myosin-binding C in the sarcomere, thereby repealing previous tenets and, in so doing, provided novel insights into the causes of HCM.

1.4.7.1.2 Diseases of the cardiac conduction system

In South Africa, the prevalence of cardiac conduction disease has not been established, barring the study by Torrington et al., (1986), which suggested that PFHBI had a prevalence of one in 500 South Africans of European descent. Because of the availability of large families in which PFHBI segregated, Brink et al., (1995) undertook a study that described mapping the first locus (in South Africa and internationally) linked to an isolated cardiac conduction disorder. Although

the PFHBI-causative gene has not been identified, the identification of the causative locus may have provided a starting point for subsequent linkage studies by other investigators. It has been reported in the publication by Schott et al., (1999) that the authors had performed a candidate locus linkage analysis that initially analysed and excluded the PFHBI locus prior to the eventual identification of PCCD-causative mutations in *SCN5A*. Therefore, studies considered to be unique to South Africa could facilitate investigations in other countries.

1.4.8 The implication of establishing common clinical and genetic links between diseases

It is not unreasonable to posit that PFHBII, which has some clinical features that are similar to the familial cardiomyopathies and cardiac conduction disorders described in this review, might well have been described in other families but under a different clinical designation. Therefore, it is essential to establish whether clinical entities with similar features may represent a disease caused by a defect in the same gene. In order to address this possibility, analyses of candidate loci that are associated with diseases that present with clinical features similar to PFHBII can be undertaken. However, once a locus has been mapped, identifying the disease-causing gene is by no means an effortless task, although resources provided by HGP databases have provided an adjunct to molecular genetic analyses. It has been suggested that HGP resources will accelerate the identification of the underlying cause of familial cardiac diseases (Roberts, 2000; Komajda and Charron, 2001). Consequently, this review will shift a gear by providing an overview of the HGP and, in particular, underline the advantages and the drawbacks of using these technologies to identify disease-causative genes.

1.5 The HGP: background

The “Human Genome Project” was initiated in 1988 and one of its primary aims was to construct *physical, genetic and transcript maps* of the human genome (Ikawa, 1991; Roberts 2000; Yaffe et

al., 2001; National Human Genome Research Institute [NHGRI] website, <http://www.nhgri.nih.gov>). A second objective of the HGP was the development of computer-based biological tools that collectively became known as *bioinformatics* (the application of information technology and *in silico* molecular biology to store and analyse one dimensional data) (Sansom and Smith, 2000; Chieurel, 2002). By 2002, molecular genetic approaches had linked approximately 25% of the estimated 4000 Mendelian genetically inherited diseases to chromosomal regions in the human genome (Perez-Iratxeta et al., 2002). However, about 450 of the linked diseases described then had no associated disease-causative gene (Perez-Iratxeta et al., 2002), although it had been widely suggested that the application of HGP data would accelerate the identification of those molecular aetiologies (Guttmacher and Collins, 2002; Perez-Iratxeta et al., 2002). A few of these successes will be illustrated in sections hereafter.

1.5.1 Identifying disease-causative genes and the HGP

1.5.1.1 The pre-HGP era: 1970-1988

For at least two decades, linkage analysis has successfully been used to map disease-causative loci (Lathrop and Lalouel, 1984; Ott, 1999). Once a chromosomal locus was mapped, positional cloning techniques such as *cDNA* selection and exon trapping were executed to identify genes and differentiate expression profiles within *contiguous* genomic clones at the locus (Duyk et al., 1990, Parimoo et al., 1991). Positional cloning techniques have been successful in identifying disease genes that are associated with cardiac disorders. The LQTS types 1, 2 and 3-causative genes, the former two encoding cardiac K^+ channels, and the latter encoding a Na^+ channel, were identified using “wet-bench” positional cloning strategies (Curran et al., 1995; Wang et al., 1995; Wang et al., 1996). Additionally, three sarcomeric protein-encoding genes in which mutations have been described that cause HCM (*MYH7*, *TNNT2* and *TPM1*) were identified by positional cloning by Geisterfer-Lowrance et al., (1990) and Thierfelder et al., (1994). The initial successes

of positional cloning studies paved the way for position-dependent and even position-independent candidate gene analyses, once common molecular aetiologies had been identified for LQTS and HCM (Watkins et al., 1995; Keating and Sanguinetti, 2001). Some examples of LQTS-associated genes identified by position-dependent or positional independent analyses are the K⁺ channel genes, *KCNE1* (Splawski et al., 1997; Tyson et al., 1997) and *KCNE2* (Abbott et al., 1999), and examples of HCM-associated genes are *TTN* (Sato et al., 1999), *MYL2* (Poetter et al., 1996) and *TNNI3* (Kimura et al., 1997).

1.5.1.2 Successful applications of HGP resources

Initially, HGP databases provided the sequences of genetic markers at loci spanning the human genome (Weissenbach et al., 1992). These genetic marker maps facilitated linkage studies and resulted in the identification of numerous disease-causative loci (Sullivan et al., 1999). Positional cloning techniques were then used to identify genes at a locus that was linked to a particular disease. However, positional cloning strategies were expensive and time-consuming, and, thus, towards the end of the 1990s, bioinformatics analyses were increasingly applied to identify novel genes at a given locus (Sullivan et al., 1999). The use of bioinformatics technologies resulted in the identification of novel genes, and, more often than not, assisted in the subsequent identification of disease-causing mutations in those genes. Examples of *in silico* analyses that have resulted in the identification of novel genes implicated as cause of cardiac disorders, ion channel defects or skeletal myopathies are listed in Table 1.3. Bioinformatics analyses have more recently been used to create transcriptional profiles of the human heart (Bortoluzzi et al., 2000), which in turn, have provided information for subsequent studies that have identified the differences in expression profiles that are associated with DCM and HCM (Dos Remedios et al., 2003; Yung et al., 2004).

Despite numerous successful applications of the HGP data and bioinformatics, there are caveats that need to be applied when utilising *in silico* disease-causing gene identification approaches instead of conventional wet-bench laboratory procedures.

Table 1.3

A representative list of disease-causative genes associated with cardiac disorders, channelopathies or skeletal myopathies that were identified using bioinformatics analysis

Gene	Protein product (associated disease in brackets)	Reference
<i>PDS</i>	Sulphur channel (Pendred syndrome)	Everett et al., 1997
<i>DTNB</i>	Dystrobrevin beta (Limb-girdle muscular dystrophy)	Puca et al., (1998)
<i>SCN1A</i>	Sodium channel (Epilepsy and febrile seizure type 2)	Escayg et al., (2000)
<i>LGMD2G</i>	Dystrophin (Limb-girdle muscular dystrophy type 2)	Moriera et al., (2000)
<i>NDRG1</i>	Oxygen regulated protein (Charcot-Marie-Tooth disease)	Kalaydjieva et al., (2000)
<i>SCA10</i>	Ataxin-10 (Spinocerebellar ataxia type 10)	Matsuura et al., (2000)
<i>GAN</i>	Gigaxonin (Giant axonal neuropathy and cytoskeletal disorganisation)	Bomont et al., (2000)
<i>SRF</i>	Serum response factor (DCM in mice)	Wang et al., (2001), Zhang et al., (2003)
<i>TNNI3K</i>	Cardiac-specific kinase (possible cardiomyopathy similar to an autosomal recessive form of DCM)	Zhao et al., (2003)

1.5.2 Limitations of the HGP: *caveat emptor*

1.5.2.1 Sequence errors

By the year 1999, it was estimated that at least 10Mb of sequence data were deposited per week into HGP databases (Felsenfeld et al., 1999). The large amounts of data that were submitted to HGP databases have made these resources susceptible to an accumulation of sequencing and annotation errors (Felsenfeld et al., 1999). Unfortunately, the enormity of the HGP makes it impossible to have error-free databases. Therefore, a number of acceptable error rates per sequence read have been set, which then, by analysing the number of errors in every 10 sequence reads of the same genomic segment, is used to categorise data into *finished* or *draft sequence*

(Sansom and Smith, 2000). As a general rule, an error rate of less than one error per 10 000 bases is set as a standard for categorising a sequenced portion of a genome as finished sequence (Sansom and Smith, 2000). Consequently, any sequence obtained from an HGP database, even so-called finished sequence, will contain errors. Therefore, in order to minimise the effects of these errors, it has been suggested that at least two separate analytical algorithms or databases should be used when performing bioinformatics analyses (Attwood, 1999) (refer to section 2.11), since sequence quality varies between databases.

1.5.2.2 Gene annotation and prediction accuracy: “real” versus “virtual” genes

The sheer volume of sequence errors and the limited capabilities of gene prediction programs have hampered the accurate determination of the total number of genes in the human genome (Istrail et al., 2004). Furthermore, the accuracy of HGP gene prediction algorithms has been questioned (Burge and Karlin, 1997; Felsenfeld et al., 1999; Bork, 2000), although Katsanis and colleagues (2001) demonstrated that gene prediction quality improved in direct proportion to an improvement in sequence quality. Certain databases, such as the Ensembl database (<http://www.ensembl.org>), have circumvented the limitations of gene prediction programs by using *ab initio* gene prediction in conjunction with nucleotide and protein similarity searches from i) different regions of the human genome and ii) other model organisms such as the mouse, rat, pufferfish, zebrafish and chimpanzee, to obtain a 95% gene prediction accuracy (Stein, 2001). Although there are sequence and annotation limitations, the gradual completion of the genome sequencing projects of model organisms is yielding important data that minimises these shortcomings (Chieurel, 2002; Ouyang et al., 2003).

1.5.3 Current sequence and annotation status of the HGP

1.5.3.1 The status of the complete human genome

Initially, the HGP had set out to complete a ten times (10x) sequence coverage of the approximately three billion bases comprising the human genome (NHGRI website). By the year 2002, a four times draft sequence coverage of the human genome was completed (Initial Human Genome Sequencing Consortium, 2001) and the following year in April 2003, a cumulative total of 93% of the human genome was completely sequenced with at least an eight to ten times sequence coverage (Ensembl website). Previously, the total number of genes in the human genome had been estimated to range between 29 000-150 000 genes (Attwood, 2000; Harrison et al., 2002; Collins et al., 2003). There have been conflicting estimates regarding the total number of human genes, although, recently, consensus was reached which hinted at a figure of approximately 27 000 genes that constitute the human genome (Pennisi, 2003). However, this data is constantly changing and more recent analyses of the Ensembl website have indicated that there are now considered to be approximately 65 000 genes in the human genome (as of June 2004). The rapid changes in the predicted number of genes highlight the problems associated with gene prediction (a discussion on these problems and their possible influence on the outcomes of the present study is provided in section 4.8.2). The basis of the present study is a description of mapping the PFHBII-causative gene to a locus on chromosome 1 and, hence, the current annotation status of that chromosome will also be reviewed.

1.5.3.2 The status of human chromosome 1

The estimated size of human chromosome 1 is approximately 246Mb (Ensembl website). As of 14 April 2004, the sequencing status of chromosome 1 was indicated to be approximately 95% complete, which corresponds to about 233Mb of finished nucleotide sequence (The Sanger Centre website, <http://www.sanger.ac.uk>). In addition, 1995 genes that have been identified by

experimental means, or by gene prediction programmes, and 119 pseudogenes have been assigned to chromosome 1 (Ensembl website). Based on these data, human chromosome 1 has an average gene density of one gene every 116 kilobases (kb). The chromosomal region described in this study, to which the PFHBII gene is mapped, has an approximate gene density of one gene every 110kb. The relevance of the sequence and annotation status of chromosome 1, relative to the present study, is further elaborated on in section 4.8.2.

1.6 The present study

A previous report by Van der Merwe et al., (1986) suggested that PFHBII was possibly a cardiac conduction disorder that was accompanied by a congestive form of cardiomyopathy. If this held true, then clues to the underlying molecular aetiology of PFHBII could be obtained by analysing the molecular pathophysiologies of both cardiac conduction disorders and cardiomyopathies.

Although cardiac conduction defects, cardiac arrhythmias and cardiomyopathies generally occur in isolation, various familial disorders have been identified in which the clinical features of the one condition may accompany features that are associated with the other. Additionally, numerous studies have identified DCM- and HCM-causative mutations in different domains of the same gene (Geisterfer-Lowrance et al., 1990; Thierfelder et al., 1994; Watkins et al., 1995; Moolman-Smook et al., 1998; Satoh et al., 1999; Kamisago et al., 2000; Olson et al., 2001; Daehmlow et al., 2002; Gerull et al., 2002). Mutations have also been identified in different domains of a cardiac ion channel (*SCN5A*) that cause the cardiac conduction defects, PCCD and CCD, as well as the arrhythmic disorders LQT3, BrS and IVF (Brugada and Brugada, 1992; Wang et al., 1995; Chen et al., 1998; Bezzina et al., 1999; Rook et al., 1999; Schott et al., 1999; Tan et al., 2001).

Thus, the first step in identifying the PFHBII-causative gene was to establish whether DCM,

and/or other cardiac-associated features, is part of the clinical spectrum of the disorder. This investigation was important in order to establish whether PFHBII demonstrated a clinical overlap with other cardiac disorders; thereby also providing a basis for subsequent molecular genetic analyses. To this end, retrospective and prospective studies were performed to obtain all available clinical data and medical histories of subjects from the PFHBII-affected family. This investigation indicated that DCM is an additional a feature of PFHBII (Fernandez et al., 2004) (see section 3.1 and 4.1 for more detail).

Overlapping clinical features and common molecular aetiologies have been identified for numerous cardiomyopathies and cardiac conduction disorders. Therefore, it was hypothesised that PFHBII may share a common aetiology with a previously mapped cardiac disorder. Thus, genetic linkage analyses of candidate loci (chromosomal regions linked to previously described disorders that were associated with conduction defects and/or DCM or HCM) were performed. Fortuitously, linkage was established to a locus on chromosome 1q32.2-q32.3, a region that overlapped a previously described DCM-associated disorder (CMD1D). Genetic fine mapping and haplotype analysis placed the PFHBII-causative locus distal to the CMD1D locus, thus supporting the identification of a novel locus associated with DCM and cardiac conduction defects. Bioinformatics analyses were employed to identify positional candidate genes, which resulted in the selection of seven plausible genes that were screened for the PFHBII-associated mutation.

In summary, the present study mapped the PFHBII-causative gene to a novel locus on chromosome 1 and redefined the clinical classification of PFHBII; the latter might result in the identification of additional affected individuals in South Africa, or in other countries. It is possible that PFHBII, although classified by a different clinical name, may have been identified

in other families in South Africa, or in other countries. Therefore, mapping the PFHBII locus might lead to other groups mapping the same disease to the locus described in the present study. In so doing, the new mapping data could be used to further refine the PFHBII locus, thus reducing the interval in which to search for the disease-causative gene. Even though the PFHBII-associated mutation was not identified in this study, the future identification of the causative gene may provide insight to understanding the pathophysiology of heart failure.

CHAPTER 2

Materials and methods

Index	Page
2.1 General	55
2.2 Selection of subjects	55
2.2.1 Members of a family in which PFHBII segregates	55
2.2.2 Unrelated subjects from the Caucasian South African Afrikaner subpopulation	58
2.3 Clinical evaluation	58
2.4 Procurement of blood	59
2.5 DNA extraction from peripheral blood	60
2.6 Data capture	61
2.7 Construction of a family pedigree	61
2.8 Genetic analysis	61
2.8.1 Selection of candidate loci	61
2.8.2 Genetic marker information at selected candidate loci	61
2.8.3 STR allele frequencies in the Caucasian Afrikaner subpopulation	63
2.8.4 Calculating STR marker heterozygosity values	63
2.8.5 Selection of genetic markers at candidate loci	63
2.8.6 Synthesis of STR oligonucleotide primers	64
2.8.7 Genotype analysis	64
2.8.7.1 PCR-based amplification of STR alleles	64
2.8.7.2 Gel electrophoresis of PCR-amplified repeat alleles	65

2.8.7.3 Assigning repeat allele genotype scores	65
2.9 Linkage analysis	66
2.9.1 Statistical parameters used for log of the odds (lod) score calculations	66
2.9.2 Fine mapping analysis	68
2.9.3 Haplotype analysis at a PFHBII-linked locus	69
2.10 Construction of a genetic, physical and transcript map	69
2.10.1 Selection of genomic clones spanning the PFHBII locus	69
2.10.2 Isolation of BAC and PAC genomic DNA	70
2.10.3 Isolation of YAC DNA	72
2.10.4 PCR-based mapping of genetic markers onto genomic clone-insert DNA sequence	74
2.11 Bioinformatics analysis	75
2.11.1 Mapping genetic markers onto <i>in silico</i> genomic clone sequences	75
2.11.2 Identifying candidate genes at the PFHBII locus	75
2.11.3 Selecting candidate genes for mutation analysis	78
2.12 Mutation analysis	79
2.12.1 Designing oligonucleotide primers	79
2.12.2 PCR-SSCP amplification conditions	79
2.12.3 PCR-SSCP gel electrophoresis	80
2.12.4 Gel purification of PCR products for sequencing analysis	81
2.12.5 Characterisation of sequence variants	82
2.12.5.1 Cycle sequencing	82
2.12.5.2 Gel electrophoresis of cycle sequencing products	82
2.12.5.3 Automated sequence analysis	83
2.12.5.4 Automated sequence interpretation	83

Materials and methods

No amount of experimentation can ever prove me right; a single experiment can prove me wrong.

- Albert Einstein

2.1 General

The University of Stellenbosch (US) Faculty of Health Sciences Ethics Committee reviewed and granted permission for the study (project number 99/067), according to guidelines set by the WHO/International Society and Federation of Cardiology Task Force (Richardson et al., 1996). Prior to commencement of the investigation, informed oral or written consent was obtained from each of the study participants. Parental or guardian consent was obtained for minors under the age of 18 years. Blood samples were drawn and clinical evaluations were performed on members of the family in which PFHBII segregates. Blood procurement and clinical evaluations were performed by Prof. P. A. Brink (Senior Specialist Internist, Tygerberg Hospital, Cape Town, RSA) or by a medical practitioner chosen by the individual. All acquired clinical data was stored at the Department of Internal Medicine, Tygerberg Hospital (Cape Town, RSA). Extracted deoxyribonucleic acid (DNA) (see section 2.5) was stored at the US/Medical Research Council (MRC) Centre for Molecular and Cellular Biology (Cape Town, RSA).

2.2 Selection of subjects

2.2.1 Members of a family in which PFHBII segregates

Individuals entered into the study were members of a family previously described by Brink and Torrington (1977). Twenty-six Caucasian South African Afrikaner subjects from the PFHBII-affected family were entered into the study (Fig. 2.1), of whom 13 individuals had been clinically examined in the original investigation by Brink and Torrington (1977).

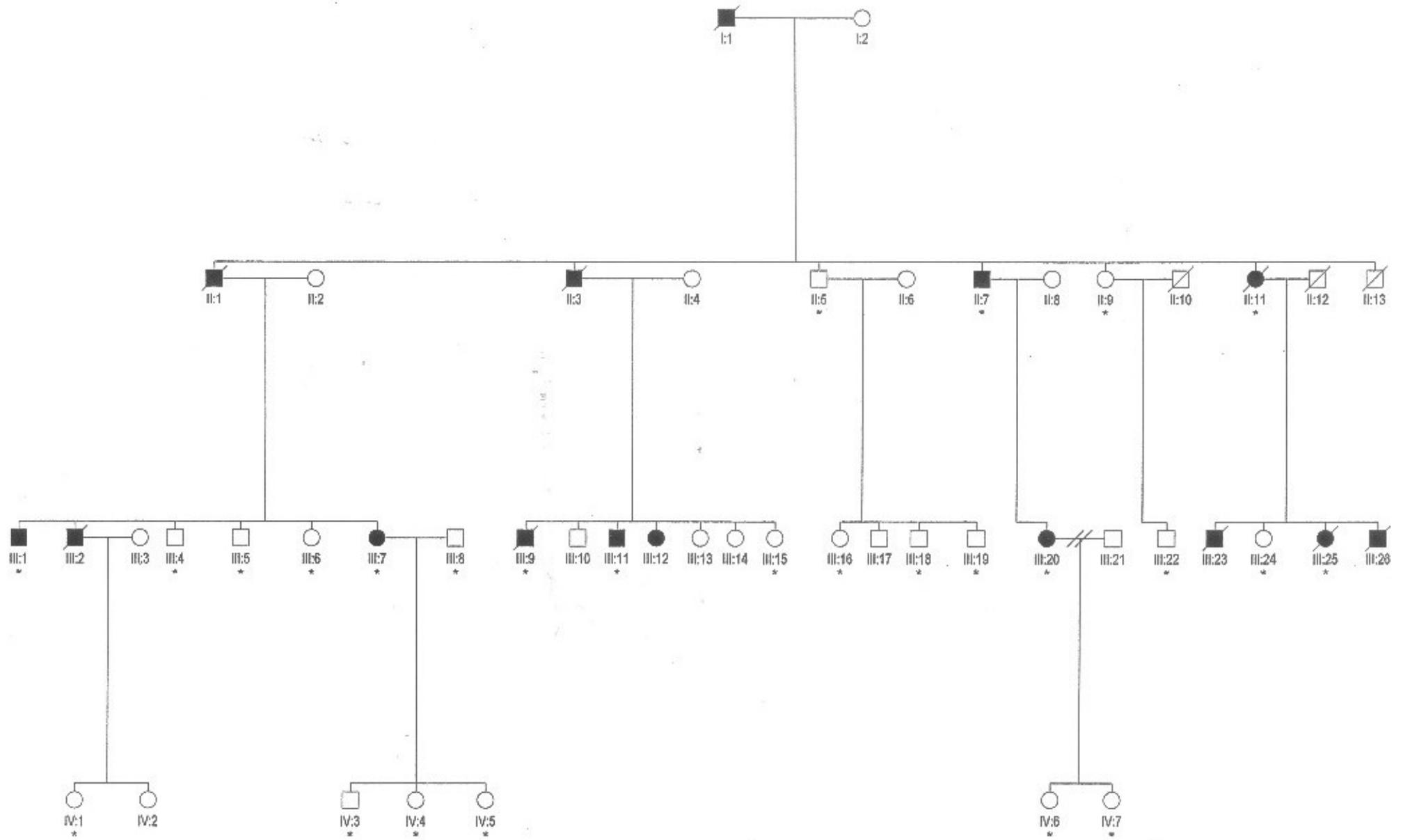


Fig. 2.1**The South African PFHBII-affected kindred.**

The multigeneration pedigree shows the PFHBII-affected family of Caucasian South African Afrikaner descent. An asterisk () indicates subjects that were entered into the present study and filled symbols indicate individuals that have been diagnosed with or have a clinical history of PFHBII. Individuals II:5, II:7, II:9, II:11, III:9, III:15, III:16, III:18, III:19 III:20, III:22, III:24 and III:25 were clinically assessed in the original study by Brink and Torrington (1977). The clinical status of all subjects from the PFHBII-affected family is provided in section 3.1.*

2.2.2 Unrelated subjects from the Caucasian South African Afrikaner subpopulation

To establish genetic marker *allele* frequencies, a panel of 30 anonymous Caucasian South African Afrikaner individuals with no clinical history of PFHBII was entered into the study. The same panel of unrelated subjects was used to estimate the frequencies of identified sequence variations (see section 2.12) in the South African Caucasian Afrikaner subpopulation.

2.3 Clinical evaluation

General assessments

Subjects were evaluated using previous clinical history, physical examination, 12-lead ECG, M-mode and 2-dimensional echocardiography. Diagnosis was based on evidence of cardiac conduction disease and/or the presence of DCM (Brink and Torrington, 1977; Fernandez et al., 2004). ECG diagnosis of affected individuals was defined by the presence of SB (a heart rate of less than 50 beats per minute [bpm]), LAHB/LPHB, first-degree heart block or CHB with narrow QRS complexes (Brink and Torrington, 1977; Fernandez et al., 2004). Minimum criteria for an echocardiograph diagnosis of DCM were an ejection fraction of less than 45% and a left ventricular end-diastolic diameter of greater than 5.6 centimetres (cm) in the absence of hypertension, ischaemic or valvular heart disease. All echocardiographic measurements were made according to the American Society of Echocardiography guidelines (Sahn et al., 1978). In some cases, family members were re-examined because their previous clinical assessment had been incomplete or had been performed more than five years prior to the start of the present study. All assessments of clinical records were performed by Prof. P Brink.

Clinical data used to establish a clinical profile of PFHBII

Clinical records of family members who opted not to give blood for the genetic linkage study were obtained, with the patients' permission, from their personal physician, or, alternately, the subject provided his/her own clinical details (A Goosen, personal communication). The clinical histories of

deceased members of the PFHBII-affected family were obtained by interviewing living first- or second-degree relatives in person or telephonically. Clinical histories of spouses of clinically affected family members were obtained in order to exclude the possibility that these subjects were clinically affected (with PFHBII), or possibly, that they presented with another familial cardiac-associated defect that could have a pathological effect on their progeny. In general, individuals with a clinical history of a pacemaker implantation or an indication of heart transplantation were assigned a clinically affected status. Clinical data for the 26 subjects that were genotyped for the linkage study were obtained by the method described in section 2.1.

Clinical designations for linkage analysis

A conservative phenotypic classification approach was used for linkage analysis, to avoid misclassifying subjects from the family in which PFHBII segregates. Progressive slowing of the heart rate may be an indication of underlying atrioventricular disease (Talajic et al., 1991). Therefore, individuals with a heart rate between 50-55 bpm were assigned a clinically uncertain status. Individuals with a previous clinical history of the disease, but for whom the clinical data could not be obtained, or for whom a second clinical examination indicated that they did not meet the criteria for classification as clinically affected, were designated clinically uncertain. These criteria were included because of concerns of setting a clinical diagnosis based on a single ECG examination. Due to the late onset of the disorder, individuals younger than 20 years were not clinically investigated and were assigned a clinically uncertain status. All clinical assignments were made prior to *genotype* allocation.

2.4 Procurement of blood

Blood was collected from family members through venous puncture of the ante cubital fossa of the forearm. Nine millilitres (ml) and 4.5ml of blood were respectively collected in a heparin tube (Vacutainer, UK) and an ethylenediamine tetraacetic acid (EDTA) tube (Vacutainer, UK).

2.5 DNA extraction from peripheral blood

DNA was extracted from EDTA-treated peripheral blood (see previous section) using the method previously described by Corfield et al., (1993) – a list of the solutions used to treat blood and for DNA extractions is provided in Appendix A. The DNA was then resuspended in 200 μ l of distilled water (dH₂O) in a 1.5ml eppendorf tube (Whitehead Scientific, RSA) and mixed at 4°C for 16 hours in a Voss rotor (Malden, UK). Ten microlitres of the resuspended DNA was aliquoted into a clean eppendorf tube, diluted with 490 μ l of dH₂O and dispensed into a quartz cuvette (Laboratory and Scientific Equipment Company [Lasec], RSA). The cuvette containing the diluted DNA was placed into a Spectronic 1201 spectrophotometer (Milton Roy, UK) that was set at a wavelength of 260 nanometers (nm) and the absorbance reading was recorded. The concentration of the extracted DNA in micrograms per microlitre (μ g/ μ l) was determined by multiplying the absorbance reading by 2.5.

Establishing cell lines from blood

Approximately 5ml of heparin treated blood (section 2.4) was used to establish lymphoblastoid cell lines using the method described by Neitzel (1986) – a list of the solutions used to establish cell lines is provided in Appendix A. On completion of the growth phase, the cells were transferred to 2ml Nunc tissue culture tube (Sterilab, RSA) and kept in a polystyrene container lined with cotton wool, which was placed at -70°C overnight in an ultra low temperature refrigerator (Revco, USA). After 16-20 hours, the tubes were taken out of refrigerator and were placed in a liquid nitrogen Union Carbide model Bx186 freezer (Washington, USA) and stored until required.

2.6 Data capture

All blood samples that were included for DNA extractions were indexed in the computer database programme, Lab Management (Microsoft Access, LABdata.mdb, Soft Craft Systems 1996-8). The programme stored names, dates of birth, sample numbers, project names and dates of entry. Cell lines were given the same accession number as blood samples, but were stored in a separate file in the database. Only authorised personnel were permitted to access the data and all information was confidential. No commercial gain was derived from the study and all data was used solely for research purposes.

2.7 Construction of a family pedigree

Genealogical data and clinical information of members from the family in which PFHBII segregates were entered into the Cyrillic 2.10 programme (Cherwell Scientific, UK) and used to construct a multigeneration family pedigree (Fig. 2.1).

2.8 Genetic analysis

2.8.1 Selection of candidate loci

Disorders associated with DCM and cardiac conduction defects, isolated DCM, HCM, isolated cardiac conduction defects and a cardiac arrhythmia, which map to chromosome 1p1-1q21, 1q32, 2q14-q22, 3p22-p25, 6q23, 7q35-q36, 9q12-q13, 11p11, 14q12, 15q11-q14, 15q22 or 19q13.3 (Table 2.1) (refer to Fig. 1.2, 1.5 and 1.7), were selected for a candidate loci genetic linkage study. For the linkage study, short tandem repeat (STR) markers at the selected candidate loci were used in polymerase chain reaction (PCR)-based genotype analyses (section 2.8.7).

2.8.2 Genetic marker information at selected candidate loci

Information for the genetic map locations of STR markers at candidate loci (Table 2.1) was obtained from the Marshfield Genetic Map Marker information database (http://research.marshfieldclinic.org/genetics/Map_Markers/maps/IndexMapFrames.html) and the Genome Database (<http://www.gdb.org>).

Table 2.1

Primer sequences and annealing temperatures of selected STR markers used in a candidate loci linkage analysis of members of the PFHBII-affected family

Candidate locus	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	T _m (°C)
Chromosome 1p1-1q21			
<i>DIS305</i>	CCAGNCTCGGTATGTTTTACTA	CTGAAACCTCTCTGTCCAAGCC	55
<i>DIS2702</i>	CCCCTTGGCATAGGGTTCAAGA	AGCCAGGCATCTGCACCTTC	55
<i>DIS176</i>	CCATAATTGTGTGTACAAC	AGTTCTTCAGAGAAACAGGA	55
Chromosome 1q32.2-q32.3			
<i>DIS456</i>	CTTCAGGATCAGGGACAACT	GCCACATACCTCCTTACCTTG	55
<i>DIS491</i>	ATCTCCCTGGCAATGATTTT	TGTCAGCAAGAANTGTGAAAGT	55
<i>DIS70</i>	TCTCGTTCCCTCCCCTGTCT	TCACGCTGAGCTCTGTTTTA	55
<i>DIS3753</i>	CATGTAATTCAGGTTAGCATGA	GAGCGATGATAGTGCCACTAC	55
<i>DIS205</i>	CTGAGCACAGCAGTGGTCTC	AAGGCTTATCAAGAGCGAGG	55
<i>DIS414</i>	GCACAGTTCAACATCCATT	TCTCTGTCATTTTAGGTCTATTTCT	55
<i>DIS2810</i>	TCAGTTCTGGTTACCGTTA	CCGTTCTTATGAGCCTTACA	55
<i>DIS2780</i>	TCTGTGTTCTAAGTATTA AAAAACC	ATCCAACCCTGACAAC	55
<i>DIS425</i>	CTAAAAACAAAACACTCAGAACCACA	ACCTAGCACAAAGGCTTCACA	55
<i>DIS505</i>	GGAAGGCTTGCTCTTG	CCTCACCCATGACACTG	55
<i>DIS217</i>	TTATAACCAACCCTGTCACAG	GCACATTTCCGGTAAAAGAA	55
Chromosome 2q14-q22			
<i>D2S2224</i>	AACTCCATCCC GGTC	ATGTGTGCCTGTGGGTTTAG	55
<i>D2S2339</i>	TGCCTTGAGGACCACT	TGCAATGTAACATAATTTGTACTAA	55
<i>D2S2215</i>	ACCTGGACTCAGAAAAGTT	TGCTGTGCCTAAAAGTATGC	55
<i>D2S112</i>	GAGTGGCGGTGAGAAGGTAT	AGCCATGCTATCTTTGAGG	55
Chromosome 3p22-p25			
<i>D3S1263</i>	CTGTTGACCCATTGATACCC	TAAAATCACAGCAGGGGTTTC	55
<i>D3S2303</i>	TTTCTGCCCTGCCTACATG	TCAGAATCACCCACAAGGG	55
<i>D3S1211</i>	CAGGGCTTGTGGGATTAGAA	ATTCAGATTTTCAGGACAAGGG	55
Chromosome 6q23			
<i>D6S262</i>	ATTCTTACTGCTGGAAAACCAT	GGAGCATAGTTACCCTTAAAATC	55
Chromosome 7q35-q36			
<i>D7S505</i>	ACTGGCCTGGCAGAGTCT	CAGCCATTTCGAGAGGTGT	55
Chromosome 9q12-q13			
<i>D9S152</i>	CCAGTTTTTACTCTCTTCTCTC	GTTATGTTAGCTTTCTAAATGGCTG	55
Chromosome 11p11			
<i>D11S1344</i>	CCCTGAACTTCTGCATTCAC	GCGCCTGGCTTGTACATATA	55
Chromosome 14q12			
<i>*MYH7int24</i>	CATCCTCTAACCCTACCC	TTGAGTTGTGGGGCCTAG	55
Chromosome 15q14			
<i>ACTC</i>	TCAGAGACAAATGGTGACAGC	GACATTGTTGGCATAACAGG	55
<i>D15S108</i>	AGGAGAGCTAGAGCTTCTAT	GTTTCAACATGAGTTTCAGA	55
Chromosome 19q13.3			
<i>D19S412</i>	GAATGAGACTCTGTGTCTCAAAAACA	ACCTCATGTGTCTCCTCCTT	52
<i>D19S902</i>	CCATCCTAATGAGGGCAA	GCACCAGTGACTGCCTGT	50
<i>D19S596</i>	GAATCCGAGAGGTGGG	GCCAGAGCCACTGTGT	50
<i>D19S246</i>	AGAGTGAGATTCCACCTTTC	GAAACACATCATTACCCAC	52

* = genetic marker supplied by Dr. J. Moolman-Smook (University of Stellenbosch, RSA); °C = degrees Celsius; T_m = annealing temperature. Underlined markers were used in genetic fine mapping analysis (see text, section 2.9.2).

2.8.3 STR allele frequencies in the Caucasian Afrikaner subpopulation

The Centre pour l'Étude du Polymorphisme Humain (CEPH) (now known as Centre Jean Dausset) maintains a database containing information on population STR marker allele frequencies that have been obtained by genotype analyses in reference families of Northern European descent (<http://www.cephb.fr/cephdb>). However, for the present study, it was speculated that the STR allele distributions and allele frequencies in Caucasian South African subpopulations may be different from the Northern European reference families. Therefore, STR allele frequencies were determined by genotype analysis of 30 unrelated ethnically matched South African Caucasian Afrikaner individuals (section 2.2.2). The observed allele frequencies were then used to calculate the heterozygosity values of the STR markers at chromosome 1q32.2-q32.3 (Table 2.1).

2.8.4 Calculating STR marker heterozygosity values

The allele frequencies for each STR marker were used to calculate the marker heterozygosity value using the equation $H = 1 - \sum f_{rs}^2$, where H is the observed heterozygosity value and $\sum f_{rs}$ is the sum of the allele frequencies at that marker (Ott J, 1999). Heterozygosity values were then used as a criterion to select the STR markers that would be used in linkage analysis study in the PFHBII-affected family (see section 2.9.1).

2.8.5 Selection of STR markers at candidate loci

Genetic markers with little variation in frequency of distribution of the individual alleles (determined in the panel of 30 Caucasian Afrikaner subjects) (see section 2.8.3) and with an observed heterozygosity value of equal to or greater than 0.60, were selected for genotype analysis of subjects from the PFHBII-affected family (Table 2.1). Allele frequencies (see Appendix and heterozygosity

values for the STR markers at chromosome 1q32.2-q32.3 are available in at the website of the US/MRC Centre for Molecular and Cellular Biology and the Department of Medical Biochemistry (Cape Town, RSA), http://www.sun.ac.za/Internet/Academic/Health/schools/Basic_appl_health/med_physbio/med_biochem/dept/Research.html#pfhb2. For the present study, population allele frequencies were only determined at chromosome 1q32.2-q32.3 (see Appendix E), because using CEPH allele data, or setting the population allele frequencies to equal, did not change the strongly negative lod scores that had been generated at the other investigated candidate loci (data not shown).

2.8.6 Synthesis of STR oligonucleotide primers

Oligonucleotide primer sequences of the STR markers selected for linkage analysis (Table 2.1) were obtained from published material and the *National Centre for Biotechnology Information (NCBI)* (<http://www.ncbi.nlm.nih.gov>), Sanger Centre (<http://www.sanger.ac.uk>), Lawrence Livermore (<http://www.bio.llnl.org>) or the Genome Database (<http://www.gdb.org>) web sites. The primers used in the PCR-based genotype analysis (Table 2.1) were synthesised at the DNA synthesis laboratory (Department of Biochemistry, University of Cape Town, Cape Town, RSA) or Research Genetics (Huntsville, Alabama, USA).

2.8.7 Genotype analysis

2.8.7.1 PCR-based amplification of STR alleles

The selected STR markers (Table 2.1) were used in a PCR-based amplification reaction to identify the segregation of tandem repeat alleles in members of a family in which PFHBII segregates. Genomic DNA obtained from 26 members of the PFHBII-affected family (section 2.2.1) and the Caucasian Afrikaner control panel of 30 subjects (section 2.2.2) was used as a template for separate PCR amplification reactions with the STR markers listed in Table 2.1. The amplification was performed in

10µl reactions containing 200 nanograms (ng) genomic DNA, 50ng each of the forward and reverse primers, 1 microcurie (µCi) [α - 32 P]-dCTP (Amersham International, UK), 25 micromolar (µM) unlabelled dCTP, 200µM of each dATP, dGTP and dTTP, 1.5mM MgCl₂, 1µl 10x Taq DNA polymerase buffer (Promega, USA) and 0.5 units (U) Taq DNA polymerase (Promega, USA). Standard parameters in a Techne GeneE thermal cycler (Cambridge, UK) for 30 cycles were as follows: denaturing at 94°C for 30 seconds, annealing at 55°C for 30 seconds, except *D19S412* and *D19S246* (52°C) and *D19S902* and *D19S596* (50°C), and extension at 72°C for 30 seconds (Table 2.1).

2.8.7.2 Gel electrophoresis of PCR-amplified repeat alleles

The PCR amplification products (see previous section) were mixed with 6µl of loading dye (Appendix A) and heat-denatured at 80°C for 2 minutes in a heating block (Techne Dri-block, UK), after which 3µl of the denatured sample was loaded onto a 6% denaturing polyacrylamide gel (Appendix A). Gel electrophoresis was performed for 2-3 hours at 1800 volts (V) (Merck, GmbH) (Appendix A). The electrophoresis apparatus was then dismantled and the gel carefully lifted from the plate using 3MM Whatmann paper (Merck, GmbH) that was cut to the same size as the glass plate (see Appendix A for glass dimensions). The gel was covered with cling wrap (GLAD™), dried on a gel drier (Drygel Slabgel Sr model se1160, Scientific Associates, RSA) for 1 hour and exposed to an X-Ray film (Protea Medical, RSA) in an autoradiograph cassette. Once exposed, the film was developed in a dark room in developer solution for 2 minutes, stop solution for 30 seconds and fixer solution for 1 minute (Appendix A).

2.8.7.3 Assigning repeat allele genotype scores

PCR amplification analysis and gel electrophoresis provides a means to identify the differences in paternally and maternally inherited repeat allele lengths at a given locus (Weber, 1990). Based on the

principle of size separation, larger sized repeat lengths will migrate more slowly through a gel matrix, whereas shorter repeat lengths migrate faster and will be positioned further from the wells into which they were loaded. The PCR-amplified repeat alleles analysed by gel electrophoresis (section 2.8.7.1 and 2.8.7.2) generated radioisotope signals at different positions on an autoradiograph. The radioisotope signals closest to the well (representing the largest repeat allele length) was assigned a value of 1 and all consecutive signals were assigned an increasing numerical value. The numerical values were then used to allocate genotype scores for each individual. The numerical values corresponding to two radioisotope signals with the strongest intensities were used as the genotype score for an individual. However, final genotype scores were assigned only after consensus was reached between two independent investigators who were blinded to the clinical status of the individual.

A sample that did not produce a radioisotope signal and, therefore, could not be assigned a genotype score was re-amplified and re-assessed by gel electrophoresis and autoradiography analysis. For the subsequent analysis, an amplification product that had previously been assigned a final genotype score served as a standard against which successive sample analyses were scored.

2.9 Linkage analysis

2.9.1 Statistical parameters used for log of the odds (lod) score calculations

Linkage computations

Pairwise (or two-point) and multipoint linkage analysis was performed using the personal computer versions of the MLINK and LINKMAP options of the LINKAGE package (version 5.2) (Lathrop and Lalouel, 1984). All lod score calculations were performed using a computer workstation with a Pentium III processor (Intel®, USA).

Setting population allele frequencies

Allele frequencies at each STR marker on chromosome 1q32.2-q32.3 were determined in a Caucasian South African Afrikaner panel, as described in section 2.8.7. The allele frequency in the Caucasian Afrikaner population was then calculated using the equation n divided by N , where n is the number of times a specific allele was observed and N is the total number of chromosomes that were analysed.

Setting disease allele frequency

The present study assumed that the disease allele was rare, since PFHBII has only been identified in one South African family (P. Brink, personal communication). Therefore, lod score calculations were performed with the frequency of the normal and disease allele assumed to be 0.9999 and 0.0001, respectively (P.A. Brink, unpublished data).

Age-dependent penetrance

The manifestation of PFHBII-associated clinical features is presumed to be age-dependent (Fernandez et al., 2004 and the present study). Therefore, clinical data obtained from members of the PFHBII-affected family were analysed to determine the percentage of individuals that presented with disease symptoms within specific ranges in age (10 year age-ranges were selected for the present study). This data was then used to calculate a disease penetrance for clinically unaffected individuals who were within a specific age-range and who carried the disease-associated haplotype. For this study, the sole unaffected individual with the disease-associated haplotype, who was included in the lod score calculation (individual IV:3), was assigned a liability (penetrance) class of 40%.

Multipoint analysis

By testing a disease-marker relationship, pairwise lod score calculations provide an indication that a disease maps near a particular marker (Ott, 1999). However, limited marker informativeness may reduce the statistical power of a pairwise linkage study (Lathrop and Lalouel, 1984; Ott, 1999). By testing the relationships between a disease and a series of markers that map close to each other, multipoint analysis minimises the influence of limited marker informativeness (Lathrop and Lalouel, 1984). For the present study, multipoint analyses were performed using genotype data from the PFHBII-affected family and intermarker distances between sets of successive markers. The information was entered into the LINKMAP programme, which generated location scores between the disease and a marker-set. Multipoint scores for a specific marker set were then calculated by dividing the location scores by 4.6 (the divisor factor is derived from the equation $2\ln [10] = 4.6$) (Lathrop and Lalouel, 1984). The calculated multipoint scores for each marker set were plotted against the genetic distances (in *centimorgan* [cM]) between markers at the locus. Thereafter, the maximum obtained lod score minus one unit (“1-lod-down”) analysis provided a 90% estimate of the genetic interval in which causative gene was likely to be located (Ott, 1999).

2.9.2 Fine mapping analysis

For a single linkage test, a lod score of +3 or more (1000:1 odds) indicates a disease is linked to a locus, whereas a score of -2 or less excludes a locus as cause of a disease (Ott, 1999). Pairwise linkage analysis was initially performed with selected markers at candidate loci (Table 2.1). If a marker was identified where a maximum lod score of greater than 1.5 was generated (in the present study it was markers at chromosome 1q32.2-q32.3), then genotype analysis proceeded with additional informative markers at that locus (fine mapping) (Table 2.1). Pairwise and multipoint lod score calculations were then assessed using the new set of markers.

The markers used for fine mapping analysis were also selected because they flanked identified *recombination* events, which were identified after haplotype construction with the STR marker data from the initial lod score calculations (see section 2.9.3).

In order to test the stability of the obtained lod values, the allele frequencies were reset to equal or the disease penetrance was varied (Hodge and Greenberg, 1992). Additionally, as linkage is dependent on robust clinical specifications, the obtained lod scores were re-calculated by changing the clinical status of one affected family member (individual III:11, Fig. 2.1). The reasons for selecting this particular subject are provided in section 3.2.2.

2.9.3 Haplotype analysis at a PFHBII-linked locus

Fine mapping analysis creates a landscape of markers that are closely linked and have well-defined chromosomal positions. For the present study, haplotypes were constructed manually taking into account the parental distribution and phase of each marker allele. It was opted not to use the Cyrillic 2.10 programme (Cherwell Scientific, UK) to construct the haplotypes because of the computational limitations of the software, namely, the inability of the software to compute phase when one of the parental alleles was absent (due to technical difficulties).

Initially, haplotypes were constructed at a locus on chromosome 1q32.2-q32.3 where preliminary analysis generated a lod score of greater than 1.5. Fine mapping analysis was then performed (see previous section), after which haplotypes were re-constructed by including the additional STR marker genotype scores.

2.10 Construction of a genetic, physical and transcript map

2.10.1 Selection of genomic clones spanning the PFHBII locus

Agar stabs of Bacterial Artificial Chromosome (BAC), P1 Artificial Chromosome (PAC) and Yeast Artificial Chromosome (YAC) vectors, containing fragments of genomic sequence derived from the locus on chromosome 1q32.2-q32.3, were obtained from the Wellcome Trust Sanger Centre (Hinxton, UK) and Research Genetics. Based on information obtained from the Sanger Centre chromosome 1 database (<http://www.sanger.ac.uk/HGP/Chr1>), ten genomic clones that spanned the PFHBII locus were selected for genetic, physical and transcript mapping analysis (Table 2.2) (Fig. 2.2).

Table 2.2

Genomic clones used in the construction of a genetic, physical and transcript map of the PFHBII locus

Clone	Size (base pairs)
CTA-55I10	90 429
RP1-272L16	157 875
RP1-28O10	97 700
RP3-434O14	135 928
CTA-447D17	26 524
CTA-321I20	19 974
RP4-782D21	84 464
CTA-11C7	109 302
RP5-879K22	137 678
YAC 937D4	~2 000 000

2.10.2 Isolation of BAC and PAC genomic DNA

An inoculating loop (Sterilab, RSA) was inserted once into the agar stab and streaked across a Luria broth agar (LA) plate (Appendix A) containing chloramphenicol (Roche, GmbH) at a final concentration of 12.5 µg/ml. The LA plate was then incubated at 37°C for 12-16 hours in a Sanyo Incubator (Vacutec, RSA). Thereafter, isolations of large clone inserts using alkaline lysis were performed (Sambrook et al., 1989). To this end, a single colony was picked from the streaked LA plate using a toothpick and inoculated into 500ml of LB broth (Appendix A) in an Erlenmeyer flask containing chloramphenicol (12.5µg/ml). The culture was incubated for 16 hours at 37°C with shaking in an orbital shaker/incubator model LM530 (YIH DER, RSA), after which it was transferred to

Sorvall centrifuge bottles and centrifuged at 4000rpm at 4°C for 15 minutes in a Sorvall HB4 rotor to harvest the cells.

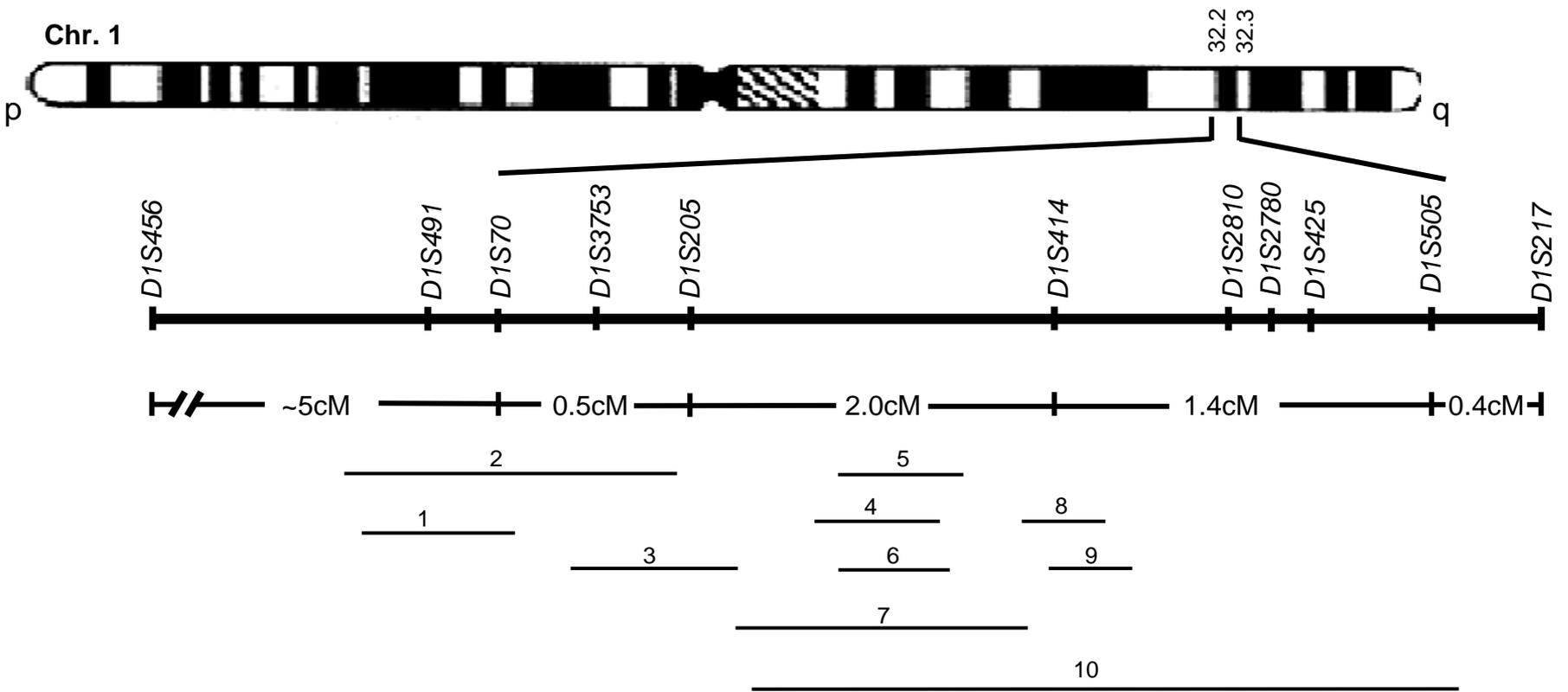


Fig. 2.2

Genetic markers and genomic clones that span the PFHBII locus on chromosome 1q32.2-q32.3.

Genetic markers and intermarker distances in centimorgans (cM) were obtained from genomic databases (see text). The BAC, PAC and YAC clones that were selected for genetic, physical and transcript mapping at the PFHBII locus are numbered: 1 = CTA-55I10; 2 = RP1-272L16; 3 = RP1-28O10; 4 = RP3-434O14; 5 = CTA-447D17; 6 = CTA-321I20; 7 = RP4-782D21; 8 = CTA-11C7; 9 = RP5-879K22; 10 = YAC 937D4.

The relative clone positions are based on draft Sanger Centre sequence data that was available at the time (June 2000).

The supernatant was then decanted and the cells were resuspended in 10ml of Solution I (Appendix A) and 20ml of Solution II (Appendix A), by inverting the centrifuge bottle several times, which was then allowed to stand at room temperature for 10 minutes. Thereafter, 15ml of ice-cold Solution III (Appendix A) was added to the cell suspension, the bottle was inverted several times, allowed to stand for 10 minutes on ice and then centrifuged at 4000rpm in a Sorvall HB4 rotor for 15 minutes at 4°C, which resulted in the cellular debris being pelleted at the bottom of the centrifuge bottle. The supernatant was transferred from the centrifuge bottle to a clean 50ml Evergreen tube (Lasec, RSA), to which 0.5 volumes of isopropanol (Merck, GmbH) was added to precipitate the genomic DNA. The suspension containing the precipitated DNA was centrifuged for 6 minutes at full speed at room temperature in a Beckman bucket centrifuge model TJ-6 (rotor diameter of 35cm) (Beckman, USA). The supernatant was discarded and 500µl of 70% EtOH was added to the pelleted DNA in the Evergreen tube, which was then resuspended by inverting the tube several times. The resuspended DNA was centrifuged at 4000rpm at room temperature for 5 minutes, after which the EtOH was discarded and the DNA pellet was allowed to dry at room temperature. The dried DNA pellet was resuspended in 100µl TE (Appendix A), transferred to a clean 1.5ml eppendorf tube and stored at 4°C.

2.10.3 Isolation of YAC DNA

An inoculating loop was inserted once into the agar stab containing copies of a YAC genomic clone, then streaked onto a YPD plate (Appendix A) and incubated at 30°C for 1-2 days. One colony was picked with an inoculating loop, re-streaked on another YPD plate and incubated at 30°C for a further 1-2 days. A single colony was then picked and inoculated into 10ml of YPD medium and incubated overnight at 30°C in an orbital shaker/incubator. The solution containing the cells was then poured into a centrifuge bottle, after which the cells were pelleted by centrifugation at 1000 rpm for 10 minutes in a Sorvall centrifuge and washed twice in 10ml of 40 millimolar (mM) EDTA/90mM 2-mercaptoethanol (Appendix A).

Thereafter, the supernatant was decanted and approximately 700µl of lysis buffer (Appendix A) was added to the lysed cells, which were then incubated at 68°C for 15 minutes. The lysate was resuspended by vortexing and then split into equal volumes in two eppendorf tubes. Six hundred microlitres of phenol:chloroform:isoamyl alcohol (PCI) (Sigma, USA) was added to each tube containing the resuspended lysate, which was then centrifuged at 12 000rpm for 10 minutes. The aqueous phase of both suspensions was pooled in a clean eppendorf tube to which 700µl of isopropanol was added. The solution was incubated at room temperature for 15 minutes and centrifuged at 12 000rpm for 15 minutes, after which the supernatant was discarded and the pellet was resuspended by vortexing and then washed with 500µl of 70% EtOH, as previously described in section 2.10.2. The DNA pellet was air-dried, resuspended in 300µl of TE and 30µg RNase (0.05 U/mg) (Roche, GmbH) and incubated at 37°C for 2 hours with intermittent vortexing to dissolve the pellet.

Seven hundred microlitres of isopropanol was then added and the mixture was incubated at room temperature for 15 minutes. The DNA, which had now precipitated in the isopropanol, was centrifuged at 12 000rpm for 15 minutes in a Beckman Microfuge® Lite bench top centrifuge (Beckman, USA) and subsequently washed twice with 500µl of 70% EtOH as previously described. The washed pellet was air-dried and resuspended in 300µl of TE and 750µl EtOH, stored on ice for 10 minutes and then centrifuged at 12 000rpm for 15 minutes. An additional 70% EtOH wash was performed as previously described after which the pellet was allowed to air-dry. The DNA pellet was resuspended in 200µl TE and stored at 4°C.

2.10.4 PCR-based mapping of genetic markers onto genomic clone-insert DNA sequence

Short tandem repeat markers were mapped onto genomic clones (Table 2.2) using a PCR-based assay, to determine whether the selected clones contained sequence that 1) spanned the PFHBII-linked locus and 2) were *contiguous* (overlapping) across the described disease interval. The 11 genetic markers at chromosome 1q32.2-q32.3 (Table 2.1) were used in an amplification assay that was similar to the method described in section 2.8.7.1, except that the PCR reaction mix contained 0.5 μ l of BAC, PAC or YAC DNA as genomic template, no radioactive isotope was added and 200 μ M of each deoxynucleotide (dATP, dCTP, dGTP and dTTP) was added.

Once the PCR amplification reaction was complete, a 1.5% gel was prepared that consisted of 0.75 milligrams (mg) agarose (Whitehead Scientific, RSA) mixed in 50ml of 1xTBE in a glass beaker (Schott, GmbH) (Appendix A) and placed in a microwave oven for 2 minutes to dissolve the agarose. Thereafter, 10mg/ml of ethidium bromide (EtBr) (Sigma, USA) was added and the solution was then poured into a Scigen casting tray (Whitehead Scientific, RSA). At one end of the casting tray, a well-forming comb was positioned into the gel. Once the gel was set, the comb was removed and the gel was placed in a Scigen electrophoresis tank (Whitehead Scientific, RSA) into which 150ml of 1xTBE was added. Thereafter, 5 μ l of the PCR amplification product (see previous paragraph) was mixed with 2 μ l of loading buffer (Appendix A) and loaded into a well. Electrophoresis then proceeded at 100 volt (v) for 20 minutes, after which the gel containing the amplified product was analysed using a 3UVTM transilluminator model LMS-26E ultra violet light source (Upland, USA) set at a wavelength of 302nm. A permanent photographic record of the gel analysis was obtained using a Sony Graphic System UP-860CE system (Sony, RSA).

2.11 Bioinformatics analyses

2.11.1 Mapping genetic markers onto *in silico* genomic clone sequences

Electronic versions of genomic clone-insert sequences (Table 2.2) and the selected STR markers (Table 2.1) at the PFHBII locus were obtained from the Sanger Centre chromosome 1 and the NCBI databases. *In silico* (computer-based) mapping analyses were then performed via the Internet (US internet service provider).

The electronic versions of the selected genomic clone-insert sequences and STR markers at the PFHBII locus were aligned to each other using the “Blast 2 sequences” option of the *BLAST* algorithm, which is available at the NCBI web address, <http://www.ncbi.nlm.nih.gov/BLAST>. The procedure for the alignment analysis is as follows: the Blast 2 sequences web page has two sequence input windows. The sequence of a genomic clone-insert was retrieved, using the “copy” option, from a database and inserted, using the “paste” option, into the “subject” window in the Blast 2 sequences web page. The sequence of the STR markers at chromosome 1q32.2-q32.3 (Table 2.1) was inserted into the “query” window using the “copy and paste” options described previously. The information was then submitted and the alignment results were returned in a separate web page.

2.11.2 Identifying candidate genes at the PFHBII locus

Using similarity searches to identify new genes

After mapping the PFHBII-causative gene to within a locus on chromosome 1q32.2-q32.3, the search for candidate genes proceeded in June 2001. In 2000, Schutte et al. reported that they had generated an extensive physical and transcript map of the Van der Woude syndrome (VWS) (MIM 119300) locus on chromosome 1q32-q41. However, later analyses of HGP databases indicated that their map did not extend beyond chromosomal band 1q32.2 (Ensembl database, Jan 2001). Initially, the present study set out to incorporate Schutte and colleagues’ mapping data of chromosome 1q32, in addition to

generating additional mapping and annotation data at the locus. However, because *in silico* sequence and annotation data of the region were available at the time (Sanger Centre chromosome 1 database, June 2001), it was opted to use HGP-generated transcript maps and similarity analyses (bioinformatics) to identify and select genes at the PFHBII locus.

The similarity analyses were performed using the “nucleotide” option of the BLAST programme. Electronic versions of genomic clone-insert sequences (see section 2.11.1) were placed into the “query” window (“copy and paste” option, see section 2.11.1) and the BLASTN search option was selected. This option queries a nucleotide sequence against a nucleotide database. The influence of low complexity or sequence repeat regions was minimised by selecting the “low complexity filter” option (also known as “masking”) (see BLAST webpage). This option was included because low complexity regions are sequences that code for transmembrane helices or coiled-coiled domains, which are so abundant in eukaryotic genomes that they can cause false-positive alignments (Huynen et al., 1998). The query sequence was then submitted and the results were returned in a separate web page.

Using databases to identify the positions and functions of annotated genes at the PFHBII locus

The Ensembl, the University of California Santa Cruz (UCSC) (<http://genome.ucsc.edu>) and NCBI databases provided an *in silico* catalogue of the genomic position, structure(s) and the function(s) of known genes and their protein products. Known genes are defined as transcripts identified by experimental analysis, which are expressed *in vivo* (Stein, 2001). By comparison, computer predicted genes are *in silico*-identified sequences that meet criteria for classification as a gene, but for which there is no experimental data to support that it has an *in vivo* biological function. These *in silico* genes are also catalogued in the HGP databases (Burge and Karlin, 1997) and were included in the bioinformatics analysis in the present study. For the present study, all gene annotation data was

obtained from NCBI builds 28 and 31 (24 December 2001 and 5 November 2002 *freeze dates*), respectively (Appendix B).

The similarity analyses identified genes that encode proteins with homology or similarity to protein signatures in which mutations have been described that cause cardiomyopathies and/or cardiac conduction disorders (see sections 1.3.2.1, 1.3.6.1 and 1.4.4). However, because this approach could result in the identification of a large number of plausible candidate genes, additional bioinformatics alignment analyses was used to shorten the list of possible causative genes.

Comparing levels of functional similarity

Functional similarity analyses use scoring matrices to assess whether an alignment is functionally significant and not purely significant because the sequences are identical (Dayhoff et al., 1978; Henikoff and Henikoff et al., 1992). The BLAST (the “protein-protein” [BLASTP] option), BLAT (<http://genome.ucsc.edu/cgi-bin/hgBlat>), *Pfam* (<http://www.sanger.ac.uk/Software/Pfam>), *BLOCKS* (<http://www.blocks.fhcrc.org>) and *PROSITE* (<http://www.expasy.org/prosite>) algorithms (Appendix C) were used to perform pairwise functional alignment analyses. If the result obtained from two or more of the algorithms differed, then the candidate gene was not included in further analyses. The cut and paste options (described in previous sections) were used to submit protein sequences to the BLAS, BLAT, Pfam and BLOCKS algorithms.

Tissue expression profiling to further prioritise candidate genes

Cardiac tissue expression was a prerequisite in the selection of candidate genes for mutation analysis. The expression profiles of genes were determined by inserting electronic versions of the nucleotide sequence of a candidate gene into the query window of the *Sequence Tag Alignment and Consensus Knowledgebase (STACK)* (<http://www.sanbi.ac.za/Dbases.html>) algorithm and then submitting it to the

“stack_heart” *expressed sequence tag (EST)* database (Appendix D). The expression data of possible candidate genes were additionally queried against the *Unigene* (<http://www.ncbi.nlm.nih.gov/Unigene>) database (Appendix D). This analysis differed from the previous bioinformatics procedure because the Unigene database uses a word or phrase-based input. The gene name or gene accession number was inserted into the “search for” box and then submitted to the Unigene EST database. For both the STACK and Unigene database submissions, the results were returned on the same webpage.

It might be relevant to mention that, although cardiac tissue expression was a requirement for selection of candidate genes, genes that were not expressed in cardiac tissue but that fulfilled all other selection criteria, were also considered for mutation analysis. The selection of these genes will be discussed in section 4.9.

2.11.3 Selecting candidate genes for mutation analysis

Threshold alignment significance levels

The algorithms used in the present study produced alignment probability (P) or Expect values (*E-values*), which indicate the difference between alignment matches that have a mathematical significance and matches that have a biological significance (Altschul et al., 1997; Attwood and Parry-Smith., 1999). A significance threshold level of 95% or greater, for a sequence of 40 nucleotides or more, was applied in this study (Altschul et al., 1997). Although 40 nucleotides represent a short sequence, masking of low complexity regions increased the probability of obtaining a match that had biological significance (Altschul et al., 1997). Generally, a protein alignment P-value closer to 1.0 (one) or an E-value closer to 0.0 (zero) represents a significant score. For the protein alignment analyses, P-values greater than 0.9 or E-values less than 0.01 were set as a significant threshold level for a gene to be considered for mutation screening (Altschul et al., 1997). For STACK expression analyses, a default E-value of 0.999 was set as a cut-off level.

2.12 Mutation analysis

2.12.1 Designing oligonucleotide primers

Non-radioactive PCR-single stranded conformation polymorphism (PCR-SSCP) analysis was performed on all the coding exons of prioritised candidate genes, according to a modified method from Orita et al. (1989). In this study, unique oligonucleotide primers were designed using data from the genomic sequence flanking individual exons and splice junctions of the selected candidate genes (Appendix F). The oligonucleotide primer sequences for the novel gene termed *VWS59* (see section 3.5.1) were obtained from Dr. B. Schutte (University of Iowa, USA) (Appendix F). Each primer's sequence consisted of 18-30 nucleotides, for which the last nucleotide of the primer at the 5' and the 3' end of the exon was positioned at least 10 nucleotides from the start of the exon. For large exons, primers were designed to analyse the exon in a series of overlapping amplification products (Appendix F). The annealing temperature of each primer was calculated using the programme DNAMAN version 4.0 (E. Rybicki, University of Cape Town, RSA). For each primer set, the lesser of the two calculated annealing temperatures minus 2°C was used as the final PCR amplification annealing temperature (Appendix F).

2.12.2 PCR-SSCP amplification conditions

For each exon, genomic DNA of three clinically affected subjects (individuals II:7, III:1 and III:9), two unaffected first-degree relatives (individuals III:4 and III:6) and an unaffected spouse of a member of a PFHBII-affected family (individual III:8) (see Table 3.1 and Fig. 3.1) was analysed by PCR-SSCP analysis (see below). Genomic DNA of 30 Caucasian Afrikaner individuals were also used in subsequent PCR amplification reactions to determine the frequency of the identified sequence variations using PCR-SSCP analysis (section 2.12.3).

The amplification of each coding region of the prioritised candidate genes was performed in a 50 μ l reaction containing 200 nanograms (ng) genomic DNA, 120ng of each of the forward and reverse primers, 200 μ M of each dATP, dCTP, dGTP and dTTP, 1.5mM MgCl₂, 1 μ l 10x Taq DNA polymerase buffer and 0.5 units (U) Taq DNA polymerase. Standard PCR-amplification parameters in a Techne GeneE thermal cycler for 30 cycles were as follows: denaturing at 94°C for 30 seconds, annealing for 30 seconds at the calculated T_m for each primer set (Appendix F) and extension at 72°C for 30 seconds. The amplification product lengths are described in Appendix F.

2.12.3 PCR-SSCP gel electrophoresis

For PCR-SSCP gel electrophoresis analysis, 8% and 10% mildly-denaturing polyacrylamide gel solutions (Merck, GmbH), each containing 5% glycerol and 15% urea, were prepared by the method described in Appendix A.

A 5 μ l aliquot of the PCR-amplified product was placed into a 1.5ml eppendorf tube containing an equal volume of SSCP loading dye (Appendix A) and heat-denatured in a heating block for 3 minutes at 95°C. The denatured samples were loaded into the wells of the PCR-SSCP gel and electrophoresis proceeded at a constant power of 25W for 16 hours at 4°C. The gel apparatus was then dismantled, the gel was silver stained (Appendix A) and air-dried for a minimum of 2 hours to avoid the gel cracking or fading. The gels were analysed and all PCR products showing electrophoresis mobility variations were re-amplified and re-assessed by electrophoresis as previously described. Samples with confirmed mobility variations were purified (see section 2.12.4) and sequenced manually using the fmol® cycle sequencing system (Promega, USA) or by using an ABI-377 automated sequencer (PE Biosystems, USA) (section 2.12.5).

Variant gel-banding patterns were assumed to represent a sequence variation at a specific nucleotide position in the coding or non-coding region of the candidate gene. Subsequent sequence analysis of the samples showing variant gel-banding patterns (section 2.12.5) determined the different nucleotide combinations (or alleles) at a specific position in the sequence. The frequency of a specific allele was then determined by comparing PCR-SSCP electrophoresis patterns of members of the family in which PFHBII segregates against 30 unrelated, unaffected South African Caucasian Afrikaner control individuals. Family members and individuals from the Caucasian Afrikaner panel with the same gel banding patterns were assumed to have the same allele combinations. Allele frequencies were calculated as described in section 2.8.3.

2.12.4 Gel purification of PCR products for sequencing analysis

The GFX™ DNA and gel band purification kit (Amersham Pharmacia Biotech, NJ, USA) was used to purify all samples in which gel mobility variants were identified. Thirty microlitres of the PCR amplification product was loaded into a well on a 1.5% agarose gel and electrophoresis proceeded for 1 hour. The gel was viewed using an ultraviolet light source at 302nm. Thereafter, using a clean scalpel blade, a slice of agarose containing the DNA band was excised and transferred to a clean pre-weighed 1.5ml eppendorf tube and the new weight recorded. Ten microlitres of GFX™ Capture buffer (Amersham Pharmacia Biotech, NJ, USA) was added for each 10mg gel slice (limited to a maximum amount of 300µl). The purification of the PCR-amplified product using the GFX™ DNA and gel band purification kit proceeded according to the manufacturer's instructions.

2.12.5 Characterisation of sequence variants

2.12.5.1 Cycle sequencing

Variant gel-banding patterns for the candidate gene *VWS59* were characterised by direct PCR cycle sequencing of using the *fmol*® sequencing kit. For each variant, the sequencing reaction mix contained 2µl of gel purified amplification product as template DNA, 5µl 5x sequencing buffer, 5% dimethylsulfoxide (Sigma Fine Chemicals, St. Louis, USA), 1 picomole (pmol) of the PCR amplification product oligonucleotide primer and 1µl α -³²P [dCTP] (Amersham, UK). Four microlitres of the above cocktail were added to four separate pre-chilled 0.5ml eppendorf tubes, each containing 2µl aliquots comprised of all four deoxynucleotides (2.5µM concentration stocks of dATP, dCTP, dGTP or dTTP) and a specific dideoxynucleotide (ddATP, ddCTP, ddGTP and ddTTP). The samples were then placed in a pre-heated GeneE Thermal Cycler (Techne Cambridge, UK) at 95°C for 2 minutes, after which, the cycling profile at the original PCR amplification conditions was performed (Appendix F). Four microlitres of sequencing stop solution (Promega, USA) was added to each tube when the PCR reaction was completed.

2.12.5.2 Gel electrophoresis of cycle sequencing products

A 6% denaturing polyacrylamide gel was prepared as described in Appendix A. The PCR-amplified samples (see previous section) were heat-denatured at 95°C for 2 minutes in a heating block, then 4µl aliquots of each sample were loaded into four separate wells and electrophoresis was performed for 2 hours at 1800V. The apparatus was switched off after 2 hours and 4µl aliquots of the same heat-denatured samples were loaded into four separate lanes and electrophoresis proceeded for a further 2 hours. After the electrophoresis had proceeded for a combined time of 4 hours, the gel apparatus was dismantled and prepared for autoradiography as described in Appendix A. The nucleotide sequence

corresponding to the radioisotope signal in each dideoxynucleotide sample lane was read from the bottom end of the autoradiograph towards the top.

2.12.5.3 Automated sequence analysis

The PCR amplification products in which gel mobility variations were identified by PCR-SSCP analysis were sequenced at the US Core Sequencing Facility (Stellenbosch, RSA). Five microlitres of the gel purified amplification product (described in section 2.12.4), which was diluted to a minimum concentration of 3.3ng/μl, was supplied with both forward and reverse PCR primers that were diluted to a concentration of 6.6pmol. Even though the PCR-SSCP technique used in the present study was proposed to have an optimum detection sensitivity equal to and greater than 95% (W de Lange, PhD thesis), a PCR-SSCP quality control test was performed in the present study. To this end, 29 exons that were previously analysed by PCR-SSCP, were subsequently re-screened by direct automated sequencing. For this analysis, a minimum of two PCR amplification products (exon and flanking splice junctions) of each of the seven prioritised candidate genes (a total of 14), which did not demonstrate PCR-SSCP gel mobility variations, was directly sequenced. The remaining 15 of the 29 sequenced exons were PCR-amplified products with lengths of less than 150bp or more than 300bp. These samples were selected because PCR-SSCP analysis has optimal detection sensitivities within the product size range of 150-250bp (Orita et al., 1989).

2.12.5.4 Automated sequence interpretation

Sequence text files (file extension: *.seq) were viewed and edited using DNAMAN version 4.0. The programme Chromas version 2.22 (Technelysium (Pty) Ltd, Australia) was used to view and edit sequence chromatograms (file extension: *.ab1). The “translate DNA to protein” analysis tool (<http://www.expasy.ch>) was used for six-frame translations of the nucleotide sequence, which was

obtained from automated sequence analysis as described in section 2.12.5.3, into protein sequence. The BLAST alignment algorithm was used to match automated sequence text files against matching database sequence entries (see section 2.11.1) and the ClustalW programme (<http://www.ebi.ac.uk>) was used to generate a graphical alignment of protein sequences. These analyses were performed using the same copy and paste options described in sections 2.11.1 and 2.11.2. All validated sequence variations were submitted to the *single nucleotide polymorphism (SNP)* database (<http://www.ncbi.nlm.nih.gov/SNP>).

CHAPTER 3

Results

Index	Page
3.1 Clinical evaluation	88
3.2 Genetic analysis	92
3.2.1 Genotype analysis at candidate loci	92
3.2.2 Lod score calculation	97
3.3 Haplotype analysis at chromosome 1q32.2-q32.3	101
3.4 Construction of an integrated genetic, physical and transcript map	101
3.5 Candidate gene search at chromosome 1q32.2-q32.3	104
3.5.1 Selecting candidate genes for mutation analysis	104
3.6 Mutation screening of seven prioritised candidate genes	107
3.6.1 PCR-SSCP analysis and characterisation of sequence variants	107

Results

Ode to LINKAGE

*Therefore, it can fairly be concluded
For PFHBII, on which we work secluded,
Twelve loci have so far been excluded
If by computation we have not been deluded
Anonymous (modified)*

3.1 Clinical evaluation

Establishing a clinical profile of PFHBII

All available clinical data (clinical histories or clinical examination records) obtained from members of the PFHBII-affected family (Table 3.1) was analysed to identify patterns of cardiac abnormalities that are associated with the disorder. Of the 43 individuals for which clinical data were obtained, 14 were identified that presented with cardiac conduction defects that ranged from SB and first-degree AV block to CHB (Table 3.1). Four of these 14 clinically affected subjects developed DCM (Table 3.1) (individuals II:11, III:2, III:9 and III:25). A possible “cardiomyopathy” was suspected in one subject (Table 3.1) (individual II:1), although the clinical evidence to support this speculation could not be obtained. Another subject (Table 3.1) (individual III:26) developed DCM without prior or accompanying conduction abnormalities, possibly indicating that DCM may be an underlying cause of PFHBII. Generally, subjects with PFHBII displayed cardiac conduction defects and accompanying DCM or demonstrated progression to DCM between 5-22 years after an initial diagnosis of conduction abnormalities. This clinical information was then used to formulate a clinical profile of PFHBII, the results of which have been published in the Cardiovascular Journal of South Africa (Fernandez et al., 2004). The three significant outcomes of this clinical study were, PFHBII is a conduction disease associated with DCM, progression from SB to CHB or DCM was observed and, the PFHBII-associated cardiac conduction defects were located in the proximity of the AV node or Bundle of His, because none of

the clinically affected subjects presented with defects in the bundle branches of the cardiac conduction system (see Fig. 1.6). A detailed discussion of these results is presented in section 4.1.

Table 3.1

Clinical characteristics of subjects from the core PFHBII-affected family studied retrospectively

Individual	Age [#] (years)	Sex	P-R interval (msec)	EF (%)	LVEDD (cm)	Comments / clinical features
I:1	61	M	nd	nd	nd	CHB, deceased at age 61 yrs
I:2	nd	F	nd	nd	nd	Clinically unaffected
II:1	40	M	nd	nd	nd	SB (HR = 26 bpm), CHB (pacemaker at age 41 yrs), deceased at age 41 yrs
II:2	nd	F	nd	nd	nd	Clinically unaffected
II:3	45	M	nd	nd	nd	CHB (pacemaker at age 45 yrs), failed PM, deceased at age 45 yrs
II:4	nd	F	nd	nd	nd	Clinically unaffected
II:5*	68	M	200	76	4.8	Clinical history of SB, HR = 53 bpm on re-evaluation, clinically uncertain
II:6	nd	F	nd	nd	nd	Clinically unaffected
II:7*	66	M	3°AVB	60	4.5	SB (HR = 48 bpm)^ψ, CHB (pacemaker at age 42 yrs),
II:8	nd	F	nd	nd	nd	Clinically unaffected
II:9*	64	F	180	65	4.4	HR = 49 bpm, HR = 58 bpm on re-evaluation, clinically uncertain
II:10	nd	M	nd	nd	nd	Clinically unaffected
II:11*	59	F	3°AVB	41	6.3	CHB (pacemaker at age 37 yrs), DCM, deceased at age 59 yrs
II:12	nd	M	nd	nd	nd	Clinically unaffected
III:1*	40	M	160	64	4.4	LAD (30° axis), LAH, diagnosed age 36 yrs
III:2	38	M	3°AVB	38	6.2	SB (HR = 50 bpm), CHB (pacemaker at age 39 yrs), DCM, deceased at age 43 yrs
III:3	nd	F	nd	nd	nd	Clinically unaffected
III:4*	47	M	140	60	4.4	Clinically unaffected
III:5*	43	M	140	60	4.4	Clinically unaffected
III:6*	45	F	140	65	4.5	Clinically unaffected

continued on overleaf

III:7*	39	F	3°AVB	52	5.1	SB (HR = 38 bpm)^ψ, AF, CHB (pacemaker at age 33 yrs)
III:8*	45	M	150	65	4.8	Clinically unaffected
III:9*	46	M	3°AVB	42	5.7	CHB (pacemaker at age 41 yrs), DCM, deceased at age 46 yrs
III:11*	36	M	200	nd	nd	SB (HR = 43 bpm)
III:12	35	F		nd	nd	CHB (pacemaker at age 35 yrs)
III:15*	32	F	140	65	4.3	HR = 54 bpm, clinically uncertain
III:16*	40	F	180	64	4.5	Slow R-wave progression; transition to dominant R-wave in V4, clinically uncertain
III:18*	33	M	160	70	4.7	Clinically unaffected
III:19*	31	M	200	63	4.9	Clinically unaffected
III:20*	38	F	2° AVB or 3°AVB	61	5.5	1°AVB[§], SB (HR = 43 bpm), Wenckebach/complete AVB
III:21	nd	M	nd	nd	nd	Clinically unaffected
III:22*	41	M	200	64	4.4	Clinically unaffected
III:23	16	M	3°AVB	nd	nd	CHB (pacemaker at age 16 yrs), deceased at age 16 yrs
III:24*	39	F	160	60	4.5	Clinically unaffected
III:25*	43	F	3°AVB	40	5.8	CHB (pacemaker at age 29 yrs), DCM, deceased at age 43 yrs
III:26	14	M	nd	nd	nd	Heart transplant, DCM, deceased at age 15 yrs
IV:1*	23	F	nd	nd	nd	Chest pain, dyspnoea, syncope, clinically uncertain
IV:2	nd	F	nd	nd	nd	Clinically unaffected
IV:3*	22	M	160	63	5.2	Clinically unaffected
IV:4*	20	F	160	70	4.3	Clinically unaffected
IV:5*	13	F	nd	nd	nd	Not examined, clinically uncertain
IV:6*	13	F	nd	nd	nd	Not examined
IV:7*	10	F	nd	nd	nd	Not examined

= age at diagnosis in years (yrs) or age when clinical history was obtained; ψ = heart rate prior to pacemaker implant; § = first-degree (°) atrioventricular block (AVB) preceded complete heart block (CHB)

AF = atrial fibrillation; AV = atrioventricular; bpm = beats per minute; cm = centimetres; DCM = dilated cardiomyopathy; EF = ejection fraction; F = female; HR = heart rate; LAH = left anterior hemiblock; LAD = left axis deviation; LVEDD = left ventricular end-diastolic diameter in centimetres; M = male; msec = milliseconds; nd = no data available or too young for clinical diagnosis; SB = sinus bradycardia.

An asterisk indicates subjects who were genotyped in the present study (refer to Fig. 2.1).

Affected individuals are shown in bold.

In an additional study, clinical examinations were performed in other branches of the core PFHBII-affected family. These investigations did not identify features that suggested that the disorder segregated in the other families, although, in one branch, a male subject was identified who presented with HCM (see section 4.5). No other subjects were identified, in the other branches of the family that presented with HCM.

Clinical designations for the linkage study

Clinical data or clinical histories (Table 3.1) were obtained for 43 out of a total of 47 individuals in the PFHBII-affected kindred (Fig. 3.1); no clinical information was available for individuals II:13, III:10, III:13 and III:14 (Fig. 3.1). In the present study, 26 members of the PFHBII-affected family were genotyped for genetic linkage analysis (these subjects are indicated in Fig. 2.1 and Table 3.1). A conservative clinical classification system was applied when assigning a clinical status to the 26 individuals (section 2.3), because false positive linkage results could be produced as a result of misclassifying subjects as clinically affected (Ott, 1999). Consequently, for this study, eight of the 26 genotyped family members were designated clinically affected and 10 were designated clinically unaffected. The eight remaining subjects were assigned a clinically uncertain status, the reasons for these designations are provided below.

Three individuals, under the age of 20 years, were not clinically evaluated, and thus were assigned a clinically uncertain status (refer to section 2.3) (Table 3.1) (Fig. 3.1). One subject (individual IV:1) consented to donate blood for genetic analysis but declined to undergo a clinical evaluation, despite reporting recurrent chest pain, dizziness and fainting spells. Because of an incomplete assessment, she was given a clinically uncertain status (Table 3.1) (Fig. 3.1).

Four other family members were designated clinically uncertain based on their current and/or past clinical features. Generally, a heart rate of less than 60 bpm is considered an indication of SB

(Josephson et al., 1991), although the present study used more stringent diagnostic criteria and only subjects with a heart rate of less than 50 bpm were considered clinically affected. However, since progression from a slow heart rate to CHB has been demonstrated in the PFHBII-affected family (Fernandez et al., 2004), additional classification criteria for subjects with atypical clinical features were included (see section 2.3). It was noted that individual II:5 (Table 3.1) (Fig. 3.1) had previously been diagnosed with SB (Brink and Torrington, 1977), although, in a recent clinical assessment, the subject presented with a heart rate of 53 bpm. Unfortunately, the previous clinical record was not available and it was not possible to determine the historic base of this diagnosis. Therefore, this subject was tentatively assigned a clinically uncertain status. Another subject (individual II:9), presented with a slow heart rate of 49 bpm, although, on re-evaluation, had presented with a heart rate of 58 bpm. Because of the difference in the diagnosis, the subject was cautiously given a clinically uncertain status (Table 3.1) (Fig. 3.1), rather than being designated clinically affected. A third family member (individual III:15) was assigned a clinically uncertain status because she presented with a heart rate of 54 bpm (Table 3.1) (Fig. 3.1). The fourth subject assigned a clinically uncertain status (individual III:16) presented with a slow R-wave progression on an evaluation by ECG (Table 3.1) (Fig. 3.1). Section 4.6 provides a further discussion of the clinical designations of the subjects that were entered into the genetic study.

3.2 Genetic analysis

3.2.1 Genotype analysis at candidate loci

Genotype analysis of members of the family in which PFHBII segregates

A PCR-based assay was used to amplify alleles at each of the selected STR markers (Table 2.1) (section 2.8.7.1). Genotype scores were assigned, using the method described in section 2.8.7.3 (Fig. 3.2), to each member of the PFHBII-affected family that was entered into the present linkage study (Fig. 3.1). The genotype data was used to calculate lod scores at the selected candidate loci.

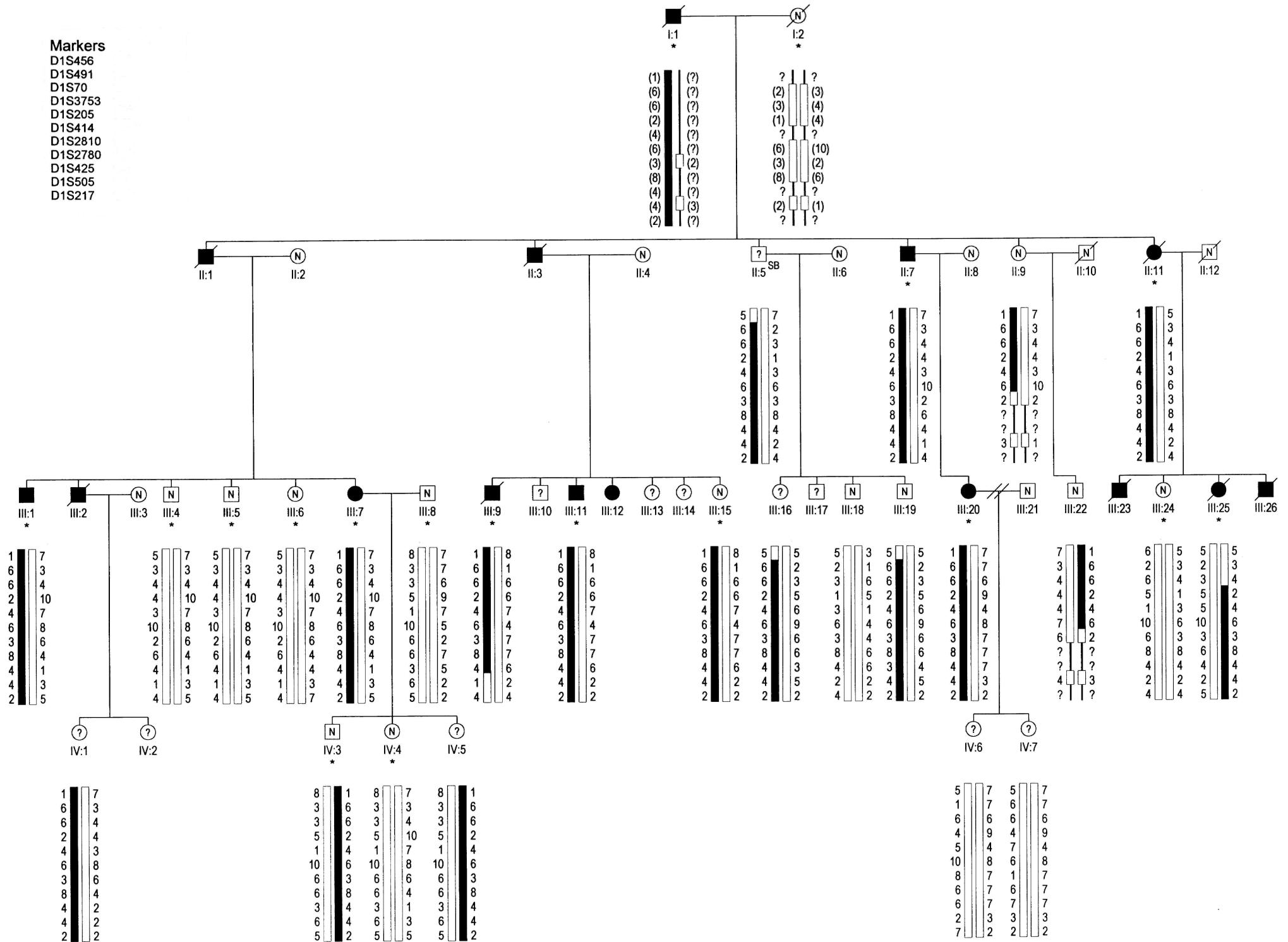
Genotype analysis in a Caucasian South African Afrikaner panel

The present study considered that the allele frequencies and STR marker heterozygosity values in the South African Afrikaner subpopulation may be different from the CEPH allele frequencies in Northern European reference families (refer to section 2.8.3). Genotype analyses were performed in a panel of 30 Caucasian Afrikaner individuals to determine the frequency of alleles at different STR markers (section 2.8.3). The allele frequencies at STR markers on chromosome 1q32.2-q32.3 are available at the US/MRC Centre for Molecular and Cellular Biology and the Department of Medical Biochemistry website (refer to section 2.8.5). The resultant allele frequencies (also used in the linkage study) were then used to calculate marker heterozygosity values in the Caucasian Afrikaner subpopulation (section 2.8.4).

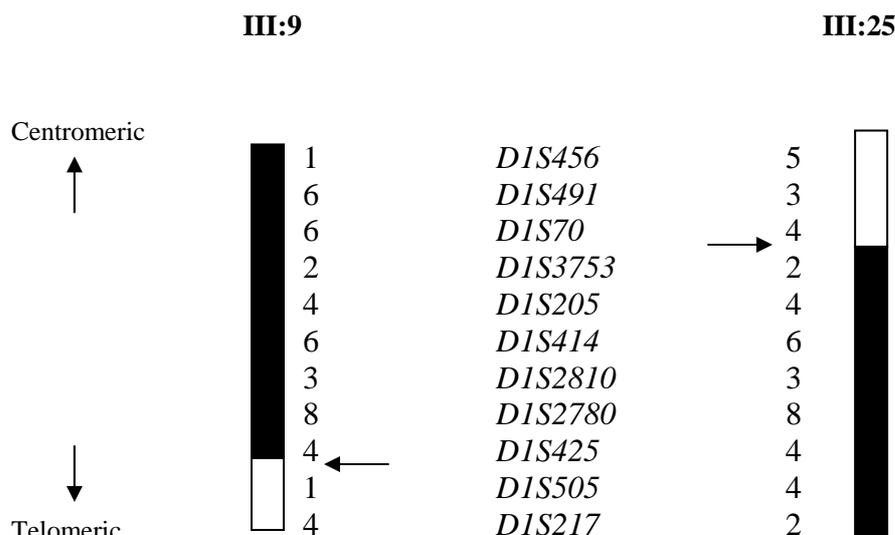
STR marker heterozygosity values at chromosome 1q32.2-q32.3

The calculated heterozygosity values for the selected STR markers at chromosome 1q32.2-q32.3 were above 0.60 (Table 3.2) (also available at the website, see section 2.8.5). The values in the Caucasian Afrikaner subpopulation were comparable to the CEPH and Marshfield Genetic Map Marker heterozygosity values, with the only heterozygosity value that differed observably was that of *DIS70*, which was 0.49 in the CEPH reference families, compared to 0.74 in the Afrikaner subpopulation (Table 3.2).

Markers
 D1S456
 D1S491
 D1S70
 D1S3753
 D1S205
 D1S414
 D1S2810
 D1S2780
 D1S425
 D1S505
 D1S217



b)

**Fig. 3.1**

Clinical designations of members of the PFHBII-affected family, genotype assignments and haplotype construction with markers at chromosome 1q32.2-q32.3.

a) A multigeneration pedigree showing the clinical designation of individuals from the family in which PFHBII segregates and haplotype data from a linked region on chromosome 1q32.2-q32.3. Marker order (centromeric to telomeric) is indicated on the top left. Filled symbols indicate clinically affected individuals; symbols with the letter N, clinically unaffected individual; symbols with a question mark, individuals with a clinically uncertain status and symbols with SB, a clinical history of sinus bradycardia. An asterisk indicates individuals used in linkage analyses. Parentheses indicate inferred genotype scores and question marks in parentheses indicate genotype scores with an indeterminate phase. Solid bars indicate the disease-associated haplotype. Thin lines represent segments of haplotypes for which the genotypes could not be assigned unambiguously.

b) An enlarged view of the disease-associated haplotype (2-4-6-3-8-4) for markers D1S3753, D1S205, D1S414, D1S2810, D1S2780 and D1S425, respectively, and, the recombination events (shown by the arrows) in individuals III:9 and III:25 (see section 3.3).

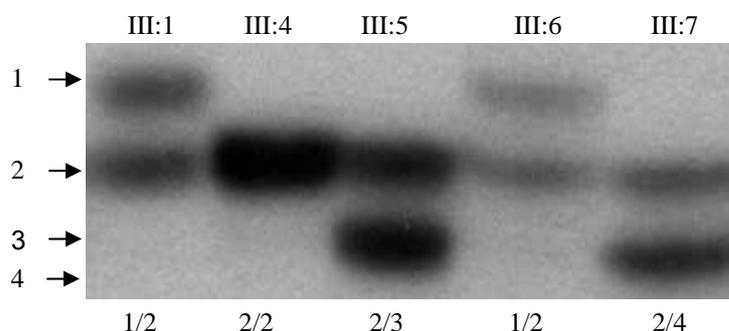


Fig. 3.2

Representative autoradiograph of a genotype analysis in a family in which PFHBII segregates.

Radioisotope signals represent the repeat alleles at marker D19S412 on chromosome 19q13.3 (one of the excluded candidate loci). Alleles were assigned numerical values (indicated to the left of the autoradiograph), as described in section 2.8.7.3. Final genotype scores for each subject are indicated below the autoradiograph. Individual pedigree identifier numbers are indicated at the top of the autoradiograph

Table 3.2

Comparison of heterozygosity values for STR markers at chromosome 1q32.2-q32.3

Genetic marker	Afrikaner population	CEPH data	Marshfield marker data
<i>DIS456</i>	0.77	0.75	0.72
<i>DIS491</i>	0.72	0.75	0.76
<i>DIS70</i>	0.74	<u>0.49</u>	0.62
<i>DIS3753</i>	0.88	No data	No data
<i>DIS205</i>	0.74	0.89	0.80
<i>DIS414</i>	0.77	0.82	0.77
<i>DIS2810</i>	0.83	0.74	0.72
<i>DIS2780</i>	0.69	0.67	0.70
<i>DIS425</i>	0.82	0.73	0.81
<i>DIS505</i>	0.86	0.82	0.75
<i>DIS217</i>	0.68	0.61	0.68

Heterozygosity score of less than 0.60 is underlined

3.2.2 Lod score calculation

Pairwise lod score analysis

Pairwise lod score calculations excluded all but one previously mapped candidate loci as cause of PFHBII, by generating lod scores of -2 or less (Table 3.3). Lod score calculations using genotype scores at the candidate locus on chromosome 1q32.2-q32.3 (Fig. 3.1) flanking the CMD1D locus generated a maximum two point lod score of 3.13 (theta $[\theta]=0.0$) at marker *DIS3753* (Table 3.4).

Multipoint analysis

Multipoint analysis generated a maximum lod score of 3.7 between the disease and markers *DIS3753* and *DIS414*. The calculated multipoint lod scores were plotted against the genetic map distances between the STR markers at chromosome 1q32.2-q32.3 (Fig. 3.3). The resultant plot provides the 90% confidence limit that the PFHBII-causative gene is localised within an interval of about 2cM flanked by the genetic markers *DIS205* and *DIS505* (Fig. 3.3).

Testing lod score stability

Setting allele frequencies to equal or varying the penetrance (Hodge and Greenberg, 1992) did not significantly alter the maximum pairwise and multipoint lod scores and only resulted in changes after the second decimal point (data not shown). Because SB is common in the general population (Josephson et al., 1991), the lod scores were re-calculated by changing the clinical status of a subject with isolated SB (individual III:11) (Table 3.1) from a clinically affected to a clinically uncertain (or [#]possibly affected) status. This analysis generated a maximum pairwise lod score of 2.83 with *DIS3753* and a multipoint score of 3.4 between markers *DIS3753* and *DIS414* (Table 3.5).

- classification used by the Cyrillic program

Table 3.3

Pairwise lod score analysis indicating exclusion of previously mapped causative loci

Locus	<u>Theta (θ)</u>				
	0.0	0.01	0.05	0.1	0.2
Chromosome 1p1-1q21^a					
<i>D1S305</i>	-16.39	-6.20	-2.98	-1.58	-0.39
<i>D1S2707</i>	-4.47	-1.43	-0.30	0.1	0.39
<i>D1S176</i>	-6.80	-1.43	-0.93	-0.25	0.19
Chromosome 2q14-q22^b					
<i>D2S2224</i>	-12.32	-6.67	-3.44	-1.98	-0.67
<i>D2S2339</i>	-0.66	-0.58	-0.38	-0.22	-0.06
<i>D2S2215</i>	-6.54	-2.09	-0.88	-0.42	-0.99
<i>D2S112</i>	-6.03	-3.00	-0.83	-0.01	0.50
Chromosome 3p22-p25^c					
<i>D3S1263</i>	-3.94	-3.55	-2.32	-1.51	-0.71
<i>D3S2303</i>	-7.89	-5.07	-2.34	-1.23	-0.34
<i>D3S1211</i>	-10.58	-4.59	-2.89	-2.07	-1.17
Chromosome 6q23^d					
<i>D6S262</i>	-4.24	-0.92	-0.26	-0.03	0.1
Chromosome 7q35-q36^e					
<i>D7S505</i>	-3.71	-2.52	-1.06	-0.92	-0.92
Chromosome 9q12-q13^f					
<i>D9S152</i>	-6.30	-1.28	-0.66	-0.37	-0.13
Chromosome 11p11^g					
<i>D11S1344</i>	-7.26	-0.97	-0.25	-0.03	0.17
Chromosome 14q12^h					
<i>MYH7^{int24}</i>	-3.64	-2.65	-1.62	-1.03	-0.43
Chromosome 15q14ⁱ					
<i>ACTC</i>	-4.77	-0.81	-0.12	-0.14	0.28
Chromosome 15q22^j					
<i>D15S108</i>	-2.21	-2.08	-1.64	-1.25	-0.72
Chromosome 19q13.3^k					
<i>D19S412</i>	-4.38	-2.81	-1.61	-1.01	-0.44
<i>D19S902</i>	-12.64	-4.07	-2.11	-1.29	-0.56
<i>D19S596</i>	-4.84	-1.13	-0.49	-0.26	-0.08
<i>D19S246</i>	-7.53	-3.23	-1.89	-1.31	-0.72

Disease-causative genes and clinical disorders (indicated in brackets below) that were excluded by linkage analysis were: a = LMNA (CMD1A); b = unknown aetiology (CMD1H); c = SCN5A (CMD1E, PCCD and CCD); d = CMD1F; e = PRKAG2 (CMH4); f = unknown aetiology (CMD1B); g = MYBPC3 (CMH4); h = MYH7 (CMH1); i = ACTC (IDC); j = TPM1 (CMH3); k = unknown aetiology (PFHBI, ICCD) (refer to Fig. 1.2, 1.5 and 1.7).

Table 3.4*Pairwise lod scores between PFHBII and markers on chromosome 1q32.2-q32.3*

Locus	Theta (θ)					Zmax (θ)
	0.0	0.01	0.05	0.1	0.2	
<i>DIS456</i>	-0.99	0.62	1.11	1.16	0.92	1.16 (0.1)
<i>DIS491</i>	-1.55	0.84	1.38	1.48	1.33	1.48 (0.1)
<i>DIS70</i>	-0.94	0.68	1.22	1.31	1.18	1.31 (0.1)
<i>DIS3753</i>	3.13	3.08	2.88	2.62	2.06	<u>3.13</u> (0.0)
<i>DIS205</i>	0.38	0.37	0.31	0.25	0.14	0.38 (0.0)
<i>DIS414</i>	2.51	2.47	2.29	2.06	1.57	2.51 (0.0)
<i>DIS2810</i>	1.97	1.94	1.82	1.66	1.30	1.97 (0.0)
<i>DIS2780</i>	1.95	1.91	1.78	1.61	1.25	1.95 (0.0)
<i>DIS425</i>	0.12	0.12	0.10	0.08	0.05	0.12 (0.0)
<i>DIS505</i>	-0.51	1.09	1.60	1.66	1.45	1.66 (0.1)
<i>DIS217</i>	-1.52	-0.74	0.39	0.74	0.82	0.82 (0.2)

*Maximum lod score is underlined***Table 3.5***Pairwise lod scores generated by changing the phenotypic status of individual III:11*

Locus	Theta (θ)					Zmax (θ)
	0.0	0.01	0.05	0.1	0.2	
<i>DIS456</i>	-1.29	0.32	0.84	0.92	0.76	0.92 (0.1)
<i>DIS491</i>	-1.85	0.55	1.10	1.22	1.12	1.22 (0.1)
<i>DIS70</i>	-1.24	0.39	0.94	1.06	0.98	1.06 (0.1)
<i>DIS3753</i>	2.83	2.79	2.61	2.37	1.86	<u>2.83</u> (0.0)
<i>DIS205</i>	0.30	0.29	0.24	0.19	0.09	0.30 (0.0)
<i>DIS414</i>	2.43	2.39	2.22	2.00	1.53	2.43 (0.0)
<i>DIS2810</i>	1.90	1.87	1.75	1.59	1.26	1.90 (0.0)
<i>DIS2780</i>	1.87	1.84	1.71	1.55	1.21	1.87 (0.0)
<i>DIS425</i>	0.11	0.11	0.09	0.07	0.04	0.11 (0.0)
<i>DIS505</i>	-0.59	0.99	1.51	1.58	1.40	1.58 (0.1)
<i>DIS217</i>	-1.68	-1.00	0.12	0.49	0.63	0.63 (0.2)

Maximum lod score is underlined

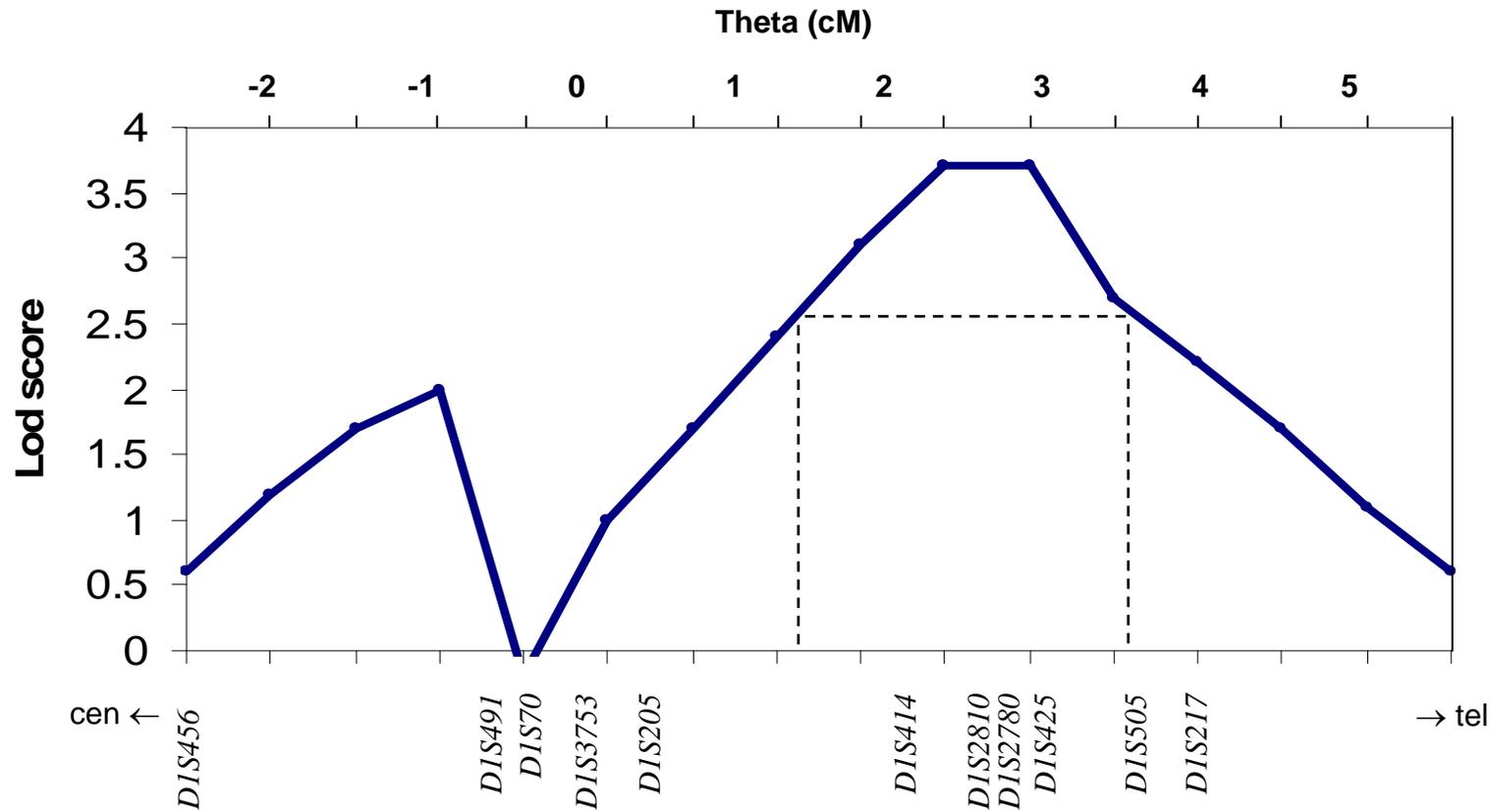


Fig. 3.3

Multipoint lod score curve at chromosome 1q32.2-q32.3.

Multipoint lod scores for the PFHBII locus versus markers used in linkage analysis. The recombination distance (θ) is plotted on the x-axis from an arbitrary origin, set at marker D1S3753. The 1-lod-down 90% support interval places the likely location of the PFHBII gene within a 2cM interval telomeric of D1S205 and centromeric of D1S505.

x-axis: 1 unit = 0.5cM; y-axis: 1 unit = 0.5 lod score units

cen = centromeric; tel = telomeric

3.3 Haplotype analysis at chromosome 1q32.2-q32.3

The eight clinically affected individuals used in the linkage study carried a common inherited chromosomal haplotype of 2-4-6-3-8-4 for markers *DIS3753*, *DIS205*, *DIS414*, *DIS2810*, *DIS2780* and *DIS425*, respectively (Fig. 3.1a and b). Recombination breakpoints, which were identified in individual III:9 and III:25, localised the PFHBII-associated gene to within a region spanning 3.9cM (2.85Mb) (physical map distances obtained from the Ensembl database) telomeric of *DIS70* and centromeric of *DIS505* (Fig. 3.1a and b). The haplotype data corresponded to the 1-lod-down likelihood analysis (Fig. 3.3), further supporting the localisation of the PFHBII-causative gene.

3.4 Construction of an integrated genetic, physical and transcript map

PCR-based mapping of markers onto genomic clones

The selected STR markers at chromosome 1q32.2-q32.3 (Table 2.1) were mapped onto the BAC, PAC and YAC clone-inserts using a PCR-based reaction that generated an amplification product (section 2.10.4) (Fig. 3.4). A PCR amplification product generated by an STR marker and genomic clone-insert DNA as a template indicated that the specific marker was physically mapped onto the insert of that genomic clone (Fig. 3.4). Thus, the PCR-based mapping data presented in this section (using marker *DIS491*) was incorporated onto a physical map of the PFHBII locus, as shown on Fig. 3.5. This iterative process was also used to map markers *DIS70*, *DIS3753*, *DIS205*, *DIS414*, *DIS2810*, *DIS2780*, *DIS425* and *DIS505* onto their respective genomic clones (see Fig. 3.5). Initially, the wet-bench mapping data for markers *DIS70*, *DIS3753*, *DIS205* and *DIS414* corresponded to previous physical and transcript data that had been generated by Schutte et al., (2000). However, during the course of the study, the availability of chromosome 1 sequence data facilitated the use of bioinformatics-based mapping strategies (see sub-section below) in favour of

the PCR-based approach. Therefore, the present study incorporated both approaches in order to generate a validated and high-resolution map of the PFHBII locus.

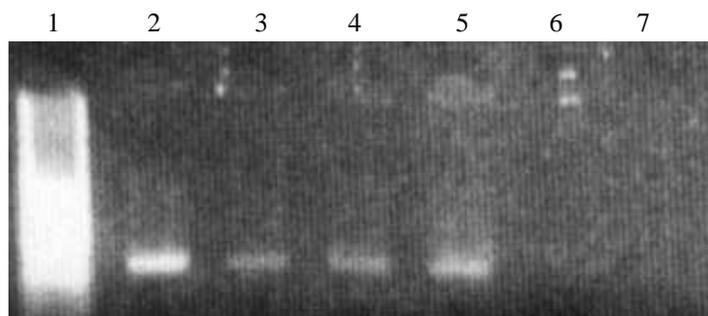


Fig. 3.4

Representative agarose gel analysis of mapping an STR marker onto genomic clone-inserts.

A 1.5% agarose gel analysis shows PCR amplification products generated using the STR marker D1S491 and template DNA extracted from genomic clones. This PCR-based result is shown in Fig. 3.5.

Lane 1 = DNA marker VI; Lane 2 = clone RP1-272L16 template DNA; Lane 3 = clone RP1-28O10 template DNA; Lane 4 = clone CTA-55I10 template DNA; Lane 5 = genomic template DNA (positive control); Lane 6 = chromosome 19 BAC template DNA (negative control); Lane 7 = water blank

Bioinformatics-based mapping of STR markers and gene sequences onto genomic clones

Bioinformatics alignment algorithms at the Sanger Centre and NCBI website (section 2.11.1) were used to map electronic versions of STR markers and gene sequences (or transcripts) onto electronic versions of the selected genomic clones at chromosome 1q32.2-q32.3. The *in silico* STR mapping data was used as an adjunct to PCR-based mapping analyses, and, the combined data was used to generate a comprehensive integrated genetic, physical and transcript map of the PFHBII locus (Fig. 3.5). The gene transcripts positioned onto the physical map of the PFHBII locus (Fig. 3.5) will be discussed in the section below.

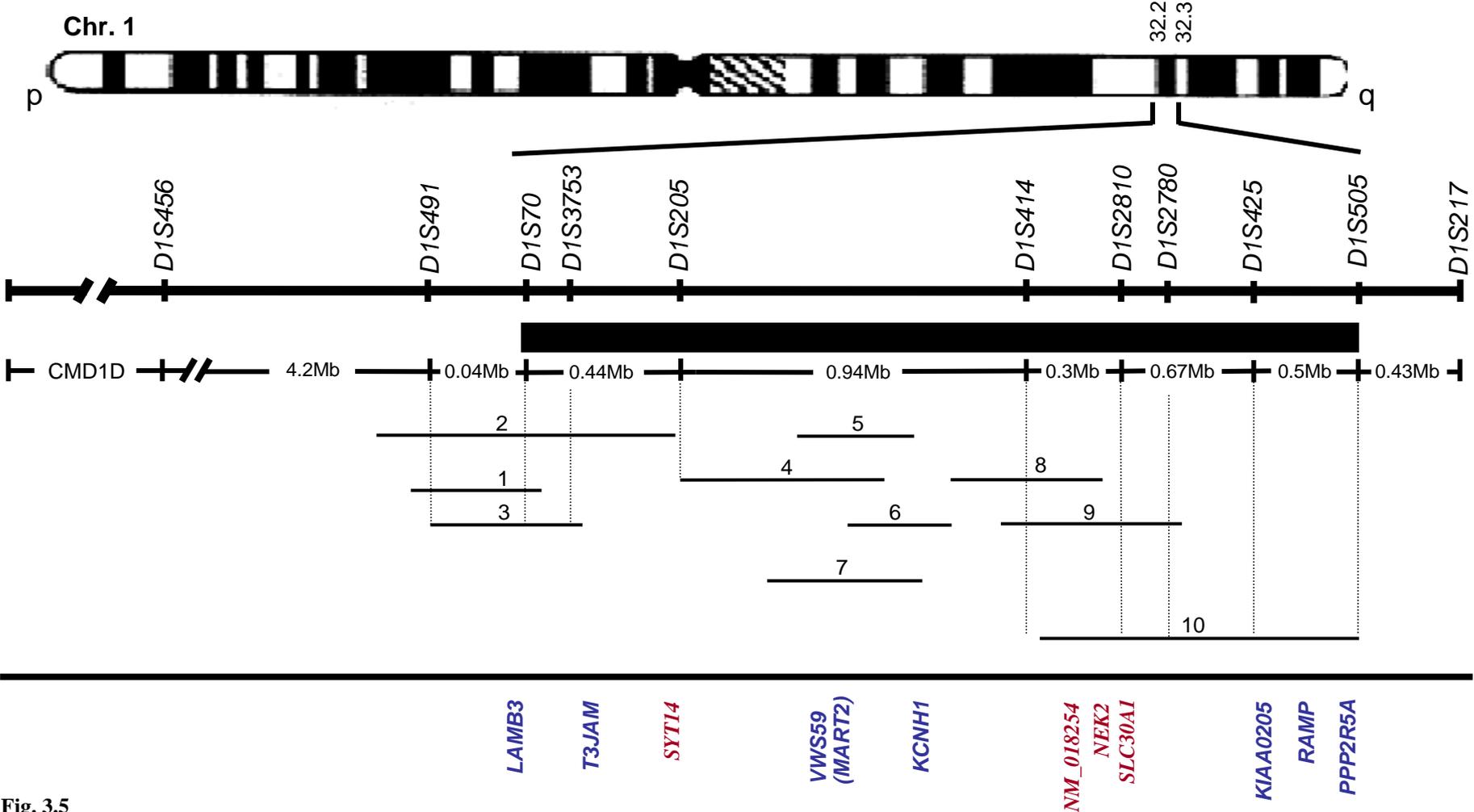


Fig. 3.5
Integrated genetic, physical and transcript map of the PFHBII locus.

The individual clones comprising the contig spanning the PFHBII locus are numbered: 1 = CTA-55110; 2 = RP1- 272L16; 3 = RP1-28010; 4 = RP3-434014; 5 = CTA-447D17; 6 = CTA-321I20; 7 = RP4-782D21; 8 = CTA-11C7; 9 = RP5-879K22; 10 = YAC 937D4. Genetic markers were positioned onto clones using wet-bench and in silico techniques. A solid bar represents the PFHBII critical interval flanked by D1S70 and D1S505. Vertical (dash) lines indicate the markers were mapped onto the genomic clone-insert by PCR-based mapping or bioinformatics analysis. The seven candidate genes indicated in blue were screened for the PFHBII-causative mutation. The candidate genes indicated in red have been prioritised for future mutation screening. Intermarker physical map distances are shown in megabases (Mb)

CMD1D = dilated cardiomyopathy locus 1D; KCNH1 = potassium channel; KIAA0205 = putative acyltransferase; LAMB3 = laminin β3; NEK2 = never in mitosis A-related kinase 2; NM_018254 = RCOR3 (REST co repressor); PPP2R5A = protein phosphatase 2A b56alpha; RAMP = retinoic acid regulated nuclear matrix associated protein; SLC30A1 = zinc transporter 1; SYT14 = synaptotgamin XIV; T3JAM = adapter protein that interacts with a cytoskeletal protein; VWS59 (MART2) = novel human gene.

3.5 Candidate gene search at chromosome 1q32.2-q32.3

3.5.1 Selecting candidate genes for mutation analysis

Database searches (NCBI database builds 28 and 31) indicated that, by genome location, 27 genes map to within the PFHBII locus (Appendix G). Pathological mutations associated with three of the 27 genes mapped to within the PFHBII locus have been implicated as cause of a skin disease (junctional epidermolysis bullosa [JEB]) (Pulkkinen et al., 1994), a disease associated with the formation of ovarian cysts (cortisone reductase deficiency [CRD]) (Draper et al., 2003) and two diseases associated with orofacial defects (VWS and politeal pterygium syndrome [PPS]) (Kondo et al., 2002) (Table 3.6). None of these described defects have been linked to electrophysiological or structural defects of the heart. Furthermore, heterozygous individuals carrying the causative mutation of the recessive condition, JEB (Pulkkinen et al., 1994) (Table 3.6), do not show any cardiac conduction abnormalities. Therefore, bioinformatics similarity analyses were used to prioritise plausible candidate genes, from the 24 other genes mapped to within the PFHBII locus, for mutation analysis based on alignment significance scores (refer to section 2.11.3). However, it has to be mentioned that laminin β 3, which is encoded by *LAMB3* and in which JEB-causative mutations have been described (Table 3.6), is an extracellular matrix structural protein linked to the sarcolemma (Pulkkinen et al., 1994). Previously, Norgett et al., (2000) and Tsubata et al., (2000) had demonstrated that defects in the sarcolemmal-linked proteins desmoplakin, δ -sarcoglycan and plakoglobin could cause DCM. Thus, based on this evidence, it was opted to rank *LAMB3* as a possible PFHBII candidate gene (see paragraph hereafter), even though described defects in the gene has not been associated with any form of cardiac abnormality (see Table 3.6).

Table 3.6*Diseases caused by defects in genes mapped to within the PFHBII locus*

Gene	OMIM	Disease	Phenotypic description	Reference
<i>LAMB3</i> (laminin β 3)	150310	Junctional epidermolysis bullosa (JEB)	Skin blisters within the dermal-epidermal basement membrane	Pulkkinen et al., (1994);
<i>HSD11B1</i> (hydroxysteroid 11-beta dehydrogenase)	600713	Cortisone reductase deficiency (CRD)	Hirsutism (increased female body hair), obesity, infertility and hyperandrogenism	Draper et al., (2003)
<i>IRF6</i> (interferon regulatory factor 6)	119300, 119500	VWS, popliteal pterygium syndrome (PPS)	Lip pits with cleft lip or palate (VWS); cleft lip or palate, genital abnormalities and webbing of skin (PPS)	Kondo et al., (2002)

Based on the selection criteria described in section 2.11, the seven candidate genes selected for mutation screening (in order of prioritisation) were a gene encoding an K⁺ channel (*KCNH1*) (accession number AJ001366) (Occhiodoro et al., 1998), a protein phosphatase (*PPP2R5A*) (accession number L42373) (McCright and Virshup, 1995), an adapter protein that interacts with a cytoskeletal protein (*T3JAM*) (accession number AL049667) (Dadgostar et al., 2003), an extracellular matrix protein laminin β 3 (*LAMB3*) (accession number D37766) (Gerecke et al., 1994), a retinoic acid-regulated nuclear matrix associated protein (*RAMP*) (accession number AF195765) (Cheung et al., 2001), a putative acyltransferase (*KIAA0205*) (accession numbers D86960) (Nagase et al., 1996) and a novel human gene encoding an acyltransferase-like protein (**VWS59* or *MART2*) (accession number NM_018194) (B.C Schutte, personal communication) (Fig. 3.5). A summary of the selected candidate genes, the most significant bioinformatics similarity alignment results and the reasoning for selecting the genes are provided in Table 3.6.

* - the gene name was provided by Dr. B Schutte (University of Iowa) in October 2000. The provisional name for *VWS59* is indexed as melanoma antigen recognized by T cells 2 (*MART2*) in the NCBI database.

Table 3.7*Candidate genes selected for mutation screening*

Candidate gene	Gene/protein signature similarity	E-value (max)	Comparative disease or pathophysiological association
<i>KCNH1</i>	K ⁺ and Na ⁺ channels	0.0	K ⁺ and Na ⁺ channels are associated with arrhythmias and conduction disorders (Curran et al., 1995; Wang et al., 1995; Chen et al., 1998; Schott et al., 1999; Tan et al., 2001)
<i>PPP2R5A</i>	Protein phosphatases	1.8e-275	Mouse model studies indicate that mutations in a phosphatase (calcineurin), are associated with hypertrophy (Molkentin et al., 1998)
<i>T3JAM</i>	Adapter protein that interacts with myosin class II heavy chain (a cytoskeletal protein)	6.5e-116	Cytoskeletal proteins (e.g. desmin and cardiac muscle LIM protein) are associated with DCM and HCM (Li et al., 1999; Knöll et al., 2002; Geier et al., 2003; sections 1.3.2, 1.3.3 and 1.3.6). A mutation in an adapter protein, ankyrin 2, is associated with LQTS type 4 (LQT4) (Mohler et al., 2003)
<i>VWS59 (MART2)</i>	<i>KIAA1173</i> (acyltransferase-like protein)	3.7e-77	Acyltransferases (e.g. tafazzin) are associated with BTHS and DCM (Bione et al., 1996)
<i>LAMB3</i>	<i>LAMA2</i> (laminin α -2) (an extracellular matrix protein)	4e-39	Extracellular matrix proteins (e.g. merosin) are associated with juvenile DCM and muscular dystrophy (Spyrou et al., 1998)
<i>RAMP</i>	GTPase-protein containing a WD40 repeat	2e-32	WD40 domains act as adaptor/regulatory modules in signal transduction or are involved in pre- <i>mRNA</i> processing and cytoskeleton assembly; altered cytoskeletal architecture has been implicated in DCM (section 1.3.3.1.2)
<i>KIAA0205</i>	<i>G4.5</i> (tafazzin) (acyltransferase)	2e-13	Acyltransferases (e.g. tafazzin) are associated with BTHS and DCM (Bione et al., 1996)

Candidate genes are listed in descending order of priority, starting with the gene with the most significant E-value.

BTHS = Barth syndrome; DCM = dilated cardiomyopathy; HCM = hypertrophic cardiomyopathy; Na⁺ = sodium; K⁺ = potassium; LAMB3 = laminin β 3; LQTS = long QT syndrome; KCNH1 = K⁺ channel; KIAA0205 = a putative acyltransferase; PPP2R5A = protein phosphatase; RAMP = a retinoic acid-regulated nuclear matrix associated protein; T3JAM = an adapter enzyme that interacts with a cytoskeletal protein; VWS59 = a novel human gene encoding an acyltransferase-like protein

3.6 Mutation screening of seven prioritised candidate genes

3.6.1 PCR-SSCP analysis and characterisation of sequence variants

Mutation analyses of *KIAA0205*, *PPP2R5A*, *RAMP* and *T3JAM* did not identify any PCR-SSCP gel electrophoresis variant patterns (Fig. 3.6). Variant PCR-SSCP electrophoretic mobility patterns were identified in three selected candidate genes, namely, *KCNHI*, *VWS59 (MART2)* and *LAMB3* (Table 3.7). For each of the three genes, the variant patterns were observed in affected and unaffected members of the PFHBII family (see section 2.12.2 for individuals used in the PCR-SSCP analysis), as well as in subjects from a Caucasian Afrikaner control panel (discussed further in section below). Sequence characterisation of individual samples representative of each of the different electrophoretic mobility patterns identified three separate nucleotide substitutions. These sequence substitutions were either identified in an exon, but did not alter an amino acid residue (see the example hereafter), or were located away from the splice consensus sequences of an intron (Table 3.7). Additional direct automated sequence analyses (see section 2.12.5.3) identified a thymine (T) to cytosine (C) (T>C) substitution in the 3'-untranslated region (UTR) of *KCNHI* and a single nucleotide substitution in exon 11 of *RAMP*, the latter resulted in a substitution of a threonine with lysine amino acid residue (Table 3.7). Neither of these sequence variations was detected by PCR-SSCP analysis. The implications of these results as a reflection of the relative sensitivity of the PCR-SSCP and sequence analyses will be discussed at greater length in section 4.8.3.1.

The sequence variations identified in *KCNHI* and *VWS59* were submitted to the NCBI SNP database, and, subsequently, were all described by other investigators in other populations (Table 3.7). The nucleotide substitutions identified in the present study in *LAMB3* and *RAMP* had, at the time, already been identified in other population groups in Japan and North America (Table 3.7).

Representative examples of the process of identifying the sequence variants in *KCNH1* and *RAMP* are provided below.

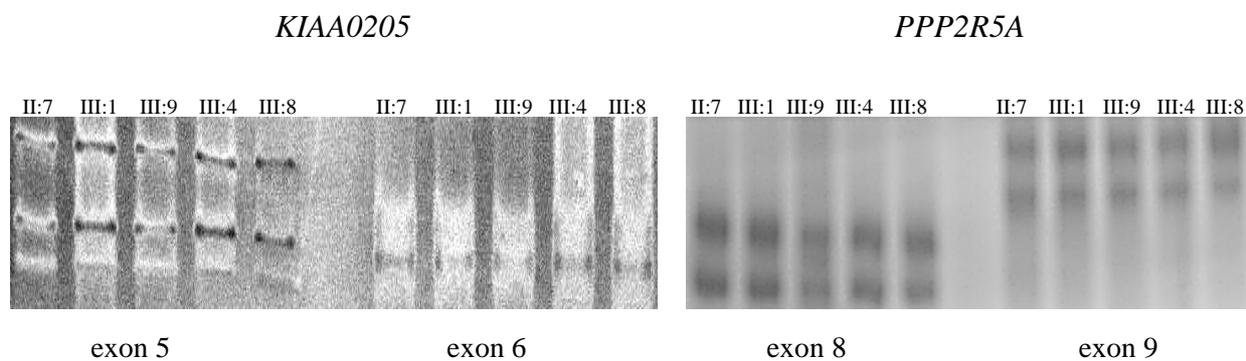


Fig. 3.6

Representative PCR-SSCP gel electrophoresis analysis of genes in which no mobility variations were identified.

No variant electrophoretic mobility patterns were identified in KIAA0205, PPP2R5A, RAMP and T3JAM. A representative of these PCR-SSCP analyses is provided in the figures above, which show analyses of exons 5 and 6 of KIAA0205 and exons 8 and 9 of PPP2R5A.

II:7, III:1 and III:9 = clinically affected subjects; III:4 and III:8 = clinically unaffected subjects (see Fig. 3.1)

Mutation analysis of KCNH1

Gel electrophoresis analysis by PCR-SSCP indicated two different electrophoretic patterns in exon 11a of *KCNH1* (Fig. 3.7a). One variant pattern (designated C/C) (Fig. 3.7a) segregated in two clinically affected subjects (individuals II:7 and III:9) and a two clinically unaffected subjects (individuals III:4 and III:6) (refer to subjects used, section 2.12.2); the other pattern (designated T/T) (Fig. 3.7a) was observed in a clinically unaffected and a clinically affected subject (individuals III:1 and III:8, respectively) (see Table 3.1). The same PCR-SSCP mobility variant patterns in exon 11a were identified in members of the Caucasian Afrikaner subpopulation (Fig. 3.7b).

Sequence analysis of an individual sample with the C/C pattern identified a T>C *transition* at nucleotide position 2225 (Fig. 3.8a and b) (Table 3.7). This sequence variant had previously been identified in a North American population of unknown ethnicity (submitted by Lee714008, NCBI SNP accession number ss1537837). Translation of the nucleotide sequence of a subject with the T>C transition indicated that the variation did not change an aspartic acid residue at the amino acid *codon* position 712, that is, it was a synonymous substitution. The PCR-SSCP analyses of 30 Caucasian subjects showed that frequency of the T- and C-alleles in the Afrikaner subpopulation were 0.900 and 0.100, respectively. No data is currently available that indicates the frequency of the T- and C-alleles in the population described by Lee714008 (Table 3.7).

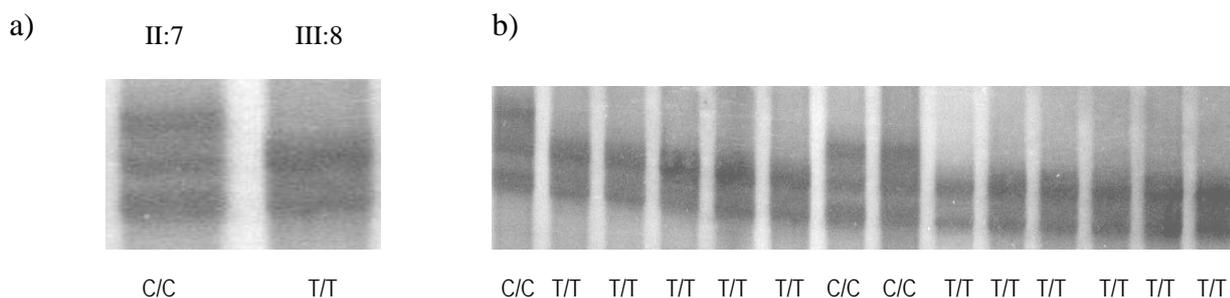


Fig. 3.7

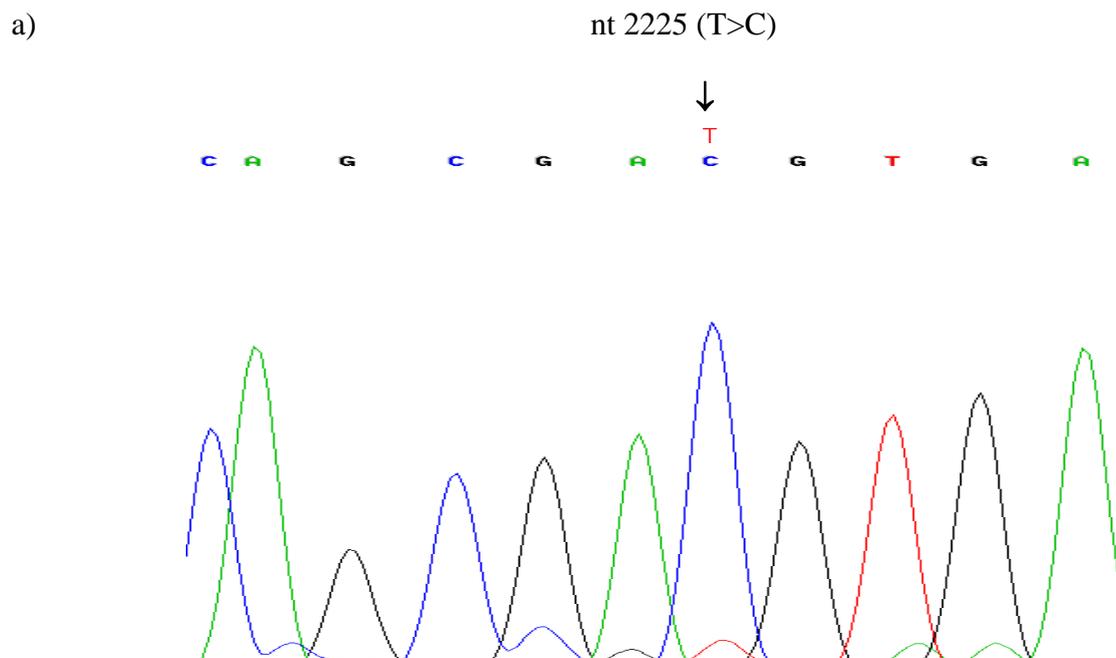
PCR-SSCP analysis of exon 11a of *KCNH1*.

*a) A 10% mildly denaturing polyacrylamide gel showing the electrophoretic variant patterns in exon 11a of *KCNH1*, which segregated in members of the *PFHBII*-affected family.*

b) A representative PCR-SSCP analysis of the segregation of the different electrophoretic mobility patterns in a Caucasian Afrikaner panel.

Based on sequence analyses (see text), the different electrophoretic patterns were designated C/C and T/T (indicated below the PCR-SSCP image) to facilitate scoring the different alleles that segregated in the Caucasian Afrikaner control panel.

C = cytosine; T = thymine



b) BLAST analysis

Ind III:1 : gattgtgtccggaagatcagcgcgctgaaacgtgaagaggaagaacgcatgaaacgaaa 79
 |||||*
 Db seq: gattgtgtccggaagatcagcgcgatgaaacgtgaagaggaagaacgcatgaaacgaaa 1778

Ind III:1: gaatgaggccccctgatcttgccccggaccacctgtccggcgctcttccagagatt 139
 |||||
 Db seq gaatgaggccccctgatcttgccccggaccacctgtccggcgctcttccagagatt 1838

Ind III:1: ccgacagcagaaaagaggccaggctggcagctgagagagggggccgggacctggatgacct 199
 |||||
 Db seq ccgacagcagaaaagaggccaggctggcagctgagagagggggccgggacctggatgacct 1898

Ind III:1 agatgtggagaagggcaatgtccttacagagcatgcctccgccaaccacagcctcgtgaa 259
 |||||
 Db seq agatgtggagaagggcaatgtccttacagagcatgcctccgccaaccacagcctcgtgaa 1958

Ind III: ggccagcgtggtcaccgtgctgaga 265
 |||||
 Db seq ggccagcgtggtcaccgtgctgaga 1984

Fig. 3.8

Sequence analysis of exon 11a of *KCNH1*.

a) Sequence analysis of individual III:1 indicated a T>C transition at nucleotide position 2225.

b) The BLAST alignment of the nucleotide sequence of individual III:1 against the *KCNH1* coding sequence obtained from the NCBI database (accession number AF078741) indicating the nucleotide mismatch (*). Nucleotide alignment co-ordinates do not correspond to the nucleotide position at which the sequence variation was identified.

Db seq = database sequence; Ind = individual identifier, nt = nucleotide

Assessment of mutation screening sensitivity and the identification of sequence variants

To test the sensitivity of the PCR-SSCP mutation-screening regimen used in the present study, 29 exons were screened by direct automated sequence analysis (see section 2.12.5.3). The following sub-sections provide a synopsis of two exons in which sequence variations were identified.

KCNHI

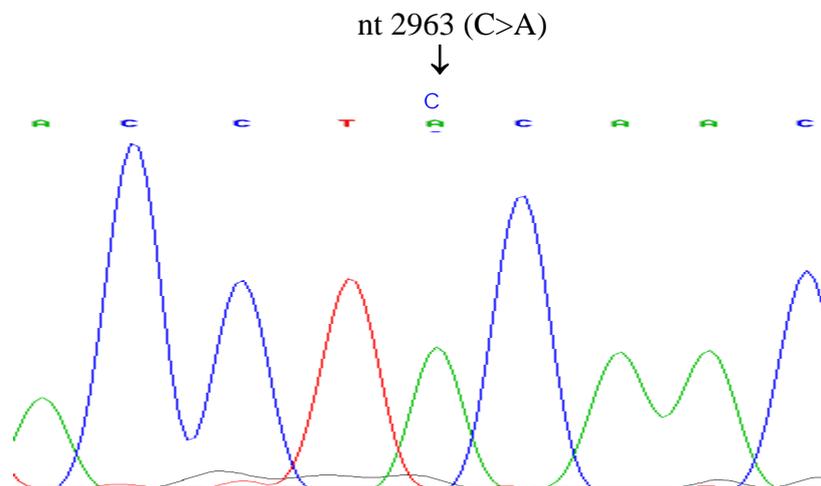
Using primer set 11d of *KCNHI* (Appendix F), sequence analysis of members of the PFHBII-affected family identified a cytosine to adenine (A) (C>A) *transversion* at nucleotide position 2963 (Fig. 3.9a and b) (Table 3.7). The sequence variation was identified in two clinically unaffected subjects (individual III:4 and III:6) (Table 3.1) and, subsequently, was submitted to the NCBI SNP database (Brincor, accession number ss4383560) (submitted on 20/02/2002). Later, a study by another group in a Japanese population identified the same sequence variant (IMS-JST147545, accession number ss4976734) (date of submission on 12/08/2002), and, in that population, the frequency of the C- and the A-alleles was 0.384 and 0.616, respectively. The frequency of the C>A nucleotide substitution was not calculated in the Afrikaner panel because it was not detected by PCR-SSCP analysis. Additionally, allele specific restriction enzyme analysis (Maeda et al., 1989) was not performed because the sequence variation did not alter a restriction enzyme site. However, such variants could be detected by other methods, such as primer extension analyses (Makridakis and Reichardt, 2001), for example.

RAMP

Direct sequence analyses of exon 11b of *RAMP* identified an A>C *transversion* at nucleotide position 34260 (Fig. 3.10a and b), which segregated exclusively in two clinically affected members of the PFHBII-affected family (individuals II:7 and III:20, father and daughter) (Fig. 3.1). The sequence variation was not identified upon sequence analyses of the other clinically affected members of the PFHBII-affected family or in five subjects from the Caucasian Afrikaner panel. A

subsequent analysis of the NCBI SNP database indicated that the A>C sequence variation had been identified in three North American populations (Table 3.7).

a)



b) BLAST analysis

```

Ind III:8   gccacagtctggaggtgaggcagagctgaaggaggacatcaaggcctaaacgcaaaa 65
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db seq     gccacagtctggaggtgaggcagagctgaaggaggacatcaaggcctaaacgcaaaa 2918

Ind III:8   atgaccaatattgagaaacagctctctgagatactcaggatattaactccagaagatcc 125
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db seq     atgaccaatattgagaaacagctctctgagatactcaggatattaactccagaagatcc 2978

Ind III:8   ttcagtctcctcaggagtggttgaaatcgcaggccacagtcccagaatcagagaga 185
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db seq     ttcagtctcctcaggagtggttgaaatcgcaggccacagtcccagaatcagagaga 3038

Ind III:8   gacattttggagccagctgagaggtctatttaaaaaaaagtcagagacagatacctac 245
           |||||||||||||||||||||||||||||||||||||||||||||||||*|
Db seq     gacattttggagccagctgagaggtctatttaaaaaaaagtcagagacagatacctacc 3098

Ind III:8   aacc 249
           ||||
Db seq     aacc 3102

```

Fig. 3.9

Sequence analysis of the 3' UTR of *KCNH1*.

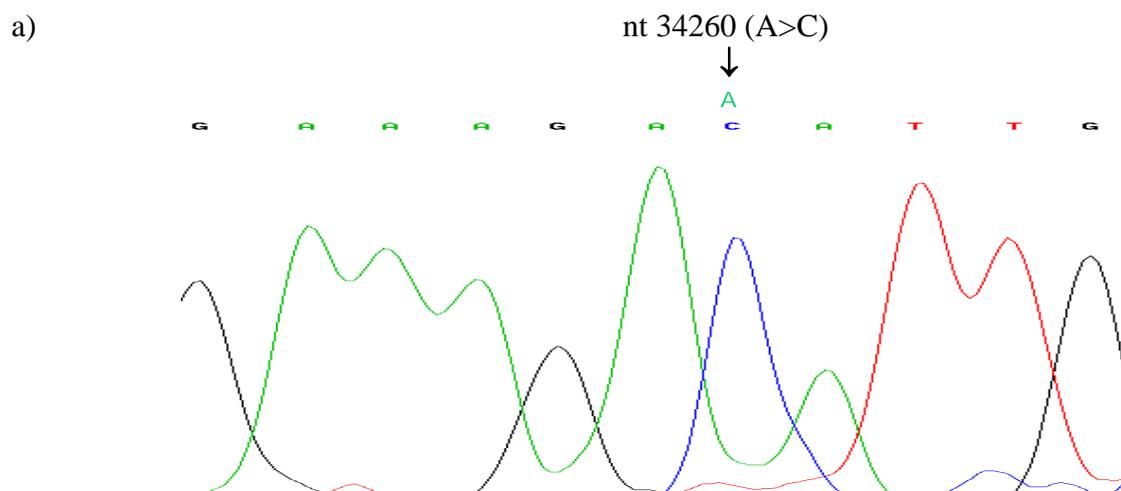
a) Direct sequence analysis (individual III:8) identified a C>A transversion at nucleotide position 2963.

b) The BLAST alignment of the nucleotide sequence of individual III:8 against the *KCNH1* coding sequence indicating the nucleotide mismatch (*). Nucleotide alignment co-ordinates do not correspond to the nucleotide position at which the sequence variation was identified.

Db seq = database sequence; *Ind* = individual identifier, *nt* = nucleotide

Alignment analysis of the RAMP variant

The NCBI SNP database indicated that the nucleotide alteration at position 34260 in *RAMP* resulted in a Thr694Lys amino acid substitution (Table 3.7). In order to assess whether the amino acid substitution possibly had a pathological consequence (from “*Guidelines for deciding whether a DNA sequence change is pathogenic*”, Strachan and Read, 1999), the L2DTL (*RAMP*) protein sequence derived from translating the nucleotide sequence of individual II:7 (see section 2.12.5.4) was compared to the human and mouse L2DTL sequences obtained from the NCBI database (accession numbers NP_057532 and AK054412). This alignment analysis showed a species conserved threonine residue at human codon position 694 (mouse equivalent codon position 693) (Fig. 3.11). Although further analyses were not pursued, this finding was interesting as it is possible that the sequence variation could play a modifying role in the pathogenesis of PFHBII (also see section 4.7).



b) BLAST analysis

```

II:7      ttgctggaaccaggaagaccttagtaaggactcttaggtcctaccaaataagcaaaa 180
          |||
Db seq    ttgctggaaccaggaagaccttagtaaggactcttaggtcctaccaaataagcaaaa 35229

II:7      ttgaaggagctggtaccagatctcagagcctccgtctcctatcagtcctatgcttcag 240
          |||
Db seq    ttgaaggagctggtaccagatctcagagcctccgtctcctatcagtcctatgcttcag 35289

II:7      aaagctgtggaacgctacctctcctttgagaccttgggagaagggtctgaaatggttag 300
          |||
Db seq    aaagctgtggaacgctacctctcctttgagaccttgggagaagggtctgaaatggttag 35349

II:7      gcaaagagaatagtccccagagaataaaaactggttggccatggcagccaaacgga 360
          |||
Db seq    gcaaagagaatagtccccagagaataaaaactggttggccatggcagccaaacgga 35409

II:7      aggctgagaatccatctccacgaagtccgtcatccagacaccaattccaggagacaga 420
          |||
Db seq    aggctgagaatccatctccacgaagtccgtcatccagacaccaattccaggagacaga 35469

II:7      gcggaagagacattgccaagcccggtaagtcagcagtggtgggaagatacatttc 474
          |||
Db seq    gcggaagaaatgccaagcccggtaagtcagcagtggtgggaagatacatttc 35523
  
```

Fig. 3.10

Sequencing analysis of exon 11b of RAMP.

A novel sequence variation was identified in exon 11b of RAMP. a) Sequence analysis of individual II:7 identified an A>C transversion at nucleotide position 34260. b) The BLAST alignment of the nucleotide sequence of individual II:7 against the RAMP coding sequence (NCBI accession number AF195765) shows the nucleotide mismatch (indicated by an asterisk).

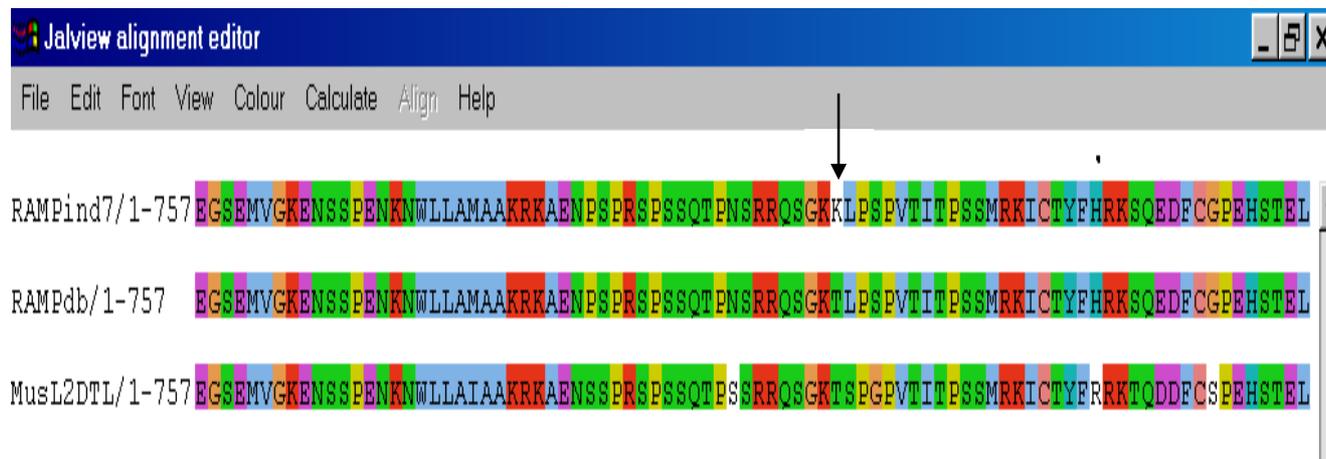


Fig. 3.11

Protein sequence alignment analysis of L2DTL.

ClustalW analysis of the alignment of the protein sequence of individual II:7 against the human and mouse L2DTL sequences, which were obtained from the NCBI database (accession numbers NP_057532 and AK054412). The position of the threonine-694 to lysine amino acid substitution identified upon sequence analyses of individual II:7 is indicated by the arrow.

Db seq = database sequence; Ind = individual identifier, nt = nucleotide; MusL2DTL = mouse RAMP protein sequence; RAMPind7 = translated protein sequence for individual II:7; RAMPdb = human RAMP protein sequence.

Table 3.8*Summary of the sequence variations identified in the present study*

Candidate Gene	Nucleotide change and position	Amino acid residue change	Allele frequency in the Afrikaner subpopulation	NCBI SNP accession number and date of submission	Other NCBI SNP entries for the same sequence variant and date of submission	Allele frequency in other populations
<i>KCNHI</i>	T>C nt 2225	Asp712Asp	T = 0.9; C = 0.1	Not submitted	Lee714008 ss1537837; 13/09/2000	No data available
<i>KCNHI</i>	C>A nt 2963	3' UTR	Not determined	Brincor ss483560 20/02/2002	IMS-JST147545 ss4976734; 12/08/2002	C = 0.384; A = 0.616
<i>VWS59</i>	G>C nt 31624	Leu178Leu	G = 0.714 C = 0.286	Brincor ss2978390 07/03/2001	YUSUKE IMS-JST041633 ss3231548; 05/09/2001 BCM_SSAHASNP ss9851849; 27/06/2003 CSHL-HAPMAP CSHL-HuFF- 200402 ss19845568; 21/02/2004	No data available
<i>LAMB3</i>	A>C nt 2591	Intronic	A = 0.525; C = 0.475	Not submitted	YUSUKE IMS-JST004433 ss2991738; 24/05/2001	A = 0.702; C = 0.298
<i>RAMP</i>	A>C nt 34260	Thr694Lys	Not determined	Not submitted	BCM_SSAHASNP ss9829563; 27/06/2003 WI_SSAHASNP ss11361027; 03/07/2003 SSAHASNP WGSA-200403 ss20430185; 18/03/2004	No data available

A = adenine; Asp = aspartic acid; C = cytosine; G = guanine; Leu = leucine; Lys = lysine; nt = nucleotide position; T = thymine; Thr = threonine; UTR = untranslated region

CHAPTER 4

Discussion

	Page
Index	117
4.1 The clinical profile of PFHBII revisited	117
4.2 Possible influences on disease presentation	119
4.3 Mapping the PFHBII-causative locus	122
4.4 Implications of mapping the PFHBII locus	125
4.5 Genealogical studies and the origins of PFHBII	126
4.6 Haplotype analysis at the PFHBII locus: diagnostic implications	129
4.7 Rationale for selecting candidate genes for mutation analyses	130
4.8 Study limitations	134
4.8.1 Linkage analysis	134
4.8.2 Bioinformatics analysis and sequence status at chromosome 1q32.2-q32.3	134
4.8.3 Limitations of laboratory techniques	136
4.8.3.1 PCR-SSCP analysis	136
4.8.3.2 Study design: screening coding and intron-exon junctions	137
4.9 Changing paradigms: implications for selecting candidate genes in future mutation screening	138
4.10 Conclusion	141

Discussion

*In theory there is no difference between theory and practice, but in practice there is.
- Anonymous*

4.1 The clinical profile of PFHBII revisited

A detailed clinical specification of a disorder is a prerequisite for accurate mapping analyses, as incomplete or incorrect clinical diagnoses may increase the chance of false linkage associations (Ott, 1999). Therefore, the present linkage study provided impetus to perform follow up examinations and reassess the PFHBII-affected family described by Brink and Torrington (1977). Three important findings that resulted from this extensive clinical investigation are discussed below.

Redefining the clinical classification of PFHBII

Since the original study by Brink and Torrington in 1977, PFHBII has been classified as a pure cardiac conduction disease, although, in 1986, A. Brink speculated that the disorder might be caused by an underlying “cardiomyopathy” (Van der Merwe et al., 1986). Consequently, the present study provides the first evidence that a cardiomyopathy is indeed a feature of the disorder, by demonstrating a familial occurrence of DCM and progression to congestive heart failure in the PFHBII-affected family (Fernandez et al., 2004). In so doing, the present study redefines the clinical classification of PFHBII, 27 years after its first description.

Familial SB and early detection

The clinical study also demonstrated that individuals with isolated SB could show progression to CHB or DCM (Fernandez et al., 2004). Although isolated SB may be considered “normal” in the general population (Josephson et al., 1991), a study by Sarachek and Leonard (1972) demonstrated a familial occurrence of this feature in a number of North American families with cardiac

conduction disease. Thus, the recognition of a familial occurrence of isolated SB in the PFHBII-affected family could facilitate the early identification of clinically affected subjects before they show progression to heart failure. The prognostic implication of early detection is important because members of the PFHBII-affected family have demonstrated progression from conduction defects to end-stage heart failure in as little as five years (Fernandez et al., 2004). By detecting the disease at an early stage, clinicians could possibly slow progression to heart failure by recommending the use of angiotensin converting enzyme (ACE) inhibitors, which, by lowering the pressure inside blood vessels, reduce the effort needed by the heart to pump blood through the body (Takano et al., 2003).

An AV conduction system defect

The third conclusion drawn from the clinical investigation was that the conduction defect in PFHBII occurs in the proximity of the AV node or Bundle of His (forming the AV conduction system) (refer to Fig. 1.6), and not in the ventricular conduction system (section 3.1) (Fernandez et al., 2004). This finding is particularly significant in a South African context, because, for many years, it has been suggested that PFHBI (a ventricular conduction system disease) (see section 1.4.3.1) and PFHBII may be clinical extremes of the same disease (P. Brink, personal communication). This misconception had also resulted in the mis-diagnosis of subjects with PFHBII as having PFHBI (P. Brink, personal communication).

Consequently, by publishing a clinical profile of PFHBII, it is envisaged that, at least in South Africa, other individuals or families with the disorder will be identified. However, this may not be a straightforward task, given that some individuals in the PFHBII-affected family have shown progression from conduction abnormalities to DCM, while others only presented with conduction defects, and one subject (individual III:26) (Table 3.1), developed DCM without preceding

conduction defects. This suggests that other factors can influence the phenotypic expression of PFHBII. The section below provides two possible reasons for the differences in the phenotype.

4.2 Possible influences on disease presentation

Geographical isolation and lifestyle

Brink and Torrington (1977) indicated that the ancestors of the family described in the present study had settled in a geographically isolated region in the Eastern Cape Province of South Africa (this district was integrated into the Southern Cape Province of South Africa in 1994). A number of subjects from the PFHBII-affected family still reside in this geographically isolated region, whereas historically, other families had relocated to present-day urban centers (A. Goosen, personal communication) (Fig. 4.1).

Members of the PFHBI-affected family that live in the isolated district have a rural lifestyle and are primarily, subsistence woodcutters (lumberjacks) (P. Brink, personal communication) - an occupation requiring a fair amount of physical activity. It is interesting that past and present clinical data or family histories have, to date, not identified DCM in any subjects from the rural Southern Cape Province community (data not shown). On the other hand, all the subjects that have been identified with DCM (individuals II:12, III:2, III:9, III:25 and III:26) (Table 3.1) are from urban centres in South Africa (Fig. 4.1). It could be argued that the availability of better medical facilities in the urban centres could account for the different diagnoses. However, all the subjects described in the present study, even those that live in the remote rural districts mentioned above, had been clinically assessed in a nearby city where there are proper ECG and echocardiographic facilities.

For all clinical examinations performed in the present study, subjects were interviewed to ascertain the influence of external influences such as exercise and alcohol consumption, the latter factor can

significantly contribute to the development of cardiomyopathy (see section 1.1) (Braunwald, 1980). It was established that the lifestyle of the rural and urban PFHBII-affected family members was similar, although the subjects that are woodcutters are physically more active. Thus, it is this writer's opinion that a physically active lifestyle may offset the progression to DCM, possibly because of the compensatory hypertrophic response that is normally associated with sustained physical exercise (Aniansson et al., 1992). This data could form the basis of an interesting future study, although, it is possible that different rural and urban dietary habits may also be a factor in the phenotypic expression of PFHBII (Dracup et al., 1994). If environmental factors do play a role in the disease presentation, the study may lead to the suggestion of lifestyle changes that could improve disease-management or prognosis.

To deviate slightly from the topic, the lives of woodcutters from the same region in which a large percentage of subjects from the PFHBII-affected family currently reside have been portrayed in the novel "*Kringe in die bos*" ("Circles in a forest") by the Dalene Matthee (Penguin Books). This narrative should provide interesting insights into the lifestyle of the family described in the present study.

Genetic modifying factors

The above hypothesis suggests external factors may contribute to variations in disease expression. However, there is an equal possibility that the differences may be due to genetic factors, or, probably, a combination of both. Although the precise causes are not yet known, it has been suggested that variability in phenotypic expression in subjects with HCM caused by the same mutation, might be caused by genetic modifying factors (Moolman-Smook et al., 2003). Thus, it is possible that modifying factors may also cause the differences in disease-progression in the PFHBII-affected family. A current investigation is being undertaken at the US/MRC Centre for

Molecular and Cellular Biology to identify possible genetic modifying factors that may play a role in the phenotypic variability in HCM and DCM (Y. Yako, personal communication), and members of the PFHBII-affected family have been entered into this important study.

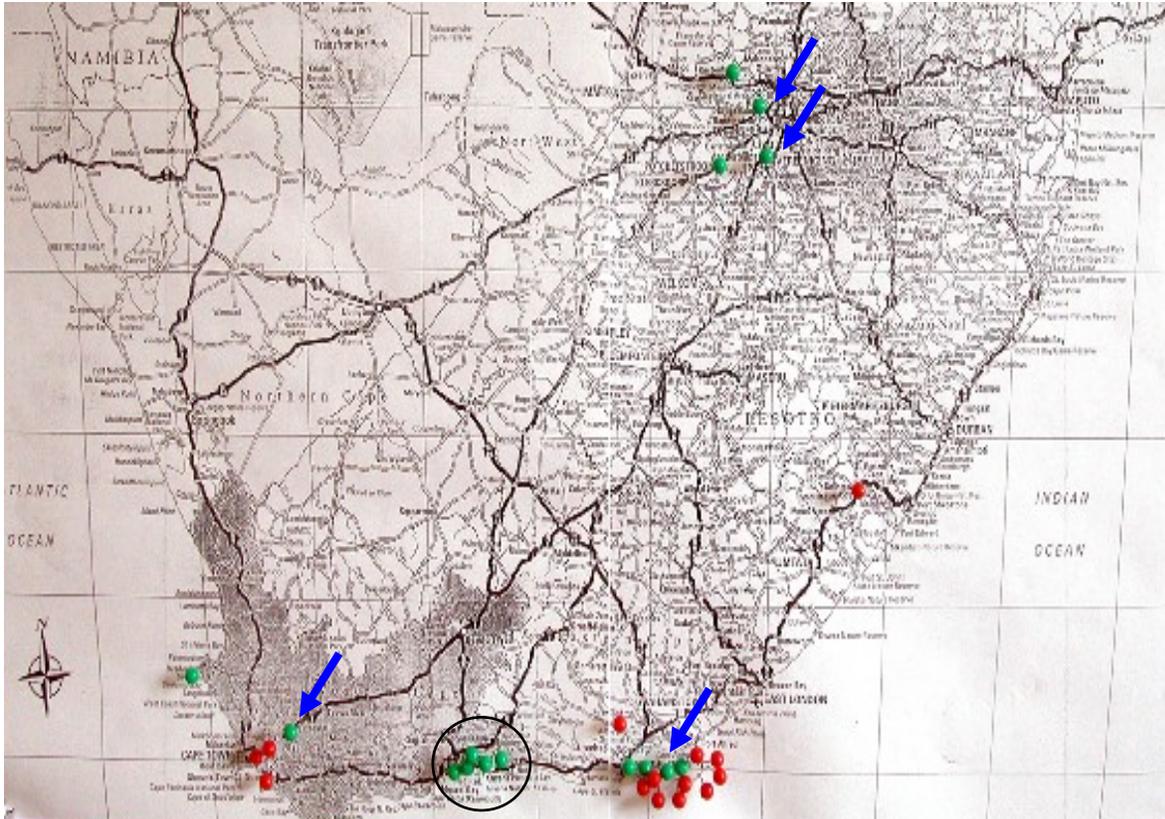


Image provided by Prof. V Corfield (US/MRC Centre for Molecular and Cellular Biology, Cape Town, RSA).

Fig. 4.1

The distribution of members of the South African PFHBII-affected family.

The green pins indicate subjects or clusters of individuals from the PFHBII-affected family and red pins indicate PFHBI-affected families (see text). A circle shows the isolated district in the Southern Cape Province of South Africa where most of the PFHBII-affected family members presently live. Arrows indicate the cities where PFHBII-affected family members with DCM have been identified.

4.3 Mapping the PFHBII-causative locus

Candidate loci linkage analyses

The literature review in Chapter 1 of this dissertation describes the clinical features and common underlying molecular aetiologies for different clinical entities, namely DCM, HCM and cardiac conduction disorders. To refresh the reader's memory and to serve as a point of reference for the discussion that follows, two examples will be recalled.

Clinically, DCM and HCM share many common clinical features, although morphological and physiological differences between familial DCM and HCM led to their classification as two separate clinical entities (Goodwin, 1970; Braunwald, 1980). Furthermore, subsequent molecular genetic analyses have resulted in the identification of common aetiologies between DCM and HCM, for example, mutations in cardiac actin and troponin T have been shown to cause both DCM (Olson et al., 1998; Kamisago et al., 2000; Li et al., 2001; Hanson et al., 2002) and HCM (Thierfelder et al., 1994; Moolman et al., 1997; Mogensen et al., 1999; Olson et al., 2000) (also refer to sections 1.3.2 and 1.3.6). Interestingly, Kamisago et al., (2000) and Hanson et al., (2002) identified the same deletion mutation (Δ Lys210) in cardiac troponin T, although the families described by Kamisago and colleagues had isolated DCM, whereas the two unrelated families described by the Hanson group presented with DCM and conduction defects. This data might demonstrate the influence of genetic modifying factors on phenotypic expression of this cardiomyopathy (see previous section).

Genetic analyses have additionally identified mutations in a cardiac Na⁺ channel that may cause the isolated cardiac conduction diseases, PCCD (Schott et al., 1999) and CCD (Tan et al., 2001), or the arrhythmic disturbances, LQT3 (Wang et al., 1995) and IVF (Chen et al., 1998) (also refer to

section 1.4.4.1). This finding is unique because conduction diseases and arrhythmias are phenotypically diverse, as the one condition is associated with a slowing or a block of the heart rate, while the other is associated with an increased or irregular heart rhythm (section 1.4.4.1).

Consequently, the present study proposed that PFHBII might share a common molecular aetiology with a previously described cardiomyopathy or cardiac conduction disease, which, thereby, provided the rationale for embarking on a candidate loci linkage study. The premise was that a genome-wide study would be undertaken if no linkage were found to a previously described disease-causative locus.

Linkage to a locus on chromosome 1q32.2-q32.3

Following a candidate locus linkage study, it was fortuitous that initial data suggested the PFHBII-causative gene was mapped to the CMD1D (*TNNT2*) locus on chromosome 1q32 (Thierfelder et al., 1994; Durand et al., 1995; Kamisago et al., 2000; Li et al., 2001; Hanson et al., 2002), because preliminary lod scores of 1.16 and 1.66 were generated with markers *DIS456* and *DIS505*, respectively (Table 3.4). However, subsequent fine mapping and haplotype analysis identified recombination events in two clinically affected subjects (Fig. 3.1a and b), placing the PFHBII-causative locus distal to *TNNT2* (Fig. 3.5).

At the time of the linkage study and fine mapping analysis, the database entries of genetic markers spanning the PFHBII locus were prone to positional rearrangements (Chromosome 1 Workshop 6 group discussion, Iowa, USA – September 2000). Therefore, it was deemed necessary to construct an integrated map of the PFHBII locus, in order to assess the correct specification of genetic map data and to update it regularly during the course of this study (section 3.4) (also refer to Le Hellard et al., 2001 and Nievergelt et al., 2004). This analysis showed that all the selected genetic markers

at chromosome 1q32.2-q32.3 (Table 2.1) were reconciled onto a physical map of the region (Fig. 3.5). In so doing, the haplotype analysis in the present study was authenticated, thereby unequivocally excluding *TNNT2* from being the PFHBII-causative gene.

Testing lod score stability at chromosome 1q32.2-q32.3

For linkage analysis, the correct clinical designation of subjects is important because assigning a clinically affected status to individuals who are, in reality, clinically unaffected may result in false-positive linkage associations (Ott, 1999). As SB may be common in the general population (Josephson et al., 1991), and because of concerns of basing a clinical diagnosis on only one ECG examination, the lod scores at chromosome 1q32.2-q32.3 were re-calculated by changing the clinical status of individual III:11 (Table 3.1) to a clinically uncertain status (see section 3.2.2). These analyses generated a maximum pairwise lod score of 2.83 and a multipoint score of 3.4 with markers *DIS3753* and *DIS414*.

Thus, re-assessment of the lod scores produced a maximum pairwise value that was below the significance level for linkage, although the maximum multipoint lod score remained above +3. The low pairwise scores generated in this study were most likely the result of the limited pedigree size or uninformative marker alleles (Ott, 1999); an example of the latter is indicated by individual II:12, who is homozygous for alleles at the markers *DIS414*, *DIS2810*, *DIS2780*, *DIS425* (Fig. 3.1). However, a multipoint analysis provides a stronger indication of linkage when pedigree size or uninformative markers cause a reduction in statistical power (Lathrop and Lalouel, 1984; Ott, 1999). Consequently, the additional tests that were performed statistically support linkage of PFHBII to the locus to chromosome 1q32.2-q32.3.

Simulated lod score analysis using an alternative phenotypic classification model

To test the robustness of the maximum lod score, various parameters were changed in order to “reduce” the score, thereby assessing whether it remained above or fell below the significance level of +3. It has to be noted that the strict (conservative) clinical classification model applied in the present study. This model was applied to avoid clinical misclassification of subjects and, in so doing, reducing the chances of generating false positive linkage. Unfortunately, the maximum obtainable lod score that would be generated by this model was also reduced.

To determine the “maximum” possible lod score that could be generated with the available data, a second (less conservative) phenotypic classification model was used in the pairwise and multipoint lod score analysis. If a less strict but accepted definition of SB was used, namely, a heart rate of less than 60 bpm (Josephson et al., 1991), then three additional subjects (individuals II:5, II:9 and III:15) (Table 3.1) (Fig. 3.1) could be assigned a clinically affected status, in that way increasing the maximum obtainable lod score (Ott, 1999). In order to assess the influence of identifying additional clinically affected subjects on the lod score calculations, the clinical status of individuals III:16 and IV:1, who present with “unusual” clinical features and carry the disease-associated haplotype (Table 3.1) (Fig. 3.1), was changed from clinically uncertain to clinically affected. This analysis generated a maximum pairwise lod score of 4.55 and a maximum multipoint score of 4.9 (data not shown). Consequently, this data indicates that the further identification of clinically affected subjects should strengthen the linkage data presented in this study.

4.4 Implications of mapping the PFHBII locus

Publication of mapping of the PFHBII locus (Fernandez et al., manuscript submitted) may lead to other investigators who have identified families with DCM and conduction defects, isolated DCM, isolated cardiac conduction abnormalities or HCM-associated disorders, to perform linkage analyses at this

locus. These analyses could map an analogous disorder to the locus described in the present study, which may help to refine the PFHBII locus. The French group that had mapped an analogous cardiac conduction disorder to the same interval on chromosome 19q13.3, to which the PFBHI-causative gene is mapped, used this approach to refine the ICCD locus (Z.Arieff, personal communication) (see section 1.4.3.1).

Additionally, the present linkage study is particularly relevant to South African families, because it demonstrates the exclusion of the PFHBI-causative locus (Brink et al., 1995; de Meeus et al., 1995) as cause of PFHBII. Historically, PFHBI and PFHBII have both been shown to segregate in different South African Caucasian Afrikaner families (Brink and Torrington, 1977) that, coincidentally, lived in the same geographical regions (Fig. 4.1). Previously, it was assumed that PFHBI and PFHBII were extremes of the same disease, and, therefore, had the same underlying molecular aetiology (P. Brink, personal communication). However, by showing that PFHBII is an AV conduction defect and by mapping the PFHBII locus, the present study has demonstrated that PFHBI and PFHBII are not only clinically distinct (see section 4.1), but also, that these clinical entities are caused by a defect in different genes.

4.5 Genealogical studies and the origins of PFHBII

Brink and Torrington (1977) speculated that the branch of the PFHBII-family, which they had investigated, descended from one of the four sons of a Dutch immigrant that arrived in South Africa in 1713, although in the article, they did not speculate about the genetic origin of the disorder. Later, Torrington (1979) traced additional branches of the PFHBII-affected family by following the family surname of individuals described by Brink and Torrington (1977). This data appeared to support the notion that there might be more South African families with the PFHBII. However, a limitation of that study was the use of a family history of “any heart condition” as a criterion to

trace “affected“ subjects.

New genealogical data suggesting a recent genetic origin of PFHBII

The present study obtained genealogical data from the kindred in which PFHBII segregates by assessing the information that Torrington had collected in 1979. Additionally, all first- and second-degree relatives and progeny of subjects that had not been described by Brink and Torrington (1977) or Torrington (1979) were interviewed telephonically or in person. Although archived records dating back to the 1700’s were obtained, family records derived from the “*Familie Bybel*” (Family Bible), which is a valuable source of genealogical information, particularly among the South African Afrikaner subpopulation (Hayden et al., 1980), provided accurate information on the personal histories of ancestral subjects from the PFHBII-affected family. This extensive investigation demonstrated that PFHBII could only be traced in five generations of the family described in the present study and not in other branches of the extended kindred (Fig. 4.2). This data might suggest that the PFHBII-causative mutation arose *de novo* in individual I:1 (Table 3.1) (Fig. 3.1) (also, the individual indicated by an arrow in Fig. 4.2).

An additional finding resulting from the extensive genealogical and clinical investigation

The extensive genealogical and clinical investigation identified a 55-year-old male subject that presented with HCM (see section 3.1). Although the subject was related to members of the PFHBII-affected family, he was shown to be in a branch of the family in which the disorder does not segregate (Fig. 4.2) (P. Brink, personal communication). Initially, it was thought that the subject might have a familial form of HCM, given the prevalence of this condition in South Africa (J. Moolman-Smook, personal communication). However, subsequent examinations of the subject’s siblings and progeny did not identify other individuals with HCM (Fig. 4.2). Therefore, it is unlikely that the subject has familial HCM, although this finding does not preclude the future

identification of HCM in this branch of family. It has to be viewed that any data derived from this study may contribute to other studies on cardiac diseases in South Africa, or even, in other countries. For this reason, the subject reported in this sub-section has given his consent to be entered into a comprehensive study to determine the underlying causes of HCM in South Africa (J. Moolman-Smook, personal communication).

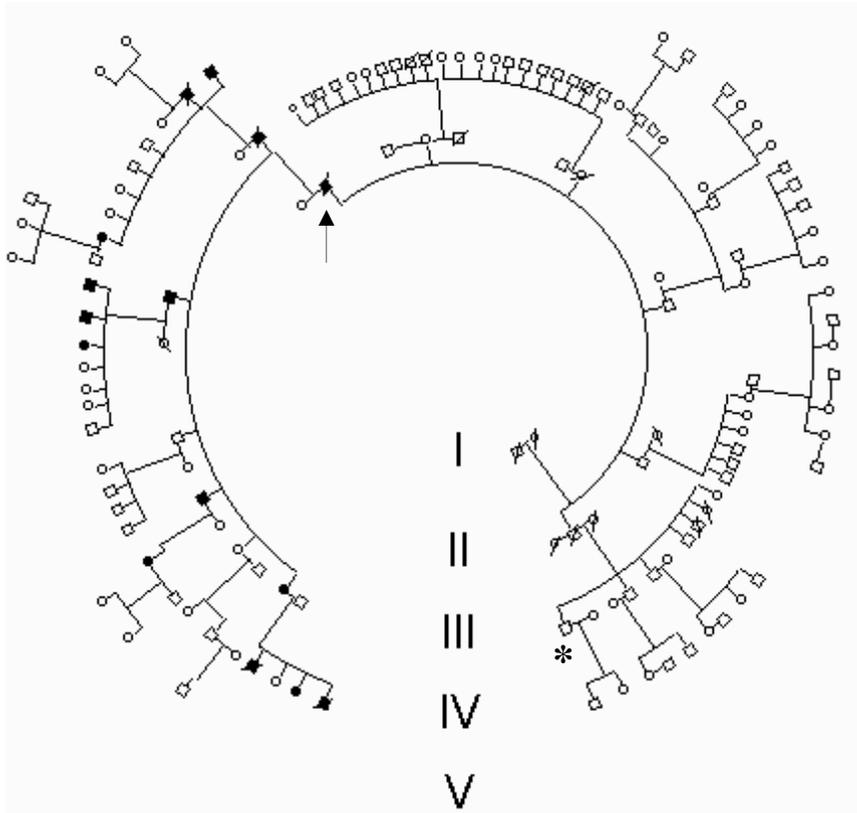


Fig. 4.2

Tracing the origins of PFHBII by extensive pedigree analysis.

The PFHBII-affected family described in the present study, and previously by Brink and Torrington (1977) is shown on the left half of the pedigree. Genealogical studies did not trace the segregation of PFHBII beyond five generations (indicated by roman numerals) in this kindred. The arrow indicates the subject previously designated individual I:1 in Table 3.1 and Fig. 3.1, who is presumed to have introduced the PFHBII-causative defect into the family described in this study.

*Solid symbols = PFHBII-affected subjects; * = subject presented with HCM.*

4.6 Haplotype analysis at the PFHBII locus: diagnostic implications

Haplotype analysis identified individuals with the disease-associated haplotype who presently do not manifest clinical signs of PFHBII (individuals II:5, II:9, III:15, III:16, III:19, III:22, IV:1, IV:3 and IV:5) (Fig. 3.1), although, for most of these subjects, “unusual” cardiac features were noted. Individual II:5 had previously been diagnosed with SB (Brink and Torrington, 1977), while a previous assessment of individual II:9 had identified a slow heart of 49 bpm. However, both these subjects had had follow up clinical examinations that indicated heart rate values of 53 bpm and 58 bpm (Table 3.1). Therefore, for the linkage analysis performed in the present study, these subjects were assigned a clinically uncertain status (Table 3.1). From the clinical aspect of their future management, however, further follow-up clinical examinations have been recommended for these subjects, because Alboni et al. (2001) suggested that a slow heart rate may be an indicator of forthcoming heart failure in individuals older than 70 years. Another haplotype carrier (individual III:15) had a heart rate of 54 bpm (Table 3.1). Additional examinations should be performed to determine whether the subject’s slow heart rate was physiological (possibly exercise-induced) or, the result of the disorder. Additionally, clinical assessment of individual IV:3 did not identify any atypical features, while his sister, individual IV:5, was not examined because she is younger than 20 years of age (see section 2.3).

In addition ECG examinations indicated that individuals III:19 and III:22 (Table 3.1) (Fig. 3.1) were at the borderline of a diagnosis of first-degree AV block (P-R intervals of 200 milliseconds) (Table 3.1), although the significance of this observation, if any, has yet to be established. Characteristics that suggested early-clinical markers of PFHBII were identified in individuals III:16 and IV:1 (Table 3.1) (Fig. 3.1). Individual III:16 displayed a slow R-wave progression, which may be associated with DCM (Fatkin and Graham, 2002), whilst individual IV:1, who had been interviewed in 1997, reported severe chest pain, dyspnoea and syncopal episodes. However, the

possible progression of this individual's symptoms have not been evaluated clinically although, contact has recently been re-established and the subject has reported a similar clinical picture. It has been recommended that continuous clinical evaluations be performed on members of the PFHBII-affected family (P. Brink, personal communication), thereby not only allowing early detection of the disease, but also, to provide a sense of reassurance for concerned members of the family.

4.7 Rationale for selecting candidate genes for mutation analyses

Mutations in ion channels and ion channel modulators have primarily been implicated in the pathogenesis of isolated cardiac conduction disorders and arrhythmias, while mutations in genes encoding cytoskeletal, sarcomeric, nuclear envelope, intracellular Ca^{2+} -regulating and energy homeostasis proteins have been identified as causes of cardiomyopathies (Schott et al., 1999; Schönberger and Seidman 2001; Seidman and Seidman 2001; Tan et al., 2001; Fatkin and Graham, 2002; Ashrafian et al., 2003; Towbin, 2003; Watkins, 2003). Equipped with this background, the following sections will provide the rationale for selecting each of the candidate genes that were screened in the present study.

LAMB3

Mutations in the gene encoding extracellular matrix protein laminin- α 2 chain of merosin (*LAMA2*) have been shown to cause early-onset DCM with muscular dystrophy in children (Spyrou et al., 1998) (Table 3.1). In addition, Bies et al., (1992) demonstrated that extracellular matrix proteins are expressed in cardiac conduction tissue. Thus, *LAMB3* was a candidate gene for PFHBII because it encodes the cardiac expressed extracellular matrix protein laminin β 3. However, no disease-causative mutation was found in *LAMB3*, although a synonymous polymorphism, which did not segregate with the disorder, was identified (Table 3.7).

T3JAM

Dadgostar et al., (2003) recently demonstrated that *T3JAM* encodes a TRAF3-interacting Jun N-terminal kinase-activating modulator protein that interacts with cytoskeletal proteins of the myosin class II heavy chain family (Table 3.6). In Chapter 1, sections 1.3.2, 1.3.3 and 1.3.6, the effect of alterations in cytoskeletal protein function and its role in the pathogenesis of DCM and HCM has been demonstrated. Another reason for selecting *T3JAM* for mutation screening is because it may also serve as an adaptor molecule that modulates the localisation of TRAF3-interacting Jun N-terminal kinase (Dadgostar et al., 2003). This finding is significant because recently Mohler et al., (2003) described a LQTS type 4-causative mutation in ankyrin-B, a membrane adaptor protein that positions the subunits of ion channels, themselves implicated as cause of LQTS (see section 1.4.4 and 1.4.5), to specific domains in the sarcolemma and endoplasmic reticulum. Therefore, the TRAF3-interacting Jun N-terminal kinase-activating modulator was screened, although these analyses did not identify the PFHBII-causing mutation.

VWS59

Transcript data derived from EST analyses suggests *VWS59* is highly expressed in heart tissue (Unigene accession number Hs.58650). Moreover, bioinformatics-based similarity analyses indicated that the protein expressed by *VWS59* might have a role similar to acyltransferases (Table 3.6). Mutations described in the tafazzin protein, a member of the acyltransferase family of proteins, have been shown to cause the DCM-associated condition known as BTHS (Bione et al., 1996). No disease-causative mutation was identified in *VWS59*, although a polymorphism was identified that segregated in the PFHBII-affected family and was present in an unaffected Caucasian Afrikaner panel (Table 3.7).

KCNH1

Ion channel activity and intracellular ion homeostasis have been shown to be crucial for the functioning of the cardiac conduction and contractile apparatus (refer to section 1.3.3.2, 1.3.6.2.2 and 1.4.4). Furthermore, voltage-gated K^+ channels have functions that include regulating neurotransmitter release, heart rate, insulin secretion, neuronal excitability, epithelial electrolyte transport and smooth muscle contraction (Occhiodoro et al., 1998). Therefore, a human-ether-a-go-go related K^+ channel gene, *KCNH1*, which is related to *KCNH2*, the gene in which mutations have been described that cause LQT2 (Curran et al., 1995), was deemed a strong PFHBII-candidate gene. The disease-associated mutation was not identified upon screening of *KCNH1*, although two polymorphisms were identified that segregated in the PFHBII-affected family and were present in an unaffected Caucasian Afrikaner panel (Table 3.7).

KIAA0205

Mutations in the G4.5 gene encoding tafazzin have been identified as cause of X-linked recessive BTHS, which may be associated with DCM (Bione et al., 1996). *KIAA0205* was considered a plausible PFHBII-candidate gene because it is predicted to encode a putative acyltransferase similar to tafazzin (Table 3.6). However, the PFHBII-causative mutation was not detected in *KIAA0205*.

RAMP

The nuclear matrix-associated gene, *RAMP*, encodes the L2DTL protein, a G-protein containing a WD40 repeat sequence (Cheung et al., 2001). In eukaryotic proteins, WD40 domains can function as adaptor/regulatory modules in signal transduction pathways, or are involved in pre-mRNA processing and cytoskeleton assembly (Garcia-Higuera et al., 1996). Additionally, a mutation has been identified in a WD40 repeat in the peroxin-seven gene (*PEX7*), which causes Refsum disease (MIM 266500), a condition characterised by progressive adult retinitis pigmentosa, peripheral

neuropathy, cerebellar ataxia, sensorineural deafness, skeletal dysplasia and cardiac arrhythmias (Braverman et al., 2002). Therefore, *RAMP* was screened for the PFHBII-causative mutation. The PFHBII-causative mutation was not identified, but a non-synonymous polymorphism, which caused a substitution of a species-conserved threonine amino acid residue, was identified in only two of the eight PFHBII-affected subjects (section 3.6.1) (Table 3.7). The sequence variation was not identified in the Afrikaner population, although it has to be noted that only five subjects from the control panel were subjected to direct sequence analysis (see section 3.6.1). It is possible that this polymorphism may have a modifying role in the pathogenesis of PFHBII in these subjects. Clinically, the two subjects with the sequence variation in *RAMP* (individuals II:7 and III:20) (Fig. 3.1) showed progression to CHB necessitating a pacemaker implantation (Table 3.1). This clinical presentation does not appear to be different from the rest of the clinically affected members of the family, although it has to be noted that individual III:20 progressed from first-degree AV block to CHB within two years (data not shown). Unfortunately, time constraints did not permit further investigation, although this data could be used in an investigation to identify potential genetic modifying factors in cardiomyopathies (see section 4.2).

PPP2R5A

Serine/threonine protein phosphatases have regulatory roles that include growth factor signaling and cell transformation (McCright and Virshup, 1995). Previously, by altering the delta 5 subunit of protein phosphatase 2A, Brewis and colleagues (2000) produced transgenic mice that presented with DCM. Additionally, the Ca^{2+} -dependent protein phosphatase, calcineurin, has been shown to induce foetal cardiac gene expression with resultant cardiac hypertrophy in mice (Molkentin et al., 1998). Thus, *PPP2R5A* was considered a good PFHBII candidate gene. However, the disease-associated mutation was not identified upon analysis of this gene.

4.8 Study limitations

4.8.1 Linkage analysis

Chance linkage associations

The present study has provided evidence that supports mapping the PFHBII-causative gene to a locus on chromosome 1q32.2-q32.3. As a candidate, rather than a whole-genome, linkage study was undertaken, the possibility that a locus that was not analysed could produce significant lod scores to suggest the causative gene lies elsewhere cannot be excluded. Generally, chance linkage associations are probable in one in twenty loci screened (Morton, 1955). However, the present investigation excluded a number of DCM, HCM and cardiac conduction-causative loci (Table 3.3). Thus, it can be presumed that for PFHBII chance associations are likely to be much less than 5%, because sequential exclusion analyses decrease the genomic interval that has to be assessed in order to map a disease to its causative locus (Morton, 1955; Ott, 1999). Furthermore, barring a whole-genome linkage analysis, the present investigation has performed numerous tests to assess the robustness of the lod score analyses at chromosome 1q32.2-q32.3 (section 3.2.2). Although a maximum lod score of +3 is used to establish linkage, it is widely viewed that a score of greater than +3.3 provides a more significant indication that a disease is linked to a locus (Ott, 1999). In the present study, the lowest simulated multipoint lod score generated was +3.4 (section 3.2.2) when applying a very cautious (conservative) assessment.

4.8.2 Bioinformatics analysis and sequence status at chromosome 1q32.2-q32.3

Gene annotation and prediction

Data obtained from the NCBI database builds 28 and 31 (freeze dates 24 December 2001 and 5 November 2002) were used to delineate candidate genes at the PFHBII locus. At the time, a total of 27 genes (known and computer predicted) were assigned to the PFHBII locus (Appendix G), of which seven genes were prioritised for mutation analysis. Furthermore, the sequence at the PFHBII

locus, between markers *DIS205* and *DIS505*, was in the draft phase (see section 1.5.2.1) when selecting the seven candidate genes for mutation screening. Recent analyses of the Ensembl database in December 2003 indicated that one computer-predicted gene (*ENSG00000173964*) (Appendix G) had been removed from the interval between *DIS70* and *DIS505* on chromosome 1q32.2-q32.3. Thus, 26 genes map to within the interval of 2.85Mb, which gives an average gene density of one gene per 110kb of genomic sequence. This figure is about the same as the overall gene density of chromosome 1, which is approximately one gene per 116kb (see section 1.5.3.2), though significantly lower than that of chromosome 19, which has a density of one gene per 26kb of genomic sequence (Grimwood et al., 2004).

Exon prediction

Due to sequencing errors, or incomplete sequence data, it has been shown that computer algorithms used to annotate genes may miss, or incorrectly predict, the exon-intron boundaries of genes (Burge and Karlin, 1997; Felsenfeld et al., 1999; Bork, 2000; Katsanis et al., 2001; Stein, 2001). This holds true for the seven candidate genes that were screened in the present study, even a well-characterised gene such as *PPP2R5A* (McCright and Virshup, 1995), for which the genomic organisation had been defined about a decade ago. Moreover, as sequence quality improves, additional exons could be identified within these genes. A recent database search indicated that four of the seven genes screened in the present study have had additional exons or alternative transcripts assigned after the 5 November 2002 freeze date (Table 4.1). Thus, it is possible that the PFHBII-causative mutation may reside within a newly identified exon or an alternative transcript. Future mutation analysis will therefore focus on screening the newly identified exons and alternative transcripts in the genes that were investigated in the present study, thereby permitting the exclusion of a PFHBII-causative mutation within the coding regions of these genes.

Table 4.1

Comparison of database entries showing changes to exon numbers of the candidate genes screened at the PFHBII locus

Gene	NCBI build 31	NCBI build 33
	5 November 2002	14 April 2003
<i>LAMB3</i>	23 exons	23 exons
<i>T3JAM</i>	17 exons	17 exons
<i>VWS59</i>	8 exons + 1 alternative transcript	12 exons + 1 alternative transcript
<i>KCNH1</i>	11 exons	11 exons
<i>KIAA0205</i>	7 exons	8 exons
<i>RAMP</i>	12 exons + 1 alternative transcript	15 exons + 1 alternative transcript
<i>PPP2R5A</i>	11 exons	13 exons

4.8.3 Limitations of laboratory techniques

4.8.3.1 PCR-SSCP analysis

Mutation analysis by PCR-SSCP gel electrophoresis was previously optimised in our laboratory to increase detection sensitivity to approximately 95-100% (W. de Lange, PhD thesis). However, this meant that, theoretically, up to 5% of sequence variations could have been missed during mutation screening. In the present study, 92 exons, or a total of 101 “amplification products”, if one considers that more than one set of primers were used to amplify across large exons, were screened by PCR-SSCP analyses (refer to Appendix E). In order to assess the possibility of missing the PFHBII-causative mutation, the sensitivity of the PCR-SSCP screening analyses was tested, as described in section 2.12.5.3. Consequently, 29 exons were screened by direct sequence analysis, of which two demonstrated a sequence variation that was not detected by PCR-SSCP gel electrophoresis analysis (refer to section 3.6.1). For this study, the sensitivity of the PCR-SSCP analyses was approximately 93% (2 out of 29 exons), which is consistent with the sensitivity level determined by W de Lange (PhD thesis), and more than the 70% and 85% sensitivity level demonstrated by Sheffield et al., (1993) and Larsen et al., (2001), respectively. However, it has to be noted that an amplified product with a known sequence variant (a positive control) was not

loaded onto any of the PCR-SSCP gels. Therefore, the mutation screening protocol did not provide a measure to assess whether there was variability in the gel-detection sensitivity over time. Consequently, in future, it may be justified to perform automated sequence analyses of the prioritised candidate genes screened in the present study, since this method is more sensitive and there is a lesser likelihood that the strategy will show detection variability (Kaczorowski and Szybalski, 1996).

4.8.3.2 Study design: screening coding and intron-exon junctions

Mutations that cause autosomal dominant cardiac conduction disease or cardiomyopathies have, by and large, been identified in the exons of genes (Marian and Roberts, 2003). Additionally, splice junctions are usually analysed during screening protocols because mutations that can alter exon splicing, and thus result in truncated protein products or non-functional proteins, have been identified as cause of conditions such as HCM (in *MyBPC3* and *TNNT2*) (Watkins et al., 1995; Watkins et al., 1996) or cardiac conduction disease (in *SCN5A*) (Schott et al., 1999) (refer to sections 1.3.6.2.1 and 1.4.5). In general, intronic sequences are not screened for mutations as they are long and, therefore, not suitable for PCR-SSCP analysis (Orita et al., 1989). Moreover, a larger number of sequence variations occur within introns (Louie et al., 2003), but, more importantly, even if a variant is identified in an intron, it is difficult to determine the significance thereof because the functional role of intronic sequences has yet to be clearly defined (Mattick, 1994). These limitations also apply to analyses of promoter regions, although, functional significance can be assigned to variations identified in canonical *motifs* such as transcription factor-, enhancer- and repressor-binding sites (Frith et al., 2004).

Because of the considerations mentioned above, the present study analysed exons and the splice junctions of each of the seven selected candidate genes. However, Towbin et al., (1993) and

Muntoni et al., (1993) demonstrated associations between a deletion mutation in the promoter sequence of dystrophin and XLCM. Furthermore, promoter mutations in genes encoding cardiotrophin 1 (Erdmann et al., 2000), phospholamban (Minamisawa et al., 2003) and serum response factor (Zhang et al., 2003) have been identified as cause of human and mouse cardiac conduction or cardiomyopathy-associated disorders. Consequently, the candidate genes screened in this study cannot conclusively be excluded as cause of PFHBII because the regulatory regions were not screened by PCR-SSCP analysis.

Because a deletion mutation in the promoter region of dystrophin has been reported to cause XLCM by Muntoni et al., (1993), it is suggested that restriction enzyme- and Southern blot analyses (Sambrook et al., 1989; Armour et al., 2002) be used in future screening of the candidate genes that were assessed in the present investigation. This analysis might not only identify large deletions, but also other chromosomal rearrangements such as inversions or duplications, or, it could fortuitously result in the identification of a point mutation that changes a restriction enzyme site used in the analysis.

4.9 Changing paradigms: implications for selecting candidate genes in future mutation screening

Initially, the available literature had demonstrated that mutations in cytoskeletal, sarcomeric, ion channel and ion channel modulating proteins (sections 1.3.3.1.3, 1.3.3.1.3, 1.3.6.2.1 and 1.4.5.1) were primarily implicated in the pathogenesis of cardiomyopathies and cardiac conduction disorders. The more recent identification of common disease mechanisms illustrating the role of Ca^{2+} handling and energy homeostasis in the pathogenesis of DCM and HCM (refer to sections 1.3.3 and 1.3.6) suggested that there could be additional molecular causes of cardiomyopathies (Watkins, 2003). Consequently, genes associated with Ca^{2+} or energy pathways, which map to the

PFHBII locus, warrant future investigation. Immediately prior to the submission of this dissertation, using bioinformatics tools, all the genes that map within the PFHBII locus, barring the seven screened in the present study, were re-assessed to include criteria for involvement in Ca^{2+} or energy homeostasis, for example Ca^{2+} or ATP-binding sites. A few of the plausible candidate genes that complied with these criteria are discussed below.

Synaptotgamin XIV (SYT14)

SYT14 (Ensembl accession number AK091517) encodes a protein composed of conserved C2 domains, which are also components of protein kinase C isozymes, phospholipases and synaptotamins (Davletov and Sudhof, 1993). The C2 domain of protein kinase C and phospholipases are thought to be involved in Ca^{2+} -dependent phospholipid binding (Davletov and Sudhof, 1993). Additionally, model animal studies have shown an association between beta-adrenergic receptors, the phospholipase/protein kinase C pathway and left ventricular remodeling (Dorn, 2002), while a C2 domain-containing protein, namely, phosphoinositide 3-kinase gamma (PI3Kgamma), has recently been shown to trigger hypertrophy, fibrosis and cardiac dysfunction in response to beta-adrenergic stimulation (Oudit et al., 2003).

REST co-repressor (NM_018254)

NM_018254 (Ensembl accession number BC031608) encodes the REST co-repressor protein (RCO3), which consists of ELM2 domains, myb (beta leucine zipper) domains and an ATP/GTP binding motif (Nagase et al., 1996). The ELM2 and myb domains of RCO3 are involved in protein-protein interactions and DNA-binding, respectively, while ATP binds to the ATP/GTP motif of the protein. The recent identification of cardiomyopathy-associated mutations in proteins involved in cellular energy homeostasis (Blair et al., 2001) has given credibility to screening genes encoding proteins with affinities for high-energy substrates such as ATP. In addition, database analyses

indicated that RCO3 is homologous to the KIAA1343 protein, which, in turn, has a weak similarity to rabbit myosin light chain kinase 2 (accession number P07313). Furthermore, Takio et al., (1985) demonstrated that rabbit myosin light chain kinase 2 regulates cardiac and skeletal muscle contraction, thus making *NM_018254* an attractive PFHBII candidate gene.

Never in mitosis A-related kinase 2 (NEK2)

NEK2 (Ensembl accession number U11050) encodes a serine/threonine protein kinase which is involved in cell cycle control and ATP-binding (Schultz et al., 1994). A binding affinity for the energy substrate ATP, and thus its potential involvement in cellular metabolism and energy homeostasis, makes *NEK2* a good candidate for PFHBII. Furthermore, database analysis indicates that *NEK2* protein domains have similarity to phosphorylase B kinase (personal observation), which catalyses the phosphorylation of serine substrates in cardiac troponin I, in which a mutation has been described that causes autosomal recessive DCM (Murphy et al., 2004).

Zinc transporter 1 (SLC30A1)

SLC30A1 (Ensembl accession number AF323590) encodes zinc ion transporter 1 (ZNT-1). Generally, it has been viewed that mutations in Na⁺ and K⁺ ion channels cause cardiac conduction disorders (CCD and PCCD) (Schott et al., 1999; Tan et al, 2001) and arrhythmic disturbances (BrS, IVF and LQTS) (Abbott et al., 1999; Chen et al., 1998; Curran et al., 1995; Splawski et al., 1997; Tyson et al., 1997; Wang et al., 1995 and 1996). However, transgenic mice with an altered zinc transporter 5 gene (*znt5*) have been shown to develop AV block, SB, osteopenia and sudden cardiac death (Inoue et al., 2002).

Preferably, all genes that map to within the PFHBII locus should be screened for the disease-causative mutation. However, such a task would be too expensive to perform in our laboratory.

Consequently, following further screening of the candidate genes already analysed in the present study, the future screening of the plausible candidate genes mentioned above, will receive priority.

4.10 Conclusion

For 27 years, PFHBII has been described as a cardiac conduction disorder that segregated in a Caucasian South African Afrikaner family. However, the present study has demonstrated that PFHBII is, in fact, a DCM-associated disorder complicated or preceded by cardiac conduction defects. This finding has been published and will undoubtedly influence the diagnoses of other individuals with PFHBII, as well as on the treatment and management of the disease, particularly in a South African setting, but maybe in other countries as well (Fernandez et al., 2004).

Additionally, the present study describes extensive genetic analyses that have resulted in mapping the PFHBII-causative gene to within a 3.9cM region on chromosome 1q32.2-q32.3 (Fernandez et al., manuscript submitted). The disease-associated mutation was not identified in an analysis of seven strong candidate genes and continuous and future mutation screening will be pursued. It is possible that mapping of the PFHBII-causative locus may allow other investigators to show linkage of an analogous disorder to the chromosomal region described in this study, which may help further refine the locus and thereby contribute to identifying the disease-causative gene. Because PFHBII is an adult-onset disorder that shows progression of clinical features, it is possible that individuals who are currently diagnosed as clinically unaffected, could, in future, present with symptoms of the disorder. Follow-up examinations have been recommended as these measures might allow the timely detection of the disease before it progresses to congestive heart failure, a stage associated with a poor prognosis. If detected early, subjects may be prescribed appropriate drug therapies that could prevent or slow the progression of the disorder.

To conclude, by identifying a novel disease-causative locus, it is proposed that the future identification of the underlying molecular aetiology of PFHBII might add weight to support existing hypotheses, or, alternatively, could result in the elucidation of a novel pathobiology for cardiomyopathies complicated by conduction defects. As a consequence, this data might provide new insights into an understanding of the complex mechanisms that result in heart failure, which in turn, might aid in the development of new therapeutic targets.

A life spent making mistakes is not only more honorable, but more useful than a life spent doing nothing.
George Bernard Shaw

APPENDIX A

EBV-medium

EBV-containing supernatant was harvested from a culture of lymphoblastoid marmoset cell line B95-8, which secretes and is permanently infected with EBV-virus. The supernatant was centrifuged at 1200rpm in a bench-top centrifuge and then filtered through a 0.45µm membrane (Millipore, USA) to remove all viable cells. The medium was stored at 4°C for up to 6 months.

Stock solutions

Na-EDTA (1L stock solution)

NaCl (Merck)	18.75ml of 4M solution
EDTA (Biorad)	250ml of 100mM solution
Mix well to total volume	

10% SDS (1 litre stock solution)

SDS (Sigma)	100g
make up to 1 litre with dH ₂ O	

Tris-Cl (1L stock solution)

Tris (Merck)	121.14g
Make up to 1L with dH ₂ O	
pH to 8.0 with HCl	

5M NaCl (1 litre stock solution)

NaCl	292.2g
Make up to 1 litre with ddH ₂ O	

4M NaOH (1 litre stock solution)

NaOH (Merck)	160g
Make up to 1 litre with ddH ₂ O	

20xSSC (1litre stock solution)

NaCl (Merck)	250.6g
NaCitrate (Merck)	176.4g
pH to 7.0, autoclave	

DNA extraction from blood

Chloroform/octonol

Chloroform (Merck)	24vol.
Octonol (Merck)	1vol.

Phenol/chloroform

Phenol (Sigma)	25vol. (1xTE saturated)
Chloroform (Sigma)	24vol.
8-hydroxyquinoline (Merck)	

Na-EDTA/SDS

Na-EDTA (stock)	450µl
10% SDS (Sigma)	50µl

Na-acetate (3M)

Na-acetate.3H ₂ O (Merck)	40.81g
dH ₂ O	50ml

Adjust to pH 5.2 with glacial acetic acid (Merck) and make up to 100ml final volume with dH₂O

Phosphate buffered saline

NaCl (Merck)	1mM
KCl (Merck)	2.68mM
Na ₂ HPO ₄ (Merck)	141.96mM
KH ₂ HPO ₄ (Merck)	141.96mM

Cell lysis buffer

Sucrose (Saarchem)	0.32mM
Triton X-100 (Sigma)	1%
MgCl ₂ (Merck, BDH)	5mM
Tris-HCl pH 7.6 (Merck)	25mM

Adjust to pH 8.0. Store at 4°C.

RPMI-medium

RPMI (Separations)	10.4g
Fetal calf serum (Centrolab)	15%
Pen/ Strep (Boehringer Mannheim)	1%

pH to 7.5, make up to 100ml and store at 4°C.

Gel electrophoresis**10xTBE (per 1000ml)**

Tris (Merck)	108g
Boric acid (Merck)	55g
EDTA (Biorad)	9.3g

Mix well and make up to 1000ml with dH₂O.

1xTBE (per 1000ml)

10xTBE stock 100ml

Add 900ml H₂O to a final volume of 1000ml

1xTE

Tris-HCl pH 8.0 (Merck)	10mM
EDTA (Biorad)	1mM

10% Ammonium persulphate (AMPS)

AMPS (Merck)	10g
--------------	-----

Add H₂O fill to 100ml

CA loading dye

Formamide (Sigma)	95%
Bromophenol blue (Merck)	0.02%
Xylene cyanol (Biorad)	0.02%

Polyacrylamide 6% denaturing (genotyping/sequencing) gel (60ml)

Urea	26.64g
10xTBE	6ml
H ₂ O	26.6
40% Acrylamide/37.5:1 Bis (Promega)	7.7ml
10% AMPS (Merck)	500µl
TEMED (Promega)	50µl

Methodolgy

Glass plates with dimensions of 30x40x0.5cm were prepared by silanising the surface of the notched plate with Wynns C-Thru (Wynns (Pty) Ltd, South Africa). Two 0.5mm plastic spacers were placed on the longitudinal edges of one plate and overlaid by the other plate. The edges of the plates were sealed in a U-shaped sealing boot (Laboratory Specialist Services). The 6% denaturing gel solution was poured between the plates with the smooth side of a shark-toothed comb placed in 5-8mm at the notched end of the solution. The gel was allowed to polymerise through the reaction of the AMPS and TEMED for at least 2 hours. Upon removal of the sealing boot and comb the gel was cleaned thoroughly to remove any remaining urea that could affect the gel running conditions. The toothed-end of the comb was carefully placed into the gel until it created separate wells. The completed setup was attached to a vertical gel electrophoresis apparatus by bulldog clamps and connected to ~3000V powerpack (James Duncan, South Africa) and pre-run in 0.5xTBE buffer at 1800V for 15-30 minutes before the samples were loaded in each well. The gel electrophoresis proceeded at 1800V for 2-4 hours, depending on the size of the amplified fragments, before being removed and lifted onto 3MM Whatmann paper (Merck) and dried on a gel-dryer (Drygel Slabgel Sr model se1160, Scientific Associates, RSA) for 1 hour. Dried gels were then placed in an autoradiography cassette, overlaid with X-ray (3MM, South Africa) and exposed depending on the radioactive signal efficiency.

Developer solution

Solution A (Polycon)	1.98L
Solution 2 (Polycon)	18.1ml (do not pour directly into container)
make up to 9.9L with dH ₂ O.	

Stop solution (6% acetic acid)

Acetic acid (Merck)	540ml
Make up to 9L with dH ₂ O.	

Fixer solution

PERFIX	1.8L
Hardner-S	225ml
Make up to 9L with dH ₂ O.	

BAC and PAC DNA extraction**Luria Broth**

Bacto tryptone (Difco)	2.5g
Yeast extract (Difco)	1.25g
NaCl (Merck)	2.5g
Make up to 250ml with dH ₂ O	
pH to 7.2	
Autoclave and let cool down to 55°C	
Chloramphenicol (12.5mg/ml stock)	250µl (final concentration ~12.5µl/ml)

LB agar (LA) plates

Bacto tryptone (Difco)	2.5g
Yeast extract (Difco)	1.25g
NaCl (Merck)	2.5g
Make up to 250ml with dH ₂ O	
pH to 7.2	
Bacto agar (Difco)	3.75g
Autoclave and let cool down to 55°C	
Chloramphenicol (12.5mg/ml stock)	250µl (final concentration ~12.5µl/ml)
Pour plates in an extractor hood and allow solidify	
Store in the dark at 4°C	

BAC and PAC mini-prep solutions**Solution I**

Glucose (Merck)	0.9g (50mM)
Tris-HCl (1M stock)	2.5ml (25mM)
EDTA (Biorad)	2ml (10mM)
Make up to 100ml with dH ₂ O	
Autoclave and store at 4°C	

Solution II (made fresh daily)

NaOH (stock)	0.4ml
10% SDS (stock)	1ml
Make up to 10ml with dH ₂ O	

Solution III

Potassium acetate (Merck)	29.4g (5M)
dH ₂ O	±30ml
pH to 4.8 with glacial acetic acid	
Make up to 100ml with dH ₂ O	

YAC DNA extraction**YPD medium**

Yeast extract	10g
Peptone	20g
Make up to 1 litre with dH ₂ O, pH to 5.8	

YPD agar

Yeast extract	10g
Peptone	20g
Make up to 1 litre with dH ₂ O, pH to 5.8	
Bacto agar	20g
40% Glucose (Merck)	50ml

2-Mercaptoethanol

0.5M EDTA	40ml
2-Mercaptoethanol (12M stock)	3.8ml
Add 456.2ml dH ₂ O	

SCE

2M Sorbitol	50ml
1M Sodium citrate	10ml
0.25M EDTA	0.24ml
Make up to 100ml with dH ₂ O	
Filter sterilise	

Polyacrylamide 12% gel (10ml)

dH ₂ O	6ml
1xTBE (stock)	1ml
40% acrylamide/37.5:1 Bis (Promega)	3ml
10% AMPS (Merck)	100µl
TEMED (Promega)	30µl

Mutation analysis**SSCP loading dye**

Formamide (Sigma)	95%
Bromophenol blue (Merck)	0.02%
Xylene cyanol (Biorad)	0.02%
NaOH (Merck)	10mM
EDTA (Biorad)	20mM

Solution B

AgNO ₃ (Merck)	1g
Make up to 1litre with H ₂ O	

Solution C

NaOH (Merck)	15g
NaBHO ₃ (Sigma)	0.1g
Formaldehyde (Merck)	4ml
Make up to 1litre with H ₂ O and constant stirring	

10 % SSCP + 5% glycerol (160ml)

40% Acrylamide/37.5:1 Bis	40.5ml
10xTBE	8ml
Glycerol (BDH)	8ml
H ₂ O	84ml
Urea (Merck)	24g
10% AMPS (Merck)	1600µl
TEMED (Promega)	160µl

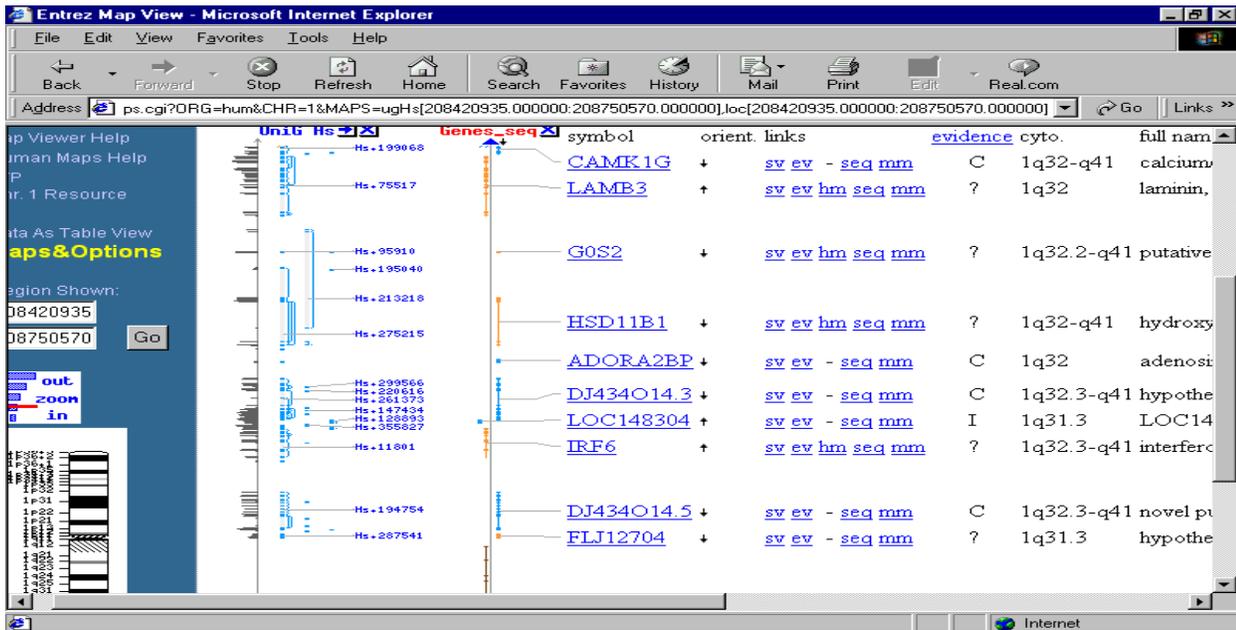
8% SSCP + 5% glycerol (160ml)

40% Acrylamide/37.5:1 Bis	32.4ml
10xTBE	8ml
Glycerol (BDH)	8ml
H ₂ O	91.8
Urea (Merck)	24g
10% AMPS (Merck)	1600µl
TEMED (Promega)	160µl

Methodology

Glass plates with dimensions of 30x40x0.5cm were prepared as before, except that a sheet of Gelbond™ (FMC Bioproducts, Rockland, Maine, USA) was adhered to the non-silanised glass plate using 70% ethanol, with the hydrophobic side facing upward. Two 1mm plastic spacers were placed at either longitudinal end of the plate and overlaid with the other plate. The plates were sealed as before and the appropriate gel solution poured in between. A flat-edged 46 well comb was placed into the gel, which was then allowed to polymerise for 2 hours. The comb was removed and the wells cleaned to remove excess urea. The gel was placed in a gel apparatus and electrophoresis proceeded at a constant power of 70W for 10 minutes at 4°C. Samples were loaded into each well and electrophoresis proceeded at constant power of 25W at 4°C for 16 hours. The gel was removed and silver in solution B for 20min with constant shaking then washed with distilled water and stained with solution C for 20 minutes with shaking. The gel was air-dried for at least 1 hour before analysing.

for the chromosomal interval between *DIS491* and *DIS205* on chromosome 1q32.2.



National Centre for Biotechnology Information (NCBI) database: <http://www.ncbi.nlm.nih.gov>

The NCBI database catalogues of the genomic clones (or genomic contigs), the chromosomal order of genes (centromeric to telomeric) and links to databases containing human/mouse genomic sequence for the specific gene. The chromosomal interval shown is between *DIS491* and *DIS205*.

APPENDIX C

UniGene Cluster Hs.261373 *Homo sapiens*

DJ434O14.3 Hypothetical protein dj434O14.3

SEE ALSO
LocusLink: [30342](#)
HomoloGene: [Hs.261373](#)

SELECTED MODEL ORGANISM PROTEIN SIMILARITIES
organism, protein and percent identity and length of aligned region

<i>H.sapiens:</i>	pir_A43336 - A43336 microtubule-vesicle linker CLIP-170 - human	24 % / 307 aa (see ProtEST)
<i>M.musculus:</i>	pir_JC5420 - JC5420 smooth muscle myosin heavy chain 1 - mouse	21 % / 333 aa (see ProtEST)
<i>R.norvegicus:</i>	pir_A39638 - A39638 plectin - rat	19 % / 509 aa (see ProtEST)
<i>C.elegans:</i>	pir_T33753 - T33753 hypothetical protein M4.1 - <i>Caenorhabditis elegans</i> (fragment)	30 % / 133 aa (see ProtEST)
<i>D.melanogaster:</i>	pir_T13030 - T13030 microtubule binding protein D-CLIP-190 - fruit fly (<i>Drosophila melanogaster</i>)	23 % / 321 aa (see ProtEST)

MAPPING INFORMATION
Chromosome: 1
Genome View: [Chromosome 1](#)

EXPRESSION INFORMATION
cDNA sources: lymph; pooled pancreas and spleen; blood; leukocyte; b-cells; b cells from burkitt lymphoma; t cells from t cell leukemia; aorta; blood; lymphocyte; prostate; tongue; pool; melanocyte+heart+uterus; lung; germ cell; pooled colon; subitary; tonsil; enriched for germinal center

Unigene database: <http://www.ncbi.nlm.nih.gov/Unigene>.

The output shows the alignment percentages and tissue distribution profiles using the entry sequence of *T3JAM* (accession number AL049667).

STACK™

BLASTN 2.2.1 [Jul-12-2001]

Reference:

Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-3402.

Query=

(1826 letters)

Database: /var/blast/heart3_0.fasta

33,206 sequences; 10,152,790 total letters

Searching.....done

	Score	E
Sequences producing significant alignments:	(bits)	Value

I708567 PROJECT:heart ; LENGTH:490bp ; DIRECTION:0 ; CLONE:IMAG... 963 0.0		
Z32860 PROJECT:heart ; LENGTH:362bp ; DIRECTION:0 ; CLONE:HEA03N... 101 5e-21		
A032194 PROJECT:heart ; LENGTH:538bp ; DIRECTION:0 ; CLONE:IMAG... 36 0.22		
cl4135ct4465 cn5825 PROJECT:heart ; COVERAGE:0.99; CRAWID: 1; TOT... 36 0.22		
cl4135ct4463 cn5823 PROJECT:heart ; COVERAGE:1.0; CRAWID: 1; TOTA... 36 0.22		
A249418 PROJECT:heart ; LENGTH:178bp ; DIRECTION:0 ; CLONE: n/a... 34 0.88		
C16976 PROJECT:heart ; LENGTH:204bp ; DIRECTION:0 ; CLONE:GEN-53... 34 0.88		
cl4435ct4775 cn6167 PROJECT:heart ; COVERAGE:0.99; CRAWID: 1; TOT... 34 0.88		
cl3224ct3521 cn4743 PROJECT:heart ; COVERAGE:0.99; CRAWID: 1; TOT... 34 0.88		
R47404 PROJECT:heart ; LENGTH:395bp ; DIRECTION:0 ; CLONE:f291-r... 32 3.5		
N56050 PROJECT:heart ; LENGTH:397bp ; DIRECTION:0 ; CLONE:J6391 ... 32 3.5		
A215854 PROJECT:heart ; LENGTH:245bp ; DIRECTION:0 ; CLONE: n/a... 32 3.5		

>AI708567 PROJECT:heart ; LENGTH:490bp ; DIRECTION:0 ;
 CLONE:IMAGE:2353662 LEN: 490 b.p. FILE gbest24.seq
 3-PRIME DEFN: as98e04.x1 Barstead aorta HPLRB6 Homo
 sapiens cDNA clone IMAGE:2353662 3', mRNA sequence. ;
 CLONELIB:Barstead aorta HPLRB6 ; TISSUE:[]
 Length = 490

Score = 963 bits (486), Expect = 0.0
 Identities = 489/490 (99%)
 Strand = Plus / Minus

Query: 1324 cgtgtggagtgccaagtccctccctaacgaagtggagcctgagggtacagggaggagaa 1383
 |||
 Sbjct: 490 cgtgtggagtgccaagtccctccctaacgaagtggagcctgagggtacagggaggagaa 431

Query: 1384 agactgggatctcagagaccagctgcaaaagaagactttgcagctccaggccaaggaaaa 1443
 |||
 Sbjct: 430 agactgggatctcagagaccagctgcaaaagaagactttgcagctccaggccaaggaaaa 371

STACK database: [http:// http://www.sanbi.ac.za/Dbases.html](http://www.sanbi.ac.za/Dbases.html)

The output shows the alignment percentages and tissue distribution profiles using the entry sequence of *T3JAM* (accession number AL049667).

APPENDIX D

BLASTN 2.2.5 [Nov-16-2002]

Reference:

Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

RID: 1049723957-03455-25496

Query=

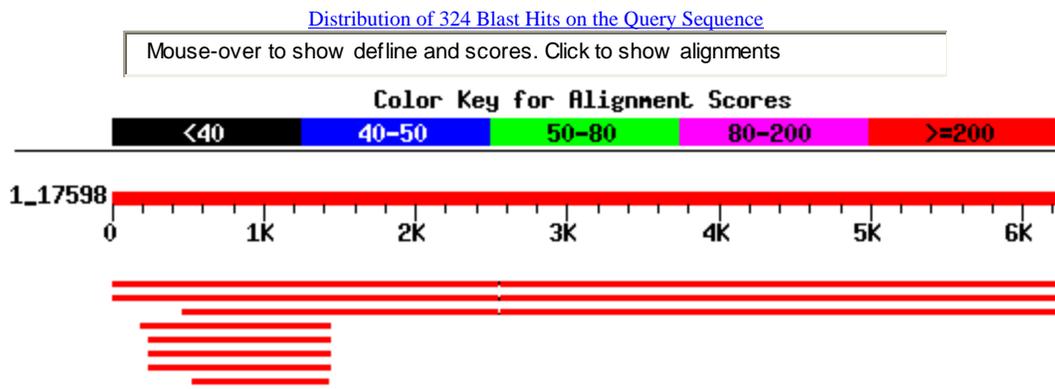
(6253 letters)

Database: All GenBank+EMBL+DDBJ+PDB sequences (but no EST, STS, GSS, or phase 0, 1 or 2 HTGS sequences)

1,726,556 sequences; 8,074,398,388 total letters

If you have any problems or questions with the results of this search please refer to the [BLAST FAQs](#)

[Taxonomy reports](#)



Sequences producing significant alignments:	Score	E
	(bits)	Value
gi 7661995 ref NM_014873.1 Homo sapiens KIAA0205 gene prod...	6901	0.0 LU
gi 1503993 dbj D86960.1 Human mRNA for KIAA0205 gene, comp...	6901	0.0 LU
gi 15668074 gb AC096637.1 Homo sapiens chromosome 1 clone ...	6877	0.0
gi 22902214 gb BC034621.1 Homo sapiens, similar to KIAA020...	2452	0.0
gi 19352306 gb AC084736.5 Homo sapiens chromosome 8, clone...	1400	0.0
gi 10048003 gb AC008066.4 Homo sapiens BAC clone RP11-293D...	1400	0.0

[gij23593471|ref|XM_129651.2|](#) Mus musculus hypothetical prot... [1326](#) 0.0 L
[gij26986566|ref|NM_172266.1|](#) Mus musculus hypothetical prot... [1326](#) 0.0 L
[gij15489110|gb|BC013667.1|](#) Mus musculus, Similar to KIAA020... [1326](#) 0.0 L U
[gij19353716|gb|BC024750.1|](#) Mus musculus, clone IMAGE:536249... [950](#) 0.0 U

Alignments

[>gij7661995|ref|NM_014873.1|](#) L U Homo sapiens KIAA0205 gene product (KIAA0205), mRNA
 Length = 6253

Score = 6901 bits (3481), Expect = 0.0
 Identities = 3619/3688 (98%)
 Strand = Plus / Plus

Query: 2558 agacaggcttactctgtcaccaggctgtagtcagtgccataatctcagctcactgca 2617
 |||
 Sbjct: 2558 agacaggcttactctgtcaccaggctgtagtcagtgccataatctcagctcactgca 2617

Query: 2618 gtctctgcctcccgggttcgagcgattctcctgcctcagcctcccagtagctgggacta 2677
 |||
 Sbjct: 2618 gtctctgcctcccgggttcgagcgattctcctgcctcagcctcccagtagctgggacta 2677

Query: 2678 caggcaccaccatcatgcttgctaatcttttagtagagatggggttcacc 2737
 |||
 Sbjct: 2678 caggcaccaccatcatgcttgctaatcttttagtagagatggggttcacc 2737

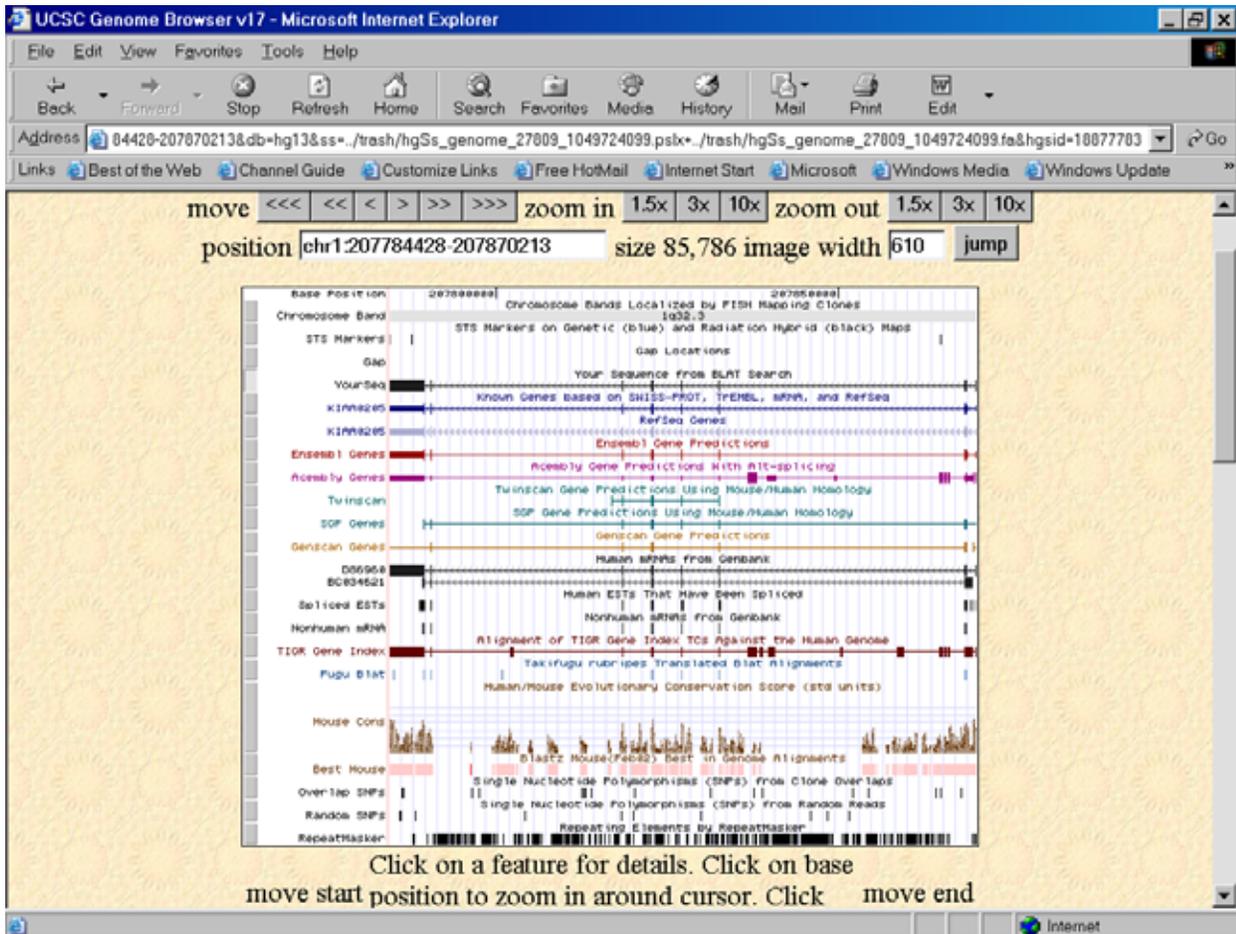
Query: 2738 atgttgccaggctggcctcgaactcctgacctcagtgatctgccacctcggcttct 2797
 |||
 Sbjct: 2738 atgttgccaggctggcctcgaactcctgacctcagtgatctgccacctcggcttct 2797

Query: 2798 aaagtgctgggattgcaggcgtgagccaccacgcccggcctcagttgctttgaagagg 2857
 |||
 Sbjct: 2798 aaagtgctgggattgcaggcgtgagccaccacgcccggcctcagttgctttgaagagg 2857

Query: 2858 gtaaagtcaagttctatttctagaaaacatttttagaaattgttgcatgtttgtac 2917
 |||
 Sbjct: 2858 gtaaagtcaagttctatttctagaaaacatttttagaaattgttgcatgtttgtac 2917

BLAST database: <http://www.ncbi.nlm.nih.gov/Blast>

The output for *KIAA0205* (accession numbers D86960) is shown. The biological/structural similarity scores and alignment sequence is provided. A hyperlinked graphic output of the alignment is provided at the top of the web page. Hyperlinks to other annotation and protein databases are also provided.



BLAT database: <http://genome.ucsc.edu/BLAT>

The output for *KIAA0205* (accession numbers D86960) is shown. The biological/structural similarity is indicated in a graphic format. Similarity scores can be obtained through clicking on the hyperlinks.

Pfam: Results for UserSeq - Microsoft Internet Explorer

Address: <http://www.sanger.ac.uk/cgi-bin/Pfam/nph-search.cgi>

[370 residues]

Potential matches - Domains with Evalues above the cutoff

Domain	Start	End	Bits	Evalue	Alignment	Mode
Acytransferase	80	225	1.00	0.0021	Align	ls

Alignments of Pfam-A domains to HMMs

Format for fetching alignments to seed:

► Alignment of [Acytransferase](#) vs UserSeq/80-225

```

      *->vevrglenlpe.kgpaiivsNHq8ylDplvlslillykrpwrplvf
UserSeq  80  v +g++ + +k+ a+ + NHq D+ l+ +l++ g+ +++
      VMEWGEDIKAVsKDEAVNLVNHQATGDVCTLMMCLQDRGLVVAQMMW 126

      iakkellggisvFlIGwlmrlagsaifidRnnhkdkpllnelylger.yq.
UserSeq 127 + ++ + +G++ g +fi+ + + r++q
      LMDHIFK---YTNFGIUSLVHGDFFIRQGR-----SYRdQQL 160

      .....ri.wkrkaygsalleyavelLrkGelvviFPEGTRs
UserSeq 161 +++ +++ ++++ +w +v+PPEG+
      lllkkhlennyrsrDRkW-----IVLPDEGGFL 188

```

Pfam database: <http://www.sanger.ac.uk/Software/Pfam>

The output for *KIAA0205* (accession numbers D86960) is shown. The biological/structural similarity score and alignment is provided as text. Hyperlinks to other annotation and protein databases are also provided.

Block Searcher Results - Microsoft Internet Explorer

Address http://www.blocks.fhrc.org/blocks-bin/blocks_search

Cutoff combined expected value for hits= 1
Cutoff block expected value for repeats/other= 1

Family	Strand	Blocks	Combined E-value
IP8002123 Phospholipid and glycerol acyltrans	1	2 of 2	0.0095
IP8000680 Borrelia lipoprotein	1	1 of 6	0.45
P800979 Tafazzin signature	1	1 of 6	0.82
P800489 Frizzled protein signature	1	1 of 7	0.98

>[IP8002123](#) 2/2 blocks Combined E-value= 0.0095: Phospholipid and glycerol acyltransferase (from 'motifs_6.msf')

Block	Frame	Location (aa)	Block E-value
IP8002123A	0	91-106	1.6
IP8002123B	0	179-187	4.1

Other reported alignments:

```

|--- 195 amino acids---|
IP8002123 AAA:!!.....:BB
Unknown AAA:.....:BB

IP8002123A <->A (1,649):90
Y205_HUMAN|Q92604 91 SKDEAVMLVNHQATGD
|||||
Unknown 91 SKDEAVMLVNHQATGD

IP8002123B A<->B (16,299):172
Y205_HUMAN|Q92604 179 IVLFPEGGF
|||||
Unknown 179 IVLFPEGGF

```

>[IP8000680](#) 1/6 blocks Combined E-value= 0.45: Borrelia lipoprotein

Block	Frame	Location (aa)	Block E-value
IP8000680C	0	231-262	0.45

Other reported alignments:

Blocks database: http://www.blocks.fhrc.org/blocks_search.html

The output for *KIAA0205* (accession numbers D86960) is shown. The biological/structural similarity score and alignment is provided as text. Hyperlinks to other annotation and protein databases are also provided.

ScanProsite

Search a sequence against PROSITE

Sequence:

```

MAITLEEAFW LGWLLVKALM RFAPFVNNL VAIPSYICTV IILQPLRVLD SKRFWYIEGI
NYKLLGKVA SNGHYAGYTV NEQGEDIKAV SKDAVNLVN HQATGDVCTL HMCLODKGLV
VAQHEWLMHD IFKYNFGIV SLVHGDFIR QGRSYRQQQL LLLKHLNN YRSRDKWIV
LPPEGGFLRK RHETSQAFK KNNLPFLTNV TLPRSGATKI ILNALVAQQK NGSPAGGDAK
ELDSKSGLQ WIIDTTIAYP KAEPIDIGTW ILGYRKPVTV HVHYRIFPIK DVPLETDDL
TULYQRVFEK EDLLSNFYET GAFFPSKGGK EAVSRENTLS NLWIFLIQSF AFLSGVMWYN
IIQYFYKCLF

```

PROSITE Release 17.24, of 27-Oct-2002

>[PDOC00001](#) [PS00001](#) [ASN_GLYCOSYLATION](#) N-glycosylation site [pattern] [Warning: pattern with a high probability of occurrence].

209 - 212 NVTL

>[PDOC00004](#) [PS00004](#) [CAMP_PHOSPHO_SITE](#) cAMP- and cGMP-dependent protein kinase phosphorylation site [pattern] [Warning: pattern with a high probability of occurrence].

191 - 194 RReT
275 - 278 RRpT

>[PDOC00005](#) [PS00005](#) [PKC_PHOSPHO_SITE](#) Protein kinase C phosphorylation site [pattern] [Warning: pattern with a high probability of occurrence].

51 - 53 SKR
154 - 156 SyR

Prosite: <http://www.expasy.org/prosite>

The output for *KIAA0205* (accession numbers D86960) is shown. The biological/structural similarity alignment is provided as text. Hyperlinks to other annotation and protein databases are also provided.

APPENDIX F

Oligonucleotide primers, annealing temperatures and amplification product sizes of candidate genes screened in the present study

<i>GENE</i> Exon	Sense oligonucleotide 5' → 3'	Antisense oligonucleotide 5' → 3'	Annealing temperature (°C)	Product size (bp)
<i>KCNHI</i>				
1	CGAGGCCAGTTTCCTGCTG	GCGGCGAGAGGTTCTGCAAT	54	229
2	GAATCTCAACTAATGTGTTGCCTCATG	CCATAATAGATTGTTCAACCAGAACAACCTC	55	250
3	CATATGCTTGTAGAATAATGAATACTTCT	TTCAAACACATAAAAGGTATGGTTCAGAATG	51	212
4	CAAATGAGAATGCAAATTACTGACCAAC	AAGGCATAAATGTATTTGTACAAGTCAG	51	214
5	GTGCACCTTTGGGATGTGGTT	CTAAAAGTGAGGCTCAAGATGAGC	51	214
6a	GACTGGTTAGATACTTGCAATTAATCC	CAATGTCCACCAAAAAGATAACATCC	53	243
6b	CAGAATAATGTGGCCTGGCTGGTT	CAGCAAAACCAATCTTCCCTGG	53	262
7a	CCAGCTAACAGAAGAGATGGCC	GACACCAAGACAATCCGCAACAACA	53	287
7b	GCCAATGTCCATCGCTAGTTGG	AGTAGTATCTTTGGTTCTTTGGCCCT	53	298
8	CAAAGACAAATGGTTTCCCTGAGCTC	GGCTCAGGGACCACTAACTTG	53	297
9	GCTAAGCCAACCCAGATACT	CTGAAGGATGTGGGCTGAAGGC	53	358
10	AGTGTACAGAGACTGTTCTTTGTGGA	CCTCTCCTGGCCAACACTTTC	52	292
11a	GTCATGTGGACATATGTGGTAGG	TACCCCATCCCTGAGCAGAC	52	311
11b	CTTCAGCTCGTGCCTCACCT	GAGGACTGGAACAAGGTGTCC	55	261
11c	CGCTTTTGTCTCTCGGGAAGT	CATGCCTCCGCCAACCACAG	55	282
11d	TCTCACGCACGGTGACCAC	ATGTGGCTAACACGCCATGTCATT	52	309
<i>PPP2R5A</i>				
1	GGAGCTGCCAAGCGTCAG	CTGCGCTCGCTGCTGGGG	51	246
2	TTCTGAGTAGGTTATTTCTAAGTTTT	CATAAATCTGAGGAACATAATTCTAG	47	276
3	AAATATAAAAACTGCCAATACGTAATTATG	TCAATTACACATACAACAACGAACGT	47	187
4	GAATATATGAGAATCTAAGACAAAAGAG	CATCATCTTAGCCTTTATTTGACTTTC	47	164
5	GGAAAAGCAATATACAATAAATGA	ATTTTTTAAAAGTAATTTCTCAGGTTTC	47	202
6	TACTTAATTCTTAAAAAGAACAAAAAACC	GTTTGATTTTGATTACATCTTGTCTTTA	47	138

7	ATAATATTCATTTTCATCTGAAATGTTTCTG	TACTTAACCCCTTCCCCTGCC	47	182
8	GCTATTAATGCAGACCAAAAACATGAA	CTCACTAATATTTATAAACTGTCCCTTA	47	135
9	CAGCTATGGAAAATACGCTGGGATTAAC	CAGCCCAAAGAAGCATGGTTTATAGC	51	203
10	CATTTGTATGGTTTCTATATTAGTTTCAC	CTAGTGTAAGTAAAGGCTTCACTG	51	328
11	CTTTCCTATTTTAGTTTTGACATTTCTCTA	CTGTAATAGGTTTCAAGTATGAGTTAG	51	272
T3JAM				
1	GACGGTACCAAGAAGGGACG	GTTGGAGTTCCCAAACCTGGAC	55	209
2	GCCTGGCTCCCTGGAAACTT	CTCCAATACTGTTGGGGCTGAG	55	208
3a	CCCCTTCTCCTCCCTCTTGG	GTGGTCACATTGGGACGGCA	55	230
3b	AGCGCAGGCAAGAGATCCGTG	CTGGCACAAGACAAGGGTTCC	55	250
4	CTCCAGGCAGTGGTTTTCAGC	CAGGCCCTGTGGAGTGAG	55	223
5	CCTCACTTCTCCTGTAAGTTC	GAGGAATGGATGGGATGTGAGAG	55	151
6	CTGCCTCCTGACTTGATTCTTTGG	CCAGAATGTCAGAGCTTAATACTTTCTG	55	145
7	CAGGGTTTCAGGGGACTTACAGA	GTAAGATTAACCTGGAGGATCGG	55	155
8	CTTACCCTCTTGCACTTTGTGCC	GAACACATAGAAGCCCAAGCTTTGT	53	179
9	CTCTTTGCAGACTAGATTTTGCCCC	TCATTTAATTCTAGATTCCTGAGATGGAAG	53	156
10	GCTCAGGATTTGGCACACAGAAG	AGCTGCCTGGAGTCCCCAG	55	247
11	GAGCCAGCTGCACAACCTCC	CATACCCTGTCTTCCCTCCC	53	227
12	CACTGACCTCCTTCTATATTGGCC	GCTGCCATGCACTCAAGGC	53	213
13	GTACTGAGTTCATTAAGTCTTGGGTTC	GCCAGGACCCCAGGAAAAC	53	177
14	GTAGTAAATATGCTAGCATAGACAAGTTCC	GCCTCTAGGCAGGTAAGAGAAG	55	166
15	CTCACTGTCACCAAGAATCTGGC	GCTATTTGCAAATGAGCCCTCAC	55	214
16	TGTCCTAGACAGAAGTGACAGTGA	GAAGGAATCGTCCCATCATCCAG	53	214
17	CCACTTTCTTCTGTCTCCCTGTC	GTTAATGTAATACACCATTGAGGAATAAAC	53	336
VWS59				
Transcript 1				
1	CATGTTTCATGGCTTAAAGTTTTG	AGGAGGTCTGAAAAGCTTGG	57	199
2	AGCCAGCCCAGTTTTGTTTA	CTGGATCGACTCCAAATACAA	57	209
3	CCCTCTGATAAACTATTTTGTCTTATT	TACCTTTCTTGCCAGCAGTG	57	147
4	ACCATTTGCAGCACAGACC	CCCAATGGTAACAGGAAAACA	57	232
5	GGCCACTCAAAGGTTGTTCT	ATCCCCAGCGGGCTCTAC	57	272
6	TTTAGATGCAGCAGCAGGAG	ATCCCACTGTTGGTGATTCC	57	236

7	GACTGGTCTGTTTGCCTTCA	TGCCTGACCAACTTTAGCAA	57	207
8	ATCTTTTATTGATATGCTTGTCTATGTCTC	AACCATAGTGTTAGCAATGTTTTCTTGACA	52	220
VWS59				
<i>Transcript 2</i>				
1	TTCAGCTCTTCAGCCAGCCC	GCTGAATCTGATACCTGGATCGAC	55	234
2	GTGTTCTGACTCAGACGTGTACTG	GCTTGGACTCTTTCTTCACAAGAAGC	55	242
LAMB3				
1	CACAGGCAGGTGGGCATTGT	CGGCCCGCACCAGGCAGG	50	314
1'	ACGGCTGCCAGATTCCTGAG	GCCACTGAGGTGCTTATCAAATGTG	55	352
2	CTCTAGGGGTTTGCCTGTTGTACC	CTGCCATATAACAGGGCCTGGA	55	156
3	GGACTCCTCCATCACACATCCC	GTGGACAAATGGCAGCTCACAC	55	222
4	CAGGACCCACCTCTGAGTTTCT	GAGGCTGTACAGTTTGAGTTCCCGT	55	200
5	AGTGGACAGTCTCTGGATGACAGTT	CCCCTCTCCTATCCACTGCCAC	55	175
6	CAGTGCCTCCCCTCTGCACA	CCCCAGGGCTACTTTTAGGATTC	55	264
7	CGGGGCACACTATTTGTCTCTTC	CCCCACCCATAGCTCCATGGAC	55	161
8	CTCCA ACTCTTCCCCTCCTTGCTG	TCCCCTTCCCACCCTAGGCT	55	256
9	GGTGCCCTTCTTTGTTGATTACTCC	AGGGGCCAGCACTCGACC	55	225
10	GGTGTGGGTGAGATGTCAGCC	GGAGATGGGGAGTAACAGACAAATG	55	278
11	GA ACTCTACTTTTCTTGTCTTA	CACCCACNTCATNCTGG	52	239
12	GTGTGTCCTCTGGGGGTGAGTT	CACTGAGGGGGCCCCCG	52	296
13	TGTGAGGCAGGGTCCAC	CCCCAGAGCCTCAGCGT	55	190
14	ACCCCTGTCTCTGACTGTC	CACAGGCTGGCGGTGGCA	55	268
14'	TATGATGCGGACCTCCGGGA	TCCAGCATAATTCTTGCAAATG	55	260
15	CCTCCTATATTTTAGTGTAAGC	CCACGGCAGTCCCATGC	55	246
16	CCTGCTCCTCCTCCATCTCCATT	CAAGCTGGGTGTGGGCTCCA	55	314
17	GTTATTGAGGGGCTTGGGTAGCCT	TCCTCCTACCTGTGCCCCAC	55	286
18	CCTTACCCCTTCTGACCTAGAGT	GCAGGACTGGAACAGCAGCC	55	223
19	GGTTAGTGGGCTCCTACAGTCAA	GGGCAGAGAGAAGTTCAGGGCC	52	283
20	TCACTGTTCACTCATTCTCTCTCC	GCCCAAATTTTCTGTTTATGCCC	52	249
21	GATCCACTATAAAGGCCAGATCTC	CACTGGCCGTACATCATTGAGCTAA	50	261
22	GCATGATGACAAATGACTGAGAC	AGCCTTCATGAATGCCCAATAGTCC	50	243
23	GAGTCTTGGGGAGTCTCACCATC	GATGAGTGGGGCAACGGGC	50	228

RAMP				
1	GTGTTGTGAGAGGCGCAAGC	GAGGTGGCGTGTGTTTCGGG	55	210
2	GACAGAGCAAGACTCCGTCTGA	ATAATAATACTTGAGGATGTGTAAAGTT	50	245
3	ACAGCTTCCAGGCATTTAAAGTTATATTA	TCCCTTATCAGTATTTACCCAGATC	50	200
4	GGTAAATACTGATAAGGGAGTGCTTTAT	CTGCAAAGAAGGTTAAGGAGATGAAA	50	203
5	ATGTGTCATTGACTGGTTTTGTAAACCT	ATTAAGTACAGGGTAATCCCAAATCTC	50	177
6	<i>GTTATATTCAGCATGTAAGAATCCTGAAAC</i>	CTATCCTACAGATAAGATCTGCACAC	50	219
7	AGATTATGATGCCCAAGATAGAATCTTC	CTCATAAAAAAGATCTCGTCGGAAG	50	168
8	CCACTAGAAATAACAAGAGCCAAG	CTTCAGTCTAACACTTGTGGGGTG	50	201
9	GGGAAATCTAATATTGCTTTGCTGAC	ACCTGTGATCTACATTTTAAACCAACT	50	186
10	GATTGAAATTTCTTAATCCCTTCTTGG	CTAGTCAAACCATAAAAAACTTGGAAG	50	221
11a	CTTCTGATCTTCATGATCCTCTCC	CCAGACAGCTTGAGTCTAGCC	50	451
11b	AGGCTTGCTCTGAGTCTAGAAATAG	GAAATGTATCTTCCCACCACTGC	50	503
12	CAGTTGAACTTCAAGGAAATCTCTG	GACCAAAGCTCAGTAACTCACT	50	180
RAMP				
Transcript 2				
1	TGCTGAAATGTGATGCTGTATTGGAATTC	GAAATAGATCATTCTGTCCAAGCCTTG	55	227
KIAA0205				
1	CACCCCAGAAAGACAGGACG	CTTATGCATTAAAGAAAAGTGTTC	50	307
2	CTTTAAAATGACAGCTTTGAACTTTTTTG	GATGTCTATCTACTGCCTTCAG	50	190
3	TCTTCTATCCAATATTATTTCTTGCTTTA	CTGCGGGGAAAAAGTGCTCA	50	164
4a	GGATATACCTGTACTTATTACGTG	AGTAACATTTGTAAGAAATGGCAAG	50	220
4b	GAGAAACAAGTCAGGCATTTGC	CCAACAAGGTAAAAAATGCTAAAAAC	50	187
5	TTGCTGCAGATAATATAAAAGTTGATTTAA	GTTGGATATATACTGAGTACTAATATAG	50	222
6	CTAAATAGTTTTTGTCTTTTCTTCTAACC	GCTATAATTTAGCAACTATTTAAAGGTT	50	182
7	ACAGTAAACAAACTAATATTTCTATCTTG	GACCTTGACAAGTCCACG	50	215

bp = base pairs; *KIAA0205* = putative acyltransferase; *KCNH1* = ether-a-go-go related potassium channel; *LAMB3* = laminin b3; *PPP2R5A* = protein phosphatase 2A b56alpha; *RAMP* = retinoic acid regulated nuclear matrix associated protein; *T3JAM* = adapter protein that interacts with a cytoskeletal protein; *VWS59* = novel gene

APPENDIX G

Genes annotated to within the PFHBII locus between *DIS70* and *DIS505* (obtained from the Ensembl database, November 2002)

Ensembl Gene ID and gene description

***ENSG00000008141.1**,LAMININ BETA-3 CHAIN PRECURSOR (LAMININ 5 BETA 3) (LAMININ B1K CHAIN) (KALININ B1 CHAIN). [Source:SWISSPROT;Acc:Q13751]
ENSG00000123689.1,PUTATIVE LYMPHOCYTE G0/G1 SWITCH PROTEIN 2. [Source:SWISSPROT;Acc:P27469]
ENSG00000117594.1,CORTICOSTEROID 11-BETA-DEHYDROGENASE, ISOZYME 1 (EC 1.1.1.146) (11-DH) (11-BETA-HYDROXYSTEROID DEHYDROGENASE 1) (11-BETA-HSD1). [Source:SWISSPROT;Acc:P28845]
***ENSG00000009790.1**,B 3 1 DJ434O14
ENSG00000162757.1,AMBIGUOUS
ENSG00000117595.1,INTERFERON REGULATORY FACTOR 6 (IRF-6). [Source:SWISSPROT;Acc:O14896]
HYPOTHETICAL 84 KD PROTEIN FROM SGA1-KTR7. [Source:RefSeq;Acc:NM_014388]
ENSG00000117597.1,NOVEL PUTATIVE PROTEIN SIMILAR TO YIL091C YEAST
ENSG00000143464.1,AMBIGUOUS
ENSG00000173964.1,LINE 1 REVERSE TRANSCRIPTASE HOMOLOG
ENSG00000143469.1,SYNAPTOTAGMIN
ENSG00000082497.1,UNKNOWN
***ENSG00000054392.1**,SIMILAR TO MELANOMA ANTIGEN RECOGNIZED BY T CELLS 2
ENSG00000152015.1,IG CHAIN V REGION
ENSG00000177394.1,TRANSFORMATION-RELATED PROTEIN (FRAGMENT). [Source:SPTREMBL;Acc:Q15662]
***ENSG00000143473.1**,POTASSIUM VOLTAGE-GATED CHANNEL SUBFAMILY H MEMBER 1 (ETHER-A-GO-GO POTASSIUM CHANNEL 1) (HEAG1) (H-EAG). [Source:SWISSPROT;Acc:O95259]
ENSG00000174660.1,40S RIBOSOMAL S25
ENSG00000174658.1,PX19
ENSG00000117625.1,AMBIGUOUS
ENSG00000082512.1,TNF RECEPTOR-ASSOCIATED FACTOR 5. [Source:RefSeq;Acc:NM_004619]
ENSG00000153363.1,UNKNOWN
ENSG00000179809.1,ARP2/3 COMPLEX 21 KDA SUBUNIT P21 ARC
ENSG00000174649.1,UNKNOWN

ENSG00000170385.1,ZINC TRANSPORTER 1 (ZNT-1).

[Source:SWISSPROT;Acc:Q9Y6M5]

ENSG00000117650.1,SERINE/THREONINE-PROTEIN KINASE NEK2 (EC 2.7.1.37) (NIMA-RELATED PROTEIN KINASE 2) (NIMA-LIKE PROTEIN KINASE 1) (HSPK 21).

[Source:SWISSPROT;Acc:P51955]

***ENSG00000123684.1**,AMBIGUOUS

***ENSG00000143476.1**,L2DTL PROTEIN. [Source:RefSeq;Acc:NM_016448]

ENSG00000162767.1,60S RIBOSOMAL PROTEIN L21. [Source:SWISSPROT;Acc:P46778]

***ENSG00000066027.1**,SERINE/THREONINE PROTEIN PHOSPHATASE 2A, 56 KDA REGULATORY SUBUNIT, ALPHA ISOFORM (PP2A, B SUBUNIT, B' ALPHA ISOFORM) (PP2A, B SUBUNIT, B56 ALPHA ISOFORM) (PP2A, B SUBUNIT, PR61 ALPHA ISOFORM) (PP2A, B SUBUNIT, R5 ALPHA ISOFORM). [Source:SWISSPROT;Acc:Q15172]

ENSG00000065600.1,AMBIGUOUS

ENSG00000117691.1,SECRETED PROTEIN OF UNKNOWN FUNCTION.

[Source:RefSeq;Acc:NM_013349]

The annotated genes (known and predicted) are indicated in bold to allow easier identification

** = genes that were screened for the PFHBII-causative mutation in the present*

REFERENCES

Abbott GW, Sesti F, Splawski I, Buck ME, Lehmann MH, Timothy KW, Keating MT, Goldstein SA. (1999). MiRP1 forms IKr potassium channels with HERG and is associated with cardiac arrhythmia. *Cell* **97**:175-187.

Aitken JK. (1932). Congenital heart block. *Lancet* **223**:1375-1379.

Alboni P, Scarò S, Fucà G. (2001). Development of heart failure in bradycardic sick sinus syndrome. *Ital Heart J* **2**:9-12.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**:3389-3402.

Anderson RH and Ho SY. (1998). The architecture of the sinus node, the atrioventricular conduction axis, and the internodal atrial myocardium. *J Cardiovasc Electrophysiol* **9**:1233-1248.

Aniansson A, Grimby G, Hedberg M. (1992). Compensatory muscle fiber hypertrophy in elderly men. *J Appl Physiol* **73**:812-816.

Arad M, Woodrow Benson D, Perez-Atayde AR, McKenna WJ, Sparks EA, Kanter RJ, McGarry K, Seidman JG, Seidman CE. (2002). Constitutively active AMP kinase mutations cause glycogen storage disease mimicking hypertrophic cardiomyopathy. *J Clin Invest* **109**:357-362.

Armour JAL, Barton DE, Cockburn DJ, Taylor GR. (2002). The detection of large deletions or duplications in genomic DNA. *Hum Mutation* **20**:325-337.

Ashrafian H, Redwood C, Blair E, Watkins H. (2003). Hypertrophic cardiomyopathy: a paradigm for myocardial energy depletion. *Trends Genet* **19**:263-268.

Attwood T. (2000). Genomics. The babel of bioinformatics. *Science* **290**:471-473.

Attwood TK and Parry-Smith DJ. (1999). Introduction. In: *Introduction to bioinformatics*, pp 1-18. Wood E, ed. London: Prentice Hall.

Aylward RD. (1928). Congenital heart block. *Br Med J* **1**:943.

Barak M, Herschkowitz S, Shapiro I, Roguin N. (1987). Familial combined sinus node and atrioventricular conduction dysfunctions. *Int J Cardiol* **15**:231-239.

Barth PG, Scholte HR, Berden JA, Van der Klei-Van Moorsel JM, Luyt-Houwen IE, Van 't Veer-Korthof ET, Van der Harten JJ, Sobotka-Plojhar MA. (1983). An X-linked mitochondrial disease affecting cardiac muscle, skeletal muscle and neutrophil leucocytes. *J Neurol Sci* **62**:327-355.

Berko BA and Swift M. (1987). X-linked dilated cardiomyopathy. *New Engl J Med* **316**:1186-1191.

Bezzina C, Veldkamp MW, van Den Berg MP, Postma AV, Rook MB, Viersma JW, van Langen IM, Tan-Sindhunata G, Bink-Boelkens MT, van Der Hout AH, Mannens MM, Wilde AA. (1999). A single Na⁺ channel mutation causing both long-QT and Brugada syndromes. *Circulation Res* **85**:1206-1213.

Bies RD, Friedman D, Roberts R, Perryman MB, Caskey CT. (1992). Expression and localisation of dystrophin in human cardiac Purkinje fibres. *Circulation* **86**:147-153.

Bione S, D'Adamo P, Maestrini E, Gedeon AK, Bolhuis PA, Toniolo D. (1996). A novel X-linked gene, *G4.5* is responsible for Barth syndrome. *Nature Genet* **12**:385-389.

Bissler JJ, Tsoras M, Goring HH, Hug P, Chuck G, Tombragel E, McGraw C, Schlotman J, Ralston MA, Hug G. (2002). Infantile dilated X-linked cardiomyopathy, *G4.5* mutations, altered lipids, and ultrastructural malformations of mitochondria in heart, liver, and skeletal muscle. *Lab Invest* **82**:335-344.

Blair E, Redwood C, Ashrafian H, Oliveira M, Broxholme J, Kerr B, Salmon A, Ostman-Smith I, Watkins H. (2001). Mutations in the gamma (2) subunit of AMP-activated protein kinase cause familial hypertrophic cardiomyopathy: evidence for the central role of energy compromise in disease pathogenesis. *Hum Mol Genet* **10**:1215-1220.

Bolhuis PA, Hensels GW, Hulsebos TJ, Baas F, Barth PG. (1991). Mapping of the locus for X-linked cardioskeletal myopathy with neutropenia and abnormal mitochondria (Barth syndrome) to Xq28. *Am J Hum Genet* **48**:481-485.

Bomont P, Cavalier L, Blondeau F, Ben Hamida C, Belal S, Tazir M, Demir E, Topaloglu H, Korinthenberg R, Tuysuz B, Landrieu P, Hentati F, Koenig M. (2000). The gene encoding gigaxonin, a new member of the cytoskeletal BTB/kelch repeat family, is mutated in giant axonal neuropathy. *Nature Genet* **26**:370-374.

Bork P. (2000). Powers and pitfalls in sequence analysis: the 70% hurdle. *Genome Res* **10**:398-400.

Bortoluzzi S, d'Alessi F, Danieli GA. (2000). A computational reconstruction of the adult human heart transcriptional profile. *J Mol Cell Cardiol* **32**:1931-1938.

Bowles KR, Gajarski R, Porter P, Goytia V, Bachinski L, Roberts R, Pignatelli R, Towbin JA. (1996). Gene mapping of familial autosomal dominant dilated cardiomyopathy to chromosome 10q21-23. *J Clin Invest* **98**:1355-1360.

Braunwald E. (1980). Pathophysiology of heart failure. In: *Heart disease: A textbook of cardiovascular medicine vol. 1*. pp 453-483. Braunwald E, ed. Philadelphia: WB Saunders.

Braverman N, Chen L, Lin P, Obie C, Steel G, Douglas P, Chakraborty PK, Clarke JT, Boneh A, Moser A, Moser H, Valle D. (2002). Mutation analysis of *PEX7* in 60 probands with rhizomelic chondrodysplasia punctata and functional correlations of genotype with phenotype. *Hum Mutation* **20**:284-297.

Brewis N, Ohst K, Fields K, Rapacciuolo A, Chou D, Bloor C, Dillmann W, Rockman H, Walter G. (2000). Dilated cardiomyopathy in transgenic mice expressing a mutant A subunit of protein phosphatase 2A. *Am J Physiol Heart Circ Physiol* **279**:H1307-H1318.

Brink AJ and Torrington. M. (1977). Progressive familial heart block: Two types. *S Afr Med J* **52**:53-59.

Brink AJ, Torrington M, van der Walt JJ. (1976). Hereditary dysrhythmic congestive cardiomyopathy. *S Afr Med J* **50**:2119-2123.

Brink PA, Ferreira A, Moolman JC, Weymar HW, Van der Merwe P-L, Corfield VA. (1995). Gene for progressive familial heart block type I maps to chromosome 19q13. *Circulation* **91**:1633-1640.

Brugada P and Brugada J. (1992). Right bundle branch block, persistent ST segment elevation and sudden cardiac death: a distinct clinical and electrocardiographic syndrome. A multicentre report. *J Am Coll Cardiol* **20**:1391-1396.

Burge C and Karlin S. (1997). Prediction of complete gene structures in human genomic DNA. *J Mol Biol* **268**:78-94.

Carrier L, Hengstenberg C, Beckmann JS, Guicheney P, Dufour C, Bercovici J, Dausse E, Berebbertrand I, Wisnewsky C, Pulvenis D, Fetler L, Vignal A, Weissenbach J, Hillaire D, Feingold J, Bouhour J-B, Hagege A, Desnos M, Isnard R, Dubourg O, Komajda M, Schwartz K. (1993). Mapping of a novel gene for familial hypertrophic cardiomyopathy to chromosome 11. *Nature Genet* **4**:311-313.

Casali C, Santorelli FM, D'Amati G, Bernucci P, DeBiase L, DiMauro S. (1995). A novel mtDNA point mutation in maternally inherited cardiomyopathy. *Biochem Biophys Res Commun* **213**:588-593.

Chen Q, Kirsch GE, Zhang D, Brugada R, Brugada J, Brugada P, Potenza D, Moya A, Borggrefe M, Breithardt G, Ortiz-Lopez R, Wang Z, Antzelevitch C, O'Brien RE, Schulze-Bahr E, Keating MT, Towbin JA, Wang Q. (1998). Genetic basis and molecular mechanism for idiopathic ventricular fibrillation. *Nature* **392**:293-296.

Cheung WM, Chu AH, Chu PW, Ip NY. (2001). Cloning and expression of a novel nuclear matrix-

associated protein that is regulated during the retinoic acid-induced neuronal differentiation. *J Biol Chem* **276**:17083-17091.

Chien K. (2000). Genomic circuits and the integrative biology of cardiac diseases. *Nature* **407**:227-232.

Chieurel M. (2002). Bioinformatics: bringing it all together. *Nature* **419**:751-757.

Collins JE, Goward ME, Cole CG, Smink LJ, Huckle EJ, Knowles S, Bye JM, Beare DM, Dunham I. (2003). Reevaluating human gene annotation: a second-generation analysis of chromosome 22. *Genome Res* **13**:27-36.

Combrink JM, Davis WH, Snyman HW. (1962). Familial bundle branch block. *Am Heart J* **64**:397-400.

Conner AC, McFadden JF, Houston BJ, Finn JL. (1959). Familial congenital complete heart block. *Am J Obstet Gynecol* **78**:75.

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.

Corfield VA, Moolman JC, Martell R, Brink PA. (1993). Polymerase chain reaction-based detection of MN blood-group specific sequences in the human genome. *Transfusion* **33**:119-124.

Curran ME, Splawski I, Timothy KW, Vincent GM, Green ED, Keating MT. (1995). A molecular basis for cardiac arrhythmia: *HERG* mutations cause long QT syndrome. *Cell* **80**:795-803.

Dadgostar H, Doyle SE, Shahangian A, Garcia DE, Cheng G. (2003). T3JAM, a novel protein that specifically interacts with TRAF3 and promotes the activation of JNK. *FEBS Letters* **553**:403-407.

Daehmlow S, Erdmann J, Knueppel T, Gille C, Froemmel C, Hummel M, Hetzer R, Regitz-Zagrosek V. (2002). Novel mutations in sarcomeric protein genes in dilated cardiomyopathy. *Biochem Biophys Res Commun* **298**:116-120.

Davies MJ. (1967). A histological study of the conduction system in complete heart block. *J Pathol Bacteriol* **94**:351-358.

Davletov BA and Sudhof TC. (1993). A single C2 domain from synaptotagmin I is sufficient for high affinity Ca^{2+} /phospholipid binding. *J Biol Chem* **268**:26386-26390.

Dayhoff MO, Schwartz RM, Orcutt BC. (1978). In: *Atlas of protein sequence and structure vol. 5*. pp 435. Dayhoff MO, ed. Washington: NBRF.

de Lange WJ. (2004). An investigation of myosin binding protein C in South Africa and a search for ligands binding to myosin binding protein C. PhD thesis, University of Stellenbosch.

de Meeus A, Stephan E, Debrus S, Jean MK, Loiselet J, Weissenbach J, Demaille J, Bouvagnet P. (1995). An isolated cardiac conduction disease maps to chromosome 19q. *Circulation Res* **77**:735-740.

Dorn GW 2nd. (2002). Adrenergic pathways and left ventricular remodeling. *J Card Fail* **8**:S370-S373.

Dos Remedios CG, Liew CC, Allen PD, Winslow RL, Van Eyk JE, Dunn MJ. (2003). Genomics, proteomics and bioinformatics of human heart failure. *J Muscle Res Cell Motil* **24**:251-260.

Drachman DA. (1975). Ophthalmoplegia plus: a classification of the disorders associated with progressive external ophthalmoplegia. In: *Handbook of Clinical Neurology*. pp 203-216. Elsevier Science Ltd.

Dracup K, Baker DW, Dunbar SB, Dacey RA, Brooks NH, Johnson JC, Oken C, Massie BM. (1994). Management of heart failure. II. Counseling, education, and lifestyle modifications. *J Am Med Assoc* **272**:1442-1446.

Draper N, Walker EA, Bujalska IJ, Tomlinson JW, Chalder SM, Arlt W, Lavery GG, Bedendo O, Ray DW, Lainge I, Malunowicz E, White PC, Hewison M, Mason PJ, Connell JM, Shackleton CH, Stewart PM. (2003). Mutations in the genes encoding 11 beta-hydroxysteroid dehydrogenase type 1 and hexose-6-phosphate dehydrogenase interact to cause cortisone reductase deficiency. *Nature Genet* **34**:434-439.

Durand JB, Abchee AB, Roberts R. (1995a). Molecular and clinical aspects of inherited cardiomyopathies. *Ann Med* **27**:311-317.

Durand JB, Bachinski LL, Bieling LC, Czernusewicz GC, Abchee AB, Yu QT, Tapscott T, Hill R, Ifegwu J, Marian AJ, Brugada R, Daiger S, Gregoritch JM, Anderson JL, Quiñones M, Towbin J,

Roberts R. (1995b). Localisation of a gene responsible for familial dilated cardiomyopathy to chromosome 1q32. *Circulation* **92**:3387-3389.

Duyk GM, Kim S, Myers RM, Cox DR. (1990). Exon trapping: a genetic screen to identify candidate transcribed sequences in cloned mammalian genomic DNA. *Proc Natl Acad Sci USA* **87**:8995-8999.

Emanuel R. (1972). Familial cardiomyopathies. *Postgrad Med J* **48**:742-745.

Erdmann J, Hassfeld S, Kallisch H, Fleck E, Regitz-Zagrosek V. (2000). Genetic variants in the promoter (983G>T) and coding region (A92T) of the human cardiotrophin-1 gene (*CTF1*) in patients with dilated cardiomyopathy. *Hum Mutation* **16**:448.

Eriksson H. (1995). Heart failure: a growing health problem. *J Intern Med* **237**:135-141.

Escayg A, Heils A, MacDonald BT, Haug K, Sander T, Meisler MH. (2000). A novel *SCN1A* mutation associated with generalised epilepsy with febrile seizures and prevalence of variants in patients with epilepsy. *Am J Hum Genet* **68**:866-873.

Everett LA, Glaser B, Beck JC, Idol JR, Buchs A, Heyman M, Adawi F, Hazani E, Nassir E, Baxevanis AD, Sheffield VC, Green ED. (1997). Pendred syndrome is caused by mutations in a putative sulphate transporter gene (*PDS*). *Nature Genet* **17**:411-422.

Fatkin D and Graham RM. (2002). Molecular mechanisms of inherited cardiomyopathies. *Physiol Rev* **82**:945-980.

Fatkin D, MacRae C, Sasaki T, Wolff MR, Porcu M, Frenneaux M, Atherton J, Vidaillet HJ Jr, Spudich S, De Girolami U, Seidman JG, Seidman C, Muntoni F, Muehle G, Johnson W, McDonough B. (1999). Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease. *New Engl J Med* **341**:1715-1724.

Felsenfeld A, Peterson J, Schloss J, Guyer M. (1999). Assessing the quality of DNA sequence from the Human Genome Project. *Genome Res* **9**:1-4.

Ferlini A, Galie N, Merlini L, Sewry C, Branzi A, Muntoni F. (1998). A novel Alu-like element rearranged in the dystrophin gene causes a splicing mutation in a family with X-linked dilated cardiomyopathy. *Am J Hum Genet* **63**:436-446.

Fernandez P, Corfield VA, Brink PA. (2004). Progressive familial heart block type II (PFHBII): a clinical profile from 1977 to 2003. *Cardiovasc J South Afr* **15**:129-132.

Fernandez P, Moolman-Smook J, Brink P, Corfield V. (2004). A gene locus for progressive familial heart block type II (PFHBII) maps to chromosome 1q32.2-q32.3. Manuscript submitted.

Franz W-M, Müller O, Katus HA. (2001). Cardiomyopathies: from genetics to the prospect of treatment. *Lancet* **358**:1627-1637.

Frith MC, Fu Y, Yu L, Chen JF, Hansen U, Weng Z. (2004). Detection of functional DNA motifs via statistical over-representation. *Nucleic Acids Res* **32**:1372-1381.

Fung YW, Wang RX, Heng HH, Liew CC. (1995). Mapping of a human LIM protein (CLP) to human chromosome 11p15.1 by fluorescence in situ hybridization. *Genomics* **28**:602-603.

Garcia-Higuera I, Fenoglio J, Li Y, Lewis C, Panchenko MP, Reiner O, Smith TF, Neer EJ. (1996). Folding of proteins with WD-repeats: comparison of six members of the WD-repeat superfamily to the G protein beta subunit. *Biochemistry* **35**:13985-13994.

Gardner RJ, Hanson JW, Ionasescu VV, Ardinger HH, Skorton DJ, Mahoney LT, Hart MN, Rose EF, Smith WL, Florentine MS. (1987). Dominantly inherited dilated cardiomyopathy. *Am J Med Genet* **27**:61-73.

Garros B, Cantrelle P, Hatton F. (1980). Cardiovascular mortality in industrial and developing countries. *Rev Epidemiol Sante Publique* **28**:69-79.

Gazes PC, Culler RM, Taber E, Kelly TE. (1965). Congenital familial cardiac conduction defects. *Circulation* **32**:32-34.

Geier C, Perrot A, Ozcelik C, Binner P, Counsell D, Hoffmann K, Pilz B, Martiniak Y, Gehmlich K, van der Ven PF, Furst DO, Vornwald A, von Hodenberg E, Nurnberg P, Scheffold T, Dietz R, Osterziel KJ. (2003). Mutations in the human muscle LIM protein gene in families with hypertrophic cardiomyopathy. *Circulation* **107**:1390-1395.

Geisterfer-Lowrance AA, Kass S, Tanigawa G, Vosberg HP, McKenna W, Seidman CE, Seidman JG. (1990). A molecular basis for familial hypertrophic cardiomyopathy: a beta cardiac myosin heavy chain gene missense mutation. *Cell* **62**:999-1006.

Gerecke DR, Wagman DW, Champlaud MF, Burgeson RE. (1994). The complete primary structure for a novel laminin chain, laminin B1k chain. *J Biol Chem* **269**:11073-11080.

Gerull B, Gramlich M, Atherton J, McNabb M, Trombitas K, Sasse-Klaassen S, Seidman JG, Seidman C, Granzier H, Labeit S, Frenneaux M, Thierfelder L. (2002). Mutations of *TTN*, encoding the giant muscle filament titin, cause familial dilated cardiomyopathy. *Nature Genet* **30**:201-204.

Giles TD. (1997). New WHO/ISFC classification of cardiomyopathies: a task not completed. *Circulation* **96**:2081-2082.

Goodwin JF. (1970). Congestive and hypertrophic cardiomyopathies. *Lancet* **I**: 731-739.

Graber HL, Unverferth DV, Baker PB, Ryan JM, Baba N, Wooley CF. (1986). Evolution of a hereditary cardiac conduction and muscle disorder: a study involving a family with six generations affected. *Circulation* **74**:21-35.

Grimwood J, Gordon LA, Olsen A, Terry A, Schmutz J, Lamerdin J, Hellsten U, Goodstein D, Couronne O, Tran-Gyamfi M, et al. (2004). The DNA sequence and biology of human chromosome 19. *Nature* **428**:529-535.

Grünig E, Tasman JA, Kucherer H, Franz W, Kubler W, Katus HA. (1998). Frequency and phenotypes of familial dilated cardiomyopathy. *J Am Coll Cardiol* **31**:186-194.

Guttmacher AE and Collins FS. (2002). Genomic medicine: a primer. *New Engl J Med* **347**:1512-1520.

Hanson EL, Jakobs PM, Keegan H, Coates K, Bousman S, Dienel NH, Litt M, Hershberger RE. (2002). Cardiac troponin T lysine 210 deletion in a family with dilated cardiomyopathy. *J Card Fail* **8**:28-32.

Hardie DG and Carling D. (1997). The AMP-activated protein kinase – fuel gauge of the mammalian cell. *Eur J Biochem* **246**:259-273.

Harrison PM, Kumar A, Lang N, Snyder M, Gerstein M. (2002). A question of size: the eukaryotic proteome and the problems in defining it. *Nucleic Acids Res* **30**:1083-1090.

Hayden MR, Hopkins HC, Macrea M, Beighton PH. (1980). The origin of Huntington's chorea in the Afrikaner population of South Africa. *S Afr Med J* **58**:197-200.

Henikoff S and Henikoff JG. (1992). Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci USA* **89**:10915-10919.

Hodge SE and Greenberg DA. (1992). Sensitivity of lod scores to changes in diagnostic status. *Am J Hum Genet* **50**:1053-1066.

Hoffmann B, Schmidt-Traub H, Perrot A, Osterziel KJ, Gessner R. (2001). First mutation in cardiac troponin C, L29Q, in a patient with hypertrophic cardiomyopathy. *Hum Mutation* **17**:524.

Hoffman EP, Brown RH Jr, Kunkel LM. (1987). Dystrophin: the protein product of the Duchenne muscular dystrophy locus. *Cell* **51**:919-928.

Hutchison CJ. (2002). Lamins: building blocks or regulators of gene transcription? *Nature Reviews* **3**:848-858.

Huynen M, Doerks T, Eisenhaber F, Orengo C, Sunyaev S, Yuan Y, Bork P. (1998). Homology-based fold predictions for *Mycoplasma genitalium* proteins. *J Mol Biol* **280**:323-326.

Ibsen HH, Baandrup U, Simonsen EE. (1985). Familial right ventricular dilated cardiomyopathy. *Br Heart J* **54**:156-159.

Ikawa Y. (1991). The Human Genome Project: past, present and future. *Science* **248**:44-49.

Initial Human Genome Sequencing Consortium. (2001). Initial sequencing analysis of the human genome. *Nature* **409**:860-921.

Inoue K, Matsuda K, Itoh M, Kawaguchi H, Tomoike H, Aoyagi T, Nagai R, Hori M, Nakamura Y, Tanaka T. (2002). Osteopenia and male-specific sudden cardiac death in mice lacking a zinc transporter gene, *Znt5*. *Hum Mol Genet* **11**:1775-1784.

Istrail S, Sutton GG, Florea L, Halpern AL, Mobarry CM, Lippert R, Walenz B, Shatkay H, Dew I, Miller JR, Flanigan MJ, Edwards NJ, Bolanos R, Fasulo D, Halldorsson BV, Hannenhalli S, Turner R, Yooseph S, Lu F, Nusskern DR, Shue BC, Zheng XH, Zhong F, Delcher AL, Huson DH, Kravitz SA, Mouchard L, Reinert K, Remington KA, Clark AG, Waterman MS, Eichler EE, Adams MD, Hunkapiller MW, Myers EW, Venter JC. (2004). Whole-genome shotgun assembly and comparison of human genome assemblies. *Proc Natl Acad Sci USA* **101**:1916-1921.

Jaenicke T, Diederich KW, Haas W, Schleich J, Lichter P, Pfordt M, Bach A, Vosberg HP. (1990). The complete sequence of the human beta-myosin heavy chain gene and a comparative analysis of its product. *Genomics* **8**:194-206.

James TN. (2002). Structure and function of the sinus node, AV node and His bundle of the human heart: part I-structure. *Prog Cardiovasc Dis* **45**:235-267.

James TN and Sherf L. (1971). Fine structure of the His bundle. *Circulation* **44**:9-28.

Jarcho JA, McKenna W, Pare JA, Solomon SD, Holcombe RF, Dickie S, Levi T, Donis-Keller H, Seidman JG, Seidman CE. (1989). Mapping a gene for familial hypertrophic cardiomyopathy to chromosome 14q1. *New Engl J Med* **321**:1372-1378.

Josephson ME, Marchlinski FE, Buxton AE. (1991). The bradyarrhythmias. In: *Principles of internal medicine*. pp 902-908. Wilson JD, Braunwald E, Isselbacher KJ, Peteresdorf RG, Martin JB, Fauci AS, Root RK, eds. New York: McGraw-Hill.

Jung M, Poepping I, Perrot A, Ellmer AE, Wienker TF, Dietz R, Reis A, Osterziel KJ. (1999). Investigation of a family with autosomal dominant dilated cardiomyopathy defines a novel locus on chromosome 2q14-q22. *Am J Hum Genet* **65**:1068-1077.

Kaczorowski T and Szybalski W. (1996). Automated four-color DNA sequencing using primers assembled by hexamer ligation. *Gene* **179**:195-198.

Kalaydjieva L, Gresham D, Gooding R, Heather L, Baas F, de Jonge R, Blechschmidt K, Angelicheva D, Chandler D, Worsley P, Rosenthal A, King RH, Thomas PK. (2000). N-myc

downstream-regulated gene 1 is mutated in hereditary motor and sensory neuropathy-Lom. *Am J Hum Genet* **67**:47-58.

Kamisago M, Sharma S, DePalma SR, Solomon S, Sharma P, McDonough B, Smoot L, Mullen MP, Woolf PK, Wigle D, Seidman JG, Seidman CE. (2000). Mutations in sarcomere protein genes as a cause of dilated cardiomyopathy. *New Engl J Med* **343**:1688-1696.

Kariv I, Szeinberg A, Fabian I, Sherf L, Kreisler B, Zeltzer M. (1966). A family with cardiomyopathy. *Am J Med* **40**:140-148.

Karkkainen S, Peuhkurinen K, Jaaskelainen P, Miettinen R, Karkkainen P, Kuusisto J, Laakso M. (2002). No variants in the cardiac actin gene in Finnish patients with dilated or hypertrophic cardiomyopathy. *Am Heart J* **143**:E6.

Kass S, Macrae C, Graber HL, Sparks EA, McNamara D, Boudoulas H, Basson CT, Baker PB 3rd, Cody RJ, Fishman MC, Cox N, Kong A, Wooley CF, Seidman JG, Seidman CE. (1994). A gene defect that causes conduction system disease and dilated cardiomyopathy maps to chromosome 1p1-1q1. *Nature Genet* **7**:546-551.

Katsanis N, Worley KC, Lupski JR. (2001). An evaluation of the draft human genome sequence. *Nature Genet* **29**:88-91.

Kearns TP. (1965). External ophthalmoplegia, pigmentary degeneration of the retina, and cardiomyopathy: a newly recognised syndrome. *Trans Ophthal Soc UK* **63**: 559-625.

Keating MT and Sanguinetti MC. (2001). Molecular and cellular mechanisms of cardiac arrhythmias. *Cell* **104**:569-580.

Kimura A, Harada H, Park JE, Nishi H, Satoh M, Takahashi M, Hiroi S, Sasaoka T, Ohbuchi N, Nakamura T, Koyanagi T, Hwang TH, Choo JA, Chung KS, Hasegawa A, Nagai R, Okazaki O, Nakamura H, Matsuzaki M, Sakamoto T, Toshima H, Koga Y, Imaizumi T, Sasazuki T. (1997). Mutations in the cardiac troponin I gene associated with hypertrophic cardiomyopathy. *Nature Genet* **16**:379-382.

Knöll R, Hoshijima M, Hoffman HM, Person V, Lorenzen-Schmidt I, Bang ML, Hayashi T, Shiga N, Yasukawa H, Schaper W, McKenna W, Yokoyama M, Schork NJ, Omens JH, McCulloch AD, Kimura A, Gregorio CC, Poller W, Schaper J, Schultheiss HP, Chien KR. (2002). The cardiac mechanical stretch sensor machinery involves a Z disc complex that is defective in a subset of human dilated cardiomyopathy. *Cell* **111**:943-955.

Koenig M, Monaco AP, Kunkel LM. (1988). The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein. *Cell* **51**:219-226.

Komajda M and Charron P. (2001). How will the human genome project change cardiovascular medicine? *Heart* **86**:123-124.

Kondo S, Schutte BC, Richardson RJ, Bjork BC, Knight AS, Watanabe Y, Howard E, Ferreira de Lima RLL, Daarck-Hirsch S, Sander A, McDonald-McGinn DM, Zackai EH, Lammer EJ, Aylsworth AS, Ardinger HH, Lidral AC, Pober BP, Moreno L, Arcos-Burgos M, Valencia C, Houdayer C, Bahuau M, Moretti-Ferreira D, Richieri-Costa A, Dixon MJ, Murray JC. (2002).

Mutations in *IRF6* cause Van der Woude and politeal pterygium syndromes. *Nature Genet* **32**: 285-289.

Krajinovic M, Pinamonti B, Sinagra G, Vatta M, Severini GM, Milasin J, Falaschi A, Camerini F, Giacca M, Mestroni L. (1995). Linkage of familial dilated cardiomyopathy to chromosome 9. *Am J Hum Genet* **57**:846-852.

Krasnow N, Qazi QH, Yermakov V. (1985). A familial dilated cardiomyopathy associated with cataracts and hip-spine disease. *Chest* **87**:56-61.

Larsen LA, Chritiansen M, Vuust. (2001). Recent developments in high-throughput mutation screening. *Pharmacogenomics* **2**: 387-399.

Lathrop GM and Lalouel JM. (1984). Easy calculations of LOD scores and genetic risks on small computers. *Am J Hum Genet* **36**:460-465.

Le Hellard S, Semple CA, Morris SW, Porteous DJ, Evans KL. (2001). Physical mapping: integrating computational and molecular genetic data. *Ann Hum Genet* **65**:221-228.

Lenègre J. (1964). Progressive cardiovascular disorders. *Prog Cardiovasc Dis* **6**:317-323.

Lev M, Kinare S G, Pick. (1970). The pathogenesis of atrioventricular block in coronary disease. *Circulation* **42**:409-425.

Li D, Czernuszewicz GZ, Gonzalez O, Tapscoft T, Karibe A, Durand JB, Brugada R, Hill R, Gregoritch JM, Anderson JL, Quinones M, Bachinski LL, Roberts R. (2001). Novel cardiac troponin T mutation as a cause of familial dilated cardiomyopathy. *Circulation* **104**:2188-2193.

Li D, Tapscoft T, Gonzalez O, Burch PE, Quinones MA, Zoghbi WA, Hill R, Bachinski LL, Mann DL, Roberts R. (1999). Desmin mutation responsible for idiopathic dilated cardiomyopathy. *Circulation* **100**:461-464.

Lin F and Worman HJ. (1993). Structural organisation of the human gene encoding nuclear lamin A and nuclear lamin C. *J Biol Chem* **268**:16321-16326.

Lynch HT, Mohiuddin S, Moran J, Kaplan A, Sketch M, Zencka A, Runco V. (1975). Hereditary progressive atrioventricular conduction defect. *Am J Cardiol* **36**:297-301

Maass A and Leinwand L. (2000). Animal models of hypertrophic cardiomyopathy. *Curr Opin Cardiol* **15**:189-196.

MacRae CA, Ghaisas N, Kass S, Donnelly S, Basson CT, Watkins HC, Anan R, Thierfelder LH, McGarry K, Rowland E, McKenna WJ, Seidman JG, Seidman CE. (1995). Familial hypertrophic cardiomyopathy with Wolff-Parkinson-White syndrome maps to a locus on chromosome 7q3. *J Clin Invest* **96**:1216-1220.

Maeda M, Murayama N, Ishii H, Uryu N, Ota M, Tsuji K, Inoko H. (1989). A simple and rapid method for HLA-DQA1 genotyping by digestion of PCR-amplified DNA with allele specific restriction endonucleases. *Tissue Antigens* **34**:290-298.

Makridakis NM and Reichardt JK. (2001). Multiplex automated primer extension analysis: simultaneous genotyping of several polymorphisms. *Biotechniques* **31**:1374-1380.

Manolio TA, Baughman KL, Rodeheffer R, Pearson TA, Bristow JD, Michels VV, Abelmann WH, Harlen WR. (1992). Prevalence and aetiology of idiopathic dilated cardiomyopathy (summary of a National Heart, Lung, and Blood Institute workshop). *Am J Cardiol* **69**:1458-1466.

Marcus FI, Fontaine GH, Guiraudon G, Frank R, Laurenceau JL, Malergue C, Grosogeat Y. (1982). Right ventricular dysplasia: a report of 24 adult cases. *Circulation* **65**:384-398.

Marian AJ and Roberts R. (2001). The molecular genetic basis for hypertrophic cardiomyopathy. *J Mol Cell Cardiol* **33**:655-670.

Marian AJ and Roberts R. (2003). To screen or not is not the question - it is when and how to screen. *Circulation* **107**:2171-2174.

Maron BJ. (1997). Hypertrophic cardiomyopathy. *Lancet* **350**:127-133.

Maron BJ, Edwards JE, Henry WL, Clark CE, Bingle GJ, Epstein SE. (1974). Asymmetric septal hypertrophy (ASH) in infancy. *Circulation* **50**:809-820.

Maron BJ and Epstein SE. (1979). Hypertrophic cardiomyopathy: a discussion of nomenclature. *Am J Cardiol* **43**:1242-1244.

Maron BJ, Gardin JM, Flack JM, Gidding SS, Kurosaki TT, Bild DE. (1995). Prevalence of hypertrophic cardiomyopathy in a general population of young adults. Echocardiographic analysis

of 4111 subjects in the CARDIA Study. Coronary artery risk development in (young) adults. *Circulation* **92**:785-789.

Maron BJ, Nichols PF 3rd, Pickle LW, Wesley YE, Mulvihill JJ. (1984). Patterns of inheritance in hypertrophic cardiomyopathy: assessment by M-mode and two-dimensional echocardiography. *Am J Cardiol* **53**:1087-1094.

Matsuura T, Yamagata T, Burgess DL, Rasmussen A, Grewal RP, Watase K, Khajavi M, McCall AE, Davis CF, Zu L, Achari M, Pulst SM, Alonso E, Noebels JL, Nelson DL, Zoghbi HY, Ashizawa T. (2000). Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. *Nature Genet* **26**:191-194.

Mattick JS. (1994). Introns: evolution and function. *Curr Opin Genet Dev* **4**:823-831.

Mayosi BM, Khogali S, Zhang B, Watkins H. (1999). Cardiac and skeletal actin gene mutations are not a common cause of dilated cardiomyopathy. *J Med Genet* **36**:796-797.

McCright B and Virshup DM. (1995). Identification of a new protein phosphatase 2A regulatory subunit. *J Biol Chem* **270**:26123-26128.

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.

McTiernan CF, Frye CS, Lemster BH, Kinder EA, Ogletree-Hughes ML, Moravec CS, Feldman AM. (1999). The human phospholamban gene: structure and expression. *J Mol Cell Cardiol* **31**:679-692.

Messina DN, Speer MC, Pericak-Vance MA, McNally EM. (1997). Linkage of familial dilated cardiomyopathy with conduction defect and muscular dystrophy to chromosome 6p23. *Am J Hum Genet* **61**:909-917.

Melov S, Shoffner JM, Kaufman A, Wallace DC. (1995). Marked increase in the number and variety of mitochondrial DNA rearrangements in aging human skeletal muscle. *Nucleic Acids Res* **23**:4122-4126.

Michels VV, Moll PP, Miller FA, Tajik AJ, Chu JS, Driscoll DJ, Barnett JC, Rodeheffer RJ, Chesebro JH, Tazelaar HD. (1992). The frequency of familial dilated cardiomyopathy in a series of patients with idiopathic dilated cardiomyopathy. *New Engl J Med* **326**:77-82.

Milan D, Jeon JT, Looft C, Amarger V, Robic A, Thelander M, Rogel-Gaillard C, Paul S, Iannuccelli N, Rask L, Ronne H, Lundstrom K, Reinsch N, Gellin J, Kalm E, Roy PL, Chardon P, Andersson L. (2000). A mutation in *PRKAG3* associated with excess glycogen content in pig skeletal muscle. *Science* **288**:1248-1251.

Minamisawa S, Sato Y, Tatsuguchi Y, Fujino T, Imamura S, Uetsuka Y, Nakazawa M, Matsuoka R. (2003). Mutation of the phospholamban promoter associated with hypertrophic cardiomyopathy. *Biochem Biophys Res Commun* **304**:1-4.

Mogensen J, Klausen IC, Pederson AK, Egebkad G, Bross P, Kruse TA, Gregorsen N, Hansen PS, Baandrup U, Borglum AD. (1999). α -Cardiac actin is a novel disease gene in familial hypertrophic cardiomyopathy. *J Clin Invest* **103**:R39-R43.

Mohler PJ, Schott J-J, Gramolini AO, Dilly KW, Guatimosim S, duBell WH, Song L-S, Haurogné K, Kyndt F, Ali ME, Rogers TB, Lederer WJ, Escande D, Le Marec H, Bennett V. (2003). Ankyrin-B causes type 4 long-QT cardiac arrhythmia and sudden cardiac death. *Nature* **421**:634-638.

Molkentin JD, Lu J-R, Antos CL, Markham B, Richardson J, Robbins J, Grant SR, Olson EN. (1998). A calcineurin-dependent transcriptional pathway for cardiac hypertrophy. *Cell* **93**:215-228.

Moller P, Lunde P, Hovig T, Nitter-Hauge S. (1979). Familial cardiomyopathy. Autosomally, dominantly inherited congestive cardiomyopathy with two cases of septal hypertrophy in one family. *Clin Genet* **16**:233-243.

Moolman JC, Brink PA, Corfield VA. (1995). Identification of a novel Ala797Thr mutation in exon 21 of the beta-myosin heavy chain gene in hypertrophic cardiomyopathy. *Hum Mutation* **6**:197-198.

Moolman JC, Corfield VA, Posen B, Ngumbela K, Seidman C, Brink PA, Watkins H. (1997). Sudden death due to troponin T mutations. *J Am Coll Cardiol* **29**:549-555.

Moolman-Smook JC, De Lange WJ, Bruwer ECD, Brink PA, Corfield VA. (1999). The origins of hypertrophic cardiomyopathy-causing mutations in two South African sub-populations: a unique profile of both independent and founder events. *Am J Hum Genet* **65**:1308-1320.

Moolman-Smook J, Flashman E, de Lange W, Li Z, Corfield V, Redwood C, Watkins H. (2002). Identification of novel interactions between domains of myosin binding protein-C that are modulated by hypertrophic cardiomyopathy missense mutations. *Circulation Res* **91**:704-711.

Moolman-Smook JC, Mayosi B, Brink P, Corfield VA. (1998). Identification of a new missense mutation in MyBP-C associated with hypertrophic cardiomyopathy. *J Med Genet* **35**:253-254.

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.

Moreira ES, Wiltshire TJ, Faulkner G, Nilforoushan A, Vainzof M, Suzuki OT, Valle G, Reeves R, Zatz M, Passos-Bueno MR, Jenne DE. (2000). Limb-girdle muscular dystrophy type 2G is caused by mutations in the gene encoding the sarcomeric protein telethonin. *Nature Genet* **24**:163-136.

Morton NE. (1955). Sequential tests for the detection of linkage. *Am J Hum Genet* **7**:277-318.

Morquio L. (1901). Sur une maladie infantile et familiale caracterisee par des modifications permanentes du pouls, des attaques syncopales et epileptiformes et la mort subite. *Arch Med Enfants* **4**:467-475.

Muntoni F, Cau M, Ganau A, Congiu R, Arvedi G, Mateddu A, Marrosu MG, Cianchetti C, Realdi G, Cao A, Melis MA. (1993). Brief report: deletion of the dystrophin muscle-promoter region associated with X-linked dilated cardiomyopathy. *New Engl J Med* **329**:921-925.

Murphy RT, Mogensen J, Shaw A, Kobu T, Hughes S, McKenna WJ. (2004). Novel mutation in cardiac troponin I in recessive idiopathic dilated cardiomyopathy. *Lancet* **363**:371-372.

Myburgh DP and Steenkamp WFJ. (1973). The hereditary nature of adult-onset heart block. *S Afr Med J* **47**:657-658.

Myers DC and Fishman GI. (2003). Molecular and functional maturation of the murine cardiac conduction system. *Trends Cardiovasc Med* **13**:289-295.

Nagase T, Seki N, Ishikawa K, Ohira M, Kawarabayasi Y, Ohara O, Tanaka A., Kotani H., Miyajima N, Nomura N. (1996). Prediction of the coding sequences of unidentified human genes. VI. The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by analysis of cDNA clones from cell line KG-1 and brain. *DNA Res* **3**:321-329.

Nasser WK, Williams JF, Mishkin ME, Childress RH, Helmen C, Merritt AD, Genovese PD. (1967). Familial myocardial disease with and without obstruction to left ventricular outflow. Clinical, hemodynamic and angiographic findings. *Circulation* **35**:638-652.

Neitzel H. (1986). A routine method for the establishment of permanent growing lymphoblastoid cell lines. *Hum Genet* **73**:320-326.

Neuwald AF. (1997). Barth syndrome may be due to an acyltransferase deficiency. *Curr Biol* **7**:R465-R466.

Nievergelt CM, Smith DW, Kohlenberg JB, Schork NJ. (2004). Large-scale integration of human genetic and physical maps. *Genome Res* **14**:1-7.

Niimura H, Bachinski LL, Sangwatanaroj S, Watkins H, Chudley AE, McKenna W, Kristinsson A, Roberts R, Sole M, Maron BJ, Seidman JG, Seidman CE. (1998). Mutations in the gene for cardiac myosin-binding protein C and late-onset familial hypertrophic cardiomyopathy. *New Engl J Med* **338**:1248-1257.

Niimura H, Patton KK, McKenna WJ, Soultis J, Maron BJ, Seidman JG, Seidman CE. (2002). Sarcomere protein gene mutations in hypertrophic cardiomyopathy of the elderly. *Circulation* **105**:446-451.

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.

Occhiodoro T, Bernheim L, Liu JH, Bijlenga P, Sinnreich M, Bader CR, Fischer-Lougheed J. (1998). Cloning of a human ether-a-go-go potassium channel expressed in myoblasts at the onset of fusion. *FEBS Letters* **434**:177-182.

Olson TM and Keating MT. (1996). Mapping a cardiomyopathy locus to chromosome 3p22-p25. *J Clin Invest* **97**:528-532.

Olson TM, Doan TP, Kishimoto NY, Whitby FG, Ackerman MJ, Fananapazir L. (2000). Inherited and de novo mutations in the cardiac actin gene cause hypertrophic cardiomyopathy. *J Mol Cell Cardiol* **32**:1686-1694.

Olson TM, Illenberger S, Kishimoto NY, Huttelmaier S, Keating MT, Jockusch BM. (2002). Metavinculin mutations alter actin interaction in dilated cardiomyopathy. *Circulation* **105**:431-437.

Olson TM, Kishimoto NY, Whitby FG, Michels VV. (2001). Mutations that alter the surface charge of alpha-tropomyosin are associated with dilated cardiomyopathy. *J Mol Cell Cardiol* **33**:723-732.

Olson TM, Michels VV, Thibodeau SN, Tai Y-S, Keating MT. (1998). Actin mutations in dilated cardiomyopathy, a heritable form of heart failure. *Science* **280**:750-752.

Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T. (1989). Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc Natl Acad Sci USA* **86**:2766-2770.

Ott J. (1999). Genetic loci and genetic polymorphisms. In: Ott J. Analysis of human genetic linkage 3rd ed. pp 24-36. Baltimore: The Johns Hopkins University Press.

Oudit GY, Crackower MA, Eriksson U, Sarao R, Koziarzki I, Sasaki T, Irie-Sasaki J, Gidrewicz D, Rybin VO, Wada T, Steinberg SF, Backx PH, Penninger JM. (2003). Phosphoinositide 3-kinase (gamma)-deficient mice are protected from isoproterenol-induced heart failure. *Circulation* **108**:1-6.

Ouyang M, Case J, Tirunagaru V, Burnside J. (2003). Five hundred sixty-five triples of chicken, human, and mouse candidate orthologues. *J Mol Evol* **57**:271-281

Ozawa T, Tanaka M, Sugiyama S, Hattori K, Ito T, Ohno K, Takahashi A, Sato W, Takada G, Mayumi B. (1990). Multiple mitochondrial DNA deletions exist in cardiomyocytes of patients with hypertrophic or dilated cardiomyopathy. *Biochem Biophys Res Commun* **170**:830-836.

Pennisi E. (2003). Human genome. A low number wins the GeneSweep Pool. *Science* **300**:1484.

Pennisi DJ, Rentschler S, Gourdie RG, Fishman GI, Mikawa T. (2002). Induction and patterning of the cardiac conduction system. *Int J Dev Biol* **46**:765-775.

Parimoo S, Patanjali SR, Shulka H, Chaplin DD, Weissman SM. (1991). cDNA selection: efficient PCR approach for the selection of cDNAs encoded in large chromosomal DNA fragments. *Proc Natl Acad Sci USA* **88**:9623-9627.

Perez-Iratxeta C, Bork P, Andrade MA. (2002). Association of genes to genetically inherited diseases using data mining. *Nature Genet* **31**:316-319.

Poetter K, Jiang H, Hassanzadeh S, Master SR, Chang A, Dalakas MC, Rayment I, Sellers JR, Fananapazir L, Epstein ND. (1996). Mutations in either the essential or regulatory light chains of myosin are associated with a rare myopathy in human heart and skeletal muscle. *Nature Genet* **13**:63-69.

Probst V, Kyndt F, Potet F, Trochu JN, Mialet G, Demolombe S, Schott JJ, Baro I, Escande D, Le Marec H. (2003). Haploinsufficiency in combination with aging causes *SCN5A*-linked hereditary Lenegre disease. *J Am Coll Cardiol* **41**:643-652.

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.

Puca AA, Nigro V, Piluso G, Belsito A, Sampaolo S, Quaderi N, Rossi E, Di Iorio G, Ballabio A, Franco B. (1998). Identification and characterisation of a novel member of the dystrobrevin gene family. *FEBS Letters* **425**:7-13.

Pulkkinen L, Christiano AM, Gerecke D, Wagman DW, Burgeson RE, Pittelkow MR, Uitto J. (1994). A homozygous nonsense mutation in the beta 3 chain of laminin 5 (*LAMB3*) in Herlitz junctional epidermolysis bullosa. *Genomics* **15**:357-360.

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.

Redwood CS, Moolman-Smook JC, Watkins H. (1999). Properties of contractile proteins that cause hypertrophic cardiomyopathy. *Cardiovasc Res* **44**:20-36.

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

Roberts R. (2000). Bioinformatics analysis of gene banks provides a treasure trove for the functional genomist. *J Mol Cell Cardiol* **32**:1917-1919.

Rook MB, Bezzina Alshinawi C, Groenewegen WA, van Gelder IC, van Ginneken AC, Jongsma HJ, Mannens MM, Wilde AA. (1999). Human SCN5A gene mutations alter cardiac sodium channel kinetics and are associated with the Brugada syndrome. *Cardiovasc Res* **44**:507-517.

Ross RS, Bulkley BH, Hutchins GM, Harshey JS, Jones RA, Kraus H, Liebman J, Thorne CM, Weinberg SB, Weech AA, Weech AA Jr. (1978). Idiopathic familial myocardopathy in three generations: a clinical and pathologic study. *Am Heart J* **96**:170-178.

Rust EM, Albayya RP, Metzger JM. (1999). Identification of a contractile deficit in adult cardiac myocytes expressing hypertrophic cardiomyopathy-associated mutant troponin T proteins. *J Clin Invest* **103**:1459-1467.

Sahn DJ, DeMaria A, Kisslo J, Weyman A. (1978). Recommendations regarding quantitation in M-mode echocardiography: results of a survey of echocardiographic measurements. *Circulation* **58**:1072-1083.

Sambrook J, Fritsch EF, Maniatis T. (1989). Molecular cloning. In: *A laboratory manual* pp 1.34-1.35. Cold Spring Harbour: CSHL Press.

Sansom CE and Smith CA. (2000). Computer applications in biomolecular sciences - part 2: bioinformatics and genome projects. *Biochemical Education* **28**:127-131.

Sarachek NS and Leonard JJ. (1972). Familial heart block and sinus bradycardia. *Am J Cardiol* **29**:451-458.

Satoh M, Takahashi M, Sakamoto T, Hiroe M, Marumo F, Kimura A. (1999). Structural analysis of the titin gene in hypertrophic cardiomyopathy: identification of a novel disease gene. *Biochem Biophys Res Commun* **262**:411-417.

Schaal SF, Seidensticker J, Goodman R, Wooley CF. (1973). Familial right bundle-branch block, left axis deviation, complete heart block, and early death. A heritable disorder of cardiac conduction. *Ann Intern Med* **79**:63-66.

Schmidt MA, Michels VV, Edwards WD, Miller FA. (1988). Familial dilated cardiomyopathy. *Am J Med Genet* **31**:135-143.

Schmitt JP, Kamisago M, Asahi M, Li GH, Ahmad F, Mende U, Kranias EG, MacLennan DH, Seidman JG, Seidman CE. (2003). Dilated cardiomyopathy and heart failure caused by a mutation in phospholamban. *Science* **299**:1410-1413.

Schönberger J, Levy H, Grünig E, Sangwatanaroj S, Fatkin D, MacRae C, Stacker H, Halpin C, Eavey R, Philbin EF, Katus H, Seidman JG, Seidman CE. (2000). Dilated cardiomyopathy and sensorineural hearing loss: a heritable syndrome that maps to 6q23-24. *Circulation* **101**:1812-1818.

Schönberger J and Seidman CE. (2001). Many roads lead to a broken heart: the genetics of dilated cardiomyopathy. *Am J Hum Genet* **69**:249-260.

Schott J-J, Alshinawi C, Kyndt F, Probst V, Hoorntje TM, Hulsbeek M, Wilde AA, Escande D, Mannens MM, Le Marec H. (1999). Cardiac conduction defects associate with mutations in *SCN5A*. *Nature Genet* **23**:20-21.

Schram G, Pourrier M, Melnyk P, Nattel S. (2001). Differential distribution of cardiac ion channel expression as a basis for regional specialization in electrical function. *Circulation Res* **90**:939-950.

Schultz SJ, Fry AM, Sutterlin C, Ried T, Nigg EA. (1994). Cell cycle-dependent expression of Nek2, a novel human protein kinase related to the NIMA mitotic regulator of *Aspergillus nidulans*. *Cell Growth Differ* **5**:625-635.

Schutte BC, Bjork BC, Coppage KB, Malik MI, Gregory SG, Scott DJ, Brentzell LM, Watanabe Y, Dixon MJ, Murray JC.(2000). A preliminary gene map for the the Van der Woude syndrome critical region derived from 900 kb of genomic sequence at 1q32-q41. *Genome Res* **10**:81-94.

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

Sekiguchi M, Hasegawa A, Ando M. (1978). A proposal for analyzing the familial occurrence of cardiomyopathy in the clinical situation. In: *Cardiomyopathy*. pp445-353. Sekiguchi M and Olson EGJ, eds. Tokyo: University of Tokyo Press.

Sheffield VC, Beck JS, Kwitek AE, Sandstrom DW, Stone EM. (1993). The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. *Genomics* **16**:325-332.

Silvestri G, Santorelli FM, Shanske S, Whitley CB, Schimmenti LA, Smith SA, DiMauro S. (1994). A new mtDNA mutation in the tRNA(Leu[UUR]) gene associated with maternally inherited cardiomyopathy. *Hum Mutation* **3**:37-43.

Siu BL, Nimimura H, Osborne JA, Fatkin D, MacRae C, Solomon S, Woodrow Benson D, Seidman JG, Seidman CE. (1999). Familial dilated cardiomyopathy locus maps to chromosome 2q31. *Circulation* **99**:1022-1026.

Solomon SD, Jarcho JA, McKenna W, Geisterfer-Lowrance A, Germain R, Salerni R, Seidman JG, Seidman CE. (1990). Familial hypertrophic cardiomyopathy is a genetically heterogeneous disease. *J Clin Invest* **86**:993-999.

Somura F, Izawa H, Iwase M, Takeichi Y, Ishiki R, Nishizawa T, Noda A, Nagata K, Yamada Y, Yokota M. (2001). Reduced myocardial sarcoplasmic reticulum Ca²⁺-ATPase mRNA expression and biphasic force-frequency relations in patients with hypertrophic cardiomyopathy. *Circulation* **104**:658-663.

Spindler M, Saupe KW, Christe ME, Sweeney HL, Seidman CE, Seidman JG, Ingwall JS. (1998). Diastolic dysfunction and altered energetics in the alphaMHC^{403/+} mouse model of familial hypertrophic cardiomyopathy. *J Clin Invest* **101**:1775-1783.

Splawski I, Tristani-Firouzi M, Lehmann MH, Sanguinetti MC, Keating MT. (1997). Mutations in the hminK gene cause long QT syndrome and suppress IKs function. *Nature Genet* **17**:338-340.

Spyrou N, Philpot J, Foale R, Camici PG, Muntoni F. (1998). Evidence of left ventricular dysfunction in children with merosin-deficient congenital muscular dystrophy. *Am Heart J* **136**:474-476.

Stein L. (2001). Genome annotation: from sequence to biology. *Nature Rev Genet* **2**:493-503.

Stephan E. (1978). Hereditary bundle branch system defect: survey of a family with four affected generations. *Am Heart J* **95**:89-95.

Strachan T and Read AP. (1999). Molecular pathology. In: *Human molecular genetics* 2nd ed. pp 377-399. Oxford: Bios Scientific Publishers.

Studer R, Reinecke H, Bilger J, Eschenhagen T, Bohm M, Hasenfuss G, Just H, Holtz J, Drexler H. (1994). Gene expression of the cardiac Na⁺-Ca²⁺ exchanger in end-stage human heart failure. *Circulation Res* **75**:443-453.

Sullivan LS, Heckenlively JR, Bowne SJ, Zuo J, Hide WA, Gal A, Denton M, Inglehearn CF, Blanton SH, Daiger SP. (1999). Mutations in a novel retina-specific gene cause autosomal dominant retinitis pigmentosa. *Nature Genet* **22**:255-259.

Sweeney HL, Feng HS, Yang Z, Watkins H. (1998). Functional analyses of troponin T mutations that cause hypertrophic cardiomyopathy: insights into disease pathogenesis and troponin function. *Proc Natl Acad Sci USA* **95**:14406-14410.

Sylvius N, Tesson F, Gayet C, Charron P, Benaiche A, Peuchmaurd M, Duboscq-Bidot L, Feingold J, Beckmann JS, Bouchier C, Komajda M. (2001). A new locus for autosomal dominant dilated cardiomyopathy identified on chromosome 6q12-q16. *Am J Hum Genet* **68**:241-246.

Takai E, Akita H, Shiga N, Kanazawa K, Yamada S, Terashima M, Matsuda Y, Iwai C, Kawai K, Yokota Y, Yokoyama M. (1999). Mutational analysis of the cardiac actin gene in familial and sporadic dilated cardiomyopathy. *Am J Med Genet* **86**:325-327.

Takano H, Hasegawa H, Nagai T, Komuro I. (2003). Implication of cardiac remodeling in heart failure: mechanisms and therapeutic strategies. *Intern Med* **42**:465-469.

Takio K, Blumenthal DK, Edelman AM, Walsh KA, Krebs EG, Titani K. (1985). Amino acid sequence of an active fragment of rabbit skeletal muscle myosin light chain kinase. *Biochemistry* **24**:6028-6037.

Talajic M, Papadatos D, Villemaire C, Glass L, Nattel S. (1991). A unified model of atrioventricular nodal conduction predicts dynamic changes in Wenckebach periodicity. *Circulation Res* **68**:1280-1293.

Tan HL, Bink-Boelkens MT, Bezzina CR, Viswanathan PC, Beaufort-Krol GC, van Tintelen PJ, van den Berg MP, Wilde AA, Balser JR. (2001). A sodium-channel mutation causes isolated cardiac conduction disease. *Nature* **409**:1043-1047.

Tanaka M, Ino H, Ohno K, Hattori K, Sato W, Ozawa T, Tanaka T, Itoyama S. (1990). Mitochondrial mutation in fatal infantile cardiomyopathy. *Lancet* **336**:1452.

Tanigawa G, Jarcho JA, Kass S, Solomon SD, Vosberg HP, Seidman JG, Seidman CE. (1990). A molecular basis for familial hypertrophic cardiomyopathy: an alpha/beta cardiac myosin heavy chain hybrid gene. *Cell* **62**:991-998.

Tardiff JC, Hewett TE, Palmer BM, Olsson C, Factor SM, Moore RL, Robbins J, Leinwand LA. (1999). Cardiac troponin T mutations result in allele-specific phenotypes in a mouse model for hypertrophic cardiomyopathy. *J Clin Invest* **104**:469-481.

Teare D. (1958). Asymmetrical hypertrophy of the heart in young adults. *Br Heart J* **20**:1-8.

Thierfelder L, MacRae C, Watkins H, Tomfohrde J, Williams M, McKenna W, Bohm K, Noeske G, Schlepper M, Bowcock A, Vosberg H-P, Seidman JG, Seidman C. (1993). A familial hypertrophic cardiomyopathy locus maps to chromosome 15q2. *Proc Natl Acad Sci USA* **90**:6270-6274.

Thierfelder L, Watkins H, MacRae C, Lamas R, McKenna W, Vosberg HP, Seidman JG, Seidman CE. (1994). Alpha-tropomyosin and cardiac troponin T mutations cause familial hypertrophic cardiomyopathy: a disease of the sarcomere. *Cell* **77**:701-712.

Tiso N, Stephan DA, Nava A, Bagattin A, Devaney JM, Stanchi F, Larderet G, Brahmabhatt B, Brown K, Baucé 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.

Torrington MA. (1979). The sociomedical and psychosocial implications of some hereditary diseases in South Africa. PhD thesis, pp 89-97. University of Stellenbosch.

Torrington M, Weymar HW, Van der Merwe P-L, Brink AJ. (1986). Progressive familial heart block. Part I. Extent of the disease. *S Afr Med J* **70**:353-355.

Towbin JA. (1998). The role of cytoskeletal proteins in cardiomyopathies. *Curr Opin Cell Biol* **10**:131-139.

Towbin JA and Bowles N. (2001). Molecular genetics of left ventricular dysfunction. *Curr Mol Med* **1**:81-90.

Towbin JA and Bowles NE. (2002). The failing heart. *Nature* **415**:227-233.

Towbin JA and Bowles NE. (2004). Heart failure as a consequence of genetic cardiomyopathy. In: *Heart failure: a companion to Braunwald's heart disease vol. 6.* pp 391-403. Mann DC and Braunwald E, eds. Philadelphia: Saunders Press.

Towbin JA, Hejtmancik JF, Brink P, Gelb B, Zhu XM, Chamberlain JS, McCabe ER, Swift M. (1993). X-linked dilated cardiomyopathy: molecular genetic evidence of linkage to the Duchenne muscular dystrophy (dystrophin) gene at the Xp21 locus. *Circulation* **87**:1854-1865.

Tsubata S, Bowles KR, Vatta M, Zintz C, Titus J, Muhonen L, Bowles NE, Towbin JA. (2000). Mutations in the human delta-sarcoglycan gene in familial and sporadic dilated cardiomyopathy. *J Clin Invest* **106**:655-662.

Tyson J, Tranebjaerg L, Bellman S, Wren C, Taylor JF, Bathen J, Aslaksen B, Sorland SJ, Lund O, Malcolm S, Pembrey M, Bhattacharya S, Bitner-Glindzicz M. (1997). IsK and KvLQT1: mutation in either of the two subunits of the slow component of the delayed rectifier potassium channel can cause Jervell and Lange-Nielsen syndrome. *Hum Mol Genet* **6**:2179-2185.

Van der Merwe P-L, Weymar HW, Torrington M, Brink AJ. (1986). Progressive familial heart block. Part II. Clinical and ECG confirmation of progression - report on 4 cases. *S Afr Med J* **70**:356-357.

Virágh S and Challice CE. (1982). The development of the conduction system in the mouse embryo heart. IV. Differentiation of the atrioventricular conduction system. *Dev Biol* **89**:24-40.

Wagner CW Jr and Hall RJ. (1967). Congenital familial atrioventricular dissociation. Report of three siblings. *Am J Cardiol* **19**:593-596

Wang D, Chang PS, Wang Z, Sutherland L, Richardson JA, Small E, Krieg PA, Olson EN. (2001). Activation of cardiac gene expression by myocardin, a transcriptional cofactor for serum response factor. *Cell* **105**:851-862.

Wang DW, Yazawa K, George AL Jr, Bennett PB. (1996). Characterisation of human cardiac Na⁺ channel mutations in the congenital long QT syndrome. *Proc Natl Acad Sci USA* **93**:13200-13205.

Wang Q, Curran ME, Splawski I, Burn TC, Millholland JM, VanRaay TJ, Shen J, Timothy KW, Vincent GM, de Jager T, Schwartz PJ, Toubin JA, Moss AJ, Atkinson DL, Landes GM, Connors TD, Keating MT. (1996). Positional cloning of a novel potassium channel gene: *KVLQT1* mutations cause cardiac arrhythmias. *Nature Genet* **12**:17-23.

Wang Q, Shen J, Splawski I, Atkinson D, Li Z, Robinson JL, Moss AJ, Towbin JA, Keating MT. (1995). *SCN5A* mutations associated with an inherited cardiac arrhythmia, long QT syndrome. *Cell* **80**:805-811.

Watkins H. (2003). Genetic clues to disease pathways in hypertrophic and dilated cardiomyopathies. *Circulation* **107**:1344-1346.

Watkins H, Conner D, Thierfelder L, Jarcho JA, MacRae C, McKenna WJ, Maron BJ, Seidman JG, Seidman CE. (1995). Mutations in the cardiac myosin-binding protein C gene on chromosome 11 cause familial hypertrophic cardiomyopathy. *Nature Genet* **11**:434-437.

Watkins H, MacRae C, Thierfelder L, Chou YH, Frenneaux M, McKenna W, Seidman JG, Seidman CE. (1993). A disease locus for familial hypertrophic cardiomyopathy maps to chromosome 1q3. *Nature Genet* **3**:333-337.

Watkins H, Seidman CE, Seidman JG, Feng HS, Sweeney HL. (1996). Expression and functional assessment of a truncated cardiac troponin T that causes hypertrophic cardiomyopathy. Evidence for a dominant negative action. *J Clin Invest* **98**:2456-2461.

Waxman NB, Catching JD, Felderhof CH, Downar E, Silver MD, Abbott MM. (1974). Familial atrioventricular heart block. An autosomal dominant trait. *Circulation* **51**:226-233.

Weber JL. (1990). Human DNA polymorphisms and methods of analysis. *Curr Opin Biotechnol* **1**:166-171.

Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, Lathrop M. (1992). A second-generation linkage map of the human genome. *Nature* **359**:794-801.

Whitfield AG. (1961). Familial cardiomyopathies. *Quart J Med* **30**:119-134.

Witt CC, Gerull B, Davies MJ, Centner T, Linke WA, Thierfelder L. (2001). Hypercontractile properties of cardiac muscle fibers in a knock-in mouse model of cardiac myosin-binding protein-C. *J Biol Chem* **276**:5353-5359.

Yaffe MB, Leparo GG, Lai J, Obata T, Volinia S, Cantley LC. (2001). A motif-based profile scanning approach for genome-wide prediction of signaling pathways. *Nature Biotechnol* **19**:348-353.

Yang Q, Sanbe A, Osinska H, Hewett TE, Klevitsky R, Robbins J. (1998). A mouse model of myosin-binding protein C human familial hypertrophic cardiomyopathy. *J Clin Invest* **102**:1292-1300.

Yung CK, Halperin VL, Tomaselli GF, Winslow RL. (2004). Gene expression profiles in end-stage human idiopathic dilated cardiomyopathy: altered expression of apoptotic and cytoskeletal genes. *Genomics* **83**:281-297.

Zhang X, Chai J, Azhar G, Sheridan P, Borrás AM, Furr MC, Khrapko K, Lawitts J, Misra RP, Wei JY. (2003). Early postnatal cardiac changes and premature death in transgenic mice overexpressing a mutant form of serum response factor. *J Biol Chem* **276**:40033-40040.

Zhao Y, Meng XM, Wei YJ, Zhao XW, Liu DQ, Cao HQ, Liew CC, Ding JF. (2003). Cloning and characterisation of a novel cardiac-specific kinase that interacts specifically with cardiac troponin I. *J Mol Med* **81**:297-304.

Cardiovascular Topics

Progressive familial heart block type II (PFHBII): a clinical profile from 1977 to 2003

P. FERNANDEZ, V.A. CORFIELD, P.A. BRINK

Summary

An evaluation of a 38-year-old Caucasian woman, who was referred to Tygerberg Hospital (Western Cape Province, RSA) with Wenckebach second-degree or possibly complete atrioventricular (AV) block that had progressed from first-degree AV block, identified a family history of the cardiac conduction system disorder progressive familial heart block type II (PFHBII). This prompted a retrospective clinical review of the subjects described in the original study, as well as additional family members who had not been examined in the original study.¹ Progression of clinical features was observed, but more importantly, PFHBII was clinically redefined as an AV nodal disorder, which may progress to dilated cardiomyopathy (DCM).

Cardiovasc J South Afr 2004; 15: 129–132.

www.cvjsa.co.za

In January 2003, a 38-year-old Caucasian woman with Wenckebach second-degree or possibly complete atrioventricular (AV) block and a heart rate of 43 beats per minute (bpm) (Fig. 1) (Table I, individual IV:20) was admitted to the Cardiology Unit of Tygerberg Hospital (Western Cape Province). The subject had been examined previously (P.A. Brink, May 2000) and diagnosed with first-degree AV block with a P–R interval of 240 milliseconds (Fig. 2). Genealogy studies indicated that the subject was the daughter of the index case of the previously described South African family with progressive familial heart block type II (PFHBII).¹

In 1977, Brink and Torrington characterised PFHBII as an inherited autosomal dominant cardiac conduction system

disorder, which segregated in a South African Caucasian Afrikaner family from the Eastern Cape Province.¹ The ECG features of PFHBII were at that time defined by isolated sinus bradycardia (SB), isolated left posterior hemiblock (LPHB) or complete heart block (CHB) with narrow QRS complexes. In this report, we describe a retrospective study of the family members examined by Brink and Torrington in 1977. Clinical data are also presented of additional subjects from the same family who were not examined in the original study. The results of the study reiterated the progressive nature of the disorder and prompted a redefinition of the clinical profile of PFHBII.

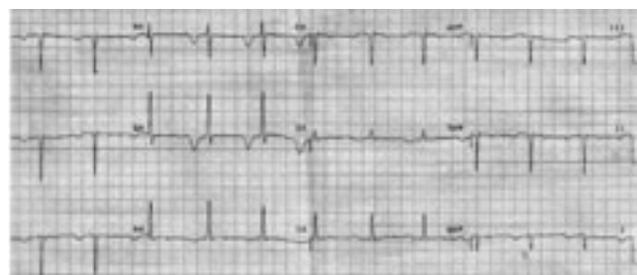


Fig. 1. Twelve-lead ECG strip for individual IV:20, taken in January 2003, showing a slow, irregular sinus rhythm with no consistent ventricular response. The AV nodal delay could represent a Wenckebach phenomenon or alternately, complete AV dissociation. The subject subsequently had a pacemaker implanted.

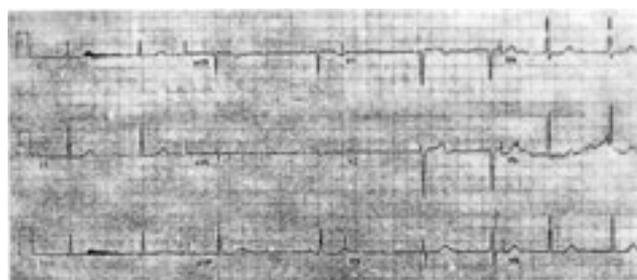


Fig. 2. Twelve-lead ECG strip for individual IV:20, showing first-degree AV block (P–R interval = 240 milliseconds). The ECG was taken in May 2000.

US/MRC Centre for Molecular and Cellular Biology, Faculty of Health Sciences, University of Stellenbosch, Tygerberg

P. FERNANDEZ, B.Sc. Hons.
V.A. CORFIELD, Ph.D.

Department of Internal Medicine, Faculty of Health Sciences, University of Stellenbosch and Tygerberg Hospital, Tygerberg

P.A. BRINK, M.D., Ph.D.

TABLE I. EVALUATION OF FOLLOW-UP CLINICAL DATA OF THE PFHBII FAMILY MEMBERS CLINICALLY ASSESSED IN THE ORIGINAL STUDY

<i>Individual</i>	<i>^aAge (y)</i>	<i>ECG</i>	<i>EF</i>	<i>LVEDD</i>	<i>Comment</i>
III:6	68	normal	76	4.8	unaffected
III:8	42	3° AVB	60	4.5	^bSB = 48 bpm, CHB (PM 42 y); present age = 69 y
III:10	64	normal	65	4.4	unaffected
III:12	37	3° AVB	41	6.3	CHB (PM 37 y), DCM, ^c59 y
IV:9	41	3° AVB	42	5.7	CHB (PM 41 y), DCM, ^c46 y
IV:11	35	3° AVB	nd	nd	CHB (PM 35 y); present age = 36 y
IV:14	32	normal	65	4.3	unaffected
^c IV:16	40	normal	64	4.5	slow R-wave progression; transition to dominant R-wave in V4 – otherwise unaffected
IV:18	33	normal	70	4.7	unaffected
IV:19	31	normal	63	4.9	unaffected
IV:20	38	2° or 3° AVB	61	5.5	1° AVB (P–R interval = 240 m/s, May 2000); SB = 43 bpm, Wenckebach/complete AVB (PM 38 y), January 2003; present age = 38 y
IV:21	41	normal	64	4.4	unaffected
IV:24	39	normal	60	4.5	unaffected
IV:25	29	3° AVB	40	5.8	CHB (PM 29 y), DCM, ^c43 y
IV:26	15	nd	nd	nd	heart transplant, DCM, ^c15 y

Not shown in the table are individuals II:3, III:7, III:9, III:11; IV:12, IV:13, IV:15, IV:17 and IV:22, who were designated clinically unaffected based on family history or a family physician's medical report. No further data could be obtained for these individuals.

^aAge at examination in years (y); ^bHeart rate prior to pacemaker (PM) implant; ^cIndividuals with atypical accompanying features; ^dAge at death

1° = first-degree; 2° = second-degree; 3° = third-degree; AVB = atrioventricular block; bpm = beats per minute; CHB = complete heart block; DCM = dilated cardiomyopathy; EF = ejection fraction; LVEDD = left ventricular end diastolic diameter; m/s = milliseconds; nd = no data available; (PM y) = PM age at implant; SB = sinus bradycardia

Clinically affected individuals are indicated in bold

Methods

Ethical approval

The present study formed part of a project approved by the Ethics Committee of the Faculty of Health Sciences at the University of Stellenbosch.

Genealogy and extent of the disease

The South African Caucasian Afrikaner family presented in this study was previously described.¹ Extensive genealogy studies were performed to determine the segregation of PFHBII in other branches of the family.

Patient panel

Twelve-lead ECG and two-dimensional echocardiographic follow-up examinations were acquired (1977 to present) from Tygerberg Hospital or from the personal physicians of subjects described in the original PFHBII study.¹ Clinical data were also obtained for the children (born after 1977) of these subjects, as well as other available members who were not examined in the original study.¹ Where the records were unavailable, if possible, clinical histories were obtained from the subjects themselves or close relatives. Clinical histories of deceased individuals were also included in the study.

Electrocardiographic and echocardiographic criteria

The ECG diagnostic criteria for PFHBII were previously

established.¹ In addition, the study included minimum criteria for an echocardiographic diagnosis of dilated cardiomyopathy (DCM), defined by an ejection fraction of less than 45% and a left ventricular end-diastolic diameter of greater than 5.6 cm in the absence of hypertension and valvular heart disease or a history consistent with ischaemic heart disease. All measurements were made according to the American Society of Echocardiography guidelines.²

Results

Genealogy and extent of the disease

The study by Brink and Torrington in 1977 obtained information on 140 members of a South African family, of which 24 members of one family branch were examined by ECG. We performed extensive family studies questioning living first-, second- and third-degree relatives of individual II:2 (Fig. 3) (not including his progeny). Wherever possible, ECGs were performed, and occasionally echocardiographic assessment. We did not detect disease with features of PFHBII, although one third-degree relative had hypertrophic cardiomyopathy. Additionally, we questioned relatives of individual II:3 (Fig. 3) and did not identify evidence of the disease in her lineage. We concluded that individual II:2 was the true carrier of PFHBII (Fig. 3), as his children reported him to have had a heart block. We do not therefore have evidence of segregation of the disease in ancestors or siblings and their progeny of II:2 and II:3, limiting known disease to the current pedigree (Fig. 3).

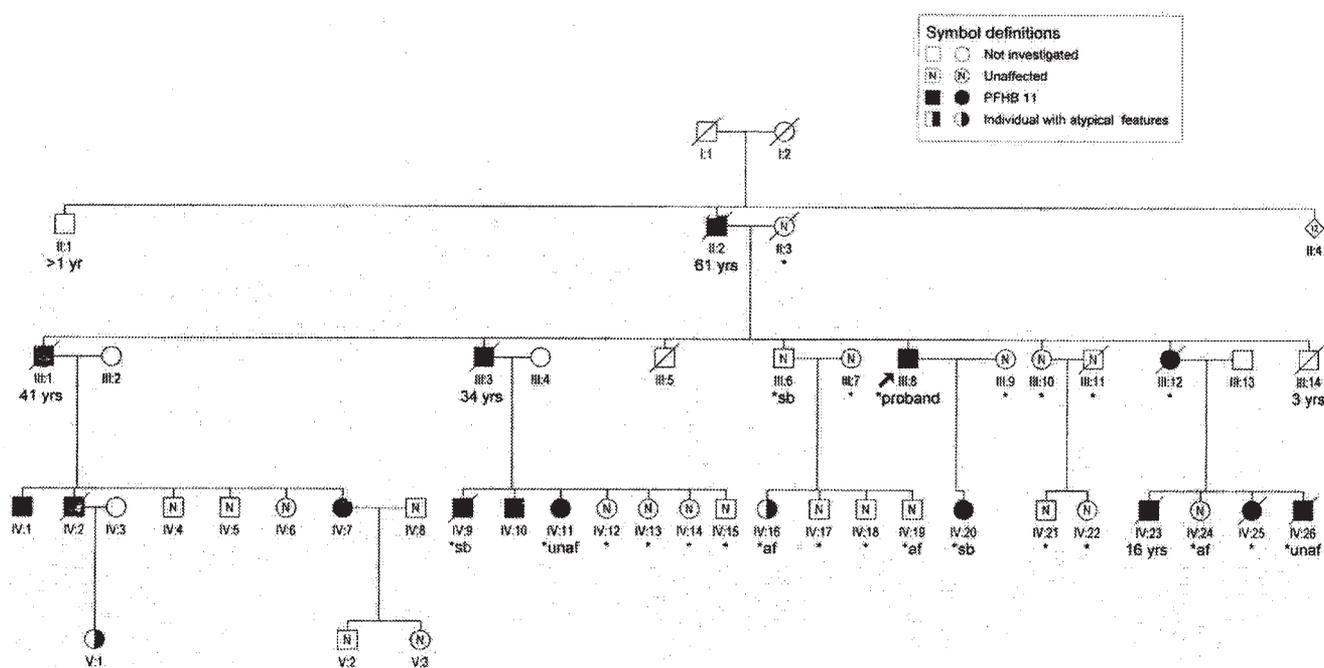


Fig. 3. A family tree showing all kindred of an individual with PFHBII. Squares represent males and circles represent females. An asterisk indicates the individuals examined in the original study. The clinical assessments of subjects whose present diagnoses differ from the original study¹ are shown below the respective symbols (af, subject previously diagnosed as affected; sb, subject previously with sinus bradycardia; unaf, subject previously unaffected).

Clinical data

Twenty-four subjects were assessed in the original study,¹ of whom six were designated clinically affected, and three were identified with SB. Follow-up clinical data or family histories were obtained for the 24 subjects described in the original study¹ (Table I). Our data indicated that two of the three subjects previously identified with SB¹ (Fig. 3, individuals IV:9 and IV:20) subsequently developed AV block and had pacemakers implanted (Table I). Furthermore, three subjects (Fig. 3, individuals III:12, IV:9 and IV:25) showed progression to DCM (Table I). A fourth subject (Fig. 3, individual IV:26), who had previously shown no conduction abnormalities,¹ required a heart transplant as a result of juvenile-onset DCM (Table I). Our follow-up data also differed from the previous diagnoses of three family members. These subjects (Fig. 3, individuals IV:16, IV:19 and IV:24), who had previously been designated clinically affected, presently did not meet our strict diagnostic criteria, although individual IV:16 did present atypical associated features on examination (Table I).

In addition, we obtained clinical or family information for 15 subjects who were not clinically examined in the original study or who were born after 1977 (Fig. 3). Eight of the 15 subjects were assigned a clinically affected status (Table II). Seven of the eight clinically affected individuals developed conduction defects that ranged from atrial fibrillation and CHB to left anterior hemiblock (LAH) (Table II). The eighth affected subject (Table II, individual IV:2) showed progression from AV block to DCM. An additional family member (Table II, individual V:1) presented with atypical features that included severe chest pain with frequent dyspnoea and syncopal episodes. However, the

subject has declined to be examined and her clinical status remains uncertain.

Discussion

Evaluation of clinical and family data spanning 26 years confirmed the progression of features in PFHBII. More importantly, the study highlighted the progression to DCM, which was not apparent in the original study. The follow-up study showed that SB preceding CHB occurred frequently in the family, although it must be noted that SB is fairly common in the general population, particularly among athletes.³ However, upon questioning, most of the affected family members indicated that they had fairly sedentary lifestyles and SB was absent in unaffected subjects with similar living patterns. Therefore, in this family, a slow heart rate of less than 50 bpm is considered familial and a diagnostic criterion of PFHBII.

Generally, individuals with first-degree AV block are asymptomatic and diagnosis thereof is usually incidental. The present study also aimed to identify clinical or physical characteristics occurring in the family that could be indicative of underlying disease. Two subjects with potentially early clinical markers for PFHBII were identified. The first subject (individual IV:16), who was examined in the original study and re-evaluated by us in 2000, was, unlike the previous study, assigned a clinically unaffected status. Our evaluation did indicate a slow R-wave progression, where the transition to a dominant R-wave occurs in V4. A slow R-wave progression is often associated with DCM,⁴ although lead placement could also play a role in the ECG assessment. However, we speculate that in this family, the atypical feature could be an early indication of impending

TABLE II. EVALUATION OF CLINICAL DATA OF PFHBII FAMILY MEMBERS WHO WERE NOT CLINICALLY ASSESSED IN THE ORIGINAL STUDY

<i>Individual</i>	<i>Age (y)</i>	<i>ECG</i>	<i>EF</i>	<i>LVEDD</i>	<i>Comment</i>
II:2	61	nd	nd	nd	HB (type unknown), ^a61 y
III:1	40	nd	nd	nd	HR < 50 bpm, CHB (PM 41 y), ^a41 y
III:3	45	nd	nd	nd	CHB (PM 45 y), failed PM, ^a45 y
IV:1	36	LAD	64	4.4	LAD (-30° axis), LAH; present age = 41 y
IV:2	38	3° AVB	38	6.2	^bSB = 50 bpm, AF, CHB (PM 39 y), DCM, ^a43 y
IV:4	49	normal	60	4.4	unaffected
IV:5	45	normal	60	4.4	unaffected
IV:6	47	normal	65	4.5	unaffected
IV:7	33	3° AVB	52	5.1	1° AVB (P-R interval = 220 m/s, March 1990), ^cSB = 38 bpm, AF, CHB (PM 33 y, October 1996); present age = 38 y
IV:8	45	normal	65	4.8	unaffected
IV:10	36	SB	nd	nd	SB = 43 bpm; present age = 39 y
IV:23	16	3° AVB	nd	nd	CHB (PM 16y), ^a16 y
^c V:1	23	nd	nd	nd	chest pain, dyspnoea, syncope
V:3	22	160	63	5.2	unaffected
V:4	20	160	70	4.3	unaffected

^aAge at examination in years (y); ^bHeart rate prior to pacemaker (PM) implant; ^cIndividuals with atypical accompanying features; ^dAge at death

1° = first-degree; 3° = third-degree; AF = atrial fibrillation; AVB = atrioventricular block; bpm = beats per minute; CHB = complete heart block; DCM = dilated cardiomyopathy; LVEDD = left ventricular end diastolic diameter; HB = heart block; HR = heart rate; LAH = left anterior hemiblock; LAD = left axis deviation; EF = ejection fraction; m/s = milliseconds; nd = no data available; (PM y) = PM age at implant; SB = sinus bradycardia

Clinically affected individuals are indicated in bold

cardiac abnormalities. The second subject (individual V:1) had reported severe chest pains, shortness of breath and fainting episodes. Recent correspondence with individual V:1 indicated a persistence of features, despite treatment with Atenolol (P.A. Brink, personal communication). Interestingly, there is a history of sudden death at a relatively young age in the subject's family, with both her grandfather and father (Table II, individuals III:1 and IV:2) having died aged 41 and 43 years, respectively.

The importance of this retrospective study is exemplified by our redefinition of the clinical profile of PFHBII. Barring the one individual with LAH (individual IV:1), we propose that the absence of ventricular conduction delay makes it highly likely that the conduction block occurs close to the origins of the Bundle of His. This suggests that PFHBII is not a disease of the ventricular conduction system. Drawing on observations from ECG data, we have amended the diagnostic criteria and characterise PFHBII as an AV nodal disorder with clinical onset between the fourth and sixth decade. In addition, four family members were identified who progressed from conduction block to DCM (Tables I and II, individuals III:12, IV:2, IV:9 and IV:25). For these subjects, the progression from complete AV block to congestive heart failure (CHF) ranged from five years (individuals IV:2 and IV:9) to 22 and 14 years (individual III:12 and IV:25), respectively. Furthermore, a fifth subject (Table I, individual IV:26) with no prior conduction defects developed DCM in early adolescence. Consequently, these data indicate that in this family, on average, death as a result of CHF occurred before the sixth decade. Unfortunately, the available data did not permit assessing whether the prognosis for individuals who only developed conduction defects was better than for those subjects with DCM. We cannot explain why particular individuals advanced

from AV block to DCM, while others only developed conduction defects, but this pattern of clinical features was consistent with other described disorders.^{5,6} Consequently, we performed a genetic linkage study and excluded the lamin A/C⁵ and CMD1H⁶ loci as cause of PFHBII (data not shown).

Conclusions

The data presented emphasise the importance of establishing a family history when making patient diagnoses. The progressive nature of the disorder was reiterated, but more importantly, a familial DCM component was demonstrated in PFHBII. Consequently, the study provides a more comprehensive clinical profile of PFHBII, which may assist clinicians to identify other South African families or individuals with this potentially life-threatening disorder.

References

1. Brink AJ, Torrington M. Progressive familial heart block – two types. *S Afr Med J* 1977; **52**: 53–59.
2. Sahn DJ, DeMaria A, Kisslo J, Weyman A. Recommendations regarding quantitation in M-mode echocardiography: results of a survey of echocardiographic measurements. *Circulation* 1978; **58**: 1072–1083.
3. Josephson ME, Marchlinski FE, Buxton AE. The bradyarrhythmias. In: Wilson JD, Braunwald E, Isselbacher KJ, Peteresdorf RG, Martin JB, Fauci AS, Root RK, eds. *Principles of Internal Medicine*. New York: McGraw-Hill, 1991: 902–908.
4. Fatkin D and Graham RM. Molecular mechanisms of inherited cardiomyopathies. *Physiol Rev* 2002; **82**: 945–980.
5. Fatkin D, MacRae C, Sasaki T, et al. Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease. *N Engl J Med* 1999; **341**: 1715–1724.
6. Jung M, Poepping I, Perrot A, et al. Investigation of a family with autosomal dominant dilated cardiomyopathy defines a novel locus on chromosome 2q14-q22. *Am J Hum Genet* 1999; **65**: 1068–1077.