UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Genetic heterogeneity in South African facioscapulohumeral muscular dystrophy (FSHD) families

BY

ANNELIZE VAN DER MERWE

Thesis submitted in partial fulfilment of the requirements for the Magister Scientia (M.Sc.) degree in Human Genetics

Department of Human Genetics and Developmental Biology
Faculty of Medicine, University of Pretoria,
South Africa

SUPERVISOR : Prof. ANTONEL OLCKERS
Centre for Genome Research,
Potchefstroom University for Christian Higher Education, South Africa
(Formerly from the Department of Human Genetics and Developmental Biology,
Faculty of Medicine, University of Pretoria, South Africa)

CO-SUPERVISOR : Dr. CLARA SCHUTTE`
Department of Neurology, Faculty of Medicine,
University of Pretoria, South Africa

February 2002

© University of Pretoria

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

TO MY PARENTS

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

ABSTRACT

FSHD is the third most common inherited disorder of muscle after Duchenne and Myotonic dystrophy. On a clinical level FSHD is characterised by progressive weakening and muscle atrophy. Initially the face, shoulder-girdle and upper arm muscles are affected, and other skeletal muscles may also become involved over time. Clinical diagnosis of the FSHD phenotype is complex due to the extreme variability and penetrance. The expression varies in the severity, rate of progression and the age of onset.

The FSHD phenotype segregates as an autosomal dominant trait. Linkage was established in 1990 to the sub-telomeric region of chromosome 4q35. This sub-telomeric region was observed to contain 3.3 kb tandem repeats. In 1993 it was concluded that FSHD is caused by a deletion of an integral number of these 3.3 kb repeats. Probe p13E-11 is utilised to detect the FSHD associated DNA rearrangements, but was observed to cross-hybridise to chromosome 10q26 as well as the Y-chromosome. Restriction mapping of chromosome 10q26 indicated that this region contains similar 3.3 kb repeat units to those on chromosome 4q35. Comparison of sequences between 4q35 and 10q26 fragments indicated the presence of an unique chromosome 10 specific *Bln* I site, allowing discrimination between chromosome 4 and 10 alleles. Translocation events between the repeats on chromosomes 4 and 10 were demonstrated via the presence of *Bln* I sites within the 3.3 kb repeats on chromosome 10q26. The interchromosomal exchanges have implications for the specificity and sensitivity of diagnostic DNA testing of FSHD.

This was the first study to investigate the molecular aetiology of FSHD in the South African population. Five extended FSHD families (F10, F20, F30, F40 and F60), consisting of 100 individuals, from the South African population were selected for this study. Haplotype analyses were performed to study the segregation of nine short tandem repeat polymorphism (STRP) markers, mapped for the first time here, in the 4q35 region with the FSHD phenotype. The FSHD associated DNA rearrangements at the D4Z4 locus were detected via Southern blot analyses utilising probe p13E-11.

Results of this study confirmed the presence of DNA rearrangements in the five FSHD families investigated. No evidence for genetic heterogeneity was therefore observed in the selected population group. Identical FSHD-associated haplotypes were observed in three

families (F10, F30 and F40), co-segregating with a *Bln* I resistant deletion fragment of 24 kb. However, an FSHD-associated haplotype, different than the one observed in F10, F30 and F40, was observed in the two other families (F20 and F60). A *Bln* I resistant deletion fragment of 34 kb co-segregated with this FSHD-associated haplotype in these families. Individual 15-31 was found to be a compound heterozygote with regard to the two FSHD-associated genotypes that segregated in the South African families investigated. This individual from family F10 therefore represents the link between the two groups of South African families.

Two distinct haplotypes were observed, each co-segregating with a specific *Bln* I resistant deletion fragment. A dual Founder Effect was therefore observed in this unique population and excludes the possibility of a single ancestral mutation in the South African FSHD population.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

OPSOMMING

Fasioskapulohumerale spierdistrofie (FSHD) is die derde algemeenste oorerflike spiertoestand na Duchenne en Miotoniese distrofie. FSHD word geklassifiseer op 'n kliniese vlak deur progressiewe verswakking en spieratrofie. Die gesig, skouergordel en bo-arm spiere word gewoonlik eerste aangetas, maar ander skeletale spiere kan ook mettertyd aangetas word. Die kliniese diagnose van die FSHD fenotipe is kompleks as gevolg van ekstreme variasie en penetrasie. Die ekspressie varieer in die graad van aantasting, tempo van progressie en ouderdom van presentering.

Die FSHD fenotipe segregeer as 'n outosomale dominante toestand. Koppeling is gevind in die sub-telomeriese gebied van chromosoom 4q35 in 1990 en bevat 3.3 kb direk-herhalende volgordes. Dit is in 1993 gepostuleer dat FSHD veroorsaak word deur 'n delesie van 'n aantal van hierdie 3.3 kb herhalings. Peiler p13E-11 is gebruik om die FSHD-geassosieerde DNA herrangskikkings te herken, maar is waargeneem om ook te hibridiseer met chromosoom 10q26 so wel as met die Y-chromosoom. Restriksiekartering van chromosoom 10q26 het aangedui dat hierdie gebied dieselfde 3.3 kb herhalingseenhede bevat as chromosoom 4q35. Vergelyking tussen die DNA volgorde van 4q35 en 10q26 fragmente het die teenwoordigheid van 'n unieke chromosoom 10 spesifieke *Bln* I setel aangedui wat dit moontlik maak om tussen die allele van chromosoom 4 en 10 te kan onderskei. Die *Bln* I setels in elke 3.3 kb herhaling van 10q26 het die teenwoordigheid van translokasiegebeurtenisse tussen die herhalings op chromosoom 4 en 10 aangedui. Die interchromosomale uitruilings het implikasies vir die spesifisiteit en sensitiwiteit van diagnostiese DNA toetsing vir FSHD.

Hierdie was die eerste studie om die molekulêre etiologie van FSHD in die Suid-Afrikaanse populasie te ondersoek. Vyf uitgebreide FSHD families (F10, F20, F30, F40 en F60), bestaande uit 100 individue, van die Suid-Afrikaanse populasie is geselekteer. Haplotipe analise is onderneem om die segregasie van nege direk-herhalende polimorfiese merkers (STRP), wat ook gekarteer is in hierdie studie, in die 4q35 gebied met die FSHD fenotipe te bestudeer. Die FSHD-geassosieerde DNA herrangskikkings by die D4Z4 lokus is waargeneem met Southern klad analise deur gebruik te maak van peiler p13E-11.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Resultate van hierdie studie het die teenwoordigheid van DNA herrangskikkings in die vyf FSHD families bevestig. Geen bewys vir genetiese heterogeniteit is waargeneem in hierdie geselekteerde populasie nie. Identiese FSHD-geassosieerde haplotipes is waargeneem in drie families (F10, F20 en F30) wat saam met 'n 24 kb *Bln* I bestande delesiefragment segregeer. 'n Verskillende FSHD-geassosieerde haplotipe as die wat waargeneem is in F10, F30 en F40 is waargeneem in die ander twee families (F20 en F60). 'n *Bln* I bestande delesiefragment van 34 kb het saam met dié FSHD-geassosieerde haplotipe in hierdie twee families gesegregeer. Individu 15-31 is gevind om 'n saamgestelde heterosigoot ten opsigte van die twee FSHD-geassosieerde genotipes wat in die Suid-Afrikaanse bevolking waargeneem is, te wees. Hierdie individu van familie F10 is dus die skakel tussen die twee groepe van Suid-Afrikaanse families.

Twee kenmerkende haplotipes, wat elkeen met 'n spesifieke *Bln* I bestande delesiefragment segregeer, is waargeneem. 'n Dubbele stigterseffek is waargeneem in hierdie unieke populasie en dit sluit dus enige moontlikheid van 'n enkele voorouerlike mutasie vir die Suid-Afrikaanse FSHD populasie uit.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND SYMBOLS i
LIST OF FIGURES AND GRAPHS viii
LIST OF TABLES x
ACKNOWLEDGEMENTS xi

CHAPTER ONE

INTRODUCTION 1

CHAPTER TWO

THE AETIOLOGY AND PATHOGENESIS OF FACIOSCAPULO-HUMERAL MUSCULAR DYSTROPHY 3

2.0 THE MUSCULAR DYSTROPHIES 3
2.1 FACIOSCAPULOHUMERAL MUSCULAR DYSTROPHY 6
2.1.1 CLINICAL ASPECTS OF FSHD 7
2.1.1.1 Presenting symptoms 7
2.1.1.1.1 The facial muscles 8
2.1.1.1.2 The upper extremities, shoulder girdle and neck muscles 9
2.1.1.1.3 The truncal muscles 11
2.1.1.1.4 The lower extremities and the pelvic girdle muscles 11
2.1.1.1.5 Asymmetry of muscle involvement 11
2.1.1.1.6 Extramuscular involvement 11
2.1.1.2 Clinical heterogeneity in the FSHD phenotype 13
2.1.1.2.1 Infantile FSHD 16
2.1.1.3 Current treatments for FSHD 17
2.1.2 GENETIC ASPECTS OF FSHD 20
2.1.2.1 Linkage of FSHD to chromosome 4q35 20
2.1.2.2 The FSHD locus on chromosome 4q35 22
2.1.2.2.1 Translocation events between chromosomes 4q and 10q 26
2.1.2.2.2 Hybrid repeat arrays and deletion of p13E-11 hybridisation site 28
2.1.2.2.3 The *Bgl* II – *Bln* I dosage test 30
2.1.2.3 Somatic and germline mosaicism 31
2.1.2.4 Anticipation 33
2.1.2.5 Female and male transmission effects 34
2.1.2.6 Sporadic FSHD 34
2.1.2.7 Phenotypic-Genotypic correlation 35
2.1.2.8 Prenatal diagnosis 37
2.1.2.9 Genetic heterogeneity 38
2.1.2.10 Candidate genes 39
2.1.2.10.1 Actinin-associated LIM protein gene (ALP) 39
2.1.2.10.2 Adenine nucleotide translocator gene (ANT) 39

2.1.2.10.3 Double homeobox gene 4 (DUX4) 40
2.1.2.10.4 Fibroblast growth factor (FGF) and FGF receptors 41
2.1.2.10.5 The FSHD region gene 1 (FGR1) 41
2.1.2.10.6 The FSHD region gene 2 (FRG2) 42
2.1.2.10.7 Human beta-tubulin gene (TUBB4Q) 43
2.1.2.11 Molecular models proposed for the aetiology of FSHD 43
2.1.2.11.1 Homeodomain 43
2.1.2.11.2 Position effect variegation 44
2.1.2.12 Objectives of this study 46

CHAPTER THREE

MATERIALS AND METHODS 47

3.1 FSHD FAMILIES 47
3.1.1 Family F10 48
3.1.2 Family F20 49
3.1.3 Family F30 50
3.1.4 Family F40 51
3.1.5 Family F60 51
3.2 DNA ISOLATION 52
3.2.1 Isolation of genomic DNA using sodium perchlorate and chloroform 52
3.2.2 Isolation of genomic DNA utilising the Wizard® Genomic kit 53
3.3 HAPLOTYPE ANALYSIS 55
3.3.1 STRP marker information 56
3.3.1.1 Marker UT1366 at locus D4S1523 57
3.3.1.2 Marker GATA5B02 at locus D4S1652 57
3.3.1.3 Marker AFMa224xh1 at locus D4S2930 58
3.3.1.4 Marker ATA22F02 at locus D4S2390 59
3.3.1.5 Marker UT5785 at locus D4S2299 60
3.3.1.6 Marker UT2219 at locus D4S2283 60
3.3.1.7 Marker UT7694 at locus D4S2688 61
3.3.1.8 Marker AFMa190zf5 at locus D4S2921 62
3.3.1.9 Marker AFM238ve3 at locus D4S426 63
3.3.2 5' - End labelling of PCR primers 64
3.3.3 The polymerase chain reaction (PCR) 64
3.3.4 Multiplex PCR 65
3.3.5 Single stranded DNA sequencing 65
3.3.6 Denaturing gel electrophoresis and autoradiography 66
3.4 SOUTHERN BLOT ANALYSIS 67
3.4.1 Restriction fragment length polymorphism (RFLP) analysis 67
3.4.2 Agarose gel electrophoresis 67
3.4.3 Genomic DNA Transfer 68
3.4.4 Isolation of p13E-11 68
3.4.5 Radio-active detection 70
3.4.5.1 Labelling of probe p13E-11 and molecular weight markers 71
3.4.5.2 Hybridisation conditions 71
3.4.5.3 Wash conditions 71
3.4.5.4 Autoradiography 72
3.4.6 Non radio-active detection 72

3.4.6.1 Labelling of probe p13E-11 72
3.4.6.2 Quantification of labelled probe 72
3.4.6.3 Hybridisation conditions 74
3.4.6.4 Wash conditions 74
3.4.6.5 Detection 74

CHAPTER FOUR

RESULTS AND DISCUSSION 75

4.1 SHORT TANDEM REPEAT POLYMORPHISM (STRP) ANALYSIS 75
4.1.1 Short tandem repeat marker UT1366 at locus D4S1523 76
4.1.2 Short tandem repeat marker GATA5B02 at locus D4S1652 77
4.1.3 Short tandem repeat marker AFMa224xh1 at locus D4S2930 78
4.1.4 Short tandem repeat marker ATA22F02 at locus D4S2390 79
4.1.5 Short tandem repeat marker UT5785 at locus D4S2299 80
4.1.6 Short tandem repeat marker UT2219 at locus D4S2283 81
4.1.7 Short tandem repeat marker UT7694 at locus D4S2688 82
4.1.8 Short tandem repeat marker AFMa190zf5 at locus D4S2921 83
4.1.9 Short tandem repeat marker AFM238ve3 at locus D4S426 85
4.2 SOUTHERN BLOT ANALYSIS 86
4.3 RADIO-ACTIVE VERSUS NON RADIO-ACTIVE DETECTION 89
4.4 MOLECULAR ANALYSIS OF SPECIFIC FSHD FAMILIES 91
4.4.1 FSHD family F10 91
4.4.2 FSHD family F20 95
4.4.3 FSHD family F30 99
4.4.4 FSHD family F40 102
4.4.5 FSHD family F60 103

CHAPTER FIVE

CONCLUSION 107

5.1 SOUTH AFRICAN FSHD FAMILIES INVESTIGATED 107
5.2 GENETIC HETEROGENEITY 110
5.3 DUAL FOUNDER EFFECT 110
5.4 FUTURE DIRECTIONS IN FSHD RESEARCH 111

REFERENCES 113

APPENDIX A

CONFERENCES AND MEETINGS AT WHICH RESEARCH CONTAINED IN THIS THESIS WERE PRESENTED 122

A.1 RESEARCH PRESENTED AT INTERNATIONAL CONFERENCES 122
A.2 RESEARCH PRESENTED AT NATIONAL CONFERENCES 123

A.3 RESEARCH PRESENTED AT THE FACULTY OF MEDICINE, UNIVERSITY OF PRETORIA 124
A.4 PUBLISHED ABSTRACTS IN INTERNATIONAL PEER-REVIEWED JOURNALS 124

APPENDIX B

MODE OF INHERITANCE OF DIFFERENT TYPES OF MUSCULAR DYSTROPHIES 125

APPENDIX C

DIAGNOSTIC CRITERIA FOR FACIOSCAPULOHUMERAL MUSCULAR DYSTROPHY 128

APPENDIX D

NUCLEOTIDE SEQUENCE OF REPEAT UNITS AND FLANKING REGIONS AT THE D4Z4 LOCUS 132

APPENDIX E

COMPARISON OF NUCLEOTIDE SEQUENCE FROM ONE *Kpn* I REPEAT UNIT DERIVED FROM CHROMOSOMES 4q35 AND 10q26 139

APPENDIX F

EXTENDED FAMILY PEDIGREES 144

LIST OF ABBREVIATIONS AND SYMBOLS

Abbreviations are listed in alphabetical order.

9B6A	probe complimentary to the homeobox sequences within each 3.3 kb repeat unit
10qter	telomeric region of the long arm of chromosome 10
α	alpha
α^{32}P-dCTP	dCTP labelled in the α position with ^{32}P isotope
A or a	adenine (in DNA sequence)
A_{260}/A_{280}	ratio of absorbency measured at 260 nm and 280 nm
ACD	acid citrate dextrose
acrylamide	C_3H_5NO
ACTA	actin alpha skeletal muscle
AD	autosomal dominant
ADP	adenosine diphosphate
ALP	actinin-associated LIM protein
ANT	adenine nucleotide translocator
ANT1	adenine nucleotide translocator isoform 1
ANT2	adenine nucleotide translocator isoform 2
ANT3	adenine nucleotide translocator isoform 3
APS	ammonium persulfate: $(NH_4)S_2O_8$
AP-SA	alkaline phosphatase-labelled streptavidin
AR	autosomal recessive
ATP	adenosine triphosphate
β	beta
Bam HI	restriction endonuclease isolated from an *E. coli* strain that carries the cloned *Bam* HI gene from Bacillus amyloliquefaciens H, with recognition site 5'-G↓GATCC-3'
Bgl II	restriction endonuclease isolated from an *E. coli* strain that carries the cloned *Bgl* II gene from *Bacillus globigii*, with recognition site 5'-A↓GATCT-3'
bisacrylamide	N,N'-methylene-bis-acrylamide:$C_7H_{10}O_2N_2$
Bln I	restriction endonuclease isolated from from an *E. coli* strain that carries the cloned *Bln* I gene from *Brevibacterium linens*, with recognition site 5'-C↓CTAGG-3'
BMD	Becker muscular dystrophy
boric acid	boracic acid: H_3BO_3
bp	base pair
BPB	bromophenol blue (3',3",5',5"-tetrabromophenolsulfonephthalein): $C_{14}H_{10}BrO_5S$
BSA	bovine serum albumin
C or c	cytosine (in DNA sequence)
°C	degrees centigrade
%C	percentage crosslinking monomer
ca.	circa: approximately
CAPN3	calpain-3
CAV3	caveolin-3
CCD	central core disease
cDNA	complementary DNA
CEB8	probe complementary to locus D4F35S1
CEN	centromere

CEPH	Centre dÉtude du Polymorphisme Humain (Centre for the Study of Human Polymorphisms)
CHLC	Co-operative Human Linkage Centre
chr	chromosome
CI	cardiac involvement
Ci	curie: quantity of any radioactive nuclide in which there are 3.7 x 10^{10} disintegrations per second
CK	creatine kinase
cm	centimeter: 10^{-2} meter
cM	centimorgan
CMD1B	congenital muscular dystrophy with secondary merosin deficiency
CNS	central nervous system
COL6A1	collagen type VI subunit α1
COL6A2	collagen type VI subunit α2
COL6A3	collagen type VI subunit α3
CPD-Star®[1]	disodium 4-chloro-3-(methoxyspiro {1,2-dioxetane-3,2'-(5'-chloro) tricyclodecan}-4-yl)phenyl phosphate: $C_{18}H_{19}Cl_2O_7PNa_2$
CS	clinical severity
CT-scan	computed tomography scan
δ	delta
dATP	2'-deoxyadenosine-5'-triphosphate
DBP	detector block powder
dCTP	2'-deoxycytidine-5'-triphosphate
ddATP	2',3'-dideoxyadenosine-5'-triphosphate
ddCTP	2',3'-dideoxycytidine-5'-triphosphate
ddGTP	2',3'-dideoxyguanosine-5'-triphosphate
ddH_20	double distilled water
ddNTP	2',3'-dideoxynucleotide
ddTTP	2',3'-dideoxythymidine-5'-triphosphate
DEPC	diethyl pyrocarbonate
DES	desmin
DGC	dystrophin-glycoprotein complex
dGTP	2'-deoxyguanosine-5'-triphosphate
DMD	Duchenne muscular dystrophy
DMPK	Myotonin-protein kinase gene
DMRV	distal myopathy with rimmed vacuoles
DNA	deoxyribonucleic acid
DNR	dinucleotide repeat
dNTP	2'-deoxynucleotide triphosphate
DRM	desmin related myopathy
dsDNA	double stranded DNA
DTT	dithiothreithol: threo-1,4-dimercapto-2,3-butanediol: $C_4H_{10}O_2S_2$
dTTP	2'-deoxythymidine-5'-triphosphate
DUX1	double homeobox gene 1
DUX2	double homeobox gene 2
DUX3	double homeobox gene 3
DUX4	double homeobox gene 4
ε	epsilon
Eco RI	restriction endonuclease isolated from an *E. coli* strain that carries the cloned *Eco* RI gene from *Escherichia coli* RY 13, with recognition site 5'-G↓AATTC-3'
EDMD	Emery-Dreifuss muscular dystrophy
EDMD-AD	autosomal dominant Emery-Dreifuss muscular dystrophy

[1] CPD-Star® is a registered trademark of Tropix Inc., Bedford, MA, U.S.A.

EDTA	ethylenediamine tetraacetic acid: $C_{10}H_{16}N_2O_8$
EMD	X-linked recessive Emery-Dreifuss muscular dystrophy
EMG	electromyography
EST	expressed sequence tag
EtBr	ethidium bromide (2,7-Diamino-10-ethyl-9-phenyl-phenanthridinium bromide): $C_{21}H_{20}BrN_3$
EtOH	ethanol: CH_3CH_2OH
FCMD	Fukuyama congenital muscular dystrophy
FER-1	dysferlin
FGF	fibroblast growth factor
FGF-R1	fibroblast growth factor receptor 1
FGF-R3	fibroblast growth factor receptor 3
FISH	fluorescence *in situ* hybridisation
formamide	carbamaldehyde: CH_3NO
FRG1	FSHD region gene 1
FRG1P	FRG1 protein
FRG2	FSHD region gene 2
FSHD	facioscapulohumeral muscular dystrophy
γ	gamma
γ^{32}P-dATP	dATP labelled in the γ position with ^{32}P isotope
G	gram
G or g	guanine (in DNA sequence)
GDB	Genome database
gDNA	genomic DNA
Genbank	Genbank® [1] : United States repository of DNA sequence information
Gm	allotype associated with IgG heavy chains
H buffer	high salt buffer [10X buffer contains: 100 mM Tris-HCl (pH 7.5), 100 mM $MgCl_2$, 10 mM Dithiothreitol, 1000 mM NaCl]
H_2O	water
HCl	hydrochloric acid
HET	heterozygosity
Hz	hertz
hhsmp3	human DNA insert showing sperm-specific hypomethylation
HIBM	hereditary inclusion body myopathy
HLA	human leukocyte antigens
Hmix	Xenopus mesoderm induced homeobox
HmprD	*Drosophila* paired
HSPG	heparan sulfate proteoglycan
IAA	isoamyl alcohol
IgG	immunoglobulin G
IQ	intelligence quotient
ITGA7	integrin α7
K buffer	potassium buffer [10X buffer contains: 200 mM Tris-HCl (pH 8.5), 100 mM $MgCl_2$, 10 mM Dithiothreitol, 1000 mM KCl]
K-acetate	potassium acetate: CH_3COOK
kb	kilo (10^3) base pair
KCl	potassium chloride
KPL	Kirkegaard & Perry Laboratories
Kpn I	restriction endonuclease isolated from an *E. coli* strain that carries the cloned *Kpn* I gene from *Klebsiella pneumoniae*, with recognition site 5'-GGTAC↓C-3'
L buffer	low salt buffer [10X buffer contains: 100 mM Tris-HCl (pH 7.5), 100 mM $MgCl_2$, 10 mM Dithiothreitol]

[1] Genbank® is a registered trademark of the National Institutes of Health, Bethesda, MD, U.S.A.

LAMA2	laminin α2 chain of merosin
LGMD	limb-girdle muscular dystrophy
LGMD1A	limb-girdle muscular dystrophy type 1A
LGMD1B	limb-girdle muscular dystrophy type 1B
LGMD1C	limb-girdle muscular dystrophy type 1C
LGMD1D	limb-girdle muscular dystrophy type 1D
LGMD1E	limb-girdle muscular dystrophy type 1E
LGMD2A	limb-girdle muscular dystrophy type 2A
LGMD2B	limb-girdle muscular dystrophy type 2B
LGMD2C	limb-girdle muscular dystrophy type 2C
LGMD2D	limb-girdle muscular dystrophy type 2D
LGMD2E	limb-girdle muscular dystrophy type 2E
LGMD2F	limb-girdle muscular dystrophy type 2F
LGMD2G	limb-girdle muscular dystrophy type 2G
LGMD2H	limb-girdle muscular dystrophy type 2H
LGMD2I	limb-girdle muscular dystrophy type 2I
LIM	Lin-11/Isl-1/Mec-3
LINE	long interspersed nuclear elements
LMNA	lamin A/C (gene encoding two components of the nuclear lamina, lamins A and C)
LOD	logarithm of the odds
Lsau	long *Sau* 3A DNA repeats
μ	micro: 10^{-6}
μCi	micro Curie
μg	microgram
μl	microlitre
μM	micromolar
m	milli: 10^{-3}
M	molar: moles per litre
M buffer	medium salt buffer [10X buffer contains: 100 mM Tris-HCl (pH 7.5), 100 mM $MgCl_2$, 10 mM Dithiothreitol, 500 mM NaCl]
M13mp18	vector number 18 of the mp series of bacteriophage M13
MD	muscular dystrophy
MD-EBS	epidermolysis bullosa simplex associated with late-onset muscular dystrophy
MEAX	myopathy with excessive autophagy
MFD	Marshfield Research Foundation
mg	milligram
mg	magnesium
Mg^{2+}	magnesium ion
Mg-acetate	magnesium acetate: $C_4H_6O_4Mg.4H_2O$
$MgCl_2$	magnesium chloride
Mhox	muscle specific homeodomain protein
MIM	mendelian inheritance in man
min	minutes
ml	millilitres
mm	millimetre
mM	millimolar
MM	Miyoshi myopathy
MPD1	autosomal dominant distal myopathy
MPRM1	autosomal dominant myopathy with proximal weakness and early respiratory muscle involvement
mRNA	messenger RNA
MTMX	myotubular myopathy
n	nano: 10^{-9}

Na-citrate	citric acid trisodium salt: $C_6H_5Na_3O_7$
NaCl	sodium chloride
Na_2EDTA	disodium EDTA: $C_{10}H_{14}N_2Na_2O_8.2H_20$
NaH_2PO_4	sodium phosphate monobasic
NaOH	sodium hydroxide
ng	nanogram
NCBI	National Center for Biotechnology Information, U.S.A.
NEM	nemaline myopathy
NIH	National Institutes of Health, U.S.A.
nm	nanometer: 10^{-9} meter
nM	nanomolar
NMR	nuclear magnetic resonance
No	number
OD	optical density
OMIM	online mendelian inheritance in man
OPMD	oculopharyngeal muscular dystrophy
orange G	7-hydroxy-8-phenylazo-1,3-naphthalenedisulfonic acid: $C_{16}H_{10}N_2O_7S2Na_2$
ORF	open reading frame
Otx	orthodenticle homeobox gene
%	percent
p	pico: 10^{-12}
^{32}P	phosphorus isotope: maximum β emission energy 1.71 MeV: half-life 14.3 days
p13E-11	probe complimentary to 3.3 kb repeat units at locus D4Z4
PAB	phosphatase assay buffer
PABP2	poly(A) binding protein 2
PAGE	polyacrylamide gel electrophoresis
Pax	paired box gene
Pax 3	paired box gene 3
Pax 6	paired box gene 6
PCR	polymerase chain reaction
PDGF	platelet-derived growth factor
PDGF-Rα	platelet-derived growth factor receptor α
PDGF-Rβ	platelet-derived growth factor receptor β
PDZ	proteins comprising the Postsynaptic density protein, Disc-large tumor suppressor and the Zonula occludens protein
PEG	polyethylene glycol: $HO(C_2H_4O)_nH$
PEV	position effect variegation
PFGE	pulsed field gel electrophoresis
pH	indicates acidity: numerically equal to the negative logarithm of H+ concentration expressed in molarity
pH30	probe complimentary to locus D4S139
pmol	pico mole
prd	paired gene
P/S	ratio of probe to standard
Pu	purine
Py	pyrimidine
q	long arm of a chromosome
qter	telomeric region of the long arm of a chromosome
RFLP	restriction fragment length polymorphism
RMD	rippling muscle disease
RNA	ribonucleic acid
rpm	revolutions per minute
RSA	relative specific activity

RSMD-1	congenital muscular dystrophy with rigid spine
RT-PCR	reverse transcriptase PCR
RYR1	ryanodine receptor
Sac I	restriction endonuclease isolated from an *E. coli* strain that carries the cloned *Sac* I gene from *Streptomyces achromogenes*, with recognition site 5'-GAGCT↓C-3'
Sau 3AI	restriction endonuclease isolated from an *E. coli* strain that carries the cloned *Sau* 3AI gene from *Staphylococcus aureus 3A*, with recognition site 5'-↓GATC-3'
SCK	serum creatine kinase
SDS	sodium dodecyl sulphate: $C_{12}H_{25}NaSO_4$
sec	seconds
Sequenase	Sequenase®[1] Version 2.0 T7 DNA Polymerase
Sequenase buffer	200 mM Tris-HCl (pH7.5); 100 mM $MgCl_2$; 250 mM NaCl
SERCA1	sarcoplasmic reticulum Ca^{2+} ATPase
SGC	sarcoglycan complex
SGCA	α-sarcoglycan
SGCB	β-sarcoglycan
SGCD	δ-sarcoglycan
SGCG	γ-sarcoglycan
SJS	Schwartz-Jampel syndrome
spermidine	N-[3-Aminopropyl]-1,4-butanediamine: $C_7H_{19}N_3$
SSC	saline-sodium-citrate buffer: 0.15 M NaCl, 15mM Na-citrate, pH 7.0
ssDNA	single stranded DNA
SSPE	saline-sodium-phosphate-EDTA buffer: 0.15 M NaCl, 10 mM NaH_2PO_4, 1 mM EDTA, pH 7.4
stop buffer	95% formamide; 0.05% xylene cyanol FF; 0.05% bromophenol blue; 20mM EDTA
STRP	short tandem repeat polymorphism
STS	sequence tagged site
T or t	thymine (in DNA sequence)
T buffer	Tris acetate buffer [10X contains: 330 mM Tris-acetate (pH 7.9), 100 mM Mg-acetate, 5 mM Dithiothreitol, 660 mM K-acetate]
T_a	annealing temperature
T_m	melting temperature
Taq polymerase	deoxynucleosidetriphosphate: DNA deoxynucleotidyltransferase, EC 2.7.7.7, from *Thermus aquaticus* BM, recombinant (*E. coli*)
TBE buffer	Tris borate-EDTA buffer: 89.15mM Tris®[2] (pH 8.0), 88.95 mM boric acid, 2.498 mM Na_2EDTA
TCAP	telethonin
TE	10 mM Tris-HCl (pH 7.5); 1 mM EDTA
TEL	telomere
TEMED	N,N,N,N'-tetramethylethylenediamine: $C_6H_{16}N_2$
temp	temperature
TetNR	tetranucleotide repeat
TMD	tibial muscular dystrophy
TPM3	α tropomyosin
TriNR	trinucleotide repeat
Tris	Tris®: tris(hydroxymethyl)-amino-methane: 2-amino-2-(hydroxymethyl)-1,3-propanediol: $C_4H_{11}NO_3$
Tris-HCl	2-amino-2-2(hydroxymethyl)-1,3-propanediol hydrochloride: $C_4H_{11}NO_3.H_2O$
Triton X-100	triton X-100®[3]: octylphenolpoly(ethylene-glycolether)$_n$: $C_{34}H_{62}O_{11}$, for n = 10
TUBB4Q	human beta-tubulin gene
U	units

[1] Sequenase® is a registered trademark of United States Biochemical Corporation, Cleveland, OH, U.S.A.
[2] Tris ® is a registered trademark of Rohm & Haas Company, Philadelphia, PA, U.S.A.
[3] Triton X-100® is a registered trademark of Rohm & Haas Company, Philadelphia, PA, U.S.A.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

UK	United Kingdom
Urea	CH_4N_2O
USA	United States of America
UV	ultraviolet
V	volt
VNTR	variable number of tandem repeats
VPDMD	vocal cord and pharyngeal weakness with ausotomal dominant distal myopathy
W	watt
Xap I (isoschizomer *Apo* I)	restriction endonuclease isolated from an *E. coli* strain that carries the cloned *Apo* I gene from *Arthrobacter protophormiae* with recognition site 5'-Pu↓AATTPy-3'
XC	xylene cyanole FF: $C_{25}H_{27}N_2O_6S_2Na$
xg	gravitational acceleration
XR	X-linked recessive
YAC	yeast artificial chromosome
■ / ●	male/female: tested FSHD normal
■ / ●	male/female: FSHD equivocal
□ / ○	male/female: never tested for FSHD (phenotypical status unknown)
■ / ●	male/female: tested FSHD positive
■ / ●	obligate gene carrier
□ / ○	male/female: deceased
◇	sex unknown
◇	multiple individuals, exact number of individuals unknown
—//—	divorced
↖	proband
⌃	dizygotic twins
⌃	monozygotic twins
►	recombination event

LIST OF FIGURES AND GRAPHS

Figure 2.1: Distribution of muscle groups predominantly affected in various muscular dystrophies ... 5
Figure 2.2: Muscle membrane proteins ... 6
Figure 2.3: Facial muscles affected in FSHD ... 8
Figure 2.4: Upper extremities, lower extremities, shoulder girdle and truncal muscles affected in FSHD ... 10
Figure 2.5: Scapular fixation in FSHD ... 18
Figure 2.6: Partial map indicating the relative position and loci and restriction endonuclease recognition sites within the 4q35 region ... 23
Figure 2.7: Restriction endonuclease map of the 4q35 and 10q26 regions... 25
Figure 2.8: Schematic representation of hybrid chromosomes ... 29
Figure 2.9: Schematic representation of the *Bgl* II – *Bln* I dosage test ... 31
Figure 3.1: Excerpt from a pedigree of family F10 ... 49
Figure 3.2: Excerpt from a pedigree of family F21 ... 50
Figure 3.3: Excerpt from a pedigree of family F30 ... 50
Figure 3.4: Excerpt from a pedigree of family F40 ... 51
Figure 3.5: Excerpt from a pedigree of family F60 ... 52
Figure 4.1: Representative autoradiograph of marker UT1366 ... 77
Figure 4.2: Representative autoradiograph of marker GATA5B02 ... 78
Figure 4.3: Representative autoradiograph of marker AFMa224xh1 ... 79
Figure 4.4: Representative autoradiograph of marker ATA22F02 ... 80
Figure 4.5: Representative autoradiograph of marker UT5785 ... 81
Figure 4.6: Representative autoradiograph of marker UT2219 ... 82
Figure 4.7: Representative autoradiograph of marker UT7694 ... 83
Figure 4.8: Representative autoradiograph of marker AFMa190zf5 ... 84
Figure 4.9: Representative autoradiograph of marker AFM238ve3 ... 85
Figure 4.10: Representative autoradiograph of the optimised Southern blot analysis ... 89
Figure 4.11: Comparison between radio-active and non radio-active detection 90
Figure 4.12: Haplotypes of selected individuals from family F10 ... 94
Figure 4.13: Southern blot analysis of family F15 ... 95
Figure 4.14: Haplotypes of selected individuals from family F20 ... 98
Figure 4.15: Haplotypes of selected individuals from family F30 ... 101
Figure 4.16: Haplotypes of selected individuals from family F40 ... 103
Figure 4.17: Haplotypes of selected individuals from family F60 ... 104
Figure 4.18: Southern blot analysis of family F60 ... 106
Figure 5.1: Relationship between five FSHD families investigated in this study 111
Figure F.1: Full pedigree of South African FSHD family F10 ... 145
Figure F.2: Full pedigree of South African FSHD family F11 ... 148
Figure F.3: Full pedigree of South African FSHD family F12 ... 151
Figure F.4: Full pedigree of South African FSHD family F13 ... 152
Figure F.5: Full pedigree of South African FSHD family F14 ... 153
Figure F.6: Full pedigree of South African FSHD family F15 ... 156

Figure F.7: Full pedigree of South African FSHD family F21 157
Figure F.8: Full pedigree of South African FSHD family F30 159
Figure F.9: Full pedigree of South African FSHD family F40 163
Figure F.10: Full pedigree of South African FSHD family F60 164

Graph 2.1: Proportion of patients with different clinical severity scores and *Eco* RI deletion fragment sizes 37

LIST OF TABLES

Table 2.1: Penetrance of FSHD 13
Table 2.2: Composition of translocated arrays 27
Table 2.3: Clinical Severity Scale for FSHD 36
Table 3.1: Average DNA yield from various amounts of starting material 53
Table 3.2: Genetic map of human chromosome 4q35 55
Table 3.3: Primers for short tandem repeat polymorphism markers located on chromosome 4q35 56
Table 3.4: Allele frequencies for marker UT1366 at locus D4S1523 57
Table 3.5: Partial gDNA sequence at marker UT1366 at locus D4S1523 57
Table 3.6: Allele frequency for marker GATA5B02 at locus D4S1652 57
Table 3.7: Partial gDNA sequence at marker GATA5B02 at locus D4S1652 ... 58
Table 3.8: Allele frequency for marker AFMa224xh1 at locus D4S2930 58
Table 3.9: Partial gDNA sequence at marker AFMa224xh1 at locus D4S2930 59
Table 3.10: Allele frequency for marker ATA22F02 at locus D4S2390 59
Table 3.11: Partial gDNA sequence at marker ATA22F02 at locus D4S2390 59
Table 3.12: Allele frequency for marker UT5785 at locus D4S2299 60
Table 3.13: Partial gDNA sequence at marker UT5785 at locus D4S2299 60
Table 3.14: Allele frequency for marker UT2219 at locus D4S2283 61
Table 3.15: Partial gDNA sequence at marker UT2219 at locus D4S2283 61
Table 3.16: Allele frequency for marker UT7694 at locus D4S2688 61
Table 3.17: Partial gDNA sequence at marker UT7694 at locus D4S2688 62
Table 3.18: Allele frequency for marker AFMa190zf5 at locus D4S2921 62
Table 3.19: Partial gDNA sequence at marker AFMa190zf5 at locus D4S2921 63
Table 3.20: Allele frequency for marker AFM238ve3 at locus D4S426 63
Table 3.21: Partial gDNA sequence at marker AFM238ve3 at locus D4S426 63
Table 4.1: Optimised conditions for nine short tandem repeat polymorphism markers located on chromosome 4q35 75
Table 4.2: Relative enzyme activity in Amersham™ buffers 86
Table B.1: Mode of inheritance of different types of muscular dystrophies 125
Table C.1: Diagnostic criteria for FSHD 129
Table D.1: Nucleotide sequence of repeat units at D4Z4 locus 132
Table E.1: Comparison of one *Kpn*I repeat unit nucleotide sequence derived from 4q35 and 10q26 139

ACKNOWLEDGEMENTS

I would like to express my appreciation towards the following individuals and institutions:

The **FSHD families** for their participation in this project. Their interest and contribution to this project was invaluable and without their involvement this study would have been impossible.

My supervisor, **Prof. Antonel Olckers**, for her never-ending support and for sharing her wisdom, experience and insight with me to enable me to fulfil my dreams. I am especially grateful to her for providing me with opportunities of a lifetime, and for making a tremendous impact on my scientific career. She is truly committed to the education of her students, and not just involved. **Dr. Clara Schutte**, my co-supervisor, for the clinical examinations of the patients and for partial funding to attend the 6th World Muscle Society meeting and the 51st American Society of Human Genetics meeting. The **Department of Neurology** for hosting of the weekly FSHD clinics and for assisting with compilation of the initial family pedigrees. **Dr. Engela Honey** for the genetic counseling and blood collection at the FSHD clinics. **Mr. Francois Honiball** for running the FSHD support group. **Debbie Prosser**, a colleague and true friend, for her encouragement and support throughout the years. **Marco Alessandrini** for joining the FSHD team and for sharing his ideas and enthusiasm with us. I am also grateful for the chromosome 10 haplotype analysis he performed which allowed us to determine the chromosomal origin of some deletion fragments. To the other members of the team: **Bronwyn Wedge, Coenie Goosen, Dr. Francois Maree** and **Wayne Towers**, for their encouragement and support. The sense of humour and laughter in the lab ensured an uplifting environment every day. **Debbie**, **Marco** and **Bronwyn** for proof reading parts of my thesis. The **MDF National Office** and **Gauteng Branch** for financial assistance to attend both national and international conferences and for the grants awarded to Dr. Antonel Olckers to fund the FSHD project and to enable completion of this study. **Intergrated DNA Technologiec Inc.** (especially Ric Devor, Mark Behlke and Vicky Struzynski-Olson) for accepting me onto the J1-Gene Quantification programme and allowing me to gain experience and expertise to broaden my scientific knowledge. **Prof. Doug Wallace**, for inviting me to his lab to gain experience in pulse field gel electrophoresis. **Prof. Rune Frants** and **Prof.**

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Silvére van der Maarel for the kind donation of the p13E-11 probe and the original Southern blot protocol. **Martie van der Walt** and other librarians of the Faculty of Medicine for assistance. The former **Department of Human Genetics and Developmental Biology**, the **Faculty of Medicine** and the **University of Pretoria** for providing the infrastructure to complete this study. The research committee of the University of Pretoria (NAVKOM) for the grants awarded to me in 2000 and 2001. A former colleague, **Yolande Havenga,** for the DNA isolation of selected patients. **Vincent Procaccio** for the French translation of the title of the first FSHD article.

My **parents** and immediate family (**Ronel, Wilna, Marlize, Kayla, Riaan and Chris**) for loving and believing in me and for their patience, genuine encouragement and support. I am especially grateful for the ability and endurance I received from **God** to complete my study.

CHAPTER ONE

INTRODUCTION

The muscular dystrophies are defined as a group of genetic disorders with progressive muscle wasting and weakness. At least 1 in 3,000 individuals are affected by an inherited neuromuscular disorder, of which the muscular dystrophies make up a considerable proportion (Emery, 1998). The exact pathogenesis of many of the muscular dystrophies is, however, still unknown.

Facioscapulohumeral muscular dystrophy (FSHD) is the third most common inherited disorder of muscle after Duchenne and Myotonic dystrophy, with a prevalence of at least 1 in 20,000 (Padberg, 1982). FSHD is characterised on a clinical level by progressive weakening and atrophy of the face, shoulder-girdle and upper arm, but other skeletal muscles may also become involved with progression of the disorder.

FSHD is also known as Landouzy-Déjérine disorder after the two physicians, who described it in 1884 (Landouzy and Déjérine, 1884). Duchenne de Boulogne was, however, the first to give a classical description of FSHD in the late 1800's, stating its myopathic nature and inheritance pattern (Kazakov *et al.*, 1974). A historical discussion therefore arose toward the end of the last century between Erb and Landouzy-Déjérine regarding the first report of "facio-scapulo-humeral muscular dystrophy" (FSHD), as FSHD was for the first time described by Duchenne de Boulogne under the name 'progressive muscular atrophy of childhood' and Erb described a 'juvenile shoulder-girdle' type. Landouzy and Déjérine described patients who differed from those described by Duchenne with regard to the progression of the muscles affected from the upper to the lower part of the body. The patients described by Landouzy and Déjérine corresponded to the cases described by Erb, but a new name (FSH type) was used by these two authors. The forms described by Duchenne and Erb were therefore not distinct from the FSH type, but were rather included into this type. Eventually the first report of FSHD was attributed to the report by Landouzy and Déjérine in 1884.

More than a century after the first report of FSHD, the FSHD locus was assigned to chromosome 4q35 through linkage analysis (Wijmenga *et al.*, 1990). This region consists of 3.3 kb repeat arrays (D4Z4) and a deletion of an integral number of these repeats has been observed in individuals affected with FSHD (van Deutekom *et al.*, 1993). To date no gene has been identified for FSHD although the molecular defect of FSHD has been defined. A detailed history of FSHD is presented in chapter two.

This is the first extensive molecular study to characterise individuals affected with FSHD in the South African population. Haplotype analysis conducted with nine short tandem repeat polymorphism (STRP) markers located on chromosome 4q35 and Southern blot analysis were performed utilising probe p13E-11 to detect DNA rearrangements of the D4Z4 repeat arrays. In chapter three and four the protocols and results obtained via these analyses are described.

Conclusions drawn from the results obtained in this study are presented in chapter five. The results obtained in this study to date excluded the possibility of the existence of genetic heterogeneity in the South African FSHD population, but suggested a Dual Founder Effect for FSHD in the South African population. Some aspects of the genetic aetiology of FSHD in this population could therefore be determined in this study.

CHAPTER TWO

THE AETIOLOGY AND PATHOGENESIS OF FACIOSCAPULOHUMERAL MUSCULAR DYSTROPHY

Muscular dystrophy is not a single disorder, but various types have been defined over the years. The muscular dystrophies are a large and heterogeneous group of inherited disorders and the one feature that the muscular dystrophies as a group share is progressive muscle wasting and weakness.

2.0 THE MUSCULAR DYSTROPHIES

Walton and Nattrass (1954) originally classified the muscular dystrophies into three main groups based on the mode of inheritance (Appendix B) and the distribution of the muscle groups that are predominantly affected:

- Duchenne (X-linked recessive type),
- Facioscapulohumeral (autosomal dominant), and
- Limb girdle (autosomal recessive and sporadic limb-girdle type).

These authors also distinguished three relatively uncommon, but clinically and genetically distinctive forms of muscular dystrophy, namely: Distal, Oculopharyngeal, and Congenital muscular dystrophy. The distribution of muscles that are predominantly affected in six of the muscular dystrophies are illustrated in Figure 2.1. These muscular dystrophies can be distinguished on both the clinical and molecular levels. The six muscular dystrophies presented in Figure 2.1 are discussed in the subsequent paragraphs.

Duchenne muscular dystrophy (DMD) [MIM 310200] and Becker muscular dystrophy (BMD) [MIM 300376] are the dystrophin associated muscular dystrophies. DMD is a more severe form than BMD with DMD being the most common form of muscular dystrophy. Both are inherited as X-linked recessive disorders and therefore predominantly affect boys. DMD and BMD are caused by deletions of one, or many, exons in the dystophin gene that was mapped to Xp21.2. The most distinctive feature of DMD is proximal muscle

weakness with characteristic pseudohypertrophy of the calves. Disease onset is typically before the age of three and individuals become wheel-chair dependent by 12 and are generally deceased by the age of 20. Individuals affected by BMD are usually affected in their twenties or thirties and generally have a normal life span. (Emery 1998; OMIM 2001a)

Dreifuss and Emery described Emery-Dreifuss muscular dystrophy (EDMD) [MIM 310300, 181350] in the mid-nineteen sixties. Significant features are cardiomyopathy, generally presenting as atrioventricular block, weakness with a humeroperoneal distribution in the early stages and early contractures of the elbows and spine. Muscles of the lower extremities are usually first affected by the age of four or five. By the early teens individuals develop a waddling gait with increased lumbar lordosis, and weakness of the shoulder girdle muscles appears later. EDMD is inherited as an X-linked recessive disorder and is caused by a mutation in the emerin gene located on Xq28. An autosomal dominant form of EDMD also exists which results from mutations in the lamin A/C (LMNA) gene, encoding two components of the nuclear lamina (lamins A and C), located on chromosome 1. (Emery 1998; OMIM 2001b)

Limb-girdle muscular dystrophy (LGMD) [MIM 159000, 159001, 253600, 253601, 253700, 254110, 600119, 601287, 601173, 601954, 603511, 604286] is a clinically and genetically heterogeneous group of disorders. Onset is generally in childhood, but it can also present later in life. Muscle involvement is generally first noted in either the pelvic, or the shoulder girdle areas. Asymmetry is often observed when the upper limbs are first involved. LGMD is inherited in either an autosomal dominant or autosomal recessive form. At least six dominant and nine recessive sub-types have been identified. LGMD2A is caused by a muscle specific protease (calpain 3) deficiency, and four other recessive types have been found to be caused by deficiencies of sarcoglycans, which form part of the dystrophin associated protein complex of the muscle membrane (Figure 2.2). (Emery 1998; OMIM 2001c)

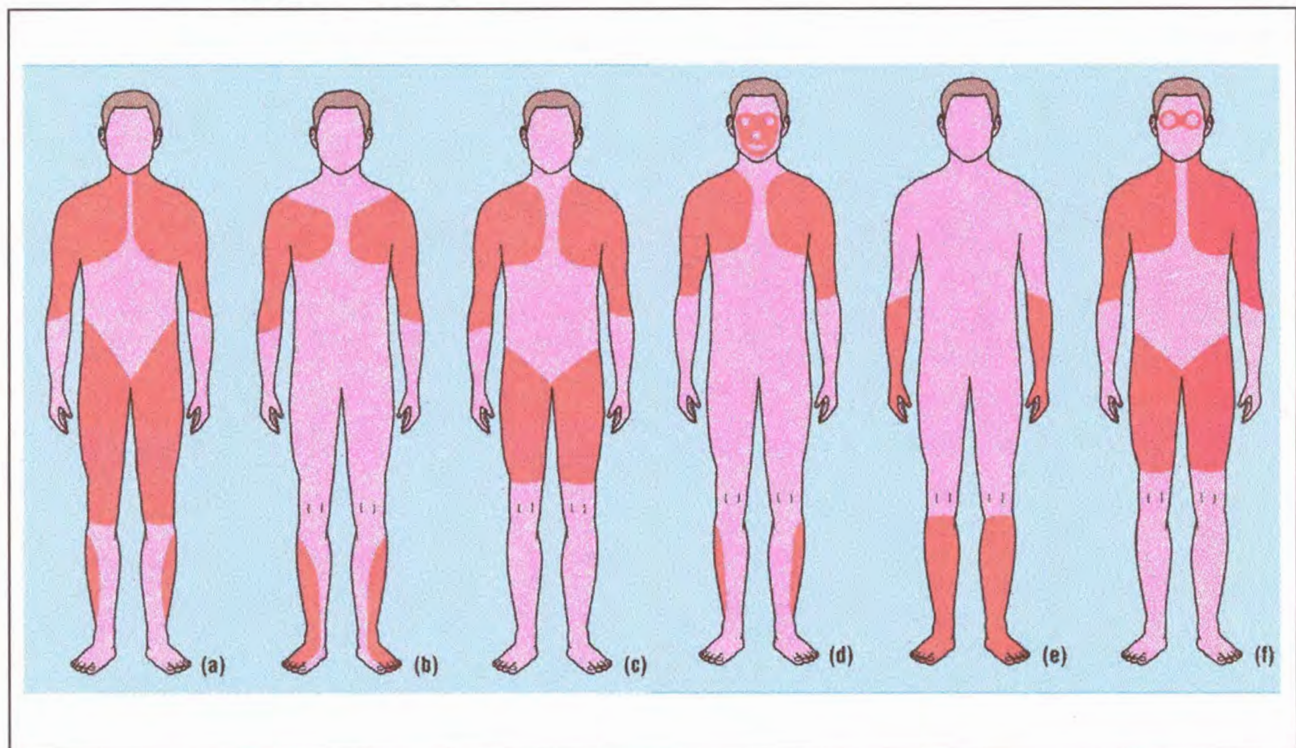
Distal myopathies [MIM 254130] are relatively rare and are associated with wasting and weakness of the distal muscles. Individuals affected with a distal myopathy are generally only mildly affected, however, some may develop problems in walking. (Emery 1998; OMIM 2001d)

Oculopharyngeal muscular dystrophy (OPMD) [MIM 164300] presents late in life and is characterised by dysphagia and progressive ptosis of the eyelids. Weakness of the

muscles of the neck and proximal upper limbs can also be present. OPMD is caused by an expansion of a $(GCG)_6$ repeat encoding a polyalanine tract located at the N terminus of the poly(A)-binding protein 2 located on chromosome 14. (Emery 1998; OMIM 2001e)

Facioscapulohumeral muscular dystrophy (FSHD) [MIM 158900] is characterised by progressive muscle wasting of the facial, shoulder, and upper arm muscles, however, other muscles such as the abdominal, foot extensor and pelvic girdle muscles may also be involved. FSHD is an autosomal dominant disorder caused by a DNA rearrangement on chromosome 4q35. The gene involved in FSHD has not been determined yet. The study presented here focuses on facioscapulohumeral muscular dystrophy (FSHD), which is discussed in section 2.0 (Emery, 1998; OMIM 2001f).

Figure 2.1: Distribution of muscle groups predominantly affected in various muscular dystrophies

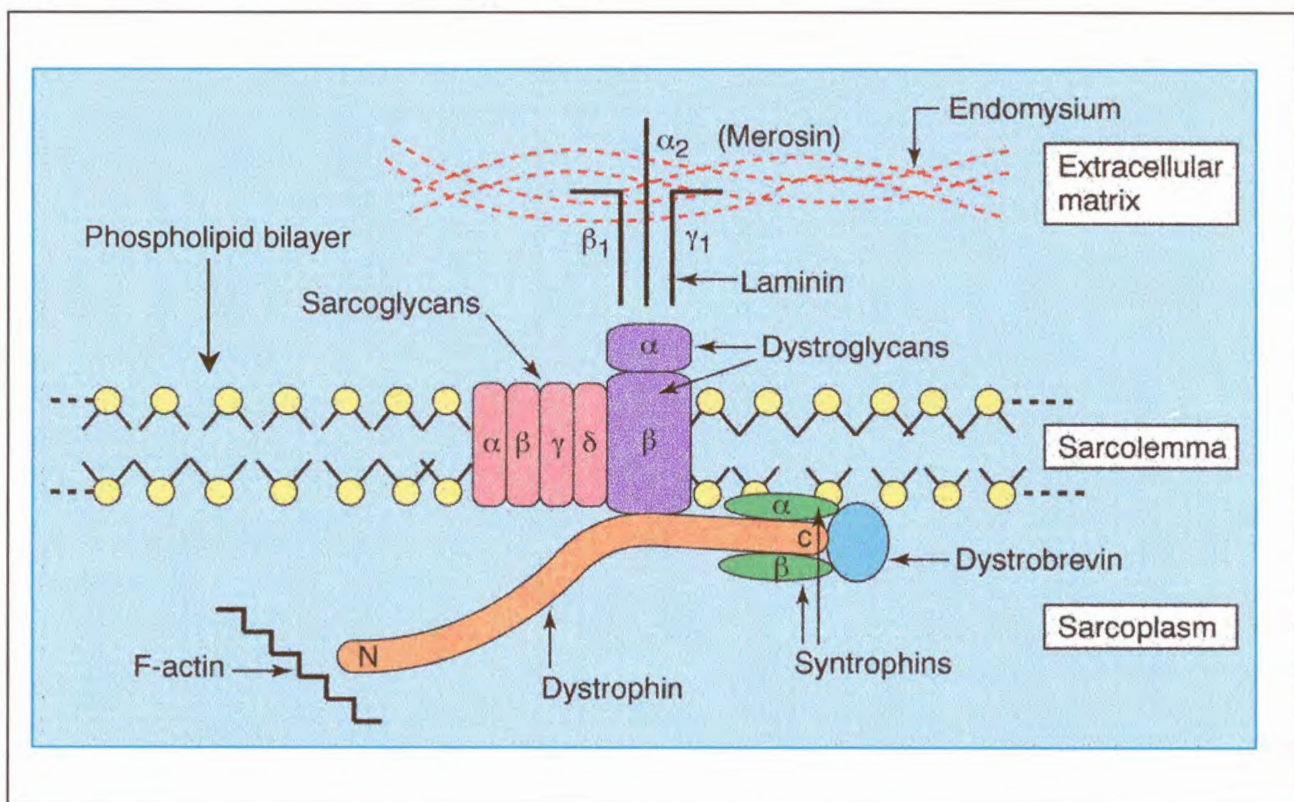


Different types of muscular dystrophies: a = Duchenne and Becker; b = Emery-Dreifuss; c = Limb girdle; d = Facioscapulohumeral; e = Distal; f = Oculopharyngeal. Adapted from Emery *et al.* (1998).

Immunohistochemical techniques enabled the identification of specific deficiencies of various membrane proteins, including dystrophin and sarcoglycan referred to as dystrophin associated glycoproteins. This identification resulted in the classification of some of the dystrophies on a histological level into dystrophinopathy which included Duchenne and Becker muscular dystrophy and sarcoglycanopathy consisting of some of the Limb girdle type muscular dystrophies. The sarcoglycan complex (SGC) consists of five units: α-, β-, γ-, δ-, ε-sarcoglycans, with the first four being associated with muscular

dystrophies (Figure 2.2). The SGC comprises the dystrophin-glycoprotein complex (DGC) together with dystrophin and the dystroglycan complex which are membrane-associated muscle proteins that span the muscle sarcolemma and form the link between the extracellular matrix (endomysium) and intracellular F-actin as illustrated in Figure 2.2. The dystrophin-glycoprotein complex is suggested to have a role in the maintenance of the stability, integrity and strength of the muscle membrane. Disruption of the complex could therefore cause a cascade of events resulting in muscle weakness (Emery 1998; Mak *et al.,* 2001). However, the precise mechanism as to how the absence of these proteins results in muscle weakness remains unresolved.

Figure 2.2: Muscle membrane proteins



Adapted from Emery *et al.* (1998)

2.1 FACIOSCAPULOHUMERAL MUSCULAR DYSTROPHY

Facioscapulohumeral muscular dystrophy (FSHD) is part of a group of inherited disorders characterised by progressive muscle wasting and weakness. FSHD is the third most common inherited disorder of muscle after Duchenne and Myotonic dystrophy, with a prevalence of at least 1:20,000 in Caucasians, including familial and sporadic cases (Padberg, 1982).

2.1.1 CLINICAL ASPECTS OF FSHD

Landouzy and Déjérine, after whom FSHD was initially named, published their first article in the "Revue de Medicine" (Review of Medicine) entitled "La myopathie atrophique progressive; myopathie sans neuropathie debutant d'ordinaire dan l'enfance, par la face", (Progressive atrophic myopathy, myopathy without neuropathy starting usually in childhood, by the face) in 1884 (Landouzy and Déjérine, 1884). In this paper the term 'facioscapulohumeral' was mentioned for the first time. They described an autopsy on a man who died of tuberculosis when he was 24 years old. Atrophy of the facial muscles was the only symptom noted at the age of three while atrophy of the shoulder girdle and upper arm muscles developed at the age of 17. The weakness progressed slowly to the muscles of the trunk and pelvic girdle. The muscles of the tongue, pharynx, larynx masseter, temporal and pterygoid muscles were not affected. No abnormalities on the brain, spinal cord, peripheral nerves and intramuscular nerve endings were present. The younger brother and sister of the patient were similarly affected and an autosomal dominant pattern of inheritance was observed. The father of the patient developed muscle weakness in the shoulder girdle at the age of 26 and facial weakness was observed when he was 32 years old (Padberg, 1982). The complex and difficult diagnosis of FSHD on the clinical level was thus already observed in the first reported case, more than a century ago. Unfortunately the difficulties associated with the clinical diagnosis of FSHD still complicate diagnosis today and are mainly due to the complex clinical phenotype of this disorder.

2.1.1.1 Presenting symptoms

This disorder is characterised at onset by asymmetrical weakness and atrophy of the facial and shoulder girdle muscles (Padberg 1982). The following muscles may also be affected: abdominal, foot extensor, upper arm, pelvic girdle and lower arm muscles. However, the degree to which the muscles can be affected varies dramatically, which forms the basis of the extremely heterogeneous nature of the FSHD phenotype (Padberg 1982).

Diagnostic criteria for FSHD have been defined by the International FSHD Consortium and are provided in Appendix C. In summary the following criteria define FSHD (Padberg *et al.,* 1991):

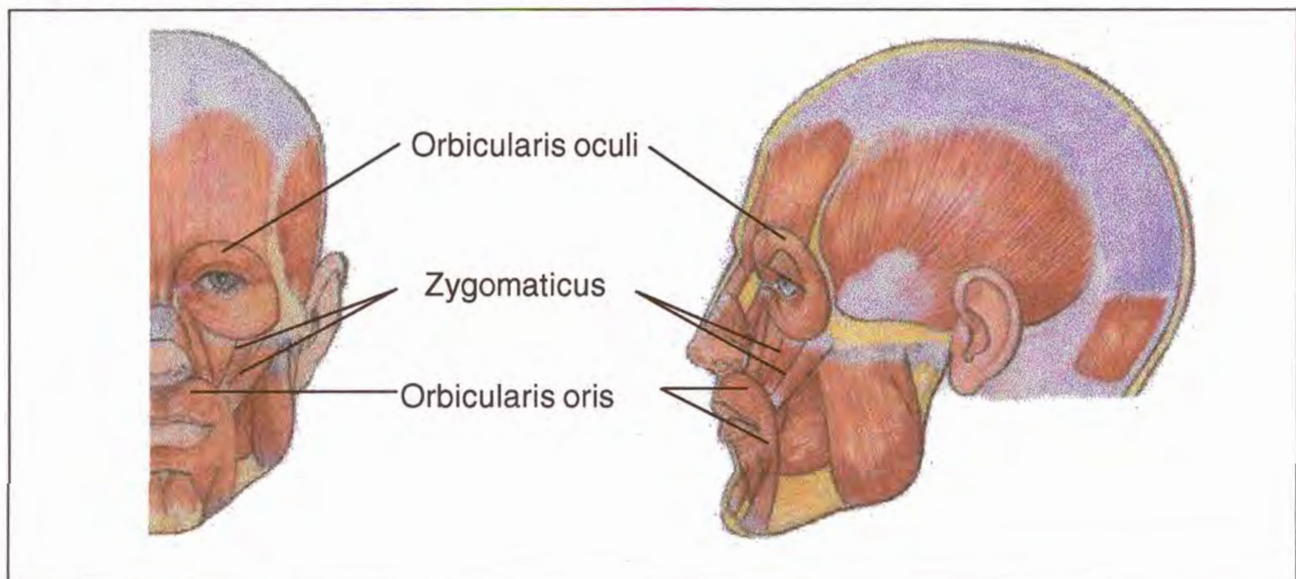
1. weakness of the facial or shoulder girdle muscles, but sparing of the extra-ocular, pharyngeal and lingual muscles and the myocardium,

2. facial weakness in more than 50% of the affected family members,
3. autosomal dominant inheritance in familial cases, and
4. presence of a myopathic disorder in both the electromyography (EMG) and muscle biopsy in at least one affected family member.

2.1.1.1.1 The facial muscles

Facial weakness is generally present in the early stages of the disorder and the zygomaticus, orbicularis oculi and orbicularis oris muscles are generally affected (Figure 2.3). Weakness of the zygomaticus muscles results in the inability to raise the corners of the mouth and patients express a grin instead of a smile. The orbicularis oris allows pursing of the lips, whistling and retaining of air under pressure. Patients find it difficult to bury the eyelashes completely when forced closed due to the weakness of the orbicularis oculi muscles. If the weakness progresses, a small piece of the sclera may be visible when attempting to close the eyes. Individuals are often seen sleeping with open eyes, and blinking is slowed and incomplete (Padberg, 1982). Other facial muscles may also become involved, resulting in an unlined forehead and a smooth, expressionless face, resulting in the term "facies myopathica" or myopathic face (Landouzy and Déjérine, 1885).

Figure 2.3: Facial muscles affected in FSHD



Adapted from Dorland's Pocket Medical Dictionary, (1995).

Facial weakness is, however, not obligatory in the diagnosis of FSHD. Felice *et al.* (2000) examined seventeen unrelated patients with facial-sparing scapular myopathy. Fourteen of the seventeen patients agreed to DNA testing and ten of the fourteen (71%) had a *Bln* I

resistant deletion fragment, which defines FSHD on the molecular level (see paragraph 2.1.2.2). Felice *et al.* (2001) also reported three individuals with no visible facial weakness. The first individual was at first diagnosed with facial-sparing scapular myopathy with symptoms of scapular winging, humeral and quadriceps atrophy together with a waddling-type gait. The second individual had difficulty climbing stairs during high school and had limited movement of wrist and finger flexors and extensors, as well as foot and toe flexors. There were no facial, scapular winging, or other characteristic features of FSHD. This individual was eventually diagnosed with limb-girdle muscular dystrophy. The last individual displayed asymmetric scapular winging and mild right trapezius atrophy. His limb-muscle strength was normal, and there was no facial weakness. FSHD was, however, confirmed on a molecular level in all three of these individuals.

Muscle disorders, clinically similar to FSHD include scapuloperoneal and scapulohumeral dystrophies, which are associated with minimal or complete absence of facial weakness, and may represent milder forms of the disorder since there is evidence for involvement of the same genetic locus (Jardine *et al.,* 1994a). The scapuloperoneal syndrome is difficult and sometimes even impossible to distinguish from FSHD on the clinical level. Identification of the FSHD gene might elucidate whether the disorders are indeed different on the genetic and molecular levels.

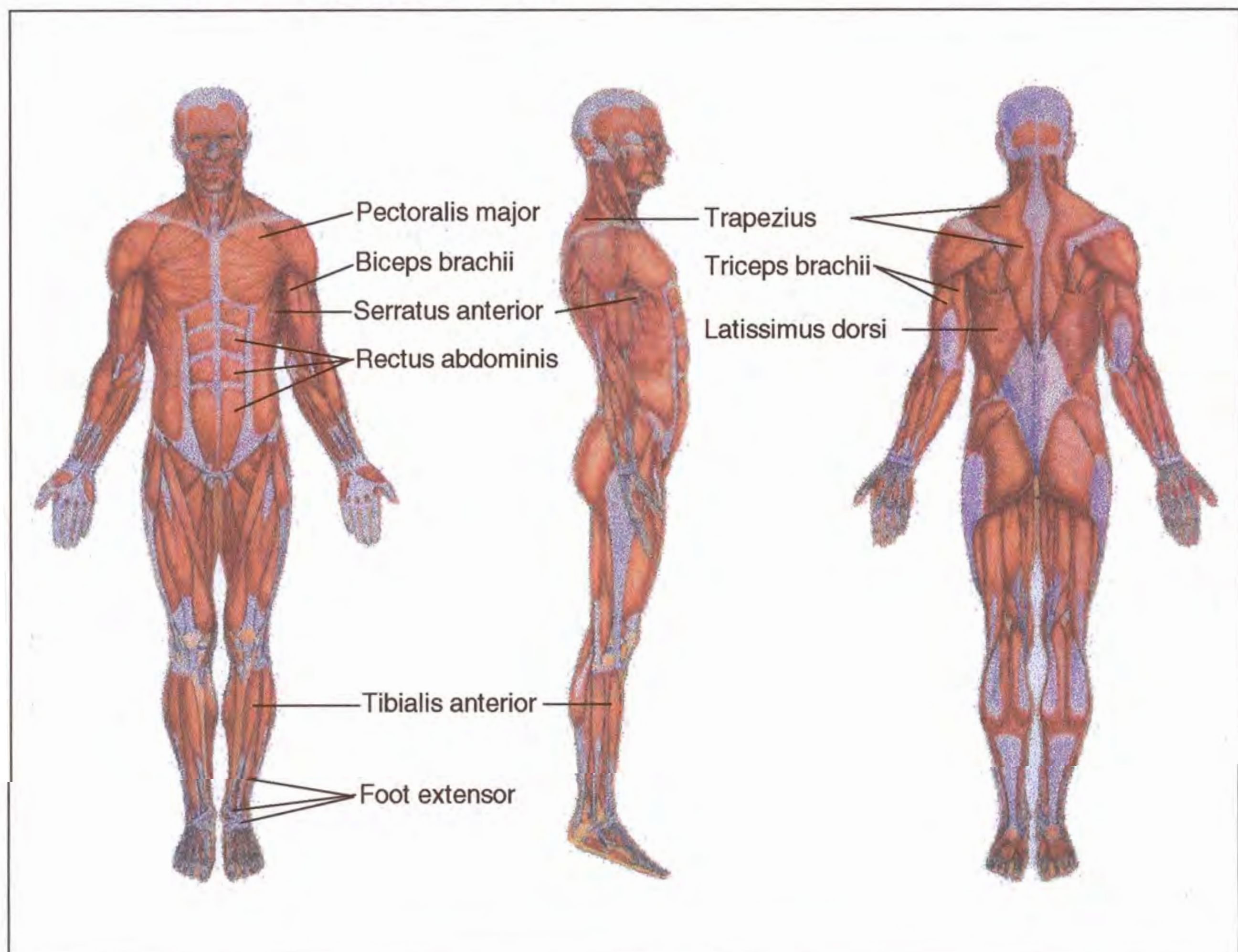
2.1.1.1.2 The upper extremities, shoulder girdle and neck muscles

Apart from the weakness of the facial muscles, one of the earliest symptoms in FSHD is the gradual loss of fixation of the scapula. The rhomboids, the lower part of the trapezius and the serratus anterior muscles, which stabilise the scapula to the torso, are all involved (Padberg, 1982). Figure 2.4 displays the muscles involved, except for the rhomboid muscle, which is part of the deep muscles and is therefore not visible on this figure. Weakness of these muscles results in allowing the scapulae to rotate slightly laterally, and to move upwardly, laterally and anteriorly over the thorax. If the rhomboid and serratus weakness progresses, scapulae alatae appear. The positional change of the scapulae contributes to the development of drooping shoulders. The clavicles lose their normal upward slope, rotate anteriorly and may ultimately even slope downwards. If the scapula fixation becomes weaker, the arms cannot be raised completely and can, for instance, be swung upwards to catch an object that is above shoulder height. When the scapular fixation, and especially the serratus function worsens further, elevation of the arm above shoulder level becomes impossible. When attempting abduction of the arms, the scapulae

ride upward over the back and their upper borders rise high up into the normal location of the trapezius muscles.

Weakness of the latissimus dorsi and the sternocostal part of the pectoralis muscles may also occur (Figure 2.4). Wasting of the latter will result in a flattened outline of the anterior thoracic wall, with a change in the direction of the axillary crease, running more horizontally instead of vertically, and pointing to the sternoclavicular joint (Padberg, 1982). The time between the onset of shoulder girdle weakness and the onset of upper arm weakness may be quite variable. Atrophy of the biceps and triceps in the upper arms may become quite severe even early in the course of the disorder resulting in the upper arm appearing thinner than the lower arm (often referred to as "Popeye" arms), due to the relative sparing of the lower arm muscles.

Figure 2.4: Upper extremities, lower extremities, shoulder girdle and truncal muscles affected in FSHD



Adapted from Dorland's Pocket Medical Dictionary, (1995).

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

2.1.1.1.3 The truncal muscles

Padberg (1982) reported abdominal muscle weakness in 58% of the patients who were investigated. Weakness of the abdominal muscles (Figure 2.4) was observed to add to the pelvic tilt and the increased lumbar lordosis caused by the anterior convexity of the lumbar spine. The lumbar lordosis was also observed to become more severe when an individual was wheelchair dependent.

2.1.1.1.4 The lower extremities and the pelvic girdle muscles

Landouzy and Déjérine (1885) and Chyatte *et al.* (1966) observed the early weakness of the anterior tibial muscles in FSHD patients. Weakness of the foot extensors (Figure 2.4) interferes with walking, resulting in a steppage gait, referring to the pattern of how an individual walks, and an inability to run (Padberg, 1982). Steppage gait is characterised by the dropping of the foot, where the foot hangs with the toes pointing down, causing the toes to touch the ground while walking. Patients tend to trip easily over small objects, falling forward onto their knees. If pelvic girdle weakness develops, a waddling gait characterised by a distinctive duck-like walk will be visible and gradual rising from a chair or climbing stairs becomes increasingly difficult.

2.1.1.1.5 Asymmetry of muscle involvement

Distinct asymmetry of muscle involvement is an important and common feature of FSHD (Padberg, 1982). The asymmetry can be present in the facial as well as in the shoulder girdle muscles, and the extremities. No correlation was observed between right or left handed individuals, the side of muscle involvement, or the severity of affected muscles (Tawil *et al.*, 1994).

2.1.1.1.6 Extramuscular involvement

Sensorineural deafness and retinal vascular abnormalities have been described by many authors to be associated with FSHD (Gieron *et al.*, 1985; Gurwin *et al.*, 1985; Korf *et al.*, 1985; Matsuzaka *et al.*, 1986; Voit *et al.*, 1986; Fitzsimons *et al.*, 1987; Yasukohchi *et al.*, 1988; Brouwer *et al.*, 1991; Pauleikhoff *et al.*, 1992; Padberg *et al.*, 1995a). High-frequency hearing loss has been observed in 50 to 64% of FSHD patients (Brouwer *et al.*, 1991; Padberg *et al.*, 1995a). Brouwer *et al.* (1991) observed a significant difference in

hearing level between 4,000 Hertz (Hz) and 6,000 Hz in FSHD patients when compared to controls. The hearing loss was also observed to be progressive and, with time, tends to involve lower frequencies of the spectrum. These authors observed that the severity of the hearing loss varied between individuals and that it was not age dependent.

Originally, retinal vascular abnormalities were mainly reported in severe cases with early-onset FSHD (Gurwin *et al.*, 1985). Fitzsimons *et al.* (1987) and Brouwer *et al.* (1993) observed exudative retinal vasculopathy, with capillary telangiectasis, microaneurysms, and capillary closure to be present in 50 to 75% of FSHD patients. This is characterised by damage or breakdown of the retina due to subretinal fluid accumulation and the capillaries being longer, wider and fewer in number than normal. No correlation between the severity of the muscular weakness, and the severity of the hearing loss or the retinal vasculopathy was observed. The pathogenic mechanism of hearing loss and retinal vasculopathy in FSHD is still unknown.

Cardiac and respiratory muscles are generally not affected in FSHD patients (Padberg 1982; Tawil *et al.*, 1998). Emery-Dreifuss muscular dystrophy, a phenotypically similar but genetically distinct disorder, is distinguished from FSHD on the clinical level by the presence of cardiac muscle involvement. Stevenson *et al.* (1990) provided the first evidence of cardiac involvement in FSHD, unfortunately, no molecular data was available to confirm the clinical diagnosis of FSHD. Laforet *et al.* (1998) reported the presence of cardiac involvement in genetically confirmed FSHD patients. This study included 100 patients of whom five had conduction defects or arrhythmia. It is therefore evident that patients with FSHD may have cardiac involvement, although to date it has only been observed in a small number of cases. Finsterer *et al.* (2000) investigated the presence of cardiac involvement (CI) in several myopathies, such as DMD, BMD, EDMD, FSHD, sarcoglycanopathies, myotonic dystrophies type 1 and 2 and mitochondrial myopathies, and observed CI to be present in all the disorders. CI can lead to symptoms such as: impulse generation defects, impulse conduction defects, thickened myocardium, left ventricular hypertrabeculation, dilatation of the cardiac cavities, intracardial thrombus formation, and heart failure with systolic and diastolic dysfunction. The above findings argue strongly in favour of FSHD patients being investigated for possible cardiac involvement.

Inflammatory cellular infiltrates are frequently (40 to 80%) observed in the skeletal muscle of FSHD patients. General histological changes in FSHD muscle are rather non-specific

with variations in fibre size and scattered small angulated fibres being the most commonly observed anomalies (Arahata *et al.*, 1995).

Involvement of the lingual muscle is considered to be one of the exclusion criteria of FSHD (Appendix C). However, Korf *et al.* (1985) reported four patients with tongue abnormalities, such as atrophic changes and movement disturbances. No findings on the genetic studies of these patients were, however, reported. Tongue atrophy was also reported by Yamanaka *et al.* (2001) in seven Japanese patients, all belonging to a group of early-onset FSHD patients with small *Eco* RI deletion fragments (10 to 17 kb). Miura *et al.* (1998) also reported a female sporadic case with early-onset FSHD with tongue atrophy with a 10 kb *Eco* RI deletion fragment. These studies therefore suggest that individuals with FSHD can have involvement of the lingual muscles and that this clinical aspect should not be used as an exclusion criterium of FSHD. The occurrence of tongue abnormalities in FSHD, especially in early onset FSHD, should, however, be investigated in larger populations.

2.1.1.2 Clinical heterogeneity in the FSHD phenotype

Clinical diagnosis of the FSHD phenotype is complex due to extreme variability in various aspects of the phenotype, including its penetrance. The phenotypical expression varies in severity, rate of progression and the age of onset, ranging from almost asymptomatic individuals, to almost 20% of patients who will be wheelchair dependent by the age of 40 years or older (Lunt *et al.,* 1991). Variability is observed both between and within families, thus complicating genetic counselling. The penetrance of the FSHD gene was estimated to be different for various age groups as listed in Table 2.1. The life expectancy of individuals affected with FSHD is not shortened since cardiac and respiratory functions are generally unaffected (Padberg 1982).

Table 2.1: Penetrance of FSHD

<5%	0 to 4 years
21%	5 to 9 years
58%	10 to 14 years
86%	15 to 19 years
95%	20 years.

Adapted from Lunt *et al.* (1989a).

Tawil *et al.* (1993a) studied monozygotic twins with extreme clinical variability. One of the twins had progressive shoulder girdle weakness from the age of 10. Upon examination, at the age of 27, severe facial, scapular, peroneal and abdominal muscle weakness was observed. The asymptomatic

twin brother showed only mild weakness of the orbicularis oculi muscles. Zygosity was first determined through red cell phenotyping as well as human leukocyte antigen (HLA) typing. High-resolution cytogenetic analysis was performed to exclude any cytogenetic abnormalities. Five markers (D4S139, D4S163, D4S171, D4S130 and D4F35S1), closely linked to FSHD, were also studied to provide further evidence of monozygosity and to exclude the possibility that submicroscopic rearrangements of distal 4q had occurred. One possible explanation for the observed variability is the presence of a somatic mutation in one of the twins. Another explanation is that both twins have FSHD and that their phenotypes represent an extreme case of variability of expression within the family. The second was the most likely explanation and, therefore, further molecular studies were necessary to confirm the presence of a deletion fragment in both individuals. Tawil *et al.* (1993b) performed the additional studies and reported the presence of a unique 4q35 DNA rearrangement in the affected individual. Results from the second study confirmed that in this particular case, the discordance was probably due to a *de novo* postzygotic mutation after the twinning process had occurred.

Griggs *et al.* (1995) re-examined the monozygotic twins discordant for FSHD, first examined by Tawil in 1993, as well as two additional sets of monozygotic twins concordant for FSHD. The twins concordant for FSHD were similarly affected in terms of age of onset, overall degree of disability, and quantitative tests of muscle, but extreme differences in the symmetry of the muscle involved were observed. The inheritance pattern of the discordant twins could not be established, since sufficient symptomatic family members were not available for examination. Upon molecular analysis, deletion fragments were observed in the two twin pairs concordant for FSHD. A deletion fragment was also present in the affected individual of the other twin pair. No deletion fragment was, however, present in the twin with the mild facial weakness. These authors also concluded that the *de novo* postzygotic mutation or mitotic crossover probably occurred during the twinning process. Based on the results from the concordant twin pairs the authors suggested that the deletion fragment determines the age of onset and severity of the disorder, and that the asymmetries observed in FSHD patients are due to other factors, which may be for instance, environmental.

Tupler *et al.* (1998) described two monozygotic male twins affected by FSHD, carrying an identical *de novo Eco* RI deletion fragment. Neurological examination of their parents was normal, with no sign of muscular dystrophy. Haplotype analysis identified that the *de novo* rearrangement on chromosome 4q had occurred in paternal gametogenesis or

postzygotically in the paternal chromosome 4 before twinning. The genetic identity of the two twins did however not concur with their respective clinical phenotypes. One twin was severely affected and the other one was almost asymptomatic. The medical history was the same for the two brothers, except for an antirabies vaccination performed at the age of five in the more severely affected twin. Tupler *et al.* (1998) hypothesised that the vaccination might have triggered an inflammatory immune reaction, which contributed to the more severe phenotype. The pathogenesis of this external factor with relation to the severe FSHD phenotype has, however, not been elucidated.

Miura *et al.* (1998) reported two unrelated, severely affected female sporadic cases diagnosed with early-onset scapulohumeral muscular dystrophy with mental impairment (intelligence quotients (IQs) ranging from 33 to 45) and epilepsy. One patient suffered from epilepsy since her second birthday and muscle weakness of the face, shoulder girdle, and upper arms was observed from the age of four years. Lack of facial expression was noticed in the second patient from the age of one. She developed epilepsy at the age of nine years and weakness of her lower limbs progressed from the age of ten, eventually becoming wheelchair dependent by the age of fourteen. She also had moderate sensorineural hearing loss and tongue atrophy. Genetic analysis confirmed the presence of a 10 kb *Eco* RI deletion fragment on chromosome 4q35 in both individuals.

Van der Kooi *et al.* (2000) reported six sporadic patients with symptoms and signs that initially caused confusion upon clinical examination. FSHD was subsequently confirmed via genetic analysis. Three patients presented with foot extensor weakness, one with thigh weakness, one with calf muscle weakness resulting in an inability to walk on his toes, and the other with mild shoulder symptoms, such as tiredness and muscle pain. None of the patients had visible facial weakness upon initial examination. An expert physical examination, however, indicated the presence of an abnormality in the facial expressions of the patients and abnormal shoulder posture or scapular winging when lifting the arms. This study highlights the need for expert clinical examination of FSHD patients. Moreover it is preferable if one clinician can investigate all patients included in a study to ensure consistency in diagnosis of this complex clinical phenotype.

A patient reported by Felice *et al.* (2001) was diagnosed with late-onset autosomal dominant distal myopathy, and presented with progressive bilateral foot drop, and developed difficulties when climbing stairs. Other problems included mild, late-onset

sensorineural hearing loss, occasional heart palpitations, cataracts and arthritis. Neurological examination revealed mild eye closure weakness, mild hearing loss, anterior foreleg muscle atrophy and a steppage gait pattern. There were no visible signs of scapular winging or humeral atrophy. FSHD was, however, confirmed on a molecular level.

The studies of Miura *et al.* (1998), Van der Kooi *et al.* (2000) and Felice *et al.* (2001) confirm the clinical variability of FSHD, and strengthen the need for DNA analysis, as well as examination by a clinician with experience in the clinical diagnosis of this clinically heterogeneous disorder.

2.1.1.2.1 Infantile FSHD

Infantile FSHD, also known as early-onset FSHD, is the most severe form of this disorder (McGarry *et al.,* 1983). Individuals affected by this form of FSHD are generally symptomatic before the age of five, display signs and symptoms of shoulder girdle weakness before the age of ten and are wheelchair dependent by the age of nine or ten (McGarry *et al.,* 1983; Brouwer *et al.*, 1993). Brooke was the first to describe infantile FSHD as a special form of the disorder and even suggested a specific clinical course and mode of inheritance (Brouwer *et al.*, 1994).

This form of FSHD accounts for less than 5% of affected cases (Brouwer *et al.*, 1993). Early facial weakness is visible during the first two years of life, and progressive weakness of the shoulder causing scapular winging, upper-arm and foot-extensor muscle weakness, characterised by foot drop, is observed, followed later by development of pelvic girdle weakness presenting as lumbar lordosis (Korf *et al.,* 1985; Bailey *et al.,* 1986). The presence of high frequency hearing loss, retinal vascular abnormalities and mental retardation have also been reported in many of the infantile FSHD patients (Korf *et al.,* 1985; Fitzsimons *et al.,* 1987; Brouwer *et al.*, 1995).

Death as a result of FSHD was reported by McGarry *et al.* in 1983. This child died at the age of five after progressive weakness and recurrent pneumonias. Her creatine phosphokinase, lactate dehydrogenase and aldolase were all elevated four to ten fold on different occasions. There was no family history of any neurological or muscle disorders. Bailey *et al.* (1986) also reported the death of four individuals, all with onset in infancy,

before the age of 20. Sporadic and familial cases occur, but on average the sporadic cases seem to be more severely affected than the familial cases (Brouwer *et al.*, 1994). Brouwer *et al.* (1995) concluded that infantile FSHD does not differ clinically, or genetically, from adult onset FSHD.

Funakoshi *et al.* (1998) examined 140 Japanese FSHD patients from 91 unrelated families, of whom twenty patients were classified as early onset FSHD. Nine of the twenty patients had small (10-11 kb) *Eco* RI fragments, and a high frequency of epilepsy (44%) and mental retardation (89%) was observed. These authors concluded that FSHD patients with a large deletion in the FSHD region tend to have a higher chance of clinical phenotypes being associated with central nervous system (CNS) abnormalities. Okinaga *et al.* (1997) also reported the presence of two early-onset FSHD cases in the Japanese population. Both children were observed to sleep with their eyes slightly open and demonstrated an inability to smile before the age of one. FSHD was confirmed on the molecular level in these two children by the presence of deletion fragments (13 kb and a 15 kb).

2.1.1.3 Current treatments for FSHD

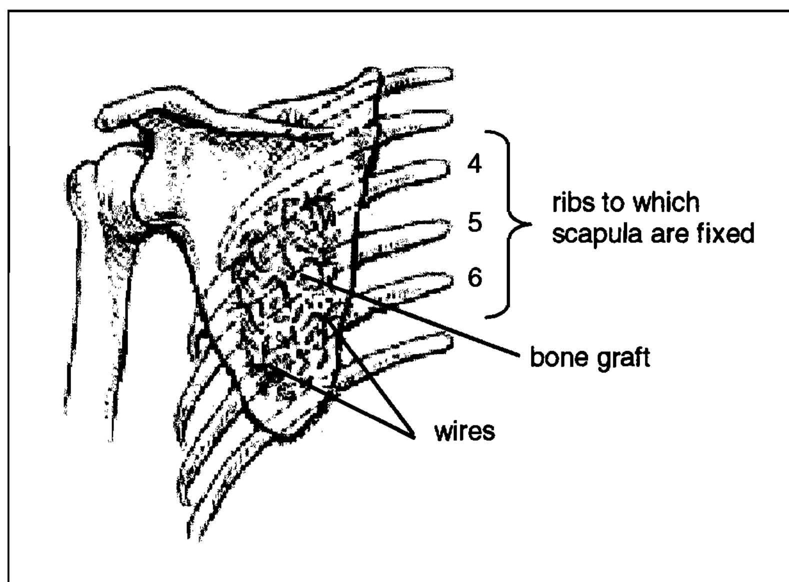
Individuals affected with any form of muscular dystrophy should consider a few general principles, which could assist them in adapting to the muscle weakness and to continue their everyday activities. Physical therapy, including light exercise, helps preserve flexibility and prevents contractures that result from immobility. Swimming is the preferred form of light exercise as it puts little strain on the muscles. A good diet is necessary to avoid unnecessary weight gain and reduces stress on the already weakened muscles.

Currently there is no treatment or cure for FSHD, but there are aids that can provide symptomatic relief. Individuals affected with FSHD have a relatively normal life expectancy and therefore sometimes require procedures that will provide prolonged relief.

In a procedure known as scapulothoracic arthrodesis, the scapula of some FSHD individuals has been fixed to the thoracic wall, to improve stability for activities involving the upper limbs. Bunch *et al.* (1993) fixed the scapula to the ribs in twelve FSHD patients and did re-examinations from three to twenty one years after the procedure. Instability of the scapula due to weakness of the muscles that stabilise the scapula results in an inability

to raise the arms forward or above the head. This group of FSHD patients did, however, have sufficient strength in the supraspinatus and deltoid muscle to abduct the arm if the scapula had been stabilised. Approximately 50% of the individuals examined by Bunch *et al.* (1993) had preserved deltoid function. The scapula is positioned over the seventh rib and fixation to ribs four, five and six is generally adequate, but as many as five ribs have been used as depicted in Figure 2.5. Holes are drilled through the scapula, and two one-millimetre-diameter stainless-steel wires are placed under each rib and through the drill-holes in the scapula. The wires are pulled tightly to compress the scapula against the graft. The resistance of the tightened wires secures the position of the scapula. The wire-ends are bent and buried within the infraspinatus muscles.

Figure 2.5: Scapular fixation in FSHD



Adapted from Bunch *et al.* (1993).

A solid fusion between the ribs and scapula is necessary, as the wire loops do not provide a rigid enough fixation, as they can slide on the ribs. Some form of external support is therefore required to ensure solid fusion. For example, a shoulder spica, which holds the arm abducted in a salute position for two months. Subsequently the arm should be placed in a figure-of-eight dressing until radiographs show a solid fusion. The twelve patients described by Bunch *et al.* (1993) all obtained solid fusion, and all of them experienced a more stable shoulder while they were carrying and lifting objects. All but one patient was capable of forward flexion and abduction to 90 degrees or more. These patients had an average of 30 degree flexion preoperatively, which increased to 65-125 degrees postoperatively. Bunch *et al.* (1993) favour this method as it is technically relatively easy and is associated with few complications.

Andrews *et al.* (1998) also performed scapulothoracic arthrodesis in six FSHD patients. In general the range of abduction and flexion increased and all the patients could continue their work after recovery. The scapula was clinically and radiologically fused to the chest wall in all the patients. Scapulothoracic arthrodesis was also performed by Letournel *et al.*

(1990) in fifteen patients. Flexion increased on average by 33 degrees and abduction by 25 degrees and, upon a sixty-nine month follow-up, the results had not deteriorated. Scapulothoracic arthrodesis offers a good long-term benefit to FSHD patients and also contributes to an increase in quality of life.

Steroids and other anti-inflammatory drugs are generally reserved for the treatment of the inflammatory myopathies, eg. polymyositis. There is, however, not a large amount of literature available on the use of steroids in FSHD. Munsat *et al.* (1972) found that creatine kinase (CK) levels dropped and symptoms improved in a small number of FSHD patients. In a more recent study by Tawil *et al.* (1997), eight FSHD patients were treated with prednisone for 12 weeks. There were, however, no significant changes in the muscle strength or mass. It was therefore concluded that treatment with prednisone had no benefit for FSHD patients, however, the effect on disease progression could not be evaluated over this short period of time.

Kissel *et al.* (1998) investigated the effect of albuterol, a β_2-agonist, in FSHD patients. β_2-agonists have been shown to induce satellite cell proliferation, increase muscle protein production, inhibit muscle proteolysis and retard the loss of muscle mass due to muscle injury, denervation, disuse, steroid atrophy, malnutrition, tumor, sepsis, and surgery. Kissel *et al.* (1998) treated fifteen FSHD patients for three months and reported improved muscle mass and overall improvement of 12% in strength. The results were encouraging and the effects of albuterol were further evaluated in a larger, randomised, double-blind, placebo-controlled trial by these authors. From this study, Kissel *et al.* (2001) reported that although treatment with albuterol for one year did not improve global strength or function in patients affected with FSHD, a significant increase in muscle mass and grip strength was observed. This indicated that albuterol does have some anabolic effect in the treated patients. The authors are, however, not sure as to why the increase in muscle mass did not translate into increased strength.

Bushby *et al.* (1998) reported four patients with FSHD in whom pain was the most distinctive symptom. Treatment of the pain was difficult, since there was a poor response to conventional anti-inflammatory therapy and eventually morphine was prescribed for one of the patients. The other three patients experienced improvement from swimming. All patients complained of a feeling of frustration together with depression and irritability. Two of the patients were treated with antidepressants and an improvement in their mood and partial improvement in their pain control was reported.

2.1.2 GENETIC ASPECTS OF FSHD

FSHD is an autosomal dominant, neuromuscular disorder and has a prevalence of 1 in 20,000 individuals in the European population and a prevalence of 1 in 14,763 individuals in Utah, U.S.A. (Walton, 1955; 1956; Lunt 1989a; Flanigan *et al.,* 2001). A higher prevalence was reported for the Utah population, as the entire geographical region contains only one muscular dystrophy referral centre, which examines almost all the reported cases within this region. It is also known that the Utah population consists of very large families. A relatively high proportion of new mutations (10 to 30%) have also been reported in FSHD (Padberg 1982; Padberg *et al.,* 1995b; Zatz *et al.,* 1995; Tawil *et al.,* 1996; Lunt 1998; Zatz *et al.,* 1998).

2.1.2.1 Linkage of FSHD to chromosome 4q35

Possible linkage was observed for the Gm locus encoding the constant region of the heavy chains of the IgG immunoglobulins (Padberg *et al.,* 1984). Gm was known to be located on the long arm of chromosome 14 (near 14q32). A subsequent study by Padberg in 1988, utilising DNA-probe D14S1, confirmed the linkage between Gm and D14S1, but no linkage was observed between FSHD and the D14S1 loci. This excluded the FSHD locus from the distal part of chromosome 14.

An international collaboration of research groups involved in the linkage analysis of FSHD was organised to reduce unnecessary duplication through the pooling of linkage data and to construct an exclusion map. Four groups consisting of 20 participants contributed linkage data: Padberg and Frants (Leiden); Upadhyaya, Sarfarazi, Lunt and Noades (Cardiff, Manchester, and London); Lucotte (Paris), and Pericak-Vance, Siddique and Shaw (Durham, NC and Manchester). A total of 57 markers on various autosomes were tested for possible linkage to FSHD, but no significant linkage between any marker and the disorder was detected (Lucotte *et al.,* 1989; Lunt 1989b; Sarfarazi *et al.,* 1989; Siddique *et al.,* 1989; Upadhyaya *et al.,* 1989). The overall data indicated that the likelihood for the location of the FSHD gene to be on chromosome 11 was 6.47 times more than being on any other chromosome. Other possible chromosomes were 19 (likelihood of 2.67) and 5 (likelihood of 2.22). Insufficient data was available for chromosomes 3, 10 and 15 and therefore remained unexcluded. Eventually participants of the international consortium

excluded almost 80% of the genome (Sarfarazi *et al.,* 1989). Jacobsen *et al.* (1990) excluded chromosomes 1, 2, 5, 7, 10 and 16 by using a panel of restriction fragment length polymorphism (RFLP) markers that were evenly spaced at approximately 20 centimorgan (cM) intervals.

Linkage was subsequently established in 1990 between the FSHD phenotype and the Mfd22 short tandem repeat polymorphism (STRP) marker at locus D4S171 in the sub-telomeric region of chromosome 4q35 (Wijmenga *et al.*, 1990). Ten multigeneration Dutch families consisting of 69 affected, 58 unaffected sibs and 25 spouses were investigated. Sixty microsatellite loci were analysed for linkage to FSHD. A maximum lod score of 6.34, at a theta value of 0.13, was observed for marker Mfd22. Only one of the ten families was uninformative for this marker. The microsatellite marker Mfd22 had previously been assigned to chromosome 4, with the use of a somatic cell hybrid panel (Weber and May, 1990).

Upadhyaya *et al.* (1990) confirmed the location of the FSHD locus to the subtelomeric region of chromosome 4, through the identification of a flanking marker, more closely linked to the FSHD locus. This variable number of tandem repeat (VNTR) marker pH30, at the D4S139 locus, was demonstrated to be tightly linked to the FSHD locus. The D4S139 locus was mapped to the distal portion of the long arm of chromosome 4 by *in situ* hybridisation (Milner *et al.*, 1989). Wijmenga *et al.* (1991) also mapped D4S139 to chromosome 4q35-qter by *in situ* hybridisation. It was thus possible to construct a primary map of the area surrounding the FSHD locus with D4S171 and D4S139, to estimate the most likely position of the gene. Several recombination events in ten families favoured the following locus order: D4S171-D4S139-FSHD-TEL. This paved the way for presymptomatic and prenatal diagnosis in an independent linkage panel of 24 families with FSHD from Great Britain that were also reported to be closely linked to the D4S139 locus (Upadhyaya *et al.,* 1991).

Members of the International FSHD Consortium for linkage analysis of the FSHD gene have pooled data in an attempt to determine the precise location of the FSHD gene as well as the order of four DNA markers on 4q35. Two additional markers (F11 and D4S163) to those used by Wijmenga in 1991 were utilised to form a linkage group consisting of four polymorphic loci in the area of the FSHD locus (D4S171, F11, D4S163 and D4S139) and covering a total distance of ca. 20 cM. Recombinants have been identified with these four

i 16609539X
b15465366

markers, which allowed the order of the FSHD locus and marker loci to be determined. The most likely locus order and the relative position of the FSHD gene was thus: CEN-D4S171-F11-D4S187-D4S163-D4S139-FSHD-TEL (Gilbert *et al.,* 1992; Mathews *et al.,* 1992; Mills *et al.,* 1992; Sarfarazi *et al.,* 1992; Upadhyaya *et al.,* 1992; Weiffenbach *et al.,* 1992a; Wijmenga *et al.,* 1992a, Winokur *et al.,* 1993). Identification of flanking markers in the 4q35 region contributed towards establishing a reliable diagnostic test for FSHD and will eventually facilitate in achieving the ultimate goal of cloning the FSHD gene.

2.1.2.2 The FSHD locus on chromosome 4q35

Wijmenga *et al.* (1992b) isolated cosmid clone 13E in a search for homeobox genes. The clone was mapped to 4q35, distal to the D4S139 locus through *in situ* hybridisation. Cosmid 13E was subcloned and a 0.8 kb probe, designated p13E-11, was obtained. Probe p13E-11, at locus D4F104S1 (also known as D4S810), was observed to detect a DNA rearrangement in individuals affected with FSHD. Fragment sizes between 14 to 28 kb were observed to segregate in ten Dutch FSHD families upon *Eco* RI digestion. The DNA rearrangements detected by probe p13E-11 were thus identified to play a role in the aetiology of FSHD.

Restriction mapping of cosmid 13E indicated that the *Eco* RI fragment detected by p13E-11 contained 3.3 kb tandem repeats with each 3.3 kb repeat flanked by *Kpn* I sites as illustrated in Figure 2.6 and Table D.1 (Wijmenga *et al.*, 1992b). The human genome contains hundreds of copies of this 3.3 kb family of tandem repeats in regions associated with heterochromatin. Different members of this long interspersed nuclear element (LINE) repeat family have been found on the short arms of all the acrocentric chromosomes, in the heterochromatic regions adjacent to the ribosomal DNA gene clusters, and on chromosomes 1, 2, 3, 4, 10, 18 and Y (Lyle *et al.,* 1995; Winokur *et al.,* 1994; Winokur *et al.,* 1996).

Hewitt *et al.* (1994) determined the sequence of the 3.3 kb repeat units and reported the presence of a double homeobox, and two repetitive sequences, namely *Lsau* and *hhspm3* within each repeat, as illustrated in Figure 2.6 and Table D.1. *Lsau* is a middle repetitive 68% GC rich element associated with β satellite DNA and found in heterochromatic regions of the genome while *hhspm3* is a low copy GC rich repeat element. Lee *et al.*

2.1.1.1.3 The truncal muscles

Padberg (1982) reported abdominal muscle weakness in 58% of the patients who were investigated. Weakness of the abdominal muscles (Figure 2.4) was observed to add to the pelvic tilt and the increased lumbar lordosis caused by the anterior convexity of the lumbar spine. The lumbar lordosis was also observed to become more severe when an individual was wheelchair dependent.

2.1.1.1.4 The lower extremities and the pelvic girdle muscles

Landouzy and Déjérine (1885) and Chyatte *et al.* (1966) observed the early weakness of the anterior tibial muscles in FSHD patients. Weakness of the foot extensors (Figure 2.4) interferes with walking, resulting in a steppage gait, referring to the pattern of how an individual walks, and an inability to run (Padberg, 1982). Steppage gait is characterised by the dropping of the foot, where the foot hangs with the toes pointing down, causing the toes to touch the ground while walking. Patients tend to trip easily over small objects, falling forward onto their knees. If pelvic girdle weakness develops, a waddling gait characterised by a distinctive duck-like walk will be visible and gradual rising from a chair or climbing stairs becomes increasingly difficult.

2.1.1.1.5 Asymmetry of muscle involvement

Distinct asymmetry of muscle involvement is an important and common feature of FSHD (Padberg, 1982). The asymmetry can be present in the facial as well as in the shoulder girdle muscles, and the extremities. No correlation was observed between right or left handed individuals, the side of muscle involvement, or the severity of affected muscles (Tawil *et al.*, 1994).

2.1.1.1.6 Extramuscular involvement

Sensorineural deafness and retinal vascular abnormalities have been described by many authors to be associated with FSHD (Gieron *et al.*, 1985; Gurwin *et al.*, 1985; Korf *et al.*, 1985; Matsuzaka *et al.*, 1986; Voit *et al.*, 1986; Fitzsimons *et al.*, 1987; Yasukohchi *et al.*, 1988; Brouwer *et al.*, 1991; Pauleikhoff *et al.*, 1992; Padberg *et al.*, 1995a). High-frequency hearing loss has been observed in 50 to 64% of FSHD patients (Brouwer *et al.*, 1991; Padberg *et al.*, 1995a). Brouwer *et al.* (1991) observed a significant difference in

hearing level between 4,000 Hertz (Hz) and 6,000 Hz in FSHD patients when compared to controls. The hearing loss was also observed to be progressive and, with time, tends to involve lower frequencies of the spectrum. These authors observed that the severity of the hearing loss varied between individuals and that it was not age dependent.

Originally, retinal vascular abnormalities were mainly reported in severe cases with early-onset FSHD (Gurwin *et al.*, 1985). Fitzsimons *et al.* (1987) and Brouwer *et al.* (1993) observed exudative retinal vasculopathy, with capillary telangiectasis, microaneurysms, and capillary closure to be present in 50 to 75% of FSHD patients. This is characterised by damage or breakdown of the retina due to subretinal fluid accumulation and the capillaries being longer, wider and fewer in number than normal. No correlation between the severity of the muscular weakness, and the severity of the hearing loss or the retinal vasculopathy was observed. The pathogenic mechanism of hearing loss and retinal vasculopathy in FSHD is still unknown.

Cardiac and respiratory muscles are generally not affected in FSHD patients (Padberg 1982; Tawil *et al.,* 1998). Emery-Dreifuss muscular dystrophy, a phenotypically similar but genetically distinct disorder, is distinguished from FSHD on the clinical level by the presence of cardiac muscle involvement. Stevenson *et al.* (1990) provided the first evidence of cardiac involvement in FSHD, unfortunately, no molecular data was available to confirm the clinical diagnosis of FSHD. Laforet *et al.* (1998) reported the presence of cardiac involvement in genetically confirmed FSHD patients. This study included 100 patients of whom five had conduction defects or arrhythmia. It is therefore evident that patients with FSHD may have cardiac involvement, although to date it has only been observed in a small number of cases. Finsterer *et al.* (2000) investigated the presence of cardiac involvement (CI) in several myopathies, such as DMD, BMD, EDMD, FSHD, sarcoglycanopathies, myotonic dystrophies type 1 and 2 and mitochondrial myopathies, and observed CI to be present in all the disorders. CI can lead to symptoms such as: impulse generation defects, impulse conduction defects, thickened myocardium, left ventricular hypertrabeculation, dilatation of the cardiac cavities, intracardial thrombus formation, and heart failure with systolic and diastolic dysfunction. The above findings argue strongly in favour of FSHD patients being investigated for possible cardiac involvement.

Inflammatory cellular infiltrates are frequently (40 to 80%) observed in the skeletal muscle of FSHD patients. General histological changes in FSHD muscle are rather non-specific

with variations in fibre size and scattered small angulated fibres being the most commonly observed anomalies (Arahata *et al.*, 1995).

Involvement of the lingual muscle is considered to be one of the exclusion criteria of FSHD (Appendix C). However, Korf *et al.* (1985) reported four patients with tongue abnormalities, such as atrophic changes and movement disturbances. No findings on the genetic studies of these patients were, however, reported. Tongue atrophy was also reported by Yamanaka *et al.* (2001) in seven Japanese patients, all belonging to a group of early-onset FSHD patients with small *Eco* RI deletion fragments (10 to 17 kb). Miura *et al.* (1998) also reported a female sporadic case with early-onset FSHD with tongue atrophy with a 10 kb *Eco* RI deletion fragment. These studies therefore suggest that individuals with FSHD can have involvement of the lingual muscles and that this clinical aspect should not be used as an exclusion criterium of FSHD. The occurrence of tongue abnormalities in FSHD, especially in early onset FSHD, should, however, be investigated in larger populations.

2.1.1.2 Clinical heterogeneity in the FSHD phenotype

Clinical diagnosis of the FSHD phenotype is complex due to extreme variability in various aspects of the phenotype, including its penetrance. The phenotypical expression varies in severity, rate of progression and the age of onset, ranging from almost asymptomatic individuals, to almost 20% of patients who will be wheelchair dependent by the age of 40 years or older (Lunt *et al.*, 1991). Variability is observed both between and within families, thus complicating genetic counselling. The penetrance of the FSHD gene was estimated to be different for various age groups as listed in Table 2.1. The life expectancy of individuals affected with FSHD is not shortened since cardiac and respiratory functions are generally unaffected (Padberg 1982).

Table 2.1: Penetrance of FSHD

<5%	0 to 4 years
21%	5 to 9 years
58%	10 to 14 years
86%	15 to 19 years
95%	20 years.

Adapted from Lunt *et al.* (1989a).

Tawil *et al.* (1993a) studied monozygotic twins with extreme clinical variability. One of the twins had progressive shoulder girdle weakness from the age of 10. Upon examination, at the age of 27, severe facial, scapular, peroneal and abdominal muscle weakness was observed. The asymptomatic

twin brother showed only mild weakness of the orbicularis oculi muscles. Zygosity was first determined through red cell phenotyping as well as human leukocyte antigen (HLA) typing. High-resolution cytogenetic analysis was performed to exclude any cytogenetic abnormalities. Five markers (D4S139, D4S163, D4S171, D4S130 and D4F35S1), closely linked to FSHD, were also studied to provide further evidence of monozygosity and to exclude the possibility that submicroscopic rearrangements of distal 4q had occurred. One possible explanation for the observed variability is the presence of a somatic mutation in one of the twins. Another explanation is that both twins have FSHD and that their phenotypes represent an extreme case of variability of expression within the family. The second was the most likely explanation and, therefore, further molecular studies were necessary to confirm the presence of a deletion fragment in both individuals. Tawil *et al.* (1993b) performed the additional studies and reported the presence of a unique 4q35 DNA rearrangement in the affected individual. Results from the second study confirmed that in this particular case, the discordance was probably due to a *de novo* postzygotic mutation after the twinning process had occurred.

Griggs *et al.* (1995) re-examined the monozygotic twins discordant for FSHD, first examined by Tawil in 1993, as well as two additional sets of monozygotic twins concordant for FSHD. The twins concordant for FSHD were similarly affected in terms of age of onset, overall degree of disability, and quantitative tests of muscle, but extreme differences in the symmetry of the muscle involved were observed. The inheritance pattern of the discordant twins could not be established, since sufficient symptomatic family members were not available for examination. Upon molecular analysis, deletion fragments were observed in the two twin pairs concordant for FSHD. A deletion fragment was also present in the affected individual of the other twin pair. No deletion fragment was, however, present in the twin with the mild facial weakness. These authors also concluded that the *de novo* postzygotic mutation or mitotic crossover probably occurred during the twinning process. Based on the results from the concordant twin pairs the authors suggested that the deletion fragment determines the age of onset and severity of the disorder, and that the asymmetries observed in FSHD patients are due to other factors, which may be for instance, environmental.

Tupler *et al.* (1998) described two monozygotic male twins affected by FSHD, carrying an identical *de novo Eco* RI deletion fragment. Neurological examination of their parents was normal, with no sign of muscular dystrophy. Haplotype analysis identified that the *de novo* rearrangement on chromosome 4q had occurred in paternal gametogenesis or

postzygotically in the paternal chromosome 4 before twinning. The genetic identity of the two twins did however not concur with their respective clinical phenotypes. One twin was severely affected and the other one was almost asymptomatic. The medical history was the same for the two brothers, except for an antirabies vaccination performed at the age of five in the more severely affected twin. Tupler *et al.* (1998) hypothesised that the vaccination might have triggered an inflammatory immune reaction, which contributed to the more severe phenotype. The pathogenesis of this external factor with relation to the severe FSHD phenotype has, however, not been elucidated.

Miura *et al.* (1998) reported two unrelated, severely affected female sporadic cases diagnosed with early-onset scapulohumeral muscular dystrophy with mental impairment (intelligence quotients (IQs) ranging from 33 to 45) and epilepsy. One patient suffered from epilepsy since her second birthday and muscle weakness of the face, shoulder girdle, and upper arms was observed from the age of four years. Lack of facial expression was noticed in the second patient from the age of one. She developed epilepsy at the age of nine years and weakness of her lower limbs progressed from the age of ten, eventually becoming wheelchair dependent by the age of fourteen. She also had moderate sensorineural hearing loss and tongue atrophy. Genetic analysis confirmed the presence of a 10 kb *Eco* RI deletion fragment on chromosome 4q35 in both individuals.

Van der Kooi *et al.* (2000) reported six sporadic patients with symptoms and signs that initially caused confusion upon clinical examination. FSHD was subsequently confirmed via genetic analysis. Three patients presented with foot extensor weakness, one with thigh weakness, one with calf muscle weakness resulting in an inability to walk on his toes, and the other with mild shoulder symptoms, such as tiredness and muscle pain. None of the patients had visible facial weakness upon initial examination. An expert physical examination, however, indicated the presence of an abnormality in the facial expressions of the patients and abnormal shoulder posture or scapular winging when lifting the arms. This study highlights the need for expert clinical examination of FSHD patients. Moreover it is preferable if one clinician can investigate all patients included in a study to ensure consistency in diagnosis of this complex clinical phenotype.

A patient reported by Felice *et al.* (2001) was diagnosed with late-onset autosomal dominant distal myopathy, and presented with progressive bilateral foot drop, and developed difficulties when climbing stairs. Other problems included mild, late-onset

sensorineural hearing loss, occasional heart palpitations, cataracts and arthritis. Neurological examination revealed mild eye closure weakness, mild hearing loss, anterior foreleg muscle atrophy and a steppage gait pattern. There were no visible signs of scapular winging or humeral atrophy. FSHD was, however, confirmed on a molecular level.

The studies of Miura *et al.* (1998), Van der Kooi *et al.* (2000) and Felice *et al.* (2001) confirm the clinical variability of FSHD, and strengthen the need for DNA analysis, as well as examination by a clinician with experience in the clinical diagnosis of this clinically heterogeneous disorder.

2.1.1.2.1 Infantile FSHD

Infantile FSHD, also known as early-onset FSHD, is the most severe form of this disorder (McGarry *et al.,* 1983). Individuals affected by this form of FSHD are generally symptomatic before the age of five, display signs and symptoms of shoulder girdle weakness before the age of ten and are wheelchair dependent by the age of nine or ten (McGarry *et al.,* 1983; Brouwer *et al.*, 1993). Brooke was the first to describe infantile FSHD as a special form of the disorder and even suggested a specific clinical course and mode of inheritance (Brouwer *et al.*, 1994).

This form of FSHD accounts for less than 5% of affected cases (Brouwer *et al.*, 1993). Early facial weakness is visible during the first two years of life, and progressive weakness of the shoulder causing scapular winging, upper-arm and foot-extensor muscle weakness, characterised by foot drop, is observed, followed later by development of pelvic girdle weakness presenting as lumbar lordosis (Korf *et al.,* 1985; Bailey *et al.,* 1986). The presence of high frequency hearing loss, retinal vascular abnormalities and mental retardation have also been reported in many of the infantile FSHD patients (Korf *et al.,* 1985; Fitzsimons *et al.,* 1987; Brouwer *et al.*, 1995).

Death as a result of FSHD was reported by McGarry *et al.* in 1983. This child died at the age of five after progressive weakness and recurrent pneumonias. Her creatine phosphokinase, lactate dehydrogenase and aldolase were all elevated four to ten fold on different occasions. There was no family history of any neurological or muscle disorders. Bailey *et al.* (1986) also reported the death of four individuals, all with onset in infancy,

before the age of 20. Sporadic and familial cases occur, but on average the sporadic cases seem to be more severely affected than the familial cases (Brouwer *et al.*, 1994). Brouwer *et al.* (1995) concluded that infantile FSHD does not differ clinically, or genetically, from adult onset FSHD.

Funakoshi *et al.* (1998) examined 140 Japanese FSHD patients from 91 unrelated families, of whom twenty patients were classified as early onset FSHD. Nine of the twenty patients had small (10-11 kb) *Eco* RI fragments, and a high frequency of epilepsy (44%) and mental retardation (89%) was observed. These authors concluded that FSHD patients with a large deletion in the FSHD region tend to have a higher chance of clinical phenotypes being associated with central nervous system (CNS) abnormalities. Okinaga *et al.* (1997) also reported the presence of two early-onset FSHD cases in the Japanese population. Both children were observed to sleep with their eyes slightly open and demonstrated an inability to smile before the age of one. FSHD was confirmed on the molecular level in these two children by the presence of deletion fragments (13 kb and a 15 kb).

2.1.1.3 Current treatments for FSHD

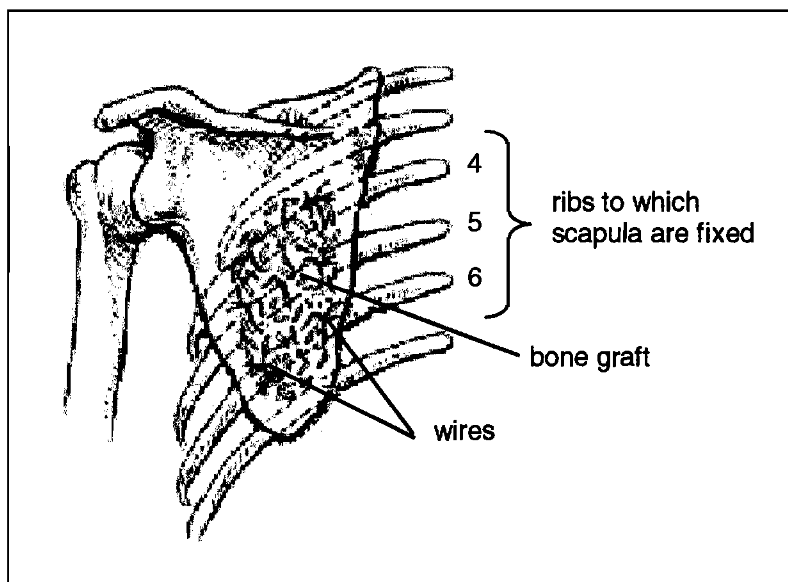
Individuals affected with any form of muscular dystrophy should consider a few general principles, which could assist them in adapting to the muscle weakness and to continue their everyday activities. Physical therapy, including light exercise, helps preserve flexibility and prevents contractures that result from immobility. Swimming is the preferred form of light exercise as it puts little strain on the muscles. A good diet is necessary to avoid unnecessary weight gain and reduces stress on the already weakened muscles.

Currently there is no treatment or cure for FSHD, but there are aids that can provide symptomatic relief. Individuals affected with FSHD have a relatively normal life expectancy and therefore sometimes require procedures that will provide prolonged relief.

In a procedure known as scapulothoracic arthrodesis, the scapula of some FSHD individuals has been fixed to the thoracic wall, to improve stability for activities involving the upper limbs. Bunch *et al.* (1993) fixed the scapula to the ribs in twelve FSHD patients and did re-examinations from three to twenty one years after the procedure. Instability of the scapula due to weakness of the muscles that stabilise the scapula results in an inability

to raise the arms forward or above the head. This group of FSHD patients did, however, have sufficient strength in the supraspinatus and deltoid muscle to abduct the arm if the scapula had been stabilised. Approximately 50% of the individuals examined by Bunch *et al.* (1993) had preserved deltoid function. The scapula is positioned over the seventh rib and fixation to ribs four, five and six is generally adequate, but as many as five ribs have been used as depicted in Figure 2.5. Holes are drilled through the scapula, and two one-millimetre-diameter stainless-steel wires are placed under each rib and through the drill-holes in the scapula. The wires are pulled tightly to compress the scapula against the graft. The resistance of the tightened wires secures the position of the scapula. The wire-ends are bent and buried within the infraspinatus muscles.

Figure 2.5: Scapular fixation in FSHD



Adapted from Bunch *et al.* (1993).

A solid fusion between the ribs and scapula is necessary, as the wire loops do not provide a rigid enough fixation, as they can slide on the ribs. Some form of external support is therefore required to ensure solid fusion. For example, a shoulder spica, which holds the arm abducted in a salute position for two months. Subsequently the arm should be placed in a figure-of-eight dressing until radiographs show a solid fusion. The twelve patients described by Bunch *et al.* (1993) all obtained solid fusion, and all of them experienced a more stable shoulder while they were carrying and lifting objects. All but one patient was capable of forward flexion and abduction to 90 degrees or more. These patients had an average of 30 degree flexion preoperatively, which increased to 65-125 degrees postoperatively. Bunch *et al.* (1993) favour this method as it is technically relatively easy and is associated with few complications.

Andrews *et al.* (1998) also performed scapulothoracic arthrodesis in six FSHD patients. In general the range of abduction and flexion increased and all the patients could continue their work after recovery. The scapula was clinically and radiologically fused to the chest wall in all the patients. Scapulothoracic arthrodesis was also performed by Letournel *et al.*

(1990) in fifteen patients. Flexion increased on average by 33 degrees and abduction by 25 degrees and, upon a sixty-nine month follow-up, the results had not deteriorated. Scapulothoracic arthrodesis offers a good long-term benefit to FSHD patients and also contributes to an increase in quality of life.

Steroids and other anti-inflammatory drugs are generally reserved for the treatment of the inflammatory myopathies, eg. polymyositis. There is, however, not a large amount of literature available on the use of steroids in FSHD. Munsat *et al.* (1972) found that creatine kinase (CK) levels dropped and symptoms improved in a small number of FSHD patients. In a more recent study by Tawil *et al.* (1997), eight FSHD patients were treated with prednisone for 12 weeks. There were, however, no significant changes in the muscle strength or mass. It was therefore concluded that treatment with prednisone had no benefit for FSHD patients, however, the effect on disease progression could not be evaluated over this short period of time.

Kissel *et al.* (1998) investigated the effect of albuterol, a β_2-agonist, in FSHD patients. β_2-agonists have been shown to induce satellite cell proliferation, increase muscle protein production, inhibit muscle proteolysis and retard the loss of muscle mass due to muscle injury, denervation, disuse, steroid atrophy, malnutrition, tumor, sepsis, and surgery. Kissel *et al.* (1998) treated fifteen FSHD patients for three months and reported improved muscle mass and overall improvement of 12% in strength. The results were encouraging and the effects of albuterol were further evaluated in a larger, randomised, double-blind, placebo-controlled trial by these authors. From this study, Kissel *et al.* (2001) reported that although treatment with albuterol for one year did not improve global strength or function in patients affected with FSHD, a significant increase in muscle mass and grip strength was observed. This indicated that albuterol does have some anabolic effect in the treated patients. The authors are, however, not sure as to why the increase in muscle mass did not translate into increased strength.

Bushby *et al.* (1998) reported four patients with FSHD in whom pain was the most distinctive symptom. Treatment of the pain was difficult, since there was a poor response to conventional anti-inflammatory therapy and eventually morphine was prescribed for one of the patients. The other three patients experienced improvement from swimming. All patients complained of a feeling of frustration together with depression and irritability. Two of the patients were treated with antidepressants and an improvement in their mood and partial improvement in their pain control was reported.

2.1.2 GENETIC ASPECTS OF FSHD

FSHD is an autosomal dominant, neuromuscular disorder and has a prevalence of 1 in 20,000 individuals in the European population and a prevalence of 1 in 14,763 individuals in Utah, U.S.A. (Walton, 1955; 1956; Lunt 1989a; Flanigan *et al.,* 2001). A higher prevalence was reported for the Utah population, as the entire geographical region contains only one muscular dystrophy referral centre, which examines almost all the reported cases within this region. It is also known that the Utah population consists of very large families. A relatively high proportion of new mutations (10 to 30%) have also been reported in FSHD (Padberg 1982; Padberg *et al.,* 1995b; Zatz *et al.,* 1995; Tawil *et al.,* 1996; Lunt 1998; Zatz *et al.,* 1998).

2.1.2.1 Linkage of FSHD to chromosome 4q35

Possible linkage was observed for the Gm locus encoding the constant region of the heavy chains of the IgG immunoglobulins (Padberg *et al.,* 1984). Gm was known to be located on the long arm of chromosome 14 (near 14q32). A subsequent study by Padberg in 1988, utilising DNA-probe D14S1, confirmed the linkage between Gm and D14S1, but no linkage was observed between FSHD and the D14S1 loci. This excluded the FSHD locus from the distal part of chromosome 14.

An international collaboration of research groups involved in the linkage analysis of FSHD was organised to reduce unnecessary duplication through the pooling of linkage data and to construct an exclusion map. Four groups consisting of 20 participants contributed linkage data: Padberg and Frants (Leiden); Upadhyaya, Sarfarazi, Lunt and Noades (Cardiff, Manchester, and London); Lucotte (Paris), and Pericak-Vance, Siddique and Shaw (Durham, NC and Manchester). A total of 57 markers on various autosomes were tested for possible linkage to FSHD, but no significant linkage between any marker and the disorder was detected (Lucotte *et al.,* 1989; Lunt 1989b; Sarfarazi *et al.,* 1989; Siddique *et al.,* 1989; Upadhyaya *et al.,* 1989). The overall data indicated that the likelihood for the location of the FSHD gene to be on chromosome 11 was 6.47 times more than being on any other chromosome. Other possible chromosomes were 19 (likelihood of 2.67) and 5 (likelihood of 2.22). Insufficient data was available for chromosomes 3, 10 and 15 and therefore remained unexcluded. Eventually participants of the international consortium

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

excluded almost 80% of the genome (Sarfarazi *et al.,* 1989). Jacobsen *et al.* (1990) excluded chromosomes 1, 2, 5, 7, 10 and 16 by using a panel of restriction fragment length polymorphism (RFLP) markers that were evenly spaced at approximately 20 centimorgan (cM) intervals.

Linkage was subsequently established in 1990 between the FSHD phenotype and the Mfd22 short tandem repeat polymorphism (STRP) marker at locus D4S171 in the sub-telomeric region of chromosome 4q35 (Wijmenga *et al.*, 1990). Ten multigeneration Dutch families consisting of 69 affected, 58 unaffected sibs and 25 spouses were investigated. Sixty microsatellite loci were analysed for linkage to FSHD. A maximum lod score of 6.34, at a theta value of 0.13, was observed for marker Mfd22. Only one of the ten families was uninformative for this marker. The microsatellite marker Mfd22 had previously been assigned to chromosome 4, with the use of a somatic cell hybrid panel (Weber and May, 1990).

Upadhyaya *et al.* (1990) confirmed the location of the FSHD locus to the subtelomeric region of chromosome 4, through the identification of a flanking marker, more closely linked to the FSHD locus. This variable number of tandem repeat (VNTR) marker pH30, at the D4S139 locus, was demonstrated to be tightly linked to the FSHD locus. The D4S139 locus was mapped to the distal portion of the long arm of chromosome 4 by *in situ* hybridisation (Milner *et al.*, 1989). Wijmenga *et al.* (1991) also mapped D4S139 to chromosome 4q35-qter by *in situ* hybridisation. It was thus possible to construct a primary map of the area surrounding the FSHD locus with D4S171 and D4S139, to estimate the most likely position of the gene. Several recombination events in ten families favoured the following locus order: D4S171-D4S139-FSHD-TEL. This paved the way for presymptomatic and prenatal diagnosis in an independent linkage panel of 24 families with FSHD from Great Britain that were also reported to be closely linked to the D4S139 locus (Upadhyaya *et al.,* 1991).

Members of the International FSHD Consortium for linkage analysis of the FSHD gene have pooled data in an attempt to determine the precise location of the FSHD gene as well as the order of four DNA markers on 4q35. Two additional markers (F11 and D4S163) to those used by Wijmenga in 1991 were utilised to form a linkage group consisting of four polymorphic loci in the area of the FSHD locus (D4S171, F11, D4S163 and D4S139) and covering a total distance of ca. 20 cM. Recombinants have been identified with these four

i 1609 539X
b15465366

markers, which allowed the order of the FSHD locus and marker loci to be determined. The most likely locus order and the relative position of the FSHD gene was thus: CEN-D4S171-F11-D4S187-D4S163-D4S139-FSHD-TEL (Gilbert *et al.,* 1992; Mathews *et al.,* 1992; Mills *et al.,* 1992; Sarfarazi *et al.,* 1992; Upadhyaya *et al.,* 1992; Weiffenbach *et al.,* 1992a; Wijmenga *et al.,* 1992a, Winokur *et al.,* 1993). Identification of flanking markers in the 4q35 region contributed towards establishing a reliable diagnostic test for FSHD and will eventually facilitate in achieving the ultimate goal of cloning the FSHD gene.

2.1.2.2 The FSHD locus on chromosome 4q35

Wijmenga *et al.* (1992b) isolated cosmid clone 13E in a search for homeobox genes. The clone was mapped to 4q35, distal to the D4S139 locus through *in situ* hybridisation. Cosmid 13E was subcloned and a 0.8 kb probe, designated p13E-11, was obtained. Probe p13E-11, at locus D4F104S1 (also known as D4S810), was observed to detect a DNA rearrangement in individuals affected with FSHD. Fragment sizes between 14 to 28 kb were observed to segregate in ten Dutch FSHD families upon *Eco* RI digestion. The DNA rearrangements detected by probe p13E-11 were thus identified to play a role in the aetiology of FSHD.

Restriction mapping of cosmid 13E indicated that the *Eco* RI fragment detected by p13E-11 contained 3.3 kb tandem repeats with each 3.3 kb repeat flanked by *Kpn* I sites as illustrated in Figure 2.6 and Table D.1 (Wijmenga *et al.*, 1992b). The human genome contains hundreds of copies of this 3.3 kb family of tandem repeats in regions associated with heterochromatin. Different members of this long interspersed nuclear element (LINE) repeat family have been found on the short arms of all the acrocentric chromosomes, in the heterochromatic regions adjacent to the ribosomal DNA gene clusters, and on chromosomes 1, 2, 3, 4, 10, 18 and Y (Lyle *et al.,* 1995; Winokur *et al.,* 1994; Winokur *et al.,* 1996).

Hewitt *et al.* (1994) determined the sequence of the 3.3 kb repeat units and reported the presence of a double homeobox, and two repetitive sequences, namely *Lsau* and *hhspm3* within each repeat, as illustrated in Figure 2.6 and Table D.1. *Lsau* is a middle repetitive 68% GC rich element associated with β satellite DNA and found in heterochromatic regions of the genome while *hhspm3* is a low copy GC rich repeat element. Lee *et al.*

YACs were digested with several common and rare cutter restriction enzymes for the production of a fine restriction map around the locus D4S104S1, orientating the rearranged fragment detected by the probe p13E-11. Wijmenga *et al.* (1993a) and Weiffenbach *et al.* (1994) determined the order of the loci around the D4S104S1 locus as: CEN-D4F35S1-D4S1101-D4S104S1-D4Z4-TEL. The homeobox probe 9B6A was identified to map to each copy of the tandem repeat, therefore confirming the presence of a homeobox sequence in each copy of the repeat motif.

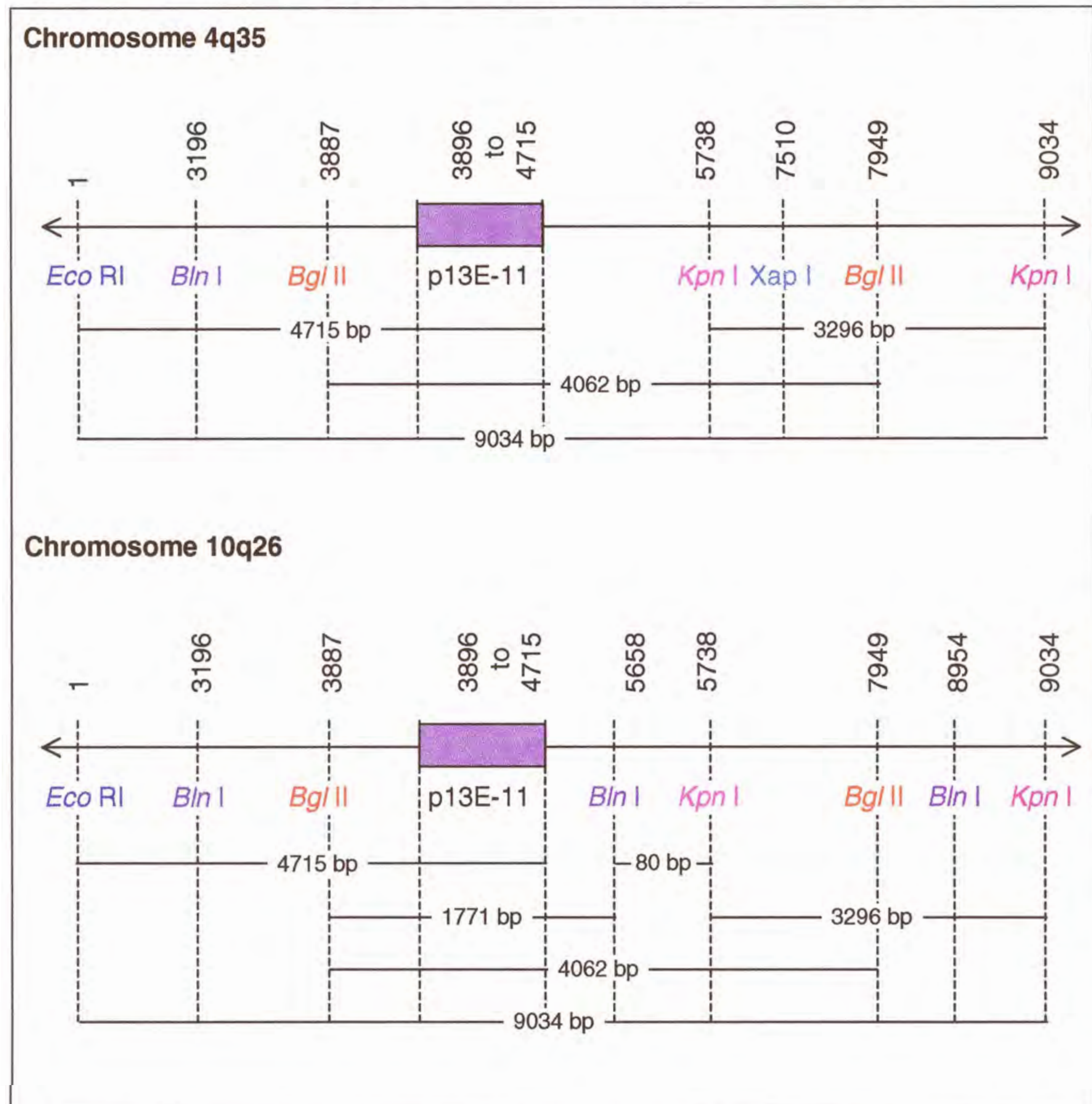
A decrease in the number of 3.3 kb tandem repeats was detected in individuals with FSHD (Van Deutekom *et al.,* 1993). The variability in fragment size was caused by the deletion of an integral number of the repeats. This was in contrast to three other muscular dystrophies, X-linked spinal and bulbar muscular atrophy, oculopharyngeal muscular dystrophy and myotonic dystrophy, in which an expansion of unstable repeats are the cause of the disorder (Lieberman and Fischbeck, 2000). FSHD negative individuals have 10-100 copies of the chromosome 4 repeat elements, corresponding to 35-300 kb fragments and patients with FSHD have less than 10 copies, thus resulting in deletion fragments which are less than 35 kb in size (Van Deutekom *et al.,* 1993).

Probe p13E-11 (D4F104S1) detects the rearranged *Eco* RI fragments. The interpretation of Southern blot analyses were, however, complicated since probe p13E-11 revealed two polymorphic loci as well as a 10 kb Y-specific fragment due to cross-hybridisation (Wijmenga *et al.,* 1993b and Weiffenbach *et al.,* 1993). Haplotype analysis could assign one of the loci to chromosome 4q35, but the location of the second locus remained unknown. Potential candidate regions were identified upon cross hybridisation of cosmid 13E using FISH, and included chromosomes 1q and 10q26, in addition to satellites of all the acrocentric chromosomes. The non-4q35 polymorphic fragment was subsequently shown to segregate with 10q telomeric microsatellite markers (Bakker *et al.,* 1995).

Deidda *et al.* (1995) also cloned a 13 kb non-4q35 fragment segregating in an FSHD Italian family and confirmed the localisation of the fragment to 10q26 via haplotype and *in situ* hybridisation. Restriction mapping of this region indicated that the 10q26 region contains a similar arrangement of *Kpn* I tandemly repeated units and flanking sequences than the FSHD region on 4q35. Comparison of sequences between the 4q35 and 10q26 fragments (Figure 2.7 and Table E.1), identified the presence of a unique chromosome 10 specific *Bln* I site, 80 nucleotides upstream of the *Kpn* I site (Deidda *et al.,* 1996).

Chromosome 10 derived repeat units are sensitive to *Bln* I, thus allowing *Bln* I to discriminate between chromosome 4 and chromosome 10 alleles. Cacurri *et al.* (1998) compared a total of 4 kb of the 4q35 sequences with the 10qter sequences and found the degree of sequence homology to be 98 to 100% between these two regions.

Figure 2.7: Restriction endonuclease maps of the 4q35 and 10q26 regions



This figure is not drawn to scale. Colours and nucleotide numbers correspond with those used to identify specific sequences in Table D.1 and Table E.1.

Upadhyaya *et al.* (1997) screened 200 control and 130 FSHD patients and indicated that double digestion with *Eco* RI and *Bln* I has a sensitivity of 95% and is therefore a valuable diagnostic test for FSHD. Orrell *et al.* (1999) also investigated 82 unrelated FSHD cases

to determine whether the *Eco* RI and *Bln* I double digestion would improve the molecular diagnosis of FSHD. Utilising the double digestion method a definitive molecular diagnosis could be made in all the affected individuals investigated. Orrell *et al.* (1999) therefore concluded that the combination of double digestion with *Eco* RI and *Bln* I, followed by pulsed field gel electrophoresis (PFGE) is the most reliable protocol for the molecular diagnosis of FSHD individuals.

Lemmers *et al.* (2000) identified a *Xap* I restriction endonuclease site on chromosome 4q35 within each 3.3 kb repeat unit (Figure 2.7; Table D.1 and Table E.1). *Xap* I displayed the opposite characteristic of *Bln* I, by uniquely digesting 4-type repeat units, leaving the 10-type units undigested. The combination of *Eco* RI, *Eco* RI/*Bln* I and *Eco* RI/*Xap* I digestions therefore allows the characterisation of each allele, even after translocation events (paragraph 2.1.2.2.1) between 4-type and 10-type repeats has occurred. The combination of these restriction endonucleases thus represents an optimal strategy for diagnosis of FSHD on the molecular level.

2.1.2.2.1 Translocation events between chromosomes 4q and 10q

Translocation events between the repeats on chromosome 4 and chromosome 10 could be demonstrated via the presence of *Bln* I sites within the *Kpn* I repeats from 10q26. The presence of 10-type repeats on chromosome 4, and *vice versa*, was observed in 160 independent Dutch familial or isolated cases (Van Deutekom *et al.*, 1996a). The frequency of these exchanges was determined via the screening of 50 unrelated control samples. A relatively high translocation frequency of 20% was observed, suggesting that it is unlikely that the FSHD gene resides within the repeated units at chromosome 4q35. Only short repeat arrays on chromosome 4 are pathogenic, irrespective of the origin of the repeat unit. A short array of 4-type or 10-type repeats on chromosome 4 therefore causes FSHD. Short repeat arrays located on chromosome 10 are, however, nonpathogenic (Cacurri *et al.*, 1998, Lemmers *et al.*, 1998).

Van der Maarel *et al.* (2000) indicated a numerical excess of 4-type repeats on chromosome 10. It was postulated that this excess is a significant, if not the major predisposing factor that gives rise to the FSHD-type deletion. A further hypothesis was formulated which stated that the FSHD deletion occurs mainly via somatic

interchromosomal gene conversion, in which the presence of a fully homologous repeat array on a non-homologous chromosome is a predisposing factor.

Van Overveld *et al.* (2000) examined the 3.3 kb D4Z4 repeat array configurations on chromosome 4 and its homologue on chromosome 10 to study the behaviour of these subtelomeric domains. PFGE was performed in 208 Dutch blood donors, and the alleles were assigned to their chromosomal origin based on their *Bln* I sensitivity and/or resistance. A standard pattern (4-type arrays on chromosome 4 and 10-type arrays on chromosome 10) was observed in 76% of the individuals. A non-standard configuration due to translocated repeat arrays has been observed in 21% of the individuals. Four-type repeat arrays on chromosome 10 occur as frequently (12%) as the reverse configuration (9%) of 10-type arrays on chromosome 4, but the composition of these translocated arrays differs significantly (Table 2.2).

Table 2.2: Configuration and composition of translocated arrays

Configuration of arrays			
76%	Standard configuration: 4-type arrays on chromosome 4 and 10-type arrays on chromosome 10		
21%	Non-standard configuration		
3%	Unpredictable configuration		
Composition of arrays			
10-type repeat arrays on chromosome 4		**4-type repeat arrays on chromosome 10**	
Homogeneous	Heterogeneous (hybrid arrays)	Homogeneous	Heterogeneous (hybrid arrays)
17%	33%	44%	10%

Adapted from Van Overveld *et al.* (2000).

The 10-type repeat arrays on chromosome 4 were more heterogeneous and the 4-type arrays on chromosome 10 were more homogeneous. This suggested a biological difference between these homologous repeat arrays, resulting in a preference of 4-type repeat arrays on chromosome 4 (Van Overveld *et al.*, 2000). It is known that repetitive sequences in the genome undergo concerted evolution. This process homogenises repetitive sequences and is thought to be important for the maintenance and integrity of each repeat unit. However, the high level of heterogeneity observed for the repeat units at the D4Z4 locus suggests that these loci escaped concerted evolution and evolved by inter- and intrachromosomal recombination. This may imply that the open reading frame,

present in each repeat unit, may have lost its function due to the rapid expansion of this repeat array (refer to paragraph 2.1.2.10.3).

An unexpected high frequency (3%) of 4-type repeat arrays shorter than 38 kb was also observed in this control population. These individuals were not included with the previously described 76% that harboured a classical standard configuration. If these short fragments reside on chromosome 4, they should be associated with FSHD. The authors proposed several possibilities for this unexpected finding:

- It is possible that the short array does not reside on chromosome 4. As reported by Van Deutekom *et al.* (1996a) translocations occur between chromosomes 4 and 10 in 20% of the Dutch population. The presence of 'double exchanged' alleles is therefore predicted to be ca. 1%.
- One of the individuals carried three 4-type repeat arrays. It is therefore possible that the short array may reside on chromosome 10.
- 30% of gene carriers are asymptomatic, and a correlation has been established between the residual repeat size and the age of onset and severity of FSHD. The array in three of the individuals was larger than 30 kb, and these individuals could therefore have subclinical characteristics or display non-penetrance.

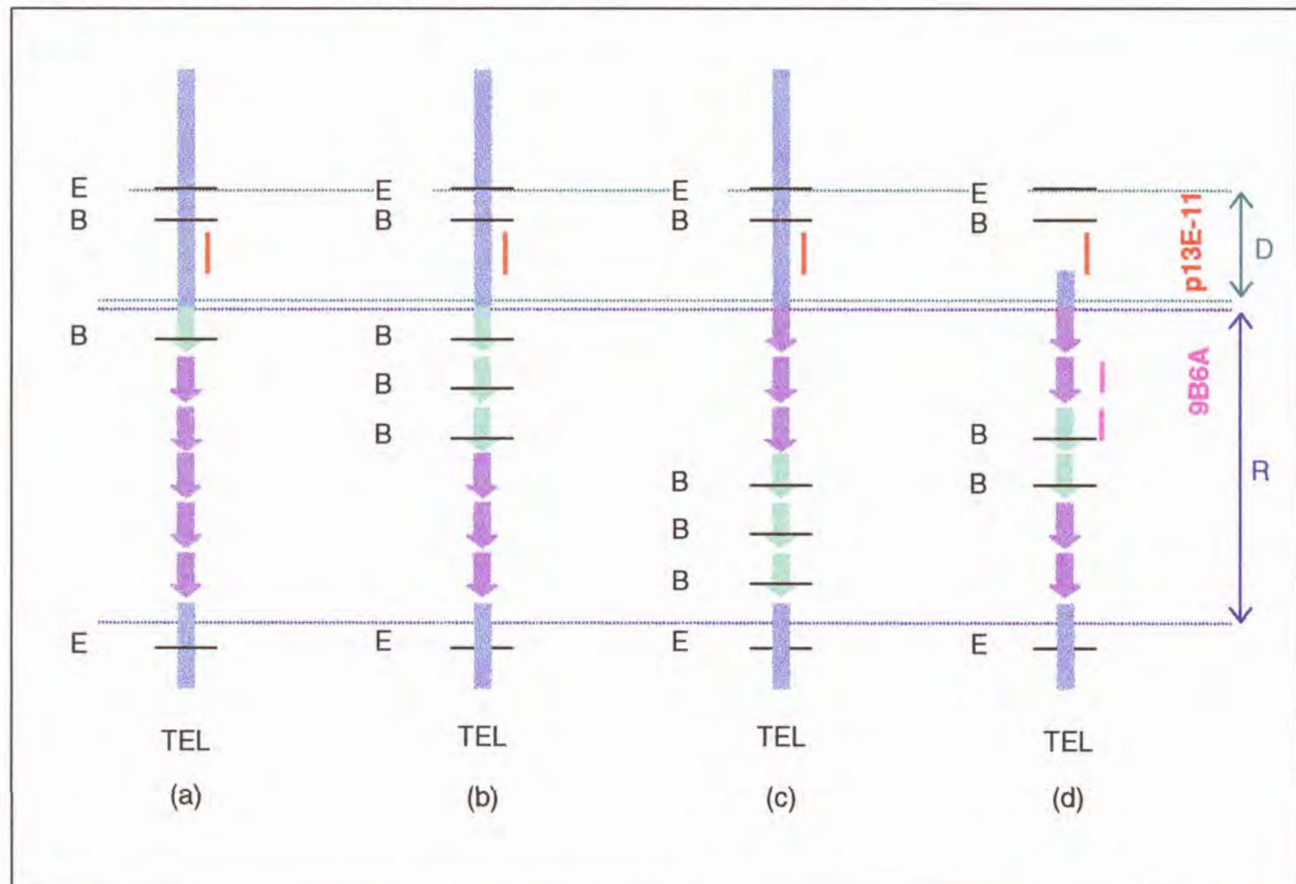
The presence of interchromosomal exchanges between the repeats on chromosome 4 and chromosome 10 has implications for both the specificity and sensitivity of diagnostic DNA testing of FSHD. The high frequency of 20% observed in the Dutch population therefore suggested the need to determine the percentage of interchromosomal exchanges in other populations as well.

2.1.2.2.2 Hybrid repeat arrays and deletion of p13E-11 hybridisation site

The molecular diagnosis of FSHD was further complicated by the identification of a deletion of the p13E-11 hybridisation site and hybrid fragments, comprising repeats of both chromosome 4 and 10 (Van Deutekom *et al.*, 1996a; Lemmers *et al.*, 1998). This was observed after hybridisation with probe 9B6A, which is complimentary to the repeat unit (D4Z4) itself. Individuals who have a deletion of the p13E-11 site will appear to have no small fragment on 4q35 when hybridised with p13E-11 as illustrated in Figure 2.8(d). Upon PFGE only 3 fragments will be observed, the fourth fragment will be visible after hybridisation with marker 9B6A, since marker 9B6A hybridises to the homeobox sequences within each repeat unit.

Hybrid repeat arrays, consisting of 10-type repeats attached to the distal end of a 4-type repeat array as depicted in Figure 2.8(c), could lead to a false positive diagnosis. The *Eco* RI/*Bln* I double digest will yield a small *Bln* I resistant fragment, but this will only correspond to the length of the 4-type repeats. The 10-type repeats will be digested with *Bln* I. On the other hand, hybrid repeat arrays, consisting of 4-type repeats attached to the distal end of a 10-type repeat array as displayed in Figure 2.8(a, b), could lead to false negative diagnosis. The *Eco* RI/*Bln* I double digest will show no deletion fragment, since the *Bln* I site within each chromosome 10-type repeat will lead to the digestion of these repeats. PFGE and the hybridisation with probe 9B6A will be able to detect the 4-type repeats on the distal end. To minimise the risk of false negative diagnosis due to hybrid repeat arrays or deletion of the p13E-11 recognition site the utilisation of PFGE, probe p13E-11 and probe 9B6A is essential in the molecular diagnosis of FSHD.

Figure 2.8: Schematic representation of hybrid chromosomes



Purple arrows represent 4-type repeat units, green arrows represent 10-type repeat units, a = Hybrid repeat arrays consisting of 4-type repeats attached to the distal end of a 10-type repeat array, b = Hybrid repeat arrays consisting of 4-type repeats attached to the distal end of 10-type repeat arrays, c = hybrid repeat arrays consisting of 10-type repeats attached to the distal end of a 4-type repeat arrays, d = hybrid repeat arrays with a deletion of the p13E-11 hybridisation site, B = *Bln* I restriction sites, D = region distal to repeat arrays, E = *Eco* RI restriction site, R = repeat units, TEL = telomere. The probe recognition sites for probe p13E-11 and probe 9B6A are indicated. Adapted from Lemmers *et al.* (1998).

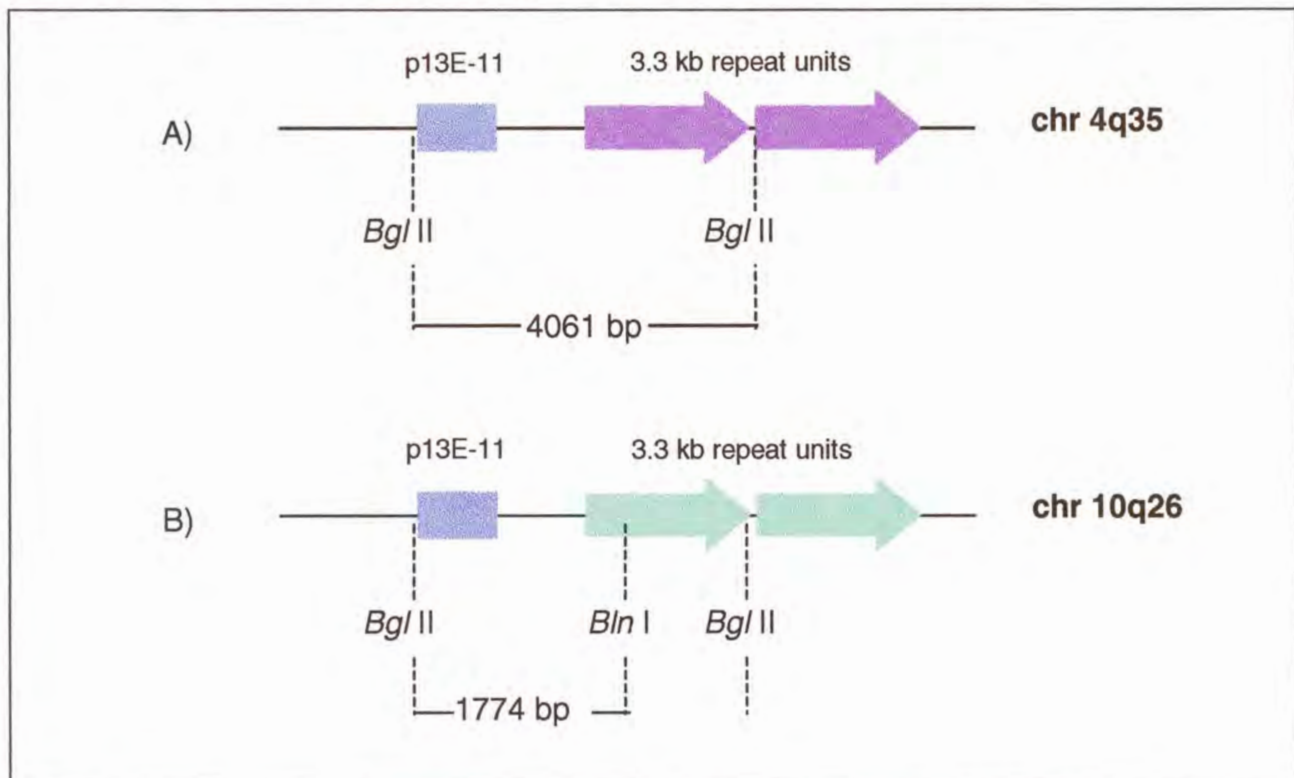
2.1.2.2.3 The *Bgl* II – *Bln* I dosage test

The diagnosis of FSHD with PFGE therefore has the advantage that all four alleles, as well as the presence of translocations, are visualised. The success of PFGE relies, however, on the quality of the aqueous DNA used and it is often difficult to detect alleles exceeding 200 kb. Also, many diagnostic laboratories do not have the facilities, or the expertise, to successfully carry out the PFGE analyses. Van der Maarel *et al.* (1999) therefore developed the *Bgl* II-*Bln* I dosage test (Figure 2.9) to provide a protocol to determine the ratio of 4-type to 10-type repeats. Deletions of the p13E-11 site (as described in paragraph 2.1.2.2.2) are also detected via the dosage test.

The *Bgl* II – *Bln* I dosage test utilises the *Bln* I polymorphism in the first D4Z4 repeat, and *Bgl* II instead of *Eco* RI to obtain a smaller fragment, to discriminate between chromosome 4 and chromosome 10 repeat units. A double digestion with *Bgl* II and *Bln* I will result in a 4,061 bp chromosome 4 specific fragment and a 1,774 bp chromosome 10 specific fragment, due to the extra *Bln* I site on chromosome 10. After hybridisation with p13E-11 the ratios between the signal intensities from chromosome 4 and 10 fragments should be equal (2:2) in the absence of translocation events. Individuals who carry one or three 4-type repeat arrays due to a translocation event should have ratios of 1:3 and 3:1 respectively. Individuals with only 4-type repeat arrays will show no hybridisation for the chromosome 10 fragment.

The dosage test will, however, fall short if the translocation between chromosome 4 and 10 occurred distal to the first polymorphic *Bln* I site or if an individual carries a complex rearrangement resulting in a hybrid repeat array (consisting of both chromosome 4 and 10 repeat units). Diagnosis of FSHD with PFGE using *Bln* I has a sensitivity of 96.5% and a specificity of >99%. In contrast, the sensitivity of the diagnosis utilising the conventional linear gel electrophoresis is 92% with a specificity of 99% (Van der Maarel *et al.*, 1999). If the dosage test is added to the conventional diagnosis the sensitivity and specificity will be improved to that of the PFGE diagnosis. The *Bgl* II – *Bln* I dosage test can also be utilised to study translocation events in large populations. Moreover it can be used in addition to the PFGE diagnosis to prevent false identification of deletions of the p13E-11 region due to the difficult visualisation of large fragment sizes (>200 kb) as a result of poor genomic DNA (gDNA) quality brought about by shearing of the gDNA prior to PFGE.

Figure 2.9: Schematic representation of the *Bgl* II – *Bln* I dosage test



A = Repeat array on chromosome 4, B = Repeat array on chromosome 10. Adapted from Van der Maarel *et al.* (1999).

Van der Maarel *et al.* (1999) observed via PFGE that the entire repeat array was translocated to the non-homologous chromosome in most of their cases. Only a small group of the translocations resulted in hybrid repeat units. The utilisation of the *Bgl* II-*Bln* I dosage test led to the conclusion that the translocations between chromosome 4 and 10 must have occurred proximal to the polymorphic *Bln* I site within the first repeat unit, in a recombination hotspot. The exact localisation of this translocation breakpoint proximal to the repeat arrays has implications for the localisation of the putative FSHD gene by narrowing down the candidate gene region. These authors also mentioned that recombination hotspots co-localise with open chromatin domains in yeast, such as promoter or coding sequences. A similar mechanism may play a role in recombination in vertebrates and the putative recombination hotspot proximal to the D4Z4 array may indeed indicate a new FSHD candidate gene locus.

2.1.2.3 Somatic and germline mosaicism

Somatic mosaicism is indicated by a fifth fragment upon PFGE analysis and hybridisation with probe p13E-11. This phenomenon has been reported by many authors (Griggs *et al.*, 1993; Weiffenbach *et al.*, 1993; Bakker *et al.*, 1995; Upadhyaya *et al.*, 1995; Zatz *et al.*,

1995; Köhler *et al.*, 1996; Bakker *et al.*, 1996;). Van der Maarel *et al.* (2000) investigated 35 sporadic FSHD families for which the chromosomal origin and size of each of the four repeat arrays (chromosomes 4 and 10) in both patients and parents could be determined for 23 families. The remaining 12 families had incomplete information and the authors stated that the "DNA quality was not sufficient" to assign all the alleles for one of the individuals. All the patients had a *Bln* I resistant deletion fragment of <35 kb (9 repeats), confirming their FSHD diagnosis. The deletion fragments ranged from 8 kb (1 repeat unit and flanking sequences) to 25 kb (6 repeats). One patient, however, inherited a 25 kb *Bln* I resistant deletion fragment from his clinically asymptomatic father and this is therefore a familial and not a sporadic case.

Van der Maarel *et al.* (2000) observed 14 cases of somatic mosaicism. Mosaicism for the disease allele was observed in three of the unaffected parents (two mothers and one father) from the 23 fully informative families. Five patients of the 23 families (4 males and 1 female) were mosaic. In the 12 families that had incomplete information, mosaicism was observed in one father, one mother, and four patients (3 males and 1 female). In 73% of mosaic individuals, the smallest D4Z4 allele was reduced to an FSHD-sized deletion fragment. Somatic mosaicism was therefore observed in 40% of cases, either in the patient or in an asymptomatic parent. The degree of mosaicism could, however, be much more, since complete allele information of only 23 of the 35 *de novo* families was available. It was therefore postulated that this high degree of somatic mosaicism implies that the deletion event is mainly mitotic.

As described in paragraph 2.1.2.2.1, interchromosomal repeat translocations occur between the homologous repeats from chromosomes 4 and 10. Van der Maarel *et al.* (2000) observed one or more 4-type repeat arrays on chromosome 10 in 46% of the mosaic individuals. This type of repeat array is, however, only present in 10% of the Dutch population (Van Deutekom *et al.*, 1996a). The reverse configuration was also present in 10% of the population but was not observed in the mosaic individuals. In mosaic individuals, the presence of 4-type arrays on chromosome 10, is therefore increased by almost five times. Somatic mosaicism was also observed in 3% of normal control individuals from the Dutch population (4 males and 2 females). One of the mosaic individuals carried an extra 4-type repeat array on chromosome 10 (Van Overveld *et al.*, 2000).

The presence of somatic mosaicism can only be observed upon PFGE, which emphasises the need to utilise PFGE in the diagnosis of FSHD. The presence of somatic mosaicism in FSHD patients might facilitate the elucidation of some aspects of the clinical variability of this disorder.

2.1.2.4 Anticipation

Myotonic dystrophy was the first disorder identified to display anticipation, which is characterised by an earlier onset with increase in the severity of clinical symptoms in subsequent generations. Here, the disease causing mutation is an unstable expansion of the CTG repeat in the myotonin-protein kinase (DMPK) gene on chromosome 19. Anticipation is caused by an increase in the number of trinucleotide repeats. The age of onset and clinical severity correlate directly with the size of the trinucleotide expansion. It has been observed that the size of the expansion can also increase with time resulting in the progression of the disorder from generation to generation (Mak *et al.*, 2001).

Anticipation for FSHD was reported by a few authors (Lunt *et al.*, 1995a; Zatz *et al.*, 1995; Tawil *et al.*, 1996). One shortcoming of all these papers reporting anticipation in their populations was, however, that affected parent-offspring pairs, thus only two generations, from multiple families were studied. Lunt *et al.* (1995a) studied 15 families over two to four generations, while Zatz *et al.* (1995) reported anticipation in onset among 28 parent-offspring pairs in 17 families and Tawil *et al.* (1996) studied 23 parent-offspring pairs from multiple families. Lunt *et al.* (2000) did mention that it was not clear how anticipation would be able to occur with a fixed mutation in each family.

Flanigan *et al.* (2001) investigated the presence of anticipation in 66 parent-offspring pairs and 21 grandparent-parent-child sets from a single family (homogeneous population) originally described by Tyler and Stephens in 1950. Tyler and Stephens (1950) described 1,249 descendants of a man who was originally born in England in 1775, but emigrated to Utah in 1840. Tyler and Stephens (1950) identified 159 affected individuals, either by history or examination, and reported on the examination results of 58 people. Twenty four of the 58 individuals (41%) who were asymptomatic by self-report were affected upon examination, and 13 of 18 individuals (72%) who were 20 years of age were also affected upon examination. This family provides one of the earliest descriptions of the frequency of individuals who are asymptomatic by self-report, but affected by examination. Flanigan *et al.* (2001) re-examined the large kindred reported in 1950 by Tyler and Stephens and

extended the pedigree to include 2,220 individuals. Genetic characterisation of the affected individuals was performed and a 20 kb disease-associated deletion fragment was observed to segregate over 12 meioses, in seven different branches of this extended family. No compelling evidence for anticipation, in either reported age of onset or in disease severity, was, however, observed in this extended family. Additional studies, utilising quantitative muscle testing, historical data, and genotyping will help to elucidate the aspects of anticipation, gender effects, and parent-origin effects in FSHD.

2.1.2.5 Female and male transmission effects

Padberg *et al.* (1995b) observed that a higher proportion of females are generally asymptomatic. Zatz *et al.* (1998) investigated 52 families, consisting of 172 patients (60% males and 40% females). An excess of affected males was observed in the patients examined. This might be explained by a greater proportion of asymptomatic females and a significantly greater number of affected sons than daughters of asymptomatic mothers. The penetrance at the age of 30 was 95% for males but only 69% for females. It was also observed that new mutations occurred more frequently in females than in males among somatic and/or germline mosaic cases. Severely affected cases were more commonly the result of sporadic mutations or mutations transmitted through maternal lines, including mosaic mothers. Males were on average more severely affected than females, and more clinically affected sons were observed in the offspring of asymptomatic mothers.

Van der Maarel *et al.* (2000) observed that mosaic males were typically affected, although mosaic females were more often the unaffected parent of a nonmosaic *de novo* patient. A difference in the clinical presentation of females and males was also observed by Busse *et al.* (2000). A 35 kb *Bln* I resistant deletion fragment was confirmed in an unaffected mother, two unaffected daughters and an affected son. This would imply that the two unaffected sisters have a 50% chance of transmitting the 35 kb disease causing deletion fragment to their offspring and that a son would have a higher risk of being affected than a daughter.

2.1.2.6 Sporadic FSHD

The frequency of sporadic cases was reported to be 0.33 (Zatz *et al.*, 1995). Wijmenga *et al.* (1992c; 1993b) identified the presence of *de novo* DNA rearrangements in FSHD. Within each FSHD family the size of the deletion fragment was constant and displayed

stable inheritance in future generations. The parental origin of the mutation was identified via PFGE and haplotype analysis. Griggs *et al.* (1993) observed eight sporadic patients from seven families displaying novel *Eco* RI rearrangements ranging between 15 kb and 23 kb.

Weiffenbach *et al.* (1993) and Griggs *et al.* (1993) examined affected sibling pairs with clinically unaffected parents. In these families, a small *Eco* RI fragment was observed in both of the affected children but not in the parents. It is highly unlikely that the FSHD-associated fragments of identical size in both of the affected offspring was a result of two independent chromosomal rearrangements. It was concluded that one of the parents should have been mosaic harbouring the FSHD-associated rearrangement in the sperm or ovarian cell lines. Jardine *et al.* (1993) also identified a *de novo* DNA rearrangement in a proband and his two affected children. A 15 kb DNA fragment was detected after *Eco* RI digestion and hybridisation with probe p13E-11 in the proband and his affected children, but was absent in both parents of the proband.

The presence of *de novo* DNA rearrangements was also confirmed by Jardine *et al.* (1994b), Padberg *et al.* (1995b) and Brouwer *et al.* (1994, 1995). As mentioned earlier in paragraph 2.1.2.6, the sporadic early-onset FSHD cases reported by Brouwer *et al.* (1995) were on average more severely affected than the familial cases. This can, however, be due to ascertainment bias as mildly affected sporadic cases might often go undiagnosed. Sporadic cases should, however, only be classified as such if both parents have been examined on a molecular level.

2.1.2.7 Phenotype-Genotype correlation

A direct correlation between the residual repeat length, the age of onset and the severity of FSHD was observed by several authors (Lunt *et al.*, 1995a, Lunt *et al.*, 1995b, Tawil *et al.*, 1996, Lunt, 1998, Ricci *et al.*, 1999). FSHD patients who harbour the smallest alleles of only 10 kb (thus one 3.3 kb repeat) present at a younger age and with a more severe form of the disorder. The phenotype of these patients is often accompanied by mental retardation and epilepsy, as discussed in paragraph 2.1.1.2.1 (Funakoshi *et al.*, 1998; Muira *et al.*, 1998). The presence of anticipation was also reported, as mentioned in paragraph 2.1.2.4. It was observed that either the age of onset or the degree of weakness, in consecutive generations in familial FSHD, becomes more severe. If

anticipation exists in FSHD, its molecular mechanism has not yet been elucidated, and can therefore not be utilised in phenotypic-genotypic correlation studies.

Ricci *et al.* (1999) studied the correlation between the clinical severity and the residual repeat length in 122 FSHD families, including 253 affected and 200 unaffected individuals. The authors utilised the 10-grade clinical severity (CS) scale, ranging from 0.5 to 5.0, to determine the severity of muscle weakness (Table 2.3). Higher scores (3 to 5) were assigned to individuals with pelvic and proximal lower limb muscle involvement, since the weakness of these muscles generally follows after weakness of the facial and shoulder muscles. Individuals with no proximal lower limb involvement will therefore have a CS score of less than three, whereas a CS score of 4.0 to 5.0 will imply that the individuals have severe lower limb involvement.

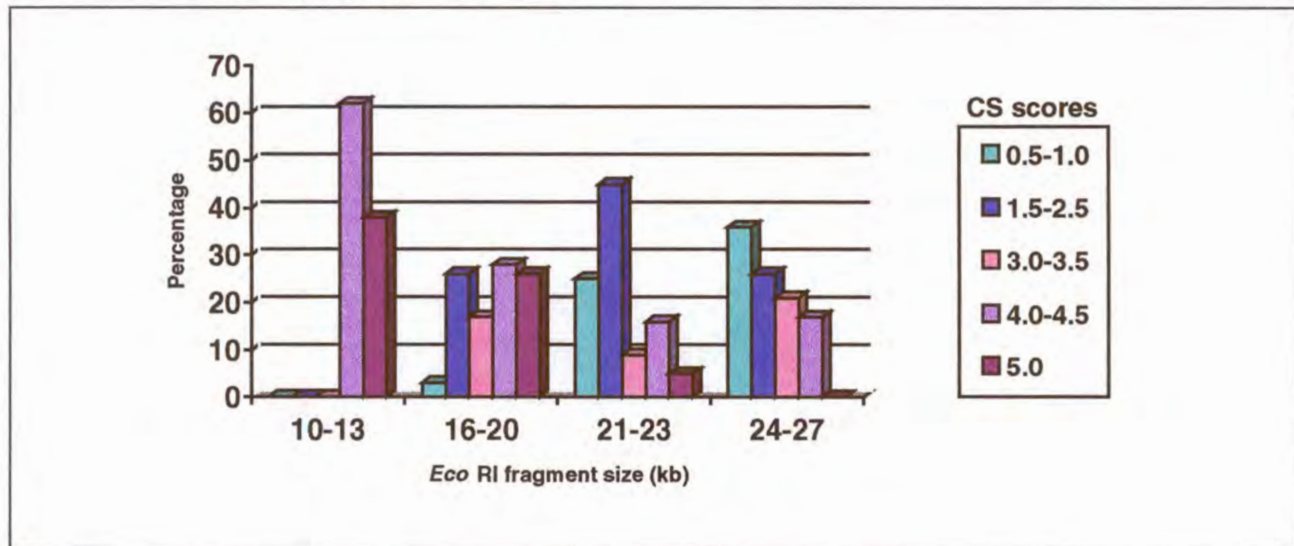
Table 2.3: Clinical Severity Scale for FSHD

Score	Clinical symptom
0.5	Facial weakness
1.0	Mild scapular involvement without limitation of arm abduction; no awareness of disorder symptoms is possible
1.5	Moderate involvement of scapular and arm muscles or both (arm abduction >60° and strength ≥3 in arm muscles); no involvement of pelvic and leg muscles
2.0	Severe scapular involvement (arm abduction <60° on at least one side); strength <3 in at least one muscular district of the arms; no involvement of pelvic and leg muscles
2.5	Tibioperoneal weakness; no weakness of pelvic and proximal leg muscles
3.0	Mild weakness of pelvic and proximal leg muscles or both (strength ≥4 in all these muscles); able to stand up from a chair without support
3.5	Moderate weakness of pelvic and proximal leg muscles or both (strength ≥3 in all these muscles); able to stand up from a chair with monolateral support
4.0	Severe weakness of pelvic and proximal leg muscles or both (strength <3 in at least one of these muscles); able to stand up from a chair with double support; able to walk unaided
4.5	Unable to stand up from a chair; walking limited to several steps with support; may use wheelchair for most activities
5.0	Wheelchair bound

Adapted from Ricci *et al.* (1999).

Graph 2.1 displays the proportion of patients with different CS scores and *Eco* RI deletion fragments. It was observed that the probability of developing a more severe form of the disorder, with a CS score of 4.0 - 5.0, is 100% in the presence of 1 - 2 repeats (10 kb), it decreased to 54% in patients harbouring 3 - 4 repeats (16 to 20 kb) and was less than 21% if the fragments were larger that 20 kb (>4 repeats). The size of the deletion fragment was therefore observed to be a major factor in determining the severity of the FSHD phenotype, thereby having an impact on the clinical prognosis, as well as genetic counselling of FSHD individuals.

Graph 2.1: Proportion of patients with different clinical severity scores and *Eco* RI deletion fragment sizes



Adapted from Ricci *et al.* (1999).

It is interesting to note from the data presented in the above graph that the percentage of severely affected individuals, with CS scores of 4.0 – 5.0, decreases significantly as the *Eco* RI fragment size increases. The percentage of individuals with CS scores of 0.5 to 3.0 increases with the increase of the *Eco* RI fragment size.

2.1.2.8 Prenatal diagnosis

Even though the gene for FSHD has not yet been identified, the detection of rearrangements at the D4Z4 locus with probe p13E-11 provides a reliable indirect method for diagnostic purposes. This therefore enables the prenatal diagnosis of FSHD.

Eggers *et al.* (1993) compiled a questionnaire about the interest in and demand for prenatal and presymptomatic diagnosis for FSHD. The questionnaire was sent to 46 patients and most of the patients indicated that they would have liked to know their diagnosis earlier to enable them to seek more efficient help, to prepare them emotionally, to avoid strenuous activities, or to choose an appropriate profession. Most of the patients also favoured prenatal diagnosis although only two indicated that they would terminate a pregnancy in the case of an affected fetus.

Upadhyaya *et al.* (1999) reported that their laboratory has been involved in the prenatal diagnosis of FSHD since 1993. Several difficulties were encountered during their study

including quality and quantity of DNA needed for the molecular diagnoses. At least 10 μg of high quality, high molecular weight DNA was required to allow good resolution, visualisation, sizing of DNA fragments up to at least 50 kb, and to avoid long autoradiographic exposure times. To obtain the quantity of DNA needed, a sufficient amount of chorion villus tissue had to be biopsied. Upadhyaya *et al.* (1999) also suggested that by completing the molecular diagnosis of an FSHD family prior to the pregnancy can improve the prenatal diagnosis as the molecular defect of the family will already be known and this will decrease the time needed for the prenatal diagnosis.

Galluzzi *et al.* (1999) reported the prenatal diagnosis of FSHD in 15 Italian families. The analysis was performed on DNA isolated from chorionic villi samples, collected at 10-11 weeks of gestation. A deletion fragment associated with the disorder was observed in nine of the fetuses. Parents chose to terminate the pregnancies in eight of the nine cases.

2.1.2.9 Genetic heterogeneity

Evidence for genetic heterogeneity has been reported by several authors (Gilbert *et al.*, 1993, Cacurri *et al.*, 1994, Deidda *et al.*, 1994, Bakker *et al.*, 1996). Gilbert *et al.* (1993) observed two FSHD families from the United States with facial weakness, scapular winging, proximal muscle weakness, and myopathic changes on muscle biopsy without inflammatory or mitochondrial pathology. The phenotype observed in these individuals therefore complied with the clinical requirements for the diagnosis of FSHD, but did not display linkage to the 4q35 region. Cacurri *et al.* (1994) investigated 19 Italian families with FSHD of whom two did not link to 4q35.

It is currently estimated that at least 5% of FSHD families do not display linkage to chromosome 4q35, implying genetic heterogeneity. The finding of heterogeneity in FSHD has important implications for genetic counselling, clinical investigations, prenatal diagnosis, and eventually the cloning of the FSHD gene. Investigation of genetic heterogeneity is therefore included as one of the top ten priorities in international FSHD research. Despite the report of genetic heterogeneity in FSHD almost a decade ago, a second locus has yet to be identified.

2.1.2.10 Candidate genes

The identification of genes associated with the chromosome 4q35 region linked to FSHD has been a difficult undertaking, largely due to the homologous regions in the genome, the high density of repeat units and the gene-poor nature of the region. Despite several attempts the gene, or gene(s) whose modification of expression culminates in the FSHD phenotype still eludes identification. The next section will describe the genes so far identified in the FSHD region.

2.1.2.10.1 Actinin-associated LIM protein gene (ALP)

Piétu *et al.* (1996) cloned a novel human specific complementary DNA (cDNA) by differential screening of a human muscle cDNA array with the aim of identifying muscle specific transcripts. The gene was mapped in the 4q34-qter region. A rat homologue was cloned, and the encoded protein was named actinin-associated LIM protein (ALP), because ALP contains a C-terminal LIM (Lin-11/Isl-1/Mec-3) domain and an N-terminal PDZ (Postsynaptic density protein, Disc-large tumor suppressor and the Zonula occludens protein) domain, which interacts with α-actinin-2 in myofiber Z-lines (Xia *et al.*, 1997). The LIM domain is a specific double-zinc finger motif and the PDZ domain is a protein module composed of β-strands and α-helices. The LIM and PDZ domains are both found in a growing variety of proteins and mediate direct protein-protein interactions. The rat sequence was used to assemble the corresponding human expressed sequence tags (ESTs) and mapped the human gene between D4S171 and the Factor XI gene (Xia *et al.*, 1997). No difference was, however, observed in the expression of ALP in the deltoid muscle from FSHD patients and controls. Bouju *et al.* (1999) therefore concluded that ALP is not a candidate gene for FSHD.

2.1.2.10.2 Adenine nucleotide translocator gene (ANT)

The adenine nucleotide translocator (ANT), also called adenine nucleotide translocase or adenosine diphosphate (ADP)/ATP translocator, is the most abundant mitochondrial protein. This protein facilitates the exchange of adenine nucleotides across the mitochondrial inner membrane. It is a 30 kD homodimer and is embedded asymmetrically in the inner membrane of the mitochondria. The dimer forms pores through which ADP can move from the matrix into the cytoplasm (Neckelman *et al.*, 1987). Three isoforms of

ANT cDNA have been isolated: human skeletal muscle (ANT1), human fibroblast cells (ANT2) and human liver (ANT3) [Neckelmann *et al.*, 1987, Houldsworth and Attardi, 1988].

Minoshima *et al.* (1989) and Li *et al.* (1989) originally mapped ANT1 to chromosome 4 and Fan *et al.* (1992) subsequently assigned ANT1 to 4q35 by fluorescent *in situ* hybridisation. Wijmenga *et al.* (1993c) and Haraguchi *et al.* (1993) mapped the ANT1 gene to 4q35 and more specifically to a site proximal to the FSHD gene. No abnormality was, however, observed after analysing the transcripts of ANT1 from several FSHD patients (Haraguchi *et al.*, 1993).

2.1.2.10.3 Double homeobox gene 4 (DUX4)

Ding *et al.* (1998) cloned a 170 amino acid protein (DUX1) containing a double homeodomain and showed that it is expressed in human rhabdomyosarcoma TE671 cells. Subsequently, two new genes containing double homeodomains and 3.3 kb repeats (DUX2 and DUX3) with similar promoters and ORFs were isolated (Gabriëls *et al.*, 1999). These genes were shown to map to the acrocentric chromosomes and are therefore not involved in FSHD (Beckers *et al.*, 2001). Gabriels *et al.* (1999) aligned the sequences of the 3.3 kb repeat units derived from D4Z4 with those derived from DUX1, DUX2 and DUX3 and identified a promoter and open reading frame encompassing the double homeobox in each 3.3 kb repeat unit and termed the putative gene DUX4 (Figure 2.6 and Table D.1). DUX4 encodes a 391 amino acid protein containing 2 homeodomains. *In vitro* transcription and translation of the ORF in a rabbit reticulocyte lysate yielded two products, corresponding to the DUX4 monomer and dimer. It was proposed that each of the 3.3 kb repeat units could therefore harbour a DUX4 gene, which in turn encodes a double homeodomain protein.

The following hypothesis for the role of the DUX4 gene in FSHD was put forward by Gabriels *et al.* (1999): The deletions of 3.3 kb repeat elements at the D4Z4 locus destabilises the heterochromatin allowing expression of the gene in some repeats in some cells. The DUX4 protein is, however, toxic to muscle cells, possibly because of its strong dimerisation potential. This hypothesis would also explain the dominant character of FSHD and the fact that the severity correlates inversely with the residual number of repeats.

Leclercq *et al.* (2001) reported the presence of a 3.3 kb repeat 40 kb centromeric of the D4Z4 locus. This repeat was, however, observed to be inverted and the authors hypothesised that this repeat might function as an enhancer.

DUX4 might therefore play a role in the pathogenesis of FSHD as partial deletion of the number of 3.3 kb repeat units may alter DUX4 expression in FSHD patients. Confirmation of this hypothesis would, however, be challenging, due to expression of the large number of homologous 3.3 kb repeats scattered all over the human genome.

2.1.2.10.4 Fibroblast growth factor (FGF) and FGF receptors

Growth factors have been indicated to play an important role in muscle regeneration or degeneration. For this reason Saito *et al.* (2000) investigated the expression of several growth factors and their receptors in FSHD patients to determine whether there is any association with FSHD pathogenesis. Saito *et al.* (2000) reported one severely affected FSHD patient with an overexpression of fibroblast growth factor (FGF) and FGF receptor 4. The other patients with FSHD, DMD, BMD, and LGMD did not show any overexpression. No significant difference in the expression levels of the platelet-derived growth factor (PDGF) and receptors, fibroblast growth factor receptor 1 (FGF-R1), FGF receptor 3 (FGF-R3), platelet-derived growth factor receptor α (PDGF-Rα), PDGF receptor β (PDGF-Rβ), and heparan sulfate proteoglycan (HSPG) were observed in the entire group of patients examined. Saito *et al.* (2000) hypothesised that the severe clinical phenotype of the FSHD patient investigated might have been due to the overexpression of FGF, which caused excessive fibrosis of the skeletal muscle. The overexpression of FGF and FGF receptor 4 in the severely affected FSHD patient therefore suggests that the above should be investigated in other FSHD patients to determine whether it is a common feature in severely affected FSHD patients.

2.1.2.10.5 The FSHD region gene 1 (FRG1)

Van Deutekom *et al.* (1996b) identified the first functional gene (FRG1) that mapped 100 kb centromeric of the repeat units on chromosome 4q35. A 1.1 kb FRG1 transcript was observed upon Northern blot analysis of adult muscle, lymphocytes, fetal brain, muscle and placenta. No homology to other known genes could be demonstrated. A polymorphism in exon one of this gene was observed and reverse transcriptase

polymerase chain reaction (RT-PCR) from lymphocytes and muscle biopsies from patients and controls indicated that both alleles were transcribed and that there was no evidence for transcription suppression. No differences in FRG1 messenger ribonucleic acid (mRNA) levels in FSHD patients compared to controls were observed. FRG1 can however not be excluded before the disease mechanism has been determined.

Grewal *et al.* (1998) investigated FRG1 on an evolutionary bases through the comparison of the genomic organisation of this gene in two species, the mouse and the Japanese puffer fish (*Fugu rubripes*). Although the puffer fish gene was found to be five times smaller than that of the mouse, it was observed that the intron and exon structure of FRG1 was identical throughout the protein coding region. The authors also reported that FRG1 was homologous in the two nematodes, *Caenorhabditis elegans* and *Brugia malayi*. It was therefore observed that the human FRG1 gene is highly conserved in both vertebrates as well as invertebrates. Furthermore, the comparison of the vertebrate homologues revealed that all the proteins contained a lipocalin sequence motif near the N-terminal. FRG1 may therefore play a role in protein transport as lipocalins are known to be a large family of extracellular proteins, which transport small hydrophobic molecules such as steroids, retinoids and lipids. The FRG1 protein (FRG1P) was also observed to localise in the nucleolus, Cajal bodies and speckles, and this could imply a fundamental role in RNA processing (Van Koningsbruggen *et al.*, 2000; Van Geel, 2001).

2.1.2.10.6 The FSHD region gene 2 (FRG2)

The FRG2 gene was identified by means of *in silico* exon prediction and confirmed in subsequent expression studies found it to map only 37 kb proximal to the D4Z4 repeat array (Van Geel, 2001). The gene consists of four exons and a strong promoter including a 5' TATA and CCAAT box was predicted. It was observed that FRG2 expression is absent in all tissues tested, however, low, but distinct levels of FRG2 expression was observed in differentiating myoblasts from FSHD patients. No expression of FRG2 could be detected in control myoblasts. FRG2 is therefore the first gene on 4qter that was found to be specifically expressed in FSHD cells. It was further observed that FRG2 is expressed in patient and control fibroblasts undergoing forced myogenesis by adenoviral *MyoD* expression, suggesting a muscle-specific role for this gene. *MyoD* plays a unique role in satellite cell activation and differentiation in myofibers. The exact function of FRG2 could not be determined but results indicated that the FRG2 protein is a nuclear protein

involved in myogenesis and that its transcriptional deregulation is related to FSHD. FRG2 therefore remains an important candidate gene for FSHD. (Van Geel, 2001)

2.1.2.10.7 Human beta-tubulin gene (TUBB4Q)

The TUBB4Q gene maps 80 kb proximal to the FSHD associated D4Z4 repeats on chromosome 4q35 (Van Geel *et al.*, 2000). The gene contains four exons, encoding a protein of 434 amino acids. Although the genomic structure shows all functional aspects of a gene, no transcripts could be detected upon RT-PCR analysis on RNA samples from a wide variety of human adult and fetal tissues. No evidence has therefore been generated to prove that expression of TUBB4Q is involved in FSHD (Van Geel, 2001).

2.1.2.11 Molecular models proposed for the aetiology of FSHD

The molecular mechanisms underlying FSHD are still unclear. It was proposed that the tandem repeats at the D4Z4 locus promote non-homologous unequal recombination between regions within the genome showing homology to it, such as the repeats within the telomeric region of 10qter, thus causing the deletions associated with FSHD. Two hypotheses are postulated in paragraphs 2.1.2.11.1 and 2.1.2.11.2 that might explain the unique pathogenesis of FSHD.

2.1.2.11.1 Homeodomain

Homeobox genes play a crucial role in establishing the anterior-posterior axis during embryogenesis in vertebrates and invertebrates. These genes are therefore candidates to determine the differences observed in specific muscle groups in FSHD (Fischbeck and Garbern, 1992). The two homeobox sequences observed in each 3.3 kb repeat are 67 to 68% homologous on a nucleotide level and were 50 to 52% identical in amino acid sequences (Hewitt *et al.*, 1994; Lee *et al.*, 1995b). Open reading frames have been observed within the homeobox sequences, but not through the whole repeat. The two homeodomain sequences were similar to that of the human paired-type homeodomains, (Paired box gene [Pax] 3, Pax 6, and Orthodenticle [Otx] 1), *Xenopus* mesoderm induced homeobox [Hmix], *Drosophila* paired (prd) [HmprD] and the muscle specific homeodomain protein [Mhox]. Proteins coded by homeodomains display sequence-specific DNA binding

and play an important role in the process of transcription as transcription regulation factors. (Lee *et al.*, 1995b)

The human Pax 3 gene is important in muscle differentiation and mutations in this gene cause Waardenburg syndrome types I and III. Mhox is only expressed in muscle tissue and regulated the expression of muscle-specific creatine kinase and other factors, by binding to the upstream regulatory sequences of the muscle creatine kinase gene (Lee *et al.*, 1995a).

The homeobox sequences within the repeat units provided evidence for a possible gene within the repeat unit. The homeodomain proteins bind DNA and regulate other genes. If the FSHD gene is located at the D4Z4 locus within the repeat units, a minimum number of the repeats may be necessary for normal gene function and deletions of the repeats could result in the loss of the FSHD gene, thus resulting in the disorder. However, no transcripts have been identified from these homeodomains, making it less likely that the FSHD gene is located within these repeat units.

2.1.2.11.2 Position effect variegation

A position effect is displayed when a change in the level of gene expression is due to a change in the position of the gene relative to its normal chromosomal environment (Kleinjan and Van Heyningen, 1998). The expression of genes is influenced by their position in the genome, thus whether located in heterochromatic or euchromatic regions. Chromosomal rearrangements can lead to the alteration of the gene's environment and may therefore change the expression of the gene which is referred to as a position effect.

Winokur *et al.* (1993) and Hewitt *et al.* (1994) postulated that the deletion of the D4Z4 repeats results in a position effect, by disrupting the local chromosomal structure. Deletions of the repeats could result in the expansion of telomeric heterochromatin into adjacent euchromatin altering the expression of the gene. Position effect variegation (PEV) has been observed in *Drosophila* and yeast and it was shown that physical proximity to the centromere, to the telomere or to other heterochromatic regions may actively suppress gene expression (Fisher and Upadhyaya 1997). The down regulation is presumably as a result of an alteration of the structure of euchromatic domains by the

heterochromatic regions. A gradient effect can be observed as genes that are closer to the rearrangement breakpoint are more severely affected than those further away.

Alternatively, the appropriate expression of the gene could require a heterochromatic environment showing a position effect if moved to an euchromatic region. It has been observed that several *Drosophila* genes require a heterochromatic environment for normal function showing position effects if placed within euchromatin. Locus D4Z4 contains *Lsau* repeats that are usually present within heterochromatic regions. The deletions at this locus could therefore result in a position effect on euchromatic gene(s) being moved into heterochromatic regions or alternatively heterochromatic genes being placed in a euchromatic environment as the deletion of GC-rich repeats could result in the loosening of the heterochromatin structure causing a position effect (Winokur *et al.*, 1994).

Position effects occur over several hundred kilobases, therefore, the FSHD gene could be located at quite a distance from D4Z4. The disease causing mutation might be the deletion of the 3.3 kb repeats which lead to FSHD due to the altered transcriptional activity of a gene(s), rather than a mutation within the gene itself. The median repeat array on chromosome 4 was found to be 21 kb larger than that on chromosome 10. This might reflect the requirement for a larger subtelomeric domain on chromosome 4 to prevent gene silencing, or expression, of critical genes in the region proximal to the telomere (Van Overveld *et al.*, 2000).

Genotype-phenotype correlation studies indicated an inverse relationship between the residual repeat length and the clinical severity (paragraph 2.1.2.7). PEV could explain this since the FSHD candidate region is moved further into the 4q heterochromatic telomeric region as the deletion increases, resulting in an increased disease severity. Furthermore the effect on the chromosomal structure from heterochromatic to euchromatic increases as the deletion increases, influencing gene expression correspondingly.

This is the first study to investigate FSHD on both the clinical and molecular levels. FSHD is clinical extremely heterogeneous which results in a complex clinical diagnosis. It is, however, possible to detect the presence of the disease causing DNA rearrangement on chromosome 4q35. This therefore highlighted the need and importance of a molecular study in the South African FSHD population.

2.1.2.12 Objectives of this study

This is the first study to investigate the molecular basis of FSHD in the South African population. Five extended Caucasian FSHD families, consisting of 100 individuals, from the South African population were selected for this study. Haplotype analyses will be performed to study the segregation of STRP markers in the 4q35 region with the FSHD phenotype; and the DNA rearrangements at the D4Z4 locus, associated with FSHD, will be detected via Southern blot analyses utilising probe p13E-11.

CHAPTER THREE

MATERIALS AND METHODS

The study presented in this thesis was approved by the Ethics Committee of the Faculty of Medicine of the University of Pretoria and Pretoria Academic Hospitals, and was performed after the appropriate approval was obtained (protocol number 91/96). Prior to the collection of blood samples, informed consent was obtained from all individuals included in this study.

3.1 FSHD FAMILIES

A group of five Caucasian FSHD families (families F10, F20, F30, F40 and F60) were investigated in this study. Core families, including a total of 74 individuals, were selected from the extended family pedigrees (Appendix F). The core families consisted of individuals who were relevant to the study, and from whom blood samples could be obtained. An additional 26 individuals were also included and a total of 100 individuals were thus investigated in this study. The individuals were numbered according to the family number (F10, F20, F30, F40 or F60) and the unique pedigree number, as presented in Appendix F. For example, individual 1 of family F10 would be presented in the pedigrees in Chapter three and the haplotypes in Chapter four as 10-1. The pedigree numbers of the individuals selected for this study are underlined in the pedigrees presented in this chapter.

A national referral centre for FSHD was established in South Africa as a joint collaboration between the Departments of Neurology and Human Genetics, at the University of Pretoria. The FSHD clinic, for the first time, provided the opportunity for individuals affected with FSHD, as well as their families to receive up to date clinical and molecular information. One specialised neurologist, Dr. Clara Schutte, clinically evaluated the individuals attending the clinic, therefore excluding any possibility of diagnostic bias during the clinical examination. The individuals received genetic counselling as well as information regarding the molecular project.

The majority of the individuals included in this study have, however, not been clinically diagnosed by Dr. Schutte, but by a different neurologist ca. 5-6 years ago. However, only clinical data generated by Dr. Schutte was utilised in the detailed discussion of the clinical phenotype in paragraphs 3.1.1 to 3.1.5, and for this reason detailed clinical data is only available for selected individuals included in the study

The pedigrees, listed in various figures throughout the thesis, contain clinical diagnoses, which were not made by Dr. Schutte. The main difference between the respective clinical diagnostic protocols is that our current protocol disregard clinical diagnoses of individuals below the age of 20 – as these clinical diagnoses are not reliable due to delayed onset and only 95% penetrance of the phenotype at the age of 21. However, since the clinical diagnostic data of young individuals was available it could not be excluded from this study. For this reason a high level of genotype-phenotype discordance was expected in the group of young individuals included in this study.

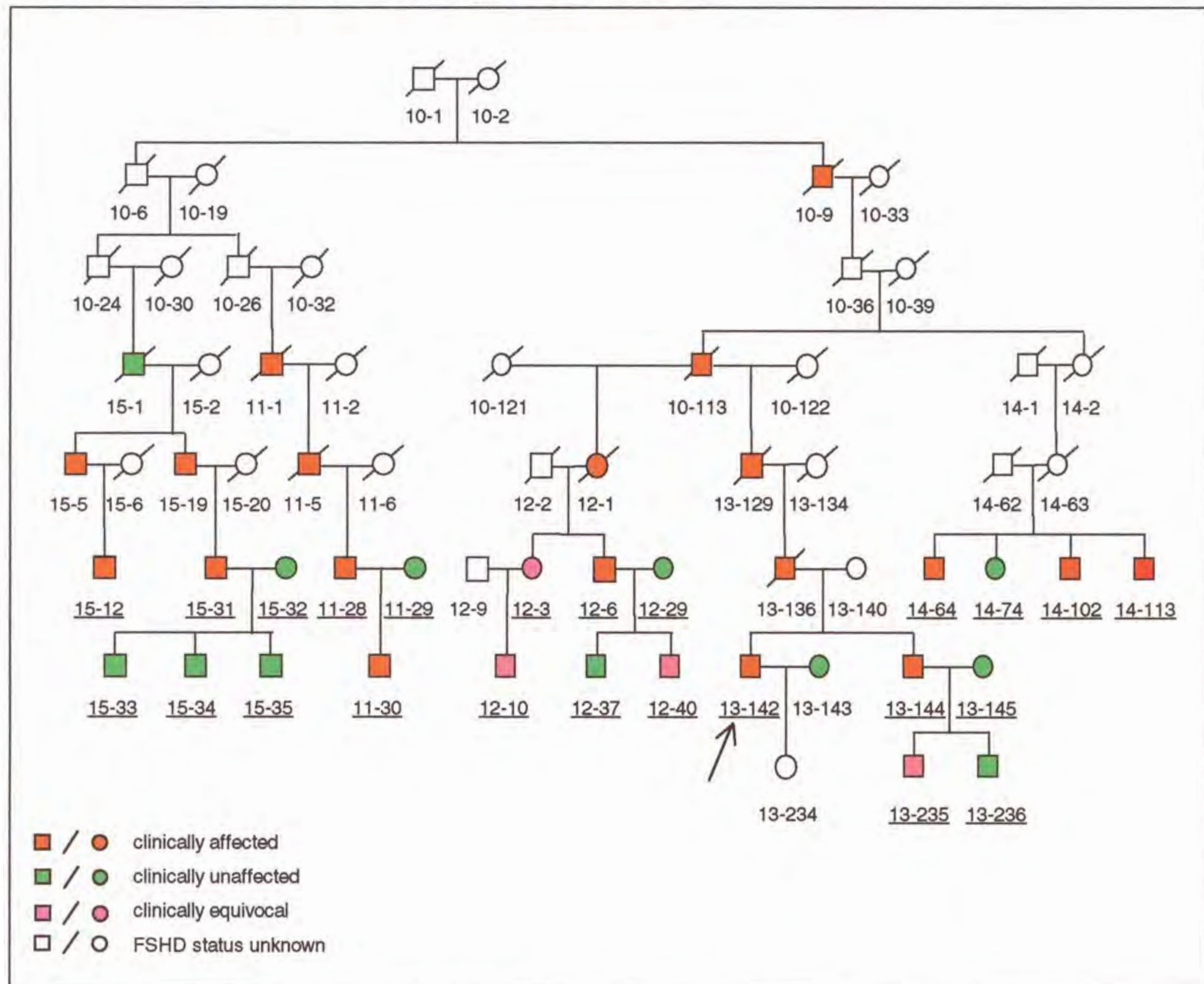
3.1.1 Family F10

Family F10 is the biggest FSHD family included in this study and was subdivided into five sub-families (F11, F12, F13, F14 and F15) for easy referral. Figure 3.1 presents an excerpt of the eight generation pedigree of family F10 which includes 480 individuals of whom 24 were genotyped. Individuals from all five of the sub-families have been included in this study.

The proband (individual 13-142 in Figure 3.1) of family F10, was clinically diagnosed with extreme facial weakness, scapula winging and high riding of the scapula. The muscles of his arms, upper legs, feet and hips were, however, only mildly affected. Two individuals not included in this phase of the FSHD project, have attended the FSHD clinic. Individual 12-31 (Figure F.3) is a 28 year old male with mild weakness of the mouth, eye and feet muscles and extreme weakness of the muscles around the shoulder area, resulting in profound high riding and winging of the scapula. His upper arm muscles are also relatively weak, no weakness of the lower arms, upper leg, hips or hand muscles are present. Individual 12-13 (Figure F.3) is 55 years old with mildly affected facial muscles but extreme weakness of the muscles around the shoulder area. Scapula winging and high riding was observed upon clinical examination. The muscles of the hips, upper legs and upper arms

were only mildly affected, while the muscles of the lower arms, hands and feet were unaffected.

Figure 3.1: Excerpt from a pedigree of family F10

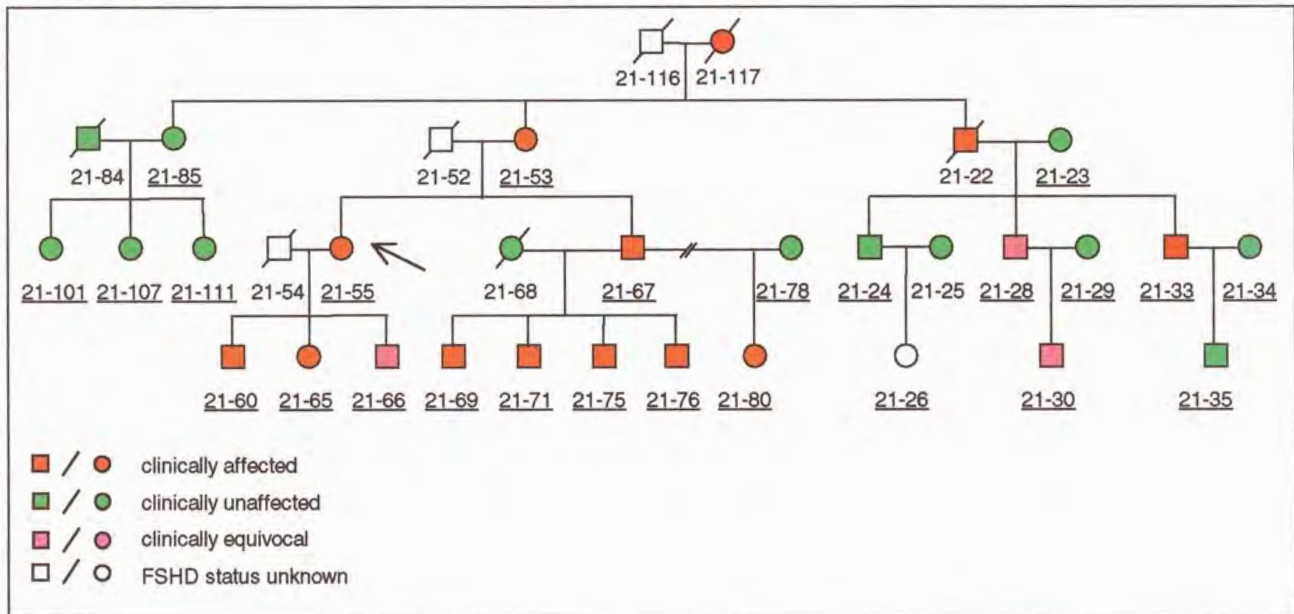


A full pedigree of family F10 is presented in Figure F.1. The full pedigrees of the sub-families are presented in Figures F.2 to F.6, Appendix F.

3.1.2 Family F20

Family F20, which include 395 individuals, was also subdivided into several sub-families for easy reference, but blood samples could only be obtained for three of the sub-families. For the purpose of this study, 25 individuals were selected from only one sub-family (F21) to represent family F20. Figure 3.2 presents an excerpt of the pedigree of family F21 and the individuals selected for this study. The proband of this family 21-55 is a 62 year old female and was classified as FSHD positive as her clinical phenotype met the clinical criteria set by the International FSHD Consortium (Appendix C).

Figure 3.2: Excerpt from a pedigree of family F21

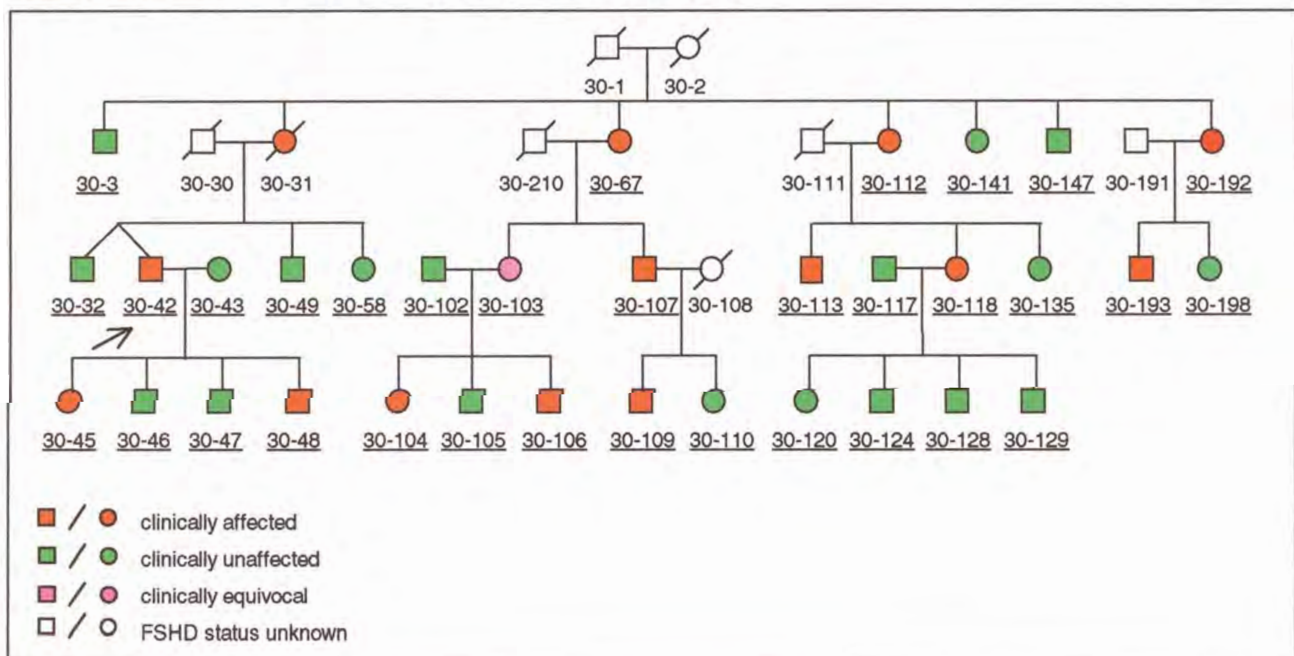


A full pedigree of this family is presented in Figure F.7.

3.1.3 Family F30

Figure 3.3 presents the excerpt from a pedigree of family F30 which includes a 65 year old male proband (individual 30-42), who was diagnosed with extreme weakness of his shoulder and surrounding muscles and pelvic girdle muscles resulting in a distinct waddling gait. This individual also has visible facial weakness. The family consists of four generations with 195 individuals of whom 33 have been genotyped in this study.

Figure 3.3: Excerpt from a pedigree of family F30

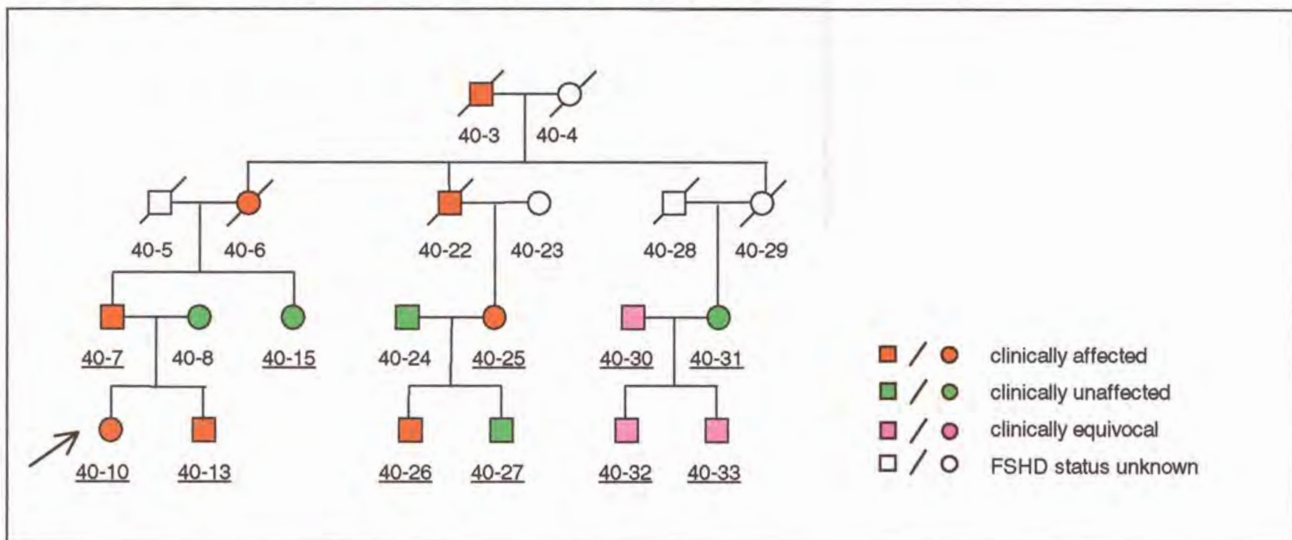


A full pedigree of this family is presented in Figure F.8.

3.1.4 Family F40

The proband of family F40 (individual 40-10 in Figure 3.4) is a middle aged female with mildly affected facial and upper arms muscles, but extreme weakness of the muscles around the scapula which resulted in extreme winging of the scapula. Her hips and upper-legs were also observed to be severely affected upon clinical examination. This individual walks with a distinct waddling gait due to the weakness of the pelvic girdle muscles. The brother of the proband (individual 40-13) has, however, only mild weakness of the eyes, mouth, upper arms and the proximal part of the lower arms. Mild weakness of the scapula resulting in only slight winging and high riding of the scapula was observed. This family consists of 40 individuals of whom 11 were genotyped.

Figure 3.4: Excerpt from a pedigree of family F40

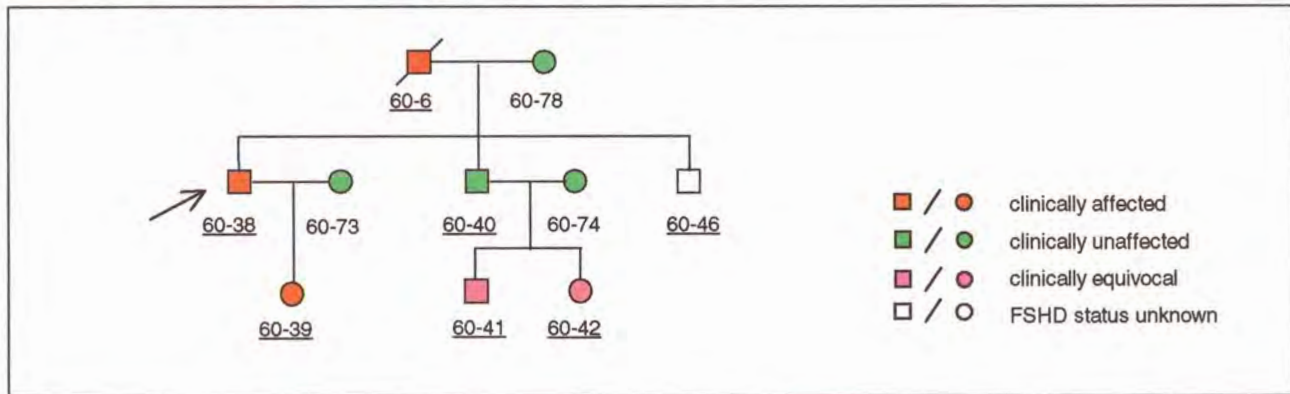


A full pedigree of this family is presented in Figure F.9.

3.1.5 Family F60

Family F60, displayed in Figure 3.5, includes a middle aged (49 years old) male proband (individual 60-38) with atrophic pectoralis major muscles and facial weakness. One deceased individual (60-6) could be genotyped as blood samples were collected from him prior to his death. Family F60 contains 75 individuals of whom seven were genotyped in this study.

Figure 3.5: Excerpt from a pedigree of family F60



A full pedigree of this family is presented in Figure F.10.

3.2 DNA ISOLATION

Whole blood was collected in tubes containing either ethylenediamine tetraacetic acid (EDTA) or acid citrate dextrose (ACD) as the preservative. The samples were divided into aliquots of 10 milliliters (ml) each. The blood was stored at -70°C until the DNA was isolated. The DNA was initially isolated with the use of sodium perchlorate and chloroform and subsequently by the Wizard Genomic DNA Purification Kit® [1] [Promega].

3.2.1 Isolation of genomic DNA using sodium perchlorate and chloroform

DNA was extracted from the 10 ml aliquots of whole blood using a modified method of the one described in 1989 by Johns and Paulus-Thomas (Olckers, 1997). If frozen, the blood was first thawed on ice and transferred to a 50 ml polypropylene centrifuge tube [Elkay or Sterilin]. Thirty five ml cell lysis buffer (0.32 M sucrose; 10 mM Tris-HCl (pH 8.0); 5 mM $MgCl_2$; 1% Triton X-100®) was added to the 10 ml whole blood and incubated on ice for 10 minutes. The solution was centrifuged at 1000 xg for 20 minutes at 4°C after incubation. The supernatant was discarded and the pellet resuspended in 9.5 ml suspension buffer (50 mM Tris-HCl (pH 8.0); 150 mM NaCl; 100 mM EDTA). Half a millilitre of a 30% SDS solution was added to the suspension, adjusting the final SDS concentration to 1.5%. Subsequently, 2.5 ml of a freshly prepared 5 M $NaClO_4$ solution and an equal volume of chloroform:IAA (24:1) were added to the suspension. The

[1] Wizard® is a registered trademark of the Promega Corporation, Madison, WI, U.S.A.

suspension was mixed on an orbital shaker for 30 minutes to extract the DNA. After the extraction, it was centrifuged at 1,000 xg for 10 minutes at 20°C.

Two volumes of ice cold absolute ethanol (stored at -20°C) were added to the aqueous phase, which was transferred to an Erlenmeyer flask, to precipitate the DNA. The DNA was spooled onto a sterile glass rod and resuspended in 5 ml of suspension buffer. The DNA was precipitated once more, where after the DNA was washed with 70% ethanol, air dried, transferred to Eppendorf®[1] tubes and dissolved in 1.0 ml of TE buffer (10 mM Tris-HCl (pH7.5), 1 mM EDTA) overnight.

The average DNA yield per 10 ml whole blood obtained with the method as described by Johns and Paulus-Thomas, (1989) was 250-450 µg and the A_{260}/A_{280} ratio for all samples was between 1.6 and 2.1. DNA preparations were treated with proteinase K (100 $\mu g.ml^{-1}$) at 60°C overnight when protein contamination was present.

3.2.2 Isolation of genomic DNA utilising the Wizard® Genomic kit

The Wizard® Genomic DNA purification kit enables the isolation of DNA from white blood cells and various other sample types. Different amounts of starting material may be used depending on the application and DNA yield required (Table 3.1). The Wizard® system was selected due to the fact that it yields DNA which is suitable for a variety of applications, including PCR, RFLP and Southern Blot.

Table 3.1: Average DNA yield from various amounts of starting material

Amount of whole blood	Typical DNA yield[a]
300 µl	5–15 µg
1.0 ml	25–50 µg
10.0 ml	250–500 µg

a = Yield depends on the quantity of white blood cells present. Adapted from Wizard® Genomic DNA purification kit, technical manual, 1998.

The kit is based on a four-step protocol. The red blood cells were lysed in the first step with the addition of 30 ml of the Cell Lysis Solution to a sterile 50 ml polypropylene tube [Elkay or Sterilin] for 10 ml whole blood samples. The tube of blood was gently rocked

---

[1] Eppendorf® is a registered trademark of Eppendorf, Hamburg, Germany.

until thoroughly mixed, transferred to the 50 ml tube containing the Cell Lysis Solution, inverted 5-6 times, and incubated for 10 minutes at room temperature during which the solution was mixed twice by inversion. After centrifugation at 2,000 xg for 10 minutes at room temperature as much supernatant as possible was removed without disturbing the visible white pellet.

To ensure efficient cell lysis of the white blood cells and their nuclei in the second step of the protocol, the tube was vigorously vortexed until all the cells were resuspended. Ten ml of Nuclei Lysis Solution was added to the tube containing the resuspended cells and mixed 5-6 times to lyse the white blood cells. The solution became very viscous and if clumps of cells were visible after mixing it was incubated at 37°C until the clumps were dissolved.

The cellular proteins were removed in the third step, leaving the genomic DNA in solution. 3.3 ml of the Protein Precipitation Solution was added to the lysed white blood cell solution and vortexed vigorously until small protein clumps were visible. The samples were centrifuged at 2,000 xg for 10 minutes at room temperature to collect the cellular proteins as a brown protein pellet at the bottom of the centrifuge tube.

Finally, the high molecular weight genomic DNA was precipitated and desalted by an isopropanol precipitation step. The supernatant was transferred to a 50 ml polypropylene tube containing 10 ml of isopropanol at room temperature. The tube was gently mixed until the white thread-like strands of DNA formed a visible mass. After centrifugation at 2,000 xg for one minute at room temperature the DNA was visible as a small white pellet. The supernantant was decanted and one sample volume of 70% ethanol (at room temperature) was added to the DNA. The tube was gently inverted several times to wash the DNA and the sides of the centrifuge tube. This was followed by centrifugation at 2,000 xg for one minute. The ethanol was carefully aspirated and the pellet air dried for 10–15 minutes after which it was rehydrated in 500-830 µl of DNA rehydration solution. For rapid DNA rehydration the DNA was incubated at 65°C for 1 hour, or alternatively the DNA was incubated overnight at room temperature or 4°C.

The average DNA yield per 10 ml whole blood obtained with the Wizard® Genomic kit was 450-800 µg and the A_{260}/A_{280} ratio for all samples was between 1.7 and 2.1. The concentrated stock DNA was stored at -20°C after working dilutions of 50 ng.μl^{-1} (used for

PCR applications) and 250 ng.μl^{-1} (used for Southern blotting) were made. The working dilutions were stored at 4°C.

3.3 HAPLOTYPE ANALYSIS

Five short tandem repeat polymorphism (STRP) markers were initially identified for haplotype analysis as illustrated in Table 3.2. Four additional markers were, however, also included for more conclusive results (Table 3.2). A total of nine STRP markers, consisting of three dinucleotide, one trinucleotide and five tetranucleotide repeat markers on chromosome 4q35, were therefore included for haplotype analysis in this study. The sex average map was utilised, since both male and female individuals were included in the study.

Table 3.2: Genetic map of human chromosome 4q35

Locus[a), b)]	**DNA marker**	**Sex average map (cM)**
D4S426*	AFM238ve3	206.98
D4S2921*	AFMa190zf5	206.98
D4S2688□	UT7694	208.07
D4S2283□	UT2219	208.07
D4S2299□	UT5785	208.07
D4S2390*	ATA22F02	208.07
D4S2930*	AFMa224xh1	208.07
D4S1652*	GATA5B02	208.07
D4S1523□	UT1366	211.65

a) * Markers initially identified for this study. b) □ Additional markers utilised. Table adapted from MFD, 2001.

Table 3.3 lists the markers, their respective loci, primer sequence, melting temperature (T_m), product size, number of alleles, as well as the heterozygosity value for each marker. The information listed in table 3.3 was obtained from several genome mapping centers: The Co-operative Human Linkage Center (CHLC, 2001), The Centre for the Study of Human Polymorphisms (CEPH, 2001), The Genome Database (GDB, 2001) and The Center for Medical Genetics Marshfield Medical Research Foundation (MFD, 2001). The T_m was calculated for each primer as described by Thein and Wallace (1986), utilising the following equation:

$$T_m = 2(A + T) + 4(G + C)$$

Table 3.3: Primers for short tandem repeat polymorphism markers located on chromosome 4q35

Locus	DNA marker[a] (repeat type)	Primer Sequence[b]	T_m	Product size (bp)	No. of alleles[d]	HET[e]
D4S1523	UT1366 (TetNR)[1]	F: 5'-tctactcacatgcggctgg-3' R: 5'-tagtgtttggtggaatttgca-3'	60 58	285[c]	7	67%
D4S1652	GATA5B02 (TetNR)[2]	F: 5'-aatccctgggtacattatatttg-3' R: 5'-cagacattctttattctttacctcc-3'	62 64	122-158	4	69%
D4S2930	AFMa224xh1 (DNR)[3]	F: 5'-cctcatggtaggttaatcccacg-3' R: 5-tattgaatgcccgccatttg-3'	70 58	217-233	9	86%
D4S2390	ATA22F02 (TriNR)[4]	F: 5'-ctcattttcccctttccact-3' R: 5-gtggttttcatcatgagatgc-3'	58 60	102-120	6	76%
D4S2299	UT5785 (TetNR)[5]	F: 5'-tgagcatgtgaaccaatgc-3' R: 5'-ctcacttcattcccaactg-3'	56 56	201-244	9	61%
D4S2283	UT2219 (TetNR)[6]	F: 5'-ccccgttatttttccatctac-3' R: 5'-ctaaagcaaaatgcagacaca-3'	60 58	383[c]	4	59%
D4S2688	UT7694 (TetNR)[7]	F: 5'-agaatgtttgtgacagatgta-3' R: 5'-cagggatgaagtaacagaag-3'	54 58	246[c]	6	52%
D4S2921	AFMa190zf5 (DNR)[8]	F: 5'-tccttcaggaactggtg-3' R: 5'-ttaaaaatctacagacaagggc-3'	52 60	141-163	8	56%
D4S426	AFM238ve3 (DNR)[9]	F: 5'-atacactgcatccatatatacaag-3' R: 5'-acattgtgaaatgaccacagtcaag-3'	68 66	177-191	6	78%

a = References for primer sets are as follows: 1) Gerken *et al.* (1993a); 2) Murray *et al.* (1995a); 3) Dib *et al.* (1996); 4) Murray *et al.* (1995b); 5) Gerken *et al.* (1993b); 6) Gerken *et al.* (1993c); 7) Gerken *et al.* (1994); 8) Dib *et al.* (1996); 9) Weissenbach *et al.* (1992). b = F and R indicate forward and reverse primers respectively. c = No allele size range was available for markers UT1366, UT2219 and UT7694. d = The number of published alleles are listed. e = The heterozygosity value for each marker is listed in this column.

3.3.1 STRP marker information

The allele frequencies and partial sequences on chromosome 4q35, encompassing each STRP marker utilised in this study are presented in the following paragraph. The allele frequencies correspond with those reported on the CEPH database (CEPH, 2001). The sequences were retrieved from the database of the National Center for Biotechnology Information with accession numbers as listed in the footnote for each marker (NCBI, 2001).

3.3.1.1 Marker UT1366 at locus D4S1523

The allele sizes of this marker have not been reported and the allele frequencies of only five of the seven alleles were reported since this marker has only been genotyped on four CEPH families (Ballard, 1999). The reported allele frequencies ranged from 0.077 to 0.346, as presented in Table 3.4.

Table 3.4: Allele frequencies for marker UT1366 at locus D4S1523

Allele (bp)	---	---	---	---	---	---	---
Frequency	0.192	0.231	0.346	0.154	0.077	0.000	0.000

No allele sizes were reported (---).

The UT1366 marker amplifies a polymorphic tetranucleotide repeat at the D4S1523 locus as presented in Table 3.5. The clone of this marker contains nine (AGAT) repeats and the amplicon is 285 bp (Gerken *et al.*, 1993a).

Table 3.5: Partial gDNA sequence at locus D4S1523 encompassing marker UT1366

Nucleotide number[a]	Genomic DNA sequence[b]
1	cttctactc acatgcggct ggaaaaaaca acagctcaca aaggaaagaa aacctaaata
61	atagcttatt tatatagtaa ttagatgata gatgatagat agatagatag atagatagat
121	agatagatag atgagagaga gagagagagg tgaggaaaac tctttaaaat atcaccatgg
181	taaatataat atcagagatt cttcatattt gtttgttttg cttttggaag tactcatatg
241	tcaccatcta gcaggactag atggtgatgc aaattccacc aaacactaaa aagaagagtc
301	tattttcttg agactgagaa aaactcatac cctcaaagac tgaagactaa tctttgattt
361	gctgaaagct tgcctaagag ttgga

a = Nucleotide positions are as reported by Gerken *et al.* (1993a). b = The sequence was deposited in Genbank with accession number L16394. The positions of the primer set are indicated by underlined text (xxx). The STRP at this locus is indicated by double underlined text (yyy).

3.3.1.2 Marker GATA5B02 at locus D4S1652

Four alleles ranging from 138 to 150 bp have been reported for marker GATA5B02, as listed in Table 3.6. The allele frequencies of the four alleles are not equal with the 138 bp allele being the most frequent and the 150 bp allele the least frequent.

Table 3.6: Allele frequency for marker GATA5B02 at locus D4S1652

Allele (bp)	138	142	146	150
Frequency	0.357	0.286	0.268	0.089

Table 3.7 presents the gDNA sequence encompassing marker GATA5B02. The marker amplifies a $(CTAT)_n$ repeat at locus D4S1652. An amplicon of 140 bp is obtained for a clone consisting of nine $(CTAT)_n$ repeats (Murray *et al.*, 1995a).

Table 3.7: Partial gDNA sequence at locus D4S1652 encompassing marker GATA5B02

Nucleotide number[a]	Genomic DNA sequence[b]
1	ggaatccctg ggtacattat atttgttaaa ttttagaatc tatcatctat ctatttctct
61	atcatctatc tatctatcta tctatctatc tatctatcta tctacttcta gtgagtggga
121	ggtaaagaat aaagaatgtc tggctaatgg aaaattctcg tataatatct attcatatca
181	caagngtaag aagattttga tagaaaaaca aacacatttg atcattcagt gtgaactccc
241	tttgctcagt tggaatttgn tttcaacata tctgggtttt ctcatcagcc tctagttaaa
301	acagtganat tttgtatagg aa

a = Nucleotide positions are as reported by Murray *et al.* (1995a). b = The sequence was deposited in Genbank with accession number G08375. The positions of the primer set are indicated by underlined text (xxx). The STRP at this locus is indicated by double underlined text (yyy).

3.3.1.3 Marker AFMa224xh1 at locus D4S2930

Marker AFMa224xh1 generates nine alleles ranging from 217 to 233 bp at locus D4S2930 as presented in Table 3.8. Equal frequencies were reported for the 225, 231, 229 and 233 bp alleles. The 225 and 231 bp alleles are the least frequent with the 221 bp allele being the most frequent.

Table 3.8: Allele frequency for marker AFMa224xh1 at locus D4S2930

Allele (bp)	217	219	221	223	225	227	229	231	233
Frequency	0.125	0.143	0.339	0.179	0.018	0.036	0.071	0.018	0.071

This marker amplifies a polymorphic dinucleotide repeat $(CA)_n$ within the D4S2930 locus. The clone of this marker contains fourteen dinucleotide repeats corresponding to a 219 bp amplicon as illustrated in Table 3.9 (Dib *et al.*, 1996). The nine reported alleles therefore contain from twelve to twenty eight dinucleotide repeats.

Table 3.9: Partial gDNA sequence at locus D4S2930 encompassing marker AFMa224xh1

Nucleotide number[a]	Genomic DNA Sequence[b]
1	cctcatggta ggttaatccc acgctttgcc aaagattacc gtatctttac tcacatcccc
61	ttcctcatgt gtcctttcca cttattatca ttcnttgggt ttttctgtca cctcacccct
121	gtcatacatt ctcacacaca cacacacaca cacacacaca cgtgtatata caataagttt
181	ctattagggt agcattgtac aaatggcggg cattcaataa atgnttaata tgcattttaa
241	aacagtgtat gcaagngtaa tcatcnaaag tcatatgact cttctctttt catgtagcat
301	ccttcagaac tgcattatag aagggagata gct

a = Nucleotide positions are as reported by Dib *et al.* (1996). b = The sequence was deposited in Genbank with accession number Z52598. The positions of the primer set are indicated by underlined text (xxx). The STRP at this locus is indicated by double underlined text (yyy).

3.3.1.4 Marker ATA22F02 at locus D4S2390

This marker generates six alleles as presented in Table 3.10. The reported alleles range from 102 to 120 bp with the allele frequency of the 120 bp allele being twenty times less than that of the most frequent allele (111 bp).

Table 3.10: Allele frequency for marker ATA22F02 at locus D4S2390

Allele (bp)	102	108	111	114	117	120
Frequency	0.135	0.212	0.385	0.096	0.154	0.019

Marker ATA22F02 is the only trinucleotide marker included in this study. This marker amplifies a $(ATA)_n$ repeat at the D4S2390 locus. The deposited sequence in Genbank, presented in Table 3.11, contains twelve trinucleotide repeats, corresponding to an amplicon size of 111 bp (Murray *et al.*, 1995b).

Table 3.11: Partial gDNA sequence at locus D4S2390 encompassing marker ATA22F02

Nucleotide number[a]	Genomic DNA Sequence[b]
1	aactcctaag tttggctaga tacttaatgg cccaaactct ttgcaaataa gtgtggttat
61	attttagaca atagaatgca aacggatacc aacaccagtt tgtgattgtn tttttataag
121	gaggctgtgt gtctttctct ctcattttcc cctttccact aattgcaaaa tgatacataa
181	taataataat aataataata ataataataa tanncnggag gagttgcctt gcatctcatg
241	atgaaaacca caatcaaagg atgaaagagt caccctacta gcttgaattc cttgattatg
301	ttgtnggg na

a = Nucleotide positions are as reported by Murray *et al.* (1995b). b = The sequence was deposited in Genbank with accession number G08314. The positions of the primer set are indicated by underlined text (xxx). The STRP at this locus is indicated by double underlined text (yyy).

3.3.1.5 Marker UT5785 at locus D4S2299

The allele frequencies of five of the eight alleles have been reported on the CEPH-database as presented in Table 3.12. The allele sizes of the eight alleles were, however, not reported for this marker. The presence of 10 alleles in an extended family of 150 individuals, ranging from 201-244 bp have, however, been reported by the Eccles Institute of Human Genetics (Ballard, 1999). No allele frequency information was available for these 10 alleles.

Table 3.12: Allele frequency for marker UT5785 at locus D4S2299

Allele (bp)	---	---	---	---	---	---	---	---
Frequency	0.200	0.233	0.200	0.233	0.133	0.00	0.00	0.00

No allele sizes were reported (---).

Marker UT5785 at locus D4S2299 amplifies a polymorphic tetranucleotide repeat $(AGAT)_n$. The sequence presented in Table 3.13 includes twelve tetranucleotide repeat units generating an amplicon of 209 bp after PCR amplification (Gerken *et al.*, 1993b).

Table 3.13: Partial gDNA sequence at locus D4S2299 encompassing marker UT5785

Nucleotide number[a]	**Genomic DNA Sequence**[b]
1	gatcagccct cctacttttg gactcaggca gaattacacc acaggctttc ttgggtctcc
61	agctcacaga gggcagatgg tgggtcttct cagcttccat gagcatgtga accaatgccc
121	atgttaatta gtgtgtgcat atgtatacac acataccgat agattagata gacagatgat
181	agatgataga cagatgatag atgatagaca gatagatgat agatagatag atagatagat
241	agatagatag atagatagat agatagattc tggagaacct tgactaatac agttgggaat
301	gaagtgagtt gtcaaatctg acagaggtag tgtatgtaga ctgagaaaaa ttagaaaatg
361	aagcttcacc agtagaacgg cacaaaatta taagtgacca ttaaggctgg gagagtttga
421	tc

a = Nucleotide positions are as reported by Gerken *et al.* (1993b). b = The sequence was deposited in Genbank with accession number L18154. The positions of the primer set are indicated by underlined text (xxx). The STRP at this locus is indicated by double underlined text (yyy).

3.3.1.6 Marker UT2219 at locus D4S2283

As previously mentioned for markers UT1366 (locus D4S51523) and UT5785 (locus D4S2299), no allele sizes were available for the four alleles reported for marker UT2219

(Table 3.14). The frequencies of the four alleles range from 0.062 to 0.40. The allele with the lowest frequency is therefore six and a half (6.5) times less frequent than the most frequent allele.

Table 3.14: Allele frequency for marker UT2219 at locus D4S2283

Allele (bp)	---	---	---	---
Frequency	0.406	0.375	0.156	0.062

No allele sizes were reported (---).

A tetranucleotide polymorphic repeat $(CTAT)_n$ at the D4S2283 locus is amplified with marker UT2219. The sequence presented in Table 3.15 contains eleven tetranucleotide repeats which correspond with a 383 bp amplicon after PCR analysis (Gerken *et al.*, 1993c).

Table 3.15: Partial gDNA sequence at locus D4S2283 encompassing marker UT2219

Nucleotide number[a]	Genomic DNA Sequence[b]
1	cagccccgt tatttttcca tctactaatt tttgggacat agcatttcta cagcagaaat
61	gcatttacat ggcctctacc tgtgtggttg catctatcta tctatctatc tatctatcta
121	tctatctatc tatctatcca tccatccatc cattctccca tccatccatc cacccactca
181	cccactcatc tatctgtctt ctgaaataaa agtttaataa aatgacattt tcttttttg
241	taatatattt gtataatttc aattatatcc tggtcgattt ctttttaaaa aaaatgctgg
301	taattagtaa ctaaatttat tttacaactt atttttgtaa atcagtggca taaatagtag
361	cagttatgtg tctgcatttt gctttagatt

a = Nucleotide positions are as reported by Gerken *et al.* (1993c). b = The sequence was deposited in Genbank with accession number L17998. The positions of the primer set are indicated by underlined text (xxx). The STRP at this locus is indicated by double underlined text (yyy).

3.3.1.7 Marker UT7694 at locus D4S2688

Six alleles, with frequencies as presented in Table 3.16, have been reported for marker UT7694 at locus D4S2688. The allele sizes of the six alleles were, however, not reported. The frequencies of only five of the six alleles have been reported with frequencies range from 0.062 to 0.375.

Table 3.16: Allele frequency for marker UT7694 at locus D4S2688

Allele (bp)	---	---	---	---	---	---
Frequency	0.375	0.281	0.219	0.062	0.062	0.00

No allele sizes were reported (---).

Marker UT7694 amplifies a tetranucleotide repeat $(AGAT)_n$ at locus D4S2688 as illustrated in Table 3.17. An amplicon of 246 bp, consisting of six tetranucleotide repeats, is generated after PCR analysis (Gerken *et al.*, 1994). A second repeat sequence (AGAC) can be observed in the sequence presented in Table 3.17. Upon STRP analysis it was, however, observed that this repeat is not polymorphic.

Table 3.17: Partial gDNA sequence at locus D4S2688 encompassing marker UT7694

Nucleotide number[a]	Genomic DNA Sequence[b]
1	tcgaactcct gacctcaggt gatccacccg cttcagcctc ccaaagtgct gggaatatag
161	gcgtgagcan ctgcacctgg cccacacatt aagtttacaa aacaagacac agaaggggaa
121	actatggaat aacaggacta gaaagagtag gagattttgt aaatgttttt attatgaaaa
181	tatttgatac acataagaga atgtttgtga cagatgtata tttatataca cacaaacata
241	tacatctgta gataggatgg atagataaga ttagatagaa gatagataga tagagagata
301	gatagataga cagacagaca gacagatagc agntggggta ggttttcatg aaacaaacac
361	cattgacccc ttcacccaca ctaagacatg gaacgtcact gttatctctg aagctcctgt
421	tcgcttctgt tacttcatcc ctggccctgg gttttctcca agagttccat tcacatacgt
481	gttgcacagc tgttaataca ctgcctagtg tgcttggctt tgaatttata atacagtatc
541	atgctgcata tctttccaca acttgcttcc acaatgtatt aactttctga gtttcga

a = Nucleotide positions are as reported by Gerken *et al.* (1994). b = The sequence was deposited in Genbank with accession number L30309. The positions of the primer set are indicated by underlined text (xxx). The STRP at this locus is indicated by double underlined text (yyy). A non polymorphic repeat is indicated by dashed underlined text (zzz).

3.3.1.8 Marker AFMa190zf5 at locus D4S2921

Table 3.18 displays the allele sizes and frequencies of the eight alleles reported for marker AFMa190zf5 at locus D4S2921. The allele sizes range from 141-163 bp, with the 151 bp allele being 36 time more frequent than the 153 bp and 163 bp alleles. The smallest allele (141 bp) is the second most frequent and is six time less frequent than the most frequent allele (151 bp).

Table 3.18: Allele frequency for marker AFMa190zf5 at locus D4S2921

Allele (bp)	141	151	153	155	157	159	161	163
Frequency	0.107	0.643	0.018	0.036	0.036	0.071	0.071	0.018

This STRP marker amplifies a $(CA)_n$ repeat at the D4S2921 locus. Tale 3.19 displays the sequence of the clone for this marker containing fourteen dinucleotide repeat units concurring with a 151 bp fragment after PCR analysis (Dib *et al.*, 1996).

Table 3.19: Partial gDNA sequence at locus D4S2921 encompassing marker AFMa190zf5

Nucleotide number[a]	Genomic DNA Sequence[b]
1	cctcttcctt caggaactgg tgcantttca tgtattctga cacaanaacc tgtttacaag
61	tctttgtttc ntcctgccaa aacacacaca cacacacaca cacacacaca cgagataaag
121	ttagtaagag aaatgccctt gtctgtagat ttttaancct tacaagtcct agagtttctt
181	tctttcnttt tttttttttt ttgagatgga gtctcactct gtcgccagtc tgaagtgcag
241	tggcgcaatc tcggctcact gcaacctctg cctcctgagt tcaaatgatt ctcctgcctc
301	agcctncctg agtagct

a = Nucleotide positions are as reported by Dib *et al.* (1996). b = The sequence was deposited in Genbank with accession number Z52419. The positions of the primer set are indicated by underlined text (xxx). The STRP at this locus is indicated by double underlined text (yyy).

3.3.1.9 Marker AFM238ve3 at locus D4S426

The allele frequencies of the six reported alleles are listed in Table 3.20. The alleles range from 177 to 191 bp, with the 189 bp allele being the least frequent and the 179 bp allele the most frequent of the six alleles. The 181 bp and 191 bp have equal reported frequencies.

Table 3.20: Allele frequency for marker AFM238ve3 at locus 426

Allele (bp)	177	179	181	187	189	191
Frequency	0.232	0.375	0.107	0.125	0.054	0.107

This dinucleotide marker amplifies a $(CA)_n$ repeat at locus D4S426. The sequence presented in Table 3.21 includes sixteen dinucleotide repeats generating an amplicon of 189 bp upon PCR analysis (Weissenbach *et al.*, 1992).

Table 3.21: Partial gDNA sequence at locus D4S426 encompassing marker AFM238ve3

Nucleotide number[a]	Genomic DNA Sequence[b]
1	agcttctact cnnttagaca atttnnagta tacactgcat ccatatatac aaggnggcag
61	tgaatacttg aaattgtctg agacagtaga ncttnggtgt cctaaccaca tacacataat
121	acacanacac acacacacac acacacacac acacacacag aggtaactat gtgtgatgat
181	gattgttaaa taacttgact gtggtcattt cacaatgtaa acatatctca aattaccata
241	ttgtattact tgaatatata taattttaat ttatcaatta tacctaaata aagct

a = Nucleotide positions are as reported by Weissenbach *et al.* (1992). b = The sequence was deposited in Genbank with accession number Z17062. The positions of the primer set are indicated by underlined text (xxx). The STRP at this locus is indicated by double underlined text (yyy).

3.3.2 5' - End labelling of PCR primers

Sixty pico moles (pmol) of the selected primer were 5'-end labelled with γ^{32}P-dATP and polynucleotide kinase [Promega]. Reaction mixtures contained 70 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 5 mM Dithiothreitol (DTT), 5 units polynucleotide kinase and 60 μCi γ^{32}P-dATP [Amersham Life Science Inc.][1] in a total volume of 14 μl. After incubation at 37°C for one hour, 26.0 μl ddH_2O was added. The end labelled primers were stored at -20°C, until required. All of the markers were end labelled, except markers D4S2390 and D4S2930 for which multiplex PCR and internal incorporation were utilised.

3.3.3 The polymerase chain reaction (PCR)

PCR conditions were optimised for each primer set with regard to the annealing temperature and the magnesium chloride ($MgCl_2$) concentrations. The optimisation reactions were performed under non-radioactive conditions.

A modified method described by Mullis and Faloona (1986) was utilised to perform PCR. The PCR reactions were prepared in a total volume of 12.5 μl in 0.5 ml thin wall PCR tubes [Hybaid] or 96 well plates [Sterilab]. The PCR reagents were thawed on ice and thoroughly mixed by vortexing prior to use. PCR were performed in reactions containing 10 mM Tris-HCl (pH 8.3), 1.0-1.5 mM $MgCl_2$, 50 mM KCl, 200 μM of each nucleotide [2'-deoxyadenosine-5'-triphosphate (dATP), 2'-deoxycytidine-5'-triphosphate (dCTP), 2'-deoxyguanosine-5'-triphosphate (dGTP) and 2'-deoxythymidine-5'-triphosphate (dTTP)], 60 nM (0.75 pmol) ^{32}P end labelled primer, 40 nM (5.0 pmol) unlabelled primer, 0.5 units (U) Taq polymerase [Promega]. 100 ng gDNA was added and the reactions were overlaid with a drop of mineral oil to prevent evaporation.

A standard PCR program consisted of: denaturation at 94°C for ten minutes, and 30 cycles consisting of denaturation at 94°C for 30 seconds (sec), annealing at the optimised temperature for 30 sec and extension at 72°C for 30 sec. A final extension of 7 minutes at 72°C ensured that all PCR products were full length products. The PCR reactions were performed in a Hybaid Touchdown™ [2] thermocycler.

---

[1] Amersham™ is a trademark of Amersham International plc., Buckinghamshire, UK.
[2] Touchdown™ is a trademark of Hybaid Limited, Ashford, Middlesex, UK.

After PCR amplification was completed an equal volume of 2X stop buffer (95% formamide; 0.05% xylene cyanol FF (XC); 0.05% bromophenol blue (BPB); 20 mM EDTA) was added to each sample. The samples were denatured at 85°C for 5 minutes and placed on ice before loading on a denaturing polyacrylamide gel. Whenever the PCR products could not be electrophoresed immediately, the products were stored at 4°C.

3.3.4 Multiplex PCR

Multiplex PCR was performed for three of the nine STRP markers (D4S2921, D4S2390 and D4S2930). The primers of these markers were not end-labelled, but internal incorporation was performed instead. The PCR reaction mixture was prepared as described in paragraph 3.3.2, without adjusting any of the reagents except for using α^{32}P-dCTP instead of the end-labelled primer. The PCR program was also the same as described in paragraph 3.3.2, except for increasing the annealing time to 60 sec. Stop buffer was added to each sample as described in paragraph 3.3.2.

3.3.5 Single stranded DNA sequencing

DNA sequencing was performed according to the chain termination protocol described by Sanger *et al.* (1977). Sequencing was performed with the Sequenase® Version 2.0 (Sequenase) kit [USB]. The control DNA included with the kit was sequenced as size reference for accurate allele sizing. The included DNA is a single-stranded phage DNA from a clone of a *Sau* 3AI fragment of bacteriophage λ DNA inserted at the *Bam* HI site of M13mp18.

Annealing of the template and primer was accomplished by adding 1 µg of the template, 3.33 pmol of the –40 forward 23-mer primer and ddH_2O, up to a final volume of 10 µl. The mixture was incubated at 99.9°C for 3 minutes and placed on ice for 5 minutes. The rapid cooling promotes primer annealing over template re-annealing. The labelling solution contained 2 µl of a 5X T7 Sequenase reaction buffer (200 mM Tris-HCl (pH 7.5), 100 mM $MgCl_2$, 250 mM NaCl), 0.1 µmol DTT suspension, 3 µM dATP, 3 µM dGTP, 3 µM dTTP, 5 µCi α^{32}P-dCTP and 3.2 U Sequenase enzyme.

The labelling solution was added to the 10 µl annealing mixture and incubated at room temperature for 3 minutes. Two and a half microlitres of the four termination mixtures (80 µM of each dNTP, 8.0 µM of one ddNTP, 50 mM NaCl) were aliquoted into four separate 0.5 ml eppendorf tubes. The labelling reactions were terminated by adding 3.5 µl of each sample to the four termination tubes. The sequencing reactions were incubated at 37°C for 7 minutes after which 4 µl of a 2X stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF) were added to each tube. The reactions were stored at 4°C until loaded onto a denaturing polyacrylamide gel.

3.3.6 Denaturing gel electrophoresis and autoradiography

A gel stock solution containing an acrylamide monomer concentration of 6% was prepared, containing 75 ml of a 40% acrylamide mixture (acrylamide [Stratagene]/ bis-acrylamide [Promega] (19:1) and 210 g (7 M) urea [USB] dissolved in TBE buffer (89.15 mM Tris base [USB]; 88.95 mM Boric acid [USB]; 2.498 mM Na_2EDTA [ACE], at a pH of 8.0-8.3). A catalyst (30 µl TEMED) and an oxidising agent (500 µl of a 10% ammonium persulphate solution) were added to 75 ml gel stock solution prior to casting of the 0.4 mm thick gel to achieve polymerisation.

The gel was left for at least an hour to allow complete polymerisation, after which it was pre-run at 60 watts until the temperature of the gel reached 45-50°C (30 min-45 min). Samples were loaded and electrophoresed in 1X TBE buffer at 45 watts for the required length of time. After termination of electrophoresis, the gels were transferred to Whatman®[1] 3MM Chromatography blotting paper, covered with Cling Wrap and vacuum-dried at 80°C for approximately 45 min. The dried gels were exposed to X-Ray film at room temperature for the required length of time.

[1] Whatman® is a registered trademark of Whatman Scientific Ltd., Kent, UK.

3.4 SOUTHERN BLOT ANALYSIS

Southern blot analysis was utilised to detect the FSHD associated DNA rearrangements at the D4Z4 locus in the South African FSHD families. Probe p13E-11 was utilised for the detection of the deletion fragments and was kindly donated to the South African FSHD research project by Dr. Silvère van der Maarel, Department of Human Genetics, Leiden University, Medical Center.

3.4.1 Restriction fragment length polymorphism (RFLP) analysis

Restriction fragment length polymorphism (RFLP) analysis was performed on high molecular weight gDNA. Two restriction enzyme digestion reactions were performed for each patient. Five micrograms of gDNA were utilised in each reaction. Reactions were incubated at 37°C for at least six hours, in the presence of 25 U *Eco* RI restriction endonuclease [Amersham], 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 100 mM NaCl, 1 mM dithiothreithol and 3.3 mM spermidine [Sigma[1]] in a final volume of 30 µl. The double restriction enzyme reaction contained 25 U *Eco* RI restriction endonuclease [Amersham] and 25 U *Bln* I restriction endonuclease [Amersham], 2X one-phor-all buffer (20 mM Tris-acetate (pH7.5), 20 mM magnesium acetate, 100 mM potassium acetate), to prevent star activity [Amersham], and 3.3 mM spermidine in a final volume of 40 µl.

3.4.2 Agarose gel electrophoresis

Two and a half microlitres of the restriction endonuclease digestion reaction was mixed with 1 µl of a 2X loading buffer (0.04 % Orange G [Sigma®] and 50 % glycerol) prior to loading on a gel. The 0.5% agarose gel contained 1.25 g molecular grade agarose [Promega] and 1.268 µM EtBr [Sigma®]. The agarose was dissolved in 1X TBE buffer in a total volume of 25 ml. The mini submarine agarose gel was electrophoresed in 1X TBE buffer at 100 V for 30 minutes and then visualised via UV transillumination. After confirmation of complete digestion the remaining product was loaded onto a maxi or midi (both 23 cm in length) 0.5% submarine agarose gel along with a molecular weight marker to determine the fragment sizes. The agarose gel was electrophoresed in 1X TBE buffer at 35V (1.5V.cm^{-1}) for 30-48 hours. After 24 hours, the 1X TBE buffer was replaced.

[1] Sigma® is a registered trademark of Sigma Chemical Company, St. Louis, MO, U.S.A.

3.4.3 Genomic DNA Transfer

The agarose gel was removed from the gel apparatus and stained for 20 minutes in 2 μg.ml^{-1} EtBr dissolved in 1X TBE buffer (if not added directly to gel mixture) with gentle shaking after which it was visualised under an UV light before blotting. The gDNA was subsequently nicked with 0.25 M HCl for 15 minutes and then denatured with fresh alkaline blotting solution (0.4 M NaOH, 1.6 M NaCl) twice for 20 minutes with gentle shaking. Hybond™[1] N^+ membrane [Amersham] was simultaneously prewet for 10 minutes in distilled water and subsequently 15 minutes in alkaline blotting solution. The DNA was transferred to the Hybond N^+ membrane overnight utilising the alkaline blotting solution and capillary action. After Southern transfer, the blot was neutralised for 5 minutes in neutralization solution (0.2 M Tris-HCl (pH 8.0), 2X SSC - 1.75% NaCl [ACE], 0.88% Na-citrate [Merck]). Fixation of the DNA was accomplished by baking the membrane for 2 hours at 80°C. The membrane was subsequently heat sealed in a plastic bag and stored at 4°C until hybridisation. The transfer efficiency was determined by staining the gel after the overnight transfer with 2 μg.ml^{-1} EtBr dissolved in 1X TBE buffer for 30 minutes followed by visualisation via UV light.

3.4.4 Isolation of p13E-11

Probe p13E-11 was cloned as an 800 bp insert into the pBluescript plasmid. Five millilitres of LB-broth containing ampicillin was inoculated with the p13E-11 stab culture and incubated with shaking overnight at 37°C. Five millilitres of the overnight culture was added to 250 ml of LB broth containing ampicillin. The larger culture was incubated overnight at 37°C with shaking. The Wizard® Purefection Plasmid DNA Purification system was utilised for the isolation of the plasmid. Two reactions of 40 ml each were performed and the remaining culture was frozen at 70°C in the presence of 15% glycerol for future use.

The bacterial cells were pelleted by centrifugation at 10,000 xg for 10 minutes at 22°C. The supernatant was discarded and the excess liquid was blotted on a paper towel. Six millilitres of Cell Resuspension Solution was added to each tube and the cell pellet

[1] Hybond™ is a trademark of Amersham International plc., Buchinghamshire England.

resuspended by vortexing. The cells were lysed by adding 6 ml of Cell Lysis Solution and mixed thorougly by inverting the tube 6-8 times. The mixture was incubated at 22°C for 5 minutes after which 8 ml of Neutralization Solution was added and the tube again thoroughly mixed by inverting it 6-8 times. The bacterial lysate was centrifuged twice at 10,000 xg for 20 minutes at 22°C and the supernatant was then transferred to a clean 50 ml centrifuge tube.

After thorough resuspension of the Endotoxin Removal Resin 1 ml was added to the supernatant. The reaction was incubated for 10 minutes at 22°C, with vigorous shaking for 5 seconds at several intervals during the incubation. The tube was placed onto the MagneSil™[1] Magnetic Separation Unit for 30 seconds until the solution turned clear. While keeping the tube on the magnet, the supernatant was transferred to a new tube. Four and a half millilitres of 5 M guanidine thiocyanate was added to the supernatant followed by an addition of 3.5 ml of MagneSil™ Paramagnetic Particles. The reaction was mixed and incubated at 22°C for 3 minutes. The tube was placed onto the magnetic unit and the solution allowed to clear for 30 seconds before the supernatant was discarded. The tube was placed on the magnetic unit for another 3 minutes and the residual liquid discarded.

The particles, to which the plasmid DNA was bound, were then washed with 4.5 ml of 4.2 M guanidine-HCl/40% isopropanol solution. The particles were completely resuspended by vortexing for 10 seconds. The tube was placed on the magnetic unit and allowed to clear for 30 seconds after which the supernatant was discarded and the tube placed onto the magnetic unit for another 3 minutes and the residual liquid discarded.

Ten millilitres of 80% ethanol wash solution was added and the particles completely resuspended by vortexing for 10 seconds. The tube was placed on the magnetic unit and the solution allowed to clear for 30 seconds after which the supernatant was discarded. The washing step was repeated for a total of 3 washes. After the final wash, the 80% ethanol wash solution was carefully removed and the tube was left open for 10 minutes on the magnet to allow evaporation of any remaining ethanol. Any residual liquid was removed from the bottom of the tube. The tube was removed from the magnet and 6.0 ml

---

[1] MagneSil™ is a trademark of Promega Corporation, Madisin, WI, U.S.A.

high quality water (freshly distilled and free from enzymes) was added to the particles and vortexed for 10 minutes. The reaction was incubated at 22°C for 1 minute and then placed onto the magnet where the solution was allowed to clear. The supernantant was transferred to a sterile 15 ml centrifuge tube.

The DNA was precipitated by the addition of 0.5 volume of 7.5 M ammonium acetate and 2.5 volume of 95 % ethanol. The DNA was collected by centrifugation at 14,000 xg for 15 minutes at 22°C and the supernatant discarded. The pellet was then rinsed in 250 µl 70% ethanol followed by centrifugation at 14,000 xg for 5 minutes. The ethanol was carefully aspirated and the pellet air-dried for 5 minutes. The DNA was resuspended in 1.5 ml high quality water and the concentration determined.

The insert was excised from the plasmid through a double digestion with *Sac* I [Amersham] and *Eco* RI [Amersham]. The restriction enzyme digestions were performed in reactions containing 1.7 µg plasmid DNA, 20 U *Sac* I, 10 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$ and 1 mM dithiothreitol. The reactions were incubated at 37°C for 2 hours followed by inactivation of the enzyme by incubation at 65°C for 15 minutes. Subsequently 30 U *Eco* RI, 50 mM Tris-HCl (pH 8.5), 10 mM $MgCl_2$, 1 mM dithiothreitol and 100 mM NaCl was added. The reactions were incubated at 37°C for 2 hours. A portion of the reaction was loaded on a 1% agarose gel to determine the efficiency of the digestions. Upon verifying that the plasmid was completely digested the remaining product was loaded on a 0.8 % low melt agarose gel [FMC]. Agarose gel electrophoresis was performed as described in paragraph 3.4.2. The 0.8 kb insert was excised from the gel with a razor blade and was placed in an eppendorf tube. Three millilitres water per gram of gel was added and the mixture placed in a boiling water bath for 7 minutes to melt the gel and denature the DNA. The DNA was divided into aliquots and stored at -20°C.

3.4.5 Radio-active detection

The fragments on the membrane were detected after hybridisation with a radio-actively labelled probe. The membrane was washed to remove any non-specific binding and exposed to X-ray film for visualisation.

3.4.5.1 Labelling of probe p13E-11 and molecular weight markers

Twenty five nanograms of the p13E-11 probe DNA was labelled by the random nonamer priming method (Multiprime Kit, Amersham) with α-^{32}PdCTP for each hybridisation reaction. The DNA was denatured by heating to 95°C for one minute then chilled on ice. Ten µl reaction buffer containing dATP, dGTP, and dTTP in a concentrated buffer solution with Tris-HCl (pH 7.8), magnesium chloride and 2-mercaptoethanol (the exact concentrations were not provided in the kit), 5 µl primer (the primer concentration was not provided in the kit), 2 U enzyme, 50 µCi α-^{32}PdCTP and the appropriate amount of water to give a final reaction volume of 50 µl were added to the denatured DNA on ice. The reactions were incubated at room temperature overnight.

DNA molecular weight marker XV [Roche] and DNA molecular weight marker III [Roche] were utilised for allele sizing in this study. Molecular weight marker XV consists of fragments ranging from 2,392 to 48,502 bp, while molecular weight marker III consists of fragments ranging from 125 to 21,226 bp. The molecular weight markers were initially also labelled by random priming, but the signal intensity was too strong, and end-labelling was therefore utilised instead.

3.4.5.2 Hybridisation conditions

The membrane was prehybridised for at least 30 minutes at 65°C in 40 ml hybridisation solution (0.125 M Na_2PO_4 (pH 7.2) [USB], 0.25 M NaCl [ACE], 1 mM EDTA [ACE], 7% SDS [USB], 10% PEG-6000 [ACE] and supplemented with 10 $\mu g.ml^{-1}$ denatured salmon sperm DNA [Sigma]. The labelled probe and molecular weight markers were denatured for 10 minutes at 95°C and placed on ice prior to hybridisation. After addition of the denatured probe and molecular weight markers to the prehybridisation solution, the blot was incubated overnight (16 hours) at 65°C in a Techne hybridisation oven.

3.4.5.3 Wash conditions

After overnight incubation with the denatured probe and molecular weight markers, the blot was washed three times for 5 minutes with washing solution (2X SSC (1.75% NaCl, 0.88% Na-citrate), 1% SDS) at 65°C. The membrane was heat sealed in a plastic bag, after removal of excess air bubbles.

3.4.5.4 Autoradiography

The hybridised membrane was exposed to X-ray film (Fuji RX-U) utilising an intensifying screen at -70°C for the appropriate time. The X-ray film was developed and fixed to allow the visualisation of the fragments.

3.4.6 Non radio-active detection

The Detector™[1] Random Primer DNA biotinylation kit and the DNADetector™ Genomic Southern Blotting kit [Kirkegaard & Perry Laboratories] were utilised for the non radio-active detection of the DNA rearrangements at the D4Z4 locus. The random primer kit utilises six base random sequence oligonucleotides as primers for the replication of the template DNA. The probes are labelled through the incorporation of biotin-N_4-dCTP during the random-primer extension by Klenow DNA polymerase. The labelled probe is then detected with the DNADetector™ kit.

3.4.6.1 Labelling of probe p13E-11

One hundred nanograms of probe DNA was dissolved to a final volume of 24 µl in DEPC treated water and mixed with 15 µg of 2.5X random primer solution. The DNA was denatured by heating the sample to 95°C for 5 minutes and immediately transferred to ice. 5 µl of the 10X dNTP mixture (0.1 mM biotin–N_4-dCTP, 0.1 mM dCTP, 0.2 mM dATP, 0.2 mM dGTP and 0.2 mM dTTP in 1 mM Tris-HCl (pH 7.5) and 0.1 mM Na_2EDTA) and 10 U Klenow polymerase was added to obtain a final volume of 50 µl. The reaction was gently mixed and incubated at 37°C for 1-4 hours. One molar Na_2EDTA (pH 7.5) was subsequently added to terminate the reaction.

3.4.6.2 Quantification of labelled probe

The amount of labelled probe was quantified utilising the DNADetector™ Genomic Southern Blotting Kit [KPL]. This kit was designed for the hybridisation and

---

[1] Detector™ is a trademark of Kirkegaard & Perry Laboratories, Inc, Gaithersburg, MD, U.S.A.

chemiluminescent detection of biotinylated DNA probes. An initial 1:200 dilution, with 2X SSC/0.1% SDS as the diluent, of the labelled probe was made after which a series of two-fold dilutions from both the 1/200 probe dilution and the undiluted biotinylated DNA standard were prepared. One microlitre of each dilution was spotted on a membrane which was pre-soaked in water and 5X SSC for 10 minutes where after it was dried for 10 minutes. The spotted dots were subsequently dried for 10 minutes and fixed onto the membrane by baking for 30 minutes at 80°C.

A blocking/diluent solution was prepared by diluting the 5X Detector Block solution 1:5 in ddH_2O and adding 0.2% (w/v) Detector Block Powder (DBP). The DBP was completely dissolved in a 50°C waterbath and cooled to room temperature before use. The dot blot was incubated for 30 minutes with 0.1 ml diluted blocking solution per cm^2 membrane after which the dot blot was incubated for 30 minutes in 0.1 ml per cm^2 membrane phosphatase-labelled streptavidin (AP-SA) diluted 1:10,000 in blocking/diluent solution. A 1X Phosphatase Wash Solution was prepared and the membrane was washed 3 times, for 5 minutes each, using at least 1 ml of wash solution per cm^2 membrane. The membrane was subsequently rinsed twice for 2 minutes each in 1X Phosphatase Assay Buffer (PAB) using at least 0.5 ml per cm^2 membrane. The excess PAB was drained off and the membrane was incubated in CDP-Star®[1] Chemiluminescent Substrate using at least 0.05 ml per cm^2 membrane for 5 minutes. Excess CDP-Star was removed by blotting the membrane on filter paper after which the membrane was heat sealed and exposed to X-ray film for ca. 2 minutes.

The total yield of biotinylated DNA probe was then determined by the comparison of the signal intensities of the biotinylated DNA standard and the biotinylated DNA probe. The ratio of probe to standard (P/S) was determined by the following equation:
P/S = probe endpoint dilution / standard endpoint dilution, after which the relative specific activity of the labelled probe was calculated by: Probe RSA = standard RSA x P/S. The relative specific activity (RSA) of the labelled probe is defined by the relative amount of biotin attached to the DNA. The total yield of biotinylated DNA probe could then be determined by: Total yield = probe RSA x reaction volume. The labelling reaction generally generates a 5 to 10 fold amplification of the template after one hour, and 10 to 50 fold amplification after 4 hours.

---

[1] CPD-Star® is a registered trademark of Tropix Inc., Bedford, MA, U.S.A.

3.4.6.3 Hybridisation conditions

The target DNA was immobilised on a Hybond N^+ membrane and prehybridised in 0.10 ml of Formamide Hybridisation Buffer (supplemented with 200 $\mu g.ml^{-1}$ Salmon Sperm DNA) per cm^2 membrane for 1 hour at 42°C. Fifty $ng.ml^{-1}$ of the biotinylated probe was denatured at 95°C for 10 minutes and immediately placed on ice. The denatured probe was added directly into the prehybridisation solution and incubated for 16 hours at 42°C.

3.4.6.4 Wash conditions

Following hybridisation the membrane was washed to remove any excess probe. The membrane was washed twice for 15 minutes in at least 1 ml per cm^2 of membrane 2X SSPE [0.3 M NaCl, 20 mM NaH_2PO_4, 2 mM EDTA (pH 7.4)]/0.1% SDS at room temperature. A second wash with 0.2X SSPE [0.03 NaCl, 2 mM NaH_2PO_4, 0.2 mM EDTA (pH 7.4)]/0.1% SDS at 55°C was performed twice for 15 minutes. Residual SDS was removed with a final wash with 2X SSPE at room temperature for 5 minutes.

3.4.6.5 Detection

The membrane was subsequently incubated in 1X blocking/diluent solution (1X detector block solution supplemented with 1% Detector Block Powder) for 45 minutes (0.25 ml 1X blocking/diluent solution per cm^2 membrane) and subsequently incubated for 30 minutes in 0.25 ml per cm^2 membrane AP-SA solution (AP-SA conjugate diluted 1:10,000 in 1X blocking/diluent solution). A 1X Phosphatase Wash Solution was utilised to wash the membrane three times for 5 minutes each using 0.75 ml of wash solution per cm^2 membrane. The membrane was rinsed for 5 minutes in 1X Phosphatase Assay Buffer using 0.4 ml per cm^2 membrane after which the CDP-Star was added (0.04 ml per cm^2 membrane) and incubated for 5 minutes. The excess CDP-Star was removed by blotting the membrane on filter paper after which it was heat sealed in a hybridisation bag. The membrane was exposed for 10 minutes to X-ray film at room temperature where after the X-ray film was developed and fixed.

CHAPTER FOUR

RESULTS AND DISCUSSION

In this chapter haplotype and Southern blot analyses of five South African FSHD families are presented. The haplotypes were constructed by genotyping short tandem repeat polymorphism (STRP) markers located at chromosome 4q35. Southern blot analysis was performed to determine the DNA rearrangements at the D4Z4 locus. Results presented in this chapter have been presented at several national and international conferences and selected abstracts have been published in international peer-reviewed journals (Appendix A).

4.1 SHORT TANDEM REPEAT POLYMORPHISM (STRP) ANALYSIS

Nine STRP markers were genotyped in five FSHD families. Primer sets were optimised with regard to the annealing temperature and $MgCl_2$ concentration. Annealing temperatures were calculated for each of the two primers in a primer set, with the formula reported by Thein and Wallace (1986). Table 4.1 lists the calculated and optimised annealing temperature and $MgCl_2$ concentration for all the markers. The primer sequences and other relevant data for these markers were presented in Table 3.3 and in paragraph 3.3.1.

Table 4.1: Optimised conditions for nine short tandem repeat polymorphism markers located on chromosome 4q35

Locus	DNA marker	Calculated T_a of forward primer	Calculated T_a of reverse primer	Optimised T_a	Optimised $MgCl_2$ concentration
D4S1523F	UT1366	60	58	55	1.5
D4S1652F	GATA5B02	62	64	55	1.5
D4S2930F	AFMa224xh1	70	58	55	1.5
D4S2390F	ATA22F02	58	60	55	1.5
D4S2299F	UT5785	56	56	56	1.0
D4S2283F	UT2219	60	58	58	1.5
D4S2688F	UT7694	54	58	55	1.5
D4S2921F	AFMa190zf5	52	60	55	1.5
D4S426F	AFM238ve3	68	66	55	1.5

T_a = annealing temperature.

A representative autoradiograph of each of the nine STRP markers utilised in this study is presented in subsequent paragraphs. The STRP markers are discussed in order of their genetic map position from the most telomeric to the most centromeric (Table 3.2). Size standards, consisting of the A, C, G and T sequencing reaction mixture of bacteriophage M13mp18, were loaded on each gel to enable precise sizing of marker alleles. Two of the sequencing reaction mixtures were mixed prior to loading on the gel (A with G and C with T). Allele size standards were also generated by mixing a portion of the PCR products of ca. six unrelated individuals. The allele size standard was loaded three or four times on the gel as a standardised reference throughout the gel to facilitate in the scoring of the alleles.

4.1.1 Short tandem repeat marker UT1366 at locus D4S1523

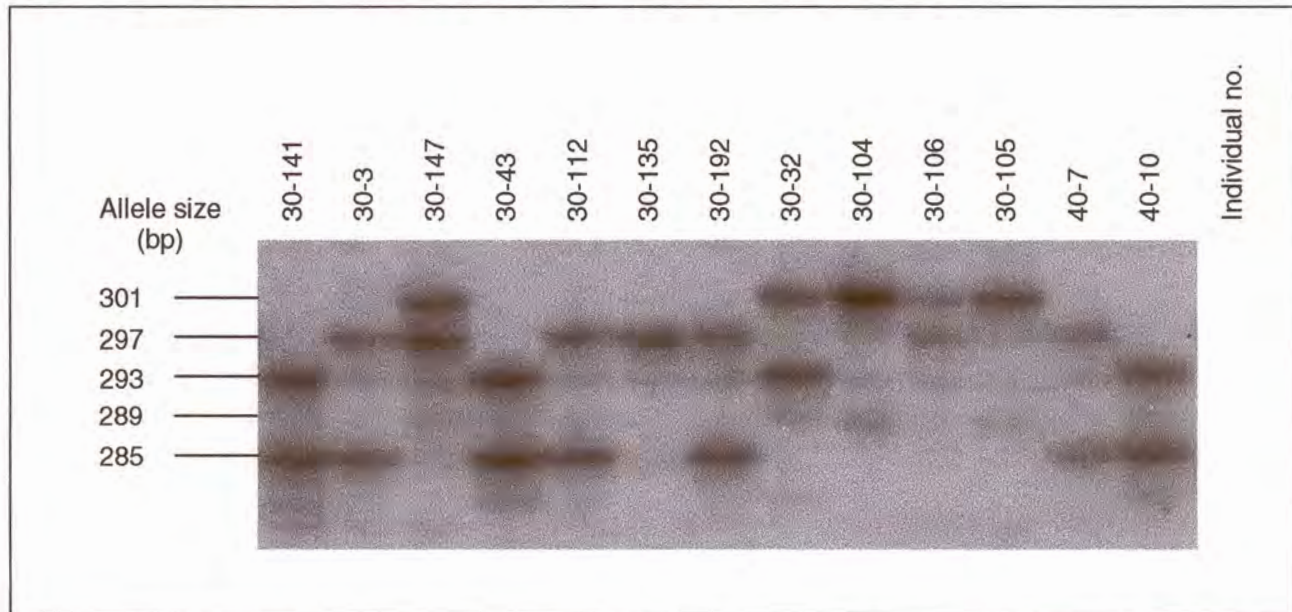
The primers of this tetranucleotide STRP marker were annealed at 55°C with the reverse primer selected for end labelling. Six alleles were observed in this study. The allele sizes of the seven reported alleles have never been published. The allele frequencies of only five of the alleles have, however, been reported (as discussed in paragraph 3.3.1.1). It was therefore not possible to correlate the six alleles obtained with the seven reported alleles. The allele sizes of marker UT1366 were thus determined for the first time through this investigation. Once the allele sizes have been reported on the CEPH databases, CEPH samples, genotyped with this marker, can be obtained to correlate the allele sizes generated in this study with those reported.

A representative autoradiograph of the amplification products generated with this marker is presented in Figure 4.1. Marker UT1366 generated strand slippage products that differ in size by four base pairs from the true allele. The secondary amplification products did not complicate scoring, as the allele motif of this marker could be discerned, and the alleles could be identified with ease.

A constant secondary fragment of 293 bp can be observed for all the individuals presented in Figure 4.1. Individuals 30-104 and 30-105 were scored to be homozygous for the 301 bp allele. A less intense fragment of 289 bp was also observed for both individuals. However, this was not scored as a true allele, since it was observed that this fragment

(289 bp) was of equal intensity to the 301 bp allele in heterozygous individuals (not presented in Figure 4.1).

Figure 4.1: Representative autoradiograph of marker UT1366



Autoradiograph of the amplification products generated with marker UT1366 described by Gerken *et al.* (1993a). The reverse primer was end labelled with γ^{32}P-dATP and the products electrophoresed on a 6% polyacrylamide gel in 1X TBE buffer for 3.5 hours. Individuals genotyped are indicated at the top and allele sizes are indicated to the left of the autoradiograph.

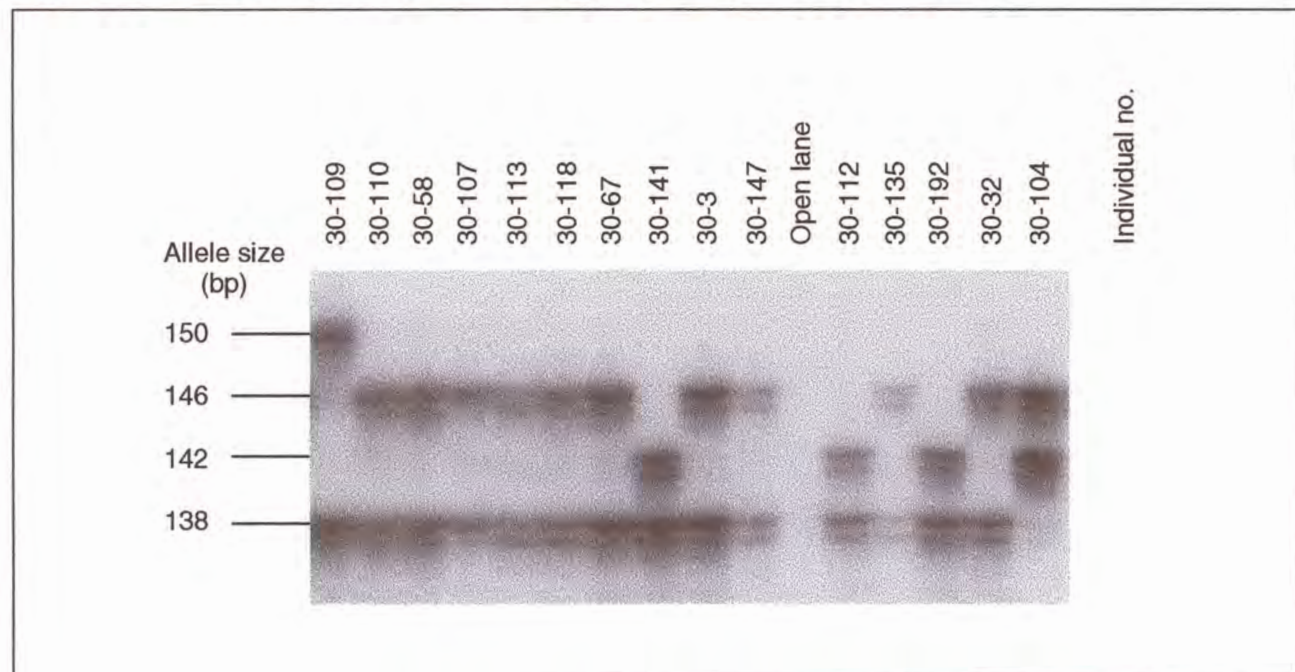
4.1.2 Short tandem repeat marker GATA5B02 at locus D4S1652

Marker GATA5B02 amplifies a tetranucleotide repeat $(CTAT)_n$ at locus D4S1652. This marker generates four alleles, ranging from 138 to 150 bp, and all four alleles were observed in this study (Figure 4.2). A heterozygosity value of 69 % was reported for this marker. As presented in Figure 4.2 most of the individuals investigated were heterozygous at this locus.

The primer set was annealed at 55°C, with the forward primer chosen for end labelling. It is evident from the autoradiograph presented here that a secondary amplification product, also referred to as "shadow bands" or "strand slippage", was generated for each allele, which migrated one base pair faster than the true allele. This marker therefore generates two fragments per allele. Hauge and Litt (1993) reported that the extra fragments are generated during PCR as a result of slipped strand mispairing. During slipped strand mispairing the primer-template complex misaligns therefore forming a heteroduplex DNA causing the two strands of DNA to differ in length by one or multiple repeat units. The extra fragments can complicate scoring when superimposed on true alleles. However, in

the case of this marker, strand slippage did not interfere with the scoring and the allele motif was discernible with ease.

Figure 4.2: Representative autoradiograph of marker GATA5B02



Autoradiograph of the amplification products generated with marker GATA5B02 described by Murray *et al.* (1995a). The forward primer was end labelled with γ^{32}P-dATP and the products electrophoresed on a 6% polyacrylamide gel in 1X TBE buffer for 2 hours. Individuals genotyped are indicated at the top and allele sizes are indicated to the left of the autoradiograph.

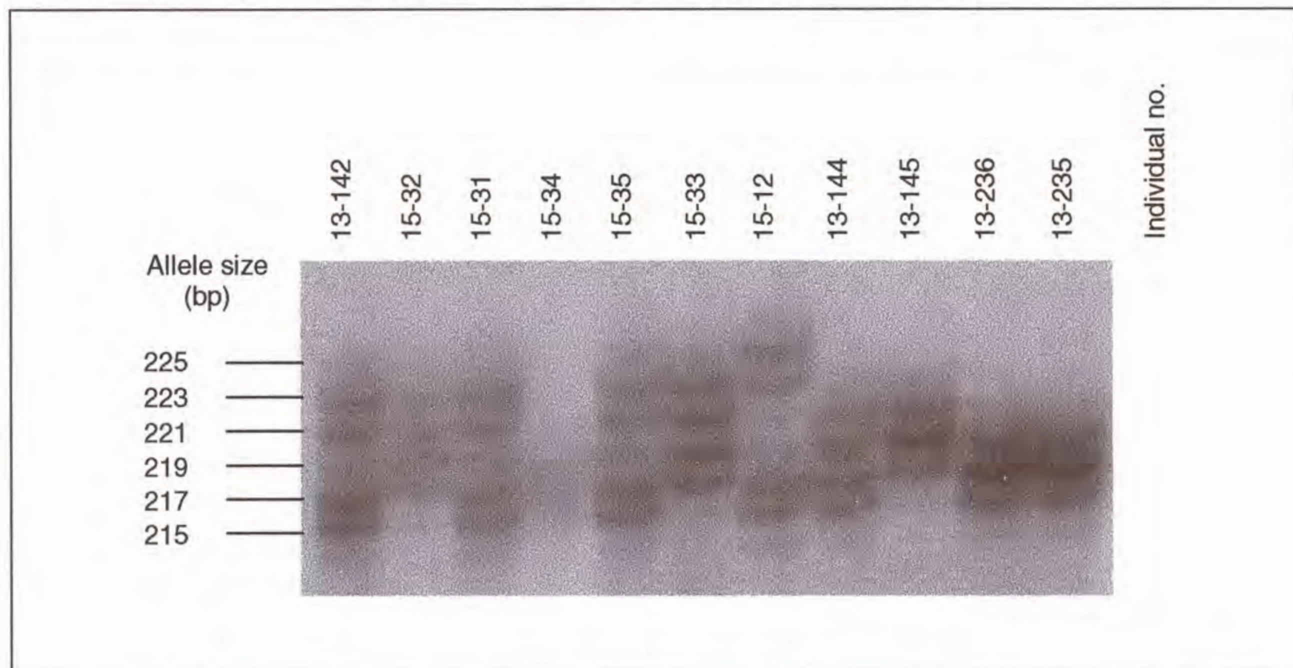
4.1.3 Short tandem repeat marker AFMa224xh1 at locus D4S2930

The primers of this marker were annealed at 55°C in a multiplex PCR reaction with the primers of markers AFMa190zf5 and ATA22F02. All nine of the reported alleles of marker AFMa224xh1, ranging from 217 to 233 bp, were observed in this study. Five of the alleles are presented in Figure 4.3, which shows a representative autoradiograph of the dinucleotide STRP at this locus.

Internal labelling of the amplification products was achieved by substituting one of the dNTPs in the PCR reaction mixture with an isotope labelled nucleotide. Both the strands are therefore visible on the gel, with the CA strand migrating faster than the GT strand, at a rate of one repeat (2 bp) difference between the two strands (Weber and May 1989). The scoring of the alleles could be obscured in this instance, as each allele is represented by two fragments. However, it did not obscure the results obtained with this marker.

The genotype (217-219) scored for individual 15-34 was confirmed on a longer exposure. The 219 bp allele for individual 13-145, and the 217 bp allele for individuals 13-236 and 13-235 appear darker than the other alleles. This is due to the strand slippage product of the GT strand being superimposed on the true allele generated from the CA strand. The strand slippage products did, however, not obscure the scoring of the alleles.

Figure 4.3: Representative autoradiograph of marker AFMa224xh1

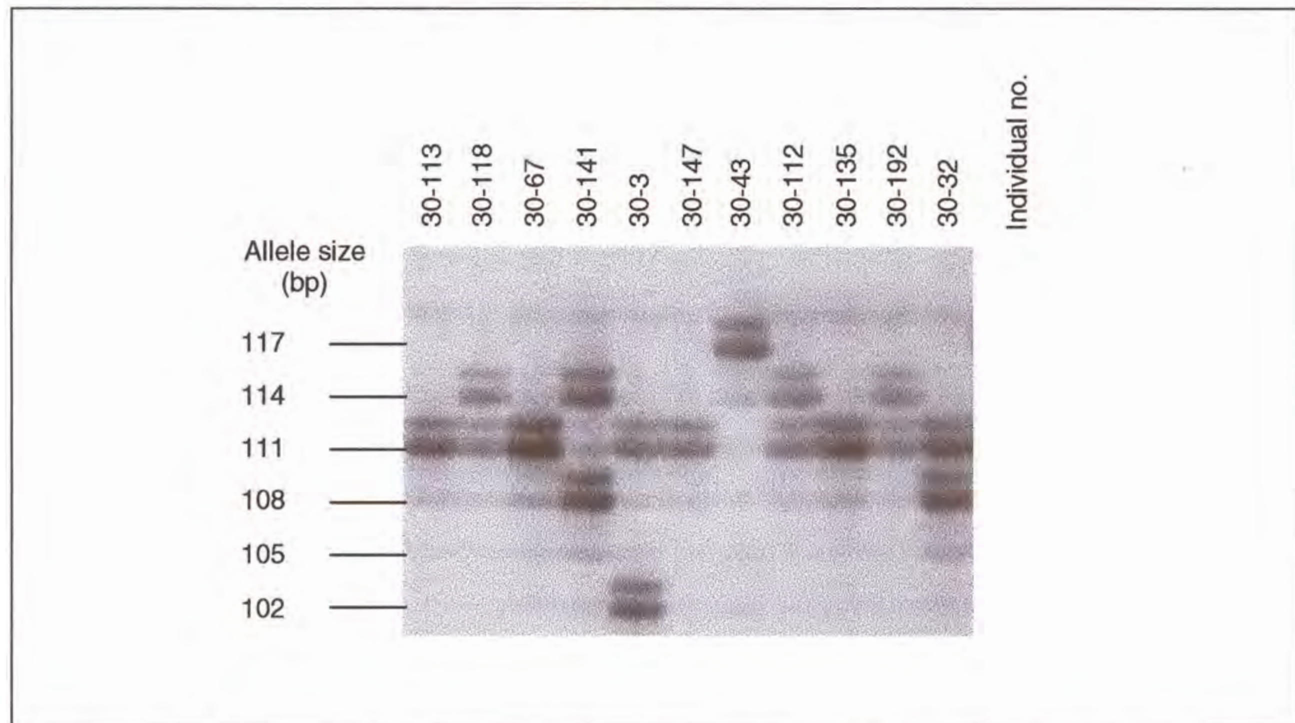


Autoradiograph of the amplification products generated with marker AFMa224xh1 described by Dib *et al.* (1996). Internal labelling was performed with α^{32}P-dCTP and the products electrophoresed on a 6% polyacrylamide gel in 1X TBE buffer for 2.5 hours. Individuals genotyped are indicated at the top and allele sizes are indicated to the left of the autoradiograph.

4.1.4 Short tandem repeat marker ATA22F02 at locus D4S2390

In this study all six of the previously reported alleles were observed and five of the alleles are presented in Figure 4.4. The amplification products were internally labelled in the multiplex reaction with markers AFMa224xh1 and AFMa190zf5 as described in paragraph 4.1.3. The sense strand, containing the $(ATA)_n$ repeat, migrated one base pair faster than the anti-sense strand, with the $(TAT)_n$ repeat. Two fragments per allele were therefore generated with this marker. It is also evident from Figure 4.4 that several constant secondary fragments were generated resulting in an allele motif of at least three fragments per allele. The sense strand was used for scoring of the alleles for this marker.

Figure 4.4: Representative autoradiograph of marker ATA22F02



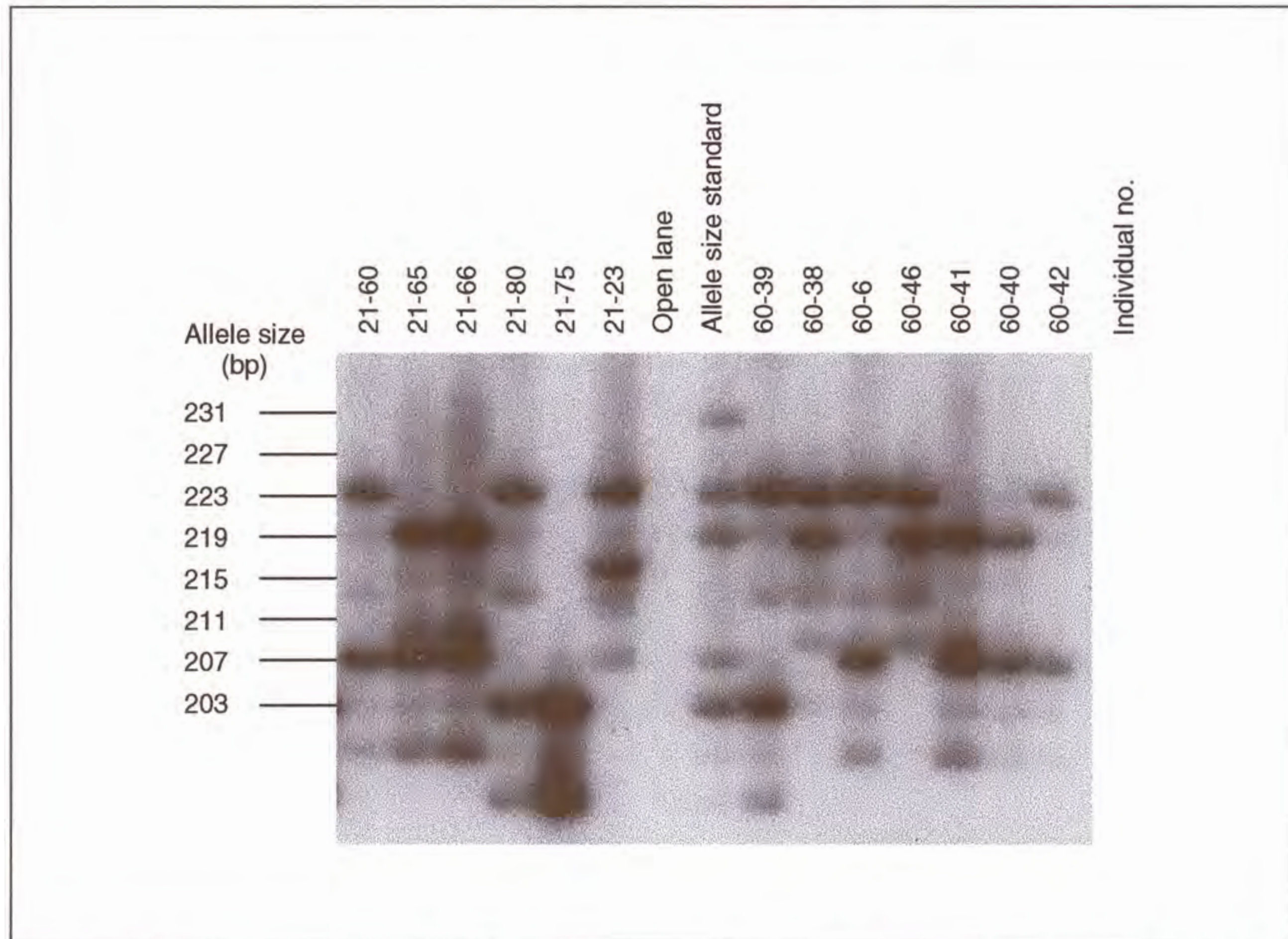
Autoradiograph of the amplification products generated with marker ATA22F02 described by Murray *et al.* (1995b). Internal labelling was performed with α^{32}P-dCTP and the products electrophoresed on a 6% polyacrylamide gel in 1X TBE buffer for 2.5 hours. Individuals genotyped are indicated at the top and allele sizes are indicated to the left of the autoradiograph.

4.1.5 Short tandem repeat marker UT5785 at locus D4S2299

All of the nine reported alleles, ranging from 203 to 235 bp, were observed in the five families included in this study. Figure 4.5 displays a representative autoradiograph of the amplification products generated with this marker. Six of the nine alleles are presented in Figure 4.5. The 231 bp allele was not observed in any of the individuals in Figure 4.5 but is presented in the allele size standard which consists of a mixture of PCR products of ca. six unrelated individuals.

The primer set of this marker was annealed at 56°C in the presence of 1 mM $MgCl_2$ with an end labelled reverse primer. Two secondary amplification products were generated for each allele, which migrated two base pairs faster than the true allele. This marker therefore generates three fragments per allele and the allele motif thus consisted of three distinct fragments, differing in size by 2 bp. Although not ideal, it did not obscure the scoring of the alleles.

Figure 4.5: Representative autoradiograph of marker UT5785



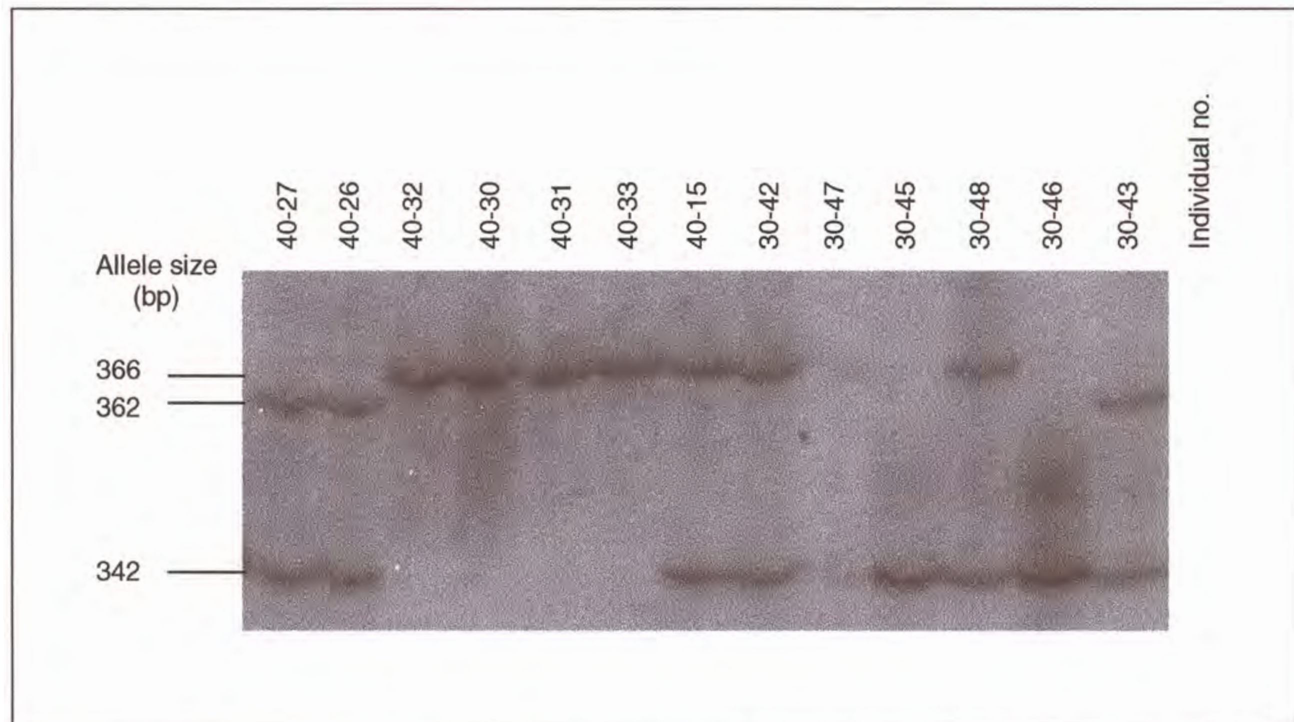
Autoradiograph of the amplification products generated with marker UT5785 described by Gerken *et al.* (1993b). The reverse primer was end labelled with γ^{32}P-dATP and the products electrophoresed on a 6% polyacrylamide gel in 1X TBE buffer for 3 hours. Individuals genotyped are indicated at the top and allele sizes are indicated to the left of the autoradiograph.

4.1.6 Short tandem repeat marker UT2219 at locus D4S2283

The reverse primer was selected for end labelling and the primers were annealed at 58°C. Four alleles have been reported for this marker but the allele sizes have not yet been reported (as mentioned in paragraph 3.3.1.6). Five alleles (342 bp and 358 to 370 bp) were generated with this marker in the five families investigated. Three of the alleles are presented in Figure 4.6, which displays a representative autoradiograph of this marker. The reported heterozygosity value for this marker is 59%. As presented in Figure 4.6, low levels of heterozygosity were also observed in the families investigated. The alleles generated with this marker were discernible with ease. The genotype for individual 30-47 was not scored from this autoradiograph, but from another autoradiograph that was exposed for a longer period of time.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Figure 4.6: Representative autoradiograph of marker UT2219



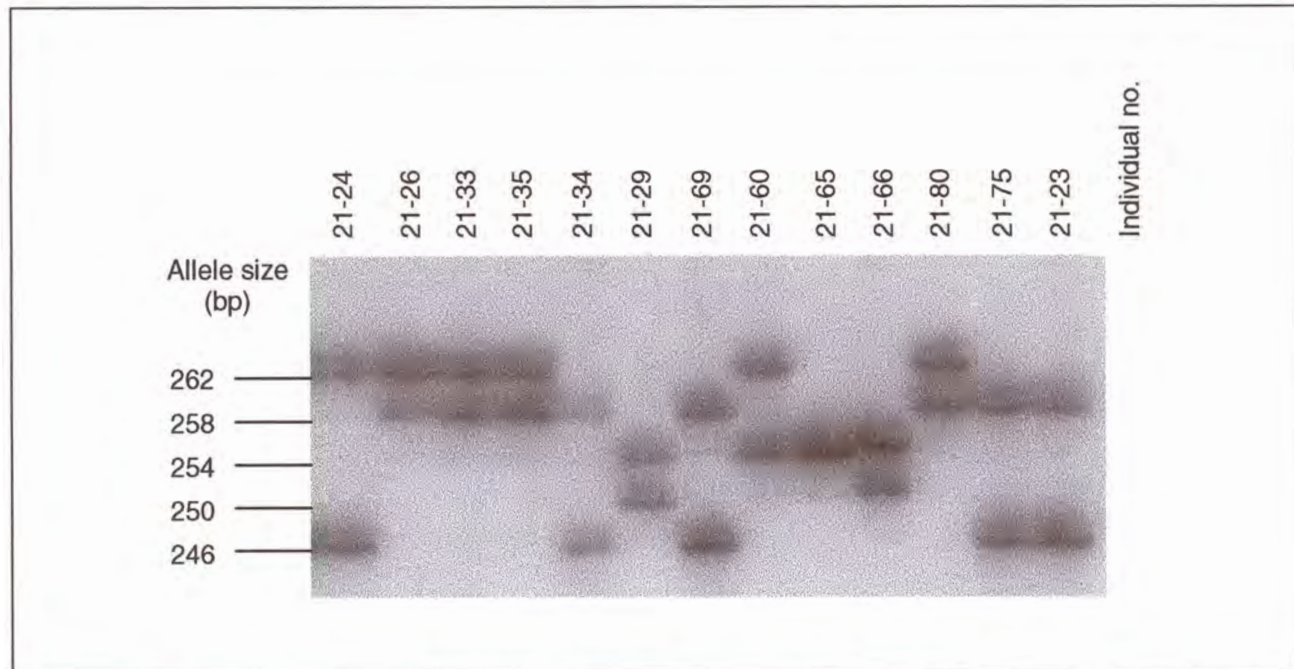
Autoradiograph of the amplification products generated with marker UT2219 described by Gerken *et al.* (1993c). The reverse primer was end labelled with γ^{32}P-dATP and the products electrophoresed on a 6% polyacrylamide gel in 1X TBE buffer for 5 hours. Individuals genotyped are indicated at the top and allele sizes are indicated to the left of the autoradiograph. The genotype for individual 30-47 was not scored from this autoradiograph.

4.1.7 Short tandem repeat marker UT7694 at locus D4S2688

Figure 4.7 displays a representative autoradiograph of the amplification products generated with this marker. All six alleles ranging from 246 to 266 bp were observed in this study of which five are presented in Figure 4.7. The primers of this marker were annealed at 55°C with the forward primer selected for end labelling.

A heterozygosity value of 52% was reported for marker UT7694. In this study, however, a relatively high level of heterozygosity was observed as displayed in Figure 4.7. No strand slippage products were generated with this marker and alleles could therefore be scored with ease.

Figure 4.7: Representative autoradiograph of marker UT7694



Autoradiograph of the amplification products generated with marker UT7694 described by Gerken *et al.* (1994). The forward primer was end labelled with γ^{32}P-dATP and the products electrophoresed on a 6% polyacrylamide gel in 1X TBE buffer for 3.5 hours. Individuals genotyped are indicated at the top and allele sizes are indicated to the left of the autoradiograph.

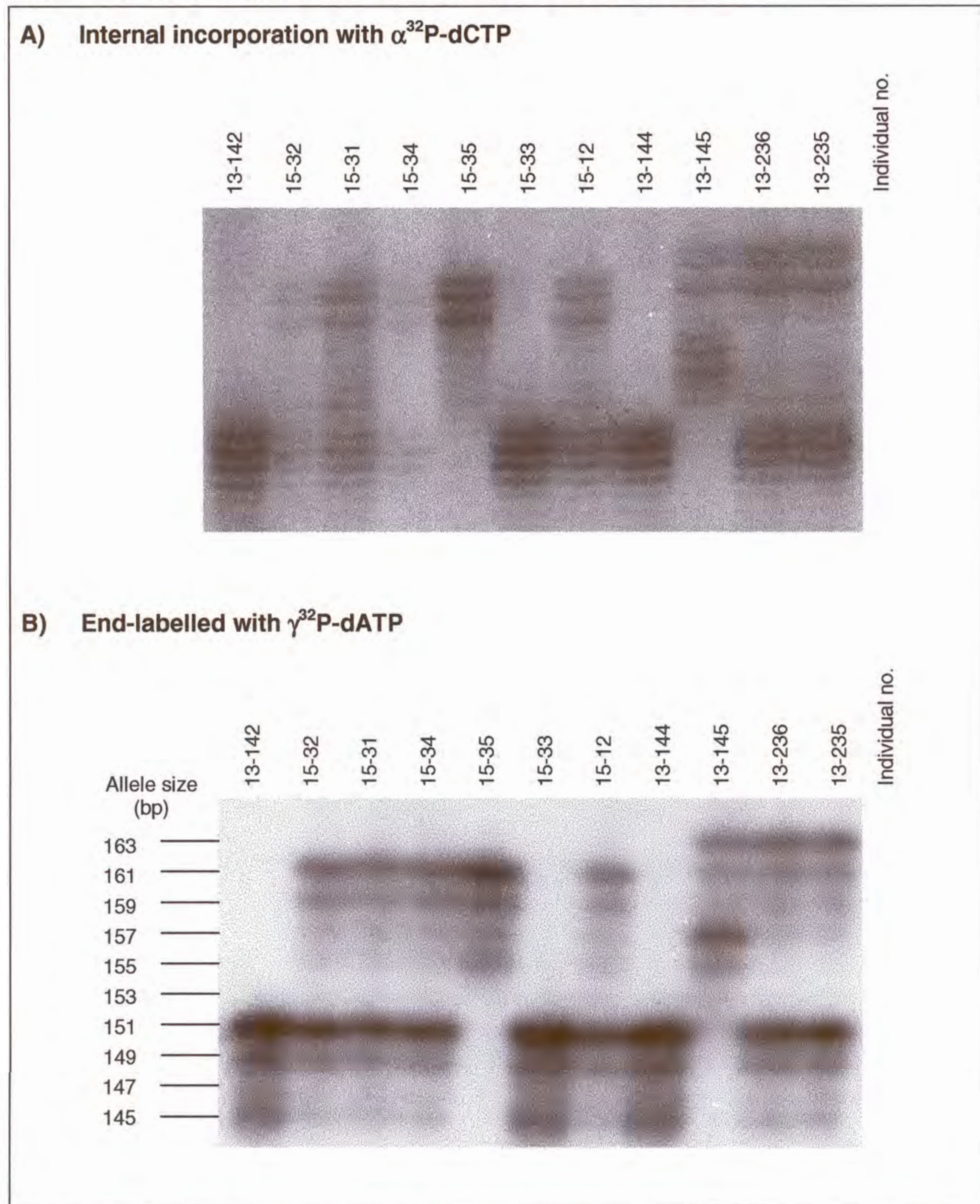
4.1.8 Short tandem repeat marker AFMa190zf5 at locus D4S2921

A representative autoradiograph of the alleles generated with this marker is displayed in Figure 4.8. As mentioned in paragraph 4.1.3 internal incorporation was at first utilised for the labelling of the amplification products and marker AFMa190zf5 was amplified in a multiplex PCR reaction with markers ATA22F02 and AFMa224xh1. Multiple strand slippage products were, however, observed which complicated the scoring of alleles (Figure 4.8A). For this reason end labelling of the primers was subsequently utilised for this marker.

The forward primer was end labelled and the marker was amplified in a standard PCR reaction. Although multiple strand slippage products were still visible after end labelling, these secondary products did not obscure the scoring of alleles. All eight of the reported alleles were observed, in the 100 individuals genotyped, of which four are presented in Figure 4.8B. Three additional alleles (147, 149 and 165 bp), not previously reported, were also observed in the families investigated but are not presented in Figure 4.8B. New markers are generally typed in a CEPH panel of individuals. It is therefore possible that alleles not previously reported in these panels, could be observed when the marker is

used in other populations, or in larger group of individuals. The possibility of observing new alleles in our unique population is therefore high.

Figure 4.8: Representative autoradiograph of marker AFMa190zf5

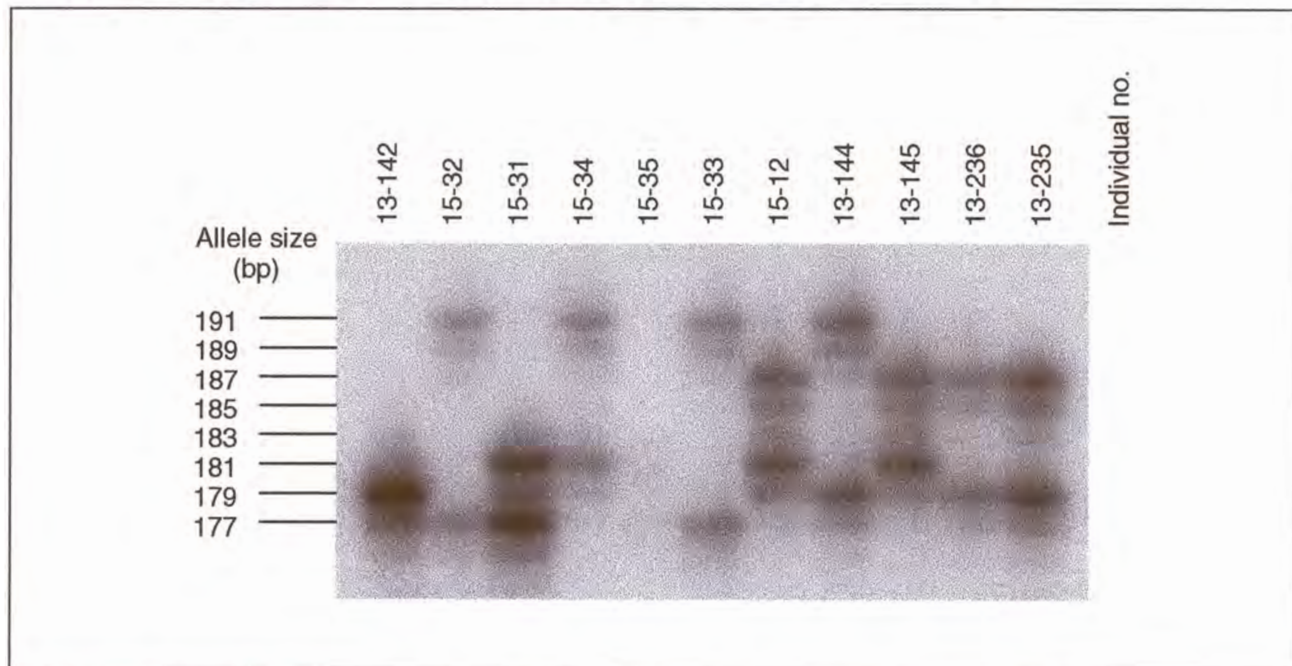


Autoradiographs of the amplification products generated with marker AFMa190zf5 described by Dib *et al.* (1996). A) = Internal labelling with α^{32}P-dCTP and the products electrophoresed on a 6% polyacrylamide gel in 1X TBE buffer for 2.5 hours. B) = The forward primer was end labelled with γ^{32}P-dATP and the products electrophoresed on a 6% polyacrylamide gel in 1X TBE buffer for 2.5 hours. Individuals genotyped are indicated at the top and allele sizes are indicated to the left of the autoradiograph.

4.1.9 Short tandem repeat marker AFM238ve3 at locus D4S426

The primers of this STRP marker were annealed at 55°C, with the forward primer selected for end labelling. Alleles generated with this marker are presented in Figure 4.9. This dinucleotide marker $(CA)_n$ generated six alleles ranging from 177-191 bp (excluding 183 bp and 185 bp). All six of the reported alleles were observed in the five families investigated. Five of the six alleles are presented in Figure 4.9, which display a representative autoradiograph for marker AFM238ve3. It is evident from the autoradiograph presented here that a secondary amplification product, 2 bp smaller than the true allele, was generated for each allele. This did, however, not obscure the scoring of the alleles. The genotype for individual 15-35 was scored from an autoradiograph that was exposed for a longer period of time.

Figure 4.9: Representative autoradiograph of marker AFM238ve3



Autoradiograph of the amplification products generated with marker AFM238ve3 described by Weissenbach *et al.* (1992). The forward primer was end labelled with γ^{32}P-dATP and the products electrophoresed on a 6% polyacrylamide gel in 1X TBE buffer for 2.5 hours. Individuals genotyped are indicated at the top and allele sizes are indicated to the left of the autoradiograph. The genotype for individual 15-35 was not scored from this autoradiograph.

Genotypes generated with the nine STRP markers could be scored with ease. Even though all the markers did not have high heterozygosity values they were all sufficiently informative for the construction of haplotypes in the group of families selected for this study.

4.2 SOUTHERN BLOT ANALYSIS

Southern blotting was performed to determine whether DNA rearrangements at the D4Z4 locus were present in the South African FSHD population. A Southern blot protocol was provided by Dr. Silvère van der Maarel (Department of Human Genetics, Leiden University Medical Centre, the Netherlands) [Lemmers *et al.,* 1998 and Church and Gilbert, 1984]. Previously, molecular studies had been performed for the proband of family F30 and two of his children at the Leiden University Medical Centre. This proband was utilised as a positive control for the optimisation and standardisation of Southern blot analysis.

The restriction endonucleases (*Eco* RI and *Bln* I) utilised for the whole genome digest did not have compatible buffers (Table 4.2) and two subsequent reactions had to be performed for each enzyme in its corresponding buffer. A protocol for a restriction enzyme buffer that was compatible with both enzymes was provided by Dr. van der Maarel for the double restriction endonuclease reaction and consisted of 5 mM Tris-HCl (pH 7.5), 1 mM $MgCl_2$, 10 mM NaCl supplemented with 3.3 mM spermidine. Ca. 75-85% complete digestion was achieved with this buffer. The preparation of the buffer, however, presented technical difficulties. A large volume of buffer had to be prepared to ensure measurable amounts of reagents, and resulted in the storage of large amounts of buffer, which would expire before it could be utilised. The preparation of this buffer was therefore not time or cost effective.

Table 4.2: Relative enzyme activity in Amersham buffers

Amersham buffer	L	M	H	K	T (+ BSA)
Eco RI	20[a]	100	100	120[a]	80[a]
Bln I	<20	20	40	100	20

L = low salt buffer, M = medium salt buffer, H = high salt buffer, K = potassium buffer, T = Tris acetate buffer, BSA = bovine serum albumin, a = weak star activity is detected. Adapted from restriction endonuclease product information sheet, Amersham (1996).

Subsequently the one-phor-all buffer supplied with restriction endonucleases from Amersham was utilised. The compatibility and an optimal 2X buffer concentration, to prevent star activity, with the use of *Eco* RI digestion reactions were previously determined by Amersham. The compatibility and concentration of the buffer for use with *Bln* I had, however, not previously been determined. Utilising the one-phor-all buffer at a 2X final

concentration with *Eco* RI and *Bln* I in a single restriction endonuclease reaction resulted in 100% complete digestion.

Very faint or even no fragments were visible on the autoradiograph after utilising the original Southern blot protocol (Lemmers *et al.,* 1998). The fragments of the molecular weight markers utilised to determine the sizes of the fragments were, however, visible on the autoradiograph. It was therefore evident that some conditions had to be adapted to enable the visualisation of the deletion fragments. Conditions that could be optimised were: 1) to ensure sufficient gDNA transfer, 2) fixation of the DNA onto the membrane, 3) temperature at which hybridisations were performed and 4) the stringency for washing of the membrane after hybridisation. Each of these aspects are discussed below.

Fragments larger than 20 kb cannot be transferred efficiently without nicking prior to transfer onto the membrane. In the original protocol the DNA fragments were nicked by UV irradiation at a wavelength of 312 nm as recommended. The transilluminator should be regularly calibrated as the dose emitted can vary with the age of the UV bulbs in the instrument. The management and control of the dose emitted by the transilluminator could therefore result in insufficient cleavage and incomplete gDNA transfer. The transilluminator that was available during this study was not calibrated and to exclude any variable factors, the DNA fragments were instead depurinated with 0.25 M HCl.

The original protocol utilised UV crosslinking for 4 minutes for the fixing of the DNA onto the nylon membrane. As stated above, the UV dose emitted is important to ensure optimal results. For the fixing of the DNA onto the membrane, baking in an oven at 80°C for 2 hours was performed instead.

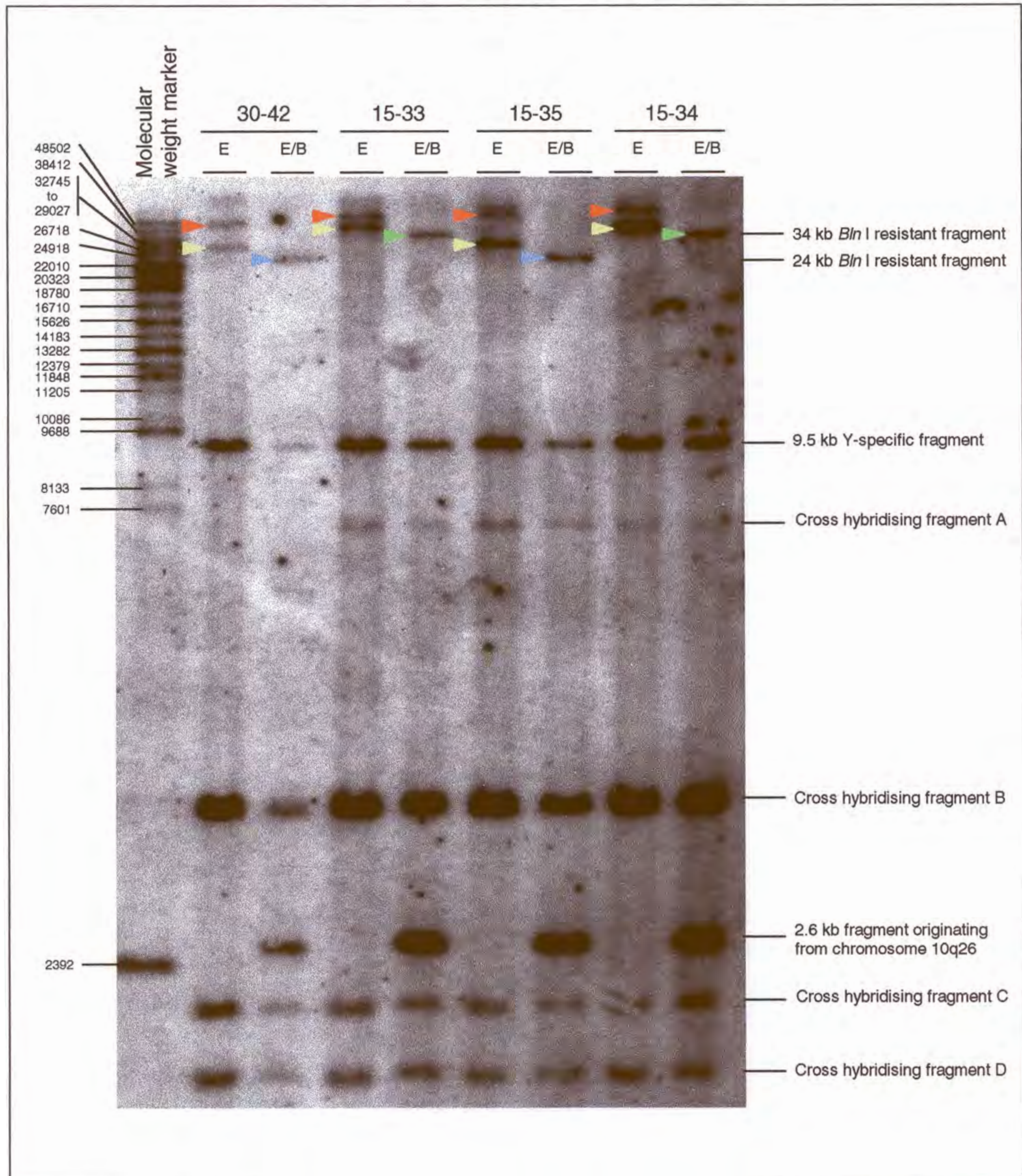
Hybridisations were performed at 65°C for 16 hours after which the membrane was washed with 2X SSC and 0.1% SDS (pre heated to 65°C) three times for 5 minutes. As previously stated, these conditions resulted in faint or no visible fragments after autoradiography. The hybridisation temperature and/or washing stringency were therefore too high and had to be lowered. The hybridisations were performed at 55°C for 16 hours and the membrane washed with 2X SSC and 1% SDS washing solution at 55°C. After autoradiography more fragments were visible, but much more background was also observed across the entire autoradiograph. This observation led to the conclusion that the stringency of the hybridisation conditions was too low. The next membrane was

hybridised at 65°C for 16 hours and washed with 2X SSC and 1% SDS washing solution at 65°C. Distinct deletion fragments as well as a Y-chromosome specific 9.5 kb fragment, a 2.6 kb *Eco* RI/*Bln* I fragment and other non-specific fragments were observed upon autoradiography. The Southern blot analysis was therefore optimised by utilising a ten times less stringent washing solution, thus at a stringency of 2X SSC and 1% SDS.

The same molecular weight markers that were loaded on the agarose gels were also random primed and hybridised with the membrane for precise sizing of the alleles. Twenty five ng of each molecular weight marker was labelled and different amounts (250 ng to 1 ng) were loaded onto the gels to determine the optimal concentration. The optimal concentration could, however, not be standardised. The signal intensity was either too intense, or no fragments were visible. The molecular weight markers were therefore end labelled instead. Fifty ng of each molecular weight marker was loaded per lane on the agarose gels and then hybridised with 500 ng of end-labelled molecular weight marker. Equal signal intensity to that of the fragments on the gels was obtained, and could be used for sizing of the fragments.

A representative autoradiograph of Southern blot analysis is displayed in Figure 4.10. Two different *Bln* I resistant deletion fragments of 24 kb and 34 kb, a 9.5 kb Y-chromosome specific fragment, a 2.6 kb chromosome 10q26 specific fragment, as well as several cross hybridising fragments are visible on this autoradiograph. The cross hybridising fragments can originate from homologous regions on the short arms of all the acrocentric chromosomes, as well as from chromosomes 1, 2, 3, 4, 10, 18, and Y, as mentioned in paragraph 2.1.2.2. The intensity of cross hybridising fragment A, just smaller than 7601 bp, is the lowest of the four cross hybridising fragments. It is evident from Figure 4.10 that all the fragments observed for individual 30-42 are of much lower intensity when compared to fragments observed in any of the other individuals. This might explain the absence of cross hybridising fragment A in individual 30-42. The dark specks visible in the lanes designated for individual 15-34 between 9.5 kb and 34 kb are due to artefact. It is evident that only samples from male individuals were included in Figure 4.10, since all of the samples displayed a 9.5 kb Y-specific fragment.

Figure 4.10: Representative autoradiograph of the optimised Southern blot analysis



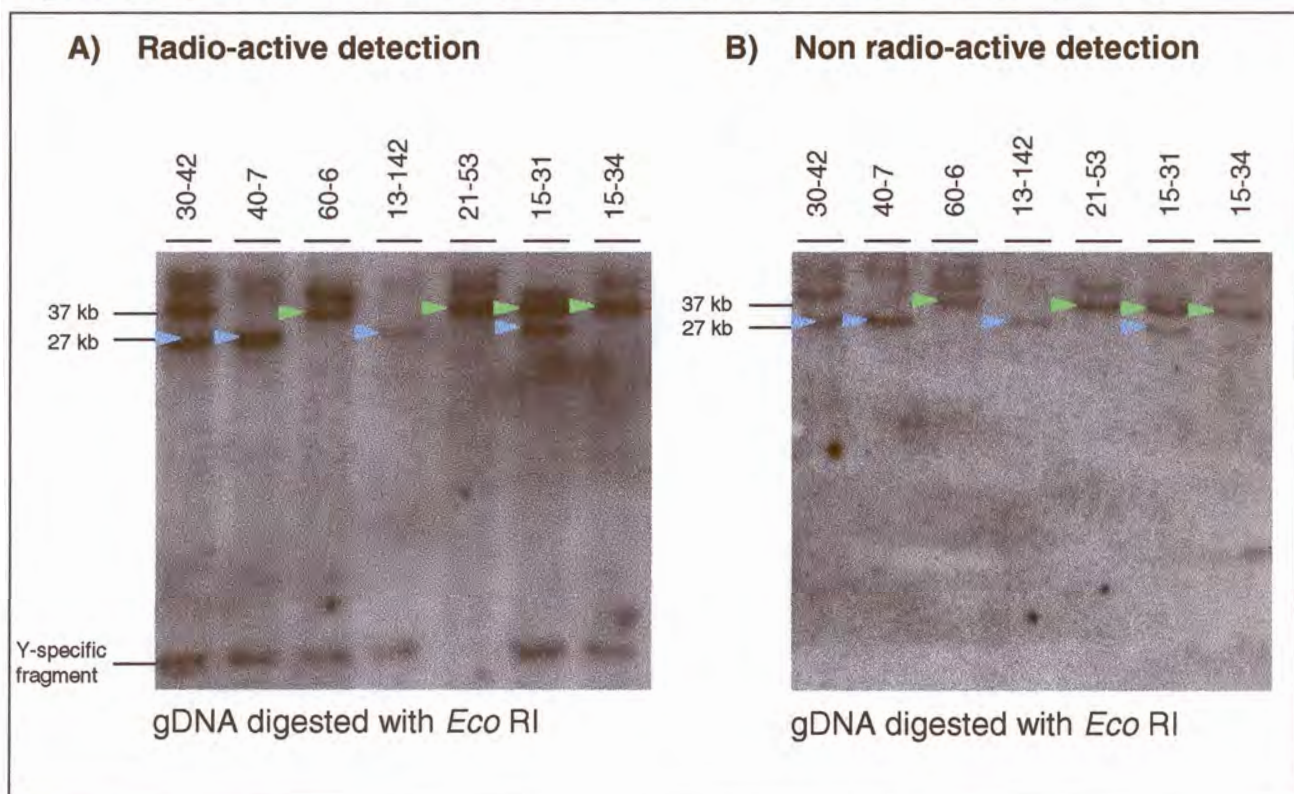
E = *Eco* RI , E/B = *Eco* RI/*Bln* I double digestion. ▶ = chromosome 10q26 fragment, ▶ = chromosome 4q35 fragment, ▶ = 24 kb *Bln* I resistant deletion fragment verifying chromosome 4q35 origin, ▶ = 34 kb *Bln* I resistant deletion fragment verifying chromosome 4q35 origin.

4.3 RADIO-ACTIVE VERSUS NON RADIO-ACTIVE DETECTION

Standard diagnosis for FSHD worldwide is Southern blot analysis utilising a radio-active labelled p13E-11 probe. The Detector™ Random Primer DNA biotinylation kit as well as

the DNADetector™ Genomic Southern Blotting kit (KPL) were utilised to standardise a non radio-active detection method for FSHD. These kits were subsequently tested on the standard optimised Southern blot protocol utilised for the radio-active detection and the results obtained are presented in Figure 4.11. In this figure two identical Southern blots are presented with different detection strategies. The non radio-active method detected the same *Eco* RI deletion fragments as the radio-active method. However, the Y-specific fragment was not detected with the non radio-active protocol. An explanation for this is that the non radio-active protocol utilised a much more stringent wash condition. The radio-active membrane was washed 3 times after hybridisation for 5 minutes in a 2X SSC/1% SDS wash solution, while the non radio-active membrane was washed twice for 15 minutes each in 2X SSPE/0.1% SDS followed by two washes for 15 minutes in 0.2X SSPE/0.1% SDS and subsequently in 2X SSPE washing solution for 5 minutes. The first wash of the non radio-active protocol (2X SSPE/0.1% SDS) is therefore a ten times more stringent solution than that utilised for the radio-active protocol. It is thus evident that with further optimisation the non radio-active detection protocol can be utilised in the detection of deletion fragments in FSHD. The radio-active protocol was utilised for the molecular diagnosis performed in this study.

Figure 4.11: Comparison between radio-active and non radio-active detection



= 37 kb *Eco* RI deletion fragments and;

= 27 kb *Eco* RI deletion fragments.

4.4 MOLECULAR ANALYSIS OF SPECIFIC FSHD FAMILIES

Markers described in paragraph 4.1 were utilised to generate genotypes for the individuals selected for this study. The genotypes generated with the nine STRP markers and the Southern blots were combined in order to elucidate the molecular basis of FSHD in the South African population. The haplotypes are indicated by coloured blocks, as indicated in Figures 4.12, 4.14, 4.15, 4.16 and 4.17, which encompass the specific allele sizes of the different markers included in the haplotype. The combined genotype analyses of the selected families are presented in the following paragraphs.

4.4.1 FSHD family F10

Family F10 was subdivided into five sub-families (F11, F12, F13, F14 and F15) for easy reference as described in paragraph 3.1.1. Twenty five individuals from family F10 were genotyped with the nine STRP markers and Southern blot marker p13E-11 as discussed in paragraphs 4.1 and 4.2 respectively. Five of the individuals are related to their respective families by marriage. Of the remaining twenty individuals, ten were clinically phenotyped to be FSHD positive, six were phenotyped to be FSHD negative and four were equivocal on a clinical level. Haplotypes for all of these individuals are presented in Figure 4.12.

The exact order of the nine STRP markers utilised in this study has not yet been mapped (Table 3.2). It is evident from Table 3.2 that marker UT1366 (locus D4S1523) is the most distal, with markers AFM238ve3 (locus D4S426) and AFMa190zf5 (locus D4S2921) being the most proximal. The exact order of the remaining six markers has not been determined. Several recombination events within this family, however, facilitated in determining the order of these markers. As indicated in Figure 4.12, recombination events in the ancestors of individuals 14-64, 14-102, 13-142, 13-144, indicated that marker UT7694 (locus D4S2688) is distal to the markers at loci D4S426 and D4S2921 and recombination events in the ancestors of individuals 12-3 and 12-6 indicated that the marker at locus D4S2390 (ATA22F02) is distal to marker UT7694 (locus D4S2688). The recombination event observed in the ancestors of individual 11-28 led to the conclusion that marker UT5785 (locus D4S2299) is distal to the marker at locus D4S2390 and that the marker at locus D4S2930 is distal to the marker at locus D4S2299. Finally the recombination event in individual 15-34 facilitated the mapping of marker UT2219 (locus D4S2283) proximal to marker UT1366 (locus D4S1523) and distal to marker AFMa224xh1 (locus D4S2930), mapping the last marker GATA5B02 (locus D4S1652) distal to marker

AFMa224xh1 and proximal to marker UT2219. The only two markers that could not be mapped through haplotype analysis in family F10 were markers AFM238ve3 (locus D4S426) and AFMa190zf5 (locus D4S2921). These two markers were, however, mapped by a recombination event in another South African family (F40). In this family, individual 40-25, as illustrated in Figure 4.16, displayed recombination allowing marker AFMa190zf5 (locus D4S2921) to be mapped proximal to marker AFM238ve3. The order in which the nine STRP markers utilised in this study, presented in the haplotypes, is thus as follows: TEL - UT1366 (D4S1523), UT2219 (D4S2283), D4S1652 (GATA5B02), AFMa224xh1 (D4S2930), UT5785 (D4S2299), ATA22F02 (D4S2390), UT7694 (D4S2688), AFM238ve3 (D4S426) and AFMa190zf5 (D4S2921) - CEN, as indicated in Figure 4.12.

In this family the 297-342-138-217-223-111-246-181-161 haplotype (indicated in the blue block in Figure 4.12), or a portion of the haplotype, was present in sixteen individuals from the five sub-families (F11, F12, F13, F14 and F15). This haplotype survived transmission through 32 meiotic events spanning eight generations. Upon Southern blot analysis, a *Bln* I resistant deletion fragment of 24 kb (3 kb shorter than the 27 kb *Eco* RI deletion fragment) was observed to segregate in this family. All of the individuals who inherited the haplotype marked in blue also inherited this 24 kb deletion fragment.

Nine of the ten clinically affected individuals (15-12, 15-31, 11-28, 11-30, 12-6, 13-142, 13-144, 14-64 and 14-102) inherited the FSHD associated haplotype and the 24 kb deletion fragment. Individual 14-113 did, however, not inherit the haplotype or the deletion fragment that segregated in this family. This individual was clinically diagnosed in his late fifties, an age at which the phenotype should have been fully penetrant. He was, however, not diagnosed by Dr. Schutte, a specialist in FSHD, and might therefore have been mis-diagnosed, as the clinical signs are often subtle and difficult to identify. Re-evaluation of individual 14-113, is therefore pertinent.

The four individuals diagnosed to be equivocal on a clinical level (12-3, 12-10, 12-40, and 13-235) all inherited the "blue" haplotype as well as the 24 kb deletion fragment. The FSHD status of these individuals was therefore confirmed upon molecular analysis. Two of the clinically unaffected individuals (15-35 and 13-236) also inherited the blue haplotype and the deletion fragment. These individuals are still younger than twenty-one, the age at which 95% of affected individuals generally start to show clinical symptoms. Upon clinical examination they could therefore still have been asymptomatic. Individuals 15-35 and 13-236 should also be clinically re-evaluated. It is possible that they could have been

asymptomatic at the time of analysis, but developed the symptoms associated with FSHD since that time.

Two *Bln* I resistant deletion fragments (24 kb and 34 kb) were observed for individual 15-31 upon Southern blot analysis as presented in Figure 4.13 (Van der Merwe *et al.*, 2001a). Both the deletion fragments are in the disease causing range, and being *Bln* I resistant, behave as chromosome 4-type fragments. This would imply that this individual is a compound heterozygote. It is, however, evident from the molecular data generated for family F10, that the 24 kb deletion fragment segregates with the FSHD phenotype in this family, and is therefore a true chromosome 4 type fragment. The chromosomal origin of the 34 kb deletion fragment was, however, not conclusive, since it is known that translocations occur between chromosome 4 and chromosome 10 type fragments. As depicted in Figure 4.13, individual 15-35 inherited the 24 kb deletion fragment as well as the "blue" haplotype from his father (15-31). His brothers, individuals 15-33 and 15-34, both inherited the 34 kb deletion fragment together with the distal portion of the "green" haplotype. A 9.5 kb Y-specific fragment is visible for all the male individuals, while this fragment is absent in the one female individual, 15-32, presented in Figure 4.13.

Verification of the chromosomal origin was therefore crucial for accurate molecular diagnosis of these two children, individuals 15-33 and 15-34. Haplotype analysis on chromosome 10q was performed in a different study[1] and confirmed that the 34 kb deletion fragment originated from chromosome 4 (Olckers *et al.*, 2001; Van der Merwe *et al.*, 2001b). Upon further analysis it was observed that the 34 kb deletion fragment as well as the haplotype indicated in green observed in this individual and his two children (15-33 and 15-34) also segregated in a second FSHD family, F20 (paragraph 4.3.2). The chromosomal origin of the 34 kb deletion fragment was therefore confirmed via chromosome 4 as well as chromosome 10 haplotype analyses. Individuals 15-33 and 15-34 were clinically phenotyped in their mid-teens to be unaffected but molecular analysis indicated that both individuals inherited the second disease causing deletion fragment (34 kb). These individuals might have been asymptomatic upon clinical examination, and it is thus critical that they be clinically re-evaluated at an appropriate age in the future.

---

[1] Chromosome 10q26 haplotype analysis performed by M. Alessandrini.

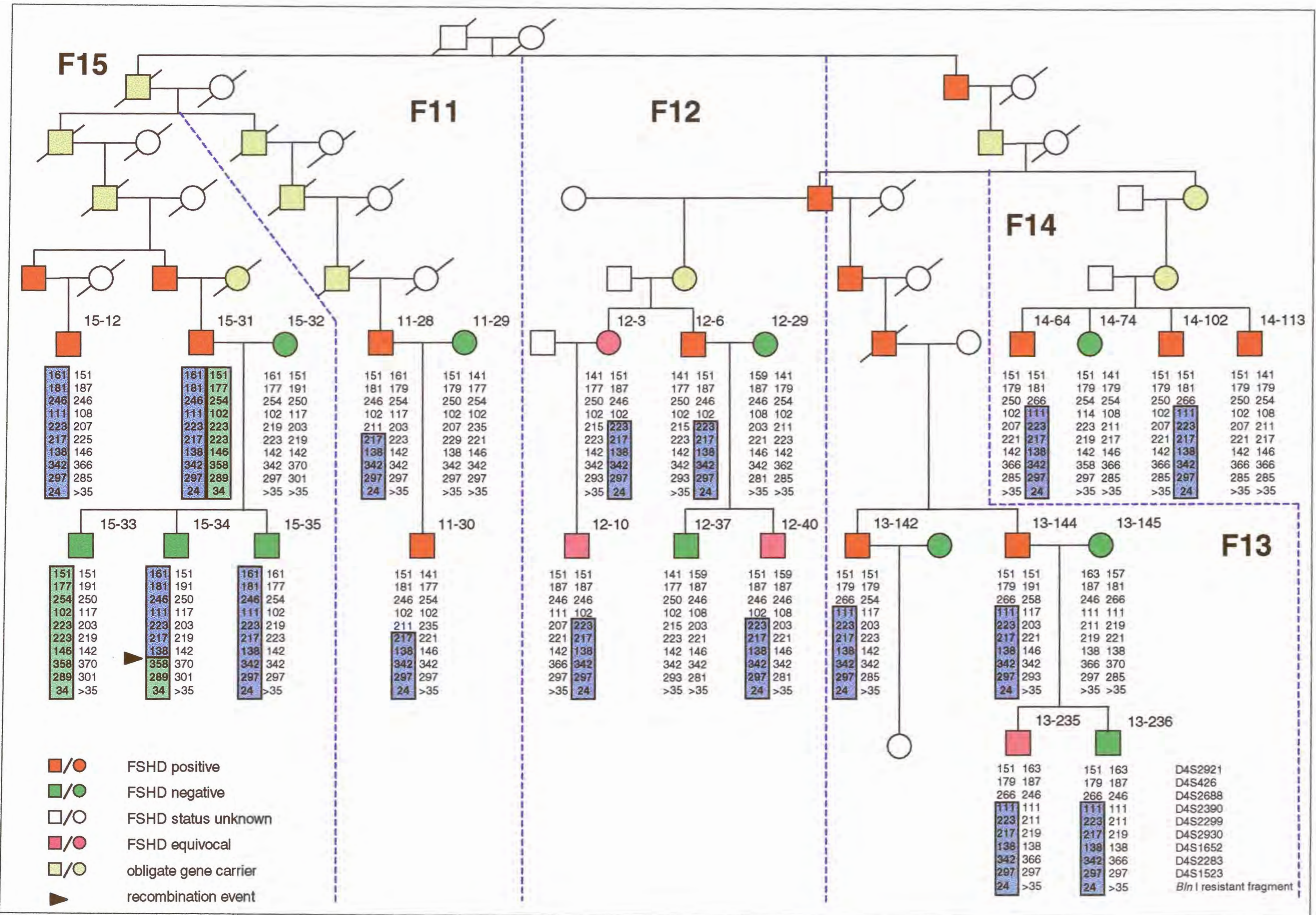
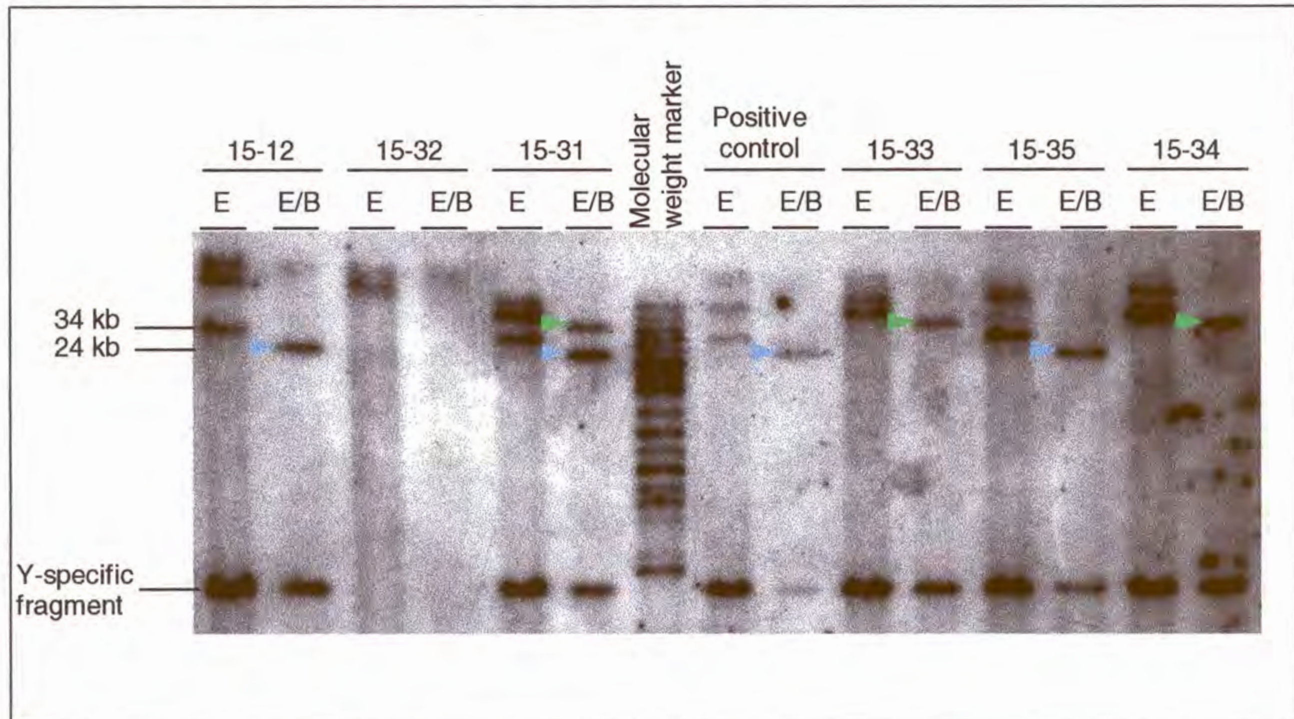


Figure 4.12: Haplotypes of selected individuals from family F10

Figure 4.13: Southern blot analysis for family F15



= 34 kb *Bln* I resistant deletion fragment;

= 24 kb *Bln* I resistant deletion fragment.

Several recombination events were observed within family F10 and facilitated in determining the order of seven of the nine STRP markers utilised in this study. FSHD could be confirmed via the presence of a *Bln* I resistant deletion fragment of 24 kb that segregated in family F10. The first South African compound heterozygote individual (15-31) for FSHD was observed in this family, as identified by the presence of two *Bln* I resistant deletion fragments. This might be useful in the elucidation of the phenotypic-genotypic correlation in FSHD due to the presence of two different sizes of *Bln* I resistant deletion fragments within one family. If the phenotype can be correlated with the genotype, the father, being a compound heterozygote, should be severely affected and the son with the 24 kb deletion fragment ought to be more severely affected than his brothers harbouring the 34 kb deletion fragments. The clinical diagnosis of these individuals is therefore essential to elucidate the correlation between the phenotype and the genotype.

4.4.2 FSHD family F20

Of the twenty five individuals genotyped from the three generations in this family, eleven were FSHD positive, ten were FSHD negative, three were phenotyped to be equivocal on a clinical level, and the FSHD status of one individual was unknown. A haplotype

indicated on a green background in Figure 4.14 was observed to segregate in family F20. As mentioned in paragraph 4.3.1 this haplotype was also observed in individuals 15-31, 15-33 and 15-34.

The proband of this family (21-55) inherited the TEL-289-358-146-223-223-102-CEN haplotype from her clinically affected mother (21-53). Four other FSHD positive individuals 21-60, 21-67, 21-80 and 21-33 also inherited this haplotype. The haplotype was also observed to segregate in two of the FSHD negative individuals (21-24 and 21-35), two of the individuals who were found to be equivocal on a clinical level (21-28 and 21-30) and in individual 21-26, whose FSHD status was unknown prior to this study. No haplotype was observed in individuals 21-65, 21-69, 21-71, 21-75 and 21-76 who were clinically phenotyped to be affected, or in individual 21-66 who was diagnosed as clinically equivocal. Haplotypes, or portions thereof, were inferred for those individuals (21-116, 21-117, 21-84, 21-52, 21-22, 21-54 and 21-68) of whom blood samples were not available. The haplotype observed to segregate in this family survived transmission through 12 meiotic events spanning four generations.

Southern blot analysis detected the presence of a 37 kb *Eco* RI deletion fragment in all of the individuals who inherited the FSHD associated haplotype. This 37 kb *Eco* RI fragment was reduced by 3 kb resulting in a 34 kb *Bln* I resistant deletion fragment. Genotype-phenotype discordance was, however, observed for individuals 21-65, 21-66, 21-69, 21-71, 21-75 and 21-76 who did not inherit this 34 kb deletion fragment. All of these individuals were younger than twenty upon clinical examination. As pointed out in family F10 the clinical diagnosis of individuals before the age of twenty is complex and may be misleading, and for this reason these individuals need to be re-examined on the clinical level.

As mentioned previously no 34 kb *Bln* I resistant deletion fragment was observed in individuals 21-69 and 21-75. However, Southern blot analysis revealed a 37 kb *Eco* RI deletion fragment in both individuals. The 37 kb *Eco* RI deletion fragment was, however, completely digested with *Bln* I, thus displaying a chromosome 10 restriction endonuclease pattern. The Southern blot analysis of individual 21-75 was repeated twice, but the 37 kb *Eco* RI fragment was *Bln* I sensitive in both analyses. It is therefore highly unlikely that

this could have been due to experimental errors. The possibility of a translocation event is also unlikely but needs to be excluded with either PFGE or the *Bgl* II/*Bln* I dosage test, as all four alleles as well as the presence of translocations will be visible via these methods of analysis.

The 146 bp allele for marker GATA5B02 at locus (D4S1652) was observed to segregate in the haplotype indicated in green. However, individual 21-28, who displays the haplotype, inherited a 150 bp allele. The son of individual 21-28 (21-30) inherited the haplotype with this 150 bp allele from his father. The genotypes of these individuals (21-28 and 21-30) were repeated multiple times, with identical results. The father of individual 21-28 could not be genotyped as he was deceased before a sample could be obtained. If the father had a 150 bp allele at locus D4S1652, individual 21-28 could have inherited this allele due to a recombination event between loci D4S2930 and D4S1652, and loci D4S1652 and D4S2283. This is, however, very unlikely due to the short genetic distances between these markers. The mutation rate of tetranucleotides was reported to be as high as 1.8×10^{-3} (Xu *et al.*, 2000). It is therefore most likely that the 146 bp allele mutated to a 150 bp allele in individual 21-28 as both his brothers (21-24 and 21-33) inherited a 146 bp allele from their father and the mother of individual 21-28 (21-23) does not harbour a 150 bp allele at this locus.

A different *Bln* I resistant deletion fragment and FSHD associated haplotype were observed to segregate in family F20 than in family F10. This indicated the presence of at least two *Bln* I resistant deletion fragments and FSHD associated haplotypes in the South African FSHD population. Although FSHD could be confirmed on a molecular level for family F20, genotype-phenotype discordance was, however, also observed. The individuals displaying discordance were all clinically diagnosed before the age of 20 and should therefore be re-examined as they could have developed symptoms since their last examination.

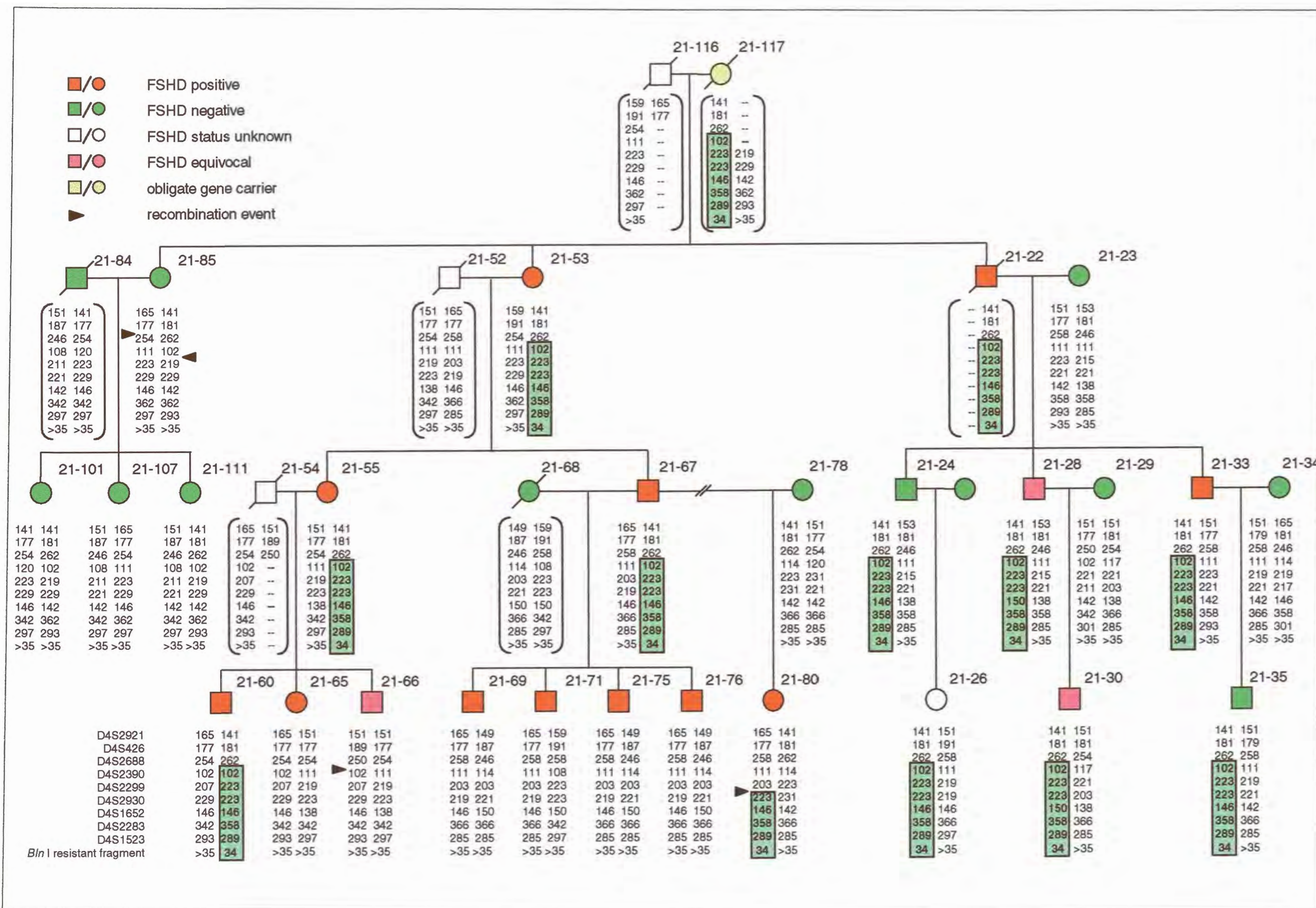


Figure 4.14: Haplotypes of selected individuals from family F20

4.4.3 FSHD family F30

Thirty three individuals were genotyped in family F30. Three of the individuals are only related to their respective families by marriage and of the remaining thirty individuals, twelve were clinically diagnosed to be FSHD positive, sixteen were diagnosed to be negative on a clinical level and two were equivocal upon clinical examination. Haplotypes for the thirty three individuals are presented in Figure 4.15.

The FSHD associated haplotype, indicated in blue in Figure 4.12, which segregated in family F10 was also observed to segregate in family F30. The entire haplotype, consisting of all nine STRP markers, that was observed in sub-family F15 also segregated in family F30. The haplotypes were deduced for those individuals (30-1, 30-2, 30-111 and 30-108), of whom blood samples were not available. Only a portion of the haplotypes of individuals 30-30 and 30-210 could be deduced as their children inherited the sample haplotype in this region. This haplotype survived transmission through 14 meiotic events spanning four generations. Nine of the twelve clinically affected individuals inherited this FSHD associated haplotype and the remaining three individuals, 30-48, 30-104 and 30-106 as well as two individuals (30-103 and 30-193) who were equivocal on the clinical level did not inherit this haplotype.

Four clinically unaffected individuals (30-3, 30-147, 30-135 and 30-46) also inherited the FSHD associated haplotype. Individual 30-3 was in his nineties upon clinical examination, individual 30-147 in his late seventies and individual 30-135 in her mid forties. These individuals were therefore clinically diagnosed at an age at which the FSHD phenotype was expected to have been fully penetrant, yet they displayed no clinical FSHD phenotype and were in fact diagnosed as asymptomatic. Individual 30-46 was clinically diagnosed to be unaffected in his mid twenties, the age at which the phenotype is reported to be 95% penetrant. This individual might, however, still have been asymptomatic by that age, and should therefore be clinically re-evaluated in the future, as symptoms might have developed since his last clinical evaluation. Molecular data therefore indicated that the four clinically unaffected individuals, 30-3, 30-147, 30-135 and 30-46, are FSHD positive on a genetic level. Genetic counselling is therefore essential for these clinically asymptomatic individuals who are treated as affected individuals, since their molecular diagnosis was confirmed to be positive.

An *Eco* RI deletion fragment of 27 kb was observed to segregate in this family. The chromosome 4 origin of this deletion fragment was verified with a *Bln* I digestion. The

Eco RI deletion fragment was reduced by 3 kb, resulting in the presence of a 24 kb *Bln* I resistant deletion fragment. All the individuals who inherited the FSHD associated haplotype, indicated in blue in Figure 4.15, inherited the 24 kb *Bln* I resistant deletion fragment. Individual 30-48 who was phenotyped to be FSHD positive, but who did not inherit the FSHD associated haplotype did, however, inherit the 24 kb *Bln* I resistant deletion fragment from his clinically affected father. This is the result of a recombination event between the D4S1523 and D4Z4 loci and therefore argues against the use of haplotype analysis alone for the diagnosis of FSHD and confirms the importance of Southern blot analysis.

Individuals 30-104 and 30-106 were clinically diagnosed with FSHD but did not inherit the haplotype or the 24 kb *Bln* I resistant deletion fragment. Both of them were clinically examined before the age of 20 and will have to be re-examined. Phenotype-genotype discordance was observed in families F10, F20 and F30. It was, however, evident that the discordant individuals were generally present in the youngest generation. The discordance might therefore be due to clinical bias as the clinical phenotype is extremely heterogeneous and complex, especially in young children. The clinical diagnosis by a clinician specialised in FSHD is therefore essential and it will be optimal if all the individuals could be diagnosed by one clinician.

Genomic DNA from individual 30-3 was amplified with the nine STRP markers described in paragraph 4.1. Individual 30-3 inherited the 297-342-138-217-223-111-246-181-161 haplotype, indicated in a blue block in Figure 4.15, from his mother. Three of the nine markers utilised in this study (D4S2930, D4S2390 and D4S2921), generated alleles that were incompatible with Mendelian segregation of alleles from his father (individual 30-1). These markers were scored again to confirm that the discrepancies observed were not due to scoring errors. It is also possible that incorrect labelling of the original tubes of whole blood might have resulted in the observed discrepancies. To eliminate this possibility, a fresh blood sample should be obtained from individual 30-3 to re-type these markers. If the alleles indicated in Figure 4.15 for individual 30-3 is, however, correct it would seem that non-paternity is indicated between individuals 30-1 and 30-3.

Via molecular analysis the presence of the same size *Bln* I resistant deletion fragment, of 24 kb, and FSHD associated haplotype was confirmed in families F10 and F30. The presence of a recombination event between the most distal STRP marker and locus D4Z4, observed in individual 30-48, highlighted the fact that Southern blot analysis is essential in the molecular diagnosis of FSHD.

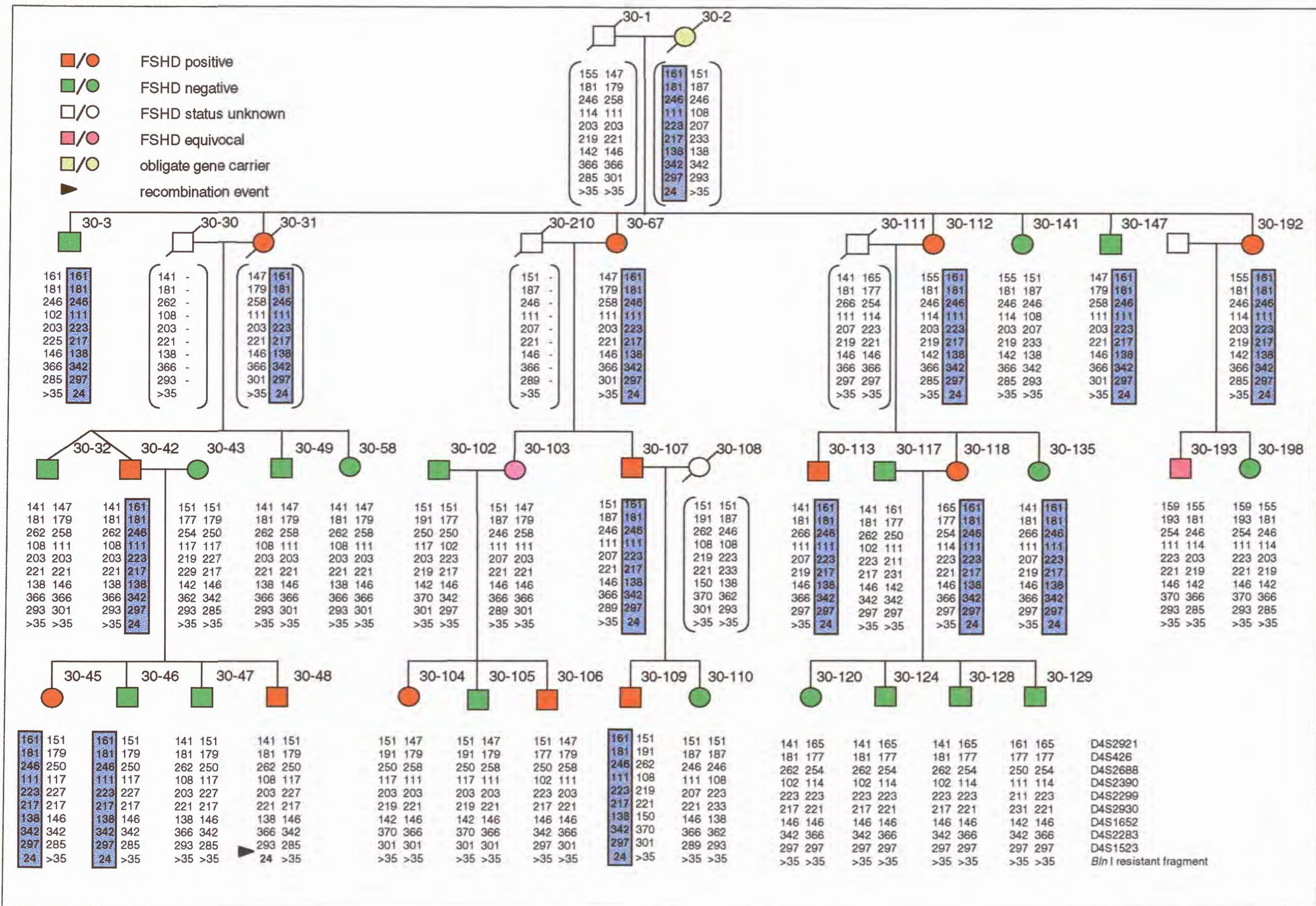


Figure 4.15: Haplotypes of selected individuals from family F30

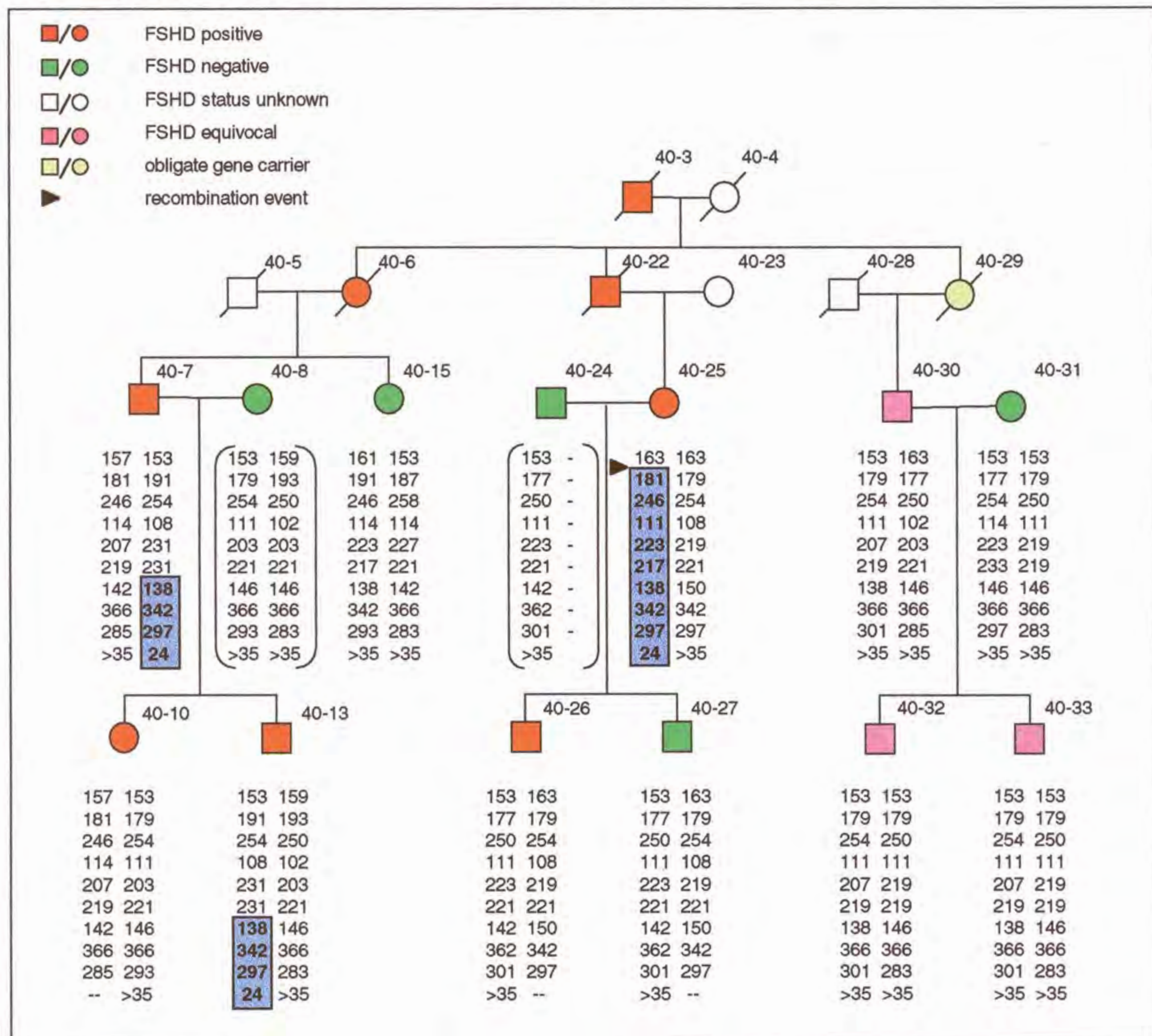
4.4.4 FSHD family F40

This family consists of 40 individuals of whom 11 were genotyped. Five of these individuals were FSHD positive, three were FSHD negative, and three were diagnosed to be equivocal on a clinical level.

Haplotypes of these 11 individuals are presented in Figure 4.16. The haplotypes were inferred for those individuals (40-8, and 40-24) of whom blood samples were not available. Only a portion of the haplotype of individual 40-24 could be inferred, as both his children inherited the same haplotype in this region. The FSHD associated haplotype segregating in family F40 is identical to the haplotype observed to segregate in families F10 and F30. As mentioned in paragraph 4.4.1 the recombination event observed in individual 40-25 facilitated elucidation of the exact order of the STRP markers. Individuals 40-7 and 40-13 inherited only a portion of the FSHD associated haplotype, including three markers, and confirmed the STRP marker order in that region. Individuals 40-10 and 40-26 who were phenotyped to be FSHD positive and the three equivocal individuals (40-30, 40-32 and 40-33) did, however, not inherit the FSHD associated haplotype.

Identical to families F10 and F30, a 24 kb *Bln* I resistant deletion fragment was observed to segregate in this family. The three individuals who inherited the FSHD associated haplotype all inherited the 24 kb *Bln* I resistant deletion fragment. The proband of family F40, individual 40-10, was clinically diagnosed in 1999 to be severely affected. Upon Southern blot analysis no *Bln* I resistant deletion fragment was observed in this individual. However, she did inherit a 27 kb *Eco* RI deletion fragment, the same as in her affected father and brother, albeit a *Bln* I sensitive fragment. The Southern blot analysis of individual 40-10 was repeated four times, but no *Bln* I resistant deletion fragment could be detected. The same scenario was observed for individuals 40-26 and 40-27. The 27 kb *Eco* RI deletion fragment observed in these three individuals, 40-10, 40-26 and 40-27, therefore displayed a chromosome 10 restriction endonuclease pattern. To exclude chromosome 4 or 10, a *Xap* I digestion needs to be performed in the future. If the fragment originates from chromosome 4 it should be *Xap* I sensitive, and *Xap* I resistance will designate it as a fragment from chromosome 10. The possibility of a translocation event is not very likely, but it can be excluded with either the *Bgl* II/*Bln* I dosage test or PFGE analysis.

Figure 4.16: Haplotypes of selected individuals from family F40



-- = no *Bln* I resistant deletion fragment was detected.

The same *Bln* I resistant deletion fragment of 24 kb and FSHD associated haplotype, indicated in blue in Figure 4.16, was observed to segregate in families F10, F30 and F40, indicating the presence of a common ancestor for these families. The presence of *Bln* I sensitive fragments in clinically affected individuals (40-10, 40-26 and 40-27) confirmed the importance of PFGE in the diagnosis of FSHD on a molecular level. This will enable the visualisation of all four alleles, two from chromosome 4 and two from chromosome 10, at the D4Z4 locus.

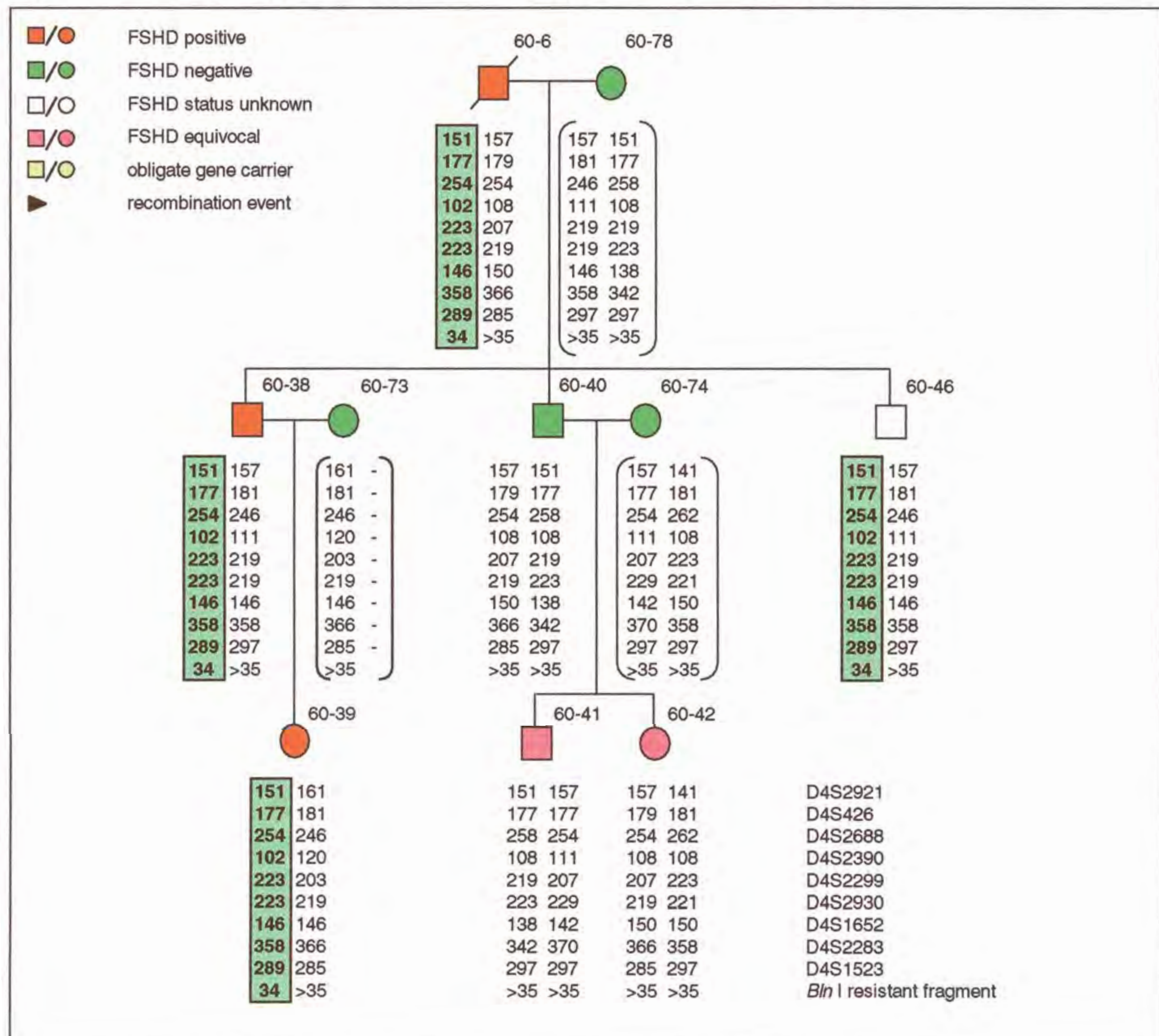
4.4.5 FSHD family F60

The core family selected for this study from family F60 included three FSHD positive, one FSHD negative, and three individuals of whom the clinical diagnosis was unknown. Figure 4.17 presents haplotypes for ten individuals from this family who were genotyped at

nine loci with the STRP markers discussed in paragraph 4.1. Deduced haplotypes are indicated for individuals 60-78, 60-73 and 60-74, of whom blood samples were not available. Only a portion of the haplotype of individual 60-73 could, however, be deduced, as only one child was genotyped.

The same FSHD associated haplotype (indicated in green in Figures 4.13 and 4.14) observed in family F20 and individuals 15-31, 15-33 and 15-34, from family F10, was observed to segregate in family F60. It is, however, evident from Figure 4.14 that only the six distal markers of the haplotype observed in family F20 are identical to the haplotype observed in family F60, whereas families F10 and F60 have the entire haplotype of nine markers in common. The three affected individuals (60-6, 60-38 and 60-39) as well as individual 60-46, whose clinical FSHD status was unknown, inherited this FSHD associated haplotype. The two clinically equivocal individuals (60-41 and 60-42) did not inherit the haplotype.

Figure 4.17: Haplotypes of selected individuals from family F60



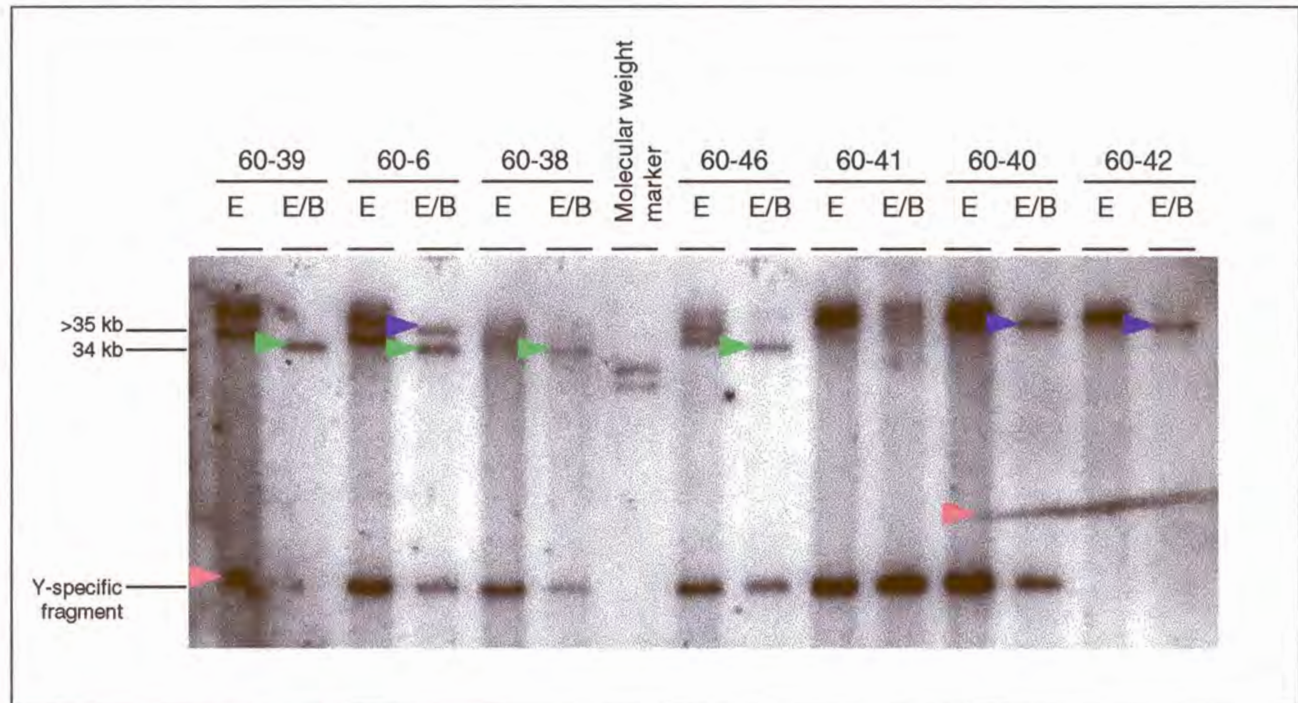
A *Bln* I resistant deletion fragment of 34 kb was observed to segregate in family F60 as presented in Figure 4.18. Complete concordance between the FSHD associated haplotype and the deletion fragment was observed. This deletion fragment was also detected to segregate with the FSHD associated haplotype observed in family F20.

In Figure 4.18, the horizontal line stretching across the last three lanes, indicated by a pink arrow is due to artefact. No 9.5 kb Y-specific fragment was observed for the female individuals 60-39 and 60-42. An artifact is, however, visible in the first lane, also indicated by a pink arrow, but it is evident that this is not a fragment, as individual 60-39 is female, and the artefact is not present in the second lane.

Two *Bln* I resistant deletion fragments are present in individual 60-6, one of 34 kb and the other of greater than 35 kb, as presented in Figure 4.18. The exact size of the larger deletion fragment could not be determined due to the limitation of the conventional gel utilised. It is however evident from Figure 4.18 that this deletion fragment is larger than the 37 kb *Eco* RI deletion fragment that was reduced to the 34 kb *Bln* I resistant fragment, and is therefore outside the reported pathogenic range. Caution should, however, be taken as the pathogenic allele size threshold for the South African population might be different than that reported for other populations (10 to 35 kb). Individual 60-40 (clinically unaffected) and his son, individual 60-42 (clinically equivocal), inherited the >35 kb *Bln* I resistant deletion fragment. Individual 60-41, the oldest son of individual 60-40, did not inherit the >35 kb *Bln* I resistant deletion fragment from his father. This was also confirmed with the haplotype analysis (Figure 4.17), as the two sons, individuals 60-41 and 60-42, inherited different chromosomes from their father, individual 60-40. It is, therefore essential that individuals 60-40 and 60-42 be clinically re-examined. The exact size of the deletion fragment should be determined with PFGE.

Family F60 harbours the same FSHD associated haplotype and *Bln* I resistant deletion fragment of 34 kb as family F20 and individuals 15-31, 15-33 and 15-34. Upon molecular analysis, it is evident that individual 15-31 forms the link between families F20 and F60.

Figure 4.18: Southern blot analysis for family F60



= 34 kb *Bln* I resistant deletion fragment;

= >35 kb *Bln* I resistant deletion fragment;

= artefact.

FSHD was confirmed for the five families investigated in this study via molecular analysis. It will therefore be possible to provide a diagnostic service for these families as well as other families in the South African population. Genotype-phenotype discordance, where individuals were clinically diagnosed to be affected, but did not inherit the FSHD associated haplotype or *Bln* I resistant deletion fragment was also observed. These individuals were, however, all diagnosed at a young age and further clinical examination is required.

To date, two different FSHD associated haplotypes and *Bln* I resistant deletion fragments (24 kb and 34 kb) have been observed to segregate in the South African population. This revealed the fact that two distinct ancestral origins are responsible for the current FSHD phenotype segregating in the South African population. It is not known how each of the FSHD genetic backgrounds evolved or how and when these mutations entered South Africa. These issues remain to be resolved in one of the next phases of this ongoing research programme.

CHAPTER FIVE

CONCLUSION

The clinical diagnosis of individuals affected with FSHD is complicated due to extreme clinical variability. Even the differentiation between various muscular dystrophies can be challenging on the clinical level, complicating specific diagnosis of a particular muscular dystrophy. This emphasises the need and relevance of molecular diagnoses of all muscular dystrophies, but in particular for those muscular dystrophies that are as clinically heterogeneous as FSHD.

In this study FSHD was investigated for the first time on a molecular level in South Africa and enabled definitive diagnoses of multiple individuals. The molecular study included both haplotype analysis, utilising nine STRP markers, and Southern blot analysis with probe p13E-11 to detect DNA rearrangements at locus D4Z4 in individuals from five extended FSHD families (F10, F20, F30, F40 and F60). Although the initial aim of this study was to genotype 74 individuals with five STRP markers, a total of 100 individuals were eventually genotyped utilising nine STRP markers. The exact order of the nine STRP markers was also mapped in this study via several recombination events observed within the families investigated.

In the group of 100 individuals, 12 individuals were related to their respective families by marriage and 83 individuals were definitively diagnosed for the first time. As discussed in paragraphs 4.4.2 and 4.4.4 a definitive diagnosis could not be made for the remaining five individuals (40-10, 40-26, 40-27, 21-69 and 21-75). These individuals were clinically diagnosed with FSHD and all harboured the *Eco* RI deletion fragment segregating in their respective families. The *Eco* RI fragments were digested with *Bln* I signifying a chromosome 10 origin. No *Bln* I resistant deletion fragments could therefore be confirmed for these individuals. It thus seems as if the deletion fragments observed in these five individuals resides on chromosome 10 and are therefore not pathogenic.

5.1 SOUTH AFRICAN FSHD FAMILIES INVESTIGATED

Family F10 was the largest South African FSHD family investigated in this study. A definitive molecular diagnosis could be made for all the individuals of this family who were included in this study. A *Bln* I resistant deletion fragment of 24 kb was observed to segregate with the FSHD associated haplotype as well as the FSHD phenotype (Figure 4.12). Genotype-phenotype discordance was observed in only one individual (14-113) who was clinically phenotyped to be affected, but who did not inherit the deletion fragment. This individual was diagnosed in his late fifties, an age at which the phenotype should have been fully penetrant. The clinical diagnosis was, however, made by a clinician not involved in the study and it is therefore suggested that this individual should be clinically re-evaluated by a clinician who specialises in FSHD. The clinical re-evaluation of individual 14-113 will be essential to substantiate his molecular diagnosis.

The first compound heterozygote individual (15-31) in the South African population was also observed in family F10 (Olckers *et al.*, 2001; Van der Merwe *et al.*, 2001b). This highlights the importance of appropriate genetic counseling for this individual as well as his children, who will all inherit FSHD from their father. The chromosomal origin of the deletion fragments was verified with chromosome 10 haplotype analysis in a separate study (Olckers *et al.*, 2001; Van der Merwe *et al.*, 2001b). The presence of two different sizes *Bln* I resistant deletion fragments within one family might be useful in the elucidation of the phenotypic-genotypic correlation in FSHD. Clinical re-evaluation by a clinician specialised in FSHD is therefore also essential for this family.

A 34 kb *Bln* I resistant deletion fragment, and a distinct FSHD associated haplotype from that in family F10 was observed to segregate in family F20, as presented in Figure 4.14. Genotype-phenotype discordance was also observed in this family. Five individuals (21-65, 21-69, 21-71, 21-75 and 21-76) who were clinically diagnosed with FSHD did not inherit the haplotype or the deletion fragment. It can therefore be argued that the phenotype, haplotype and deletion fragment observed are not associated with FSHD in this family. These five individuals were, however, clinically diagnosed at a young age (before the age of 20), and it has been observed that the clinical diagnosis of children is much more complicated than that of adults, thus necessitating re-examination of these individuals on a clinical level at a more advanced age. It is likely that these individuals, who are now in their twenties, will be negative upon clinical examination, confirming the molecular analysis. If they are, however, observed to be clinically affected, further

molecular analysis will be necessary to elucidate the molecular nature of the phenotype-genotype discordance in this family. For this reason, the clinical evaluation of these individuals is a high priority during the next phase of this ongoing project.

Sub-family F21 of family F20 was genotyped as a representative family for the extended family. It was, however, revealed by the presence of a compound heterozygote individual in sub-family F15 of family F10, that one sub-family does not provide sufficient representation of the entire family. If sub-family F15 was for example not included in this study, the compound heterozygote individual would not have been observed and the spectrum of molecular complexity present in this family might not have been revealed and appreciated. Individuals from the other sub-families of family F20 will therefore be included in an extended study in the future.

Family F30 was observed to harbour the same 24 kb *Bln* I resistant deletion fragment and FSHD associated haplotype as family F10, as illustrated in Figure 4.15. Only two individuals (30-104 and 30-106), who were clinically phenotyped to be FSHD positive, did not inherit the deletion fragment. These individuals were younger than 20 years of age upon clinical examination and need to be clinically re-evaluated to confirm the proposed discordance.

A 27 kb *Eco* RI deletion fragment corresponding to a 24 kb *Bln* I resistant deletion fragment was also observed to segregate in family F40. The 24 kb *Bln* I resistant deletion fragment was observed to segregate with the same FSHD associated haplotype, indicated in Figure 4.16, as in families F10 and F30. No definitive molecular diagnosis could be made for three individuals (40-10, 40-26 and 40-27) from this family, although these individuals all harboured a 27 kb *Eco* RI deletion fragment. This fragment was digested with *Bln* I, therefore displaying the characteristics of a chromosome 10-type deletion fragment. Individual 40-10 is, however, the proband of this family, and was recently clinically diagnosed with FSHD by a neurologist specialising in FSHD. In the diagnosis of FSHD on a molecular level, PFGE enables the visualisation of all four alleles. The importance of PFGE analysis was confirmed with the presence of *Bln* I sensitive fragments in these clinically affected individuals. Diagnosis of these individuals was thus postponed until PFGE analyses could be utilised to confirm the chromosomal origin of the 27 kb *Eco* RI deletion fragment.

Family F60 was observed to harbour the same FSHD associated haplotype and 34 kb *Bln* I resistant deletion fragment as those segregating in family F20. This indicates that families F20 and F60 should thus have a common ancestor, which will be further investigated in future studies. It was evident from the results obtained for individual 60-6, as discussed in paragraph 4.4.5 and illustrated in Figure 4.18, that it will be necessary in the future to determine the exact allele size threshold, and therefore the pathogenic range, in the South African population to provide accurate molecular diagnosis. In addition, utilisation of PFGE instead of conventional horizontal gel electrophoresis will undoubtedly increase both the specificity and sensitivity of the molecular diagnosis.

5.2 GENETIC HETEROGENEITY

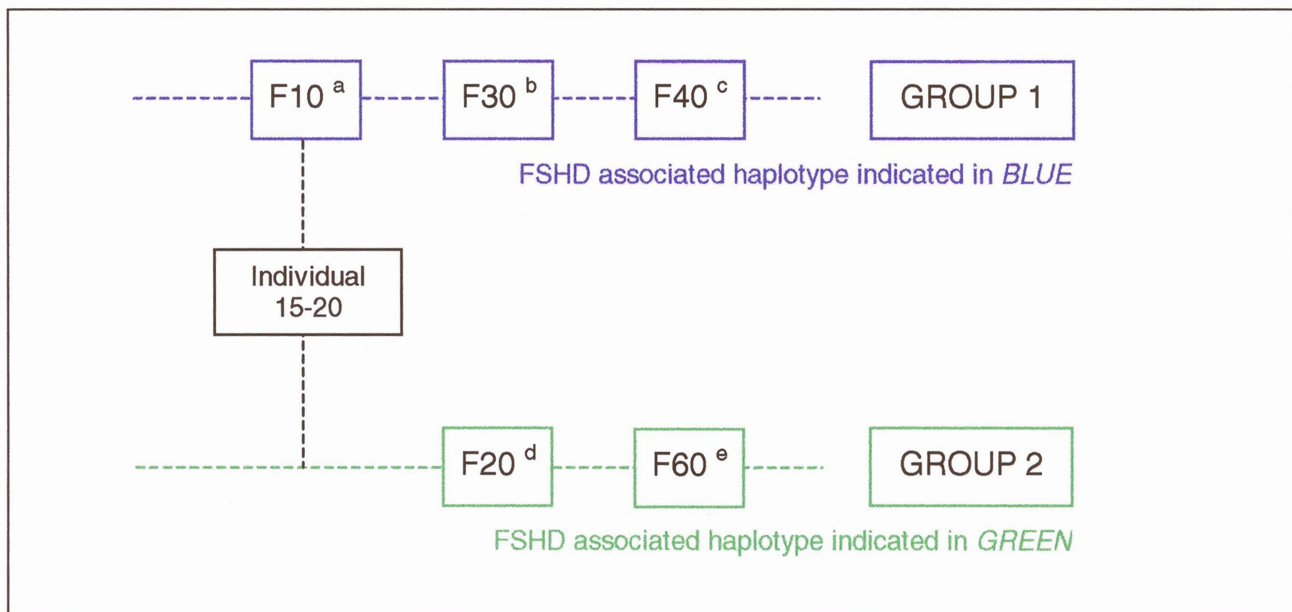
The utilisation of *Bln* I was observed to improve the molecular diagnosis of FSHD in the South African population. *Bln* I resistant deletion fragments associated with the FSHD phenotype were detected in all five families investigated (Olckers *et al.*, 2000; Van der Merwe *et al.*, 2000b). From the results generated in this study, it was concluded that no evidence for genetic heterogeneity was observed in the group of South African families included in this study. To determine whether genetic heterogeneity exists in the South African FSHD population, an even larger group of clinically well characterised families should be investigated. This is currently one of the aims of an extended FSHD research programme.

5.3 DUAL FOUNDER EFFECT

The possibility of a single ancestral mutation in the South African FSHD population was excluded by the presence of two distinct haplotypes, each co-segregating with a specific *Bln* I resistant deletion fragment (Van der Merwe *et al.*, 2000a). This indicates distinct genetic backgrounds for FSHD in South Africa, and thus two ancestral origins for the current FSHD phenotype observed in the South African population. Results to date indicate the presence of a dual Founder Effect for FSHD in the South African population. This is analogous to familial hypercholesterolemia in the South African Afrikaner population in which multiple founder mutations were observed (Vergotine *et al.*, 2001). Genotype-phenotype comparative studies between the two South African FSHD family groups will shed light on the significance of the ancestral origins of FSHD in the South African population.

Due to elucidation of some aspects of the molecular basis of FSHD in the South African population, the relationship between the five families investigated could be determined. Individual 15-31 was found to be a compound heterozygote with regard to the two FSHD-associated genotypes that segregated in the South African families investigated. Individual 15-20, the mother of 15-31, from family F10 represents the link between the two groups of South African families, namely groups 1 and 2, as indicated in Figure 5.1 (Olckers *et al.,* 2001).

Figure 5.1: Relationship between five FSHD families investigated in this study



a = Family F10 presented in Figure 4.12, b = Family F30 presented in Figure 4.15, c = Family F40 presented in Figure 4.16, d = Family F20 presented in Figure 4.14, e = Family F60 presented in Figure 4.17.

5.4 FUTURE DIRECTIONS IN FSHD RESEARCH

At the time of analysis the *Xap* I analysis had not yet been reported. In cases where no definitive diagnosis could be made, for example individuals 40-10, 40-27 an 40-27, (as discussed in paragraph 4.4.4) the utilisation of *Xap* I analysis might assist in the characterisation of the genetic defect. The translocation frequency between the 3.3 kb repeats on chromosomes 4q35 and 10q26 in the South African FSHD population and in the general population will in the future be investigated via the utilisation of the *Bgl* II/*Bln* I dosage test.

Calculation of LOD scores was not an aim of this study. It could be argued that statistical evidence might be necessary given the various cases of phenotype-genotype discordance that were observed in this study. However, the discordance was primarily observed in young individuals under the age of 20, and the presence of DNA rearrangements due to

the deletion of 3.3 kb repeat units at the D4Z4 locus, has been demonstrated to be causative of FSHD. For these reasons, the calculation of LOD scores may only become an aim of this study after the re-evaluation of these individuals by a specialised clinician.

In the future a phenotype-genotype correlation study needs to be performed on clinically well characterised individuals. As mentioned previously, the clinical diagnosis of FSHD is very difficult due to extreme clinical variability of the FSHD phenotype. It will be essential to get one clinician, specialised in FSHD diagnoses, to clinically evaluate all individuals in order to exclude clinical bias. In addition the aspect of anticipation has to be addressed. As mentioned in paragraph 2.1.2.4, a shortcoming of all the papers reporting anticipation in their populations was that they investigated only affected parent-offspring pairs, thus only two generations. It is foreseen that the extended FSHD families in the South African population will enable the investigation of anticipation over several generations, most likely substantiating the study reported by Flanigan *et al.* (2000).

Results generated in this study indicated that it was extremely useful to have utilised both haplotype and Southern blot analyses in our elucidation of the molecular basis of FSHD in the South African population (Van der Merwe *et al.,* 1999; Olckers *et al.,* 2000). The combination of haplotype and Southern blot analysis definitively identified the relationship between the families investigated. Moreover, haplotype analysis also supported the presence or absence of *Bln* I resistant deletion fragments in the families investigated. Future studies will reveal whether all South African FSHD families have only the two ancestral origins reported here. In addition, the detailed haplotype data generated via this study will allow comparison of the South African FSHD population with other populations. This will undoubtedly facilitate determination of the global ancestral origins of FSHD.

Several international groups have made significant advances in FSHD research since the identification of the FSHD associated DNA rearrangements at the D4Z4 locus in 1992. This study represents the first attempt to unravel the molecular basis of FSHD in the South African population. Only via collaboration across several disciplines will we be able to provide effective support to those affected by this complex and highly variable phenotype. To date several structures and resources including a molecular diagnostic service, an FSHD clinic, a physiotherapy clinic and a specific support group have been put in place to support individuals affected by FSHD. Identification of the dual Founder Effect of the FSHD phenotype in the South African population paved the way for an effective molecular diagnostic FSHD service in this population.

REFERENCES

Andrews C.T., Taylor T.C. and Patterson V.H. Scapulothoracic arthrodesis for patients with facioscapulohumeral muscular dystrophy. *Neuromusc. Disord.*, **8**, 580-584, 1998.

Arahata K., Ishihara T., Fukunaga H. Orimo S., Lee J.E.H., Goto K. and Nonaka I. Inflammatory response in facioscapulohumeral muscular dystrophy (FSHD): Immunocytochemical and genetic analyses. *Muscle Nerve*, **Suppl 2**, S56-S66, 1995.

Baily R.O., Marzulo D.C. and Hans M.B. Infantile facioscapulohumeral muscular dystrophy: new observations. *Acta Neurol. Scand.*, **74**, 51-58, 1986.

Bakker E., Wijmenga C., Vosen R.H.A.M., Padberg G.W., Hewitt J., Van der Wielen M., Rasmussen K. and Frants R.R. The FSHD-linked locus D4F104S1 (p13E-11) on 4q35 has a homologue on 10qter. *Muscle Nerve,* **Suppl 2**, S39-S44, 1995.

Bakker E., Van der Wielen M.J.R., Voorhoeve E., Ippel P.F., Padberg G.W., Frants R.R. and Wijmenga C. Diagnostic, predictive, and prenatal testing for facioscapulohumeral muscular dystrophy: diagnostic approach for sporadic and familial cases. *J. Med. Genet.*, **33**, 29-35, 1996.

Ballard L.W. Personal communication. Utah Center for Human Genome Research University of Utah, Dept. of Human Genetics, Eccles Institute of Human Genetics, Salt Lake City, UT, U.S.A., 1999.

Beckers M.C., Gabriëls J., Van der Maarel S., De Vriese A., Frants R.R., Collen D. and Belayew A. Active genes in junk DNA? Characterization of *DUX* genes embedded within 3.3 kb repeated elements. *Gene*, **264**, 51-57, 2001.

Bennett R.L., Steinhaus K.A., Uhrich S.B., O'Sullivan C.K., Resta R.G., Lochner-Doyle D., Markel D.S., Vincent V. and Hamanishi J. Recommendations for standardized human pedigree nomenclature. *Am. J. Hum. Genet.,* **56**, 745-752, 1995.

Bouju S., Piétu G., Le Cunff M., Cros N., Malzac P., Pellissier J-F., Pons F., Léger J-J., Auffray C. and Dechesne C.A. Exclusion of muscle specific actinin-associated LIM protein (ALP) gene form 4q35 facioscapulohumeral muscular dystrophy (FSHD) candidate genes. *Neuromusc. Disord.*, **9**, 3-10, 1999.

Brouwer O.F., Padberg G.W., Ruys C.J.M., Brand R., De Laat J.A.P.M. and Grote J.J. Hearing loss in facioscapulohumeral muscular dystrophy. *Neurology*, **41**, 1878-1881, 1991.

Brouwer O.F., Wijmenga C., Frants R.R. and Padberg G.W. Facioscapulohumeral muscular dystrophy: the impact of genetic research. *Clin. Neurol. Neurosurg.*, **95**, 9-21, 1993.

Brouwer O.F., Padberg G.W., Wijmenga C. and Frants R.R. Facioscapulohumeral muscular dystrophy in early childhood. *Arch. Neurol.*, **51**, 387-394, 1994.

Brouwer O.F., Padberg G.W., Bakker E., Wijmenga C. and Frants R.R. Early onset facioscapulohumeral muscular dystrophy. *Muscle Nerve*, **Suppl 2**, S67-S72, 1995.

Bunch W.H. and Siegel I.M. Scapulothoracic arthrodesis in facioscapulohumeral muscular dystrophy. *J. Bone Joint Surg.*, **75-A** (3), 372-376, 1993.

Bushby K.M.D., Pollitt C., Johnson M.A., Rogers M.T. and Chinnery P.F. Muscle pain as a prominent feature of facioscapulohumeral muscular dystrophy (FSHD): four illustrative case reports. *Neuromusc. Disord.*, **8**, 574-579, 1998.

Busse K., Köhler J., Stegmann K., Pongratz D., Koch M.C. and Schreiber H. An inherited 4q35-*Eco*RI-DNA-fragment of 35 kb in a family with a sporadic case of facioscapulohumeral muscular dystrophy (FSHD). *Neuromusc. Disord.*, **10**, 178-181, 2000.

Cacurri S., Deidda G., Piazzo N., Novelletto A., La Cesa I., Servidei S., Galluzzi G., Wijmenga C., Frants R.R. and Feliceti L. Chromosome 4q35 haplotypes and DNA rearrangements segregating in affected subjects of 19 Italian families with facioscapulohumeral muscular dystrophy (FSHD). *Hum. Genet.*, **94**, 367-374, 1994.

Cacurri S., Piazzo N., Deidda G., Vigneti E., Galluzzi G., Colantoni L., Merico B., Ricci E. and Felicetti L. Sequence homology between 4qter and 10qter loci facilitates the instability of subtelomeric *Kpn*I repeat units implicated in Facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.*, **63**, 181-190, 1998.

Church G.M. and Gilbert W. Genomic sequencing. *Proc. Natl. Acad. Sci. USA.,* **81**, 1991-1995, 1984.

Chyatte S.B., Vignos P.J. and Watkins M. Early muscular dystrophy: differential patterns of weakness in Duchenne, Limb-girdle and facioscapulohumeral types. *Arch. Phys. Med.* **47**, 499-503, 1966.

Deidda G.C., Cacurri S., La Cesa I., Scoppetta C. and Felicetti L. 4q35 molecular probes for the diagnosis and genetic counseling of facioscapulohumeral muscular dystrophy. *Ann. Neurol.*, **36** (1), 117-118, 1994.

Deidda G., Cacurri S., Grisanti P., Vigneti E., Piazzo N. and Felicetti L. Physical mapping evidence for a duplicated region on chromosome 10qter showing high homology with the facioscapulohumeral muscular dystrophy locus on chromosome 4qter. *Eur. J. Hum. Genet.*, **3**, 155-167, 1995.

Deidda G., Cacurri S., Piazzo N. and Felicetti L. Direct detection of 4q35 rearrangements implicated in facioscapulohumeral muscular dystrophy (FSHD). *J. Med. Genet.*, **33**, 361-365, 1996.

Dib C., Faure S., Fizames C., Samson D., Drouot N., Vignal A., Millasseau P., Marc S., Hazan J., Seboun E., Lathrop M., Gyapay G., Morissette J. and Weissenbach J. A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature,* **380**, 152-154, 1996.

Ding H., Beckers M.-C., Plaisance S., Marynen P., Collem D. and Belayew A. Characterization of a double homeodomain protein (DUX1) encoded by a cDNA homologous to 3.3 kb dispersed repeated elements. *Hum. Mol. Genet.*, **7** (11), 1681-1694, 1998.

Dorland's Pocket Medical Dictionary, 25th edition. W.B. Saunders Company (A division of Harcourt Brace & Company), Philadelphia, PA, U.S.A., 1995.

Eggers S., Passos-Bueno M.R. and Zatz M. Facioscapulohumeral muscular dystrophy: aspects of genetic counselling, acceptance of preclinical diagnosis, and fitness. *J. Med. Genet.*, **30**, 589-593, 1993.

Emery A.E.H. The muscular dystrophies: Clinical reviews. *BMJ.,* **317**, 991-995, 1998.

Fan Y-S., Yang H-M. and Lin C.C. Assignment of the human muscle adenine nucleotide translocator gene (ANT1) to 4q35 by fluorescence in situ hybridization. *Cytogenet. Cell Genet.*, **60**, 29-30, 1992.

Felice K.J., North W.A., Moore S.A. and Mathews K.D. FSH dystrophy 4q35 deletion in patients presenting with facial-sparing scapular myopathy. *Neurology*, **54**, 1927-1931, 2000.

Felice K.J. and Moore S.A. Unusual clinical presentation in patients harboring the facioscapulohumeral dystrophy 4q35 deletion. *Muscle Nerve*, **24**, 352-356, 2001.

Finsterer J. and Stöllberger C. Cardiac involvement in primary myopathies. *Cardiology*, **94**, 1-11, 2000.

Fisher J. and Upadhyaya M. Molecular genetics of facioscapulohumeral muscular dystrophy (FSHD). *Neuromusc. Disord.*, **7**, 55-62, 1997.

Fitzsimons R.B., Gurwin E.B. and Bird A.C. Retinal vascular abnormalities in facioscapulohumeral muscular dystrophy: A general association with genetic and therapeutic implications. *Brain*, **110**, 631-648, 1987.

Flanigan K.M., Coffeen C.M., Sexton L., Stauffer D., Brunner S. and Leppert M.F. Genetic characterization of a large, historically significant Utah kindred with facioscapulohumeral dystrophy. *Neuromusc. Disord.*, **11**, 525-529, 2001.

Funakoshi M., Goto K. and Arahata K. Epilepsy and mental retardation in a subset of early onset 4q35 facioscapulohumeral muscular dystrophy. *Neurology*, **50**, 1791-1794, 1998.

Gabriëls J., Beckers M.-C., Ding H., De Vriese A., Plaisance S., Van der Maarel S.M., Padberg G.W., Frants R.R., Hewitt J.E., Collen D. and Belayew A. Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. *Gene*, **236**, 25-32, 1999.

Galluzzi G., Deidda G., Cacurri S., Colantoni L., Piazzo N., Vigneti E., Ricci E., Servidei S., Merico B., Pachì, Brambati B., Mangiola F., Tonali P. and Felicetti L. Molecular analysis of 4q35 rearrangements in facioscapulohumeral muscular dystrophy (FSHD): application to family studies for a correct genetic advice and a reliable prenatal diagnosis of the disorder. *Neuromusc. Disord.*, **9**, 190-198, 1999.

Gieron M.A., Korthals J.K. and Kousseff B.G. Facioscapulohumeral dystrophy with cochlear hearing loss and tortuosity of retinal vessels. *Am. J. Med. Genet.*, **22**, 143-147, 1985.

Gilbert J.R., Stajich J.M., Speer M.C., Vance J.M., Stewart C.S., Yamaoka L.H., Samson F., Fardeau M., Potter T.G., Roses A.D. and Pericak-Vance M.A. Linkage studies in facioscapulohumeral muscular dystrophy (FSHD). *Am. J. Hum. Genet.,* **51**, 424-427, 1992.

Gilbert J.R., Stajich J.M., Wall S., Carter S.C., Qiu H., Vance J.M., Stewart C.S., Speer M.C., Pufky J., Yamaoka L.H., Rozear M., Samson F., Fardeau M., Roses A.D. and Pericak-Vance M.A. Evidence for heterogeneity in facioscapulohumeral muscular dystrophy (FSHD). *Am. J. Hum. Genet.*, **53**, 401-408, 1993.

Gerken S.C., Matsunami N., Lawrence E., Carlson M., Moore M., Ballard L., Melis R., Robertson M., Bradley P., Elsner T., Tingey A., Rodriguez P., Albertsen H., Lalouel J-M. and White R. Genetic and physical mapping of simple sequence repeat containing sequence tagged sites from the human genome. Genbank accession number L16394, 1993a.

Gerken S.C., Matsunami N., Lawrence E., Carlson M., Moore M., Ballard L., Melis R., Robertson M., Bradley P., Elsner T., Tingey A., Rodriguez P., Albertsen H., Lalouel J-M. and White R. Genetic and physical mapping of simple sequence repeat containing sequence tagged sites from the human genome. Genbank accession number L18154, 1993b.

Gerken S.C., Matsunami N., Lawrence E., Carlson M., Moore M., Ballard L., Melis R., Robertson M., Bradley P., Elsner T., Tingey A., Rodriguez P., Albertsen H., Lalouel J-M. and White R. Genetic and physical mapping of simple sequence repeat containing sequence tagged sites from the human genome. Genbank accession number L17998, 1993c.

Gerken S.C., Matsunami N., Plaetke R., Albertsen H., Ballard L., Melis R., Lawrence E., Moore M., Holik P.R., Carlson M., Shao X., Robertson M., Bradley P., Elsner T., Tingey A., Lalouel J-M. and White R. Genetic and physical mapping of simple sequence repeat containing sequence tagged sites from the human genome. Genbank accession number L30309, 1994.

Grewel P.K., Todd L.C., Van der Maarel S., Frants R.R. and Hewitt J.E. FRG1, a gene in the FSHD muscular dystrophy region on human chromosome 4q35, is highly conserved in vertebrates and invertebrates. *Gene*, **216**, 13-19, 1998.

Griggs R.C., Tawil R., Storvick D., Mendell J.R. and Altherr M.R. Genetics of facioscapulohumeral muscular dystrophy: new mutations in sporadic cases. *Neurology*, **43**, 2369-2372, 1993.

Griggs R.C., Tawil R., McDermott M., Forrester J., Figlewicz D., Weiffenbach B. and the FSH-DY group. Monozygotic twins with facioscapulohumeral dystrophy (FSHD): Implications for genotype/phenotype correlation. *Muscle Nerve*, **Suppl 2**, S50-S55, 1995.

Gurwin E.B., Fitzsimons R.B., Sehmi K.S. and Bird A.C. Retinal telangiectasis in facioscapulohumeral muscular dystrophy with deafness. *Arch. Opthalmol.*, **103**, 1695-1700, 1985.

Haraguchi Y., Chung A.B., Torroni A., Stepien G., Shoffner J.M., Wasmuth J.J., Costigan D.A., Polak M., Altherr M.R., Winokur S.T. and Wallace D.C. Genetic mapping of human heart-skeletal muscle adenine nucleotide translocator and its relationship to the facioscapulohumeral muscular dystrophy locus. *Genomics*, **16**, 479-485, 1993.

Hauge X.Y. and Litt M. A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR. *Hum. Mol. Genet.*, **2** (4), 411-415, 1993.

Hewitt J.E., Lyle R., Clark L.N., Valleley E.M., Wright T.J. Wijmenga C., Van Deutekom J.C.T., Francis F., Sharpe P.T., Hofker M., Frants R.R. and Williamson R. Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. *Hum. Mol. Genet.*, **3** (8), 1287-1295, 1994.

Houldsworth J. and Attardi G. Two distinct genes from ADP / ATP translocase are expressed at the MRNA level in adult human liver. *Proc. Natl. Acad. Sci. USA.*, **85**, 377-381, 1988.

Jacobsen S.J., Diala E.S., Dorsey B.V., Rising M.B., Graveline R., Falls K., Schultz P., Hogan C., Rediker K., D'Amico C. and Weiffenbach B. A clinically homogenous group of families with facioscapulohumeral (Landouzy-Déjérine) muscular dystrophy: linkage analysis of six autosomes. *Am. J. Hum. Genet.*, **47**, 376-388, 1990.

Jardine P., Jones M., Tyfield L., Upadhyaya M. and Lunt P. *De novo* DNA rearrangement in atypical facioscapulohumeral muscular dystrophy. *Clin. Genet.*, **44**, 167, 1993.

Jardine P.E., Upadhyaya M., Maynard J., Harper P. and Lunt P.W. A scapular onset muscular dystrophy without facial involvement: possible allelism with facioscapulohumeral muscular dystrophy. *Neuromusc. Disord.*, **4** (5/6), 477-482, 1994a.

Jardine P.E., Koch M.C., Lunt P.W., Maynard J., Bathke K.D., Harper P.S. and Upadhyaya M. *De novo* facioscapulohumeral muscular dystrophy defined by DNA probe p13E-11 (D4S104S1). *Arch. Disord. Childhood*, **71**, 221-227, 1994b.

Johns M.B. and Paulus-Thomas J.E. Purification of human genomic DNA from whole blood using sodium perchlorate in place of phenol. *Anal. Biochem.*, **180**, 276-278, 1989.

Kissel J.T., McDermott M.P., Natarajan R., Mendell J.R., Pandya S., King W.M., Griggs R.C., Tawil R. and the FSH-DY group. Pilot trial of albuterol in facioscapulohumeral muscular dystrophy. *Neurology*, **50**, 1402-1406, 1998.

Kissel J.T., McDermott M.P., Mendell J.R., King W.M., Pandya S., Griggs R.C. and Tawil R. Randomized, double-blind, placebo-controlled trial of albuterol in facioscapulohumeral muscular dystrophy. *Neurology*, **57** (8), 1434-1440, 2001.

Kleinjan D-J. and Van Heyningen V. Position effect in human genetic disorder. *Hum. Mol. Genet.*, **7** (10), 1611-1618, 1998.

Köhler J., Rupilius B., Otto M., Bathke K. and Koch M.C. Germline mosaicism in 4q35 facioscapulohumeral muscular dystrophy (FSHD1A) occurring predominantly in oogenesis. *Hum. Genet.*, **98**, 485-490, 1996.

Korf B.R., Bresnan M.J., Shapiro F., Sotrel A. and Abroms I.F. Facioscapulohumeral dystrophy presenting in infancy with facial diplegia and sensorineural deafness. *Ann. Neurol.*, **17**, 513-516, 1985.

Laforêt P., De Toma C., Eymard B., Becane H.M., Jeanpierre M., Fardeau M. and Duboc D. Cardiac involvement in genetically confirmed facioscapulohumeral muscular dystrophy. *Neurology*, **51**, 1454-1456, 1998.

Landouzy L. and Déjérine J. Progressive atrophic myopathy, starting usually in childhood, by the face, without neuropathy (De la myopathie atrophique progressive, myopathie héréditaire débutant, dans l'enfance, par la face, sans altération du systéme nerveux). *C. R. Acad. Sci, Paris*, **98**, 53-55, 1884.

Landouzy L. and Déjérine J. Progressive atrophic myopathy (De la myopathie atrophique progressive). *Rev. Med. Franc.*, 5, 81-117, 1885.

Lee J.H., Goto K., Sahashi K.O., Nonaka I., Matsuda C. and Arahata K. Cloning and mapping of a very short (10 kb) *Eco*RI fragment associated with facioscapulohumeral muscular dystrophy (FSHD). *Muscle Nerve,* **Suppl 2**, S27-S31, 1995a.

Lee J.H., Goto K., Matsuda C. and Arahata K. Characterization of a tandemly repeated 3.3 kb *Kpn*I unit in the facioscapulohumeral muscular dystrophy (FSHD) gene region on chromosome 4q35. *Muscle Nerve,* **Suppl 2**, S6-S13, 1995b.

Lemmers R.J.L.F., Van der Maarel S.M., Van Deutekom J.C.T., Van der Wielen M.J.R., Deidda G., Dauwerse H.G., Hewitt J., Hofker M., Bakker E., Padberg G.W. and Frants R.R. Inter- and intrachromosomal sub-telomeric rearrangements on 4q35: implications for facioscapulohumeral muscular dystrophy (FSHD) aetiology and diagnosis. *Hum. Mol. Genet.*, **7** (8), 1207-1214, 1998.

Lemmers R.J.L.F., De Kievit P., Van Geel M., Van der Wielen M.J.R., Bakker E., Padberg G.W., Frants R.R. and Van der Maarel S.M. *Xap* I improves diagnosis of facioscapulohumeral muscular dystrophy (FSHD). FSHD International Consortium Research Meeting, Philadelphia, PA., U.S.A., 2000.

Letournel E., Fardeau M., Lytle J.O., Serrault M. and Gosselin R.A. Scapulothoracic arthrodesis for patients who have facioscapulohumeral muscular dystrophy. *J. Bone Joint Surg.*, **72-A** (1), 78-84, 1990.

Li K., Warner C.K., Hodge J.A., Minoshima S., Kudoh J., Fukuyama R., Maekawa M., Shimizu Y., Shimizu N. and Wallace D.C. A human muscle adenine nucleotide translocator gene has four exons, is located on chromosome 4, and is differentially expressed. *J. Biol. Chem.*, **264** (24), 13998-14004, 1989.

Lieberman A.P. and Fischbeck K.H. Triplet repeat expansion in neuromuscular disorder. *Muscle Nerve*, **23**, 843-850, 2000.

Lucotte G., Berriche S. and Fardeau M. Linkage analysis of French families with facioscapulohumeral muscular dystrophy. *J. Med. Genet.*, **26**, 485-486, 1989.

Lunt P.W., Compston D.A.S. and Harper P.S. Estimation of age dependent penetrance in facioscapulohumeral muscular dystrophy by minimising ascertainment bias. *J. Med. Genet.*, **26**, 755-760, 1989a.

Lunt P.W. A workshop on facioscapulohumeral (Landouzy-Déjérine) disorder, Manchester, 16 to 17 November 1988. *J. Med. Genet.* **26**, 535-537, 1989b.

Lunt P.W. and Harper P.S. Genetic counselling in facioscapulohumeral muscular dystrophy. *J. Med. Genet.*, **28**, 655-664, 1991.

Lunt P.W., Jardine P.E., Koch M.C., Maynard J., Osborn M., Williams M., Harper P.S. and Upadhyaya M. Correlation between fragment size at D4F104S1 and age at onset or at wheelchair use, with a possible generational effect, accounts for much phenotypic variation in 4q35 – facioscapulohumeral muscular dystrophy (FSHD). *Hum. Mol. Genet.*, **4** (5), 951-958, 1995a.

Lunt P.W., Jardine P.E., Koch M., Maynard J., Osborn M., Williams M., Harper P.S. and Upadhyaya M. Phenotypic-genotypic correlation will assist genetic counseling in 4q35-facioscapulohumeral muscular dystrophy. *Muscle Nerve*, **Suppl 2**, S103-S109, 1995b.

Lunt P.W. 44th ENMC International Workshop: Facioscapulohumeral muscular dystrophy: Molecular studies. *Neuromusc. Disord.*, **8**, 126-130, 1998.

Lunt P.W. Facioscapulohumeral muscular dystrophy: Diagnostic and molecular aspects. *Monogr. Clin. Neurosci.*, **18**, 44-60, 2000.

Lyle R., Wright T.J., Clark L.N. and Hewitt J.E. The FSHD-Associated repeat, D4Z4, is a member of a dispersed family of homeobox-containing repeats, subsets of which are clustered on the short arms of the acrocentric chromosomes. *Genomics*, **28**, 389-397, 1995.

Mak W., and Ho S.L. The impact of molecular biology on clinical neurology. *H.K.M.L.,* **7**, 40-49, 2001.

Mathews K.D., Mills K.A., Bosch E.P., Ionasescu V.V., Wiles K.R., Buetow K.H. and Murray J.C. Linkage localization of facioscapulohumeral muscular dystrophy (FSHD) in 4q35. *Am. J. Hum. Genet.,* **51**, 428-431, 1992.

Matsuzaka T., Sakuragawa N., Terasawa K. and Kuwabara H. Facioscapulohumeral dystrophy associated with mental retardation, hearing loss, and tortuosity of retinal arterioles. *J. Child Neurol.*, **1**, 218-223, 1986.

McGarry J., Garg B. and Silbert S. Death in childhood due to facioscapulohumeral dystrophy. *Acta Neurol. Scand.*, **68**, 61-63, 1983.

Mills K.A., Buetow K.H., Xu Y., Ritty T.M., Mathews K.D., Bodrug S.E., Wijmenga C., Balazs I. and Murray J.C. Genetic and physical mapping on chromosome 4 narrows the localization of the gene for facioscapulohumeral muscular dystrophy (FSHD). *Am. J. Hum. Genet.*, **51**, 432-439,1992.

Milner E.C.B., Lotshaw C.L., Van Dijk K.W., Charmley P., Concannon P. and Schroeder H.W. Isolation and mapping of a polymorphic DNA sequence pH30 on chromosome 4 (HGM provisional no. D4S139). *Nucleic Acid Res.*, **17** (10), 4002, 1989.

Minoshima S., Kudoh J., Fukuyama R., Maekawa M., Shimizu Y., Li K., Wallace D.C. and Shimizu N. Mapping of the human muscle adenine nucleotide translocator gene (ANT1) to chromosome 4. *Cell. Genet.*, **51**, 1044-1045, 1989.

Miura K., Kumagai T., Matsumoto A., Iriyama E., Watanabe K., Goto K. and Arahata K. Two cases of chromosome 4q35-linked early onset facioscapulohumeral muscular dystrophy with mental retardation and epilepsy. *Neuropediatrics*, **29**, 239-241, 1998.

Mullis K., Faloona F., Scharf S., Saiki R., Horn G. and Erlich H. Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harbor Symp. Quant. Biol.*, **L1**, 263-273, 1986.

Munsat T.L., Piper D., Cancilla P. and Mednick J. Inflammatory myopathy with facioscapulohumeral distribution. *Neurology*, **22**, 335-347, 1972.

Murray J., Sheffield V., Weber J.L., Duyk G. and Buetow K.H. Cooperative Human Linkage Center. Genbank accession number G08375, 1995a.

Murray J., Sheffield V., Weber J.L., Duyk G. and Buetow K.H. Cooperative Human Linkage Center. Genbank accession number G08314, 1995b.

Neckelmann N., Li K., Wade R.P., Shuster R. and Wallace D.C. cDNA sequence of a human skeletal muscle ADP/ATP translocator: Lack of a leader peptide, divergence from a fibroblast translocator cDNA, and coevolution with mitochondrial DNA genes. *Proc. Natl. Acad. Sci. USA.*, **84**, 7580-7584, 1987.

Neuromuscular disorders: gene location. *Neuromusc. Disord.*, **11** (1), 104-106, 2001.

Okinaga A., Mutsuaka T., Umeda J., Yanagihara I., Inui K., Nagai T. and Okada S. Early-onset facioscapulohumeral muscular dystrophy: two case reports. *Brain Dev.*, **19**, 563-567, 1997.

Olckers A. Genetic heterogeneity in malignant hyperthermia susceptibility: Ph.D. thesis, University of Pretoria, South Africa, 1997.

Olckers A., Van der Merwe A., Van der Maarel S.M., Honey E., Frants R.R. and Schutte S-M. Segregation of DNA rearrangements at the D4Z4 locus in South African FSHD families. FSHD International Consortium Research Meeting, Philadelphia, PA., U.S.A., 2000.

Olckers A., Van der Merwe A., Alessandrini M., Walalce D.C., Van der Maarel S.M., Honey E., Frants R.R. and Schutte C-M. FSHD in the South African population: Evidence to date suggest a Dual Founder Efffect. FSHD International Consortium Research Meeting, San Diego, CA., U.S.A., 2001.

Orrell R.W., Tawil R., Forrester J., Kissel J.T., Mendell J.R. and Figlewicz D.A. Definitive molecular diagnosis of facioscapulohumeral muscular dystrophy. *Neurology*, **52**, 1822-1826, 1999.

Padberg G.W. Facioscapulohumeral disease: Ph.D. thesis, Leiden University, The Netherlands, 1982.

Padberg G.W., Eriksson A.W., Volkers W.S., Bernini L., Van Loghem E., Meera Khan P., Nijenhuis L.E., Pronk J.C. and Schreuder G.M.Th. Linkage studies in autosomal dominant facioscapulohumeral muscular dystrophy. *J. Neurol. Sci.*, **65**, 261-268, 1984.

Padberg G.W., Klasen E.C., Volkers W.S., De Lange G.G. and Wintzen A.R. Linkage studies in facioscapulohumeral muscular dystrophy. *Muscle Nerve*, **11**, 833-835, 1988.

Padberg P.W., Lunt P.W., Koch M. and Fardeau M. Diagnostic criteria for facioscapulohumeral muscular dystrophy (workshop report). *Neuromusc. Disord.*, **1** (4), 231-234, 1991.

Padberg G.W., Brouwer O.F., De Keizer R.J.W., Dijkman G., Wijmenga C., Grote J.J. and Frants R.R. On the significance of retinal vascular disorder and hearing loss in facioscapulohumeral muscular dystrophy. *Muscle Nerve*, **Suppl 2**, S73-S80, 1995a.

Padberg G.W., Frants R.R., Brouwer O.F., Wijmenga C., Bakker E. and Sandkuijl L.A. Facioscapulohumeral muscular dystrophy in the Dutch population. *Muscle Nerve*, **Suppl 2**, S81-S84, 1995b.

Padberg G.W., Lunt P.W., Koch M. and Fardeau M. Facioscapulohumeral muscular dystrophy. In: Diagnostic criteria for neuromuscular disorders, 2nd edition. Emery A.E.H., editor. London, Royal Society of Medicine Press, 9-15, 1997.

Pauleikhoff D., Bornfeld N., Bird A.C. and Wessing A. Severe visual loss associated with retinal telangiectasis and facioscapulohumeral muscular dystrophy. *Graefe's Arch Clin. Exp. Ophthalmol.*, **230**, 362-365, 1992.

Ricci E., Galluzzi G., Deidda G., Cacurri s., Colantoni L., Merico B., Piazzo N., Servidei S., Bigneti E., Pasceri V., Silvestri G., Mirabella M., Mangiola F., Tonali P. and Felicetti L. Progress in the Molecular Diagnosis of Facioscapulohumeral muscular dystrophy and correlation between the number of KpnI repeats at the 4q35 locus and clinical phenotype. *Ann. Neurol.*, **45**, 751-757, 1999.

Saito A., Higuchi I., Nakagawa M., Saito M., Uchida Y., Inose M., Kasai T., Niiyama T., Fukunaga H., Arimura K. and Osame M. An overexpression of fibroblast growth factor (FGF) and FGF receptor 4 in a severe clinical phenotype of facioscapulohumeral muscular dystrophy. *Muscle Nerve*, **23**, 490-497, 2000.

Sarfarazi M., Upadhyaya M., Padberg G., Pericak-Vance M., Siddique T., Lucotte G. and Lunt P. An exclusion map for facioscapulohumeral (Landouzy-Déjérine) disorder. *J. Med. Genet.*, **26**, 481-484, 1989.

Sarfarazi M., Wijmenga C., Upadhyaya M., Weiffenbach B., Hyser C., Mathews K., Murray J., Golbert J., Pericak-Vance M., Lunt P., Frants R.R., Jacobsen S., Harper P.S. and Padberg G.W. Regional mapping of facioscapulohumeral muscular dystrophy gene on 4q35: combined analysis of an international consortium. *Am. J. Hum. Genet.*, **51**, 396-403, 1992.

Siddique T., Roper H., Pericak-Vance M.A., Shaw J., Warner K.L., Hung W-Y., Phillips K.L., Lunt P., Cumming W.J.K. and Roses A.D. Linkage analysis in the spinal muscular atrophy type of facioscapulohumeral disorder. *J. Med. Genet.*, **26**, 487-489, 1989.

Stevenson W.G., Perloff J.K., Weiss J.N. and Anderson T.L. Facioscapulohumeral muscular dystrophy: Evidence for selective genetic electrophysiologic cardiac involvement. *J. Am. Coll. Cardiol.*, **15**, 292-299, 1990.

Tawil R., Storvick D., Feasby T.E., Weiffenbach B. and Griggs R.C. Extreme variability of expression in monozygotic twins with FSH muscular dystrophy. *Neurology*, **43**, 345-348, 1993a.

Tawil R., Storvick D., Weiffenbach B., Altherr M.R., Feasby T.E. and Griggs R.C. Chromosome 4q DNA rearrangement in monozygotic twins discordant for facioscapulohumeral muscular dystrophy. *Hum. Mutat.*, **2**, 492-494, 1993b.

Tawil R., McDermott M.P., Mendell J.R., Kissel J., Griggs R.C. and the FSHD-DY group. Facioscapulohumeral muscular dystrophy (FSHD): Design of natural history study and results of baseline testing. *Neurology*, **44**, 442-446, 1994.

Tawil R., Forrester J. Griggs R.C., Mendell J., Kissel J., McDermott M., King W., Weiffenbach B. and Figlewicz D. Evidence for anticipation and association of deletion size with severity in facioscapulohumeral muscular dystrophy. *Ann. Neurol.*, **39**, 744-748, 1996.

Tawil R., McDermott M.P., Pandya S., King W., Kissel J., Mendell J.R., Griggs R.C. and the FSH-DY group. A pilot trial of prednisone in facioscapulohumeral muscular dystrophy. *Neurology*, **48**, 46-49, 1997.

Tawil R., Figlewicz D.A., Griggs R.C., Weiffenbach B. and the FSH Consortium. Facioscapulohumeral dystrophy: A distinct regional myopathy with a novel molecular pathogenesis. *Ann Neurol.,* **43**, 279-282, 1998.

Thein S.L. and Wallace R.B. The use of synthetic oligonucleotides as specific hybridization probes in the diagnosis of genetic disorders. In: Human genetic diseases: a practical approach. Davies K.E., editor. Oxford, IRL Press, 33-50, 1986.

Typler R., Berardinelli A., Barbierato L, Frants R., Hewitt J.E., Lanzi G., Maraschio P. and Tiepolo L. Monosomy of distal 4q does not cause facioscapulohumeral muscular dystrophy. *J. Med. Genet.*, **33**, 366-370, 1996.

Tupler R., Barbierato L., Memmi M., Sewry C.A., De Grandis D., Maraschio P., Tiepolo L. and Ferlini A. Identical de novo mutations at the D4F104S1 locus in monozygotic male twins affected by facioscapulohumeral muscular dystrophy (FSHD) with different clinical expression. *J. Med. Genet.*, **35**, 778-783, 1998.

Tyler F.H and Stephens F.E. Studies in disorders of muscle. II. Clinical manifestations and inheritance of facioscapulohumeral dystrophy in a large family. *Ann. Intern. Med.*, **32**, 640-660, 1950.

Upadhyaya M., Sarfarazi M., Lunt P.W., Broadhead W. and Harper P. A genetic linkage study of facioscapulohumeral (Landouzy-Déjérine) disease with 24 polymorphic probes. *J. Med. Genet.*, **26**, 490-493, 1989.

Upadhyaya M., Lunt P.W., Sarfarazi M., Broadhead W., Daniels J., Owen M. and Harper P.S. DNA marker applicable to presymptomatic and prenatal diagnosis of facioscapulohumeral disorder. *Lancet*, **336**, 1320-1321, 1990.

Upadhyaya M., Lunt P.W., Sarfarazi M., Broadhead W., Daniels J., Owen M. and Harper P.S. A closely linked DNA marker for facioscapulohumeral disease on chromosome 4q. *J. Med. Genet.*, **28**, 665-671, 1991.

Upadhyaya M., Lunt P., Sarfarazi M., Broadhead W., Farnham J. and Harper P.S. The mapping of chromosome 4q markers in relation to facioscapulohumeral muscular dystrophy (FSHD). *Am. J. Hum Genet.*, **51**, 404-410, 1992.

Upadhyaya M., Maynard J., Osborn M., Jardine P. and Harper P.S. Germinal mosaicism in facioscapulohumeral muscular dystrophy (FSHD). *Muscle Nerve*, **Suppl 2**, S45-S49, 1995.

Upadhyaya M., Maynard J., Rogers M.T., Lunt P.W., Jardine P., Ravine D. and Harper P.S. Improved molecular diagnosis of facioscapulohumeral muscular dystrophy (FSHD): validation of the differential double digestion of FHSD. *J. Med. Genet.*, **3**, 475-479, 1997.

Upadhyaya M., MacDonald M. and Ravine D. Prenatal diagnosis for facioscapulohumeral muscular dystrophy (FSHD). *Prenatal Diagnosis*, **19**, 959-965, 1999.

Van der Kooi A.J., Visser M.C., Rosenberg N., Van den Berg-Vos R., Wokke J.H.J., Bakker E. and De Visser M. Extension of the clinical range of facioscapulohumeral dystrophy: report of six cases. *J. Neurol. Neurosurg. Psychiatry*, **69**, 114-116, 2000.

Van der Maarel S.M., Deidda G., Lemmers R.J.L.F., Bakker E., Van der Wielen M.J.R., Sandkuijl L., Hewitt J.E., Padberg P.W. and Frants R.R. A new dosage test for subtelomeric 4;10 translocations improves conventional diagnosis of facioscapulohumeral muscular dystrophy (FSHD). *J. Med. Genet.*, **36**, 823-828, 1999.

Van der Maarel S.M., Deidda G., Lemmers R.J.L.F., Van Overveld P.G.M., Van der Wielen M., Hewitt J.E., Sandkuijl L., Bakker B., Van Ommen G-J.B., Padberg G.W. and Frants R.R. *De novo*

Facioscapulohumeral muscular dystrophy: Frequent somatic mosaicism, sex-dependent phenotype, and the role of mitotic transchromosomal repeat interaction between chromosomes 4 and 10. *Am. J. Hum. Genet.*, **66**, 26-35, 2000.

Van der Maarel S.M., Deidda G., Van Geel M., Rijkers T., Lemmers R.J.L.F., Van Deutekom J.C.T., Figlewicz D., Forrester J., Hewitt J.E., Padberg G.W. and Frants R.R. FRG2 is transcriptionally up-regulated in differentiating primary myoblast cultures of FSHD patients. Submitted for publication.

Van der Merwe A., Schutte C-M., Honey E., and Olckers A. Haplotype analysis of South African facioscapulohumeral muscular dystrophy families. FSHD International Consortium Research Meeting, San Francisco, U.S.A., 1999.

Van der Merwe A., Schutte C-M., Honey E., and Olckers A. Molecular investigation of facioscapulohumeral muscular dystrophy (FSHD) in South African families (abstract). *Neuromusc. Disord.*, **10**(4,5), 376, 2000a.

Van der Merwe A., Schutte C-M., Van der Maarel S.M., Honey E., Frants R.R. and A. Olckers. Molecular analysis of Facioscapulohumeral muscular dystrophy (FSHD) in the South African population, (abstract). *Am. J. Hum. Genet.*, **67**(4), 314, 2000b.

Van der Merwe A., Schutte C-M., Van der Maarel S.M., Honey E., Frants R.R. and Olckers A. Deletion fragment analysis in South African FSHD families (abstract). *Neuromusc. Disord.*, **11**(6,7), 633-634, 2001a.

Van der Merwe A., Schutte C-M., Van der Maarel S.M., Alessandrini M., Honey E., Frants R.R. and Olckers A. DNA rearrangements at the D4Z4 locus in South African facioscapulohumeral muscular dystrophy families (abstract). *Am. J. Hum. Genet.*, **69**(4), 637, 2001b.

Van Deutekom J.C., Wijmenga C., Van Tienhoven E.A., Gruter A.M., Hewit A.M., Padberg G.W., Van Ommen G.J., Hofker M.H. and Frants R.R. FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. *Hum. Mol. Genet.*, **2**, 2037-2042, 1993.

Van Deutekom J.C., Bakker E., Flemmers P.J., Van der Wielen M.J., Bik E., Hofker M.H., Padberg G.W. and Frants R.R. Evidence for subtelomeric exchange of 3.3 kb tandemly repeated units between chromosomes 4q35 and 10q26: implications for genetic counselling and etiology of FSHD1. *Hum. Mol. Genet.*, **5** (12), 1997-2003, 1996a.

Van Deutekom J.C.T., Lemmers R.J.L.F., Grewal P.K., Van Geel M., Romberg S., Dauwerse H.G., Wright T.J. Padberg G.W., Hofker M.H., Hewitt J.E. and Frants R.R. Identification of the first gene (FRG1) from the FSHD region on human chromosome 4q35. *Hum. Mol. Genet.*, 5 (5), 581-590, 1996b.

Van Geel M., Van Deutekom J.C.T., Van Staalduiden A., Lemmers R.J.L.F., Dickson M.C., Hofker M.H., Padberg G.W., Hewitt J.E., De Jong P.J. and Frants R.R. Identification of a novel β-tubulin subfamily with one member (TUBB4Q) located near the telomere of chromosome region 4q35. *Cytogenet. Cell Genet.*, **88**, 316-321, 2000.

Van Geel M. Facioscapulohumeral muscular dystrophy, a complex human disease entangled in dynamic molecular evolution: Ph.D thesis, Leiden University, The Netherlands, 2001.

Van Koningsbruggen S., Deidda G., Dirks R.W., Mommaas A.M., Onderwater J., Padberg G.W., Frants R.R. and van der Maarel S.M. The localization of FRG1P in the nucleolus, Cajal Bodies and speckles implies a fundamental role in RNA processing. *Am. J. Hum. Genet.*, **67** (4), 181, 2000.

Van Overveld P.G.M., Lemmers R.J.F.L, Deidda G., Sandkuijl L., Padberg G.W., Frants R.R. and Van der Maarel S.M. Interchromosomal repeat array interaction between chromosomes 4 and 10: a model for subtelomeric plasticity. *Hum. Mol. Genet.*, **9** (19), 2879-2884, 2000.

Vergotine J., Thiart R. and Kotze M.J. Clinical versus molecular diagnosis of heterozygous familial hypercholesterolaemia in the diverse South African population. *S. Afr. Med. J.*, **91** (12), 1053-1059, 2001.

Voit T., Lamprecht A., Lenard H.G. and Goebel H.H. Hearing loss in facioscapulohumeral dystrophy. *Eur. J. Pediatr.*, **145**, 280-285, 1986.

Walton J.N. and Nattrass F.J. On the classification, natural history and treatment of them myopathies. *Brain*, **77** (2), 169-231, 1954.

Walton J.N. On the inheritance of muscular dystrophy. *Ann. Hum. Genet.*, **20**, 1-13, 1955.

Walton J.N. The inheritance of muscular dystrophy: further observations. *Ann. Hum. Genet.*, **21**, 40-58, 1956.

Weber J.L. and May P.E. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.*, **44**, 388-396, 1989.

Weber J.L. and May P.E. Dinucleotide repeat polymorphism at the D4S171 locus. *Nucleic Acids Res.*, **18** (8), 2202, 1990.

Weiffenbach B., Bagley R., Falls K., Hyser C., Storvick D., Jacobsen S.J., Schultz P., Mendell J., Van Dijk K.W., Milner E.C.B. and Griggs R. Linkage analyses of five chromosome 4 markers localizes the facioscapulohumeral muscular dystrophy (FSHD) gene to distal 4q35. *Am. J. Hum. Genet.*, **51**, 416-423, 1992a.

Weiffenbach B., Dubois J., Storvick D., Tawil R., Jacobsen S.J., Gilbert J., Wijmenga C., Mendell J.R., Winokur S., Altherr M.R., Schultz P., Olandt S., Frants R.R., Pericak-Vance M. and Griggs R.C. Mapping the facioscapulohumeral muscular dystrophy gene is complicated by chromosome 4q35 recombination events. *Nat. Genet.*, **4**, 165-169, 1993.

Weiffenbach B., Dubois J., Manning S., Ma N.S., Schutte B.C., Winokur S.T., Altherr M.R., Jacobsen S.J., Stanton V.P., Yokoyama K. and Moir D. YAC contigs for 4q35 in the region of the facioscapulohumeral muscular dystrophy (FSHD) gene. *Genomics*, **19**, 532-541, 1994.

Weissenbach J., Gyapay G., Dib C., Vignal A., Morissette J., Millasseau P., Vaysseix G. and Lathrop M. A second-generation linkage map of the human genome. *Nature,* **359**, 794-801, 1992.

Wijmenga C., Frants R.R., Brouwer O.F., Moerer P., Weber J.L. and Padberg G. Location of facioscapulohumeral muscular dystrophy gene on chromosome 4. *Lancet.*, **336**, 651-653, 1990.

Wijmenga C., Padberg G.W., Moerer P., Wiegant J., Liem L., Brouwer O.F., Milner E.C.B., Weber J.L., Van Ommen G-J.B., Sandkuyl L.A. and Frants R.R. Mapping of facioscapulohumeral muscular dystrophy gene to chromosome 4q35-qter by multipoint linkage analysis and in situ hybridization. *Genomics*, **9**, 570-575, 1991.

Wijmenga C., Sandkuijl L.A., Moerer P., Van der Boorn N., Bodrug S.E., Ray P.N., Brouwer O.F., Murray J.C., Van Ommen G.J.B., Padberg G.W. and Frants R.R. Genetic linkage map of facioscapulohumeral muscular dystrophy of five polymorphic loci on chromosome 4q35-qter. *Am. J. Hum. Genet.*, **51**, 411-415, 1992a.

Wijmenga C., Hewitt J.E., Sandkuijl A., Clark L.N., Wright T.J., Dauwerse H.G., Gruter A.M., Hofker M.H., Moerer P., Williamson R., Van Ommen G.J.B., Padberg G.W. and Frants R.R. Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat. Gen.*, **2**, 26-30, 1992b.

Wijmenga C., Brouwer O.F., Padberg G.W. and Frants R.R. Transmission of de-novo mutation associated with facioscapulohumeral muscular dystrophy. *Lancet*, **340**, 985-986, 1992c.

Wijmenga C., Wright T.J., Baan M.J., Padberg G.W., Williamson R., Van Ommen G.J.B., Hewitt J.E., Hofker M.H. and Frants R.R. Physical mapping and YAC-cloning connects four genetically distinct 4qter loci (D4S163, D4S139, D4F35S1 and D4F104S1) in the FSHD gene-region. *Hum. Mol. Genet.*, **2** (10), 1667-1672, 1993a.

Wijmenga C., Frants R.R., Hewitt J.E., Van Deutekom J.C.T., Van Geel M, Wright T.J., Padberg G.W., Hofker M.H. and Van Ommen G-J.B. Molecular genetics of facioscapulohumeral muscular dystrophy. *Neuromusc. Disord.*, **3** (5/6), 487-491, 1993b.

Wijmenga C., Winokur S.T., Padberg G.W., Skraastad M.I., Altherr M.R., Wasmuth J.J., Murray J.C., Hofker M.H. and Frants R.R. The human skeletal muscle adenine nucleotide translocator gene maps to chromosome 4q35 in the region of the facioscapulohumeral muscular dystrophy locus. *Hum. Genet.*, **92**, 198-203, 1993c.

Winokur S.T., Schutte B., Weiffenbach B., Washington S.S., McElligott D., Chakravarti A., Wasmuth J.H. and Altherr M.R. A radiation hybrid map of 15 loci on the distal long arm of chromosome 4, the region containing the gene responsible for facioscapulohumeral muscular dystrophy (FSHD). *Am. J. Hum. Genet.*, **53**, 874-880, 1993.

Winokur S.T., Bengtsson U., Feddersen J., Mathews K.D., Weiffenbach B., Bailey H., Markovich R.P., Murray J.C., Wasmuth J.J., Altherr M.R. and Schutte B.C. The DNA rearrangement associated with facioscapulohumeral muscular dystrophy involves a heterochromatin-associated repetitive element: implications for a role of chromatin structure in the pathogenesis of the disorder. *Chromosome Res.*, **2**, 225-234, 1994.

Winokur S.T., Bengtsson U., Vargas J.C., Wasmuth J.J. and Altherr M.R. The evolutionary distribution and structural organization of the homeobox-containing repeat D4Z4 indicates a functional role for the ancestral copy in the FSHD region. *Hum. Mol. Genet.*, **5** (10), 1567-1575, 1996.

Wizard® Genomic DNA Purification kit, Technical Manual. Promega Corporation, Madison, WI, U.S.A., 1998.

Wright T.J., Wijmenga C., Clark L.N., Frants R.R., Williamson R. and Hewitt J.E. Fine mapping of the FSHD gene region orientates the rearranged fragment detected by the probe p13E-11. *Hum. Mol. Genet.*, **2** (10), 1673-1678, 1993.

Xia H., Winokur S.T., Kuo W-L., Altherr M.R. and Bredt D.S. Actinin-associated LIM Protein: Identification of a domain interaction between PDZ and spectrin-like repeat motifs. *J. Cell Biol.*, **139** (2), 507-515, 1997.

Xu X., Peng M., Fang Z. and Xu X. The direction of microsatellite mutations is dependent upon allele length. *Nat. Genet.*, **24**, 396-399, 2000.

Yamanaka G., Goto K., Matsumura T., Funakoshi M., Komori T., Hayashi Y.K. and Arahata K. Tongue atrophy in facioscapulohumeral muscular dystrophy. *Neurology*, **57**, 733-735, 2001.

Yasukohchi S., Yagi Y., Akabane T., Terauchi A., Tamagawa K. and Mizuna Y. Facioscapulohumeral dystrophy associated with sensorineural hearing loss, tortuosity of retinal arterioles, and an early onset and rapid progression of respiratory failure. *Brain Dev.*, **10**, 319-324, 1988.

Zatz M., Marie S.K., Passos-Bueno M.R., Vainzof M., Campiotto S., Cerqueira A., Wijmenga C., Padberg G. and Frants R.R. High proportion of new mutations and possible anticipation in Brazilian facioscapulohumeral muscular dystrophy families. *Am. J. Hum. Genet.*, **56**, 99-105, 1995.

Zatz M., Marie S.K., Cerqueira A., Vainzof M., Pavanello R.C.M. and Passos-Bueno M.R. The Facioscapulohumeral muscular dystrophy (FSHD1) gene affects males more severely and more frequently than females. *Am. J. Med. Genet.*, **77**, 155-161, 1998.

Internet references

CEPH, Centre dÉtude du Polymorphisme Humain (Centre for the Study of Human Polymorphisms), http://www.cephb.fr, 2001.

CHLC, The Co-operative Human Linkage Center, http://www.chlc.org, 2001.

GDB, The Genome Database, http://www.gdb.org, 2001.

MFD, Center for Medical Genetics, Marshfield Medical Research Foundation, http://research.marshfieldclinic.org/genetics, 2001.

NCBI, National Center for Biotechnology Information, National Institutes of Health, http://www.ncbi.nlm.nih.gov, 2001.

OMIM, Online Mendelian Inheritance in Man, Duchenne muscular dystrophy and Becker muscular dystrophy, 310200. http://www.ncbi.nlm.nih.gov, 2001a.

OMIM, Online Mendelian Inheritance in Man, Emery-Dreifuss muscular dystrophy, 310300, 181350. http://www.ncbi.nlm.nih.gov, 2001b.

OMIM, Online Mendelian Inheritance in Man, Limb-girdle muscular dystrophy, 159000, 159001, 253600, 253601, 253700, 254110, 600119, 601287, 601173, 601954, 603511, 604286. http://www.ncbi.nlm.nih.gov, 2001c.

OMIM, Online Mendelian Inheritance in Man, Distal myopathies, 254130, http://www.ncbi.nlm.nih.gov, 2001d.

OMIM, Online Mendelian Inheritance in Man, Oculopharyngeal muscular dystrophy, 164300, http://www.ncbi.nlm.nih.gov, 2001e.

OMIM, Online Mendelian Inheritance in Man, Facioscapulohumeral muscular dystrophy, 158900, http://www.ncbi.nlm.nih.gov, 2001f.

APPENDIX A

CONFERENCES AND MEETINGS AT WHICH RESEARCH CONTAINED IN THIS THESIS WERE PRESENTED

Results presented in this thesis were presented at the following national and international meetings. The presenting author's name is underlined in each case.

A.1 RESEARCH PRESENTED AT INTERNATIONAL CONFERENCES

A.1.1 **51th Annual meeting of the American Society of Human Genetics**: San Diego, U.S.A., October 2001.
Van der Merwe A., Schutte C-M., Van der Maarel S.M., Alessandrini M., Honey E., Frants R.R. and Olckers A. DNA rearrangements at the D4Z4 locus in South African facioscapulohumeral muscular dystrophy families (poster presentation).

A.1.2 **FSHD International Consortium Research Meeting**: Philadelphia, U.S.A., October 2001.
Olckers A., Van der Merwe A., Alessandrini M., Wallace D.C., Van der Maarel S.M., Honey E., Frants R.R. and Schutte C-M. FSHD in the South African population: Evidence to date suggest a dual founder effect.

A.1.3 **6th Congress of the World Muscle Society**: Salt Lake City, Utah, U.S.A., September 2001.
Van der Merwe A., Schutte C-M., Van der Maarel S.M., Honey E., Frants R.R. and Olckers A. Deletion fragment analysis in South African FSHD families (poster presentation).

A.1.4 **50th Annual meeting of the American Society of Human Genetics**: Philadelphia, U.S.A., October 2000.
Van der Merwe A., Schutte C-M., Van der Maarel S.M., Honey E., Frants R.R. and Olckers A. Molecular analysis of facioscapulohumeral muscular dystrophy (FSHD) in the South African population (poster presentation).

A.1.5 **FSHD International Consortium Research Meeting**: Philadelphia, U.S.A., October 2000.
Olckers A., Van der Merwe A., Van der Maarel S.M., Honey E., Frants R.R. and Schutte C-M. Segregation of DNA rearrangements at the D4Z4 locus in South African FSHD families.

A.1.6 **5th Congress of the World Muscle Society**: White River, South Africa, June 2000.
Van der Merwe A., Schutte C-M., Honey E. and Olckers A. Molecular investigation of Facioscapulohumeral muscular dystrophy in South African families (platform presentation).

A.1.7 **FSHD International Consortium Research Meeting**: San Francisco, U.S.A., October 1999.
Van der Merwe A., Schutte C-M., Honey E., and Olckers A. Haplotype analysis of South African facioscapulohumeral muscular dystrophy families.

A.2 RESEARCH PRESENTED AT NATIONAL CONFERENCES

A.2.1 **9th South African Society of Human Genetics Congress**: The Kruger National Park, South Africa, August 2001.
Van der Merwe A., Schutte C-M., Van der Maarel S.M., Honey E., Frants R.R. and Olckers A. Southern blot and haplotype analysis of selected facioscapulohumeral muscular dystrophy (FSHD) families (platform presentation).

A.2.2 **Annual congress of the Neurological association of South Africa**: Wild Coast Sun, KwaZulu-Natal, March 2001.
Van der Merwe A., Schutte C-M., Van der Maarel S.M., Honey E., Frants R.R. and Olckers A. Molecular analysis of D4Z4 rearrangements in South African Facioscapulohumeral muscular dystrophy (FSHD) families (platform presentation).

A.2.3 **Muscular Dystrophy Symposium of the Muscular Dystrophy Foundation of South Africa**: Strand, Cape Town, March 1999.
Van der Merwe A. and Olckers A. Facioscapulohumeral muscular dystrophy: the significance of a diagnostic service in South Africa (poster presentation).

A.3 RESEARCH PRESENTED AT THE FACULTY OF MEDICINE, UNIVERSITY OF PRETORIA

A.3.1 **Faculty Day 2001: Faculty of Medicine, University of Pretoria, Pretoria, South Africa**

Van der Merwe A., Schutte C-M., Honey E. and Olckers A. Facioscapulohumeral muscular dystrophy (FSHD) in the South African population: a molecular investigation. August 2000 (platform presentation).

A.4 PUBLISHED ABSTRACTS IN INTERNATIONAL PEER-REVIEWED JOURNALS

A.4.1 Van der Merwe A., Schutte C-M., Van der Maarel S.M., Alessandrini M., Honey E., Frants R.R. and Olckers A. DNA rearrangements at the D4Z4 locus in South African facioscapulohumeral muscular dystrophy families (abstract). *Am. J. Hum. Genet.,* **69**(4), 637, 2001.

A.4.2 Van der Merwe A., Schutte C-M., Van der Maarel S.M., Honey E., Frants R.R. and Olckers A. Deletion fragment analysis in South African FSHD families (abstract). *Neuromusc. Disord.,* **11**(6,7), 633-634, 2001.

A.4.3 Van der Merwe A., Schutte C-M., Van der Maarel S.M., Honey E., Frants R.R. and A. Olckers. Molecular analysis of Facioscapulohumeral muscular dystrophy (FSHD) in the South African population (abstract). *Am. J. Hum. Genet.*, **67**(4), 314, 2000.

A.4.4 Van der Merwe A., Schutte C-M., Honey E., and Olckers A. Molecular investigation of facioscapulohumeral muscular dystrophy (FSHD) in South African families (abstract). *Neuromusc. Disord.*, **10**(4,5), 376, 2000.

APPENDIX B

MODE OF INHERITANCE OF DIFFERENT TYPES OF MUSCULAR DYSTROPHIES

This appendix contains information on the mode of inheritance, gene locus, gene symbol and gene product for all the muscular dystrophies.

Table B.1: Mode of inheritance of different types of muscular dystrophies

Disorder	Inheritance	Gene locus	Gene symbol	Gene product
Muscular dystrophies:				
Duchenne/Becker	XR	Xp21.2	DMD	Dystrophin
Emery Dreifuss	XR	Xq28	EMD	Emerin
Emery Dreifuss	AD	1q11-q23	EDMD-AD (LMNA)	LaminA/C
Facioscapulohumeral	AD	4q35	FSHD	---
Limb-girdle muscular dystrophies (LGMD)				
LGMD, dominant	AD	5q31	LGMD1A	Myotilin
LGMD with cardiac involvement	AD	1q11-21	LGMD1B (LMNA)	LaminA/C
LGMD, dominant	AD	3p25	LGMD1C (CAV3)	Caveolin-3
Familial dilated cardiomyopathy with conduction defect and adult-onset LGMD	AD	6q23	LGMD1D	---
LGMD, dominant	AD	7q	LGMD1E	---
Vocal cord and pharyngeal weakness with autosomal dominant distal myopathy	AD	5q31	VPDMD (possibly allelic to LGMD1A)	---
LGMD, recessive	AR	15q15.1-q21.1	LGMD2A (CAPN3)	Calpain 3
LGMD, recessive	AR	2p13	LGMD2B (FER-1)	Dysferlin
LGMD, recessive	AR	13q12	LGMD2C (SGCG)	γ-sarcoglycan
LGMD, recessive	AR	17q12-q21.33	LGMD2D (SGCA)	α-sarcoglycan

continued ...

Table B.1: continued ...

Disorder	Inheritance	Gene locus	Gene symbol	Gene product
LGMD, recessive	AR	4q12	LGMD2E (SGCB)	β-sarcoglycan
LGMD, recessive	AR	5q33-q34	LGMD2F (SGCD)	δ-sarcoglycan
LGMD, recessive	AR	17q11-q12	LGMD2G (TCAP)	Telethonin
LGMD, Hutterite type	AR	9q31-q34.1	LGMD2H	---
LGMD, recessive	AR	19q13.3	LGMD2I	---
Distal myopathy:				
Autosomal dominant distal myopathy	AD	14	MPD1	---
Distal myopathy with rimmed vacuoles	AR	9p1-q1	DMRV	---
Hereditary inclusion body myopathy	AR	9p1-q1	HIBM	---
Miyoshi myopathy	AR	2p12-14	MM	Dysferlin
Tibial muscular dystrophy	AD	2q31	TMD	Titin?
Other myopathies:				
Autosomal dominant myopathy with proximal muscle weakness and early respiratory muscle involvement (Edström)	AD	2q24-31	MPRM1	---
Autosomal dominant myopathy with proximal muscle weakness and early respiratory muscle involvement	AD	2q21	MPRM2	---
Bethlem myopathy	AD	21q22.3	COL6A1	Collagen VI α1
Bethlem myopathy	AD	21q22.3	COL6A2	Collagen VI α2
Bethlem myopathy	AD	2q37	COL6A3	Collagen VI α3
Desmin-related myopathy	AD	11q22	DRM (CRYAB)	αB-crystallin
Desmin-related myopathy	AD	2q35	DES	Desmin
Epidermolysis bullosa simplex associated with late-onset muscular dystrophy	AR	8q24-qter	MD-EBS	Plectin
Myopathy with excessive autophagy	XR	Xq28	MEAX	---
Oculopharyngeal	AD	14q11.2-q13	OPMD (PABP2)	Poly A binding protein 2
Congenital myopathies:				
Central core	AD	19q13.1	CCD (RYR1)	Ryanodine receptor
Congenital myopathy with excess of thin myofilaments	AR	1q42,1	ACTA1	Actin alpha, skeletal muscle

continued ...

Table B.1: continued ...

Disorder	Inheritance	Gene locus	Gene symbol	Gene product
Congenital muscular dystrophy with merosin deficiency	AR	6q2	LAMA2	Laminin α2 chain of merosin
Congenital muscular dystrophy with secondary merosin deficiency	AR	1q42	CMD1B	---
Congenital muscular dystrophy with integrin deficiency	AR	12q13	ITGA7	Integrin α7
Congenital muscular dystrophy with rigid spine	AR	1p35-36	RSMD-1	---
Fukuyama CMD	AR	9q31-q33	FCMD	Fukutin
Myotubular myopathy	XR	Xq28	MTMX	Myotubularin
Nemaline myopathy	AD	1q21-q23	NEM1 (TPM3)	α tropomyosin
Nemaline myopathy	AR	2q21.1-q22	NEM2	Nebulin
Nemaline myopathy	AD	1q42.1	ACTA1	Actin alpha, skeletal muscle
Myotonic syndromes				
Brody disease	AR	16p12	SERCA1	Sarcoplasmic reti-culum Ca^{2+} ATPase
Myotonic dystrophy (Steinert)	AD	19q13	DM	Myotonin-protein kinase
Myotonic dystrophy type 2	AD	3q	DM2	---
Myotonia, dominant (Thomsen)	AD	7q35	CLC-1	Muscle chloride channel
Myotonia, recessive (Becker)	AR	7q35	CLC-1	Muscle chloride channel
Rippling muscle disease	AD	1q41	RMD	---
Schwartz-Jampel syndrome	AR	1p34-p36.1	SJS	---

XR = X-linked recessive; AR = autosomal recessive; AD = autosomal dominant, --- = no information available. (Adapted from Neuromuscular disorders: gene location, 2001).

APPENDIX C

DIAGNOSTIC CRITERIA FOR FACIOSCAPULOHUMERAL MUSCULAR DYSTROPHY

The following criteria define facioscapulohumeral muscular dystrophy (FSHD) on a clinical level:

1. Onset of the disorder is in the facial or shoulder girdle muscles; sparing of the extra-ocular, pharyngeal and lingual muscles and the myocardium.
2. Facial weakness in more than 50% of the affected family members.
3. Autosomal dominant inheritance in familial cases.
4. Evidence of a myopathic disorder in electromyography (EMG) and muscle biopsy in at least on affected member without biopsy features specific to alternative diagnoses.

The following clinical definitions apply:

i. **Non-penetrance** refers to an obligate gene carrier without symptoms or signs relating to the disorder.

ii. **Pre-symptomatic** indicated that an individual has no complaints (symptoms) related to the disorder, but has muscle atrophy and weakness upon physical examination.

iii. **Symptomatic** refers to patients with complaints and clinical symptoms related to the weakness and muscle atrophy of FSHD.

The International FSHD consortium has defined the clinical, genetic, laboratory criteria and the criteria for an individual to be included in linkage analysis for FSHD (Padberg *et al.,* 1991 and Padberg *et al.,* 1997). In Table C.1, in the left hand column, the inclusion criteria are indicated with an “I”, the exclusion criteria with an “E” and additional comments with a “C”.

Table C.1: Diagnostic criteria for FSHD

1.0	**CLINICAL CRITERIA**
1.1	**Onset**
I	Onset of the disorder is in the facial or shoulder-girdle muscles. Presenting symptoms usually relate to weakness or wasting of these muscles.
E	Onset in pelvic girdle muscles suggests alternative diagnoses; although subsequent pelvic girdle involvement is not uncommon in FSHD.
C	Clinically recognizable age at onset is very variable; age at symptomatic presentation is even more so. The mean age at recognizable onset is in the second decade. Onset before the age of 5 years, although rare in families, is not uncommon in the more severe proven new mutation cases, and does not exclude the diagnosis. Infantile or early childhood onset requires facial weakness to be present, since a clinical diagnosis cannot otherwise be reliably made.
1.2	**Facial**
I	Facial weakness affecting eye closure (orbicularis oculi) and peri-oral muscles (orbicularis oris) occurs in the vast majority of patients. In the absence of facial weakness, a diagnosis of FSHD can be accepted only if the majority of affected family members have facial weakness.
E	Extra-ocular, masticatory, pharyngeal and lingual muscle weakness is not part of the disorder.
C	Facial weakness may be very subtle and is sometimes noticeable by asymmetry of facial expression only. There is also some evidence that a dominant scapulohumeral presentation without facial weakness may be due to the same mutation mechanism at 4q35.
1.3	**Shoulders**
C	The scapular fixators are the muscles more prominently involved. Also the pectoralis major muscles will become affected early in most cases. The deltoid muscles remain unaffected for a long period of time and often have a particular pattern of atrophy, i.e. partial and proximal.
1.4	**Asymmetry**
I	Asymmetry of involvement in the shoulder girdle muscle is the rule, usually affecting the right side first.
C	Symmetrical weakness and atrophy at presentation is unusual and necessitates increased caution before accepting the diagnosis as FSHD. Asymmetrical involvement of the facial muscles occurs frequently.
C	NMR, ultrascan or CT-scan may be of help to detect asymmetry of muscle atrophy.
1.5	**Progression**
I	Progression is inevitable, albeit at a rate which is highly variable and in some cases virtually imperceptible.
E	Regression of symptoms and signs does not occur and would exclude the diagnosis.
C	The rate of progression and severity level reached tend to correlate inversely with age at onset.
C	Progression of the disorder usually includes involvement of abdominal and foot extensor muscles at an early stage; pelvic girdle weakness and upper arm weakness may occur at any time after the onset of shoulder girdle weakness.
C	Neck extensor, intrinsic hand and triceps surae muscle weakness is uncommon but can be observed occasionally within families and is not dependent on advance age or severe involvement.
1.6	**Severity**
C	At any age the disorder has a wide range of severity. Five aspect of note are:
	1) Overall, between 10-20% of cases have eventual requirement for a wheelchair. 2) Severity in recognised isolated new mutation cases tend to be greater than in large families. 3) Presymptomatic cases occur at any age and appear to comprise approximately 30% of all cases in large families. 4) Once symptomatic, the disorder is progressive in the majority of cases. The rate of progression is variable, although faster rates tend to be seen with earlier ages at onset. Rarely, there can be long periods of apparent arrest of progression.

continued ...

Table C.1: continued ...

	5) There is broad correlation in 4q35 cases between greater clinical severity and smaller residual DNA fragment size at D4F104S1; it is currently uncertain whether this may also be influence by possible generational anticipation.
C	There appears to be no difference in mean age at death between patients and their non-affected sibs.
1.7	**Contractures**
C	Contractures and pseudohypertrophy of muscles may be present.
E	Severe and diffuse contractures exclude the diagnosis of FSHD.
1.8	**Cardiac disorder**
E	Cardiomyopathy is not part of the disorder. When present it suggests and alternative diagnosis.
1.9	**Hearing loss**
C	Hearing loss is part of the disorder; it starts with high tone perceptive deafness and my progress to involve all frequencies. The severity of the hearing loss varies between subjects at any age, but tends to be progressive. It is recommended that the results of hearing assessments be documented for several affected members in each family.
1.10	**Retinal disorder**
C	A retinal vasculopathy with capillary telangiectasis, microaneurysms and capillary closure has been reported in some members of some FSHD families. At present it is unclear whether this is a specific association. It should not be used for diagnostic purposes.
1.11	**Mental retardation**
C	A few cases have been reported with mental retardation. It is recommended that investigation of any such case should include chromosome analysis, concentrating on the distal long arm of chromosome 4. However, no causally associated cytogenetic abnormalities have yet been recorded, and haploinsuffiency of the 4q35 region does not seem to cause FSHD.
2.0.	**GENETIC CRITERIA**
I	The pattern of inheritance in familial cases is autosomal dominant.
C	Sporadic cases occur, their frequency is unknown, but they are not rare. Only if both parents have been examined can a case be accepted as “sporadic”.
E	There is no substantiated evidence for recessive inheritance.
C	The mutation rate is unknown due to many uncertainties related to prevalence, penetrance and ascertainment.
C	Published estimates of prevalence remain approximations, the literature suggests widely variable regional differences. A prevalence of 1 in 20,000 in Europe appears a reasonable figure.
C	Penetrance is almost complete. Non-penetrance is estimated at less than 5% beyond the age of 20.
3.0.	**LABORATORY CRITERIA**
C	Serum creatine kinase (SCK) levels can be normal, but are often elevated, though rarely exceed five times the upper limit of normal. Persistently high CK values above this level warrant exclusion of other neuromuscular diagnoses.
C	EMG often shows short duration, low amplitude polyphasic potentials. Some neurogenic features such as high amplitude potentials and positive sharp waves are present occasionally, but do not characterise individual families. Motor and sensory nerve conduction velocities are normal.
E	Giant potentials are not a feature of the disorder.
C	Muscle biopsies may exhibit any of the standard myopathic criteria. In addition, small angular fibres are not uncommon and moth-eaten fibres are frequently found. An occasional small group of atrophic fibres may be observed, in which case another biopsy in the same patient or an affected sib is desirable. Cellular infiltrates are not uncommon in FSHD and can be extensive. Their significance is unknown. In these cases, either an autosomal dominant pattern of inheritance of a deleted DNA fragment at 4q35 is required to establish the diagnosis of FSHD.

continued ...

Table C.1: continued ...

4.0	**PHENOTYPIC-GENOTYPIC ANALYSIS**
I	Individuals who have been examined by a physician familiar with this disorder, and classified as affected according to the above criteria.
I	unaffected family members aged 20 yr and over, who have been examined as above.
I	unrelated spouses, whether or not examined.
E	any subject whose clinical status remains in dispute.
E	apparently unaffected individuals under age 20 year.
E	An unaffected individual with a CK level repeatedly above the normal range in the absence of a proven alternative explanation for this.
5.0	**RECOMMENDED INVESTIGATIONS IN AT LEAST ONE MEMBER OF EACH FAMILY INCLUDED IN LINKAGE STUDIES**
The following are recommended investigations:	
a)	Fully documented history and clinical examination,
b)	Serum creatine kinase,
c)	EMG,
d)	Muscle biopsy from an affected muscle for routine analysis,
e)	Audiometry, and
f)	Lymphoblast cell line and/or high molecular weight DNA sample suitable for pulsed field gel studies, and tested for persistence of DNA fragment of size 40 kb at locus D4F104S1, following double digestion of DNA with restriction enzymes *Eco* RI and *Bln* I.

APPENDIX D

NUCLEOTIDE SEQUENCE OF REPEAT UNITS AND FLANKING REGIONS AT THE D4Z4 LOCUS

The nucleotide sequence of the 3.3 kb repeat units and the flanking regions at the D4Z4 locus on chromosome 4q35 is presented in Table D.1.

Table D.1: Nucleotide sequence of repeat units at D4Z4 locus

Nucleotide position	Genomic DNA Sequence
1	gaattctatc tggtacccag agggaagggg gttcccagtg agggcaggac caggcttcat
61	gcacctcttc aggaatgttc tcctcatagt ccagcctcaa ggtgtgcatc ctctgtgtgc
121	atggagtcca tggcaggctc tgcctgggga gccgtccagc tgcacacctg caatgtggtg
181	gtgaccctca tgaatgggtg gttctgggcc ccatggctgg cagcagagag ggagatgttc
241	agccaccaag cccagagccc tgccacaggc ttctgtgagg cctccatctg ctctgggttc
301	ttgccctgag aggctgccct gaagtcaaac agaagcaggt gggcctctct tccagggctg
361	ctctctcccc cactgacagc tccctagagg gagactcaga cagcggggac agattcctca
421	ggcataagca ctggagttta ggctggccag ttcattccat acgcccacat gacatgacac
481	aaggcagagg ctgtgggaca aaggtattgc cttttcttct ggcatgagga atggcttagg
541	aagcagggga tggtggggct ggggttgagt gatgggctgt gggccacaag gagtgggtgg
601	gcgctgagaa agtgtcctgg ttgtctgtcc atagacgcag aatgagtggc atcccaggag
661	cctgtgaggg gctggcagag acttactggt tccagtaaaa gccccatgtg gatgcagtaa
721	tgctgcctgc tggtccttgg ctgtaattac aaacaggtac atgaggtacc catgcatctt
781	gaagctctca gggagtgggt tccagctgct catggtaggc acttttagtc actgaacatg
841	cttcaggcat gtccaagctt gattaagcca ggcatcttgc tgtgaggccc tccacttcac
901	taagaacact cttccttgct tcccctggaa gttggacctt ccagttctgg ttctggagac
961	acgatggccc ctcctggacc cctgggagaa tgtgctcagg tgacacacag ttgatggggc
1021	ccatttccaa gccattcttc catttcccac tgtttgaggg acccgaggcc ggtgacaagc
1081	acagagccac ccaaggccag ctgtctgcac ctaaatgtga tgcttgtctg gatgtctcag
1141	ggccagaacc ctccaggtga gatggcctgg tcctcaccac ctggcgtccg tgctcccttt
1201	tcctctgttc aatcctggcg ccaatgcctc cctcaactct caggtcacca ttggagaaga
1261	tgctcaggaa gaacaagcag ctgcagttaa ccctgctgaa agtggcagat gggtccaggc
1321	tcttgagctc gtcttggaca tggaacatgt ggatacaggc tttgagcagt gtgtgtagct

continued ...

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Table D.1: continued ...

Nucleotide position	Genomic DNA Sequence
1381	ctttcaggaa ggaagggaaa agggtgttac ccgggtccta caccctggaa cgacccttct
1441	cagacagtaa atagttggca gggtgcggtc atgtgtgatt ttagttttca actttaggct
1501	ttcattttca aattccacaa taaacacata aggtggagtt ctggtttcag cacacacaca
1561	cacacacaca aacacacaca cacacacaca cacagtctct ctctctgtat gtctctttct
1621	gtctctctct ctctccttcc tgcaaggatc cttgttaaca agaaaccttc tgccaaatgc
1681	ctctgaagca caggcaggtc ttggggagcc acaaggccac ttcctctttg tgcactagtg
1741	tcttgggtag gcatagcttt cagagctctg gggcctccac aaccttgccc tgctgtccag
1801	gggcagccct catgcagggg tgtcctaaga acttttcagg atgcacaagt tcagcactgt
1861	cttccaatgt gtgtttcacg atattttaat ggtggttctt ttgggaaaaa ggaaaggttc
1921	tgtgatcaat tatgggacac attgagctac agatcttttt cacaattgct cttaacaagc
1981	aggtagaccc tgagaacatg agtagcttcc ccgcaggtaa cttgagtgca tgagaacttt
2041	tgctttacaa ccatgccaat ctcacctcag cagttggcag tgctgcacgg ggcagacttc
2101	cctactcaaa ggctgtgaag cttttctttc ttttttttta aacattattt ttctttatag
2161	aattttgttg ggctgatatc aagcctggct tggtactgcc tcattttttt tggaatcaga
2221	acgctgttct ttaactcacg ggttgtgaag ttagaaggtg ctggtgtgac agcctgacaa
2281	gcagagcgca gctccaatcc caccttcatg ctctcatctg acgcagagcc ctcagagaag
2341	tggggaagtg cttcctggcc ctgcttctgg gggccgtccc caaggcagtc caccgaactt
2401	ccaaaacagc cttccctcac acacagccct gagccctcct gccgctcctc aatgttgcac
2461	atctctgaga agtggtccag catgttgctg tccaggggca gtgagaagca ggtgcggtga
2521	cacatgtctt cacggaccat gagcaccggg taaatctcct gcacaatctc cttgggggac
2581	accttgaggg agaaagcccc aacaactgat ggcatgccac atggcagaaa gcaaagactt
2641	acccttaccc cagcccaaag tcctgagaat catgccaaaa atccttggtt tcccactttt
2701	taaaaatttt aaaattaaaa tcccaggttc cgcgtataca tgccatgccc acctgcacct
2761	gtgtgtgtgt gtgtgtttgt gcacgcagga cagagcctgg cccattgact attcctgcag
2821	accaagaaaa atccctatgc agagtaaggg gagatggaag aaacgaggga gagaaaatgg
2881	cagccttgcc tcctcccttg cccagtgcta aggtccccag ggcaaatggc ttttgccttc
2941	aacttcacct taacaacata caaaatatat tcatttttac ttccgtcact ttcttaacat
3001	tacaaattgt atctttatat atgatttgta ttttcacaga gatttaagaa tttaatgcac
3061	cattatagta gaaaattgta tatctgtgta tatatttaca ttgaacagag agctttatat
3121	tttcatgtgg ttttatgatg ctgtccagca tcatttaatt tttcaacata attaactctc
3181	tttagcattt tttttcctag ggttattcta gtagttaaca acctcagctt ttttatttta
3241	atctttgaaa gtctttattt ttttctaatt tttgaaatac agtatttccc agatcaatta
3301	ttattggttg ctagtatttt ttctttcatc gctttgaaat ctggaaagtt cttagcatcc
3361	ccgctttttc tctgaaataa tgtttatgcc attttctccc tctattcttt ttaaaagact
3421	ctatctctga atgtattggt ctacttgatg gtgtccagta agtcttatat ttcacccgta
3481	attttcccat tctttaaaat attagtttcc aggactcaat atttgtgaat aatatatgtt
3541	caattttctt ttttctgctc cattgtttgc tgttgtgtct ctgtagtgaa tttataaaac
3601	tcagttatta tattcttcaa ctctatgatt tctgttgggt ttttaaaaat agtttttatc

continued ...

Table D.1: continued ...

Nucleotide position	Genomic DNA Sequence
3661	tctttgttga tattttgctc attcgttatt tttaaatttc actcagttgt ctctttctgt
3721	tatagttttg ctcactgaga atgcataaga tgattatttt aagttctcca tcagatatgc
3781	aaaaatcttt atttgttaaa attcagtttc tgaatattta tgtttttctt ccaatgggga
3841	atattttctg ccttctctgt gtgccttgtg attttttttt ttaaagagat ctgggatct
3901	atacagcact catcaaatct agcatttaaa gactggctca gtaaaggggg ataccgacag
3961	caatagtcca ggctatagat tctaggtgct tcacaaacac attcccaaat atattttctc
4021	tggacttcgc tgtgtttcca agttaaagag aatttttct caatgtattt tagattctat
4081	tgtatatttt cttccccagt tggctgtctg tggtattgca gtttcactag tgctgtagca
4141	aacactcatc tttcttctca gcagacacaa actgtcatct atatgactcc atcatgtcct
4201	tcagcactcc acatcaggag agaaagaatc tagtcattag acaatttatc aaaaagccaa
4261	acatttcaac acatattcta ctgttttaat tctctcctga aggagatact gggagttggg
4321	catttctca ttagcccagt tactgttctg ggtgaaaaaa taaactgcag tggacaggct
4381	gtaagccaga cttcatcaaa tttctgcacc aatgaaaaaa aaatttacaa gagaaaaaca
4441	aaaaacccta ttaaacgtca cggacaaggc cagagtttga atatactgtg gtcatctctg
4501	ctccagtgca aactgtttcc agaaagccta cttctatttt ccttgctgta acagaggaac
4561	atttcctgtc ttatgtttat tctactctgc aatcccctaa ggctttttct ctccctccca
4621	gaatcttaaa gtgcattcga actcacaggc aaaatcctcc cagaatcttg tgagaacata
4681	aatgatctga ctagtttggc attgcttttg gggatctggg aaaatctgtg cacacttctg
4741	gagacccttg tcatgccatt ttttataaat ctattgtgcc tcaagtcaga agtgtgtgag
4801	gggagatggg gagacattgg gatgcgcgcg cctggggctc tcccacaggg ggctttcgtg
4861	agccaggcag cgagggccgc ccccgcgctg cagcccagcc aggccgcgac ggcagagggg
4921	gtctcccaac ctgccccggc gcgcggggat ttcgcctacg ccgccccggc tcctccggac
4981	ggggcgctct cccaccctca ggctcctcgg tggcctccgc acccgggcaa aagccgggag
5041	gaccgggacc cgcagcgcga cggcctgccg ggcccctgcg cggtggcaca gcctgggccc
5101	gctcaagcgg ggccgcaggg ccaaggggtg cttgcgccac ccacgtccca ggggagtccg
5161	tggtggggct ggggccgggt ccccaggtcg ccggggcggc gtgggaaccc caagccgggg
5221	caagcttcca cctccccagc ccgcgccccc ggacgcctcc gcctccgcgc ggcaggggca
5281	gatgcaaggc atcccggcgc cctcccaggc gctccaggag ccggcgccct ggtctgcact
5341	cccctgcggc ctgctgctgg atgagctcct ggcgagcccg gagtttctgc agcaggcgca
5401	acctctccta gaaacggagg ccccggggga gctggaggcc tcggaagagg ccgcctcgct
5461	ggaagcaccc ctcagcgagg aagaataccg ggctctgctg gaggagcttt aggacgcggg
5521	gttgggacgg ggtcgggtgg ttcggggcag ggcggtggcc tctctttcgc ggggaacacc
5581	tggctggcta cggaggggcg tgtctccgcc ccgcccctc caccgggctg accggcctgg
5641	gattcctgcc ttctaggtct aggcccggtg agagactcca caccgcggag aactgccatt
5701	ctttcctggg catcccgggg atcccagagc cggcccaggt accagcaggt gggccgccta
5761	ctgcgcacgc gcgggtttgc gggcagccgc ctgggctgtg ggagcagccc gggcagagct
5821	ctcctgcctc tccaccagcc caccccgccg cctgaccgcc ccctccccac cccccacccc
5881	ccacccccgg aaaacgcgtc gtcccctggg ctgggtggag acccccgtcc cgcgaaacac

continued ...

Table D.1: continued ...

Nucleotide position	Genomic DNA Sequence
5941	cgggccccgc gcagcgtccg ggcctgacac cgctccggcg gctcgcctcc tatgcgcccc
6001	cgcgccaccg tcgcccgccc gcccgggccc ctgcagccgc ccaggtgcca gcacggagcg
6061	cctggcggcg gaacgcagac cccaggcccg gcgcacaccg gggacgctga gcgttccagg
6121	cgggagggaa ggcgggcaga gatggagaga ggaacgggag acctagaggg gcggaaggac
6181	gggcggaggg acgttaggag ggagggaggg aggcagggag gcagggagga acggagggaa
6241	agacagagcg acgcagggac tgggggcggg cgggagggag ccggggaacg gggggaggaa
6301	ggcagggagg aaaagcggtc ctcggcctcc gggagtagcg ggaccccgc cctccgggaa
6361	aacggtcagc gtccggcgcg ggctgagggc tgggcccaca gccgccgcgc cggccggcgg
6421	ggcaccaccc attcgccccg gttccgtggc ccagggagtg ggcggtttcc tccgggacaa
6481	aagaccggga ctcgggttgc cgtcgggtct tcacccgcgc ggttcacaga ccgcacatcc
6541	ccaggctgag ccctgcaacg cggcgcgagg ccgacagccc cggccacgga ggagccacac
6601	gcaggacgac ggaggcgtga ttttggtttc cgcgtggctt tgccctccgc aaggcggcct
6661	gttgctcacg tctctccggc ccccgaaagg ctggccatgc cgactgtttg ctcccggagc
6721	tctgcgggca cccggaaaca tgcagggaag ggtgcaagcc cggcacggtg ccttcgctct
6781	ccttgccagg ttccaaaccg gccacactgc agactcccca cgttgccgca cgcgggaatc
6841	catcgtcagg ccatcacgcc ggggaggcat ctcctctctg gggtctcgct ctggtcttct
6901	acgtggaaat gaacgagagc cacacgcctg cgtgtgcgag accgtcccgg caacggcgac
6961	gcccacaggc attgcctcct tcacggagag agggcctggc acactcaaga ctcccacgga
7021	ggttcagttc cacactcccc tccaccctcc caggctggtt tctccctgct gccgacgcgt
7081	gggagcccag agagcggctt cccgttcccg cgggatccct ggagaggtcc ggagagccgg
7141	cccccgaaac gcgcccccct ccccctccc ccctctcccc cttcctcttc gtctctccgg
7201	ccccaccacc accaccgcca ccacgccctc cccccccacc cccccccccc accaccacca
7261	ccaccacccc gccggccggc cccaggcctc gacgccctgg ggtcccttcc ggggtggggc
7321	gggctgtccc agggggctc accgccattc atgaaggggt ggagcctgcc tgcctgtggg
7381	cctttacaag ggcggctggc tggctggccg gctgtccggg caggcctcct ggctgcacct
7441	gccgcagtgc acagtccggc tgaggtgcac gggagcccgc cggcctctct ctgcccgcgt
7501	ccgtccgtga aattccggcc ggggctcacc gcgatggccc tcccgacacc ctcggacagc
7561	accctccccg cggaagcccg gggacgagga cggcgacgga gactcgtttg gaccccgagc
7621	caaagcgagg ccctgcgagc ctgctttgag cggaacccgt acccgggcat cgccaccaga
7681	gaacggctgg cccaggccat cggcattccg gagcccaggg tccagatttg gtttcagaat
7741	gagaggtcac gccagctgag gcagcaccgg cgggaatctc ggccctggcc cgggagacgc
7801	ggcccgccag aaggccggcg aaagcggacc gccgtcaccg gatcccagac cgccctgctc
7861	ctccgagcct ttgagaagga tcgctttcca ggcatcgccg cccgggagga gctggccaga
7921	gagacgggcc tcccggagtc caggattcag atctggtttc agaatcgaag ggccaggcac
7981	ccgggacagg gtggcagggc gcccgcgcag gcaggcggcc tgtgcagcgc ggcccccggc
8041	gggggtcacc ctgctccctc gtgggtcgcc ttcgcccaca ccggcgcgtg gggaacgggg
8101	cttcccgcac cccacgtgcc ctgcgcgcct ggggctctcc cacaggggc tttcgtgagc
8161	caggcagcga gggccgcccc cgcgctgcag cccagccagg ccgcgccggc agaggggatc

continued ...

Table D.1: continued ...

Nucleotide position	Genomic DNA Sequence
8221	tcccaacctg ccccggcgcg cggggatttc gcctacgccg ccccggctcc tccggacggg
8281	gcgctctccc accctcaggc tcctcggtgg cctccgcacc cgggcaaaag ccgggaggac
8341	cgggacccgc agcgcgacgg cctgccgggc ccctgcgcgg tggcacagcc tgggcccgct
8401	caagcggggc cgcagggcca aggggtgctt gcgccaccca cgtcccaggg gagtccgtgg
8461	tggggctggg gccgggtcc ccaggtcgcc ggggcggcgt gggaacccca agccggggca
8521	gctccacctc cccagcccgc gcccccggac gcctccgcct ccgcgcggca ggggcagatg
8581	caaggcatcc cggcgccctc ccaggcgctc caggagccgg cgccctggtc tgcactcccc
8641	tgcggcctgc tgctggatga gctcctggcg agcccggagt ttctgcagca ggcgcaacct
8701	ctcctagaaa cggaggcccc gggggagctg gaggcctcgg aagaggccgc ctcgctggaa
8761	gcacccctca gcgaggaaga ataccgggct ctgctggagg agctttagga cgcggggttg
8821	ggacggggtc gggtggttcg gggcagggcg gtggcctctc tttcgcgggg aacacctggc
8881	tggctacgga ggggcgtgtc tccgccccgc cccctccacc gggctgaccg gcctgggatt
8941	cctgccttct aggtctaggc ccggtgagag actccacacc gcggagaact gccattcttt
9001	cctgggcatc ccggggatcc cagagccggc ccaggtacca gcaggtgggc cgcctactgc
9061	gcacgcgcgg gtttgcgggc agccgcctgg gctgtgggag cagcccgggc agagctctcc
9121	tgcctctcca ccagcccacc ccgccgcctg accgccccct ccccaccccc cacccccac
9181	ccccggaaaa cgcgtcgtcc cctgggctgg gtggagaccc ccgtcccgcg aaacaccggg
9241	ccccgcgcag cgtccgggcc tgacaccgct ccggcggctc gcctcctatg cgcccccgcg
9301	ccaccgtcgc ccgcccgccc gggcccctgc agccgcccag gtgccagcac ggagcgcctg
9361	gcggcggaac gcagacccca ggcccggcgc acaccgggga cgctgagcgt tccaggcggg
9421	agggaaggcg ggcagagatg gagagaggaa cgggagacct agaggggcgg aaggacgggc
9481	ggagggacgt taggagggag ggagggaggc agggaggcag ggaggaacgg agggaaagac
9541	agagcgacgc agggactggg ggcgggcggg agggagccgg ggaacggggg gaggaaggca
9601	gggaggaaaa gcggtcctcg gcctccggga gtagcgggac ccccgccctc cgggaaaacg
9661	gtcagcgtcc ggcgcgggct gagggctggg cccacagccg ccgcgccggc cggcggggca
9721	ccacccattc gccccggttc cgtggcccag ggagtgggcg gtttcctccg ggacaaaaga
9781	ccgggactcg ggttgccgtc gggtcttcac ccgcgcggtt cacagaccgc acatccccag
9841	gctgagccct gcaacgcggc gcgaggccga cagccccggc cacggaggag ccacacgcag
9901	gacgacggag gcgtgatttt ggtttccgcg tggctttgcc ctccgcaagg cggcctgttg
9961	ctcacgtctc tccggccccc gaaaggctgg ccatgccgac tgtttgctcc cggagctctg
10021	cgggcacccg gaaacatgca gggaagggtg caagcccggc acggtgcctt cgctctcctt
10081	gccaggttcc aaaccggcca cactgcagac tccccacgtt gccgcacgcg ggaatccatc
10141	gtcaggccat cacgccgggg aggcatctcc tctctggggt ctcgctctgg tcttctacgt
10321	cagttccaca ctcccctcca ccctcccagg ctggtttctc cctgctgccg acgcgtggga
10381	gcccagagag cggcttcccg ttcccgcggg atccctggag aggtccggag agccggcccc
10441	cgaaacgcgc ccccctcccc cctccccccct ctccccccttc ctcttcgtct ctccggcccc
10501	accaccacca ccgccaccac gccctccccc cccccccccc ccccccacca ccaccaccac

continued ...

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Table D.1: continued ...

Nucleotide position	Genomic DNA Sequence
10201	ggaaatgaac gagagccaca cgcctgcgtg tgcgagaccg tcccggcaac ggcgacgccc
10261	acaggcattg cctccttcac ggagagaggg cctggcacac tcaagactcc cacggaggtt
10561	caccccgccg gccggcccca ggcctcgacg ccctgggtcc cttcggggt gggcgggct
10621	gtcccagggg ggctcaccgc cattcatgaa ggggtggagc ctgcctgcct gtgggccttt
10681	acaagggcgg ctggctggct ggctggctgt ccgggcaggc ctcctggctg cacctgccgc
10741	agtgcacagt ccggctgagg tgcacgggag cccgccggcc tctctctgcc cgcgtccgtc
10801	cgtgaaattc cggccggggc tcaccgcgat ggccctcccg acaccctcgg acagcaccct
10861	ccccgcggaa gcccggggac gaggacggcg acggagactc gtttggaccc cgagccaaag
10921	cgaggccctg cgagcctgct ttgagcggaa cccgtacccg ggcatcgcca ccagagaacg
10981	gctggcccag gccatcggca ttccggagcc cagggtccag atttggtttc agaatgagag
11041	gtcacgccag ctgaggcagc accggcggga atctcggccc tggcccggga gacgcggccc
11101	gccagaaggc cggcgaaagc ggaccgccgt caccggatcc cagaccgccc tgctcctccg
11161	agcctttgag aaggatcgct ttccaggcat cgccgcccgg gaggagctgg ccagagagac
11221	gggcctcccg gagtccagga ttcagatctg gtttcagaat cgaaggccca ggcacccggg
11281	acagggtggc agggcgcccg cgcaggcagg cggcctgtgc agcgcggccc ccggcggggg
11341	tcaccctgct ccctcgtggg tcgccttcgc ccacaccggc gcgtggggaa cggggcttcc
11401	cgcaccccac gtgccctgcg cgcctggggc tctcccacag gggctttcg tgagccaggc
11461	agcgagggcc gccccgcgc tgcagcccag ccaggccgcg ccggcagagg ggatctccca
11521	acctgccccg gcgcgcgggg atttcgccta cgccgccccg gctcctccgg acggggcgct
11581	ctcccaccct caggctcctc gctggcctcc gcacccgggc aaaagccggg aggaccggga
11641	cccgcagcgc gacggcctgc cgggcccctg cgcggtggca cagcctgggc ccgctcaagc
11701	ggggccgcag ggccaagggg tgcttgcgcc acccacgtcc caggggagtc cgtggtgggg
11761	ctggggccgg ggtccccagg tcgccggggc ggcgtgggaa ccccaagccg gggcagctcc
11821	acctccccag cccgcgcccc cggacgcctc cgcctccgcg cggcaggggc agatgcaagg
11881	catcccggcg ccctcccagg cgctccagga gccggcgccc tggtctgcac tcccctgcgg
11941	cctgctgctg gatgagctcc tggcgagccc ggagtttctg cagcaggcgc aacctctcct
12001	agaaacggag gccccggggg agctggaggc ctcggaagag gccgcctcgc tggaagcacc
12061	cctcagcgag gaagaatacc gggctctgct ggaggagctt taggacgcgg ggttgggacg
12121	gggtcgggtg gttcggggca gggccgtggc ctctctttcg cggggaacac ctggctggct
12181	acggaggggc gtgtctccgc cccgccccct ccaccgggct gaccggcctg ggattcctgc
12241	cttctaggtc taggcccggt gagagactcc acaccgcgga gaactgccat tctttcctgg
12301	gcatcccggg gatcccagag ccggcccagg tacctgcgca cgcgcgggtt tgcgggcagc
12361	cgcctgggct gtgggagcag cccgggcaga gctctcctgc ctctccacca gcccaccccg
12541	ctccgctccg gcggctcgcc tcctgtgtgc ccccgcgcca ccgtcgcccg cccgcccggg
12601	cccctgcagc ctcccagctg ccagcgcgga gctcctggcg gtcaaaagca tacctctgtc
12661	tgtctttgcc cgcttcctgg ctagacctgc gcgcagtgcg caccccggct gacgtgcaag
12721	ggagctcgct ggcctctctg tgcccttgtt cttccgtgaa attctggctg aatgtctccc
12421	ccgcctgacc gcccccctccc caccccccac cccccacccc cggaaaacgc gtcgtcccct

continued ...

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Table D.1: continued ...

Nucleotide position	Genomic DNA Sequence
12481	gggctgggtg gagacccccg tcccgcgaaa caccgggccc cgcgcagcgt ccgggcctga
12781	cccaccttcc gacgctgtct aggcaaacct ggattagagt tacatctcct ggatgattag
12841	ttcagagata tattaaaatg ccccctccct gtggatccta tagaagattt gcatcttttg
12901	tgtgatgagt gcagagatat gtcacaatat cccctgtaga aaaagcctga aattggttta
12961	cataacttcg gtgatcagtg cagatgtgtt tcagaactcc atagtagact gaacctagag
13021	aatggttaca tcacttaggt gatcagtgta gagatatgtt aaaattctcg tgtagacaga
13081	gcctagacaa ttgttacatc acctagtgat cagtgcaggg ataagtcata aagcctcctg
13141	taggcagagt gtaggcaagt gttccctccc tgggctgatc agtgcagaga tatctcacaa
13201	agcccctata agccaaacct tgacaagggt tacatcacct gtttgagcag tggaaatata
13261	tatcacaaag ccccctgtag acaaagccca gacaattttt acatctcctg agtgagcatt
13321	ggagagatct gtcacaatgc ccctgtaggc agagcttaga caagtgttac atcacctggg
13381	tgatcagtgc agagatatgt caaaacgctc ctgtagtctg aacctagaca ggagttacat
13441	caccttgggg atcagtgcag agatacgtga gaattcc

a) DNA sequence was retrieved from Genbank with accession number AF117653 as reported by Gabriels *et al.*, (1999). b) The sequence of p13E-11 is indicated by thick underlining (nt. 3896-4715). c) The start of the 3.3 kb repeat units are indicated by arrows (nt. 5738-9052 and 9034-12329). d) The Lsau and hhspm3 repeats are shown by single (nt. 5719-6036) and dashed underlining (nt. 7055-7523) respectively. e) The double homeodomains are indicated by double underlining (nt. 7588-7767, 7813-7992 and 10883-11062, 11108-11287). f) The black box represents the GC-signal (nt. 7311-7319 and 10607-10615). g) The TATA signal (nt. 7385-7389 and 10681-10685) is represented by the blue box. h) The putative DUX4 gene is represented by the green dot-dash underlining (nt. 7534-8808 and 10829-12103). i) The restriction enzymes are indicated by: gaattc for *Eco* RI; cctagg for *Bln* I; agatct for *Bgl* II; ggtacc for *Kpn* I; and aaattc for *Xap* I.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

APPENDIX E

COMPARISON OF NUCLEOTIDE SEQUENCE FROM ONE *Kpn* I REPEAT UNIT DERIVED FROM CHROMOSOMES 4q35 AND 10q26

Sequences derived from the homologous *Kpn* I repeat units on chromosome 4q35 and chromosome 10q26 were compared and are presented in Table E.1.

Table E.1: Comparison of one *Kpn*I repeat unit nucleotide sequence derived from 4q35 and 10q26

Chr 4:	agatctgggg atctatacag cactcatcaa atctagcatt taaagactgg ctcagtaaag
Chr 10:	agatctgggg atctatacag cactcatcaa atctagcatt taaagactgg ctcagtaaag
	Bgl II p13E-11
	ggggataccg acagcaatag tccaggctat agattctagg tgcttcacaa acacattccc
	ggggataccg acagcaatag tccaggctat agattctagg tgcttcacaa acacattccc
	aaatatattt tctctggact tcgctgtgtt tccaagttaa agagaatttt ttctcaatgt
	aaatatattt tctctggact tgggtgtgtt tccaagttaa agagaatttt ttctcaatgt
	attttagatt ctattgtata ttttcttccc cagttggctg tctgtggtat tgcagtttca
	attttagatt ctattgtcta ttttcttccc cagttggctg tctgtggtat tgcagtttca
	ctagtgctgt agcaaacact catctttctt ctcagcagac acaaactgtc atctatatga
	ctagtgctgt agcaaacact catctttctt ctcagcaggc acaaactctc atctatatga
	ctccatcatg tccttcagca ctccacatca ggagagaaag aatctagtca ttagacaatt
	ctccatcatg tccttcagca ctccacatca ggagagaaag aatctagtca ttagacaatt
	tatcaaaaag ccaaacattt caacacatat tctactgttt taattctctc ctgaaggaga
	tatcaaaaag ccaaacattt caacacatat tctactgttt taattctctc ctgaaggaga
	tactgggagt tgggcatttt ctcattagcc cagttactgt tctgggtgaa aaaataaact
	tactgggagt tgggcatttt ctgattagcc cagttactgt tctgggtgaa aaaagaaact
	gcagtggaca ggctgtaagc cagacttcat caaatttctg caccaatgaa aaaaaaattt
	gcagtggaca ggctgtaagc cagacttcat caaatttctg cagcaatgaa aaaaaaattt
	acaagagaaa aacaaaaaac cctattaaac gtcacggaca aggccagagt ttgaatatac
	acaagagaaa aacaaaaaac cctattaaac gtcacggaca aggccagagt ttgaatatac
	tgtggtcatc tctgctccag tgcaaactgt ttccagaaag cctacttcta ttttccttgc
	tgtggtcatc tccgctccag tgcaaactgt tttcagaaag cctgctttta ttttccttgc
	tgtaacagag gaacatttcc tgtcttatgt ttattctact ctgcaatccc ctaaggcttt
	tgtaacagag gaacatttcc tgtcttatgc ttattctact ctgcagtccc ctaaggcttt

continued ...

Table E.1: continued ...

Chr 4:	ttctctccct cccagaatct taaagtgcat tcgaactcac aggcaaaatc ctcccagaat
Chr 10:	ttctctccct cccagaatct taaagtgcat tcgaacgcac aggcaaaatc ctcccagaaa
	cttgtgagaa cataaatgat ctgactagtt tggcattgct tttggggatc tgggaaaatc
	cttgtgaaaa cataaatgat ctgactagtt tggcattgct tttggggatc tgggaaaatc
	tgtgcacact tctggagacc cttgtcatgc catttttat aaatctattg tgcctcaagt
	tgtgcacact tctggagacc cttgtcaggc catttttat aaatctattg tgcctcaagt
	cagaagtgtg tgaggggaga tggggagaca ttgggatgcg cgcgcctggg gctctcccac
	cagcagtgtg tgaggggaga tggggagaca ttgggatgcg cgcgcctggg gctctcccac
	agggggcttt cgtgagccag gcagcgaggg ccgcccccgc gctgcagccc agccaggccg
	agggggcttt cgtgagccag gcagcgaggg ccgcccccgc gctgcagccc agccaggccg
	cgacggcaga gggggtctcc caacctgccc cggcgcgcgg ggatttcgcc tacgccgccc
	cgccggcaga ggggatctcc caacctgccc cggcgcgcgg ggatttcgcc tacgccgccc
	cggctcctcc ggacggggcg ctctcccacc ctcaggctcc tcggtggcct ccgcacccgg
	cggctcctcc ggacggggcg ctctcccacc ctcaggctcc tcggtggcct ccgcacccgg
	gcaaaagccg ggaggaccgg gacccgcagc gcgacggcct gccgggcccc tgcgcggtgg
	gcaaaagccg ggaggaccgg gacccgcagc gcgacggcct gccgggcccc tgcgcggtgg
	cacagcctgg gcccgctcaa gcggggccgc agggccaagg ggtgcttgcg ccacccacgt
	cacagcctgg gcccgctcaa gcggggccgc agggccaagg ggtgcttgcg ccacccacgt
	cccaggggag tccgtggtgg ggctggggcc ggg*tcccca ggtcgccggg gcggcgtggg
	cccaggggag tccgtggtgg ggctggggcc ggggtcccca ggtcgccggg gcggcgtggg
	aaccccaagc cggggcaagc ttccacctcc ccagcccgcg cccccggacg cctccgcgcg
	aaccccaagc cggggca*gc t*ccacctcc ccagcccgcg cccccggacg cctccgcgcg
	gcaggggcag atgcaaggca tcccggcgcc ctcccaggcg ctccaggagc cggcgccctg
	gcaggggcag atgcaaggca tcccggcgcc ctcccaggcg ctccaggagc cggcgccctg
	gtctgcactc ccctgcggcc tgctgctgga tgagctcctg gcgagcccgg agtttctgca
	gtctgcactc ccctgcggcc tgctgctgga tgagctcctg gcgagcccgg agtttctgca
	gcaggcgcaa cctctcctag aaacggaggc cccgggggag ctggaggcct cggaagaggc
	gcaggcgcaa cctctcctag aaacggaggc cccgggggag ctggaggcct cggaagaggc
	cgcctcgctg gaagcacccc tcagcgagga agaataccgg gctctgctgg aggagcttta
	cgcctcgctg gaagcacccc tcagcgagga agaataccgg gctctgctgg aggagcttta
	ggacgcgggg ttgggacggg gtcgggtggt tcggggcagg gcggtggcct ctctttcgcg
	ggacgcgggg ttgggacggg gtcgggtggt tcggggcagg gcggtggcct ctctttcgcg
	gggaacacct ggctggctac ggaggggcgt gtctccgccc cgccccctcc accgggctga
	gggaacgcct ggctggctac ggaggggcgt gtctccgccc cgccccctcc accgggctga
	ccggcctggg attcctgcct tctaggtcta ggcccggtga gagactccac accgcggaga
	ccggcctggg attcctgcct tctaggccta ggcccggtga gagactccac acagcggaga
	Bln I
	actgccattc tttcctgggc atcccgggga tcccagagcc ggcccaggta ccagcaggtg
	actgccattc tttcctgggc atcccgggga tcccagagcc ggcccaggta ccagcaggtg
	Kpn I
	ggccgcctac tgcgcacgcg cgggtttgcg ggcagccgcc tgggctgtgg gagcagcccg
	ggccgcctac tgcgcacgcg cgggtttgcg ggcagccgcc tgggctgtgg gagcagcccg

continued ...

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Table E.1: continued ...

Chr 4: Chr 10:	ggcagagctc tcctgcctct ccaccagccc accccgccgc ctgaccgccc cctccccacc ggcagagctc tcctgcctct ccaccagccc accccgccgc ctgaccgccc cctccccacc
	ccccaccccc cacccccgga aaacgcgtcg tcccctgggc tgggtggaga cccccgtccc ccc*accccc cgcccccgga aaacgcgtcg tcccctgggc tgggtggaga cccccgtccc
	gcgaaacacc gggccccgcg cagcgtccgg gcctgacacc gctccggcgg ctcgcctcct gcgaaacacc gggccccgcg cagcgtccgg gcctgacacc gctccggcgg ctcgcctcct
	atgcgccccc gcgccaccgt cgcccgcccg cccgggcccc tgcagccgcc caggtgccag ctgcgccccc gcgccaccgt cgcccgcccg cccgggcccc tgcagccgcc caggtgccag
	cacggagcgc ctggcggcgg aacgcagacc ccaggcccgg cgcacaccgg ggacgctgag cacggagcgc ctggcggcgg aacgcagacc ccaggcccgg cgcacaccgg ggacgctgag
	cgttccaggc gggagggaag gcgggcagag atggagagag gaacgggaga cctagagggg cgttccaggc gggagggaag gcgggcagag atggagagag gaacgggaga cctagagggg
	cggaaggacg ggcggaggga cgttaggagg gagggaggga ggcagggagg cagggagg** cggaaggatg ggcggaggga cgttaggagg gagggaggga ggcagggagg cagggaggca
	******aacg gagggaaaga cagagcgacg cagggactgg gggcgggcgg gagggagccg gggaggaacg gagggaaaga cagagcgacg cagggactgg gggcgggcgg gagggagccg
	gggaacgggg ggaggaaggc agggaggaaa agcggtcctc ggcctccggg agtagcggga gggа*cgggg ggaggaaggc agggaggaaa agcggtcttc ggcctccggg agtagcggga
	cccccgccct ccgggaaaac ggtcagcgtc cggcgcgggc tgagggctgg gcccacagcc cccccgccct ccgggaaaac ggtcagcgtc cggcgcgggc tgagggctgg gcccacagcc
	gccgcgccgg ccggcggggc accacccatt cgccccggtt ccgtggccca gggagtgggc gccgcgccgg ccggcggggc accacccatt cgccccggtt ccgggggccca gggagtgggc
	ggtttcctcc gggacaaaag accgggactc gggttgccgt cgggtcttca cccgcgcggt ggtttcctcc gggacaaaag accgggactc gggttgccgt cgggttttca cccgcgcggt
	tcacagaccg cacatcccca ggctgagccc tgcaacgcgg cgcg*ggccg acagccccgg tcacagaccg cacatcccca ggctgagccc tgcaacgcgg cgcgaggccg acagccccgg
	ccacggagga gccacacgca ggacgacgga ggcgtgattt tggtttccgc gtggctttgc ccacggagga gccacacgca ggacgacgga ggcgtgattt tggtttccgc gtggctttgc
	cctccgcaag gcggcctgtt gctcacgtct ctccggcccc cgaaaggctg gccatgccga cctctgcaag gcggcctgtt gctcacgtct ctccggcccc cgaaaggctg gccatgccga
	ctgtttgctc ccggagctct gcgggcaccc ggaaacatgc agggaacagg gtgcaagccc ctgtttgctc ccggagctct gcgggcaccc ggaaacatgc aggggacagg gtgcaagccc
	ggcacggtgc cttcgctctc cttgccaggt tccaaaccgg ccacactgga ctccccacgt ggcacggtgc cttcgctctc cttgccaggt tccaaaccgg ccacactgga ctccccacgt
	tgccgcacgc gggaatccat cgtcaggcca tcacgccggg gaggcatctc ctctctgggg tgccgcacgc gggaatccat cgtcaggcca tcacgccggg gaggcatctc ctctctgggg
	tctcgctctg gtcttctacg tggaaatgaa cgagagccac acgcctgcgt gtgcgagacc tctcgctctg gtcttctacg tggaaatgaa cgagagccac acgcctgcgt gtgcgagacc
	gtcccggcaa cggcgacgcc cacaggcatt gcctccttca cggagagagg gcctggcaca gtcccggcaa cggcgacgcc cacaggcatt gcctccttca cggagagagg gcctggcaca
	ctcaagactc ccacggaggt tcagttccac actcccctcc accctcccag gctggtttct ctcaagactc ccacggaggt tcagttccac actcccctcc accctcccag gctggtttct

continued ...

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Table E.1: continued ...

Chr 4: Chr 10:	ccctgctgcc gacgcgtggg agcccagaga gcggcttccc gttcccgcgg gatccctgga ccctgctgcc gacgcgtggg agcccagaga gcggcttccc gttcccgcgg gatccctgga
	gaggtccgga gagccggccc ccgaaacgcg ccccctccc ccctcccccc tctcccctt gaggtccgga gagccggccc ccgaaacgcg ccccctccc ccctcccccc tctcccctt
	cctcttcgtc tctccggccc caccaccacc accgccacca cgccctcccc cccaccccc cctcttcgtc tctccggccc caccaccacc accgccacca cgccctcccc caccaccccc
	ccc*cc**** ********** ccaccaccac caccaccacc ccgccggccg gccccaggcc cccacccac accaccacca ccaccaccac caccaccacc ccgccggccg gccccaggcc
	tcgacgccct gggtccctt ccggggtggg gcgggctgtc ccaggggggc tcaccgccat tcgacgccct ggg*tccctt ccggggtggg gcgggctgtc ccaggggggc tcaccgccat
	tcatgaaggg gtggagcctg cctgcctgtg ggcctttaca agggcggctg gctggctggc tcatgaaggg gtggagcctg cctgcccgtg ggcctttaca agggcggctg gctggctggc
	cggctgtccg ggcaggcctc ctggctgcac ctgccgcagt gcacagtccg gctgaggtgc tggctgtccg ggcaggcctc ctggctgcac ctgccgcagt gcacagtccg gctgaggtgc
	acgggagccc gccggcctct ctctgcccgc gtccgtccgt gaaattccgg ccggggctca acgggagccc gccggcctct ctctgcccgc gtccgtccgt gaaattgcgg ccggggctca *Xap* I
	ccgcgatggc cctcccgaca ccctcggaca gcaccctccc cgcggaagcc cggggacgag ccgcgatggc cctcccgaca ccttcggaca gcaccctccc cgcggaagcc cggggacgag
	gacggcgacg gagactcgtt tggaccccga gccaaagcga ggccctgcga gcctgctttg gacggcgacg gagactcgtt tggaccccga gccaaagcga ggccctgcga gcctgctttg
	agcggaaccc gtacccgggc atcgccacca gagaacggct ggcccaggcc atcggcattc agcggaaccc gtacccgggc atcgccacca gagaacggct ggcccaggcc atcggcattc
	cggagcccag ggtccagatt tggtttcaga atgagaggtc acgccagctg aggcagcacc cggagcccag ggtccagatt tggtttcaga atgagaggtc acgccagctg aggcagcacc
	ggcgggaatc tcggccctgg cccgggagac gcggcccgcc agaaggccgg cgaaagcgga ggcgggaatc tcggccctgg cccgggagac gcggcccgcc agaaggccgg cgaaagcgga
	ccgccgtcac cggatcccag accgccctgc tcctccgagc ctttgagaag gatcgctttc ccgccgtcac cggatcccag accgccctgc tcctccgagc ctttgagaag gatcgctttc
	caggcatcgc cgcccgggag gagctggcca gagagacggg cctcccggag tccaggattc caggcatcgc cgcccgggag gagctggcca gagagacggg cctcccggag tccaggattc
	agatctggtt tcagaatcga agggccaggc acccgggaca gggtggcagg gcgcccgcgc agatctggtt tcagaatcga agggccaggc acccgggaca gggtggcagg gcgcccgcgc *Bgl* II
	aggcaggcgg cctgtgcagc gcggcccccg gcgggggtca ccctgctccc tcgtgggtcg aggcaggcgg cctgtgcagc gcggcccccg gcgggggtca ccctgctccc tcgtgggtcg
	ccttcgccca caccggcgcg tggggaacgg ggcttcccgc accccacgtg ccctgcgcgc ccttcgccca caccggcgcg tggggaacgg ggcttcccgc accccacgtg ccctgcgcgc
	ctggggctct cccacagggg gctttcgtga gccaggcagc gagggccgcc cccgcgctgc ctggggctct cccacagggg gctttcgtga gccaggcagc gagggccgcc cccgcgctgc
	agcccagcca ggccgcgccg gcagagggga tctcccaacc tgccccggcg cgcggggatt agcccagcca ggccgcgccg gcagagggga tctccca*cc tgccccggcg cgcggggatt

continued ...

Table E.1: continued ...

Chr 4:	tcgcctacgc cgccccggct cctccggacg gggcgctctc ccaccctcag gctcctcggt
Chr 10:	tcgcctacgc cgccccggct cctccggacg gggcgctctc ccaccctcag gctcctcggt
	ggcctccgca cccgggcaaa agccgggagg accgggaccc gcagcgcgac ggcctgccgg
	ggcctccgca cccgggcaaa agccgggagg accgggaccc gcagcgcgac ggcctgccgg
	gcccctgcgc ggtggcacag cctgggcccg ctcaagcggg gccgcagggc caaggggtgc
	gcccctgcgc ggtggcacag cctgggcccg ctcaagcggg gccgcagggc caaggggtgc
	ttgcgccacc cacgtcccag gggagtccgt ggtggggctg gggccggggt ccccaggtcg
	ttgcgccacc cacgtcccag gggagtccgt ggtggggctg gggccggggt ccccaggtcg
	ccggggcggc gtgggaaccc caagccgggg cagctccacc tccccagccc gcgccccccgg
	ccggggcggc gtgggaaccc caagccgggg cagctccacc tccccagccc gcgccccccgg
	acgcctccgc ctccgcgcgg caggggcaga tgcaaggcat cccggcgccc tcccaggcgc
	acgcctccg* *****cgcgg caggggcaga tgcaaggcat cccggcgccc tcccaggcgc
	tccaggagcc ggcgccctgg tctgcactcc cctgcggcct gctgctggat gagctcctgg
	tccaggagcc ggcgccctgg tctgcactcc cctgcggcct gctgctggat gagctcctgg
	cgagcccgga gtttctgcag caggcgcaac ctctcctaga aacggaggcc ccgggggagc
	cgagcccgga gtttctgcag caggcgcaac ctctcctaga aacggaggcc ccgggggagc
	tggaggcctc ggaagaggcc gcctcgctgg aagcacccct cagcgaggaa gaataccggg
	tggaggcctc ggaagaggcc gcctcgctgg aagcacccct cagcgaggaa gaataccggg
	ctctgctgga ggagctttag gacgcggggt tgggacgggg tcgggtggtt cggggcaggg
	ctctgctgga ggagctttag gacgcggggt tgggacgggg tcgggtggtt cggggcaggg
	cggtggcctc tctttcgcgg ggaacacctg gctggctacg gaggggcgtg tctccgcccc
	cggtggcctc tctttcgcgg ggaacgcctg gctggctacg gaggggcgtg tctccgcccc
	gccccctcca ccgggctgac cggcctggga ttcctgcctt ctaggtctag gcccggtgag
	gccccctcca ccgggctgac cggcctggga ttcctgcctt ctaggcctag gcccggtgag
	Bln I
	agactccaca ccgcggagaa ctgccattct ttcctgggca tcccggggat cccagagccg
	agactccaca cagcggagaa ctgccattct ttcctgggca tcccggggat cccagagccg
	gcccaggtac c
	gcccaggtac c
	Kpn I

a) The 4q35 sequence was retrieved from Genbank with accession number AF117653 as reported by Gabriels *et al.*, (1999) and the 10q26 sequence was retrieved from Genbank with accession number NT_028298. b) The restriction enzymes are indicated by: cctagg for *Bln* I; agatct for *Bgl* II; aaattc for Xap I; and ggtacc for *Kpn* I. c) The sequence of p13E-11 is indicated by single underlining. d) The sequence indicated in green and the * = differences observed in the two compared sequences.

APPENDIX F

EXTENDED FAMILY PEDIGREES

This appendix contains the full pedigrees of the core FSHD families (F10, F20, F30, F40 and F60), which were selected for genotyping as discussed in chapter three. Symbols utilised within the pedigrees are listed below and were assigned as reported by the Pedigree Standardization Task Force [PSTF] (Bennet *et al.*, 1995).

male/female: tested FSHD normal

male/female: FSHD equivocal

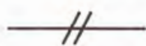
male/female: never tested for FSHD (phenotypical status unknown)

male/female: tested FSHD positive

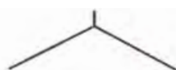
male/female: deceased

Sex unknown

multiple individuals, exact number of individuals unknown


Divorced

Proband


dizygotic twins

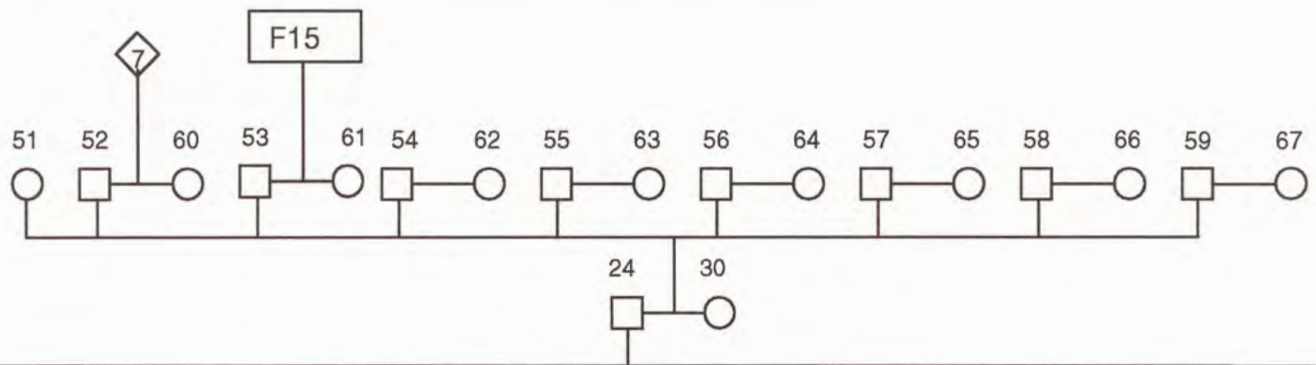
monozygotic twins

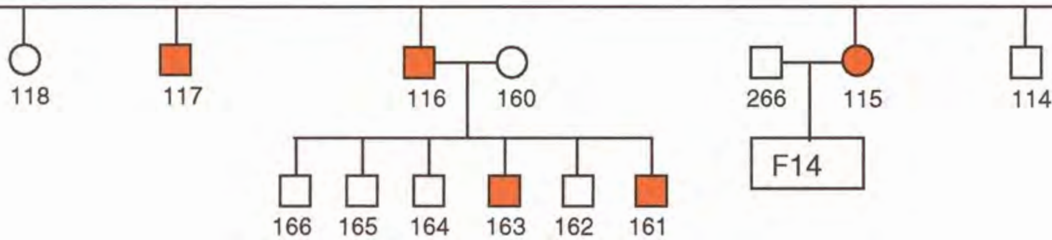
UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Figure F.1: Full pedigree of South African FSHD family F10

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Figure F.1: Full pedigree of South African FSHD family F10 continued ...





F14 : Subfamily F14 of family F10 as presented in Figure F.5

F15 : Subfamily F15 of family F10 as presented in Figure F.6

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Figure F.1: Full pedigree of South African FSHD family F10 continued ...

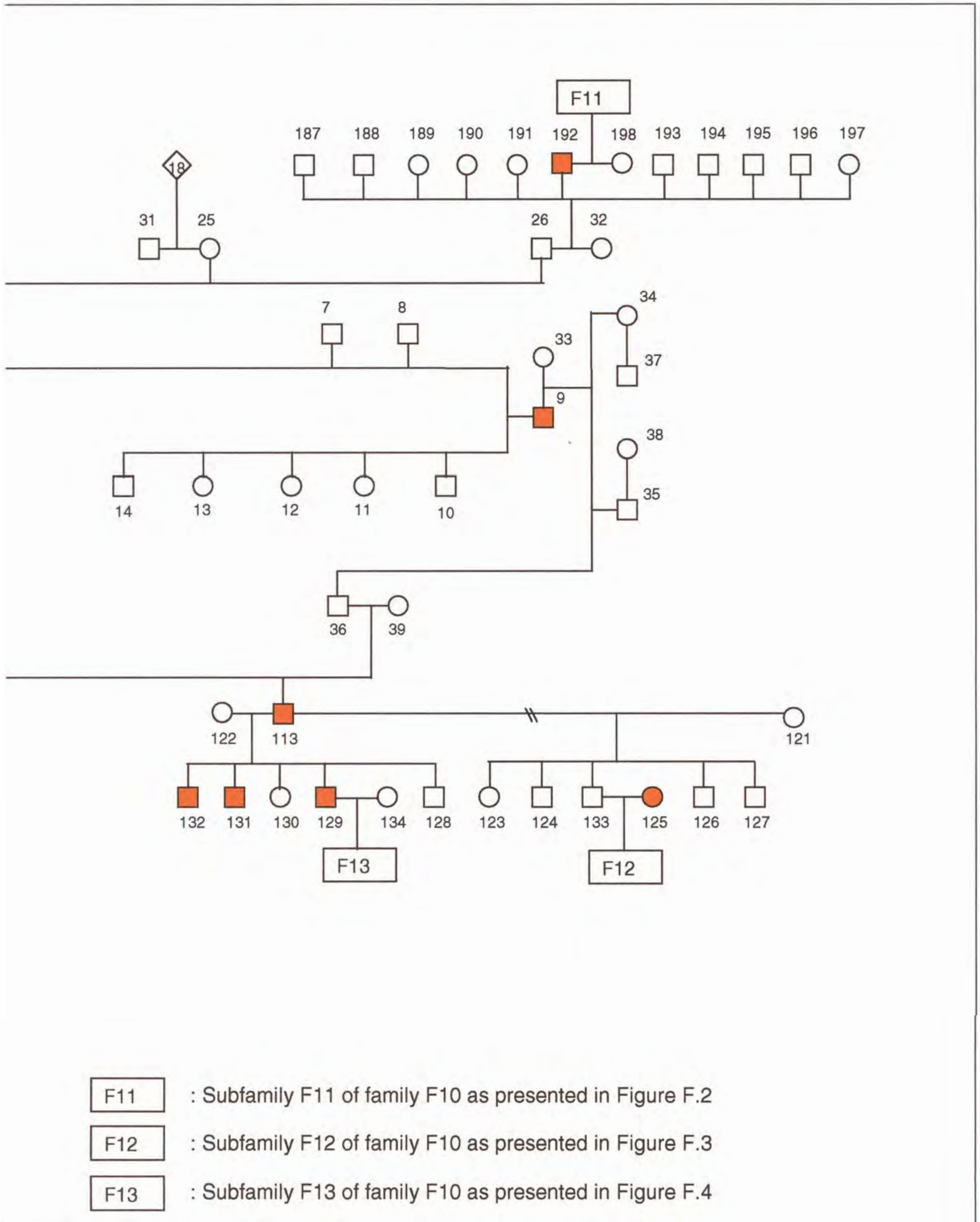
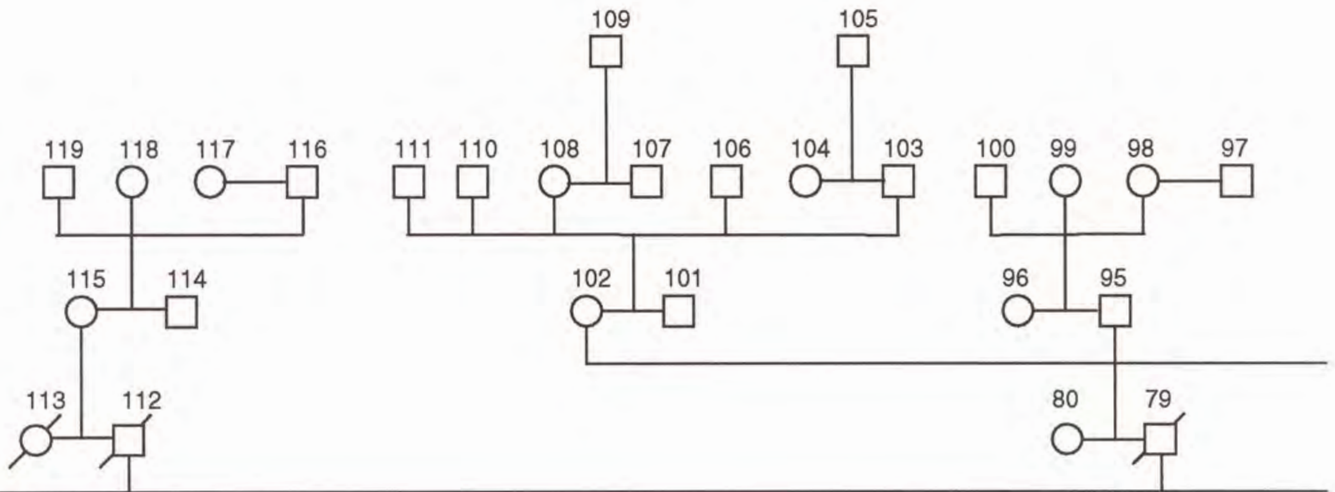


Figure F.2: Full pedigree of South African FSHD family F11

Figure F.2: Full pedigree of South African FSHD family F11 continued ...



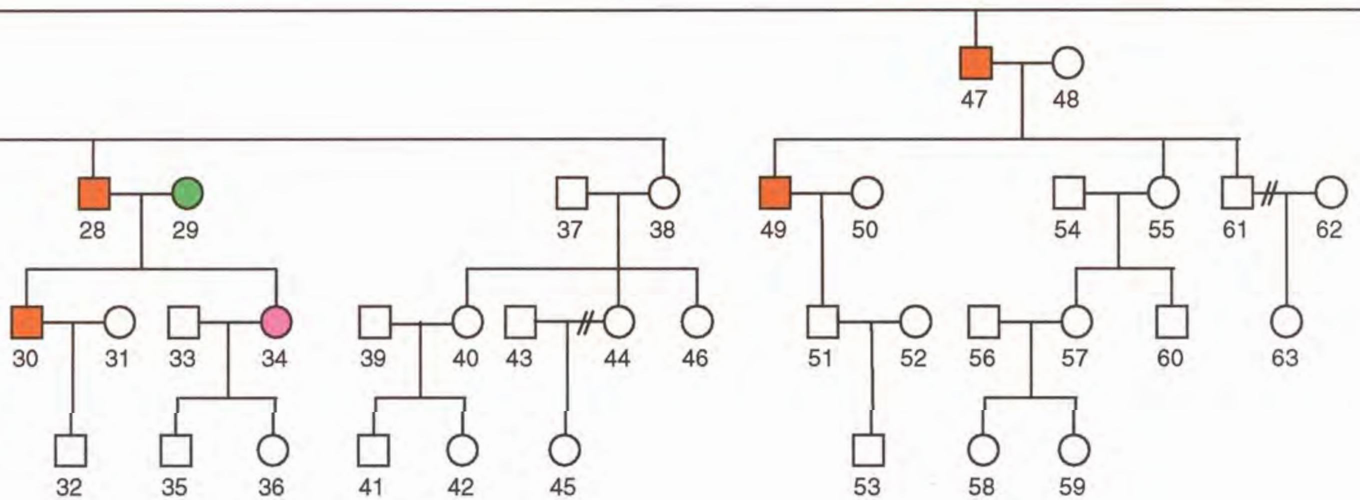
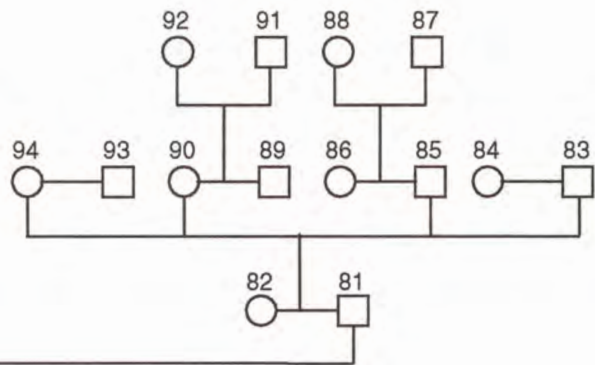
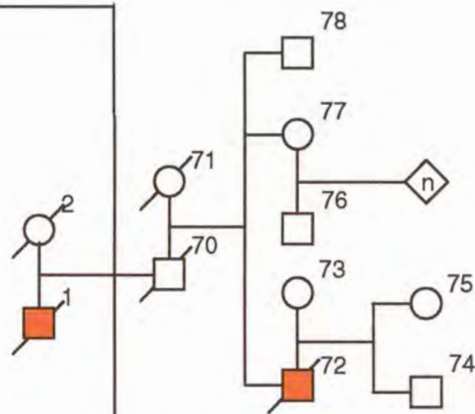


Figure F.2: Full pedigree of South African FSHD family F11 continued ...





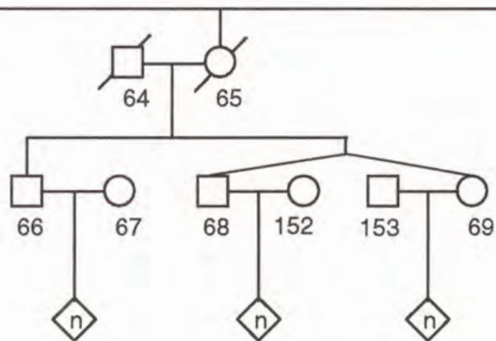


Figure F.3: Full pedigree of South African FSHD family F12

Figure F.4: Full pedigree of South African FSHD family F13

263 262 261 260 259 258 257 152

264 158 159 256 157 155 156 154 153 150 151

138 267 141 137

149 241

148 240

239

147

238

134 146

237

129

145 236

144 235

140

136 143

234

142

139 135

243 242 244 245 246 247 248

249 250 251 252 253 254

Figure F.5: Full pedigree of South African FSHD family F14

11 12 17

9 10 13 14 15 16 20 21 22 23 26 27

7 8 18 19 24 25

5 6

3 4

156 157

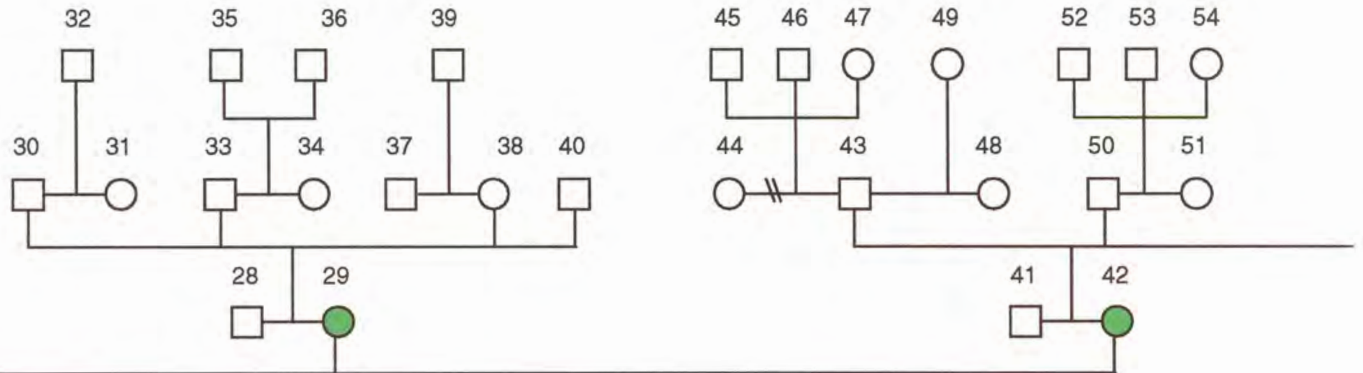
181 182 173 174 165 166 158 159

175 176 177 178 179 180 167 168 169 172 160 161 163 164

170 171 162 183

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Figure F.5: Full pedigree of South African FSHD family F14 continued …



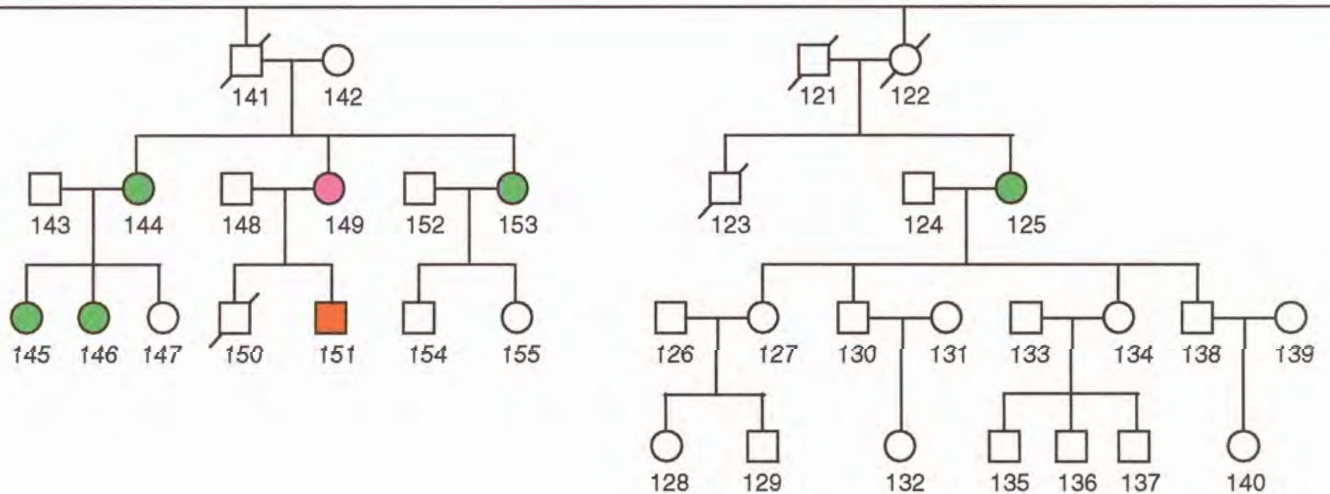


Figure F.5: Full pedigree of South African FSHD family F14 continued ...

Figure F.6: Full pedigree of South African family F15

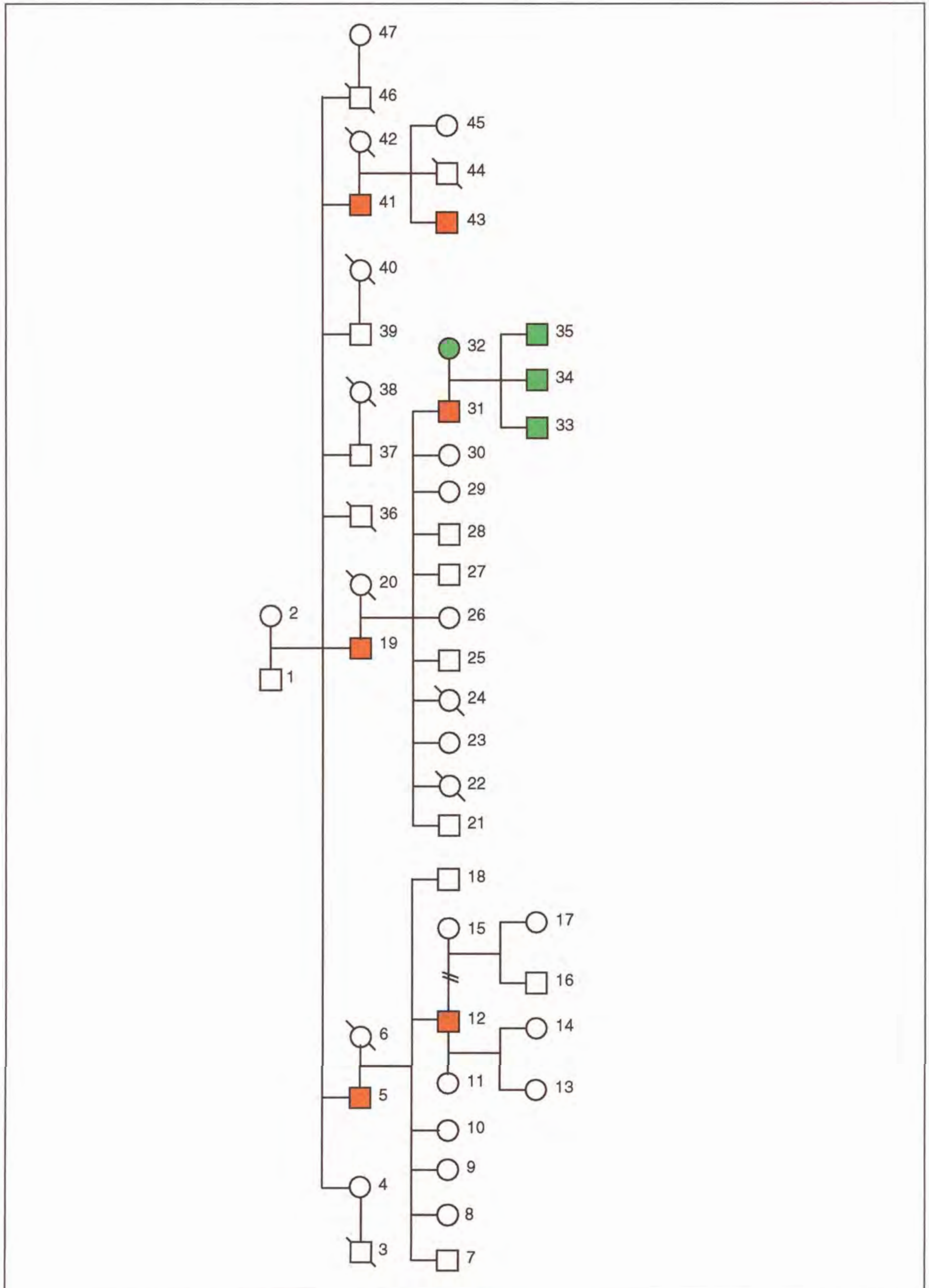
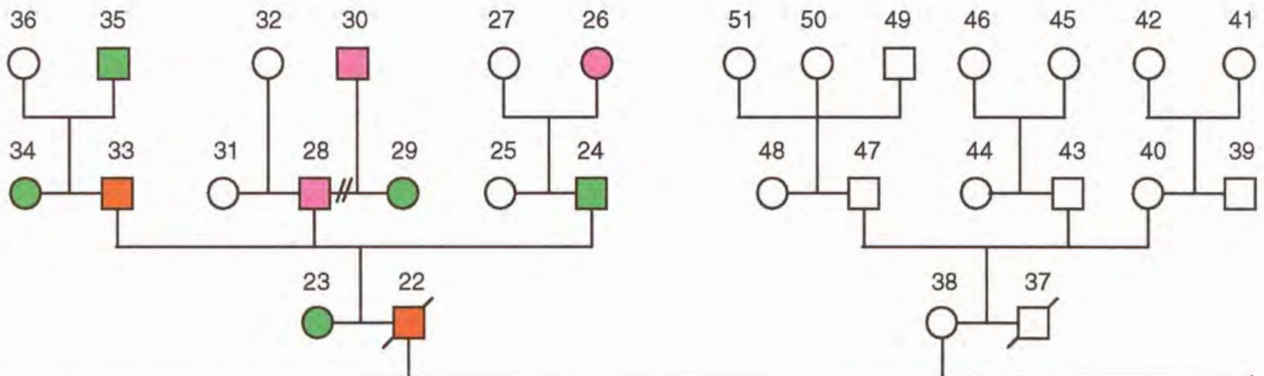
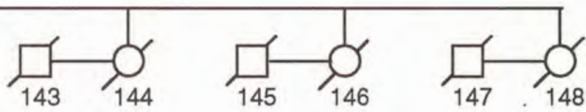


Figure F.7: Full pedigree of South African family F21

Figure F.7: Full pedigree of South African family F21 continued ...





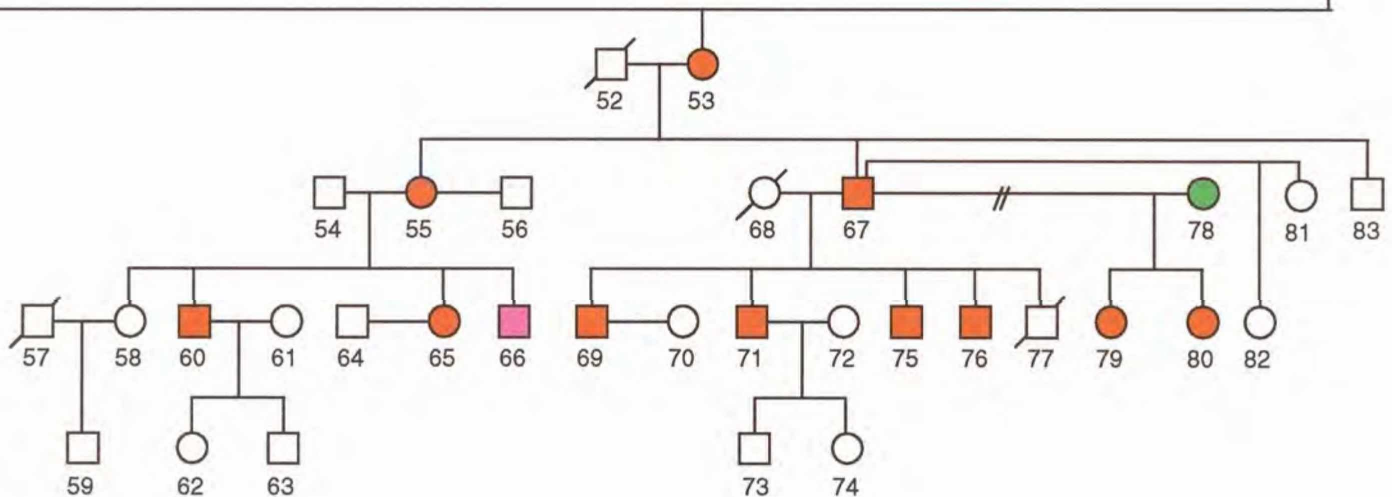
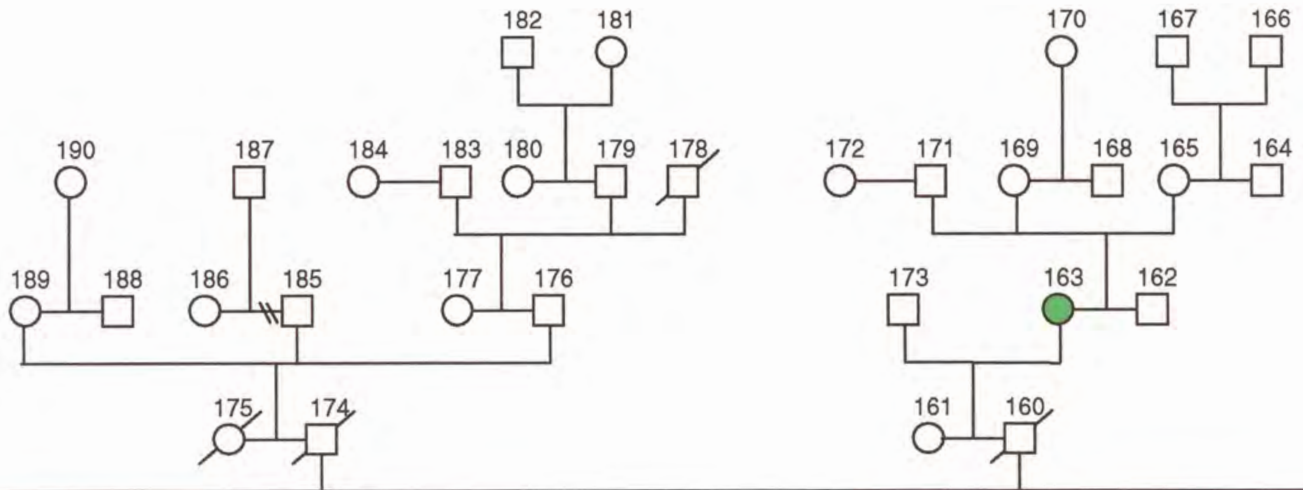


Figure F.8: Full pedigree of South African FSHD family F30

Figure F.8: Full pedigree of South African FSHD family F30 continued ...



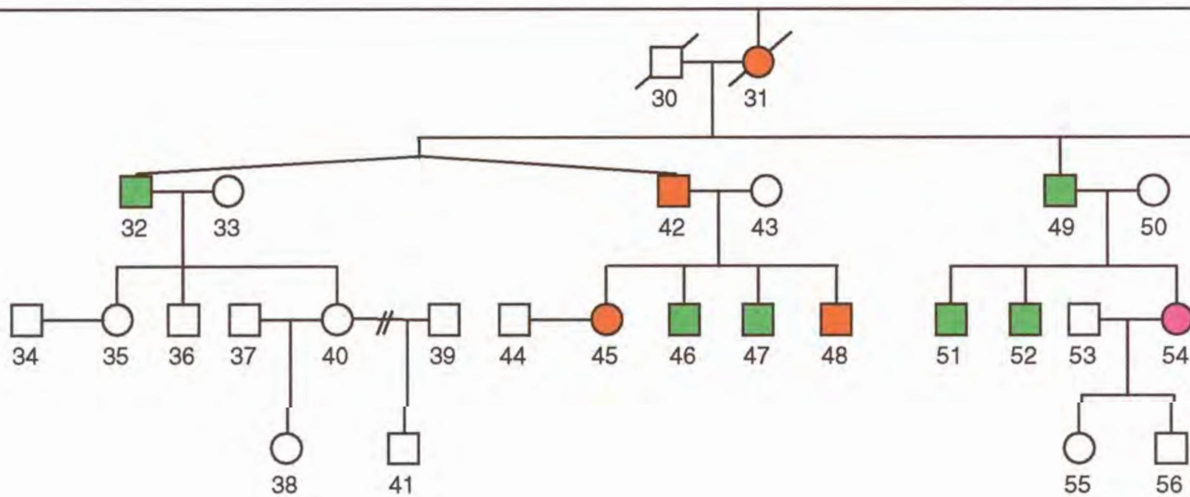
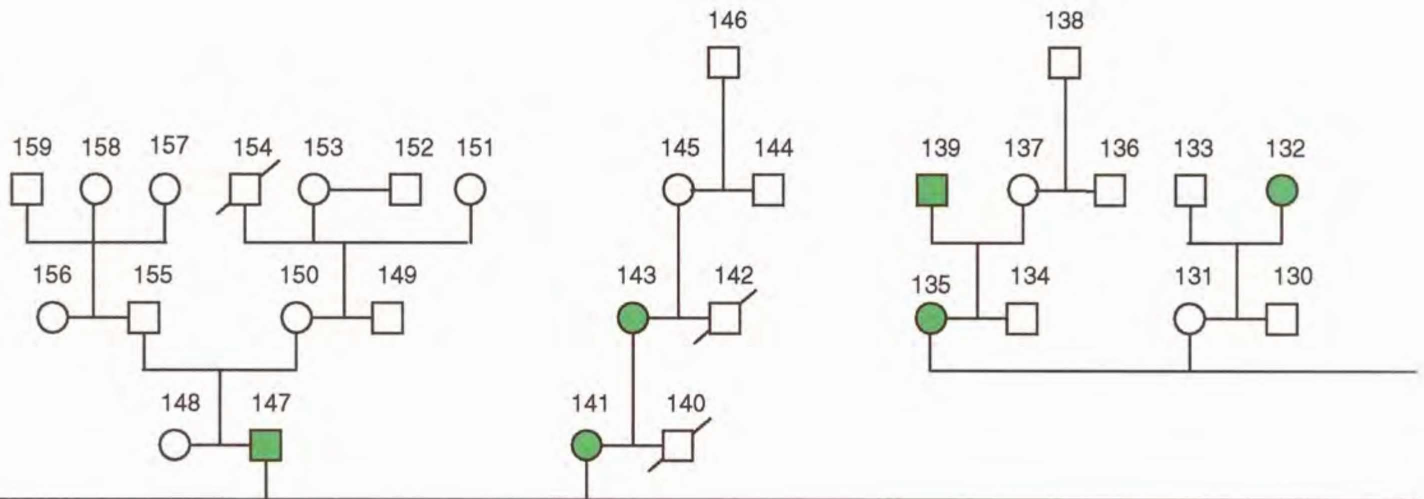


Figure F.8: Full pedigree of South African FSHD family F30 continued ...



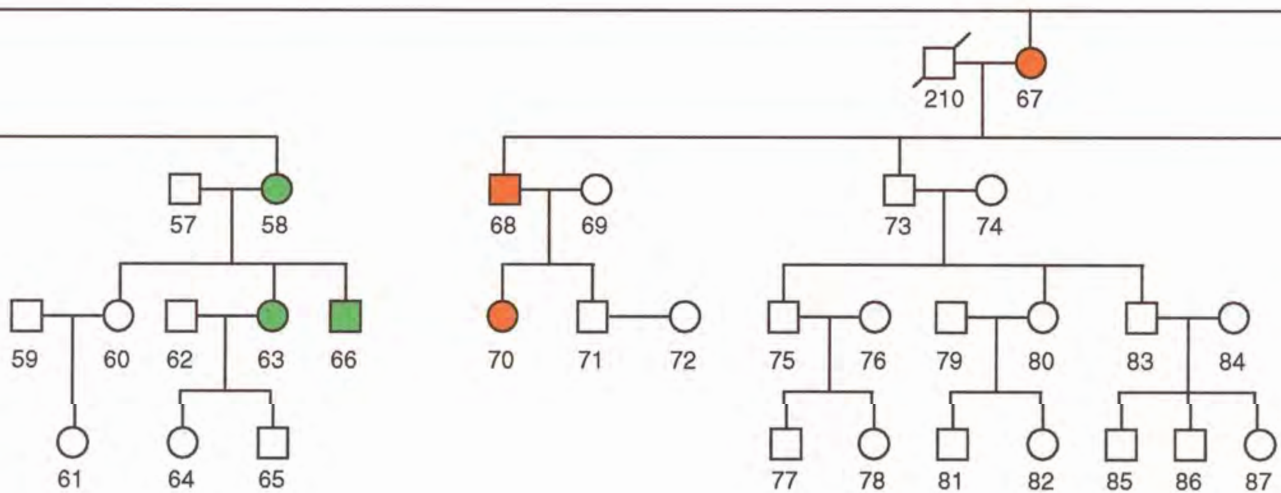


Figure F.8: Full pedigree of South African FSHD family F30 continued ...

Figure F.9: Full pedigree of South African FSHD family F40

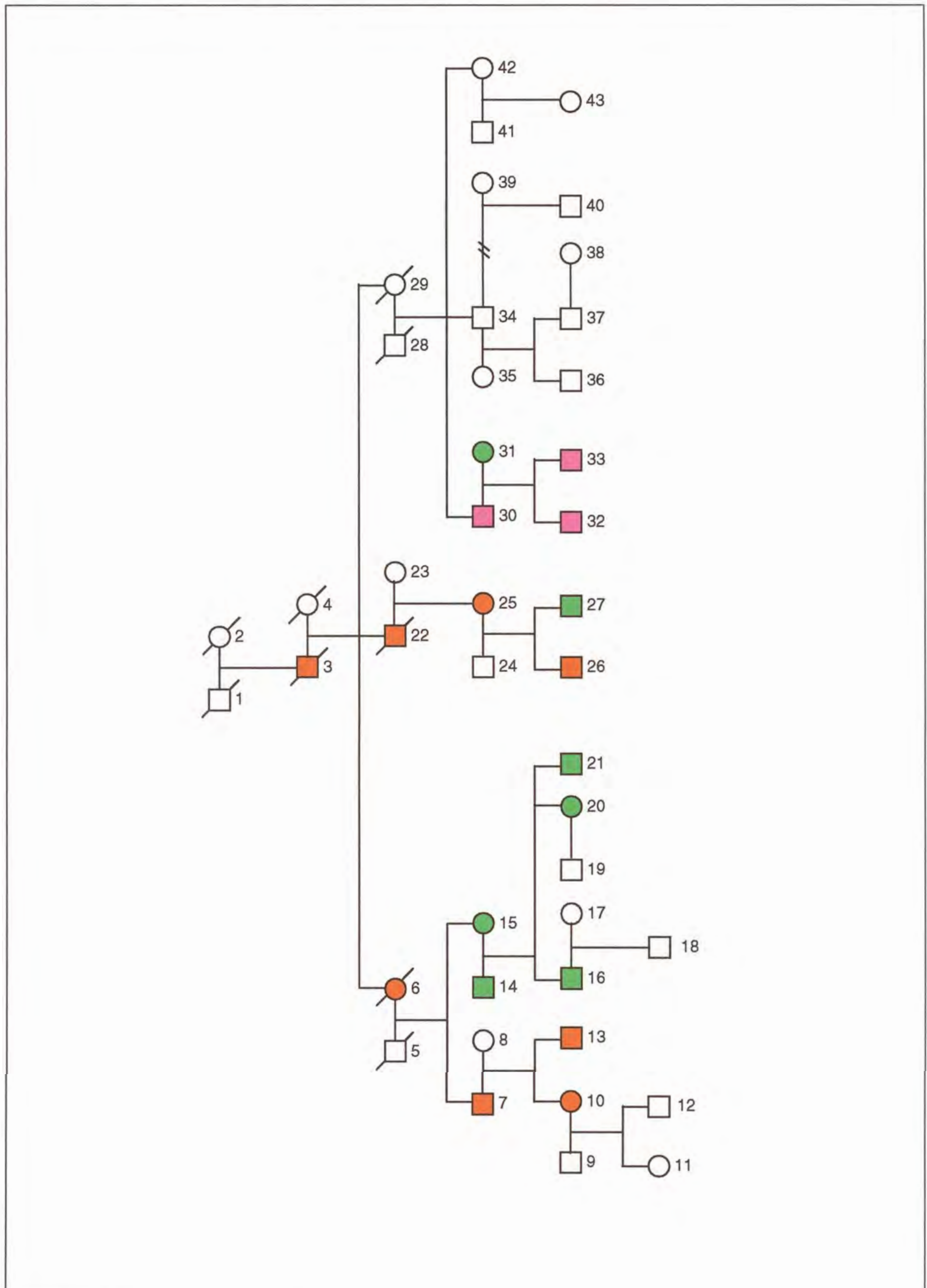


Figure F.10: Full pedigree of South African FSHD family F60

77 45 44 42 41 39

47 76 46 75 43 74 40 73 38

78 6

3 5

48 8 54 16 55

9 49 12 50 14 52 15 19 20 21 17 18

10 11 13 51 53

Figure F.10: Full pedigree of South African FSHD family F60 continued ...