Molecular analysis of the facioscapulohumeral muscular dystrophy (FSHD) associated DNA rearrangements in the South African population

BY

ANNELIZE VAN DER MERWE, M.Sc.

Thesis submitted for the degree Philosophiae Doctor in Biochemistry
at the Potchefstroomse Universiteit vir Christelike Hoër Onderwys

PROMOTER: Professor Antonel Olckers
Centre for Genome Research,
Potchefstroom University for Christian Higher Education

CO-PROMOTER: Doctor Clara-Maria Schutte
Department of Neurology, University of Pretoria

December 2003

Molekulêre analise van die fasioskapulêrehumerale spierdistrofie (FSHD) geassosieerde DNA herrangskikkings in die Suid-Afrikaanse populasie

DEUR

ANNELIZE VAN DER MERWE, M.Sc.

Proefskrif voorgelê vir die graad Philosophiae Doctor in Biochemie
aan die Potchefstroomse Universiteit vir Christelike Hoër Onderwys

PROMOTOR: Professor Antonel Olckers
Sentrum vir Genomiese Navorsing,
Potchefstroomse Universiteit vir Christelike Hoër Onderwys

MEDEPROMOTOR: Doktor Clara-Maria Schutte
Departement Neurologie, Universiteit van Pretoria

Desember 2003

TO MY PARENTS

ABSTRACT

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. FSHD is characterised by progressive weakening and atrophy of the face, shoulder-girdle and upper arm, although other skeletal muscles may also become involved with progression of the disorder.

The FSHD phenotype segregates as an autosomal dominant trait. Linkage was established to chromosome 4q35 in 1990. A deletion of an integral number of 3.3 kb repeats, localised at the D4Z4 locus, was reported to cause FSHD. Translocation events between chromosomes 4q35 and 10q26, also harbouring similar 3.3 kb repeat units, have been detected via the presence of *Bln* I sites within the 10q26 repeats.

In this study the D4Z4 locus was investigated for the first time on a molecular level in the Black South African and Khoi-San populations. The translocation frequency between chromosomes 4q35 and 10q26 was evaluated via the *Bln* I / *Bgl* II dosage test. Haplogroup analysis was utilised to determine the relative evolutionary age of the translocation event.

The Eurasian population harbours an excess of 4-on-10 fragments. This excess was postulated to be a significant, if not the major predisposing factor that gives rise to the FSHD-type deletion. The predisposed population thus has individuals that are more susceptible to FSHD.

An enrichment of 10-on-4 was observed in the Black South African population. It was postulated that this enrichment is an epigenetic protective factor for FSHD, since no FSHD case has been reported in this population to date. It is further hypothesised that the absence of FSHD cases in this population is due to this enrichment. As a consequence, the excess and enrichment of specific translocation profiles in different populations is an additional factor that affects the aetiology of FSHD within specific populations.

OPSOMMING

Fasioskapulêrehumerale spierdistrofie (FSHD) is die derde algemeenste oorerflike spiertoestand na Duchenne en Miotoniese distrofie. FSHD word gekenmerk deur progressiewe verswakking en spieratrofie van die gesig, skouergordel en bo-arm spiere, maar ander skeletale spiere kan ook mettertyd aangetas word.

Die FSHD fenotipe segregeer as 'n outosomale dominante toestand en koppeling is gevind met chromosoom 4q35 in 1990. FSHD word veroorsaak deur 'n delesie van 'n aantal van die 3.3 kb herhaalvolgordes by die D4Z4 lokus. Translokasies tussen die herhalings op chromosome 4 en 10, wat soortgelyke 3.3 kb herhalings bevat, is waargeneem deur die teenwoordigheid van *Bln* I setels binne in die 10q26 herhalings.

In hierdie studie is die D4Z4 lokus vir die eerste keer op 'n molekulêre vlak in die Swart Suid-Afrikaanse en Khoi-San populasies bestudeer. Die translokasiefrekwensie tussen chromosome 4q35 en 10q26 is bestudeer met behulp van die *Bln* I / *Bgl* II dosistoets. Haplogroep analise is gebruik om die relatiewe evolusionêre ouderdom van die translokasies te bepaal.

Die Eurasiese populasie het 'n oormaat van 4-op-10 fragmente. Hierdie oormaat is gepostuleer om 'n beduidende, indien nie die vernaamste vatbaarheidsfaktor te wees wat oorsprong gee aan die FSHD-tipe delesie. Individue in die vatbare populasie is dus meer geneig tot FSHD.

'n Verryking van 10-op-4 fragmente is waargeneem in die Swart Suid-Afrikaanse populasie. Dit is gepostuleer dat hierdie verryking 'n epigenetiese beskermingsfaktor vir FSHD is, aangesien FSHD gevalle nie in hierdie populasie gerapporteer is nie. Die hipotese is dus gevorm dat die afwesigheid van gerapporteerde FSHD gevalle in hierdie populasie 'n gevolg is van hierdie verryking. Gevolglik is die oormaat en verryking van spesifieke translokasieprofiele in verskillende populasies 'n addisionele faktor wat die etiologie van FSHD beïnvloed in spesifieke populasies.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND SYMBOLS

Abbreviations and symbols are listed in alphabetical order.

4p	short arm of chromosome 4
4qA	variant A of the long arm of chromosome 4
4qB	variant B of the long arm of chromosome 4
4qter	telomeric region of the long arm of chromosome 4
4-on-4	chromosome 4-type fragments localised on chromosome 4
4-on-10	chromosome 4-type fragments localised on chromosome 10
4:10 ratio	ratio of chromosome 4-type fragments to chromosome 10-type fragments
9B6A	probe complimentary to the homeobox sequences within each 3.3 kb repeat unit
10qter	telomeric region of the long arm of chromosome 10
10-on-4	chromosome 10-type fragments localised on chromosome 4
10-on-10	chromosome 10-type fragments localised on chromosome 10
α	alpha
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
ACE	Associated Chemical Enterprises
ACTA1	actin alpha skeletal muscle
AD	autosomal dominant
AD EDMD	autosomal dominant Emery-Dreifuss muscular dystrophy
ADP	adenosine diphosphate
ALS	amyotrophic lateral sclerosis
Alu	short interspersed nuclear element, characterised by the *Alu* I restriction enzyme
ANT	adenine nucleotide translocator
ANT1	adenine nucleotide translocator isoform 1
ANT2	adenine nucleotide translocator isoform 2
ANT3	adenine nucleotide translocator isoform 3
AP-SA	alkaline phosphatase-labelled streptavidin
AR	autosomal recessive
AR	androgen receptor
AT	amino terminal
ATP	adenosine triphosphate
AXIM	Africa X-ray Industrial and Medical Pty (Ltd)
β	beta
Bam HI	restriction endonuclease isolated from *Bacillus amyloliquefaciens H*, with recognition site 5'-G↓GATCC-3'
Bgl II	restriction endonuclease isolated 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 *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
X^2	chi-square
C1 buffer	4% v/v Triton®[1] X-100, 43 mM $MgCl_2$, 40 mM Tris and 1.3 M sucrose
ca.	circa: approximately
CACNA1A	α1A-voltage-dependent calcium channel subunit
CAPN3	calpain-3
CAV3	caveolin-3
CCD	central core disease
cDNA	complementary DNA
CDP-Star®[2]	disodium 2-chloro-5-(4-methoxyspiro[1,2-dioxetane-3,2'-(5'-chloro)- tricycle [3,3,1,1] decan]-4yl)phenyl) phosphate: $C_{18}H_{19}Cl_2O_7PNa_2$
CEB8	probe complementary to locus D4F35S1
CEN	centromere
chr	chromosome
CI	cardiac involvement
CK	creatine kinase
CLC-1	muscle chloride channel
cm	centimeter: 10^{-2} meter
cM	centimorgan
CMD	congenital muscular dystrophy
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
CRYAB	αB-crystalling
CS	clinical severity
CT	carboxy terminal
CT-scan	computed tomography scan
Cu	copper
δ	delta
D4Z4	FSHD locus harbouring 3.3 kb repeat elements
dATP	2'-deoxyadenosine-5'-triphosphate
DBE	D4Z4 binding element
DBP	detector block powder
dCTP	2'-deoxycytidine-5'-triphosphate
ddH_20	double distilled water
DEPC	diethyl pyrocarbonate
DES	desmin
DGC	dystrophin-glycoprotein complex
dGTP	2'-deoxyguanosine-5'-triphosphate
DIG	digoxigenin
DM	myotonic dystrophy
DM2	myotonic dystrophy type 2
DMD	Duchenne muscular dystrophy
DMPK	myotonin-protein kinase gene
DMRV	distal myopathy with rimmed vacuoles
DNA	deoxyribonucleic acid

[1] Triton® X-100 is a registered trademark of Rohm & Haas Company, Philadelphia, PA, U.S.A.
[2] CPD-Star® is a registered trademark of Tropix Inc., Bedford, MA, U.S.A.

DNR	dinucleotide repeat
dNTP	2'-deoxynucleotide triphosphate
DRC	D4Z4 recognition complex
DRM	desmin related myopathy
DRPLA	dentatorubral-pallidoluysian atrophy (Haw River syndrome)
dsDNA	double stranded DNA
DTT	dithiothreithol: threo-1,4-dimercapto-2,3-butanediol: $C_4H_{10}O_2S_2$
dTTP	2'-deoxythymidine-5'-triphosphate
dUTP	2'-deoxyuridine-5'-triphosphate
DUX	double homeobox gene
DUX1	double homeobox gene 1
DUX2	double homeobox gene 2
DUX3	double homeobox gene 3
DUX4	double homeobox gene 4
ε	epsilon
E	expected value (χ^2 test)
Eco RI	restriction endonuclease isolated from an *E. coli* strain that carries the cloned *ecoRIR* 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
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
et al.	*et altera*: Latin abbreviation for "and others"
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
FISH	fluorescent *in situ* hybridisation
FKRP	fukutin-related protein
Fl-dUTP	fluorescein-11-2'-deoxyuridine-5'-triphosphate
FMR1	fragile site mental retardation 1 gene
FMR2	fragile site mental retardation 2 gene
FMRP	fragile X mental retardation protein
FRAXA	fragile X syndrome
FRAXE	fragile XE syndrome
FRDA	Friedreich ataxia
FRG1	FSHD region gene 1
FRG1P	FRG1 protein
FRG2	FSHD region gene 2
FSHD	facioscapulohumeral muscular dystrophy
γ	gamma
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	immunoglobulim marker
GNE	acetylglucosamine-2-epimerase

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

H_2O	water
HCl	hydrochloric acid
HD	Huntington disease
hhsmp3	human DNA insert showing sperm-specific hypomethylation
HIBM	hereditary inclusion body myopathy
Hind III	restriction endonuclease isolated from *Haemophilus influenzae*, with recognition site 5'-A↓AGCTT-3'
HLA	human leukocyte antigens
HMGB2	high mobility group box 2 protein
Hmix	*Xenopus* mesoderm induced homeobox
HmprD	*Drosophila* paired
HSPG	heparan sulfate proteoglycan
Hz	hertz
IBM2	inclusion body myopathy type 2
IBM3	inclusion body myopathy type 3
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, 1 M KCl
K-acetate	potassium acetate: CH_3COOK
kb	kilo (10^3) base pair
KCl	potassium chloride
kDa	kilo Dalton
$KHCO_3$	potassium hydrogen carbonate
KH_2PO_4	potassium phosphate monobasic
kHz	kilo hertz
Kpn I	restriction endonuclease isolated from *Klebsiella pneumoniae*, with recognition site 5'-GGTAC↓C-3'
Ksp AI	restriction endonuclease isolated from *Kurtia species* N88, with recognition site 5'-GTT↓AAC-3'
LAMA2	laminin α2 chain of merosin
lamin A/C	gene encoding two components of the nuclear lamina, lamins A and C
LGMD	limb-girdle muscular dystrophy
LGMD1	limb-girdle muscular dystrophy type 1
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
LGMD2	limb-girdle muscular dystrophy type 2
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
LGMD2J	limb-girdle muscular dystrophy type 2J
LINE	long interspersed nuclear element
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
LTR	long terminal repeat
µ	micro: 10^{-6}
µCi	micro Curie
µg	microgram
µl	microlitre
µM	micromolar
m	milli: 10^{-3}
M	molar: moles per litre
M13mp18	vector number 18 of the mp series of bacteriophage M13
MBS®[1]	multiblock system
MD	muscular dystrophy
MDC1A	congenital muscular dystrophy type 1A
MDC1B	congenital muscular dystrophy type 1B (with secondary merosin deficiency 2)
MDC1C	congenital muscular dystrophy type 1C (with secondary merosin deficiency 2)
MD-EBS	epidermolysis bullosa simplex associated with late-onset muscular dystrophy
MEAX	myopathy with excessive autophagy
MEB	muscle-eye-brain disease
Mfd22	short tandem repeat polymorphism marker at locus D4S171
MFM/ARVC	myofibrillar myopathy with arrhythmogenic right ventricular cardiomyopathy
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
milli	10^{-3}
min	minutes
ml	millilitres
mm	millimetre
mM	millimolar
MM	Miyoshi myopathy
MPD1	autosomal dominant distal myopathy (myopathy distal type 1)
MPRM1	autosomal dominant myopathy with proximal weakness and early respiratory muscle involvement type 1
MPRM2	autosomal dominant myopathy with proximal weakness and early respiratory muscle involvement type 2
mRNA	messenger RNA
mtDNA	mitochondrial DNA
MTM1	myotubular myopathy
MYHC2A	myosin heavy chain IIa
MYOT	myotilin
n	nano: 10^{-9}
N-Lauroylsarcosine	$C_{15}H_{28}NO_3Na$
Na-citrate	trisodium citric acid 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
Na_2HPO_4	sodium phosphate dibasic

[1] MBS® is a registered trademark of Thermo Electron Corporation, Milford, MA, U.S.A.

NaOH	sodium hydroxide
ng	nanogram
NCBI	National Center for Biotechnology Information, U.S.A.
NEM1	nemaline myopathy type 1
NEM2	nemaline myopathy type 2
NEM3	nemaline myopathy type 3
NEM4	nemaline myopathy type 4
NEM5	nemaline myopathy type 5
NH_4Cl	ammonium chloride
NIH	National Institutes of Health, U.S.A.
nm	nanometer: 10^{-9} meter
nM	nanomolar
NMR	nuclear magnetic resonance
No	number
O	observed value (χ^2 test)
OD	optical density
OMIM™[1]	Online Mendelian Inheritance in Man
OPMD	oculopharyngeal muscular dystrophy
orange G	7-hydroxy-8-phenylazo-1,3-naphthalenedisulfonic acid: $C_{16}H_{10}N_2O_7S_2Na_2$
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
PBS buffer	phosphate-buffered saline buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4 (pH 7.4), 2 mM KH_2PO_4)
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
pg	picogram
PGD	preimplantation genetic diagnosis
pH	indicates acidity: numerically equal to the negative logarithm of H+ concentration expressed in molarity
pH30	probe complimentary to locus D4S139
pLAM	the most distal repeat unit at the D4Z4 locus, consisting of only 2.9 kb of a 3.3 kb repeat unit
pmol	picomole
POMGnT1	O-mannoseβ-1,2-N-acetylglucosaminyl transferase
PPP2R2B	protein phosphatase 2A

[1] OMIM™ and Online Mendelian Inheritance in Man™ are trademarks of the Johns Hopkins University, MA, U.S.A.

prd	paired gene
Pre	premutation
PROMM	proximal myotonic myopathy
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
RPMI	Roswell Park Memorial Institute
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 *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'
SBMA	spinobulbar muscular dystrophy (Kennedy disease)
SCA	spinocerebellar ataxia
SCA1	spinocerebellar ataxia type 1
SCA2	spinocerebellar ataxia type 2
SCA3	spinocerebellar ataxia type 3
SCA6	spinocerebellar ataxia type 6
SCA7	spinocerebellar ataxia type 7
SCA8	spinocerebellar ataxia type 8
SCA12	spinocerebellar ataxia type 12
SCK	serum creatine kinase
SDS	sodium dodecyl sulphate: $C_{12}H_{25}NaSO_4$
SE buffer	sodium chloride / EDTA buffer (75 mM NaCl, 25 mM EDTA at pH 8.0)
sec	seconds
SEPN1	selenoprotein N1
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
SMA	spinal muscular atrophy
SNP (s)	single nucleotide polymorphism (s)
SOD1	Cu/Zn superoxide dismutase
spermidine	N-[3-Aminopropyl]-1,4-butanediamine: $C_7H_{19}N_3$
SSC	saline-sodium-citrate buffer: 0.15 M NaCl, 15 mM Na-citrate at pH 7.0
ssDNA	single stranded DNA
SSPE	saline-sodium-phosphate-EDTA buffer: 0.15 M NaCl, 10 mM NaH_2PO_4 (pH 7.4), 1 mM EDTA
STRP	short tandem repeat polymorphism
STS	sequence tagged site
θ	theta – recombination fraction
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: 89.15 mM Tris®[1] (pH 8.0), 88.95 mM boric acid, 2.498 mM Na_2EDTA
TCAP	telethonin
TE buffer	10 mM Tris-HCl (pH 7.5); 1 mM EDTA
TEL	telomere
temp	temperature
TMD	tibial muscular dystrophy
TNN	titin
TNNT1	troponin T1
TPM2	β tropomyosin
TPM3	α tropomyosin 3
TRIM32	tripartite-motif-containing protein 32
Tris	Tris®: tris(hydroxymethyl)-amino-methane: 2-amino-2-(hydroxymethyl)-1,3-propane diol: $C_4H_{11}NO_3$
Tris-HCl	2-amino-2-2(hydroxymethyl)-1,3-propanediol hydrochloride: $C_4H_{11}NO_3.H_2O$
Triton X-100	Triton®: X-100 octylphenolpoly(ethylene-glycolether)$_n$: $C_{34}H_{62}O_{11}$, for n = 10
TUBB4Q	human tubulin beta polypeptide 4 member Q
Tween™[2]20	polyoxyethylene sorbitan monolaurate
U	units
UCMD	Ullrich congenital muscular dystrophy
UK	United Kingdom
U.S.A.	United States of America
UTR	untranslated region
UV	ultraviolet
V	volt
VNTR	variable number of tandem repeats
VPDMD	vocal cord and pharyngeal weakness with autosomal dominant distal myopathy
v/v	volume per volume
w/v	weight per volume
Xap I	restriction endonuclease isolated from *Xanthomonas ampelina* Slo 51-021 with recognition site 5'-Pu↓AATTPy-3'
XC	xylene cyanole FF: $C_{25}H_{27}N_2O_6S_2Na$
x g	relative gravitational acceleration
XR	X-linked recessive
YAC	yeast artificial chromosome
YY1	Yin Yang 1 transcription factor
Zn	zinc
ZNF9	zinc finger protein 9

[1] Tris® is a registered trademark of Rohm & Haas Company, Philadelphia, PA, U.S.A.
[2] Tween™20 is a trademark of ICI Americas Inc., Wilmington, DE, U.S.A.

LIST OF EQUATIONS

LIST OF FIGURES

Figure no. **Title of Figure** **Page no.**

LIST OF GRAPHS

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to extend my sincere gratitude towards the following individuals and institutions:

The **individuals** who participated in this project. Their contribution to this project was invaluable and without their involvement this study would have been impossible.

My supervisor, **Prof. Antonel Olckers**, for sharing her platform, wisdom, experience, insight, vision and knowledge with me and for making a tremendous impact on my scientific career. I would also like to thank her for her never-ending support and her utter commitment towards the education of her students at an exceptionally high international level. **Dr. Clara Schutte**, my co-supervisor, for providing support and encouragement and for sharing her clinical expertise with us. The CGR team: PhD colleagues: **Marco Alessandrini, Tumi Semete** and **Wayne Towers;** MSc colleagues: **Christa Mouton, Tharina van Brummelen** and **Madeleine Wessels;** and **Desire Hart** for their encouragement and support. The spirit, enthusiasm, sense of humour and laughter in the team were uplifting each and every day. I am grateful for all your assistance with the clinics. **Wayne Towers** for the proof-reading of my thesis. **Christa Mouton** for culturing of the immortal lymphoblastoid cell lines and for the preparation of the agarose plugs for the Khoi-San samples included in this study. **Martha Sebogoli** for ensuring a clean work place every week.

The Centre for Genome Research, Potchefstroom University for Christian Higher Education for providing the infrastructure in which I could complete this study and for financial assistance in the form of a post graduate bursary. In particular, I would like to thank the Dean of the Faculty of Natural Sciences, **Prof. Daan van Wyk**, and the Vice Principal (Academic), **Prof. Frikkie van Niekerk,** for their support. It is great to study in an environment where the top management are really interested and where the Dean makes the effort to get to know the students. I would also like to thank the **staff** of the **Ferdinand Postma Library** for assistance.

DNAbiotec (Pty) Ltd for the infrastructure, equipment and financial support which enabled me to complete this study. **Paul Olckers** for his interest in the education of students. I am grateful for his support, encouragement, valuable input and for all his help with the maintenance of the PFGE system. **Fifth Dimension Technologies** and **Createk** for the use of their data projector and facilities during the course of my study for numerous presentations and meetings.

Prof. Doug Wallace, for providing the facility for the preparation of the cell cultures and the agarose plugs of the Khoi-San samples. **Prof. Louise Warnich** for allowing us to utilise their PFGE system. **Biochemistry**, PUCHE for the utilisation of their hybridisation oven. **Prof. Silvére van der Maarel** and **Prof. Rune Frants** for the kind donation of the p13E-11 probe and the original Southern blot protocol. **Dr. Richard Lemmers** for all his advice with the preparation of the agarose plugs. **Sr. Chrissie Lessing, Dr. Peter Schwarz, Jutta** and **Astrid** for assisting in the collection of blood samples at the various clinics. **Nutrition,** PUCHE for providing the facilities to host one of the clinics.

Elsevier (The Lancet) and **C.A. Sewry** for the permission to reprint Figure 3 (The Lancet, 359, 687-695, 2002), and **The BMJ Publishing group** and **Louise Anderson** for the permission to reprint Figure 2 (BMJ, 317, 991-995, 1998), as illustrated on pages 7, 8 and 12 of this thesis respectively. **Bio-Rad® Laboratories,** especially **Petr Samanek** for the assistance with the Quantity One® v 4.4.1 program and for the permission to reprint Figure 1-1 presented in Figure 5.2 on page 87. **Dr. Zhi-Ying Wu** (Department of Neurology, First Affiliated Hospital, Fujian Medical University, China), for the translation of two articles (Su et al., 2003; Wang et al., 2003) published in Chinese.

My **parents** for loving and believing in me, during this project in particular, and for their patience, genuine encouragement, never ending support and financial assistance. My supportive extended family, **Ronel, Wilna, Marlize, Kayla, Riaan** and **Chris**. Without your encouragement and support none of this would have been possible. I know that you sacrificed some of your dreams to enable me to realise mine. You are truly the dream family everyone wish they had. **Hermien Kekana** for her support.

I am especially grateful for the ability and endurance I have received from **God** which enabled me to complete this study.

CHAPTER ONE

INTRODUCTION

At least 1 in 3,000 individuals are affected by an inherited neuromuscular disorder (Emery, 1998). The muscular dystrophies, which represent a considerable proportion of this group, are defined as a group of genetic disorders with progressive muscle wasting and weakness. The pathogenesis of many, but unfortunately not all, of the muscular dystrophies has been elucidated. In Chapter two several of the more common muscular dystrophies are discussed. The mode of inheritance, gene location and genes involved in most of these muscular dystrophies are presented in Appendix B.

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 the progression of the disorder. A detailed clinical description of the FSHD phenotype is presented in Chapter three. The diagnostic criteria for FSHD as defined by the International FSHD Consortium are listed in Appendix C.

FSHD is also referred to as Landouzy-Déjérine disorder after the two physicians, who described it in 1884 (Landouzy and Déjérine, 1884). However, during the past 148 years several different names for this disorder were utilised by various authors, as indicated in Appendix A. Duchenne de Boulogne was, however, the first to report the clinical description of FSHD in the middle 1800's, stating its myopathic nature and pattern of inheritance (Kazakov *et al.*, 1974).

FSHD was first identified by Duchenne de Boulogne in 1848 under the name 'progressive muscular atrophy of childhood', and Erb described a 'juvenile shoulder-girdle' type muscular dystrophy in 1886 (Kazakov *et al.*, 1974). A historical discussion arose in the late nineteenth century between Erb and Landouzy-Déjérine regarding the first report of "facio-scapulo-humeral muscular dystrophy" (FSHD). 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 (Kazakov *et al.*, 1974). The patients described by Landouzy and Déjérine corresponded to those described by Erb, but a new term, "FSH type" was used by these two authors to describe the observed phenotype. The forms of muscular dystrophy described by Duchenne and Erb were thus not distinct from the "FSH type", but were rather included in this phenotype. After several decades of debate, the first report of FSHD was eventually 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 via 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 was 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 was identified (Van Deutekom *et al.*, 1993). The sequence of the D4Z4 locus and its homologue on chromosome 10q26 was determined, as presented in Appendices D and E. The genetic aspects of FSHD are presented in Chapter four, including the complex nature of this locus on chromosome 4q35, and several of the epigenetic factors that influence this complex phenotype.

The study presented in this thesis is the first extensive molecular study to characterise the D4Z4 locus in the Black South African and the Khoi-San populations. In Chapter five the protocols utilised to fulfil the objectives outlined in Chapter four are described. The results obtained in this study are presented and discussed in Chapter six. Appendices F and G contain supportive results obtained in this study. Conclusions drawn from the results obtained in this study are presented in Chapter seven. The consensus neighbour-joining tree of 104 human mtDNA complete sequences is presented in Appendix H. Conferences and meetings at which research was presented during the period of this study are listed in Appendix I.

CHAPTER TWO

THE MUSCULAR DYSTROPHIES

The word 'dystrophy' was compiled from two Greek terms, *dys* meaning abnormal or faulty; and *trophe*, meaning food or nourishment (Emery, 2000). Muscular dystrophy therefore implies that the nourishment of the muscle is defective, and hence the disorder due to faulty nutrition of the muscles. The term 'muscular dystrophy' was therefore observed to be misleading. Muscular dystrophy is not a single disorder but is a large and heterogeneous group of inherited disorders. The one feature that the muscular dystrophies share as a group is progressive muscle wasting and weakness, which is generally symmetrical.

The muscular dystrophies were classified in 1954 by Walton and Nattrass into three main groups, Duchenne, Facioscapulohumeral, and Limb girdle, based on the mode of inheritance and the distribution of the muscle groups that are predominantly affected. These authors also distinguished between three relatively uncommon, but clinically and genetically distinctive forms of muscular dystrophy, namely, Distal, Oculopharyngeal, and Congenital muscular dystrophy. Another relatively uncommon, but also distinctive disorder, Emery-Dreifuss muscular dystrophy, was characterised in the mid-nineteen sixties (Emery, 2000).

Immunohistochemical techniques enabled the identification of specific deficiencies of various membrane proteins resulting in differentiation between the different types of muscular dystrophies (Table 2.1). 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 (Cohn and Campbell, 2000; Bushby and Beckmann, 2003).

Table 2.1: Subcellular localisation of proteins involved in different types of muscular dystrophy

Subcellular localisation	Protein involved	Type of muscular dystrophy
Nuclear membrane	Emerin	X-linked EDMD
	Lamin A/C	EDMD-AD

continue ...

Table 2.1: continue ...

Subcellular localisation	Protein involved	Type of muscular dystrophy
Cytosol	Calpain-3	LGMD2A
	Fukutin (?)[a]	Fukuyama congenital muscular dystrophy
Cytoskeleton	Dystrophin	DMD
	Dystrophin	BMD
	Telethonin	LGMD2G
Sarcolemma	α- sarcoglycan (adhalin)	LGMD2D
	β- sarcoglycan	LGMD2E
	γ- sarcoglycan	LGMD2C
	δ- sarcoglycan	LGMD2F
	Caveolin-3	LGMD1C
	Dysferlin	LGMD2B
	Dysferlin	Miyoshi myopathy
Extracellular matrix	Laminin α2	Congenital muscular dystrophy

a = the exact localisation of fukutin has not yet been determined. Adapted from Cohn and Campbell, 2000; Bushby and Beckmann, 2003.

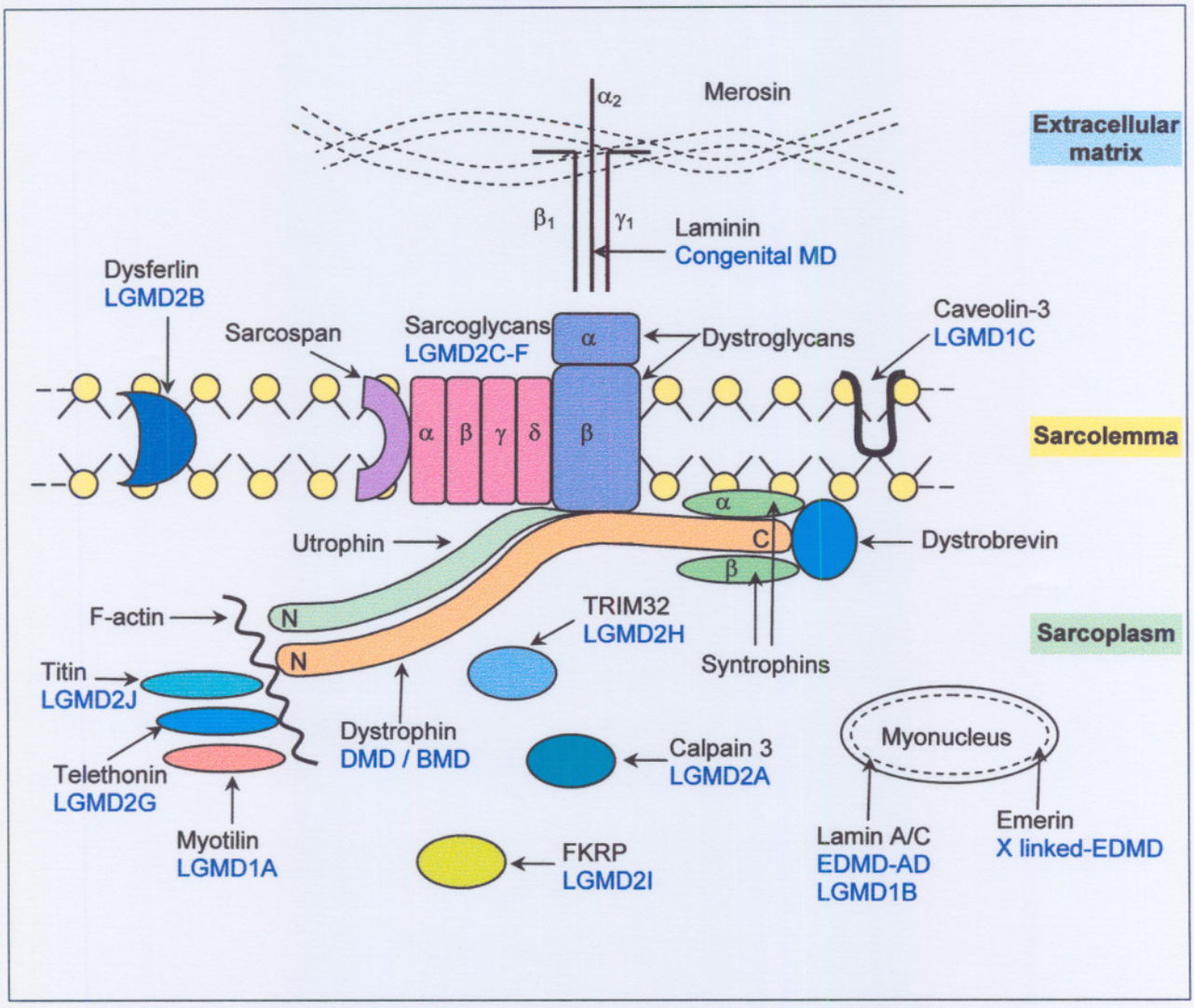
2.1 HISTOCHEMICAL AND BIOCHEMICAL ASPECTS OF THE MUSCULAR DYSTROPHIES

The dystrophin-glycoprotein complex (DGC) has been characterised biochemically and is divided into several subcomplexes that include dystroglycans, sarcoglycans, syntrophins, dystrobrevin and sarcospan, as illustrated in Figure 2.1 (McNally *et al.*, 1998; Cohn and Campbell, 2000). The DGC forms the link between the cytoplasmic actin, the membrane and the extracellular matrix of muscle.

Five sarcoglycans have been identified to date and are classified according to their molecular mass: 50 kDa (α) [Online Mendelian Inheritance in Man™ (OMIM™) 600119], also known as adhalin, 43 kDa (β) [OMIM 600900], 35 kDa (γ) [OMIM 253700], 35 kDa (δ) [OMIM 601411] and 47 kDa (ε), as illustrated in Figure 2.1 (Cohn and Campbell, 2000). The first four sarcoglycans (α to δ) are associated with muscular dystrophies. The dystroglycans consist of a 156 kDA laminin binding subunit (α) and a 43 kDa transmembrane subunit (β) and serve as links between laminin-2 and dystrophin, in that α-sarcoglycan binds to laminin-2 and β-sarcoglycan binds to the C-terminus of dystrophin (Figure 2.1). Syntrophin is a 59 kDa cytoplasmic protein, which binds directly to the C-terminal region of dystrophin (McNally *et al.*, 1998). Dystrophin binds F-actin and interacts with the DGC to form the link between the extracellular matrix (endomysium) and intracellular F-actin as illustrated in Figure 2.1 (McNally *et al.*, 1998; Cohn and Campbell, 2000; Reilly and Hanna, 2002). The DGC 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 initiate a cascade of events that result in muscle weakness (Emery, 1998;

Mak and Ho, 2001). However, the precise mechanism by which the absence of these proteins results in muscle weakness remains unresolved.

Figure 2.1: Muscle membrane proteins



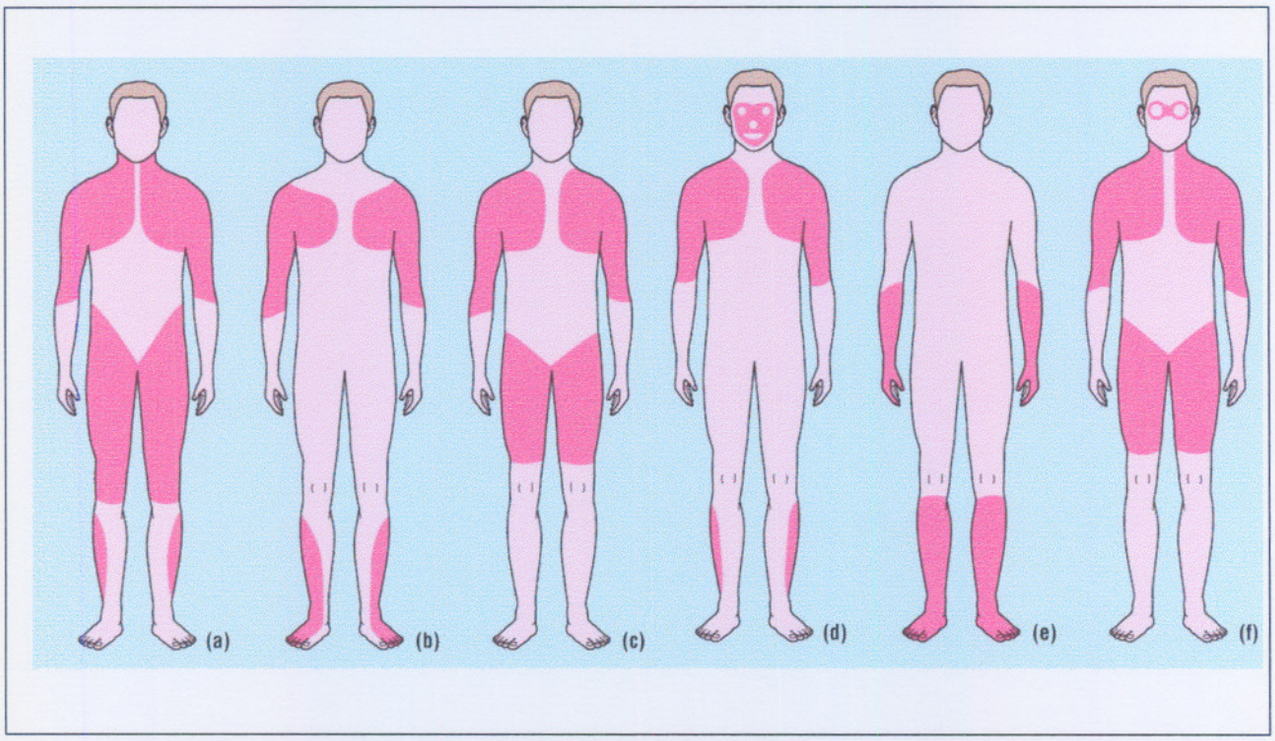
Adapted from Emery, (1998) and Zatz *et al.* (2003).

2.2 CLINICAL AND MOLECULAR ASPECTS OF THE MUSCULAR DYSTROPHIES

The muscular dystrophies can further be distinguished on both the clinical and molecular levels, since the mode of inheritance, gene location and genes for most of these muscular dystrophies have been identified (Appendix B). To highlight the striking differences but also the similarities that complicate clinical diagnoses between the muscular dystrophies, the distribution of muscles groups that are predominantly affected in six of the muscular dystrophies are presented in Figure 2.2 and are discussed in the subsequent paragraphs. Congenital muscular dystrophy, although not presented in Figure 2.2, will be discussed in

addition to the six muscular dystrophies indicated, as this form of muscular dystrophy, although rare, does form part of the muscular dystrophy group of disorders.

Figure 2.2: Distribution of muscle groups predominantly affected in various muscular dystrophies



Different types of muscular dystrophies where the affected muscles are indicated by dark pink: a = Duchenne and Becker; b = Emery-Dreifuss; c = Limb girdle; d = Facioscapulohumeral; e = Distal; f = Oculopharyngeal. Adapted from Emery, (1998).

2.2.1 DUCHENNE MUSCULAR DYSTROPHY

Duchenne muscular dystrophy (DMD) [OMIM 310200] was named after the French neurologist Duchenne de Boulogne, who described the disorder in several publications from 1861 to 1868 (Emery, 2000). There is, however, as so often in the history of medicine, some disagreement as to who actually described this disorder first. Recent historical research has in fact revealed that an English physician, Edward Meryon, actually described the disorder in great detail several years before the report of Duchenne. However, DMD is so well established, that it would be unlikely that the name of this disorder will ever change.

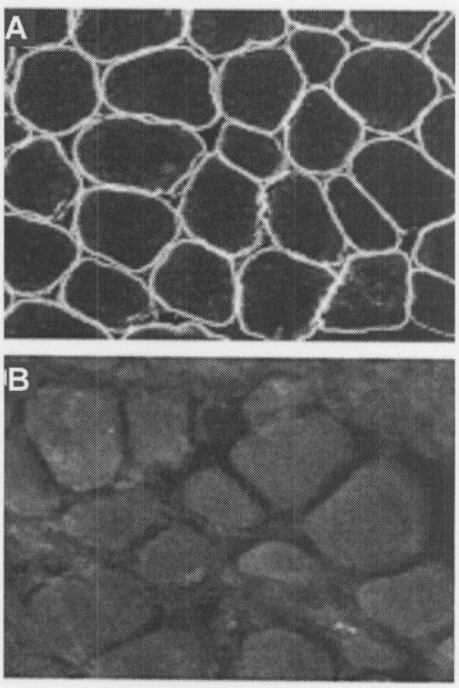
DMD is the most common form of muscular dystrophy, affecting circa (ca.) 1 in 3000 individuals, and is inherited as an X-linked recessive disorder and therefore predominantly affects boys (Ray *et al.*, 1985; OMIM, 2003a). Disease onset is typically before the age of three with individuals becoming wheel-chair dependent by 12 and are generally deceased by the age of 20 (Emery, 2002a; Wagner, 2002). The most distinctive feature of DMD is

proximal muscle weakness with characteristic pseudo-hypertrophy of the calves and individuals are generally observed to be walking on their toes (Emery, 1998; OMIM, 2002a). Due to the weakness and atrophy of the muscles of the pelvis, individuals affected by DMD are also observed to have a waddling gait and a pelvis that is tilted forward. To compensate for the weakened pelvic muscles and to retain the upright position when standing, affected individuals are generally observed to push their abdomen forward and their shoulders backwards, referred to as lordosis (Emery, 2000). The hip and shoulder muscles are also affected, but the eye muscles are always spared and chewing and swallowing are unaffected (Emery, 1998; Wagner, 2002).

Contractures and talipes, characterised by the sole of the foot turning inwards, generally develops in individuals being wheelchair dependant (Emery, 2000). However, the most serious complication of prolonged sitting in a wheelchair is that of scoliosis, which can result in serious problems with breathing and chest infections (Emery, 2000; OMIM, 2003a). The prevention of scoliosis is therefore one of the main challenges in the treatment of affected individuals who become confined to a wheelchair, since respiratory problems are the main cause of death in these individuals. Myocardial involvement, by the age of 6 years, has also been observed in a high percentage of individuals affected with DMD (OMIM, 2003a).

DMD is caused by mutations, generally large deletions or duplications in one, or many of the 79 exons, disrupting the reading frame in the dystrophin gene (2.5 Mb), that was mapped to Xp21.2 (Kunkel *et al.*, 1985; Ray *et al.*, 1985; OMIM, 2003b). This disorder was therefore the first inherited disorder in which the causative gene was located by linkage analysis. Dystrophin is a component of the multi-protein complex linking the cytoskeleton of the muscle fibre to the extracellular matrix, as discussed in paragraph 2.1 and illustrated in Figure 2.1, and is absent or abnormal in biopsies from muscles of DMD patients, as illustrated in Figure 2.3.

Figure 2.3: Frozen section of a muscle biopsy specimen of an individual with DMD



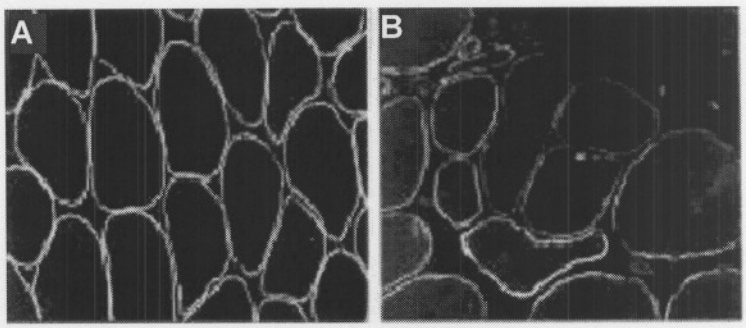
A = control individual, B = individual with DMD. The specimens were stained with labelled antibodies to dystrophin. Reprinted with permission from Elsevier, The Lancet and C.A. Sewry. (Emery, 2002b)

2.2.2 BECKER MUSCULAR DYSTROPHY

Becker muscular dystrophy (BMD) [OMIM 300376] was named after Peter Emil Becker, who first described and distinguished this disorder from the other muscular dystrophies in the mid-1950s (Emery, 2000). The distribution of muscle wasting and weakness is similar to that of DMD, but the disorder is more benign and has a frequency of 1 in 20,000. Individuals affected by BMD are usually affected in their twenties or thirties and generally have a normal life span (OMIM, 2003c). As in DMD, pseudo-hypertrophy of the calf muscles with toe walking, a waddling gait, and the development of lordosis is common in individuals affected with BMD (Emery, 1998; Emery, 2000; Wagner, 2002).

As for DMD, BMD is also caused by mutations in the dystrophin gene, however, in BMD dystrophin is only reduced in amount, as illustrated in Figure 2.4, or abnormal in size, compared to the complete absence in DMD (Wagner, 2002; OMIM, 2003b). This is due to the mutations not disrupting the reading frame, but rather resulting in portions of the protein being deleted. DMD and BMD are therefore also referred to as the dystrophin associated muscular dystrophies (OMIM, 2003c).

Figure 2.4: Frozen section of a muscle biopsy specimen of an individual with BMD



A = control individual, B = individual with BMD. The specimens were stained with labelled antibodies to dystrophin. Reprinted with permission from the BMJ Publishing group, and Louise Anderson. (Emery, 1998)

2.2.3 EMERY-DREIFUSS MUSCULAR DYSTROPHY

Dreifuss and Emery described Emery-Dreifuss muscular dystrophy (EDMD) [OMIM 310300, 181350] in the middle 1960's (OMIM, 2003d, OMIM, 2003e). Even though EDMD is an uncommon type of dystrophy, it has characteristic and distinctive symptoms and early recognition with subsequent treatment can be life saving. A significant feature is cardiomyopathy, generally presenting as an atrioventricular block (heart block), which results in an abnormal slowing of the heart rate (Cohn and Campbell, 2000; Emery, 2000; Bushby *et al.*, 2003). Progressive muscle wasting and weakness with a humeroperoneal distribution, i.e. weakness of the shoulder, upper arm, anterior tibial and peroneal muscles

of the lower legs, occurs early in the course of the disorder (Emery, 2000; Emery, 2002a; Emery, 2002b; Wagner, 2002).

Another distinguishing feature is the presence of muscle contractures even before any significant degree of muscle weakness (Emery, 2000). The muscle contractures predominantly affect three regions, the Achilles tendons (heel cords), resulting in affected individuals walking on their toes; the elbows, preventing the full extension of the elbow; and the muscle at the back of the neck (the post cervical muscles), causing difficulty in the forward bending of the neck (Emery, 2000; Emery, 2002b). Muscles of the lower extremities, i.e., the proximal limb-girdle musculature, are usually 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. There is never any calf enlargement, pseudo-hypertrophy or involvement of the nervous system.

EDMD is inherited as an X-linked recessive disorder and is caused by a mutation in the *STA* gene located on chromosome Xq28, encoding the 254 amino acid nuclear membrane protein, emerin (Emery, 2000; Emery, 2002b; OMIM, 2003d, OMIM, 2003f). Emerin is localised to the inner nuclear membrane and plays a role in membrane anchorage to the cytoskeleton (Cohn and Campbell, 2000). There is a complete absence of emerin in the muscle of most individuals with EDMD. An autosomal dominant form of EDMD also exists, clinically very similar to the X-linked form, which results from mutations in the lamin A/C (*LMNA*) gene, located on chromosome 1q21 (Emery, 2000; OMIM, 2003e, OMIM, 2003g). This gene encodes lamins A and C, two components of the nuclear lamina (a fibrous layer on the nucleoplasmic side of the inner nuclear membrane). These lamins have been observed to interact with chromatin and lamina-associated proteins and emerin (OMIM, 2003g).

2.2.4 LIMB-GIRDLE MUSCULAR DYSTROPHY

The limb-girdle muscular dystrophies (LGMD) are a clinically and genetically heterogeneous group of disorders. The proximal limb muscles, pelvic and shoulder girdle muscles are generally affected (Bushby, 1999). However, this type of muscle weakness may be due to several factors, of which congenital myopathies, spinal muscular atrophies, polymyositis, certain infections, and some drugs, such as steroids are a few. It is for this reason that LGMD was and still is considered as a collection of disorders with unknown or unspecified pathogenesis.

The LGMDs can be divided into two main groups according to inheritance, i.e., autosomal dominant, LGMD type 1 (LGMD1) versus autosomal recessive, LGMD type 2 (LGMD2), and disease severity (Bushby, 1999). Onset of the disorder is generally after childhood in the dominantly inherited forms and the symptoms tend to be milder in severity as opposed to the autosomal recessive forms with onset in childhood which tend to be more severe. At least five dominant (LGMD1A, LGMD1B, LGMD1C, LGMD1D and LGMD1E) and ten recessive (LGMD2A, LGMD2B, LGMD2C, LGMD2D, LGMD2E, LGMD2F, LGMD2G, LGMD2H, LGMD2I and LGMD2J) sub-types have been identified (Tonini *et al.*, 2002; Neuromuscular Disorders: gene location, 2003). Each of the different forms has only been described in one or a few families world wide (Bushby, 1999; Wagner, 2002; Bushby and Beckmann, 2003).

2.2.4.1 Limb-girdle muscular dystrophy type 1

LGMD1A (OMIM 159000) is characterised by proximal muscle weakness at a mean age of 27 years with progressing distal weakness (Bushby, 1999; OMIM, 2003h). A distinctive nasal, dysarthric pattern of speech has been noted in ca. half of the affected individuals (Wagner, 2002). Tightened heel cords and reduced knee and elbow deep tendon reflexes and elevated creatine kinase (CK) level, up to nine fold, are also generally observed. LGMD1A is caused by a mutation in the gene mapped to chromosome 5q31, encoding the sarcomeric protein myotilin [*MYOT*] (Zatz *et al.*, 2000).

Symmetric weakness starting in the proximal limb muscles before the age of 20 years and progressing to involve the upper-limb muscles in the third or fourth decade, is observed in LGMD1B (OMIM 159001) [OMIM, 2003i]. Cardiac involvement with atrioventricular conduction disturbance, requiring pacemaker implantation, is observed in ca. 60% of patients (Bushby, 1999; Wagner, 2002). The CK levels of individuals with LGMD1B are normal to moderately elevated. LGMD1B is caused by mutations in the *LMNA* gene located on chromosome 1q11-21, and is therefore allelic to EDMD-AD (Muchir *et al.*, 2000; OMIM, 2003g). This disorder is distinguished from EDMD, by the involvement of the Achilles tendons only occurring later in life (Muchir *et al.*, 2000; Wagner, 2002).

LGMD1C (OMIM 601253) is associated with mutations in the caveolin-3 gene (CAV3) located on chromosome 3p25 (Minetti *et al.*, 1998; OMIM, 2003j, OMIM, 2003k). Disease onset is generally during childhood, generally by the age of 5, with proximal muscle weakness, calf hypertrophy, cramping and dilated cardiomyopathy being some of the clinical symptoms (Minetti *et al.*, 1998; Bushby, 1999; Cohn and Campbell, 2000).

Linkage analysis identified a new locus on chromosome 6q23 to be associated with LGMD1D (OMIM 603511) [OMIM, 2003l]. Progressive proximal leg weakness with or without proximal arm weakness, and absence of ankle deep-tendon reflexes have been observed in individuals with LGMD1D (Beckmann *et al.*, 1999) Cardiac complications are also a significant feature of this disorder (Beckmann *et al.*, 1999; Bushby, 1999).

Onset of LGMD1E is around the third decade and muscle weakness involves proximal upper and lower limb muscles. No evidence of cardiac involvement has been reported (Beckmann *et al.*, 1999). CK levels are mildly elevated and dysphagia has been reported in a single family. LGMD1E has been mapped to chromosome 7q, but the gene and gene product involved still need to be elucidated (Speer *et al.*, 1999; Bushby and Beckmann, 2003).

2.2.4.2 Limb-girdle muscular dystrophy type 2

Most of the LGMD's are inherited as recessive disorders, and may be classified into the sarcoglycanopathies, i.e. LGMD2C to LGMD2F, and the non-sarcoglycanopathies, i.e. LGMD2A, LGMD2B, LGMD2G to LGMD2J, depending on whether the specific LGMD is associated with one of the sarcoglycans (Wagner, 2002).

LGMD2A (OMIM 253600) is the most common of the LGMDs and is caused by ca. 100 distinct mutations, including nonsense, missense and splicing mutations, in the gene encoding a muscle specific protease, calpain 3 (CAPN3), localised to chromosome 15q15.1 (Bushby, 1999; Wagner, 2002; OMIM, 2003m; OMIM, 2003n). LGMD2A is therefore also known as calpainopathy. Onset is generally during childhood, but can occur at a later age (Beckmann *et al.*, 1999). Pelvic girdle involvement is generally first noted, however, some cases with initial shoulder girdle involvement have been observed (Wagner, 2002). CK levels are generally markedly elevated. The muscle weakness has been observed to be asymmetric and spreads from the lower to upper limbs or vice versa within 20 years. The presence of contractures and facial weakness has been observed, but only occurs later in the progression of the disorder (Beckmann *et al.*, 1999).

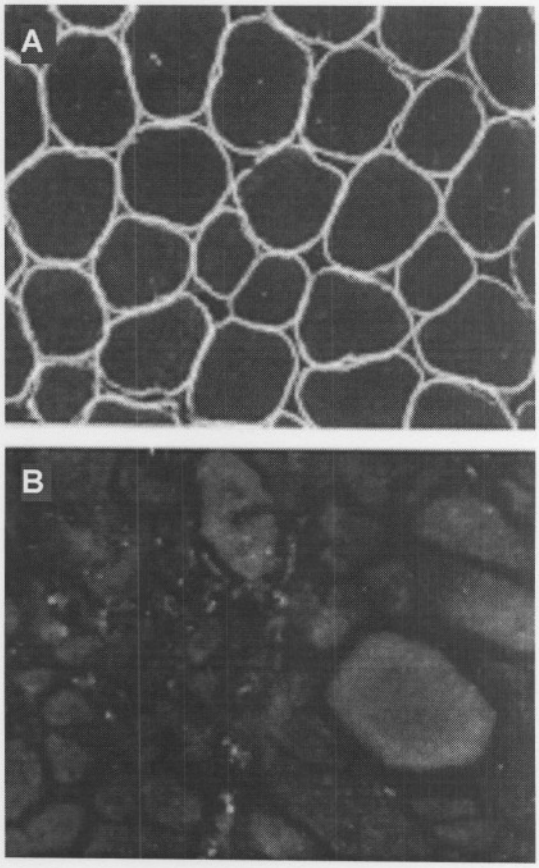
LGMD2B (OMIM 253601) is caused by mutations in the gene encoding the skeletal muscle protein, dysferlin, located on chromosome 2p13.3-p13.1, therefore also known as dysferlinopathy (Bushby, 1999; Bushby and Beckmann, 2003; OMIM, 2003o; OMIM, 2003p). Onset is generally in the late teens with predominantly proximal limb weakness, with slow progression (Cohn and Campbell, 2000; Wagner, 2002). CK levels are generally

markedly elevated. Shoulder girdle involvement and calf hypertrophy have also been observed but are rare (Beckmann *et al.*, 1999; Zatz *et al.*, 2003). Individuals with LGMD2B do not develop cardiomyopathy, respiratory or facial muscle weakness and contractures are almost never present (Beckmann *et al.*, 1999). Miyoshi myopathy has also been mapped to chromosome 2p13.3-p13.1 and is allelic to LGMD2B (Bushby, 1999).

LGMD2C (OMIM 253700) is characterised with disease onset before the age of 5 years, confinement to a wheelchair by the age of 12 years, and death before 20 years of age (Wagner, 2002; OMIM, 2003q). This disorder was previously also known as severe childhood autosomal recessive muscular dystrophy. Pseudo-hypertrophy, cardiac involvement and CK levels which are markedly elevated have also been observed in individuals with LGMD2C (Bushby *et al.*, 2003). Various mutations in the gene mapped to chromosome 13q encoding γ-sarcoglycan (35 kDa) have been observed to be associated with LGMD2C, also known as γ-sarcoglycanopathy (Wagner, 2002; OMIM, 2003q).

LGMD2D (OMIM 600119) is the most common of the sarcoglycanopathies with typical onset in childhood (Wagner, 2002; OMIM, 2003r). This disorder is characterised by early presentation of lower limb-girdle weakness; scapular winging and calf hypertrophy are frequently observed (Wagner, 2002). CK levels are raised to the greatest extent of all the limb girdle muscular dystrophies. Mutations in the gene, mapped to chromosome 17q12-q21.33, encoding adhalin (also known as α-sarcoglycan), is associated with LGMD2D (α-sarcoglycanopathy) [OMIM, 2003s). Adhalin was observed to be completely absent in individuals affected with LGMD2D, as illustrated in Figure 2.5.

Figure 2.5: Frozen section of a muscle biopsy specimen of an individual with LGMD2D



A = control individual, B = individual with LGMD2D. The specimens were stained with labelled antibodies to adhalin (α-sarcoglycan). Reprinted with permission from Elsevier, The Lancet and C.A. Sewry. (Emery, 2002b)

LGMD2E (OMIM 604286) is caused by a mutation in the gene, localised to chromosome 4q12, encoding β-sarcoglycan (43 kDa) [OMIM, 2003t; OMIM, 2003u]. Reduced expression of β-sarcoglycan was observed in the sarcolemma

of individuals with LGMD2E (β-sarcoglycanopathy). A severe phenotype is characterised by early childhood onset with loss of ambulation in the early teens (Wagner, 2002). Upper and lower limb-girdle weakness as well as calf hypertrophy is observed. However, milder phenotypes with onset in adulthood, have also been reported (OMIM, 2003t). Fanin *et al.*, (2003) observed cardiomyopathy in half of the LGMD2E cases they investigated and highlighted the point of careful cardiac monitoring in LGMD2E patients.

Mutations in the gene encoding δ-sarcoglycan, localised to chromosome 5q33, have been implicated in LGMD2F (OMIM 601287) [OMIM, 2003v; OMIM, 2003w]. The phenotype is generally severe, with early childhood onset, loss of ambulation by adolescence and death prior to the third decade (Wagner, 2002). Dilated cardiomyopathy with focal areas of necrosis is a histological hallmark of this disorder (Bushby, 1999; Cohn and Campbell, 2000; Bushby *et al.*, 2003).

LGMD2G (OMIM 601954) is characterised by mildly elevated CK levels with rimmed vacuoles present upon muscle biopsy (OMIM, 2003x). Foot drop, proximal muscle atrophy in the upper limbs and proximal and distal muscle atrophy in the lower limbs was observed in individuals with LGMD2G (Wagner, 2002). No evidence for extraocular or facial muscle weakness or any cardiac involvement was, however, detected. LGMD2G is associated with mutations in the gene localised to chromosome 17q11-12 encoding telethonin, a sarcomeric protein of 19 kDa (Vainzof *et al.*, 2002).

LGMD2H (OMIM 254110) is also known as Hutterite type muscular dystrophy, since it was observed in the Hutterite population of North America, and is associated with mutations in the tripartite-motif-containing gene 32 (TRIM32) located on chromosome 9q31-q34.1 (Wagner, 2002; OMIM, 2003y). TRIM32 (72 kDa) is a putative E3 ubiquitin ligase, expressed in several tissues, including heart and skeletal muscle (Bushby and Beckmann, 2003). E3 ubiquitin ligases generally tag proteins for degradation which in turn leads to the proteins being ligated to ubiquitin and being degraded by the proteasome pathway. A mutation in this type of gene could therefore lead to accumulation of the target protein within a specific tissue. To date there has, however, not been any evidence for protein accumulation in biopsies from patients diagnosed with LGMD2H. LGMD2H is a relatively mild, slowly progressive disorder with onset between 1 and 9 years of age with initial muscle weakness in the quadriceps and pelvic girdle (Wagner, 2002; Bushby and Beckmann, 2003). Facial muscle weakness can occur as the disorder progress.

Mutations in the fukutin-related protein (FKRP) gene localised to chromosome 19q13.3, have been observed in individuals with LGMD2I [OMIM 607155] (Bushby and Beckmann, 2003; OMIM, 2003z; OMIM, 2003aa). The age of onset is very broad, ranging from 6 months to 40 years. Clinical symptoms include predominantly pelvic girdle weakness, muscle hypertrophy and cardiomyopathy. High serum CK levels, of 10 to 50 times normal have also been observed (Driss *et al.*, 2000; Wagner, 2002).

LGMD2J is caused by mutations in the titin gene localised to chromosome 2q31. This disorder is characterised by early onset of proximal muscle weakness with high serum CK levels (Bushby and Beckmann, 2003).

2.2.5 DISTAL MUSCULAR DYSTROPHY

This disorder can be divided into two main groups: the late onset (over 40 years of age) with autosomal dominant inheritance, including Tibial muscular dystrophy (TMD) [OMIM 600334] and autosomal dominant distal myopathy (MPD1), and early onset (less than 30 years of age), with autosomal recessive inheritance, including Miyoshi myopathy (MM) [OMIM 254130], distal myopathy with rimmed vacuoles (DMRV) and hereditary inclusion body myopathy (HIBM) [Emery, 2002b; OMIM, 2003bb, OMIM, 2003cc). Apart from MM (associated with mutations in the gene located on chromosome 2p12-14 encoding dysferlin) and TMD (caused by mutations in the gene encoding titin localised to chromosome 2q31), the genes and gene products of the other distal muscular dystrophies still need to be elucidated (Emery, 2002b, Neuromuscular disorders: gene locations, 2003).

In general these disorders are relatively rare and are associated with wasting and weakness of the distal muscles. Weakness in the hands is present from the beginning and is generally characterised by clumsiness while involvement of the lower leg muscles results in tripping, and making standing on the heels difficult (Emery, 1998; OMIM, 2003bb, OMIM, 2003cc). Individuals affected with this disorder are generally only mildly affected, as it is a slowly progressive disorder and the affected individuals have a normal life expectancy.

2.2.6 OCULOPHARYNGEAL MUSCULAR DYSTROPHY

Oculopharyngeal muscular dystrophy (OPMD) [OMIM 164300] generally presents late in life, i.e. in the thirties or even later (OMIM, 2003dd). The most prominent characteristic of

OPMD is the presence of progressive ptosis of the eyelids (Emery, 1998). Diplopia has also been observed, but it is relatively uncommon. A serious symptom that may also develop as the disorder progress is the presence of dysphagia (Emery, 2002b). Weakness of the muscles of the neck and proximal upper limbs may also be present. CK levels are normal to moderately elevated. A characteristic feature of OPMD is the presence of rimmed vacuoles that are positive for ubiquitin upon muscle biopsy. OPMD, an autosomal dominant disorder, 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 gene (PABP2) located on chromosome 14 (Emery, 1998; OMIM, 2003ee).

2.2.7 CONGENITAL MUSCULAR DYSTROPHY

The congenital muscular dystrophies (CMD) are a heterogeneous group of autosomal recessively inherited disorders (Wagner, 2002). Children with a congenital muscular dystrophy present with hypotonia and weakness at birth or within the first few months of life and they are generally unable to walk. The muscle weakness is generally non-progressive, but joint contractures might result in immobility (Emery, 1998; Emery, 2002b). Several different forms, some with or without mental retardation have been recognised, of which a few are discussed in the following paragraphs.

Mutations in LAMA2, localised to chromosome 6q2, encoding for the α2 chain of laminin-2, are associated with 50% of individuals diagnosed with classic congenital muscular dystrophy (MDC1A), also known as merosin deficient CMD [OMIM 156225] (Emery, 1998; Cohn and Campbell, 2000; OMIM, 2003ff). Onset is at birth with severe muscle weakness and arthrogryposis, with a 10 fold elevation in CK levels. The majority of patients are never able to walk (Wagner, 2002).

Congenital muscular dystrophy (MDC1C) is characterised on a clinical level by onset within the first few weeks of life with severe weakness and wasting of the shoulder girdle muscles and the leg muscles. Hypertrophy and severe respiratory insufficiency have also been reported (Wagner, 2002). The CK levels of individuals with MDC1C are markedly increased. Children with MDC1C are never able to walk and are generally deceased by the second decade of life due to respiratory failure. The brain structure and intelligence of these individuals are normal. MDC1C is allelic to LGMD2I in that it is also caused by mutations in the gene encoding FKRP localised to chromosome 19q13.3 (Wagner, 2002).

Fukuyama muscular dystrophy, one of the congenital muscular dystrophies, is the second most common form of muscular dystrophy after Duchenne muscular dystrophy in Japan, but is, however, rare outside this country (Wagner, 2002). The onset is in infancy with calf pseudo-hypertrophy and the children are generally severely mentally challenged with cases of epilepsy (Muntoni and Guicheney, 2002). Weakness of the facial, neck, proximal arm and distal leg muscles is also noted (Wagner, 2002). This disorder is caused by mutations in the fukutin gene localised to chromosome 9q31.

Muscle-eye-brain disease is characterised by ocular abnormalities, brain malformations and onset during infancy. Individuals affected with this disorder are severely mentally challenged (Muntoni and Guicheney, 2002). Mutations in the gene, localised to chromosome 1p3, encoding the O-mannose β-1-2-N-acetylglucosaminyltransferase (POMGnT1) have been observed to be associated with muscle-eye-brain disease (Wagner, 2002).

Rigid spine muscular dystrophy is a relatively mild congenital disorder and is characterised by contractures of the spinal extensors causing severe limitation in movement. Early respiratory insufficiency generally requires nocturnal ventilatory support (Wagner, 2002). Mutations in the gene encoding selenoprotein N1 (SEPN1), localised to chromosome 1p36, have been associated with this disorder (Neuromuscular disorders: gene location, 2003).

2.2.8 FACIOSCAPULOHUMERAL MUSCULAR DYSTROPHY

Facioscapulohumeral muscular dystrophy (FSHD) [OMIM 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 (OMIM, 2003gg). FSHD is the exception to the general rule in that weakness in most dystrophies is generally symmetrical, but in FSHD muscle weakness it is often asymmetrical.

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 chapters three and four (Emery, 1998; OMIM, 2003gg).

CHAPTER THREE

CLINICAL ASPECTS OF FSHD

Facioscapulohumeral muscular dystrophy is the third most common hereditary disorder of muscle after Duchenne and myotonic dystrophy, with 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; Padberg, 1982; 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 and also due to the high consanguinity resulting in a possible founder effect for FSHD in this population. FSHD is a highly variable disorder with weakness occurring from infancy to late life, but generally in the second decade of life.

3.1 PRESENTING SYMPTOMS

This disorder is characterised at onset by asymmetrical weakness and atrophy of the facial (paragraph 3.1.1) and shoulder girdle (paragraph 3.1.2) muscles with sparing of bulbar extraocular and respiratory muscles (Padberg, 1982; Tawil *et al.*, 1998). Involvement of the upper arm (paragraph 3.1.3), abdominal (paragraph 3.1.4), foot extensor and pelvic girdle (paragraph 3.1.5) muscles has also been observed. The degree to which the muscles can be affected varies dramatically, which forms the basis of the extremely heterogeneous nature of the FSHD phenotype (paragraph 3.2). Extramuscular involvement (paragraph 3.2.3) as well as central nervous system involvement (paragraph 3.2.2.1) has, however, also been observed as the disorder progresses.

3.1.1 FACIAL MUSCLES

The early stages of the disorder, in the vast majority of patients, is generally characterised by facial weakness. The zygomaticus, orbicularis oculi and orbicularis oris muscles are generally affected. Weakness of the zygomaticus muscles results in the inability to raise the corners of the mouth and thus results in a transverse smile. The orbicularis oris allows pursing of the lips, whistling and retaining of air under pressure. Speech may even

become unclear if the facial weakness is severe (Emery, 2000). Orbicularis oculi muscle weakness results in the inability of the patients to bury their eyelashes completely when forced closed. Individuals are often observed to sleep with open eyes, and blinking is slowed and incomplete. Other facial muscles may also become involved, resulting in an unlined forehead and a smooth, expressionless or myopathic face (Padberg, 1982). The presence of a typical myopathic face (facies myopathica) was also observed by Landouzy and Déjérine in 1885.

Although facial weakness is present in the majority of cases, in the early course of the disorder it is, however, not obligatory in the diagnosis of FSHD. Felice *et al.* (2000) examined seventeen unrelated patients diagnosed with facial-sparing scapular myopathy. DNA testing was performed for fourteen of the seventeen patients, of which ten (71%) harboured a *Bln* I resistant deletion fragment, which defines FSHD on the molecular level (see paragraph 4.3.1). In 2001, Felice and Moore reported three individuals with no visible facial weakness. The first individual, at first diagnosed with facial-sparing scapular myopathy, had symptoms of scapular winging, humeral and quadriceps atrophy and a waddling-type gait. The second individual, diagnosed with limb-girdle muscular dystrophy, 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 was no facial, scapular winging, or other characteristic features of FSHD. The last individual displayed mild right trapezius atrophy, asymmetric scapular winging, 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. Facial-sparing scapulohumeral dystrophy was further diagnosed in three molecular confirmed FSHD patients (Krasnianski *et al.*, 2003a). Uncini *et al.* (2002) also reported two patients without any facial weakness. The diagnosis of only one of the patients could, however, be confirmed on the molecular level.

Scapuloperoneal and scapulohumeral dystrophies, which are associated with minimal or complete absence of facial weakness, are clinically similar to FSHD and it is therefore difficult and sometimes even impossible to distinguish them from FSHD on the clinical level (Jardine *et al.*, 1994a; Tawil *et al.*, 1995). These disorders may even include and represent milder forms of FSHD since there is evidence for involvement of the same genetic locus (Jardine *et al.*, 1994a). Identification of the FSHD gene may shed light on whether the disorders are indeed different on the genetic and molecular levels.

3.1.2 SHOULDER GIRDLE MUSCLES

One of the earliest symptoms in FSHD, together with the facial muscle weakness, is the gradual loss of fixation of the scapula due to weakness of the muscles which stabilise the scapula to the torso, i.e., the lower part of the trapezius, the rhomboids and the serratus anterior muscles (Padberg, 1982). Weakness of these muscles results in the scapulae to rotate slightly laterally, and to move upwardly, laterally and anteriorly over the thorax. The clavicles lose their normal upward slope, rotate anteriorly and may ultimately even slope downwards and together with the positional change of the scapulae contribute to the development of drooping shoulders. When the scapular fixation, and especially the serratus anterior function worsens further, elevation of the arm above shoulder level becomes impossible. The abduction of the arms results in the scapulae rising upward over the back into the normal location of the trapezius muscles, also known as high riding of the scapulae (Padberg, 1982). Weakness of the latissimus dorsi and the sternocostal part of the pectoralis muscles may also occur resulting in a flattened outline of the anterior thoracic wall (Tyler and Stephens, 1950, Padberg, 1982). The direction of the axillary crease is therefore changed from being vertical to being more horizontal.

3.1.3 UPPER ARM MUSCLES

The time of onset between shoulder girdle weakness and upper arm weakness may be quite variable (Padberg ,1982). It has been observed that 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 (Padberg, 1982).

3.1.4 ABDOMINAL MUSCLES

Padberg (1982) reported abdominal muscle weakness in 58% of patients investigated. It was observed that the majority of individuals had abdominal muscle weakness before any foot extensor involvement, therefore indicating that abdominal muscle weakness can be an early symptom in FSHD. Weakness of the abdominal muscles 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 be more severe when an individual was wheelchair dependent.

3.1.5 LOWER EXTREMITIES AND PELVIC GIRDLE MUSCLES

Landouzy and Déjérine (1885), Tyler and Stephens (1950) and Chyatte *et al.* (1966) observed the early weakening of the anterior tibial muscles in FSHD patients. Weakness of the foot extensors results in a steppage gait and interferes with walking, resulting in an inability to run (Padberg, 1982). The observed 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 and patients therefore tend to trip easily over small objects. Pelvic girdle weakness results in a waddling gait characterised by a distinctive duck-like walk and gradually rising from a chair or climbing stairs becomes increasingly difficult.

3.1.6 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).

3.2 CLINICAL HETEROGENEITY IN THE FSHD PHENOTYPE

The clinical diagnosis of the FSHD phenotype is complex due to extreme variability, both between and within families, 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). Tonini *et al.* (2004) reported that 20% of individuals harbouring a mutation are asymptomatic.

Diagnostic criteria for FSHD have been defined by the International FSHD Consortium and are listed 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.

As mentioned above extra-ocular muscle weakness resulting in ptosis, pharyngeal muscle weakness and involvement of the lingual muscle is considered to be exclusion criteria of FSHD (Appendix C). Korf *et al.* (1985), however, reported four patients with tongue abnormalities, such as atrophic changes and movement disturbances. No genetic analyses were, however, performed to confirm the diagnosis of these individuals. Yamanaka *et al.* (2001) also reported tongue atrophy in seven Japanese patients. These patients all belonged to a group of early-onset FSHD patients (see paragraph 3.2.2) with small *Eco* RI deletion fragments, 10 to 17 kilobase pairs (kb). In 1998, Miura *et al.* also reported the presence of tongue atrophy in a female sporadic case with early-onset FSHD, harbouring a 10 kb *Eco* RI deletion fragment. There is thus evidence identifying lingual muscle involvement in FSHD, indicating that this clinical aspect should not be used as an exclusion criterion of FSHD. The occurrence of tongue abnormalities in FSHD, especially in early onset FSHD, should, however, be investigated in a larger sample size.

Van der Kooi *et al.* (2000) also reported on the extreme clinical heterogeneity observed in FSHD after the examination of six sporadic patients with symptoms and signs that initially caused confusion upon clinical diagnosis. One patient presented with thigh weakness, three patients with foot extensor weakness, one patient with calf muscle weakness resulting in an inability to walk on his toes, and the other patient with mild shoulder symptoms, such as tiredness and muscle pain. No visible facial weakness was observed upon initial examination. An expert physical examination, however, revealed the presence of an abnormality in the facial expressions of the patients and abnormal shoulder posture or scapular winging when lifting the arms. FSHD was subsequently confirmed via molecular genetic analysis.

Felice and Moore (2001) reported a patient who presented with progressive bilateral foot drop, and experienced difficulties when climbing stairs. Other problems included occasional heart palpitations, late-onset sensorineural hearing loss, cataracts and arthritis. The patient was diagnosed with late-onset autosomal dominant distal myopathy. 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.

McGonigal *et al.* (2002) investigated a 56 year old man diagnosed with myasthenia gravis. The individual had left sided ptosis and was positive for anti-acetylcholine receptor antibody titre. Upon further analysis it was observed that this individual had a 40 year history of progressive right foot drop with recent involvement of the left foot. Mild weakness of facial and neck muscles, winging of the scapulae, a prominent lumbar lordosis and waddling gait was also observed, resulting in the diagnosis of this individual with FSHD. A 28 kb *Eco* RI deletion fragment was detected upon molecular analysis, confirming the clinical diagnosis. This was the first report to associate myasthenia gravis with FSHD.

Krasnianski *et al.* (2003a) examined three patients from a single family with atypical FSHD phenotypes harbouring 20 kb deletion fragments. All three patients presented with ptosis, chronic progressive external ophthalmoplegia, scapular winging and lumbar lordosis. The 50 year old father was further diagnosed with facial weakness, proximal arm and shoulder girdle atrophy and foot drop. However, the 15 year old boy and girl revealed no signs of any facial weakness. The involvement of a mutation in the mitochondrial DNA (mtDNA) causing the external ophthalmoplegia was excluded, since no mtDNA deletions were observed. A muscle biopsy on the 15 year old boy also indicated minimal myopathic changes without any ragged red fibres.

The studies of Korf *et al.* (1985), Miura *et al.* (1998), Van der Kooi *et al.* (2000), Felice and Moore (2001), Yamanaka *et al.* (2001) and McGonigal *et al.* (2002) therefore all confirm the clinical heterogeneity of FSHD. These studies highlight the need for expert clinical examination of FSHD patients to ensure the most accurate clinical diagnosis possible of this clinically heterogeneous disorder. It is also preferable that the same clinician investigate all patients included in a study to ensure consistency in diagnosis of this complex clinical phenotype. Moreover, the need for DNA analysis to confirm the clinical diagnosis was also strengthened and highlighted in the majority of cases.

3.2.1 MONOZYGOTIC TWIN STUDIES

Several authors studied monozygotic twins in order to elucidate the extreme clinical variability observed in FSHD. At least five sets of monozygotic twins affected with FSHD have been described (Tawil *et al.*, 1993a; Tawil *et al.*, 1993b; Griggs *et al.*, 1995; Hsu *et al.*, 1997; Tupler *et al.*, 1998).

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, asymmetric involvement of the facial, scapular, peroneal and abdominal muscles was observed. The individual had no extraocular, laryngeal or pharyngeal weakness. 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 additional molecular 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 phenotypical discordance was probably due to a *de novo* postzygotic mutation after the twinning process.

Two sets of monozygotic twins concordant for FSHD, as well as the monozygotic twins discordant for FSHD, first examined by Tawil in 1993, were re-examined by Griggs *et al.* (1995). The twins concordant for FSHD were equally 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 19.9 kb deletion fragment was present in the affected individual of the other discordant twin pair, but no deletion fragment was present in the other individual of the discordant twin pair. 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 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 may have triggered an inflammatory immune reaction, which contributed to the more severe phenotype.

Hsu *et al.* (1997) also reported a monozygotic twin pair, confirmed via HLA typing and analysis of six other loci. An *Eco* RI deletion fragment of 16 kb in size was observed in both individuals, but was not present in either of their parents. Differences in the severity were also observed between these individuals. One of the twins had more severe facial and shoulder muscle weakness together with an elevated CK level. The authors did not mention any differences in the medical history of these individuals, compared to the patients reported by Tupler *et al.* (1998).

Fitzsimons (1999) cited an important observation reported for monozygotic twins discordant for other disorders, such as type I diabetes and multiple sclerosis (Utz *et al.*, 1993; Kotzin 1993; Abbas *et al.*, 1996; Wilson *et al.*, 1998; Wilson *et al.*, 2000). In a monozygotic twin pair concordant for multiple sclerosis, the same T-cell receptor genes for V-α chains were expressed in response to certain antigens. In a different monozygotic twin pair discordant for multiple sclerosis, it was, however, observed that different T-cell receptor genes were expressed in response to the same antigens (Utz *et al.*, 1993). The differences observed in some FSHD monozygotic twins could therefore be explained by differences in T-cell receptors or other T-cell gene expression during inflammation (e.g. after vaccination during childhood). The exact pathogenesis of the external factor(s), e.g. vaccination, with relation to the severity of the FSHD phenotype, however, still remains to be elucidated.

3.2.2 EARLY-ONSET FSHD

Early-onset FSHD, also known as infantile FSHD, is characterised by onset before the age of five (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 and is the most severe form of this disorder (McGarry *et al.,* 1983; Brouwer *et al.*, 1993). Early facial weakness is visible during the first two years of life and weakness of the shoulder muscles causing scapular winging is visible before the age of ten. Progressive weakness results in the deterioration of the upper-arm and foot-extensor muscles, characterised by foot drop, followed by development of pelvic girdle weakness presenting as lumbar lordosis (Korf *et al.,* 1985; Bailey *et al.,* 1986). Individuals affected with early-onset FSHD are generally wheelchair dependent by the age of nine or ten (McGarry *et al.,* 1983; Brouwer *et al.*, 1993).

McGarry *et al.* (1983) reported death as a result of FSHD. The 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. Okinaga *et al.* (1997) 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 15 kb respectively). Early-onset FSHD does therefore not differ clinically, or genetically, from adult onset FSHD (Brouwer *et al.,* 1995; Okinaga *et al.,* 1997).

Funnell and George (2003), reported the presence of epiphora in a 5½ year old boy, due to lacrimal pump failure. During examination no ptosis and the inability to close his eyes completely, indicating orbicularis oculi muscle weakness were observed. His mother was diagnosed with FSHD and was confined to a wheelchair. The boy had difficulty to rise from a sitting position and could not run. Lumbar lordosis and mild winging of the scapulae was also observed upon clinical examination. The FSHD status of the mother

and the son was confirmed via molecular analysis. Upon re-examination at the age of 6½ high frequency hearing loss was also diagnosed in the boy. This early-onset FSHD case therefore also highlights the need to consider ear and eye involvement in the diagnosis of FSHD (paragraph 3.2.3). The presence of high frequency hearing loss, retinal vascular abnormalities and tongue atrophy in early-onset FSHD patients has also been reported by several other authors (Korf *et al.,* 1985; Matsuzaka *et al.,* 1986; Voit *et al.,* 1986; Brouwer *et al.*, 1994; Brouwer *et al.*, 1995; Miura *et al.,* 1998).

3.2.2.1 Early-onset FSHD with central nervous system involvement

Matsuzaka *et al.* (1986) investigated a 7 year old girl, diagnosed with early-onset FSHD. Facial muscle weakness, an inability to raise her arms above her shoulders and winging of her scapulae was noted at the age of 2 years. Mental impairment was confirmed at the age of 3 years, with an intelligence quotient (IQ) of 35-40. Lumbar lordosis developed at the age of 5 years associated with frequent falls. The authors could, however, not find any significant family history of any neuromuscular disorder in the family.

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. The presence of mental impairment was also observed in four of the ten early-onset FSHD cases examined by Brouwer *et al.* (1995).

Miura *et al.* (1998) reported two unrelated, severely affected sporadic cases diagnosed with early-onset scapulohumeral muscular dystrophy. Both individuals had mental impairment (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.

Sporadic and familial early-onset FSHD cases occur, but the sporadic cases were on average more severely affected than the familial cases (Brouwer *et al.,* 1994; Brouwer

et al., 1995). FSHD patients with a large deletion in the FSHD region have therefore been observed to have a higher chance of developing early-onset FSHD and clinical phenotypes being associated with central nervous system (CNS) abnormalities (Miura *et al.*, 1998; Funakoshi *et al.*, 1998).

3.2.3 EXTRAMUSCULAR INVOLVEMENT

The clinical diagnosis of the FSHD phenotype is further complicated by the involvement of the ears, eyes and heart of individuals affected with FSHD. Extramuscular involvement in FSHD has been reported by several authors and is discussed in the following paragraphs.

3.2.3.1 Sensorineural deafness and retinal vascular abnormalities

The association of sensorineural deafness and retinal vascular abnormalities with FSHD have been described by many authors (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; Brouwer *et al.*, 1995; Padberg *et al.*, 1995a; Fitzsimons 1999). High-frequency hearing loss has been observed in 50 to 64% of the FSHD patients investigated by Brouwer *et al.* (1991) and Padberg *et al.* (1995a). When comparing controls to FSHD patients, Brouwer *et al.* (1991) observed a significant difference in hearing level between 4 kilo hertz (kHz) and 6 kHz in FSHD patients and the hearing loss was also observed to be progressive and, with time, tends to involve lower frequencies of the spectrum.

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 and detachment of the retina due to massive intraretinal and subretinal lipid 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. It was also noted that the severity of the hearing loss varied between individuals and that it was not age dependent (Fitzsimons *et al.*, 1987; Brouwer *et al.*, 1991; Brouwer *et al.*, 1993; Padberg *et al.*, 1995a). The pathogenic mechanism of hearing loss and retinal vasculopathy in FSHD is, however, still unknown.

3.2.3.2 Cardiac muscle involvement

In general, involvement of the cardiac muscles is considered as an exclusion criterium for FSHD (Appendix C; Padberg, 1982). 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 (paragraph 2.2.3). The first evidence of cardiac involvement in FSHD was, however, provided by Stevenson *et al.* in 1990. The clinical diagnosis of FSHD could unfortunately not be confirmed on a molecular level. Laforêt *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 and Stöllberger (2000) investigated the presence of cardiac involvement in several myopathies, such as DMD, BMD, EDMD, FSHD, sarcoglycanopathies, myotonic dystrophies type 1 and 2 and mitochondrial myopathies, and observed cardiac involvement to be present in all of these disorders. Cardiac involvement 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 via echocardiography for possible cardiac involvement (Faustmann *et al.*, 1996; Bushby *et al.*, 2003).

3.3 CURRENT TREATMENTS FOR FSHD

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 and increase their quality of life.

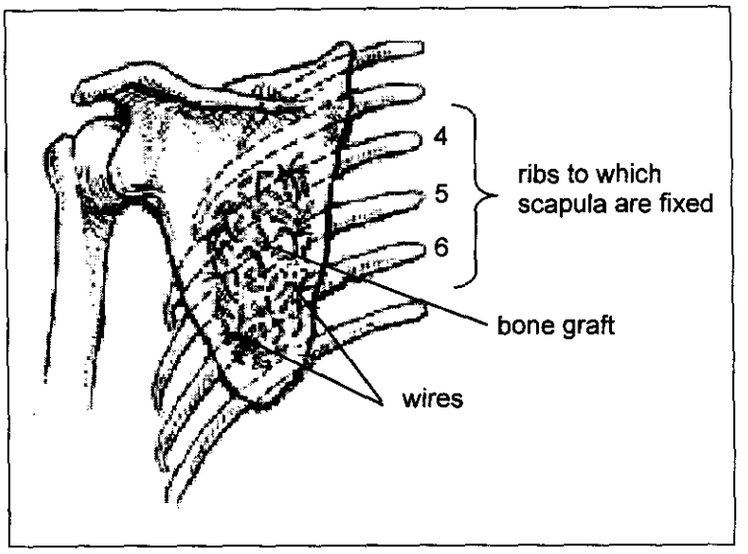
3.3.1 SCAPULOTHORACIC ARTHRODESIS

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 and Siegel (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 and Siegel (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 3.1.

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 wire-ends are bent and buried within the infraspinatus muscles.

Figure 3.1: Scapular fixation in FSHD



Adapted from Bunch and Siegel, (1993).

A solid fusion between the ribs and scapula is necessary, as the wire loops do not provide a rigid enough fixation, since 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 and Siegel (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 and Siegel (1993) favour this method as it is technically relatively easy and is associated with few complications, such as neurovascular due to brachial plexus palsy. Mackenzie *et al.*, (2003) also reported neurovascular complications after scapulothoracic arthrodesis. These authors, however, indicated that

the correct scapular position, as well as the monitoring of pulses during surgery, are important factors in preventing these 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.

3.3.2 STEROID AND ANTI-INFLAMMATORY TREATMENT

Steroids and other anti-inflammatory drugs are generally reserved for the treatment of the inflammatory myopathies, eg. polymyositis. There is thus not a large amount of literature available on the use of steroids in FSHD. Munsat *et al.* (1972) found that 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.

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.

3.3.3 ALBUTEROL TREATMENT

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 the reason for the increase in muscle mass not translating into increased strength.

3.3.4 CREATINE TREATMENT

Creatine is a dietary supplement which is utilised to increase exercise performance. The benefit of creatine supplementation on exercise lead to the investigation of the utilisation of creatine as a possible therapeutic agent in the treatment of specific disorders (Persky and Brazeau 2001). Creatine is distributed throughout the body, however, 95% of creatine is found in skeletal muscle with the remaining 5% found in the brain, liver, kidney and testes. Since the majority of creatine is localised to skeletal muscle, supplementation may be useful in the treatment of myopathies such as FSHD. Walter *et al.* (2000) evaluated the safety and efficacy of creatine in various types of muscular dystrophies, including FSHD, BMD, DMD and sarcoglycan-deficient LGMD. A mild but significant improvement of 3% was observed in muscle strength with a 10% increase in daily-life activities upon creatine supplementation. The mechanism by which creatine may benefit individuals is, however, still unclear and well controlled studies for each of the muscular dystrophies are necessary to accurately assess long-term benefits and side effects of creatine supplementation.

3.3.5 PHYSICAL THERAPY

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.

Eagle, (2002) reported the findings of a workshop that was held in Newcastle, UK, to review the use of exercise and orthotics in the management of neuromuscular disorders,

especially in DMD, BMD, FSHD, DM, spinal muscular atrophy (SMA) Types II and III, the congenital muscular dystrophies and the limb girdle muscular dystrophies. Since limited amount of published data are available to evaluate the use of exercise and orthotics in the treatment of muscular dystrophies in general, the aims of the workshop were to define the current practice in prescribing exercise for neuromuscular disorders, to evaluate the use of specific training exercises and to suggest guidelines, based on evidence and consensus of expert opinions.

3.3.5.1 Current practise

The aims of physiotherapy in neuromuscular disorder treatment are three fold, firstly to maintain or improve the muscle strength by exercise; secondly to maximise functional ability by exercise and the use of orthoses; and thirdly to minimise the development of contractures by stretching and splinting. In general exercise programmes tend to be linked with activities of daily living. Hydrotherapy is enjoyed by children, but older children and adults generally prefer land-based therapy, as they are often embarrassed and become cold easily. (Eagle 2002)

3.3.5.2 Specific training exercises

Only a few of the therapists had significant experience in the use of exercise in FSHD and a lack of literature on the role of exercise, stretching and orthoses, made it difficult to evaluate their value in FSHD. The presence of fatigue, and pain due to over exercise, especially in FSHD, was reported by several participants. It is however not known if this could be related to increased weakness. Hydrotherapy was the treatment of choice of many therapists, however, the evidence was only based on observation by the therapists, and the treatment was therefore based on regular assessment and review rather than on physical evidence.

3.3.5.3 Guidelines

An accurate diagnosis should be made first, to enable the therapist to consider possible complications during exercise. The possibility of fatigue and pain due to over exercise should be considered, although no evidence currently exists to suggest that exercise is contraindicated. Exercise should be prescribed on a regular basis to promote general physical health. Ankle-foot-orthoses may be used in patients with foot drop to improve their gait. (Eagle 2002)

Although the participants provided some evidence to suggest that exercise can improve muscle strength, it is not clear what type, frequency, or duration of the exercise is most effective. It is therefore evident that there is hardly any scientific evidence on the inclusion of exercise in FSHD. The emphasis on the importance of the establishment of multicentre trials was highlighted, since this will create the necessary infrastructure to include significant amount of individuals to produce reliable and valid results that could be utilised in the establishment of treatment based therapies for neuromuscular disorders.

CHAPTER FOUR

GENETIC ASPECTS OF FSHD

Facioscapulohumeral muscular dystrophy is an autosomal dominant, neuromuscular disorder with almost complete penetrance, i.e. 95% at the age of 20 (Lunt *et al.*, 1989b). This disorder is unique amongst genetic disorders, since the mutation causing FSHD is known but the gene(s) affected by the mutation remains unknown. Wijmenga *et al.* (1990) utilised linkage analysis to map the FSHD locus (D4F104S1) to chromosome (chr) 4q35, as presented in paragraph 4.1. Two years later the same authors (Wijmenga *et al.*, 1992b) reported the presence of 3.3 kilo base pair (kb) tandem repeat elements within this FSHD locus, discussed in paragraph 4.2. Subsequently, a deletion of an integral number of these repeats, as discussed in paragraph 4.3, was identified in individuals with FSHD by Van Deutekom *et al.* (1993). Several epigenetic and other factors have, however, also been reported by several authors (discussed in paragraph 4.3.2). To this day the gene(s) involved in FSHD eludes identification although several candidate genes have been proposed (as discussed in paragraph 4.4). In an effort to identify the FSHD gene, expression profiles (discussed in paragraph 4.5) were performed, to identify the differentially expressed genes in FSHD. Although the exact mechanism causing FSHD is not yet known, possible mechanisms are reviewed in paragraph 4.6.

4.1 LINKAGE OF THE FSHD LOCUS

Since the beginning of the 1980s linkage studies were performed in an attempt to identify the FSHD locus. Padberg *et al.* (1984) found possible linkage for the immunoglobulin marker (Gm) locus, encoding the constant region of the heavy chains of the IgG immunoglobulins, located on the long arm of chromosome 14 (near 14q32). The FSHD locus was, however, excluded from the distal part of chromosome 14 after a subsequent study by Padberg *et al.* in 1988, utilising DNA-probe D14S1.

An international collaboration between four groups consisting of 20 participants was established to advance the search for the FSHD locus via the reduction of any unnecessary duplication. The four groups consisted of: 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). The four groups pooled their linkage data to enable the construction of an exclusion map. A total of 57 markers on various autosomes were included in the linkage analyses, but no significant linkage between any marker and FSHD 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). Eventually almost 80% of the genome was excluded by participants of the international consortium. Jacobsen *et al.* (1990) further 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.

Wijmenga *et al.* (1990) established linkage between the FSHD phenotype and the Mfd22 short tandem repeat polymorphism (STRP) marker at locus D4S171 in the subtelomeric region of chromosome 4q35. Ten multigeneration Dutch families consisting of 69 affected, 58 unaffected sibs and 25 spouses were investigated. A total of 60 microsatellite loci were analysed and a maximum lod score of 6.34, at a recombination fraction (θ) of 0.13, was observed for marker Mfd22. This indicated genetic linkage to chromosome 4. The microsatellite marker Mfd22 had previously been assigned to chromosome 4, with the use of a somatic cell hybrid panel (Weber and May, 1990). Only one of the ten families was uninformative for this marker.

Flanking markers, localised to the subtelomeric region of chromosome 4 and more closely linked to the FSHD locus were utilised to confirm the location of the FSHD locus (Upadhyaya *et al.*, 1990). Marker pH30, a variable number of tandem repeat (VNTR) marker, localised at the D4S139 locus was demonstrated to be tightly linked to the FSHD locus. Milner *et al.* (1989) initially mapped the D4S139 locus to the distal portion of the long arm of chromosome 4 via *in situ* hybridisation, with subsequent confirmation by Wijmenga *et al.* (1991). A primary map of the area surrounding the FSHD locus with D4S171 and D4S139 could therefore be constructed 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).

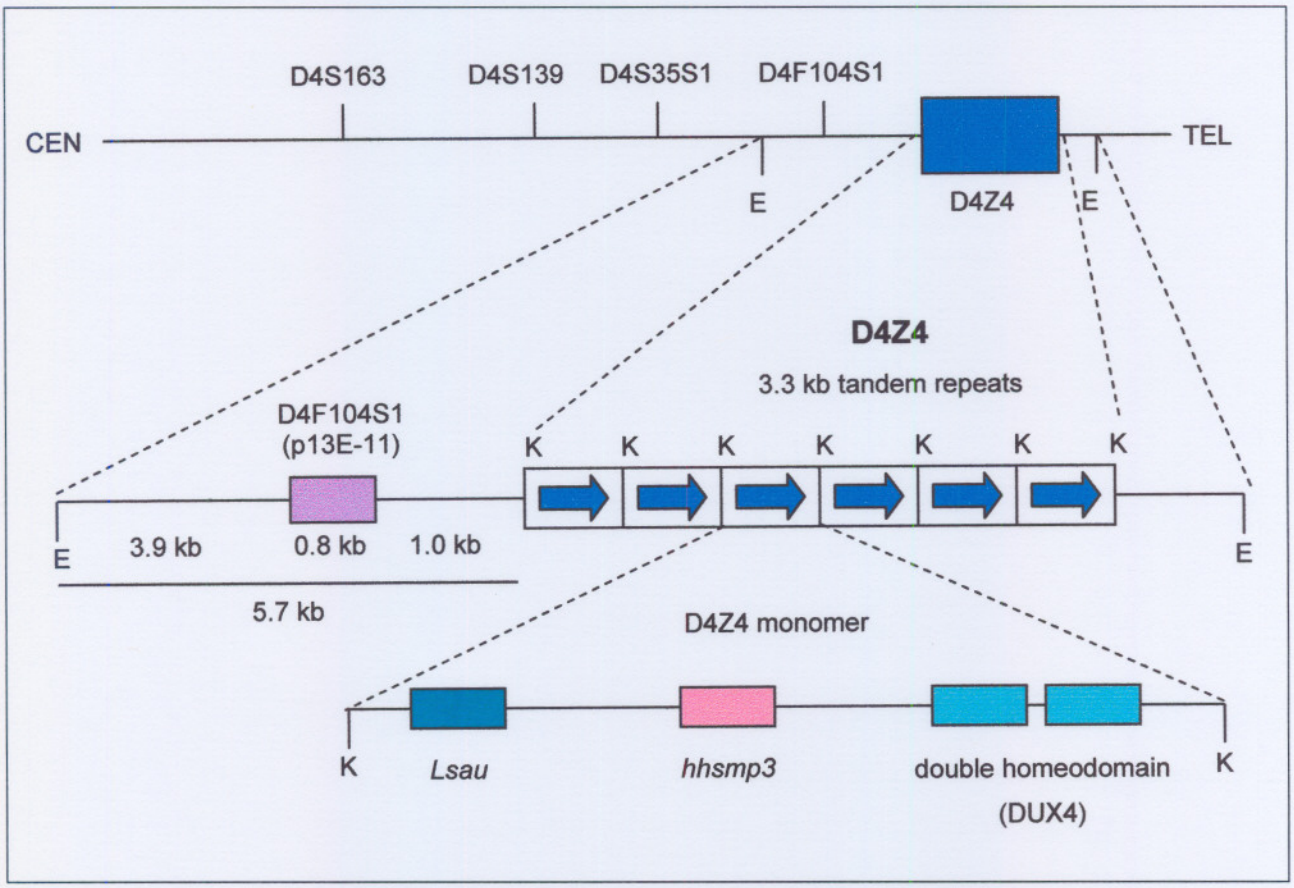
In an attempt to determine the precise location of the FSHD gene members of the International FSHD Consortium have pooled their linkage data. Two additional markers, F11 and D4S163, to those used by Wijmenga *et al.* in 1991 were utilised to form a linkage group consisting of four polymorphic loci in the area of the FSHD locus (D4S171, F11, D4S163 and D4S139). These four polymorphic loci covered a total distance of ca. 20 cM. Recombinants between these four markers have been observed, allowing 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).

4.2 GENOMIC ORGANISATION OF THE FSHD LOCUS

In an independent search for homeobox genes Wijmenga *et al.* (1992b) isolated cosmid clone 13E which mapped to chromosome 4q35, distal to the D4S139 locus. Cosmid 13E was subcloned and a 0.8 kb probe, designated p13E-11, was generated. Wijmenga *et al.* (1992b) further identified, via restriction mapping of cosmid 13E, the presence of 3.3 kb tandem repeat elements at this locus. Each of the 3.3 kb repeat units was flanked by a *Kpn* I site as illustrated in Figure 4.1 and Table D.1. It was postulated by the authors that non-homologous recombination between sister chromatids was a possible mechanism for the rearrangements observed in FSHD individuals. This concept is still considered as a mechanism resulting in the deletions observed in FSHD, as discussed in paragraph 4.3.2.3.

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 4.1 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.* (1995a) cloned a 10 kb deletion fragment (consisting of one 3.3 kb repeat unit) from a severely affected FSHD patient and confirmed the presence of two homeodomain sequences and one *Lsau*-like sequence within each repeat (Lee *et al.,* 1995b). The GC content in the 3.3 kb unit was 71 to 73%.

Figure 4.1: Partial map indicating the relative positions of loci and restriction endonuclease recognition sites within the 4q35 region



Restriction sites: E = *Eco* RI restriction site; K = *Kpn* I restriction site; CEN = centromere; TEL = telomere; *Lsau* = long *Sau* 3A DNA repeats; *hhsmp*3 = human DNA insert showing sperm-specific hypomethylation. Adapted from Fisher and Upadhyaya, 1997.

A sequence tagged site (STS) [D4F106S1], which maps 2 kb proximal to D4F104S1, was developed by Wright *et al.* (1993) and utilised to isolate yeast artificial chromosomes (YACs) for the construction of a cosmid contig. Two YACs were isolated, y25C2E (470 kb) and y956A11 (930 kb), both containing the loci D4S139, D4F35S1 and D4F104S1. Both YACs were mapped to chromosome 4q35 using fluorescent *in situ* hybridisation (FISH), and were demonstrated to be positive for markers p13E-11, CEB8 (a VNTR marker) and pH30, which respectively identify the loci D4F104S1, D4F35S1 and D4S139 via hybridisation analysis.

Two additional probes, p13G2.2E (locus D4S1101) and 9B6A (a 317 bp PCR product containing homeobox sequences from the 3.3 kb tandem repeats), isolated from I13G, a cosmid overlapping 13E, were also used for the mapping of these two YACs. The isolated 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) deduced the order of the loci around the D4S104S1 locus to be:

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.

4.3 MOLECULAR DIAGNOSIS OF FSHD

The mutation causing FSHD, as described in paragraph 4.3.1, was already identified in 1993 by Van Deutekom *et al.* However, the gene(s) affected by the mutational event is still not known. FSHD is therefore a genetic disorder for which the gene remains unknown 10 years after the mutation was identified. This is the longest time period that has ever elapsed between identification of the mutation and identification of the gene.

4.3.1 DNA REARRANGEMENTS ASSOCIATED WITH FSHD

Probe p13E-11 localised to locus D4F104S1 and was observed to detect a DNA rearrangement upon *Eco* RI digestion in individuals affected with FSHD (Van Deutekom *et al.*, 1993). Fragment sizes between 14 to 28 kb were observed to segregate in ten Dutch FSHD families. In this manner, the DNA rearrangements detected via probe p13E-11 were identified to play a role in the aetiology of FSHD.

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 3.3 kb repeats. FSHD negative individuals have 10-100 copies of the chromosome 4 repeat elements, corresponding to 38-350 kb fragments. Patients with FSHD have less than 10 copies of the repeat, thus resulting in deletion fragments which are less than 38 kb in size (Van Deutekom *et al.*, 1993). A deletion giving rise to pathology was in contrast to three other muscular dystrophies namely X-linked spinal and bulbar muscular atrophy (Kennedy's disease), oculopharyngeal muscular dystrophy and myotonic dystrophy, in which an expansion of unstable trinucleotide repeats is the cause of the respective disorders (La Spada *et al.*, 1994; Lieberman and Fischbeck, 2000).

4.3.2 EPIGENETIC AND OTHER FACTORS COMPLICATING THE MOLECULAR DIAGNOSIS OF FSHD

The molecular diagnosis of FSHD is influenced by several factors, as discussed in paragraphs 4.3.2.1 to 4.3.2.8. However, modified Southern blot based assays have been

developed to facilitate in the molecular diagnosis of FSHD (discussed in paragraphs 4.3.2.1.1, 4.3.2.2.1 and 4.3.2.3.1).

4.3.2.1 D4Z4 homologous regions in the genome

The human genome contains hundreds of copies of the 3.3 kb family of tandem repeats which are generally located in heterochromatic regions (Lyle *et al.*, 1995). 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 the pericentromeric regions of chromosomes 1, 3, 9, 10 and Y (Hewitt *et al.*, 1994; Lyle *et al.*, 1995; Winokur *et al.*, 1994; Altherr *et al.*, 1995; Winokur *et al.*, 1996).

The interpretation of Southern blot analyses were complicated since two polymorphic loci as well as a 10 kb Y-specific fragment were detected due to cross-hybridisation of probe p13E-11 (D4F104S1), which detects the rearranged *Eco* RI fragments (Wijmenga *et al.*, 1993b and Weiffenbach *et al.*, 1993). Via haplotype analysis one of the loci could be assigned to chromosome 4q35, but the location of the second locus remained unknown. Potential candidate regions were identified upon cross hybridisation of cosmid 13E utilising 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) 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 as the FSHD region on 4q35 (Figure 4.2). 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.

4.3.2.1.1 Pulsed Field Gel Electrophoresis utilising *Bln* I and *Xap* I restriction fragments

The presence of a unique chromosome 10 specific *Bln* I site, 80 nucleotides upstream of the *Kpn* I site, was identified through the comparison of sequences between the 4q35 and 10q26 fragments, as presented in Table E.1 (Deidda *et al.*, 1996). The organisation of

selected restriction enzyme cutting sites on chromosome 4q35 and chromosome 10q26 are illustrated in Figure 4.2. The *Bln* I site within each chromosome 10 repeat unit can therefore be utilised to discriminate between chromosome 4 and chromosome 10 repeat units, since the chromosome 10 repeat units are sensitive to *Bln* I.

Figure 4.2: Schematic representation of partial restriction endonuclease maps of the 4q35 and 10q26 regions

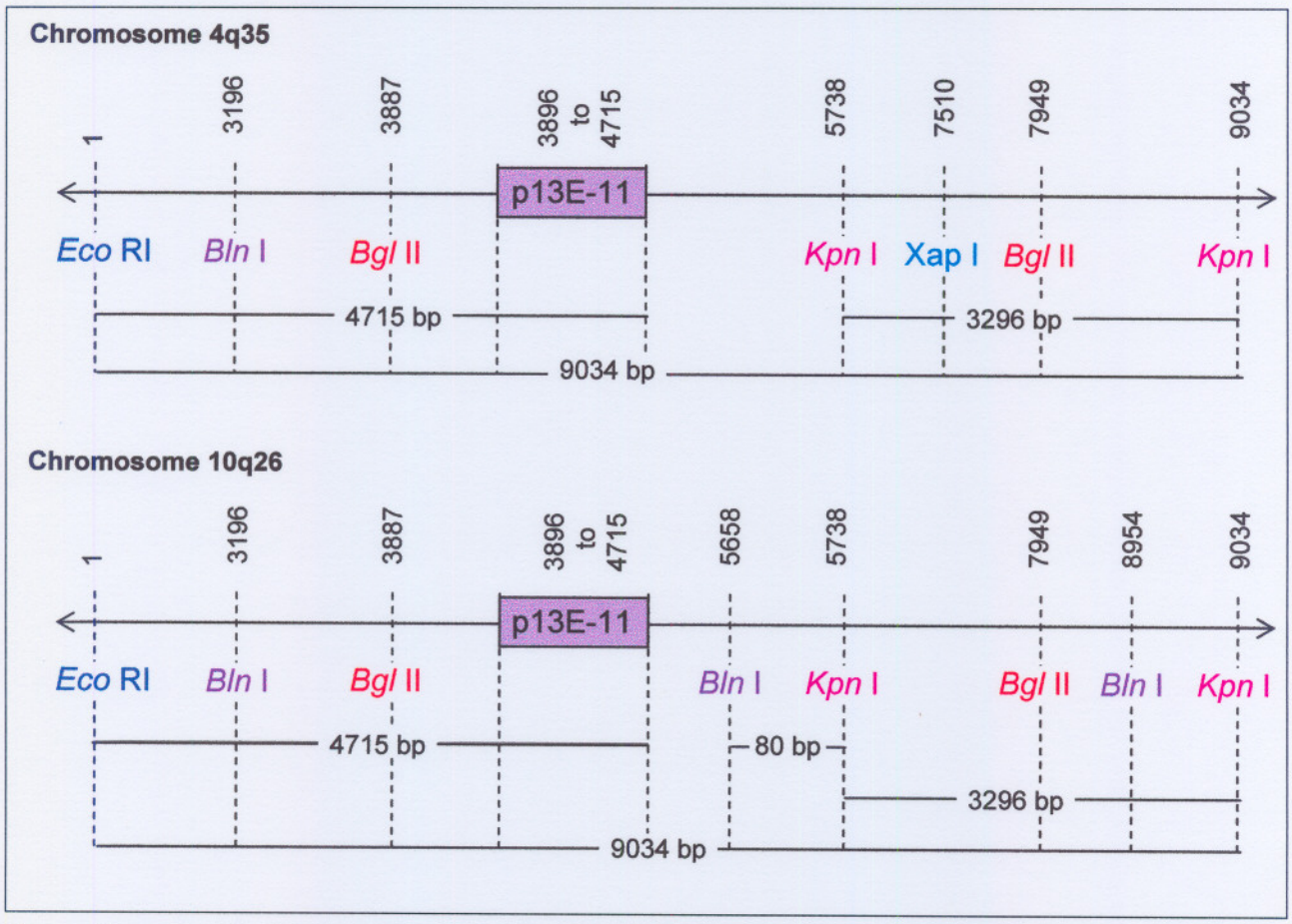


Figure not drawn to scale. Colours and nucleotide numbers correspond with those used to identify specific sequences in Table D.1 and Table E.1.

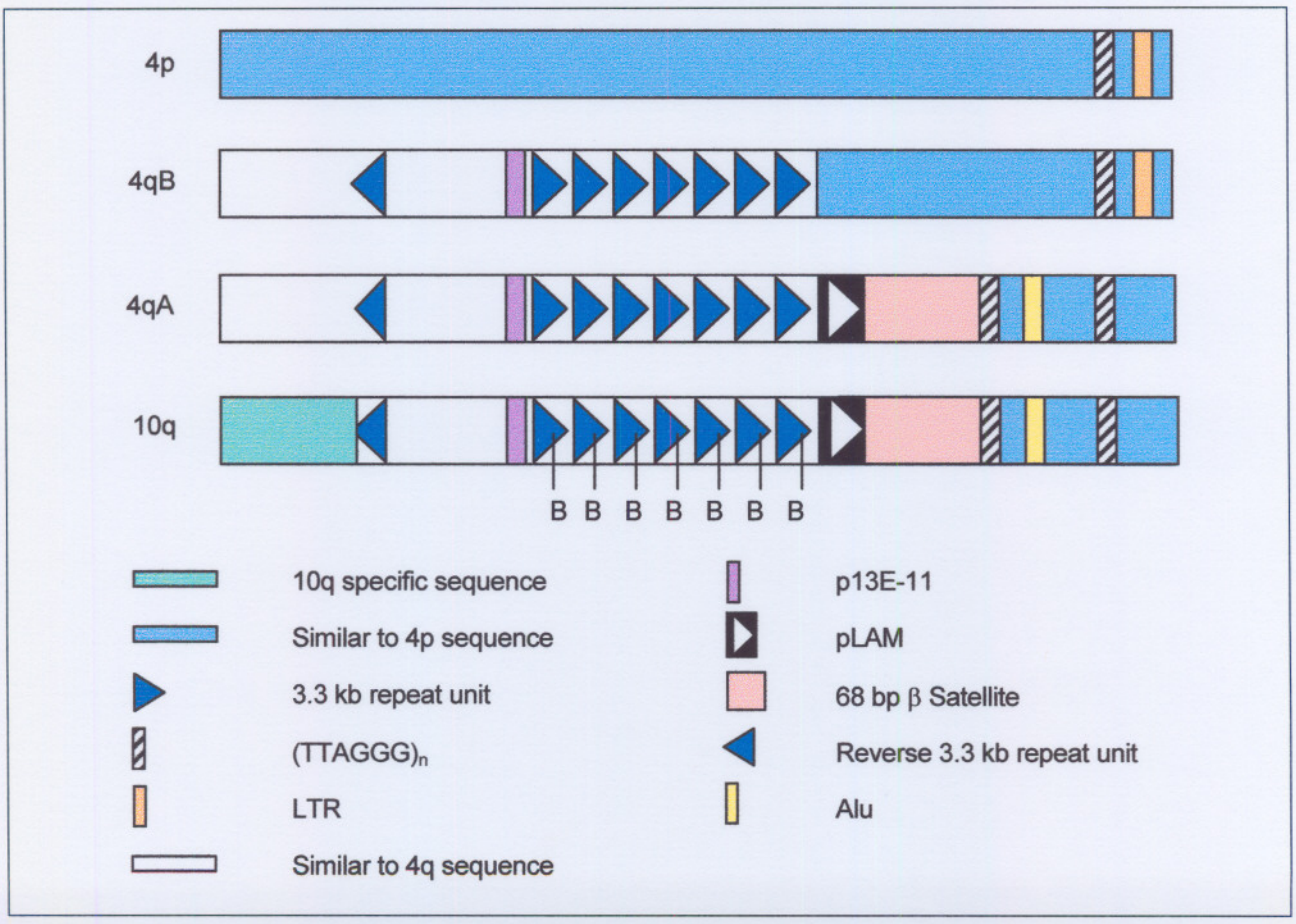
Upadhyaya *et al.* (1997) screened 200 control and 130 FSHD patients and indicated that diagnoses based on double digestion with *Eco* RI and *Bln* I has a sensitivity of 95%, and is therefore a valuable diagnostic test for FSHD. Orrell *et al.* (1999) also investigated 82 unrelated FSHD cases to determine 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, with a sensitivity of 96.5% and a specificity of >99%. In contrast, the sensitivity of the diagnosis utilising conventional linear gel electrophoresis is 92% with a specificity of 99% (Van der Maarel *et al.*, 1999).

Lemmers *et al.* (2001) further identified a *Xap* I restriction endonuclease site on chromosome 4q35 within each 3.3 kb repeat unit, as illustrated in Figure 4.2; 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 *Xap* I digestions therefore allows the characterisation of each allele, even after translocation events (paragraph 4.3.2.3) between 4-type and 10-type repeats have occurred. The combination of these restriction endonucleases with PFGE thus represents an optimal strategy for the diagnosis of FSHD on a molecular level.

4.3.2.2 Variants of the 4qtel region

The homology between the subtelomeric regions of chromosomes 4q and 10q was, however, observed not to be restricted to the 3.3 kb repeats, but rather extends both distally and proximally, as illustrated in Figure 4.3. Van Geel *et al.* (2002) sequenced different clones containing subtelomeric regions from chromosomes 10q and 4q and observed that the chromosome 4 clones were only 92% homologous to each other.

Figure 4.3 Schematic representation of comparison of the sequence organisation of the subtelomeric regions of chromosomes 4q35 and 10q26



B = *Bln* I restriction site. Adapted from Van Geel *et al.* (2002).

Two different allele variants, termed 4qA and 4qB were identified. These variants displayed significant variations in the organisation of sequence elements, as depicted in Figure 4.3. Allele 4qA was observed to harbour an 8 kb region of 68 bp satellite DNA immediately distal to the D4Z4 locus with a 1 kb $(TTAGGG)_n$ array adjacent to this. Neither of these repeats were observed in the 4qB allele. Van Geel *et al.* (2002) also observed that the terminal 3.3 kb repeat in the 4qB allele only contains the first 570 bp of a complete repeat unit. The terminal repeat of the 4qA allele variant contains a pLAM unit. This pLAM unit only consists of 2.9 kb of the 3.3 kb repeat unit (Van Deutekom *et al.*, 1993; Van Geel *et al.*, 2002).

Sequence comparison between 4q and 4p indicated 93% sequence homology between 4qA and 4p, and a 99.5% sequence homology between 4qB and 4p (Van Geel *et al.*, 2002). The presence of an *Alu* insertion within the 4qA allele, which is absent from both the 4qB allele and 4p, together with the insertion of a long terminal repeat (LTR) within 4qB and 4p results in the higher homology between 4qB and 4p (presented in Figure 4.3). Upon sequence comparison it was found that the homology between 4qA and 10q is higher, 96-98%, than between 4qA and 4qB, 91-93%. The higher homology is due to the presence of an Alu sequence but the absence of an LTR sequence on 10q (as illustrated in Figure 4.3). An inverted copy of the 3.3 kb repeat was observed to be present 42 kb upstream of the D4Z4 locus on both 4q and 10q. Van Geel *et al.* (2002) thus indicated that the 10q telomeric region distal to the 3.3 kb repeats has similar organisation than the 4q subtelomeric region.

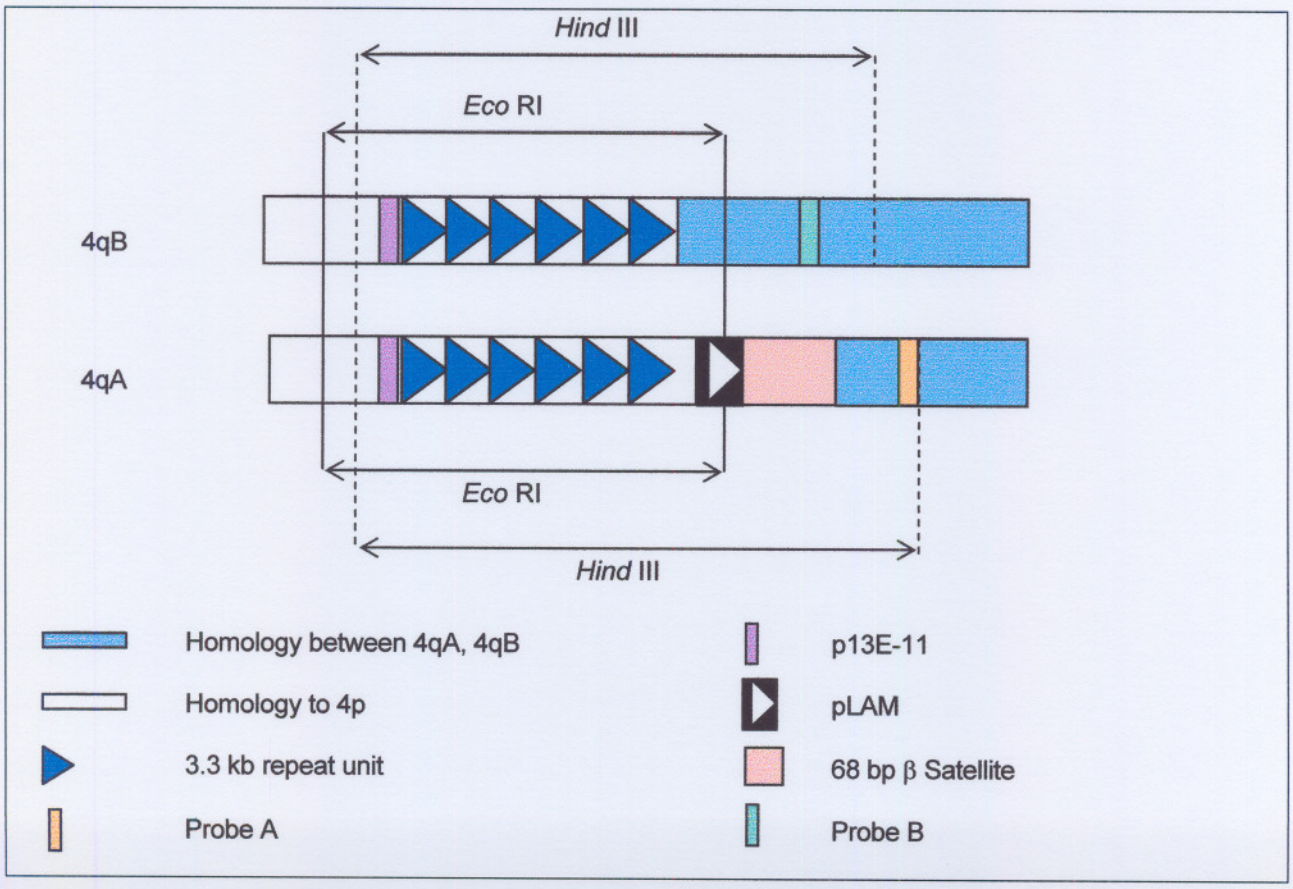
The study by Van Geel *et al.* (2002) highlighted the complex evolutionary history of the 4q telomere. The authors postulated that the 4qA allele was generated through an ancestral duplication event between 4p and 4q. Subsequent to this duplication event a section of 4q, including the 3.3 kb repeat elements, was duplicated onto chromosome 10q. Evidence for chromosome 4qter being the progenitor sequence was generated by Winokur *et al.* (1996) and Clark *et al.* (1996) through investigation of the evolution of the D4Z4 locus in primates. In all the great apes D4Z4 probes were observed to hybridise to the telomeric region of the long arm of chromosome 3, syntenic to human chromosome 4. The 4qB allele was proposed to be the result of a more recent duplication of the subtelomeric sequence from 4p onto 4q (Van Geel *et al.*, 2002).

The sequence of the subtelomeric regions of chromosomes 4q and 10q are therefore strikingly similar. It is therefore highly unlikely that the absence of a pathogenic effect, is due to a difference in the telomeric chromatin organisation if a deletion fragment is present on chromosome 10q26.

4.3.2.2.1 Differentiation between alleles 4qA and 4qB

Lemmers *et al.* (2002) investigated the frequency of the 4qA and 4qB alleles in 80 control individuals and 80 unrelated individuals affected with FSHD. Comparable frequencies of 0.42 and 0.58 for the 4qA and 4qB alleles respectively were observed. The most significant finding of their study was, however, that the deletion fragment present in the FSHD affected individuals was exclusively associated with the 4qA allele. Probes were designed to detect the *Hind* III fragments of the 4qA and 4qB alleles specifically, as illustrated in Figure 4.4. The *Hind* III fragments encompass locus D4F104S1 (probe p13E-11 hybridisation site), locus D4Z4 and either allele 4qA or allele 4qB. The association of the 4qA allele with FSHD might be explained by a functional difference between alleles 4qA and 4qB, as illustrated in Figure 4.3.

Figure 4.4: Schematic representation of the differentiation between the 4qA and 4qB alleles



Adapted from Van Geel *et al.* (2002).

4.3.2.3 Subtelomeric exchange between 4q and 10q sequences

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. 10-type repeats on chromosome 4 (10-on-4), and *vice versa* (4-on-10), was observed in a group that included 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 (Van Deutekom *et al.*, 1996a). A relatively high translocation frequency of 0.2 was observed. These translocations events were thus observed in both affected and unaffected individuals, indicating that a specific gene was not disrupted by the translocation. It is therefore unlikely that the FSHD gene is localised within the repeated units on chromosome 4q35.

Van der Maarel *et al.* (2000) indicated an excess of 4-on-10. The authors postulated that this excess is a significant, if not the major predisposing factor, that gives rise to the deletion associated with the FSHD phenotype. The presence of fully homologous repeat arrays on non-homologous chromosomes, therefore 10-on-4 and 4-on-10, may result in interchromosomal gene conversion via non-homologous recombination between these two chromosomes (Stankiewicz and Lupski, 2002). These recombination events are therefore suspected to cause the FSHD associated deletions. FSHD was proposed to be a recombination-based disorder and could thus be classified as a genomic disorder (Stankiewicz and Lupski, 2002).

Van Overveld *et al.* (2000) examined the 3.3 kb D4Z4 repeat array configurations on chromosome 4 and its homologue on chromosome 10 in an effort to elucidate the configuration of these subtelomeric domains. PFGE was performed in a control group of 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-on-4 and 10-on-10) was observed in 76% of the individuals. A non-standard configuration, due to translocated repeat arrays, was present in 21% of the individuals. 4-on-10 (12%) occur approximately as frequently as the reverse configuration of 10-on-4 (9%), but the composition of these translocated arrays differs significantly as indicated in Table 4.1.

Table 4.1: Configuration and composition of translocated repeat arrays

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

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

The 10-on-4 were more heterogeneous and the 4-on-10 were more homogeneous. It is therefore proposed that there is a preference for 4-on-4, since 33% of the 10-on-4 arrays are hybrids. The distinct configuration and composition differences observed for the 10-type and 4-type repeat arrays therefore indicate a biological difference between these homologous repeat arrays (Van Overveld *et al.*, 2000). Repetitive sequences in the genome undergo concerted evolution (Liao, 1999). This process homogenises repetitive sequences and is important for the maintenance and integrity of each repeat unit within an array. However, the high level of heterogeneity observed for the repeat units at the D4Z4 locus suggests that these loci escaped concerted evolution and instead evolved through inter- and intrachromosomal recombination, as discussed above. The authors proposed that the open reading frame (DUX4) which is present in each repeat unit (refer to paragraph 4.4.2), and for which no expression has been observed, may have lost its function during all the recombination events that caused the differences observed in the homologous repeat arrays between chromosomes 4 and 10.

An unexpected high frequency (0.3) of 4-type repeat arrays shorter than 38 kb was also observed in this control population (Van Overveld *et al.*, 2000). 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:

a. 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 Authors proposed that the prevalence of 'double exchanged' alleles should be ca. 1%.

b. One of the individuals carried three 4-type repeat arrays. In this instance it was regarded to be most likely that the deleted short array resides on chromosome 10.

c. 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 individuals, classified as control individuals, could therefore have donated blood, even though they had FSHD.

Matsumura *et al.* (2002) investigated the translocation frequency in the Japanese, Korean and Chinese populations. The authors examined 153 Japanese, 124 Korean and 114 Chinese control individuals as well as 56 Japanese FSHD patients. Upon analysis the individuals were classified into five groups according to the number of chromosomes with *Bln* I resistant 4q type repeat arrays: nullisomy (no 4q type repeat arrays), monosomy (one 4q type repeat array), disomy (two 4q type repeat arrays), trisomy (three 4q type repeat arrays) and quatrosomy (four 4q type repeat arrays). The configuration and translocation frequencies of the D4Z4 locus in these populations are presented in Table 4.2.

Table 4.2 Configuration and translocation frequencies of the D4Z4 locus in the Japanese, Korean and Chinese populations

Population	10-on-4		Standard	4-on-10		Total	Translocation frequency
	Null	Mono	Di	Tri	Quat		
Japanese controls	1	11	111	29	1	153	0.275
Korean controls	0	5	87	32	0	124	0.298
Chinese controls	0	6	92	16	0	114	0.193
Japanese FSHD patients	0	5	38	13	0	56	0.321

Null = nullisomy; Mono = monosomy; Di = disomy; Tri = trisomy; Quat = quatrosomy. Adapted from Matsumura *et al.* (2002).

Subtelomeric exchange was thus observed in the Japanese, Korean and Chinese populations, with frequencies comparable to that described in the Dutch population (20%, as discussed in paragraph 4.3.2.3). The frequency of chromosome 4-on-10 (trisomies and quatrosomies) was higher than that of chromosome 10-on-4 (nullisomies and monosomies), and is similar to what was observed in the Dutch population. No significant difference in the translocation between the control and the FSHD affected group was detected in the Asian populations.

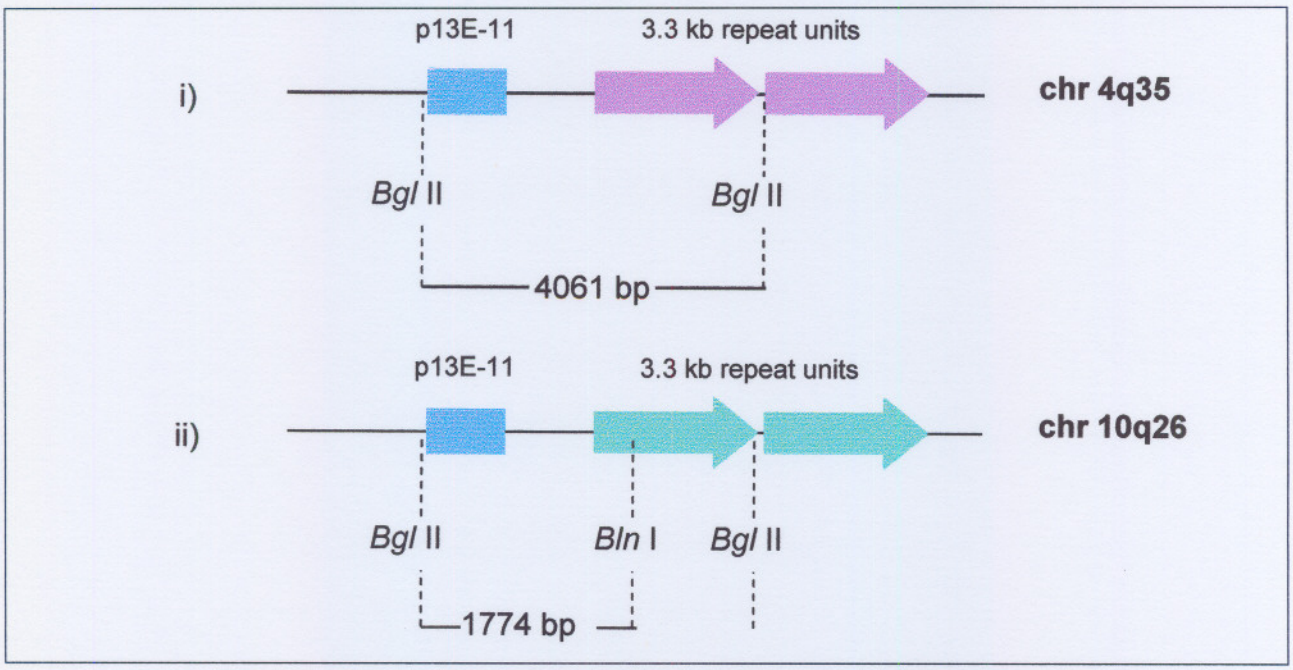
The presence of interchromosomal exchanges between the repeats on chromosome 4 and chromosome 10 has implications for both the specificity and sensitivity of diagnostic DNA testing of FSHD (as discussed in paragraph 4.3.2.1.1). Only short repeat arrays on chromosome 4 are pathogenic, irrespective of the origin of the repeat unit. Short repeat arrays located on chromosome 10 are, therefore, nonpathogenic (Cacurri *et al.*, 1998, Lemmers *et al.*, 1998). The high frequency of ca. 0.2 observed in the Dutch, Japanese, Korean and Chinese populations suggest the need to determine the percentage of interchromosomal exchanges in other populations as well.

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

The diagnosis of FSHD via PFGE (as discussed in paragraph 4.3.2.1.1) has the advantage that all four alleles, as well as the presence of translocations, are visualised. However, the success of PFGE relies on the quality of DNA utilised. Since agarose embedded DNA is not always utilised it is generally difficult to detect alleles exceeding 200 kb with aqueous DNA. Diagnostic laboratories may also not have the facilities, or the expertise, to successfully carry out PFGE analyses.

Van der Maarel *et al.* (1999) developed the *Bgl* II-*Bln* I dosage test, as presented in Figure 4.5, to determine the ratio of 4-type to 10-type repeats. Deletions of the p13E-11 site (as discussed in paragraph 4.3.2.4) are also detected via the dosage test. The *Bgl* II – *Bln* I dosage test utilises the *Bln* I polymorphism in the first D4Z4 repeat on chromosome 10q26, and *Bgl* II instead of *Eco* RI to obtain a smaller fragment, to discriminate between the 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. Subsequent to hybridisation with probe p13E-11 the ratios between the signal intensities from 4q and 10q 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 prior translocation event should exhibit ratios of 1:3 and 3:1 respectively. Individuals with only 4-type or 10-type repeat arrays will show no hybridisation for the other fragment.

Figure 4.5: Schematic representation of the *Bgl* II – *Bln* I dosage test



i = Repeat array on chromosome 4; ii = Repeat array on chromosome 10. Adapted from Van der Maarel *et al.* (1999).

The dosage test will, however, fall short if the translocation between 4q and 10q 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). Van Overveld *et al.* (2000), however, only observed translocations resulting in a hybrid repeat array, with the translocated repeats distal to the first polymorphic *Bln* I site, in 4.3% of the individuals investigated. The dosage test may therefore underestimate the translocation frequency in a particular population. However, since the dosage test can be utilised for the analysis of large sample sizes, it remains a good initial screening test to determine whether plasticity exists in a specific population, and whether further analyses are necessary.

As discussed in paragraph 4.3.2.1.1 the sensitivity and specificity of the diagnosis utilising conventional linear gel electrophoresis is 92% and 99% respectively. However, if the dosage test is added to the conventional diagnosis the sensitivity and specificity will be increased making it comparable to that of diagnoses via PFGE (Van der Maarel *et al.*, 1999). The *Bgl* II – *Bln* I dosage test can also be utilised to study translocation events in large populations. Moreover it can be employed, 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.

Van der Maarel *et al.* (1999) observed via PFGE that in the majority of cases the entire repeat array was translocated to the non-homologous chromosome. Only a small group (11%) of the translocations resulted in hybrid repeat units. Since the *Bgl* II-*Bln* I dosage test utilises the *Bln* I polymorphism in the first D4Z4 repeat, the authors concluded 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. In their study (Van der Maarel *et al.*, 1999) the authors did not localise the FSHD gene within this recombination hotspot region, however, their study did narrow down the region of interest. In the future this translocation breakpoint could be characterised via FISH analysis. The 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.

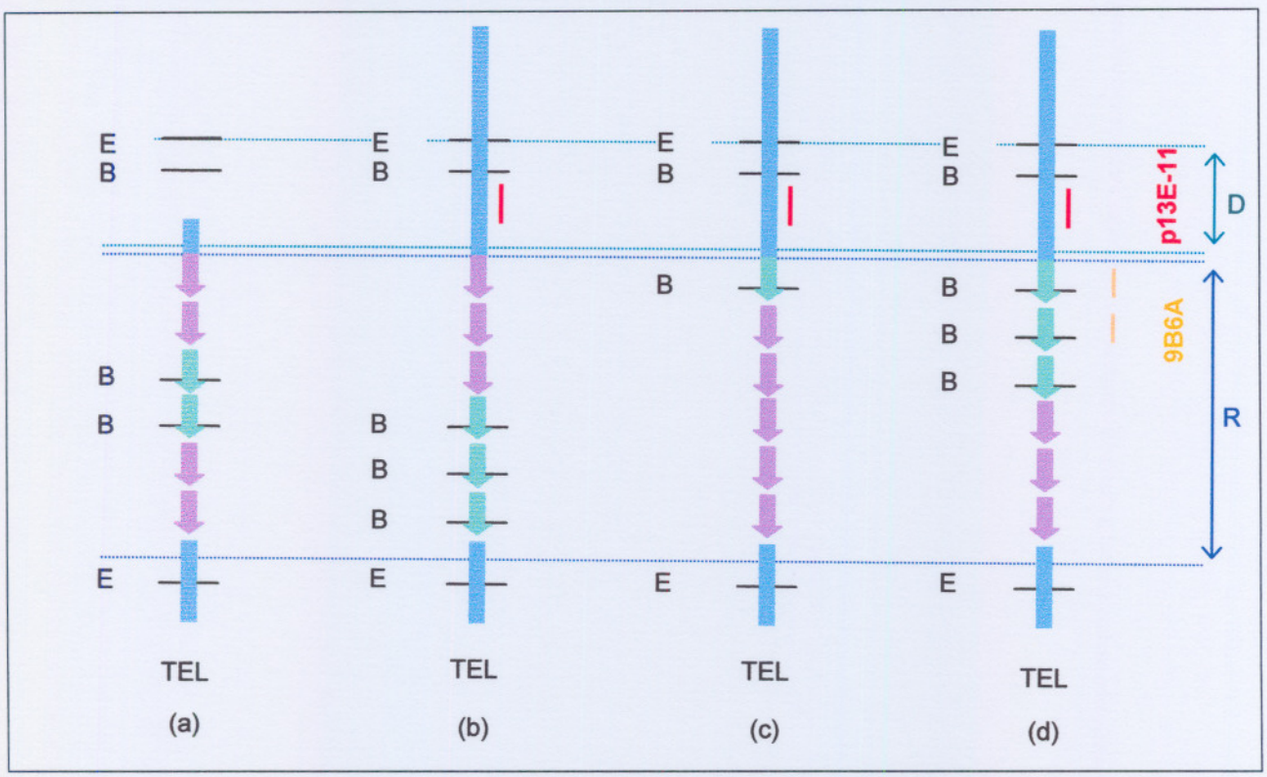
4.3.2.4 Hybrid repeat arrays and deletion of p13E-11 hybridisation site

In addition to the previously mentioned factors complicating the diagnosis of FSHD, hybrid repeat arrays comprising repeats of both chromosome 4 and 10, and deletions of the p13E-11 hybridisation site, further complicates the molecular diagnosis (Van Deutekom *et al.,* 1996a; Lemmers *et al.,* 1998; Lemmers *et al.*, 2003). The effects of these two additional factors were observed after hybridisation with probe 9B6A, which is complimentary to the repeat unit (D4Z4) itself.

As illustrated in Figure 4.6(a), individuals who have a deletion of the p13E-11 site will appear to have no deletion fragment on 4q35 when hybridised with p13E-11. However, upon PFGE only 3 fragments will be observed with p13E-11, the fourth fragment will only 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 4.6(b), 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, since 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 indicated in Figure 4.6(c, d), 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 the 10-type repeats. Via PFGE and hybridisation with probe 9B6A the 4-type repeats on the distal end will be detected. 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. Probe 9B6A was, however, not available to be utilised in this study.

Figure 4.6: Schematic representation of hybrid chromosomes



Purple arrows represent 4-type repeat units; green arrows represent 10-type repeat units; a = Hybrid repeat arrays with a deletion of the p13E-11 hybridisation site; b = Hybrid repeat arrays consisting of 10-type repeats attached to the distal end of a 4-type repeat arrays; c = Hybrid repeat arrays consisting of 4-type repeats attached to the distal end of a 10-type repeat array; d = Hybrid repeat arrays consisting of 4-type repeats attached to the distal end of 10-type repeat arrays; 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).

4.3.2.5 Somatic and germline mosaicism

Somatic mosaicism is indicated by a fifth fragment that is detected subsequent to PFGE analysis and hybridisation with probe p13E-11. This phenomenon has been reported by several 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 cases 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 harboured 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 who inherited a 25 kb *Bln* I resistant deletion fragment from his clinically asymptomatic father was identified, therefore a familial and not a sporadic case. This case was reclassified as familial and omitted from the above group of sporadic cases.

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 (4 males and 1 female) from the group of 23 families 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 a fragment within the FSHD causative range. Overall somatic mosaicism was observed in 40% of cases, either in the patient or in an asymptomatic parent. The degree of mosaicism could, however, be even higher, since complete allelic information was only available for only 23 of the 35 *de novo* families. The authors postulated that this high degree of somatic mosaicism implies that the deletion event is mainly mitotic.

As described in paragraph 4.3.2.3, 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 control 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 increased almost five fold. 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 may facilitate the elucidation of some aspects of the clinical variability of this disorder, since different muscle samples from a particular individual will contain

different levels of the disease allele. These levels could therefore be correlated with the degree of severity in these samples.

4.3.2.6 Anticipation

Myotonic dystrophy was the first disorder identified to display anticipation, which is characterised by an earlier onset with an increase in the severity of clinical symptoms in subsequent generations (Harper *et al.*, 1992; Mak and Ho, 2001). Here, the disease causing mutation is an unstable expansion of the CTG repeat in the myotonin-protein kinase (DMPK) gene on chromosome 19 (Buxton *et al.*, 1992). Anticipation in this disorder is caused by an increase in the number of trinucleotide repeats from generation to generation. Moreover, there is a direct correlation between the size of the trinucleotide expansion, the age of onset and the clinical severity (La Spada *et al.*, 1994; Mak and Ho, 2001).

At least fifteen neurological disorders have been demonstrated to be caused by an expansion of unstable trinucleotide repeats, as presented in Table 4.3. (La Spada *et al.*, 1994; Lieberman and Fischbeck, 2000; Cummings and Zoghbi, 2000). These trinucleotide repeat disorders can be categorised into two subclasses based on the location of the trinucleotide repeats in either the untranslated or exonic sequences. The location of the trinucleotide repeat in relation to the gene may play a prominent role in the unique pathogenesis of each disorder. The noncoding repeat disorders are generally multisystemic and phenotypic heterogeneity often exists. These disorders are also generally associated with clinically silent, intermediate-size expansions, or premutations, that may expand to the pathologic range after germline transmission. The disorders with expansions located in coding sequences (Huntington disease, spinobulbar muscular atrophy, dentatorubropallidoluysian atrophy, SCA types 1,2,3,6 and 7) were observed to be associated with an expansion of $(CAG)_n$ repeats coding for polyglutamine tracts. The size and variation of the repeat expansions are much greater in the noncoding-repeat disorders than in the polyglutamine repeat disorders. (Cummings and Zoghbi, 2000)

Anticipation is generally observed in these trinucleotide repeat disorders due to the instability of the repeats (La Spada *et al.*, 1994; Cummings and Zoghbi, 2000). The degree of anticipation is further dependent on parental origin of the repeat. For most of the trinucleotide repeat disorders there is a greater risk of repeat expansion upon paternal

transmission, except for Fragile X syndrome, Friedreich ataxia and congenital forms of DM, where the maternally transmitted alleles are more prone to instability (Cummings and Zoghbi, 2000).

Table 4.3: Neurological disorders caused by expansion of unstable trinucleotide repeats

Disorder	Gene	Locus	Protein	Repeat	Repeat size		Repeat location
					Normal	Affected	
FRAXA [1]	FMR1 [15]	Xq27.3	FMRP [22]	CGG	6-53	60-200 (pre) [25] >230 (full) [26]	5'-UTR [27]
FRAXE [2]	FMR2 [16]	Xq28	FMR2 protein	GCC	6-35	61-200 (pre) >200 (full)	5'-UTR
FRDA [3]	FRDA [17]	9q13-21.1	Frataxin	GAA	7-34	34-80 (pre) >100 (full)	Intron 1
DM [4]	DMPK [18]	19q13	DMPK	CTG	5-37	50-1000s	3'-UTR
SBMA [5]	*AR* [19]	Xq13-21	Androgen receptor	CAG	9-36	38-62	Coding (AT) [28]
Huntington disease	HD [20]	4p16.3	Huntington	CAG	6-35	36-121	Coding (AT)
DRPLA [6]	DRPLA	12p13.31	Atrophin-1	CAG	6-35	49-88	Coding (AT)
Spinocerebellar ataxia type 1 [7]	SCA1	6p23	Ataxin-1	CAG	6-44	39-82	Coding (AT)
Spinocerebellar ataxia type 2 [8]	SCA2	12q24.1	Ataxin-2	CAG	15-31	36-63	Coding (AT)
Spinocerebellar ataxia type 3 [9]	SCA3	14q32.1	Ataxin-3	CAG	12-40	55-84	Coding (CT) [29]
Spinocerebellar ataxia type 6 [10]	SCA6	19p13	CACNA1A [23]	CAG	4-18	21-33	Coding (CT)
Spinocerebellar ataxia type 7 [11]	SCA7	3p12-13	Ataxin-7	CAG	4-35	37-306	Coding (AT)
Spinocerebellar ataxia type 8 [12]	SCA8	13q21	---	CTG	16-37	110-<250	3'-Terminal exon
Spinocerebellar ataxia type 12 [13]	SCA12	5q31-33	PPP2R2B [24]	CAG	7-28	66-78	5'-UTR
OPMD [14]	PABP2 [21]	14q11	PABP2	GCG	6	8-13	Coding (CT)

1 = Fragile X syndrome, MIM 309550; 2 = Fragile XE syndrome, MIM 309548; 3 = Friedreich ataxia, MIM 229300 and 606829; 4 = Myotonic Dystrophy, MIM 160900; 5 = Spinobulbar muscular dystrophy or Kennedy disease; 6 = Dentatorubral-pallidoluysian atrophy or Haw River syndrome, MIM 125370 and 607462; 7 = Machado-Joseph disease, MIM 164400; 8 = MIM 183090; 9 = MIM 10915014; 10 MIM = 183086; 11 = MIM 164500; 12 = MIM 603680; 13 = MIM 604326; 14 = Oculopharyngeal muscular dystrophy, MIM 164300 and 2579509; 15 = Fragile X mental retardation 1 gene; 16 = Fragile site mental retardation 2 gene; 17 = Friedreich ataxia gene, also known as frataxin; 18 = Myotonic dystrophy protein kinase; 19 = Androgen receptor; 20 = Huntington disease gene; 21 = poly(A)-binding protein 2 gene; 22 = Fragile X mental retardation protein; 23 = α1A-Voltage-dependent calcium channel subunit; 24 = Protein phosphatase PP2A; 25 = premutation; 26 = full mutation; 27 = untranslated region; 28 = amino terminal; 29 = carboxy terminal.

Anticipation in FSHD is currently a topic of great debate and has been reported by a few authors (Lunt *et al.*, 1995a; Zatz *et al.*, 1995; Tawil *et al.*, 1996). One shortcoming of all these studies where anticipation was reported in the respective populations was, however, that affected parent-offspring pairs, thus only two generations, from multiple families were

studied. This aspect does cast doubt on the above reports. 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 stable 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 (U.S.A.) in 1840. Tyler and Stephens (1950) identified 159 affected individuals, either by history or examination, and reported on the clinical examination results of 58 people. Twenty four of the 58 individuals (41%) who were asymptomatic by self-report were diagnosed as affected upon examination. Thirteen 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, spanning seven generations in seven different branches of this extended family. However, no compelling evidence for anticipation, in either reported age of onset or in disease severity, was observed in this extended family. Additional studies in different populations, utilising quantitative muscle testing, historical data, and genotyping will help to elucidate the aspects of anticipation, gender effects, and parent-origin effects in FSHD. Until that time, anticipation associated with the FSHD phenotype will remain a controversial topic.

4.3.2.7 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 may 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 lineages, 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.

It was also observed that mosaic males were generally affected, although mosaic females were more often the unaffected parent of a nonmosaic sporadic patient (Van der Maarel *et al.,* 2000). A difference in the clinical presentation of females and males was also observed by Busse *et al.* (2000). A 35 kb *Bln* I resistant (thus 4-type) deletion fragment was confirmed in an unaffected mother (65 years of age), two unaffected daughters (40 and 27 years, respectively) and an affected son (30 years of age). 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. The phenotype differences observed between males and females were confirmed by Tonini *et al.* (2004). The authors also observed more females than males in the asymptomatic or minimally affected group of patients.

4.3.2.8 Genetic heterogeneity

The great clinical heterogeneity that exists in FSHD, as discussed in paragraph 3.2, alluded to the possibility of genetic heterogeneity in FSHD. 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, Krasnianski *et al.*, 2003b). 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 display linkage to 4q35.

Despite the report of genetic heterogeneity in FSHD almost a decade ago, a second locus has yet to be identified. It is currently estimated that at least 5% of FSHD families do not display linkage to chromosome 4q35, implying that genetic heterogeneity does exist in FSHD. The finding of heterogeneity in FSHD has important implications for genetic

counselling, clinical investigations, prenatal diagnosis, and the eventual characterisation of the FSHD gene(s).

4.3.3 SPORADIC FSHD

A relatively high proportion of new mutations (10 to 30%) have been reported in FSHD (Padberg 1982; Padberg *et al.*, 1995b; Zatz *et al.*, 1995; Tawil *et al.*, 1996; Lunt 1998; Zatz *et al.*, 1998). 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. The presence or absence of the mutation within the sporadic case and the parents respectively, was identified via PFGE and haplotype analyses. 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 seems highly unlikely that the FSHD-associated fragments of identical size in both of the affected offspring were a result of two independent chromosomal rearrangements. It was thus concluded that one of the parents was mosaic thus harbouring the FSHD-associated rearrangement in their germlines. Jardine *et al.* (1993) also identified a *de novo* DNA rearrangement in a proband which was eventually transmitted to 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.

4.3.4 CORRELATION BETWEEN THE FSHD PHENOTYPE AND GENOTYPE

A direct correlation between the residual repeat length, the age of onset and the severity of FSHD was reported 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 3.2.2.1 (Funakoshi *et al.,* 1998; Miura *et al.,* 1998). The presence of anticipation was also reported, as mentioned in paragraph 4.3.2.6 but remains controversial. It was observed that either the age of onset or the degree of weakness, in consecutive generations in familial FSHD, becomes more severe (Lunt *et al.*, 1995a). The issue of whether anticipation exists in FSHD has to be clarified first and can therefore not yet be utilised in phenotype-genotype 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, as illustrated in Table 4.4. 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.

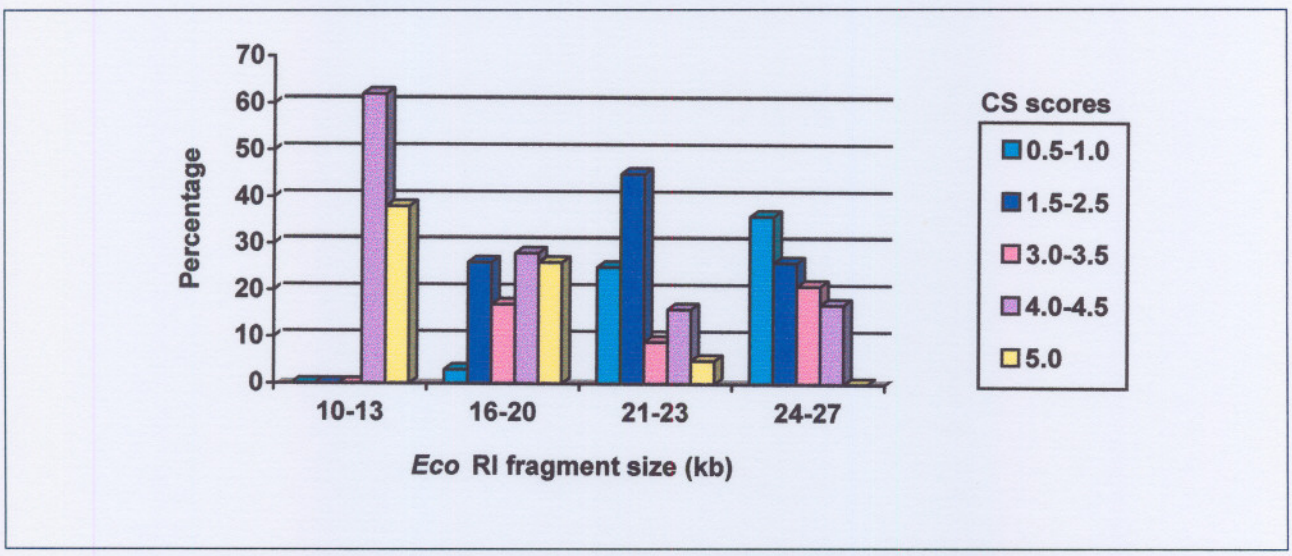
Table 4.4: Clinical severity scale for FSHD

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).

The proportion of patients with different CS scores and *Eco* RI deletion fragments is displayed in Graph 4.1. It was concluded 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 4.1: Proportion of patients with different clinical severity scores and *Eco* RI deletion fragment sizes



CS = clinical severity. CS scores were calculated as indicated in Table 4.4. Adapted from Ricci *et al.* (1999).

It is of interest 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 1.0 increases with the increase of the *Eco* RI fragment size.

A correlation between the residual repeat length, and the severity of the FSHD phenotype was also demonstrated in 506 individuals from 106 families (Tonini *et al.* 2004). However, this correlation could only be established when the total sample, including both males and females, was analysed. Upon separate investigation of the genders, the correlation was only significant for the females. The authors therefore proposed that the severity of the phenotype observed in FSHD has to be influenced by other genetic or epigenetic mechanisms.

Butz *et al.* (2003) investigated the phenotype-genotype correlation in 39 unrelated FSHD affected individuals. The individuals had confirmed molecular diagnostic results, with borderline deletion fragments corresponding to 8–14 D4Z4 repeat units and heterogeneous clinical diagnoses of FSHD, facial-sparing FSHD, atypical FSHD and non-FSHD. Although the correlation between the clinical severity and the number of 3.3 kb repeat units was indicated by several authors, Butz *et al.* (2003) highlighted the limitations of the genetic test in these borderline cases. No definite cut-off point could be established to separate FSHD and FSHD-like myopathies, since four very heterogeneous phenotypes, as described above, were classified in this borderline region. No correlation between the number of 3.3 kb repeats and the clinical severity could be observed in these individuals. The molecular test for FSHD is therefore limited to differentiate FSHD from FSHD like muscle disorders when more than 8 repeats units are present. From the data presented by Butz *et al.* (2003), it was highlighted that the FSHD phenotype can not be explained solely by the rearrangement at the D4Z4 locus. Analysis of this data also confirmed that other genetic or epigenetic mechanisms have to be involved. Although caution in the diagnostic and genetic counselling process is always important, is especially so for borderline cases.

4.3.5 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 paves the way for 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 a total of 46 patients. Most of the patients indicated that they would have liked to know their diagnosis earlier to enable them to seek more effective 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 foetus.

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. Molecular characterisation of a family prior to prenatal testing is currently standard practice.

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 foetuses. Parents chose to terminate the pregnancies in eight of the nine cases.

A preimplantation genetic diagnostic (PGD) test was designed by Marshall *et al.* (2003). Three STRP markers, D4S2299, D4S2283 and D4S1523, localised in the FSHD region were utilised. The authors performed PGD for two couples, both resulting in the birth of unaffected individuals. The approach by Marshall *et al.* (2003) therefore allowed the successful diagnosis for FSHD from single cells. However, in an earlier study Van der Merwe (2002), utilised the same three STRP markers, together with six additional STRP markers to construct haplotypes of 100 individuals from five FSHD families. The presence of a recombination event between the most distal STRP marker (D4S1523) and locus D4Z4 was observed in one affected individual (1%). Thus, if only the haplotypes were utilised for the diagnosis of this individual, he would have been diagnosed incorrectly as unaffected. Furthermore, several recombination events between the nine STRP markers were observed within the families which enabled the mapping of the exact order of the nine STRP markers. The order of the three most distal markers was defined by Van der Merwe (2002) to be D4S1652, D4S2283 and D4S1523. Therefore, the three markers utilised by Marshall *et al.* (2003) were not the three most distal markers. Although the PGD test by Marshall *et al.* (2003) correctly diagnosed two individuals, it is of the utmost importance to take the presence of possible recombination events in consideration. This highlighted the fact that Southern blot analysis remains essential in the molecular diagnosis of FSHD.

4.4 CANDIDATE GENES

The identification of genes associated with the chromosome 4q35 region linked to FSHD has been a difficult endeavour, largely due to the D4Z4 homologous regions in the genome (as discussed in paragraph 4.3.2.1), the high density of repeat units and the gene-poor nature of the FSHD region. Despite several attempts the gene(s) whose differential expression culminates in the FSHD phenotype still eludes identification. In the next section the genes so far identified in the FSHD region are discussed.

4.4.1 ADENINE NUCLEOTIDE TRANSLOCATOR 1 GENE (ANT1)

The adenine nucleotide translocator (ANT), also called adenine nucleotide translocase or adenosine diphosphate (ADP)/ATP translocator, is the most abundant mitochondrial protein (Adrian *et al.*, 1986). This protein facilitates the exchange of adenine nucleotides across the mitochondrial inner membrane. It is a 30 kDa 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 were described by Neckelmann *et al.* (1987) and Houldsworth and Attardi (1988): human skeletal muscle (ANT1), human fibroblast cells (ANT2) and human liver (ANT3).

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 via 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. Haraguchi *et al.*, 1993, however did not observe any abnormality after analysing the transcripts of ANT1 from several FSHD patients (see paragraph 4.5).

4.4.2 DOUBLE HOMEOBOX GENE 4 (DUX4)

Ding *et al.* (1998) cloned a 170 amino acid protein (DUX1) containing a double homeodomain and demonstrated 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 three DUX genes were mapped to the acrocentric chromosomes and are therefore not involved in FSHD (Beckers *et al.*, 2001). Gabriëls *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 ORF encompassing the double homeobox in each 3.3 kb repeat unit and termed the putative gene DUX4 (Figure 4.1 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.

It was hypothesised by Gabriëls *et al.* (1999) that the deletions of the 3.3 kb repeat elements at the D4Z4 locus destabilises the heterochromatin allowing expression of the gene in selected repeats of some cells. The DUX4 protein is, however, toxic to muscle cells (Gabriëls *et al.*, 1999). The authors also postulated that the dominant character of FSHD as well as the fact that the severity correlates inversely with the residual number of repeats could be explained by this hypothesis.

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

DUX4 may therefore play a role in the pathogenesis of FSHD since 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 in the human genome.

4.4.3 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, foetal 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 (Van Deutekom *et al.*, 1996b). FRG1 can however not be excluded before the disease mechanism has been elucidated as discussed in paragraph 4.6.

Grewal *et al.* (1998) investigated FRG1 on an evolutionary basis via the comparison of the genomic organisation of this gene in two species, the mouse (*Mus musculus*) 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 concluded that the human FRG1 gene is highly conserved in both vertebrates and 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, which could imply a fundamental role in RNA processing (Van Koningsbruggen *et al.*, 2000; Van Geel, 2001).

4.4.4 THE FSHD REGION GENE 2 (FRG2)

The FRG2 gene was identified by means of *in silico* exon prediction and confirmed via subsequent expression studies, mapping it to 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. Although FRG2 expression was absent in all tissues tested, low but distinct levels of FRG2 expression was observed in differentiating myoblasts from FSHD patients. No expression of FRG2 could be detected in control differentiating myoblasts. FRG2 was 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 myofibres. 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 dysregulation is related to FSHD (Van Geel, 2001). FRG2 therefore remains an important candidate gene in FSHD as discussed in paragraph 4.5.

4.4.5 HUMAN BETA-TUBULIN 4 GENE (TUBB4Q)

The TUBB4Q gene (Tubulin beta polypeptide 4 member Q) maps 80 kb proximal to the FSHD associated D4Z4 repeats on chromosome 4q35 and was therefore initially considered as a candidate gene (Van Geel *et al.,* 2000). The gene contains four exons, encoding a protein of 434 amino acids. Although the genomic structure indicated 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 foetal tissues. To date, no evidence has been generated to demonstrate that expression of TUBB4Q is involved in FSHD (Van Geel, 2001).

4.5 GENE EXPRESSION PROFILING IN FSHD

Gene expression profiles in FSHD were investigated by Tupler *et al.* (1999) via the comparison of mRNA expression patterns of FSHD and normal muscle. Muscle from three sex- and age-matched control individuals as well as one BMD and one amyotrophic lateral sclerosis (ALS) patient were included in the study. Histological examination was performed on all the muscles included in the analyses. A wide spectrum of differentially expressed genes was observed. A significant number of genes encoding transcription regulators, such as histone 4 acetyl transferase were observed to be over expressed. In contrast, five genes encoding transcription factors were identified to be under expessed in FSHD muscle. The dysregulation of transcription factors and regulators could explain the broad sel ection o f genes that were observed to be differentially expressed in FSHD muscle.

The over expression of *FRG1, FRG2* and *ANT1* was demonstrated in FSHD patients (Gabellini *et al.*, 2002). Muscle samples from DMD, LGMD and ALS patients were included as controls. The expression of *FRG2* was only detectable in FSHD muscle samples while *FRG1* and *ANT1* were expressed in all muscle samples. However, *FRG1* and *ANT1* were observed to be over expressed in FSHD muscle. The authors further identified that the level of over expression of the three genes were inversely related to the distance from the D4Z4 locus and the number of 3.3 kb repeats present at the locus.

Jiang *et al.* (2003), however, argued against the data reported by Gabellini *et al.* (2002) since no over expression of *FRG1, FRG2* or *ANT1* was observed in FSHD muscle by these authors. Quantitative real-time PCR analyses were performed by Jiang *et al.* (2003)

compared to the reverse-transcriptase PCR utilised by Gabellini *et al.* (2002). Jiang *et al.* (2003) did not observe any significant increase in the level of either *FRG2* or *ANT1* transcripts upon normalisation with four different internal RNA standards and seven muscle samples compared to the one standard and three muscle samples utilised by Gabellini *et al.* (2002). The authors observed a significant decrease in transcript levels for *FRG1,* corresponding to the findings by van Deutekom *et al.* 1996b (as discussed in paragraph 4.4.3).

The alterations in protein expression in FSHD muscle were also investigated by Laoudj-Chenivesse *et al.* (2003). Higher expression levels of Cu/Zn superoxide dismutase (SOD1) was detected in all FSHD muscles analysed. The over expression of SOD1 could therefore be an indication of an over production of reactive oxygen species causing mitochondrial dysfunction. These findings by Laoudj-Chenivesse *et al.* (2003), are in agreement with the observations of Winokur *et al.* (2003a) indicating an increased susceptibility of FSHD myoblasts to oxidative stress. An over expression of *ANT1* was also reported by Laoudj-Chenivesse *et al.* (2003), however, the levels observed were unrelated to the number of 3.3 kb repeats as demonstrated by Jiang *et al.* (2003) but in contrast to the findings of Gabellini *et al.* (2002). Laoudj-Chenivesse *et al.* (2003) proposed that the increased expression of *ANT1* results in mitochondrial dysfunction which will eventually cause atrophy of FSHD muscle fibres.

Winokur *et al.* (2003b) generated FSHD expression profiles utilising oligonucleotide microarrays. The majority of genes with altered expression profiles in FSHD muscle were identified to be direct targets of MyoD. In addition to these genes, differential expression of other genes influencing the activity of MyoD and the process of myogenic differentiation were observed. However, no increased expression of 4q35 genes was observed.

The large degree of variation observed in the gene expression data by different authors emphasised the need for standardisation between researchers. The studies should be carefully planned to ensure sufficient amount of patients and the inclusion of different types of muscle from the same patient. There is, however, currently an effort underway, supervised by Dr. Flanigan (Departments of Neurology, Pathology, Human Genetics and Program in Neuroscience, University of Utah, U.S.A) to establish a standardised protocol for gene expression in FSHD (FSHD international research consortium workshop 2003), in which researchers will exchange samples to eliminate human bias.

4.6 MOLECULAR MECHANISMS IMPLICATED IN FSHD

The exact 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. Several hypotheses are discussed in the following five paragraphs that may explain the unique pathogenesis of FSHD.

4.6.1 HOMEODOMAIN HYPOTHESIS

The homeodomain hypothesis was postulated by Lee *et al.* (1995b), in which they propose that FSHD may be due to a rearrangement in a homeobox gene. 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 identical 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 throughout the entire 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 encoded by homeodomains display sequence-specific DNA binding and play an important role in the process of transcription, as transcription regulating 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 regulates 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. 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 the repeat units.

4.6.2 POSITION EFFECT VARIEGATION HYPOTHESIS

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. They hypothesised that the deletions of the 3.3 kb repeats could result in the expansion of telomeric heterochromatin into adjacent euchromatin altering the expression of the gene.

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 (or its position in the environment) and may therefore change the expression of the gene, which is then referred to as a position effect.

Position effect variegation (PEV) has been observed in *Drosophila* and yeast and it was indicated 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 located more distally.

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 demonstrating 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. 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 (Kleinjan and Van Heyningen, 1998). The FSHD gene could thus be located quite a distance from D4Z4 if PEV was indeed the molecular mechanism implicated in FSHD. 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).

The majority of genotype-phenotype correlation studies indicate an inverse relationship between the residual repeat length and the clinical severity (paragraph 4.3.4). PEV could explain this phenomenon since the FSHD candidate region is moved further into the 4q heterochromatic telomeric region as the deletion increases, resulting in increased disease severity. Furthermore the effect on the chromosomal structure from heterochromatic to euchromatic increases as the deletion increases, influencing gene expression accordingly.

Recently Jiang *et al.* (2003) re-evaluated the PEV hypothesis by investigating histone modification and gene expression profiles for the 4q subtelomeric region. According to the PEV hypothesis it is assumed that the normal length D4Z4 arrays are heterochromatic and heterochromatinisation in unaffected individuals spreads from the array to genes involved in FSHD on 4q35. The PEV model states that the normal heterochromatinisation at 4q35 is lost in the deletion state, as observed in FSHD patients. The consequence of the loss of this heterochromatin-spreading is predicted to be an inappropriate increase in expression of critical genes in *cis* in affected individuals. Histone hypoacetylation was identified as a marker for constitutive heterochromatin. It was however observed that histone H4 acetylation levels were similar to that of unexpressed euchromatin rather than to that of constitutive heterochromatin. Furthermore, no position-dependent increase in transcript levels for *FRG1* and *ANT1* was observed, in contrast to the study by Gabellini *et al.* (2002) as discussed on page 64. These findings therefore make it unlikely that the D4Z4 arrays are the source of the *cis*-spreading heterochromatinisation. Based on this evidence Jiang *et al.* (2003) rejected the PEV hypothesis as a possible genetic mechanism implicated in FSHD.

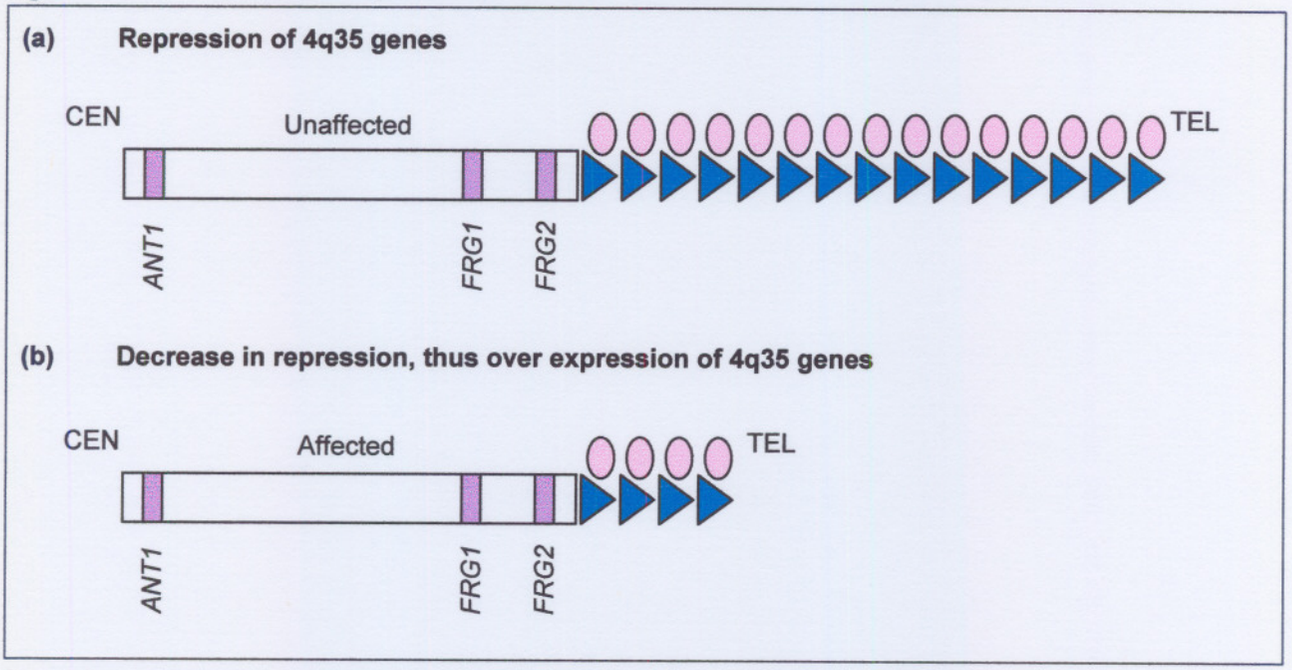
4.6.3 LONG DISTANCE LOOPING HYPOTHESIS

Jiang *et al.* (2003) proposed a new hypothesis, whereby intrachromosomal communication at the D4Z4 locus and the FSHD gene results from looping between the 3.3 kb repeats at the D4Z4 locus and a distant *cis* transcription control region. Since all FSHD patients have at least one copy of the 3.3 kb repeat element (Tupler *et al.*, 1996), it was proposed that the D4Z4 sequence itself activates the expression by long-distance looping interactions with control regions. The interaction could involve a 27 bp binding element within each D4Z4 repeat unit which was observed to bind several nuclear proteins, discussed in paragraph 4.6.4. The nuclear proteins may enable the upregulation of transcription of 4q35 genes as a result of looping interactions (Jiang *et al.*, 2003). The genotype-phenotype correlations observed in FSHD could also be explained by the looping hypothesis. It is proposed that a D4Z4 array with more than ten repeat elements are able to form stable intra-array interaction, those with five to ten copies establish less stable interactions, and arrays with one to four copies almost never form these interactions. The decrease in interaction between the D4Z4 array and the control region could thus explain the increased severity of symptoms and decreased age of onset that is associated with short D4Z4 arrays at 4q35.

4.6.4 REPRESSION COMPLEX HYPOTHESIS

Gabellini *et al.* (2002) reported the binding of a multiprotein complex (the D4Z4 recognition complex [DRC]), consisting of YY1 (Yin Yang 1 transcription factor), HMGB2 (high mobility group box 2 protein) and nucleolin, to a 27 bp sequence known as the D4Z4 binding element (DBE), located within each D4Z4 repeat unit, as depicted in Figure 4.6. The authors hypothesised that the DBE is a transcriptional repression element that leads to over expression of 4q35 genes when deleted, as in the case of FSHD. The DRC actively represses gene expression when bound to the DBE. In FSHD patients, however, an integral number of the 3.3 kb repeat elements, therefore also DBE's, are deleted, reducing the number of bound repressor complexes and a subsequent decrease in repression of genes localised on 4q35. The loss of active repression of gene expression therefore results in an over expression of genes ultimately causing FSHD. The mechanistic basis via which the DRC represses transcription is still unknown.

Figure 4.7: Schematic representation of the repression complex hypothesis



○ = D4Z4 repression complex comprising YY1, HMGB2 and nucleolin; ▶ = 3.3 kb repeat element. Adapted from Gabellini *et al.* (2002).

4.6.5 D4Z4 HYPOMETHYLATION HYPOTHESIS

Van Overveld *et al.* (2003) investigated the presence of possible epigenetic mechanisms underlying the pathogenesis of FSHD via methylation studies performed at the D4Z4 locus. It was observed that a deletion of 3.3 kb repeat elements at the D4Z4 locus results in hypomethylation of the locus. The authors induced hypomethylation at the D4Z4 locus in cultured cells resulting in the transcriptional upregulation of 4qter genes as was observed in the majority of expression studies in FSHD muscle (paragraph 4.5). Individuals clinically diagnosed with FSHD but without a 4q35 deletion fragment (non-4q-linked), were subsequently identified to be hypomethylated at the D4Z4 locus. This strongly supported an important role for D4Z4 hypomethylation in the aetiology of FSHD. The presence of hypomethylation at the D4Z4 locus in non-4q-linked individuals therefore has a significant implication for the diagnosis of FSHD in these patients. The authors therefore concluded that epigenetic factors do contribute to the pathogenicity of FSHD via the transcriptional upregulation of 4qter genes.

Several mechanisms implicated in FSHD have therefore been proposed (Winokur *et al.*, 1993; Hewitt *et al.*, 1994; Gabellini *et al.*, 2002; Jiang *et al.*, 2003; Van Overveld *et al.*, 2003). However, despite all the reports the exact defined molecular basis for FSHD still remains to be elucidated.

The FSHD locus is thus located in a complex region of the human genome. To further elucidate the role of the D4Z4 locus, as well as chromosomal plasticity in FSHD the evolutionary history of this locus will need to be investigated in the future. This highlighted the need for a molecular study in an older population. The study presented here is the first to investigate the D4Z4 locus in the Khoi-San and the Black South African populations, both more ancient than those previously reported.

The Khoi-San population cluster at the deepest node of the human phylogenetic tree, while the Black South African population form part of macrohaplogroup L (Chen *et al.*, 2000). The Black South African population therefore represents an intermediate between the Khoi-San population and other modern world populations. Investigating these older populations will allow insight into the plasticity of the D4Z4 locus over time. Through the study of the history of the repeat, it is envisioned that we would be able to gain more insight into certain phenomena of the repeat and FSHD observed today, such as: the translocations observed between chromosomes 4 and 10, the translocation frequency differences in different populations, and the process via which the FSHD associated deletion occurred.

4.7 OBJECTIVES OF THIS STUDY

The D4Z4 locus of different populations world wide have been extensively studied. However, no study has been performed in either the Khoi-San or the Black South African populations. The Khoi-San population, being the most ancient human population on earth, will be utilised to investigate the D4Z4 locus in an evolutionary context. Currently, no reports of black FSHD patients exist. The Black South African population will therefore be included to investigate the organisation of the D4Z4 locus and to compare that to other populations.

4.7.1 SPECIFIC OBJECTIVES OF THIS INVESTIGATION

Three objectives were investigated in this study:

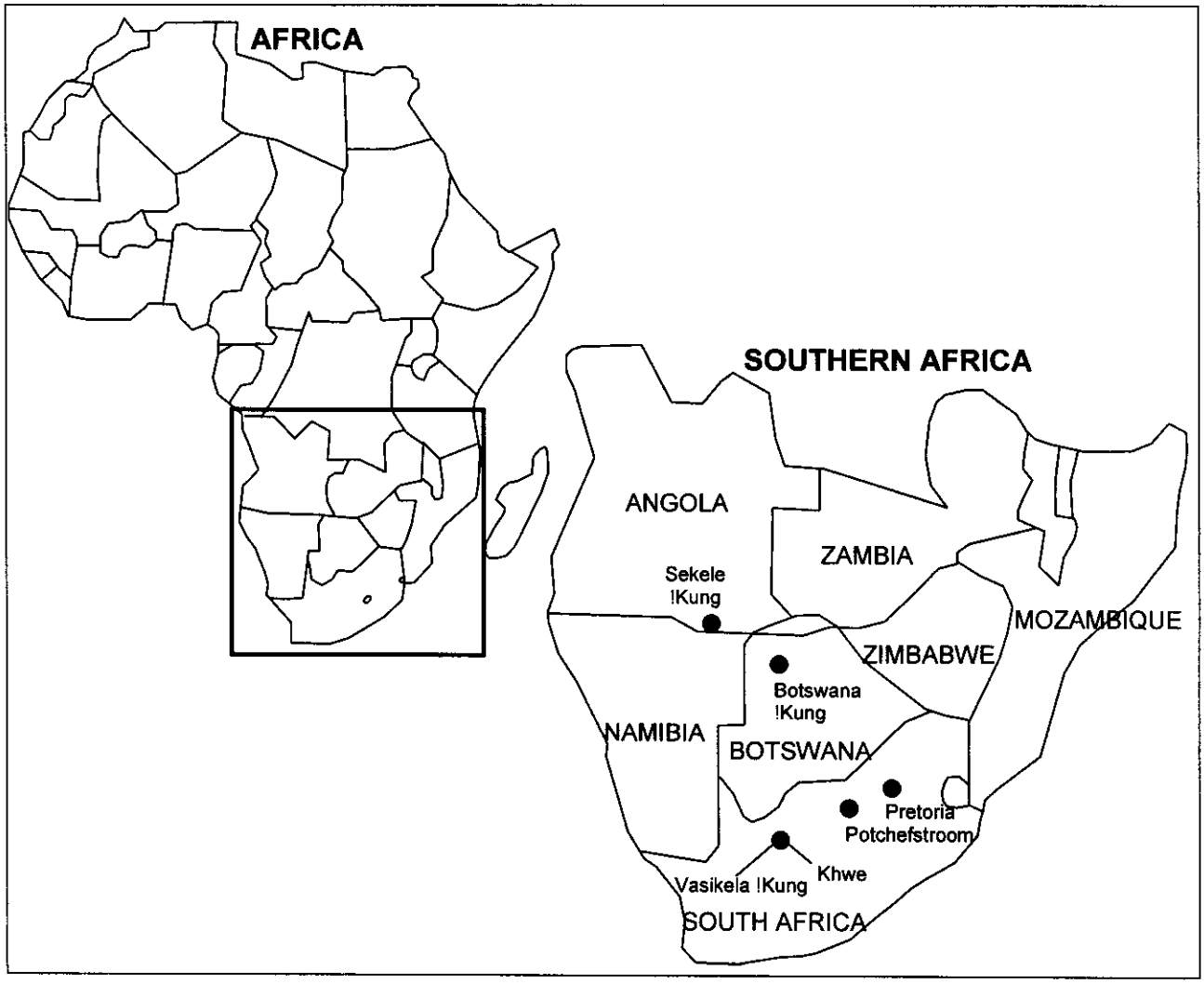
1. Determination of the translocation frequency between the homologous 3.3 kb *Kpn* I repeat units on chromosomes 4q35 and 10q26 in the Black South African population.
2. Elucidation of the translocation frequency in a sample population of the Khoi-San.
3. Investigation of the difference in the allele sizes and frequencies at the D4Z4 locus between the Black South African, the Khoi-San and other reported populations.

CHAPTER FIVE

MATERIALS AND METHODS

A group of 300 unrelated individuals was included in this study. This group consisted of 27 unrelated Khoi-San speaking !Kung and Khwe and 273 unrelated African samples from different ethnic groups (Figure 5.1). The individuals were classified according to the ethnicity listed on either their birth certificates or identity documents (Chen *et al.*, 2000). The samples of the 273 black South Africans were collected at several non-neuromuscular related clinics around Pretoria and Potchefstroom, South Africa (Figure 5.1).

Figure 5.1: Maps of Africa and Southern Africa indicating specific geographic locations



The samples of the 273 black South Africans were collected at non-neuromuscular related clinics around Pretoria and Potchefstroom. The 27 Khoi-San samples were collected from individuals of the Vasikela !Kung and Khwe populations. Adapted from Chen *et al.* (2000).

This study has been approved by the Ethics Committee of Potchefstroom University for Christian Higher Education with the approval number 02M07 for the study entitled "Molecular analysis of the South African facioscapulohumeral muscular dystrophy (FSHD) population". It was also approved previously by the Ethics Committee of the University of Pretoria with approval number 91/96. Informed consent was obtained from all individuals, or their legal guardians, prior to enrolment in the study.

The translocation frequency between chromosomes 4q35 and 10q26 was evaluated via the *Bln* I / *Bgl* II dosage test (paragraphs 5.3) and confirmed via pulsed field gel electrophoresis (paragraph 5.2). The allele sizes at the D4Z4 locus were determined via pulsed field gel electrophoresis.

5.1 ISOLATION OF HIGH MOLECULAR WEIGHT GENOMIC DNA

Different isolation methods were utilised for the isolation of high molecular weight genomic DNA from both immortal lymphoblastoid cell lines and whole blood. For PFGE the gDNA was embedded into agarose plugs (paragraphs 5.1.1 and 5.1.2) and for the dosage test liquid gDNA (paragraph 5.1.3) was isolated.

5.1.1 Culture of immortal lymphoblastoid cell lines[1]

Immortal lymphoblastoid cell lines had already been established for the Khoi-San samples (Chen *et al.*, 2000). The immortal lymphoblastoid cells were rapidly thawed at 37 degrees centigrade (°C) and the tube drenched in 70% ethanol for sterilisation purposes. The cells were placed in 9 millilitres (ml) pre-warmed RPMI (Roswell Park Memorial Institute) media [GIBCO™[2]] supplemented with 10% heat inactivated foetal bovine serum and 2 millimolars (mM) glutamine in a 15 ml tube. The cells were collected by centrifugation at 180 gravitational accelerations (x g) for 5 minutes.

5.1.1.1 Preparation of agarose plugs from cultured cells[3]

The cultured cells were washed three times via centrifugation in ice-cold phosphate-buffered saline (PBS) buffer [137 mM sodium chloride (NaCl), 2.7 mM potassium chloride

[1] The culture of the immortal lymphoblastoid cell lines was performed by C. Mouton, according to a protocol optimised by the author.
[2] GIBCO™ is a trademark of Invitrogen, Inc., Carlsbad, CA., U.S.A.
[3] The preparation of the agarose plugs from cultured cells was performed by C. Mouton, according to a protocol optimised by the author.

(KCl), 10 mM sodium phosphate dibasic (Na_2HPO_4, pH 7.4), 2 mM potassium phosphate monobasic (KH_2PO_4,)], counted and 2 x 10^7 cells resuspended in 1 ml of SE buffer (75 mM NaCl, 25 mM ethylenediamine tetraacetic acid (EDTA), at pH 8.0). The cell suspension was heated to 42°C to facilitate the resuspension of the cells. A solution of 1.4% InCert®[1] agarose [Cambrex] melted in SE buffer was prepared and incubated at 60°C to avoid coagulation of the agarose. An equal volume of SE buffer / 1.4 % agarose to that of SE buffer was added to the heated cell suspension. To ensure even distribution of the cells throughout the agarose, the cell suspension was gently mixed after which aliquots of 85 microlitres (µl) were dispersed into agarose plug molds [Bio-Rad®[2]]. The plugs were coagulated at 4°C for 30 minutes.

A solution containing 10 ml SE buffer, 1% N-Lauroylsarcosine (Sigma) and 300 µl of a 20 milligram (mg).ml^{-1} pronase solution (Roche) was prepared in a 15 ml tube. The coagulated plugs were pushed into the solution utilising the air pressure of a pipetting balloon and incubated at 37°C for two days. At day three the SE/N-Lauroylsarcosine/pronase solution was replaced with 10 ml double distilled water (ddH_20) and subsequently with 10 ml 0.5 M EDTA (pH 8.0). The agarose plugs were stored in the 0.5 molar (M) EDTA (pH 8.0) at 4°C until utilised for restriction enzyme digestion (paragraph 5.2.1).

5.1.2 Preparation of agarose plugs from whole blood samples

Whole blood was collected in tubes containing EDTA as the preservative. The samples were stored at -70°C until the DNA was isolated. Before DNA isolation the blood was rapidly thawed at 37°C upon which the EDTA tube was inverted several times to ensure a homogeneous suspension. Aliquots of 10 ml of whole blood were transferred to a sterile 50 ml tube (Sterilin) upon which the blood tube was rinsed with 1 x erythrocyte lysis buffer (155 mM ammonium chloride (NH_4Cl, pH 8.0), 10 mM potassium hydrogen carbonate ($KHCO_3$), 1 mM EDTA) and added to a 50 ml tube containing the 10 ml whole blood. Erythrocyte lysis buffer was added to a final volume of 40 ml, the solution was mixed by inversion and the tube put on ice for 5 minutes to ensure sufficient red blood cell lysis.

The lymphocytes were collected via centrifugation at 515 x g, at 15°C for 8 minutes in an Eppendorf®[3] 5810R centrifuge, and the supernatant discarded. The pellet was

[1] InCert® is a registered trademark of CBM Intellectual properties, Inc., East Rutherford, NJ., U.S.A.
[2] Bio-Rad® is a registered trademark of Bio-Rad Laboratories, Hercules, CA., U.S.A.
[3] Eppendorf® is a registered trademark of Eppendorf AG, Hamburg, Germany.

resuspended with 30 ml erythrocyte lysis buffer and the solution mixed to obtain a homogenised suspension. Half of the suspension (15 ml) was aliquoted into a 15 ml tube and were further utilised for the preparation of agarose plugs. The remaining 15 ml suspension in the 50 ml tube was utilised to prepare liquid gDNA (paragraph 5.1.3.1). Both tubes were incubated on ice for 5 minutes and the lymphocytes collected via centrifugation at 515 x g, 15°C for 8 minutes.

The pellet in the 15 ml tube was resuspended in 250 µl of SE buffer after which an equal volume of 1.4% InCert® agarose, melted in SE buffer and incubated at 60°C to avoid coagulation of the agarose, was added. The solution was gently mixed to ensure an equal distribution of cells throughout the solution and aliquots of 85 µl dispensed into plastic molds and incubated at 4°C to coagulate. The coagulated plugs were pushed into a 15 ml tube containing 10 ml SE buffer, 1% N-Lauroylsarcosine and 300 µl of a 20 $mg.ml^{-1}$ pronase solution and incubated at 37°C for two days. At day three the SE/N-Lauroylsarcosine/pronase solution was replaced with 10 ml ddH_20 and subsequently with 10 ml 0.5 M EDTA (pH 8.0). The agarose plugs were stored in the 0.5 M EDTA (pH 8.0) at 4°C until utilised for restriction enzyme digestion (paragraph 5.2.1).

5.1.3 Preparation of liquid gDNA from whole blood

Standard DNA isolation protocols could be performed for the isolation of liquid gDNA, since the fragments of interest for the dosage test are only 4,061 bp and 1,774 bp respectively. Three different isolation procedures were utilised as discussed in paragraphs 5.1.3.1 to 5.1.3.3.

5.1.3.1 Isolation of liquid gDNA

The collected lymphocytes (paragraph 5.1.2) were resuspended in a 50 ml tube with 3 ml nucleus lysis buffer (10 mM 2-amino-2-2(hydroxymethyl)-1,3-propanediol hydrochloride (Tris-HCl, pH 8.2), 2 mM EDTA, 400 mM NaCl), 150 µl of 20% sodium dodecyl sulphate (SDS) and 100 µl pronase (20 $mg.ml^{-1}$) and the suspension incubated at 37°C for 2 days. At day three 1.2 ml 5 M NaCl was added to the lysed cells and the tube thoroughly mixed. The precipitated protein was collected via centrifugation at 3,100 x g, at 15°C for 15 minutes and the supernatant transferred into a clean 15 ml tube. The DNA was precipitated with 10 ml 100% ethanol, collected with a glass Pasteur pipette and dissolved in 250 µl TE buffer (10 mM Tris-HCl (pH 7.5), 1 mM EDTA). DNase inactivation was

performed for 10 minutes at 65°C and the liquid DNA stored at 4°C until utilised for restriction enzyme digestion (paragraph 5.3.1).

5.1.3.2 Isolation of gDNA utilising the Wizard®[1] Genomic DNA Purification Kit

The Wizard® Genomic DNA purification kit [Promega] 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 5.1).

Table 5.1: 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 [Sterilin] for 10 ml whole blood samples. The tube of blood was gently rocked 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 x g 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

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

and vortexed vigorously until small protein clumps were visible. The samples were centrifuged at 2,000 x g 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 gDNA 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 x g for one minute at room temperature the DNA was visible as a small white pellet. The supernatant 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 x g 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 microgram (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 250 nanograms (ng).μl^{-1} were made. The working dilutions were stored at 4°C.

5.1.3.3 Isolation of gDNA utilising the NucleoSpin®[1] Blood kit

The NucleoSpin® Blood kit utilises a silica membrane for the specific binding of DNA. The binding process is reversible allowing the isolation of gDNA under low ionic strength conditions in an alkaline elution buffer. A modified method to that described in the kit manual was utilised as to allow for the isolation of a larger volume of blood (3 ml instead of 200 µl) per patient.

The red blood cells of 3 ml of whole blood were lysed by the addition of 3 ml of C1 buffer (4% volume per volume (v/v) Triton® X-100 [USB], 43 mM magnesium chloride ($MgCl_2$) [Merck], 40 mM Tris [USB] and 1.3 M sucrose [Sigma]), pre-chilled to 4°C, in a total

[1] NucleoSpin® is a registered trademark of Macherey-Nagel, Düren, Germany.

volume of 15 ml. The suspension was centrifuged at 4°C for 15 minutes at 1,300 x g and the supernatant subsequently discarded. The pelleted leucocytes were resuspended in 400 µl PBS buffer via vortexing for 1 minute and transferred to a 1.5 ml Eppendorf® tube. Subsequent to the transfer, 400 µl B3 buffer (supplied in the kit) and 50 µl Proteinase K (23 $\mu g.\mu l^{-1}$) were added to the resuspended leucocytes and the suspension incubated for 10 minutes at 70°C. Following incubation the suspension was vortexed for 1 minute after the addition of 420 µl 100% ethanol.

Half of the suspension was added to a NucleoSpin® blood column, centrifuged for 2 minutes at 11,000 x g and the flow-through discarded. The remaining suspension was then also added to the column, centrifuged for 2 minutes at 11,000 x g and the flow-through discarded. To ensure that the column was completely dry, it was centrifuged for an additional 2 minutes. The DNA bound to the column was subsequently washed with 700 µl B5 buffer (supplied in the kit), centrifuged at 11,000 x g for 1 minute and the flow through discarded. A second wash with 300 µl B5 buffer was performed, the column centrifuged at 11,000 x g for 2 minutes and the flow through discarded.

After washing of the DNA the NucleoSpin® blood column was transferred to a clean 1.5 ml tube and 200 µl elution buffer, pre-warmed to 70°C, added. The column was incubated at 70°C for 15 minutes followed by centrifugation at 11,000 x g for 2 minutes. The elution, incubation and centrifugation steps were repeated to ensure optimal DNA recovery.

The average DNA yield per 3 ml whole blood obtained with the NucleoSpin® Blood kit was 110 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 250 $ng.\mu l^{-1}$ were made. The working dilutions were stored at 4°C.

5.2 PULSED FIELD GEL ELECTROPHORESIS

The identification of a chromosome 10 specific *Bln* I restriction site within each chromosome 10 derived unit, and a chromosome 4 specific *Xap* I restriction site within each chromosome 4 derived unit, facilitated the discrimination between chromosome 4 and chromosome 10 derived units. The analysis with *Eco* RI / *Hind* III, *Eco* RI / *Bln* I, and *Xap* I therefore allows characterisation of each of the four alleles (two from chromosome 4

and two from chromosome 10) and enables the visualisation of specific allele sizes and translocation events between chromosomes 4q35 and 10q26. The utilisation of the triple analysis and PFGE can also be utilised to detect hybrid repeat arrays, as well as in cases where the probe region is deleted or even for somatic mosaicism for the repeat array.

5.2.1 Restriction enzyme digestion of agarose embedded gDNA for PFGE analysis

Restriction fragment length polymorphism (RFLP) analysis was performed on the high molecular weight gDNA embedded in agarose plugs. Five µg of gDNA was utilised in each reaction. After the storage in 10 ml 0.5 M EDTA the plugs were washed in 150 µl ddH_20 and subsequently equilibrated twice for 2 hours in 150 µl TE buffer before each digestion. The TE buffer was removed after the second equilibration step and replaced with the appropriate 1 x restriction enzyme buffer as presented in Table 5.2. The agarose plugs were incubated overnight at 37°C.

Table 5.2 Restriction enzyme and buffer combinations

Restriction enzyme reaction	1 x restriction enzyme buffer composition
Double digestion: *Eco* RI (MBI Fermentas) and *Hind* III (MBI Fermentas)	Buffer B (Roche) 10 mM Tris-HCl (pH 8.0), 100 mM NaCl. 5 mM $MgCl_2$, 1 mM mercaptoethanol
Double digestion: *Eco* RI (MBI Fermentas) and *Bln* I (Amersham Biosciences™[1])	Buffer K (Amersham Biosciences™) 50 mM Tris-HCl (pH 7.5), 100 mM NaCl. 10 mM $MgCl_2$, 1 mM DTT
Single digestion: *Xap* I (MBI Fermentas)	YTango buffer (MBI Fermentas) 33 mM Tris-acetate (pH 7.9 at 37°C), 10 mM magnesium acetate, 66 mM potassium acetate, 0.1 $mg.ml^{-1}$ BSA

Following overnight incubation the respective 1 x restriction enzyme buffer was replaced with identical fresh 1 x restriction enzyme buffer supplemented with 3.3 mM spermidine [Sigma], 1 mM dithiothreitol (DTT) [Promega] and 20 units (U) of the appropriate restriction enzyme. The restriction endonuclease reactions were performed in a final volume of 150 µl. Reactions were incubated at 37°C for at least six hours. The restriction enzyme buffer was removed from the agarose plug after the incubation prior to the loading on the agarose gel as described in paragraph 5.2.2.

[1] Amersham Biosciences™ is a trademark of Amersham plc, Amersham Place Little Chalfont, Buckinghamshire, UK.

5.2.2 Agarose gel electrophoresis for pulsed field analysis

The agarose plugs were carefully loaded into separate wells utilising a scalpel, subsequent to the restriction enzyme digestion. Appropriate molecular weight markers, CHEF lambda ladder [Bio-Rad] and CHEF 8-48 kb ladder [Bio-Rad] were included to determine the fragment sizes. The 1% agarose gel (Hispanagar D5) was electrophoresed in 0.5 x TBE buffer (44.5 mM Tris base [USB] (pH 8.0), 44.59 mM Boric acid [USB]; 1.25 mM disodium EDTA (Na_2EDTA) [Associated Chemical Enterprises (ACE)]), at 8.5 volts per centimetres ($V.cm^{-1}$) for 22 hours at 21°C in four identical cycles. During each cycle, the switch time was linearly increased from 1 second to 16 seconds. The analyses were performed on a Hoefer, HG 1000 Hula gel pulsed-field electrophoresis unit.

5.2.3 Genomic DNA transfer

Following electrophoresis the agarose gel was visualised with an ultraviolet (UV) light before Southern transfer. The gDNA was nicked with 0.25 M hydrochloric acid (HCl) [ACE] for 10 minutes and then denatured with fresh alkaline blotting solution (0.4 M sodium hydroxide (NaOH) [ACE], 1.6 M NaCl [ACE]) twice for 20 minutes with gentle shaking. Hybond™[1] N^+ membrane [Amersham Biosciences] was pre-wet for 10 minutes in distilled water and subsequently placed for 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 [Roche] (pH 8.0), 2 x saline-sodium-citrate (SSC) buffer - 0.3 M NaCl [ACE], 30 mM trisodium citric acid (Na-citrate) [ACE]). The transfer efficiency was determined by staining the gel after the overnight transfer with 2 $\mu g.ml^{-1}$ ethidium bromide (EtBr) [Promega] dissolved in 1 x TBE buffer for 30 minutes followed by visualisation via UV light.

5.2.4 Isolation of probe p13E-11

Probe p13E-11 was originally cloned as an 800 bp insert into the pBluescript plasmid (Wijmenga *et al.*, 1992b). The plasmid was initially isolated from an overnight culture with the Wizard® Purefection Plasmid DNA purification system [Promega] (paragraph 5.2.4.1) followed by the excision of the 800 base pairs (bp) insert from the plasmid through a

[1] Hybond™ is a trademark of Amersham Biosciences Limited, Amersham Place Little Chalfont, Buckinghamshire, UK.

double digestion with *Sac* I and *Eco* RI. As the culturing of the plasmid is time consuming the probe was subsequently isolated from the purified plasmid DNA via the Polymerase Chain Reaction (PCR) as described in paragraph 5.2.4.2. The 800 bp insert was initially excised from the amplicon through a double digestion with *Sac* I [Amersham Biosciences) and *Eco* RI [Fermentas], however, analysis indicated that the hybridisation of the probe containing the primer sequences did not differ to that of the original 800 bp isolated probe.

5.2.4.1 Isolation of probe p13E-11 from an overnight culture

Probe p13E-11 was cloned as an 800 bp insert into the pBluescript plasmid. Five millilitres of LB-broth [Promega] containing ampicillin was inoculated with the p13E-11 stab culture and incubated with shaking overnight at 37°C. Five ml 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 x g for 10 minutes at 22°C. The supernatant was discarded and the excess liquid was blotted on a paper towel. Six ml of Cell Resuspension Solution was added to each tube and the cell pellet resuspended by vortexing. The cells were lysed by adding 6 ml of Cell Lysis Solution and mixed thoroughly 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 x g 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

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

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 ml 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 high quality water 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 supernatant 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 x g 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 x g 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 via spectrophotometry (Eppendorf® Biophotometer).

5.2.4.2 Isolation of probe p13E-11 via polymerase chain reaction

A modified method described by Mullis *et al.* (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 [Costar®[1] Thermowell™[2]].

[1] Costar® is a registered trademark of Corning Incorporated, Corning, NY, U.S.A.
[2] Thermowell™ is a trademark of Corning Incorporated, Corning, NY, U.S.A.

The PCR reagents were thawed on ice and thoroughly mixed by vortexing prior to use. PCR was performed in reactions containing 10 mM Tris-HCl (pH 9.0), 1.5 mM $MgCl_2$, 50 mM KCl, 0.1% Triton® X-100, 200 micromolar (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)], 5.0 picomoles (pmol) of each primer and 0.5 U Taq polymerase [Promega]. 100 ng plasmid DNA 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 55°C 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 Thermo Hybaid®[1] MultiBlock System® (MBS®).

5.2.5 Labelling of probe p13E-11 and molecular weight marker

The *Gene Images*™[2] Random Prime labelling module (Amersham Biosciences) was utilised for the non radio-active detection of the fragments at the D4Z4 locus. The random primer kit utilises nonamers of random sequence as primers for the replication of the template DNA. The probes are labelled through the incorporation of fluorescein-11-2'-deoxyuridine-5'-triphosphate (Fl-dUTP) during the random-primer extension by Klenow DNA polymerase. The labelled probe is then detected with the *Gene Images* CDP-*Star*®[3] detection module (Amersham Biosciences) as described in paragraph 5.2.8.

One hundred ng of p13E-11 probe or molecular weight marker DNA was dissolved in a final volume of 34 μl in diethyl pyrocarbonate (DEPC) treated water. The DNA was denatured at 95°C for 5 minutes and immediately transferred to ice. To a 1.5 ml microcentrifuge tube on ice 10 μl of nucleotide mix (fluorescein-11-dUTP, dATP, dCTP, dGTP and dTTP in Tris-HCl, pH 7.8, 2-mercaptoethanol and $MgCl_2$), 5 μl of primer, the 34 μl denatured DNA and 5 U Klenow enzyme were added in a final volume of 50 μl. The reaction was gently mixed and incubated at 37°C for 1-4 hours. Na_2EDTA (pH 7.5) to a final concentration of 20 mM was subsequently added to terminate the reaction.

---

[1] Thermo Hybaid® is a registered trademark of Hybaid Limited, Ashford, Middlesex, UK.
[2] Gene Images™ is a trademark of Amersham Biosciences Limited, Amersham Place Little Chalfont, Buckinghamshire, UK.
[3] CPD-Star® is a registered trademark of Tropix Inc., Bedford, MA, U.S.A.

5.2.5.1 Monitoring of labelling efficiency

To ensure optimal signal after hybridisation, the efficiency of the labelling of the p13E-11 probe and the molecular weight markers were determined. A series of dilutions, 1/5, 1/10, 1/25, 1/50, 1/100, 1/250 and 1/500, were prepared from the 5 x nucleotide mix included in the *Gene Images*™ Random Prime labelling module (Amersham Biosciences), in 1 x TE buffer. A detection strip was prepared via the dotting of 5 µl of the labelled probe and molecular weight marker onto a Hybond N^+ membrane (Amersham Biosciences) together with 5 µl of the 1/5 dilution of the nucleotide mix. The 1/5 dilution was added to the detection strip as a negative control. The detection strip was subsequently washed in pre-heated 2 x SSC at 60°C for 15 minutes.

A reference strip was also prepared. Five µl of each of the dilutions, except the 1/5 dilution, were dotted onto a Hybond N^+ membrane (Amersham Biosciences). Both the reference and the detection strips were placed on a UV transilluminator. The fluorescent spots of the labelled probe and the molecular weight marker on the detection strip should have intensities between the 1/10 and 1/250 fluorescent spots on the reference strip. This would indicate that the probe and molecular weight marker have successfully incorporated the fluorescein label. The more comparable the signal intensity of the probe and molecular weight marker is to the 1/10 dilution, the more efficient was the labelling reaction. Subsequent to the washing step, the 1/5 dilution on the detection strip should contain little fluorescence, indicating that the fluorescence of the probe is due to incorporated fluorescein only, and not unincorporated Fl-dUTP.

5.2.6 Hybridisation

The Hybond N^+ membrane with the target DNA attached was prehybridised in 0.125 ml per cm^2 membrane hybridisation buffer (5 x SSC - 0.75 M NaCl [ACE], 75 mM Na-citrate [ACE]), 1 in 20 dilution liquid block (supplied in *Gene Images*™ kit), 0.1% weight per volume (w/v) SDS and 5% (w/v) Dextran sulphate (Amersham Biosciences) for 1 hour at 65°C in a Techne®[1] Roller-Blot Hybridiser HB-3D. The labelled p13E-11 probe and molecular weight marker were denatured at 95°C for 5 minutes and immediately placed on ice. The denatured probe and molecular weight markers were subsequently added to the prehybridisation solution to a final concentration of 10 $ng.ml^{-1}$ and incubated for 16 hours at 65°C.

[1] Techne® is a registered trademark of Techne Incorporated, Princeton, NJ, U.S.A.

5.2.7 Stringency washes

Following hybridisation the membrane was washed to remove any excess probe. The membrane was washed in at least 2 ml per cm^2 of membrane 1 x SSC (150 mM NaCl [ACE], 15 mM Na-citrate [ACE]), 0.1% (w/v) SDS, pre-warmed to 65°C for 15 minutes. A second wash with 0.5 x SSC (75 mM NaCl [ACE], 15 mM Na-citrate [ACE]), 0.1% (w/v) SDS at 65°C was performed for 15 minutes.

5.2.8 Non radio-active detection

The *Gene Images* CDP-*Star*® detection module (Amersham Biosciences) was utilised for the non radio-active detection. All the containers utilised for the detection were rinsed with 70% ethanol before use to remove any bacterial alkaline phosphatase contamination. Following the stringency washes (paragraph 5.2.7) the membrane was incubated with gentle agitation for 1 hour at room temperature in 1.0 $ml.cm^{-2}$ of a 1 in 10 dilution of liquid blocking agent (supplied in kit) in buffer A (100 mM Tris-HCl (pH 9.5), 300 mM NaCl). Buffer A was prepared fresh and was autoclaved before each detection to avoid any contamination with exogenous alkaline phosphatase. The membrane was then incubated in a 5000 x dilution of anti-fluorescein-AP-conjugate, 0.5% (w/v) bovine serum albumin (Amersham Biosciences) in buffer A with gentle agitation at room temperature for 1 hour. Any unbound conjugate was removed by washing the membrane three times for 10 minutes at room temperature in 2 $ml.cm^{-2}$ 0.3% (v/v) Tween™20 [USB] diluted with buffer A.

After the completion of the antibody incubation and subsequent washes any excess wash buffer was drained off from the membrane. The detection reagent, CDP-Star®, was added to the membrane to a volume of 30 $\mu l.cm^{-2}$ and incubated at room temperature for 5 minutes. The excess detection reagent was drained off and the membrane heat-sealed in a hybridisation bag.

5.2.9 Autoradiography

The heat-sealed membrane was placed in a film cassette [Hypercassette™[1], Amersham Biosciences] and exposed to chemiluminescent detection X-ray film [Roche]. The X-ray

[1] Hypercassette™ is a trademark of Amersham Biosciences Limited, Amersham Place Little Chalfont, Buckinghamshire, UK.

film was exposed to the membrane for an appropriate time, 1 hour to overnight, where after the X-ray film was developed (Africa X-ray Industrial and Medical (AXIM) Pty (Ltd) X-ray developer) and fixed (AXIM hi-speed X-Ray fixer) to allow the visualisation of the fragments.

5.3 THE *Bgl* II / *Bln* I DOSAGE TEST

The *Bgl* II – *Bln* I dosage test was utilised in this study. This method of analysis is generally employed to study translocation events in large populations.

5.3.1 Restriction fragment length polymorphism analysis for dosage test analyses

Liquid gDNA was available for all the individuals of the Black South African population included in this study. However, only agarose embedded gDNA was available for the Khoi-San samples. The RFLP strategies for the two different sample types are described in paragraphs 5.3.1.1 and 5.3.1.2.

5.3.1.1 Restriction enzyme digestion of liquid gDNA for dosage test analysis

RFLP analysis was performed on the isolated liquid gDNA described in paragraph 5.1.3. A total of 2.5 to 5 μg of gDNA was utilised in each reaction. Reactions were incubated at 37°C for at least six hours, in the presence of 20 U *Bln* I [Amersham Biosciences], 20 U *Bgl* II [Amersham Biosciences], 20 mM Tris-HCl (pH 8.5), 10 mM $MgCl_2$, 100 mM KCl, 1 mM DTT and 3.3 mM spermidine [Sigma] in a final volume of 30 to 40 μl.

5.3.1.2 Restriction enzyme digestion of agarose embedded gDNA for dosage test analysis

The agarose embedded gDNA was digested as described in paragraph 5.2.1. The only deviation from the protocol presented in paragraph 5.2.1 was the utilisation of 20 U *Bln* I [Amersham Biosciences] and 20 U *Bgl* II [Amersham Biosciences], 20 mM Tris-HCl (pH 8.5), 10 mM $MgCl_2$, 100 mM KCl, and 1 mM DTT instead of the enzymes and buffers listed in Table 5.2.

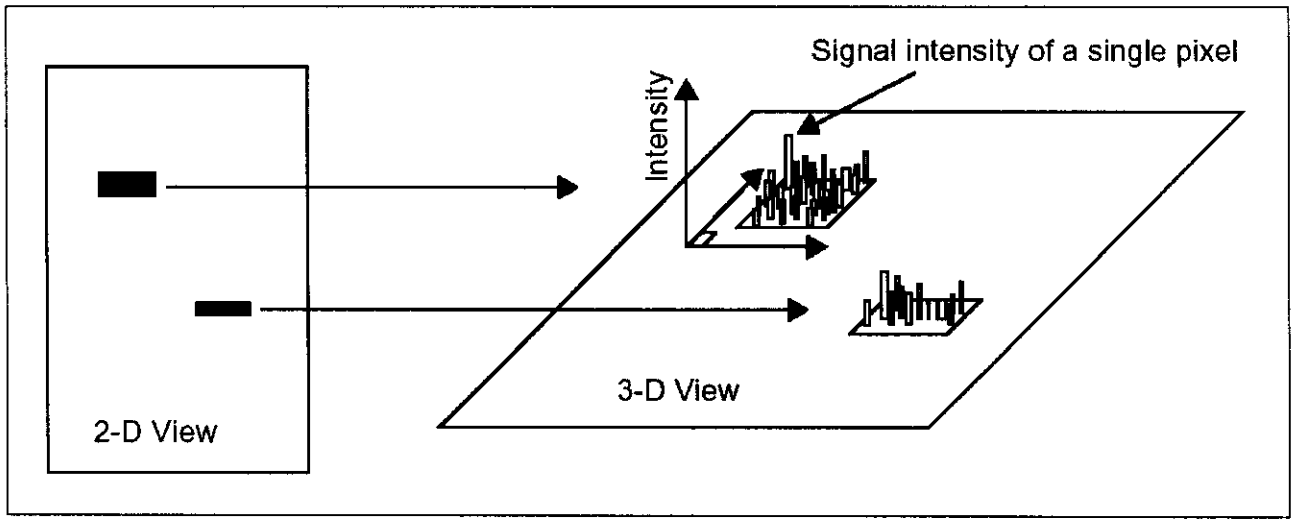
5.3.2 Agarose gel electrophoresis for *Bln* I / *Bgl* II dosage test analysis

Following incubation 2 µl of the RFLP reaction was mixed with 1 µl of a 2 x loading buffer (0.04 % Orange G [Sigma®] and 50 % glycerol) prior to loading on a 0.8% agarose gel. After confirmation of complete digestion the remaining product was loaded onto a maxi (20 cm x 25 cm) 0.8% agarose gel along with a molecular weight marker, 1 kb DNA ladder [Promega], to determine the fragment sizes. The agarose gel was electrophoresed in 1 x TBE buffer (89.15 mM Tris base; 89.95 mM Boric acid; 2.498 mM Na_2EDTA, at a pH of 8.0) at 40 V (1.6 $V.cm^{-1}$) for 24 hours. The gDNA transfer, probe labelling, hybridisation and detection were performed as described in paragraphs 5.2.3 to 5.2.9. Subsequent to autoradiography (as described in paragraph 5.2.9) the X-ray film was scanned (HP ScanJet 4300C) for further analysis via the Quantity One®[1] version 4.4.1 (v 4.4.1) software package [Bio-Rad®].

5.3.3 Fragment intensity quantification via the Quantity One® v 4.4.1 software program

Image analysis was performed with the Quantity One® v 4.4.1 program. In this study, the volume of a fragment indicated the total signal intensity inside a defined boundary. The total signal intensity of a particular fragment therefore signified the sum of the intensities of the pixels within the volume boundary multiplied by the pixel area. The analyses were thus performed on a three dimensional level, with the defined pixel area on the X and Y axes, and the signal intensity of the pixel area on the Z axis, as presented in Figure 5.2.

Figure 5.2: Schematic representation of the quantification of fragment intensities



Reprinted from the Quantity One® user guide with permission from the Imaging Technical Support, Bio-Rad® Laboratories, Hercules, CA, U.S.A..

[1] Quantity One® is a registered trademark of Bio-Rad® Laboratories, Hercules, CA, U.S.A.

The volume boundary was identified via the positioning of a rectangular block around the fragment of interest. The blocks positioned over the chromosome 4 and 10 fragments were labelled U1, U2, U3, etc., indicating that it was a fragment with unknown intensity with the number indicating the order in which the blocks were created. Rectangular blocks were also utilised to determine the background intensities of each fragment and were labelled B1, B2, B3, etc. As the total signal intensity was dependent on the pixel area, it was crucial to ensure that the volume boundary for the fragments of interest as well as the background boundary for a particular sample was identical. This was achieved via defining a fixed area which was copied and superimposed onto the remaining fragments. The signal intensities of the fragments and the backgrounds were exported as a spreadsheet to Microsoft®[1] Excel for further analysis.

5.3.4 Calculation of fragment intensities

The intensity for each fragment was corrected for the background around that particular fragment. The standard ratio (R_s) between the intensities of chromosomes 4 and 10 specific signals was obtained from three individuals harbouring normal repeat array distributions (disomic individuals). This R_s was utilised as a correction factor. For each individual the contribution of chromosome 4 specific signal to the sum of the chromosomes 4 (S4) and 10 (S10) signals was calculated via Equation 5.1.

Equation 5.1: Calculation of contribution of chromosome 4 signal

$$\text{Contribution of chromosome 4 signal} = (S4/R_s)/[(S4/R_s)+S10]$$

From Van der Maarel *et al.* (1999).

The ratio, expected value and contribution of the chromosome 4 signal for the seven possible profiles; disomy, trisomy, quatrosomy, monosomy, nullisomy and the deletion of the p13E-11 hybridisation site on either chromosome 4 or 10, are listed in Table 5.3. For an individual harbouring no translocation between chromosome 4q35 and 10q26, a disomic individual, a ratio of 2:2 with a value of 1.00 is expected. The contribution of the chromosome 4 specific signal to the total signal would therefore be 0.5. The expected

[1] Microsoft® is a registered trademark of Microsoft Corporation in the United States and other countries.

value and the contribution of the chromosome 4 signal value were utilised to classify the individuals analysed in this study according to the seven categories.

Table 5.3: Expected ratios between the chromosome 4q35 and 10q26 signal intensities

Profile		4:10 ratio	Expected value	Contribution of chromosome 4 signal[a]
Disomy		2:2	1.00	0.5
Trisomy	4-on-10	3:1	3.00	0.75
Quatrosomy		4:0	∞	1.0
Monosomy	10-on-4	1:3	0.333	0.25
Nullisomy		0:4	0	0
Deletion of p13E-11 hybridisation site on chr 4		1:2	0.5	0.33
Deletion of p13E-11 hybridisation site on chr 10		2:1	2	0.66

a = as determined via Equation 5.1.

5.3.5 Chi-square analysis

The chi-square (χ^2) test was utilised to determine whether the differences observed between the Khoi-San and the Black South African populations were statistically significant. The null hypothesis for the χ^2 analysis was that the differences observed between the two populations was not statistically significant, while the alternative hypothesis was that the difference observed was statistically significant. The χ^2 value was determined utilising Equation 5.2.

Equation 5.2: Calculation of Chi-square

$$\chi^2 = \frac{(O - E)^2}{E}$$

χ^2 = Chi-square; O = Observed value; E = Expected value. From Samuels (1989).

5.4 AMPLIFICATION OF mtDNA FOR HAPLOGROUP ANALYSIS

The primers listed in Table 5.4 were utilised for the amplification of specific regions of the mtDNA each consisting of ca. 600 bp each. The amplification of the regions was performed as described in paragraph 5.2.4.2 utilising the T_a as indicated in the table below.

Table 5.4: Primer information for haplogroup analysis

Primer	Sequence[1]	T_m	Mean T_m	Optimised T_a[2]	Size (bp)
L3073	F: 5'-aaagtcctacgtgatctgagttc-3'	53	53	55	639
H3670	R: 5'-ggcgtagtttgagtttgatgc-3'	55			
L6869	F: 5'-ccggcgtcaaagtatttagc-3'	53	53	58	578
H7406	R: 5'-gggttcttcgaatgtgtggtag-3'	53			
L9886	F: 5'-tccgccaactaatatttcactt-3'	51	50	55	617
H10462	R: 5'-aatgaggggcatttggtaaa-3'	49			
L10403	F: 5'-aaaggattagactgaaccgaa-3'	50	50.5	56	612
H10975	R: 5'-ccatgattgtgaggggtagg-3'	51			
L11486	F: 5'-aaaactaggcggctatggta-3'	51	51	56	629
H12076	R: 5'-ggagaatgggggataggtgt-3'	51			

1 = Maca-Meyer *et al.* (2001). 2 = The annealing temperatures of the primers were optimised by T. van Brummelen.

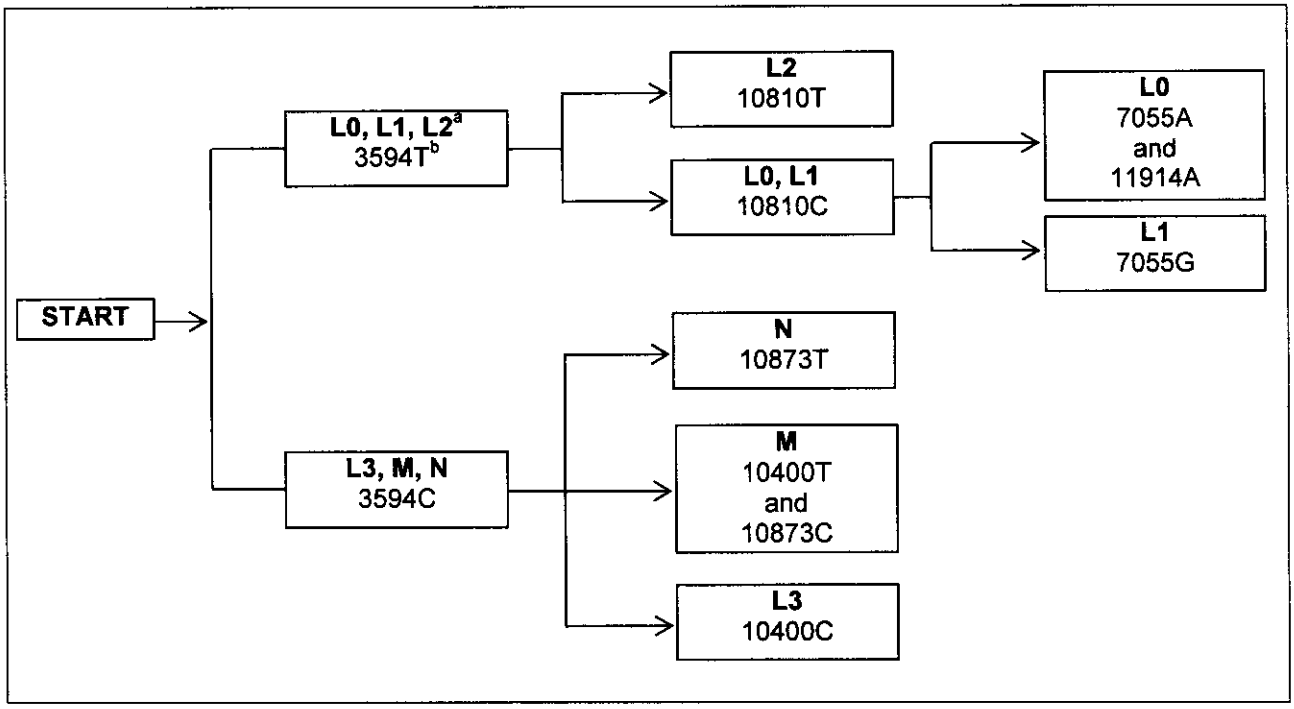
5.4.1 Agarose gel electrophoresis of amplified mtDNA for haplogroup analysis

Agarose gel electrophoresis was utilised to determine the success of the amplification of the five mtDNA regions of interest prior to RFLP analysis or cycle sequencing. Two µl of the amplified products were electrophoresed on a 2% agarose gel along with an appropriate molecular weight marker. The agarose gel was electrophoresed in 1 x TBE buffer at 100 V (5 $V.cm^{-1}$) for 30 minutes.

5.4.2 Differentiation between specific haplogroups

Five single nucleotide polymorphisms (SNPs) within the mtDNA, as illustrated in Figure 5.3 and Tables 5.5 to 5.9, were analysed to determine the haplogroups of selected individuals (Wallace, 2004). RFLP analysis was utilised for the investigation of one of the SNPs, as discussed in paragraph 5.4.2.1, while an automated sequencing strategy was followed for haplogrouping of the remaining four SNPs, discussed in paragraph 5.4.2.2.

Figure 5.3: Schematic representation of the differentiation between specific haplogroups



a = The letters in bold indicate the specific haplogroup(s), b = The numbers / letter combinations indicate the specific nucleotide and nucleotide number that should be present to characterise a particular haplogroup.

5.4.2.1 RFLP analysis for single nucleotide polymorphism 3594

RFLP analysis was utilised for the detection of the SNP at position 3594. A total of 10 µl of the 639 bp amplified product were utilised in each reaction. Reactions were incubated at 37°C for at least six hours, in the presence of 10 U *Ksp* AI [Fermentas], 10 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$ and 0.1 $mg.ml^{-1}$ BSA in a final volume of 30 µl.

Following incubation, 5 µl of the RFLP reaction were mixed with 2.5 µl of a 2 x loading buffer (0.04 % Orange G [Sigma®] and 50 % glycerol) prior to loading on a 2% agarose gel along with a molecular weight marker, 100 bp DNA ladder [Promega], to determine the fragment sizes. The agarose gel was electrophoresed in 1 x TBE buffer at 100 V (5 $V.cm^{-1}$) for 90 minutes.

The presence of a T nucleotide at position 3594, signifying haplogroups L0, L1 and L2, was demonstrated by the digestion of the 639 bp amplified product into a 543 bp and a 96 bp product. Haplogroups L3, M and N, represented by a C nucleotide at position 3594, as presented in Table 5.5, resulted in the 639 bp amplified product remaining undigested.

Table 5.5: mtDNA sequence from nucleotide 3001 to 3720, encompassing SNP3594

Nucleotide number[a]	mtDNA sequence[b]
3001	ggacatcccg atggtgcagc cgctattaaa ggttcgtttg ttcaacgatt aaagtcctac
3061	gtgatctgag ttcagaccgg agtaatccag gtcggtttct atcta-cttc aaattcctcc
3121	ctgtacgaaa ggacaagaga aataaggcct acttcacaaa gcgccttccc ccgtaaatga
3181	tatcatctca acttagtatt atacccacac ccacccaaga acagggtttg ttaagatggc
3241	agagcccggt aatcgcataa aacttaaaac tttacagtca gaggttcaat tcctcttctt
3301	aacaacatac ccatggccaa cctcctactc ctcattgtac ccattctaat cgcaatggca
3361	ttcctaatgc ttaccgaacg aaaaattcta ggctatatac aactacgcaa aggccccaac
3421	gttgtaggcc cctacgggct actacaaccc ttcgctgacg ccataaaact cttcaccaaa
3481	gagcccctaa aacccgccac atctaccatc accctctaca tcaccgcccc gaccttagct
3541	ctcaccatcg ctcttctact atgaaccccc ctccccatac ccaaccccct ggt[c]aacctc
3601	aacctaggcc tcctatttat tctagccacc tctagcctag ccgtttactc aatcctctga
3661	tcagggtgag catcaaactc aaactacgcc ctgatcggcg cactgcgagc agtagcccaa

a = Nucleotide numbers are as published on MITOMAP, 2003. b = The sequence was deposited in Genbank with accession number NC_001807. This sequence is a revised version of the Cambridge Reference Sequence, and nucleotide 3106 is maintained as a gap for nucleotide numbering purposes. The positions of the primer set are indicated by underlined text (xxx). The nucleotide of interest is indicated by a box.

5.4.2.2 Automated sequence analysis of SNPs 7055, 10400, 10810 and 11914

The SNPs located at positions 7055, 10400, 10810 and 11914 of the mtDNA was analysed via automated sequencing. The mtDNA sequence encompassing these SNPs as well as the positions of the primers utilised to amplify the respective SNPs are discussed in paragraphs 5.4.2.2.3 to 5.4.2.2.6.

5.4.2.2.1 PCR product purification for cycle sequencing

The QIAquick®[1] PCR purification kit was utilised for the purification of the amplified products from excess primers, nucleotides, polymerases, salts and mineral oil. All the centrifugation steps were performed for 1 minute at 17,949 x g in an Eppendorf® 5810R centrifuge utilising a fixed angle rotor F-45-30-11. The contents of most of the buffers included in the kit were not supplied in the manual.

One volume of each of the amplified products was mixed with 5 volumes of buffer PB. The DNA was bound to a QIAquick® spin column via centrifugation after which the flow-through was discarded. A total of 750 µl of Buffer PE, containing 80% ethanol, were utilised to remove any contaminants from the DNA. To ensure complete removal of the ethanol the spin column was centrifuged for an additional minute after the removal of the flow-through

[1] QIAquick® is a registered trademark of QIAGEN, Clifton Hill, Victoria, Australia.

of the first centrifugation step. The DNA was eluted from the spin column via the addition of 50 µl of buffer EB (10 mM Tris-HCl, pH 8.5). The efficiency of the purification procedure was determined via spectrophotometry.

5.4.2.2.2 Cycle sequencing

The ABIPrism®[1] BigDye™[2] Terminator version 3.1 (V3.1) Ready Reaction Cycle Sequencing Kit was utilised for the sequencing of the four regions. The kit included the Amplitaq®[3] DNA Polymerase FS enzyme and the 5 X sequencing buffer (Tris-HCl, pH 9.0 and $MgCl_2$).

Quarter sequencing reactions were performed. The cycle sequencing was performed in 10 µl reaction volumes containing 2 µl of the BigDye™ 5 X sequencing buffer, 2 µl of the BigDye™ Ready Reaction mix, 3.2 pmol of the primer and 20 ng of the purified PCR product. The forward primer of the specific primer set, as listed in Table 5.4, were utilised in the sequencing reaction.

The cycle sequencing program consisted of 25 cycles of: denaturation at 96°C for ten seconds, annealing at 50°C for 5 seconds, elongation at 60°C for 4 minutes and an indefinite holding step at 4°C. The sequencing reactions were performed in a Thermo Hybaid® MBS® thermocycler.

Subsequent to the cycle sequencing the sequenced product was precipitated to remove any unincorporated BigDye™ terminators, which could affect the eventual result. A sodium acetate / ethanol precipitation strategy was utilised for the precipitation. The precipitation was performed via the addition of 62.5 µl of 99.8% ethanol, 3 µl of 3 M sodium acetate solution (pH 4.6) and 14.5 µl ddH_20, to the 10 µl sequencing reaction. The sequenced product was precipitated via centrifugation at 13,000 x g for 20 minutes in an Eppendorf® 5810R centrifuge utilising a fixed angle rotor F-45-30-11. The supernatant was discarded immediately, after which 250 µl of 70% ethanol were added and the sample centrifuged for 10 minutes at 13,000 x g. The precipitated sequenced product was cleared of the supernatant and subsequently air dried for 30 minutes.

---

[1] ABI Prism® is a registered trademark of Applera Corporation, Foster City, CA, U.S.A.
[2] BigDye™ is a trademark of Applera Corporation, Foster City, CA, U.S.A.
[3] Amplitaq® is a registered trademark of Roche Molecular Systems Inc., Pleasanton, CA, U.S.A.

Electrophoresis of the sequenced products was not performed by the author, but were analysed by Inqaba Biotechnical Industries (Pty) Ltd on contract. The pelleted sequenced product was resuspended in 6 µl Hi-Di™[1] deionised formamide and 3 µl was injected into a SpectruMedix™[2] (SCE2410) Genetic Analysis System sequencer. Sequence analysis was performed utilising the BioEdit Sequence Alignment Editor version 5.0.9 programme (BioEdit, 2003).

5.4.2.2.3 Single nucleotide polymorphism 7055

Cycle sequencing, as described in paragraph 5.4.2.2.2 was utilised for analysis of the SNP at position 7055. This SNP is utilised to distinguish between haplogroups L0 and L1. Haplogroup L1 is characterised by the presence of a G nucleotide at this position, while an A nucleotide at this position signifies the L0 haplogroup. The positions of the primers utilised to amplify the region encompassing SNP7055 are indicated in Table 5.6. These primer sequences are listed in Table 5.3.

Table 5.6: mtDNA sequence from nucleotide 6841 to 7440, encompassing SNP7055

Nucleotide number[a]	mtDNA sequence[b]
6841	ctatccccac cggcgtcaaa gtatttagct gactcgccac actccacgga agcaatatga
6901	aatgatctgc tgcagtgctc tgagccctag gattcatctt tcttttcacc gtaggtggcc
6961	tgactggcat tgtattagca aactcatcac tagacatcgt actacacgac acgtactacg
7021	ttgtagccca cttccactat gtcctatcaa tagg[a]gctgt atttgccatc ataggaggct
7081	tcattcactg atttccccta ttctcaggct acaccctaga ccaaacctac gccaaaatcc
7141	atttcactat catattcatc ggcgtaaatc taactttctt cccacaacac tttctcggcc
7201	tatccggaat gccccgacgt tactcggact accccgatgc atacaccaca tgaaacatcc
7261	tatcatctgt aggctcattc atttctctaa cagcagtaat attaataatt ttcatgattt
7321	gagaagcctt cgcttcgaag cgaaaagtcc taatagtaga agaaccctcc ataaacctgg
7381	agtgactata tggatgcccc ccaccctacc acacattcga agaacccgta tacataaaat

a = Nucleotide numbers are as published on MITOMAP, 2003. b = The sequence was deposited in Genbank with accession number NC_001807. The positions of the primer set are indicated by underlined text (xxx). The nucleotide of interest is indicated by a box.

5.4.2.2.4 Single nucleotide polymorphism 10400

A 617 bp product was generated via PCR utilising the primers listed in Table 5.4. The positions of the primers are indicated in Table 5.7. The SNP at position 10400 was utilised to differentiate between haplogroups M and L3. A T nucleotide at this position is characteristic of haplogroup M, while a C nucleotide would be indicative of haplogroup L3.

[1] Hi-Di™ is a trademark of Applera Corporation, Foster City, CA, U.S.A.
[2] SpectruMedix™ is a trademark of SpectruMedix LLC., State College, PA, U.S.A.

Table 5.7: mtDNA sequence from nucleotide 9841 to 10500, encompassing SNP10400

Nucleotide number[a]	mtDNA sequence[b]
9841	caactttcct cactatctgc ttcatccgcc aactaatatt tcactttaca tccaaacatc
9901	actttggctt cgaagccgcc gcctgatact ggcattttgt agatgtggtt tgactatttc
9961	tgtatgtctc catctattga tgagggtctt actcttttag tataaatagt accgttaact
10021	tccaattaac tagttttgac aacattcaaa aaagagtaat aaacttcgcc ttaattttaa
10081	taatcaacac cctcctagcc ttactactaa taattattac attttgacta ccacaactca
10141	acggctacat agaaaaatcc accccttacg agtgcggctt cgaccctata tcccccgccc
10201	gcgtcccttt ctccataaaa ttcttcttag tagctattac cttcttatta tttgatctag
10261	aaattgccct ccttttaccc ctaccatgag ccctacaaac aactaacctg ccactaatag
10321	ttatgtcatc cctcttatta atcatcatcc tagccctaag tctggcctat gagtgactac
10381	aaaaaggatt agactgaacc gaattggtat atagtttaaa caaaacgaat gatttcgact
10441	cattaaatta tgataatcat atttaccaaa tgcccctcat ttacataaat attatactag

a = Nucleotide numbers are as published on MITOMAP, 2003. b = The sequence was deposited in Genbank with accession number NC_001807. The positions of the primer set are indicated by underlined text (xxx). The nucleotide of interest is indicated by a box.

5.4.2.2.5 Single nucleotide polymorphism 10810

Haplogroup L2 was distinguished from haplogroups L0 and L1 via the presence of a T nucleotide, instead of a C nucleotide, at position 10810. A 612 bp product was amplified, as presented in Table 5.8 harbouring the SNP at position 10810. The primer sequences are listed in Table 5.4 while the positions of the respective primers are indicated in Table 5.8

Table 5.8: mtDNA sequence from nucleotide 10381 to 11040, encompassing SNP10810

Nucleotide number[a]	mtDNA sequence[b]
10381	aaaaaggatt agactgaacc gaattggtat atagtttaaa caaaacgaat gatttcgact
10441	cattaaatta tgataatcat atttaccaaa tgcccctcat ttacataaat attatactag
10501	catttaccat ctcacttcta ggaatactag tatatcgctc acacctcata tcctccctac
10561	tatgcctaga aggaataata ctatcgctgt tcattatagc tactctcata accctcaaca
10621	cccactccct cttagccaat attgtgccta ttgccatact agtctttgcc gcctgcgaag
10681	cagcggtggg cctagcccta ctagtctcaa tctccaacac atatggccta gactacgtac
10741	ataacctaaa cctactccaa tgctaaaact aatcgtccca acaattatat tactaccact
10801	gacatgactc tccaaaaaac acataatttg aatcaacaca accacccaca gcctaattat
10861	tagcatcatc cctctactat tttttaacca aatcaacaac aacctattta gctgttcccc
10921	aacctttcc tccgaccccc taacaacccc cctcctaata ctaactacct gactcctacc
10981	cctcacaatc atggcaagcc aacgccactt atccagtgaa ccactatcac gaaaaaaact

a = Nucleotide numbers are as published on MITOMAP, 2003. b = The sequence was deposited in Genbank with accession number NC_001807. The positions of the primer set are indicated by underlined text (xxx). The nucleotide of interest is indicated by a box.

5.4.2.2.6 Single nucleotide polymorphism 11914

The SNP at position 11914 was utilised to confirm the inclusion into haplogroup L0. The presence of an A nucleotide would signify the confirmation for haplogroup L0. The 629 bp product was amplified utilising the primers listed in Table 5.4 The positions of the primers are indicated in Table 5.9

Table 5.9: mtDNA sequence from nucleotide 11461 to 12120, encompassing SNP11914

Nucleotide number[a]	mtDNA sequence[b]
11461	actcttaaaa ctaggcggct atggtataat acgcctcaca ctcattctca acccccctgac
11521	aaaacacata gcctacccct tccttgtact atccctatga ggcataatta taacaagctc
11581	catctgccta cgacaaacag acctaaaatc gctcattgca tactcttcaa tcagccacat
11641	agccctcgta gtaacagcca ttctcatcca aacccccctga agcttcaccg gcgcagtcat
11701	tctcataatc gcccacgggc ttacatcctc attactattc tgcctagcaa actcaaacta
11761	cgaacgcact cacagtcgca tcataatcct ctctcaagga cttcaaactc tactcccact
11821	aatagctttt tgatgacttc tagcaagcct cgctaacctc gccttacccc ccactattaa
11881	cctactggga gaactctctg tgctagtaac cac[g]ttctcc tgatcaaata tcactctcct
11941	acttacagga ctcaacatac tagtcacagc cctatactcc ctctacatat ttaccacaac
12001	acaatggggc tcactcaccc accacattaa caacataaaa ccctcattca cacgagaaaa
12061	caccctcatg ttcatacacc tatcccccat tctcctccta tccctcaacc ccgacatcat

a = Nucleotide numbers are as published on MITOMAP, 2003. b = The sequence was deposited in Genbank with accession number NC_001807. The positions of the primer set are indicated by underlined text (xxx). The nucleotide of interest is indicated by a box.

CHAPTER SIX

RESULTS AND DISCUSSION

This chapter contains the results of the dosage test, PFGE and haplogroup analyses performed in this study, accompanied by a discussion of the presented results. These analyses were performed to determine whether plasticity at the D4Z4 locus was present in the specific population groups investigated. The obtained results were analysed to determine whether there are molecular differences at the D4Z4 locus between the two populations under investigation, as well as those reported in the literature. It was envisaged that analysis of the plasticity of the D4Z4 locus over time would reveal characteristics that could explain how the FSHD associated deletion occurs and what the role of the translocations between chromosomes 4q35 and 10q26 is in the pathogenesis of FSHD.

6.1 DOSAGE TEST ANALYSIS

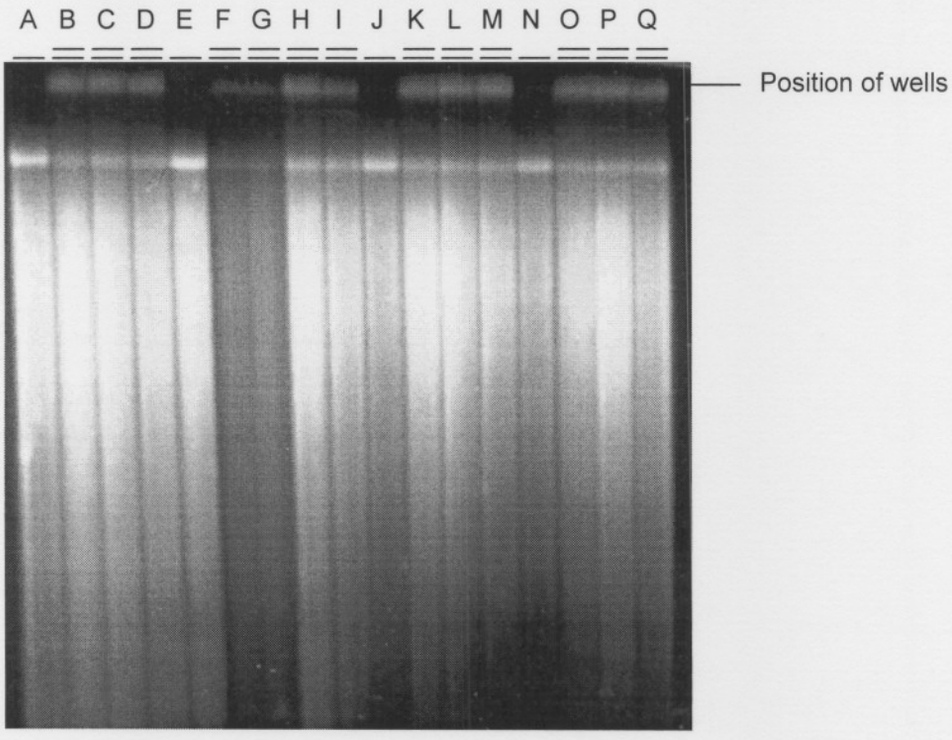
The dosage test was utilised for the initial screening for plasticity at the D4Z4 locus in the populations investigated. A total of 300 individuals, consisting of 273 individuals from the Black South African population and 27 individuals of the Khoi-San population were included in the investigation. Since this is one of the most unique populations on a genetic level only the limited number of Khoi-San individuals that were available were included in this study. The individuals from the Black South African population were not selected according to ethnicity, but selected across different ethnic groups.

6.1.1 RESTRICTION FRAGMENT LENGTH POLYMORPHISM ANALYSIS

A representative agarose gel of the RFLP reactions for dosage test analyses for selected individuals is presented in Figure 6.1. As described in paragraph 5.3.1, 2.5 to 5 μg of gDNA was utilised in each RFLP reaction. This difference in concentration is depicted in Figure 6.1. Lanes F and G contain 2.5 μg digested gDNA, while the remaining lanes contain 5 μg digested gDNA. No undigested gDNA was visible in the wells of lanes A, E, J and N, but undigested gDNA is visible in the remaining wells (wells indicated by double

lines in Figure 6.1). The amount of undigested gDNA was, however, minimal in comparison to the digested gDNA. This did not interfere with the analyses since the accuracy of the dosage test is based on the quantitative analysis of chromosome 4 and 10 fragments within a sample, and not between samples. The quantity and quality difference between gDNA samples was therefore prevented. The accuracy of the dosage test is further dependant on the digestion of both enzymes, *Bln* I and *Bgl* II. If either of the enzymes did not digest the gDNA, the fragment sizes of interest would not be obtained and would not be visible on the X-ray film. The correct size fragments for all the samples displaying incomplete digestion, (wells indicated by double lines), in Figure 6.1 were observed upon autoradiography.

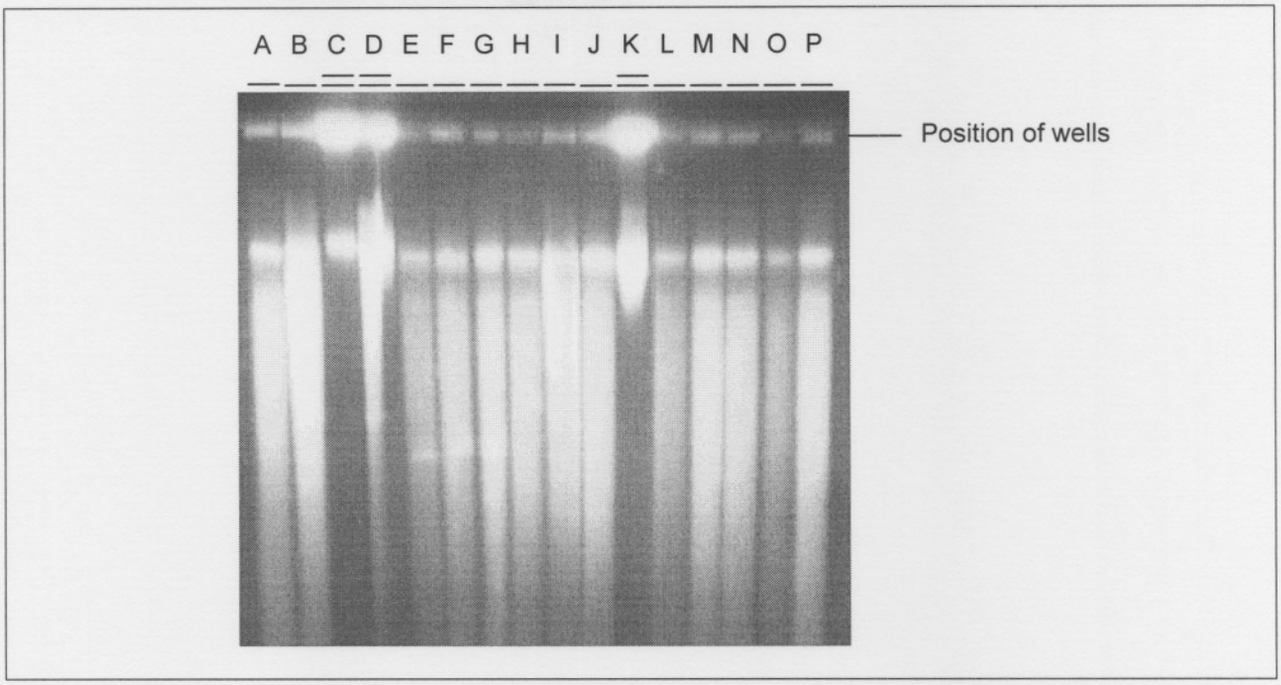
Figure 6.1: Photographic representation of the dosage test RFLP analysis gel



Fragments were separated on a 0.8% agarose gel, which was electrophoresed at 1.6 $V.cm^{-1}$ for 24 hours in 1 x TBE buffer.

As described above, although selected samples were incompletely digested, the subsequent analyses were not influenced. However, for a few samples, as depicted in Figure 6.2 lanes C, D and K (wells indicated by double lines), a significant level of incomplete digestion was observed. The level of digestion was not sufficient and these samples were thus not included for analysis. The remaining samples included in Figure 6.2 also displayed incomplete digestion, but the level was sufficient for analysis, as described for Figure 6.1.

Figure 6.2: Photographic representation of the dosage test RFLP analysis gel indicating incomplete digestion

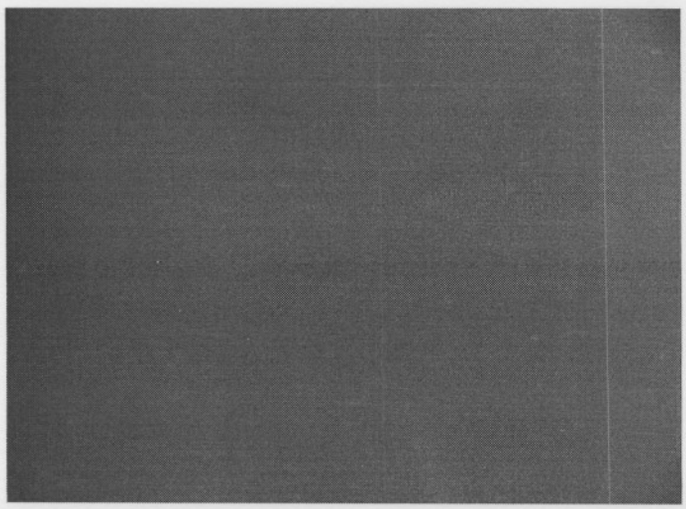


Fragments were separated on a 0.8% agarose gel, which was electrophoresed at 1.6 $V.cm^{-1}$ for 24 hours in 1 x TBE buffer. All wells indicated by double lines (C, D and K) were not utilised for analyses.

6.1.2 GENOMIC DNA TRANSFER

The efficiency of the gDNA transfer to the nylon membrane was determined via the staining of the gel with 2 $\mu g.ml^{-1}$ EtBr for 30 minutes with gentle agitation, subsequent to overnight transfer. A representative agarose gel stained with EtBr is depicted in Figure 6.3. It was evident from the gel presented below that the transfer was successful, as no gDNA was visible on the gel.

Figure 6.3: Agarose gel after Southern blot transfer

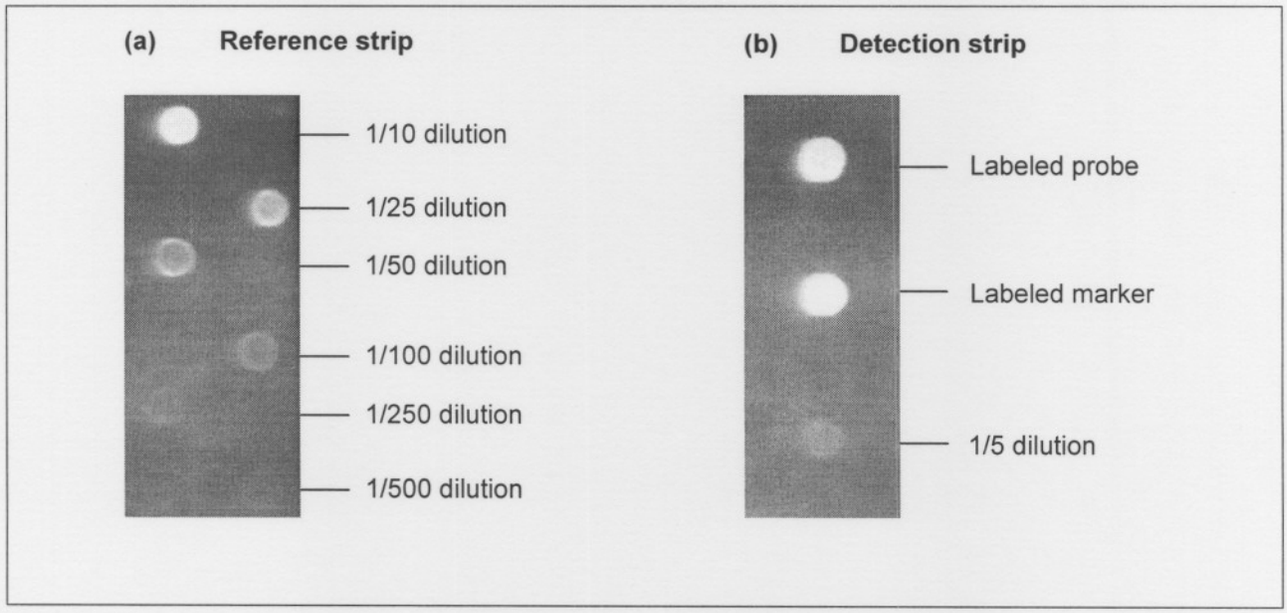


0.8 % agarose gel after overnight transfer. Stained with 2 $\mu g.ml^{-1}$ EtBr for 30 minutes with gentle agitation.

6.1.3 LABELLING OF PROBE P13E-11 AND MOLECULAR WEIGHT MARKER

The p13E-11 probe as well as the molecular weight marker utilised were labelled as described in paragraph 5.2.5. Before the utilisation of the labelled probe and molecular weight marker in hybridisation reactions, the efficiency of the labelling reactions was determined as described in paragraph 5.2.5.1. The reference and detection strips utilised for the assessment of the labelling efficiency are presented in Figure 6.4 (a) and Figure 6.4 (b) respectively. The p13E-11 probe as well as the molecular weight marker utilised in all the hybridisation reactions were successfully labelled, since the signal intensity of both was comparable to the 1/10 dilution on the reference strip. Furthermore the signal intensity of the labelled probe and molecular weight marker was due to incorporated fluorescein only, since the 1/5 dilution was very faint.

Figure 6.4: Assessment of labelling efficiency



6.1.4 NON RADIO-ACTIVE DETECTION

Worldwide Southern blot analysis utilising a radio-active labelled p13E-11 probe is utilised as the standard diagnostic protocol for deletion fragment (conventional and PFGE) as well as the dosage test. However, during this study a non-radioactive approach was developed. The digoxigenin (DIG) kit [Roche] was initially utilised for the detection of Southern blots, since the kit could be used for single gene detection protocols. After several months of investigating various optimisation strategies it was, however, concluded that the kit was not sensitive enough for the detection of the fragments of interest in FSHD.

Subsequently, the *Gene Images* CDP-*Star*® random prime labelling and detection kit [Amersham Biosciences] was utilised. The kit was tested by the manufacturer to be compatible with sensitivity levels of ^{32}P detection strategies and should be able to detect a single copy gene in human gDNA representing 0.06 pg target in 0.125 μg DNA. Probe p13E-11 and the appropriate molecular weight marker were labelled and the Southern blot detected as described in paragraphs 5.2.5 and 5.2.8 respectively.

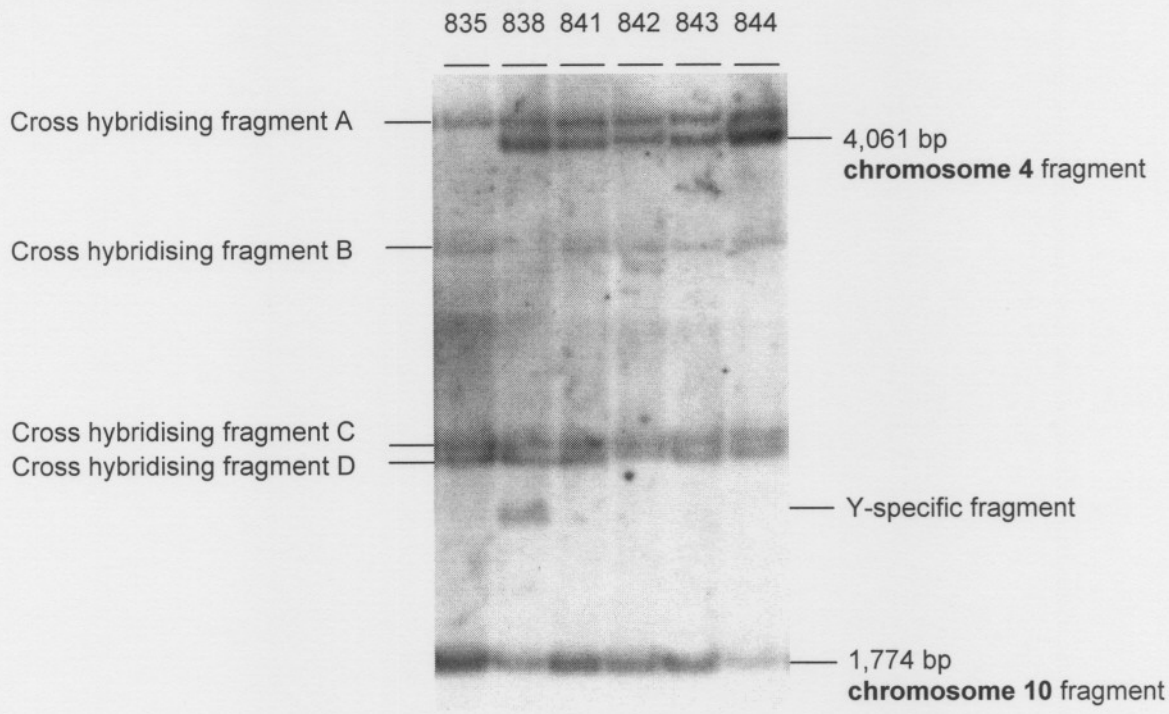
In Figure 6.5 (a) a representative autoradiograph of dosage test analyses is presented. Two fragments with a molecular weight of approximately 4 kb, corresponding to the 4,061 bp chromosome 4 fragment, were observed. It was, however, determined that the fragment with the lowest molecular weight of the two fragments corresponded to the 4,061 bp chromosome 4 fragment (Lemmers, 2004). Via the utilisation of more stringent wash conditions, increasing the stringency wash temperature from 60°C to 65°C, it was possible to remove the signal generated by cross hybridising fragment A, indicated in Figure 6.5 (a). However, in this study the cross hybridising fragment was utilised as an internal control and the less stringent wash conditions were therefore utilised.

Several other cross hybridising fragments were also observed (fragments B, C and D), but none of them overlapped in size with the fragments of interest and did therefore not interfere with the analyses. It is evident that five female individuals (835, 841, 842, 843 and 844) and one male individual (838) were included in Figure 6.5, since only individual 838 displayed the Y-specific fragment. The fragment with the lowest molecular weight on the X-ray film corresponded to the 1,774 bp fragment originating from chromosome 10.

The table presented in Figure 6.5 (b) indicates the ratio of chromosome 4 to chromosome 10 fragments of the individuals included on the autoradiograph presented in Figure 6.5 (a) after visual inspection. This table is discussed in detail in section 6.1.5.

Figure 6.5: Representative autoradiograph of dosage test analysis

(a)



(b)

835	838	841	842	843	844	Individual number
0	?	2	2	2	3	**Chromosome 4 fragments**
4	?	2	2	2	1	**Chromosome 10 fragments**
0:4	?	2:2	2:2	2:2	3:1	**Ratio 4:10**

Fourteen hour exposure.

6.1.5 ASSESSMENT OF FRAGMENT INTENSITIES

The intensities of the fragments of interest were assessed on three levels. At first the intensities were estimated via visual inspection, as described in paragraph 6.1.5.1. Subsequent to the visual estimation, the intensities were quantified via the Quantity One® v 4.4.1 software program as discussed in paragraph 6.1.5.2. Thirdly, density trace analyses were performed as presented in paragraph 6.1.5.4.

6.1.5.1 Visual estimation of intensities

Even via visual inspection of the autoradiograph presented in Figure 6.5 (a) it is clear that the intensities of the 4,061 bp chromosome 4 specific fragment and the 1,774 bp chromosome 10 specific fragment for each of the individuals were not equal. Individual 835 does not harbour a chromosome 4 specific fragment but displays a chromosome 10

specific fragment with a high intensity. The intensities of the two fragments of interest of individuals 841, 842 and 843 appears to be equal, while the intensity of the 4,061 bp fragment of individual 844 is much more intense than the 1,774 bp fragment. The fragments of individual 838 do also appear to differ in intensities.

The table depicted in Figure 6.5 (b) represents the ratio of chromosome 4 to chromosome 10 fragments (ratio 4:10) after visual inspection. As mentioned in paragraph 6.1.5, cross hybridising fragment A was utilised as an internal control. This cross hybridising fragment was visible for individual 835, although no chromosome 4 fragment was detected. The presence of cross hybridising fragment A thus acts as an internal control, indicating the successful RFLP, gDNA transfer and detection of fragments within this size range. The absence of a chromosome 4 fragment for individual 835 was therefore not due to any experimental error, but due to the absence of this specific fragment within this individual. Individual 835 therefore harbours no chromosome 4-type fragment. In this case the 4:10 ratio was thus scored as 0:4 for individual 835. Individual 838 harbours chromosome 4 and 10 fragments. Upon visual inspection these fragments did not appear to be of equal intensities, as the chromosome 4 fragment seemed to be more intense. By comparison to individual 844 it was, however, evident that the difference in intensity between the two fragments of interest for individual 838 was not as significant as that of individual 844. The chromosome 4 fragment of individual 838 appears less intense than that of individual 844. In contrast, the chromosome 10 fragment of individual 838 appears more intense when compared to that of individual 844. The ratio 4:10 for individual 838 could therefore not be estimated with confidence via visual inspection, while that of individual 844 was scored to be 3:1. The ratio 4:10 of individuals 841, 842 and 843 was scored as 2:2, since the fragments appeared to be of equal intensities.

6.1.5.2 Fragment intensity quantification via the Quantity One® v 4.4.1 software program

It was, however, evident that the ratio of chromosome 4 to 10 fragments could not be determined accurately via visual inspection. This was highlighted by the fact that the ratio of 4:10 for individual 838 could not be deduced from Figure 6.5, as described in paragraph 6.1.5.1.

The Quantity One® v 4.4.1 software program was subsequently utilised to quantify the intensities of the chromosome 4 and 10 fragments, as described in paragraph 5.3.3. The intensities of the fragments in Figure 6.6 were determine via the positioning of volume

boundaries U1, U3, U5, U7 and U9 over the chromosome 4 fragments with B1, B3, B5, B7 and B9 corresponding to the background of each sample respectively. Volume boundaries U2, U4, U6, U8 and U10 were positioned over chromosome 10 fragments with the corresponding background blocks B2, B4, B6, B8 and B10 respectively.

Figure 6.6: Quantification of fragment intensities via volume boundaries

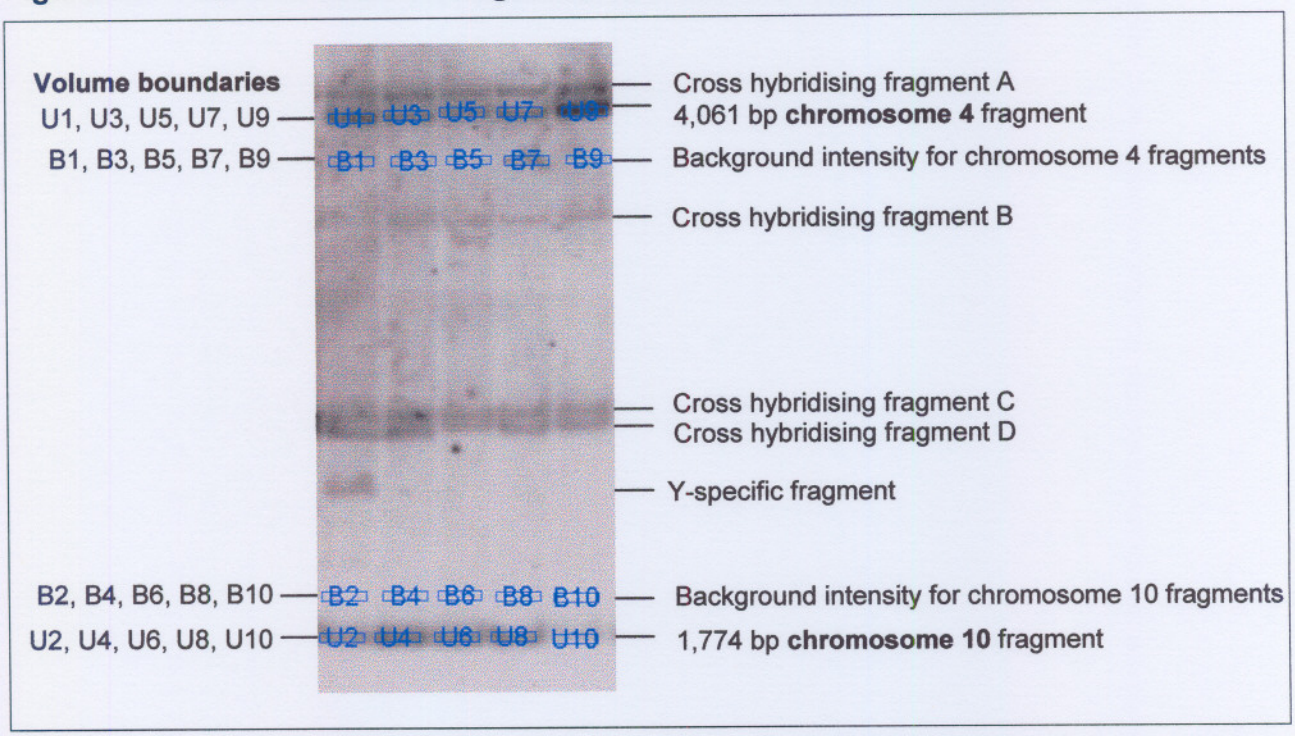


Image presented as exported from Quantity One® version 4.4.1. U1 to U10 represent the volume boundaries for the unknown samples while B1 to B10 represent the volume boundaries for the background for each sample. U1, U3, U5, U7 and U9 were positioned over the chromosome 4 fragments with B1, B3, B5, B7 and B9 corresponding to the background of each sample respectively. U2, U4, U6, U8 and U10 were positioned over chromosome 10 fragments with the corresponding background boundaries B2, B4, B6, B8 and B10 respectively.

6.1.5.3 Calculation of fragment intensities

The intensities quantified via the Quantity One® v 4.4.1 software program, as described in paragraph 6.1.5.2, were further utilised to calculate the intensities of the fragments after correcting for background, as described in paragraph 5.3.4. The contribution of the chromosome 4 specific signal to the sum of the chromosomes 4 and 10 signals were calculated via Equation 5.1 (page 88). The signal intensities of the chromosome 4 and 10 fragments with their respective backgrounds are presented in Table F.1 of Appendix F. The background intensity value was subtracted from the fragment intensity value to obtain a normalised value which was utilised to determine the contribution of the chromosome 4 specific signal via Equation 5.1. The contribution value for each of the samples included in this study is also presented in Table F.1.

6.1.5.4 Representative density traces of nullisomic, monosomic, disomic, trisomic and quatrosomic individuals

Density traces for fragments of selected samples were also determined for comparison to the obtained intensities. In the density traces, the fragments are represented by an intensity profile, as illustrated in Figure 6.7. The intensity profile was generated via calculating the average intensity of the pixels within a defined area. The vertical line, superimposed over the sample lane in Figures 6.7 to 6.11, indicates the points at which the density traces were determined.

For all the density traces presented, the peaks corresponding to the intensities of the 4,061 bp chromosome 4 fragment and the 1,774 bp chromosome 10 fragment are indicated by an F4 and an F10 respectively. The intensity peaks of the cross hybridising fragments A to D, as earlier depicted in Figure 6.5, are signified by the letters A to D respectively, whereas the peaks corresponding to the intensities of the background are indicated by the letter E.

Density traces do not, however, provide a quantitative value for the density of the fragment. Via density trace analysis it was also not possible to correct the fragment density for its background. It was, utilised to graphically illustrate the qualitative values obtained via volume boundary analysis. In addition it provided a graphical representation of the different translocation profiles that were obtained.

A representative density trace of a disomic individual is presented in Figure 6.7. The peak height of peaks F4 and F10 are comparable. The peak height correlated to the intensity of the fragments, indicating that this profile corresponded to a disomic profile for chromosomes 4 and 10 fragments. The cross hybridising fragments A, B, C and D were presented by peaks A, B, C and D respectively. Background peaks, indicated as E, can also be observed, however the intensities of these peaks were much lower than that of the true peaks and did not interfere with analysis in any of the samples analysed.

Figure 6.7: Representative density trace of a disomic individual

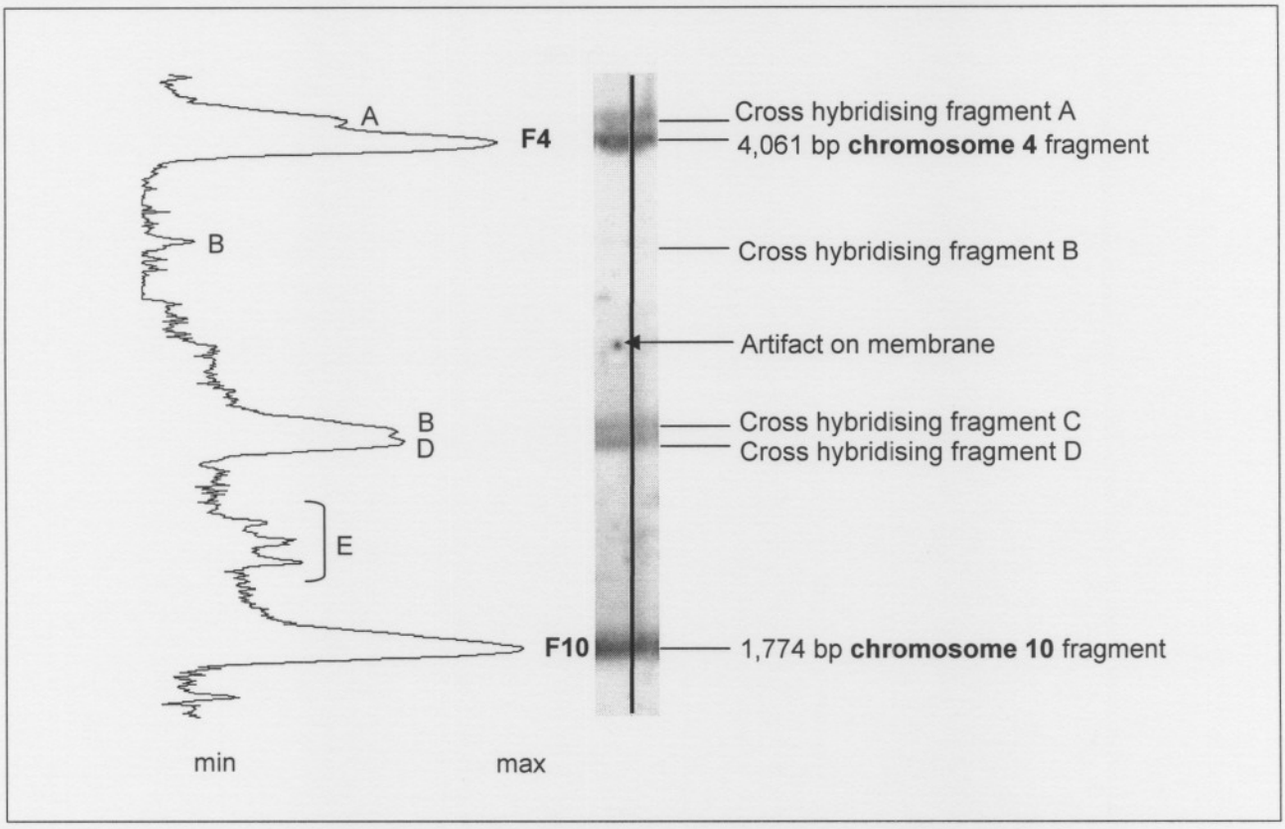


Image presented as exported from Quantity One® version 4.4.1.

Figure 6.8 is a representative density trace of an individual harbouring no chromosome 4 fragment, as illustrated for individual 835 in Figure 6.5. Via comparison to the density trace presented in Figure 6.7, it was evident that there was no peak visible below peak A which characterised cross hybridising fragment A.

Cross hybridising fragments B, C and D is represented by peaks B, C and D respectively. The absence of a peak for the chromosome 4 specific fragment therefore confirmed that this individual was nullisomic for the chromosome 4 fragment. This individual thus harboured four chromosome 10 type fragments.

Figure 6.8: Representative density trace of a nullisomic individual

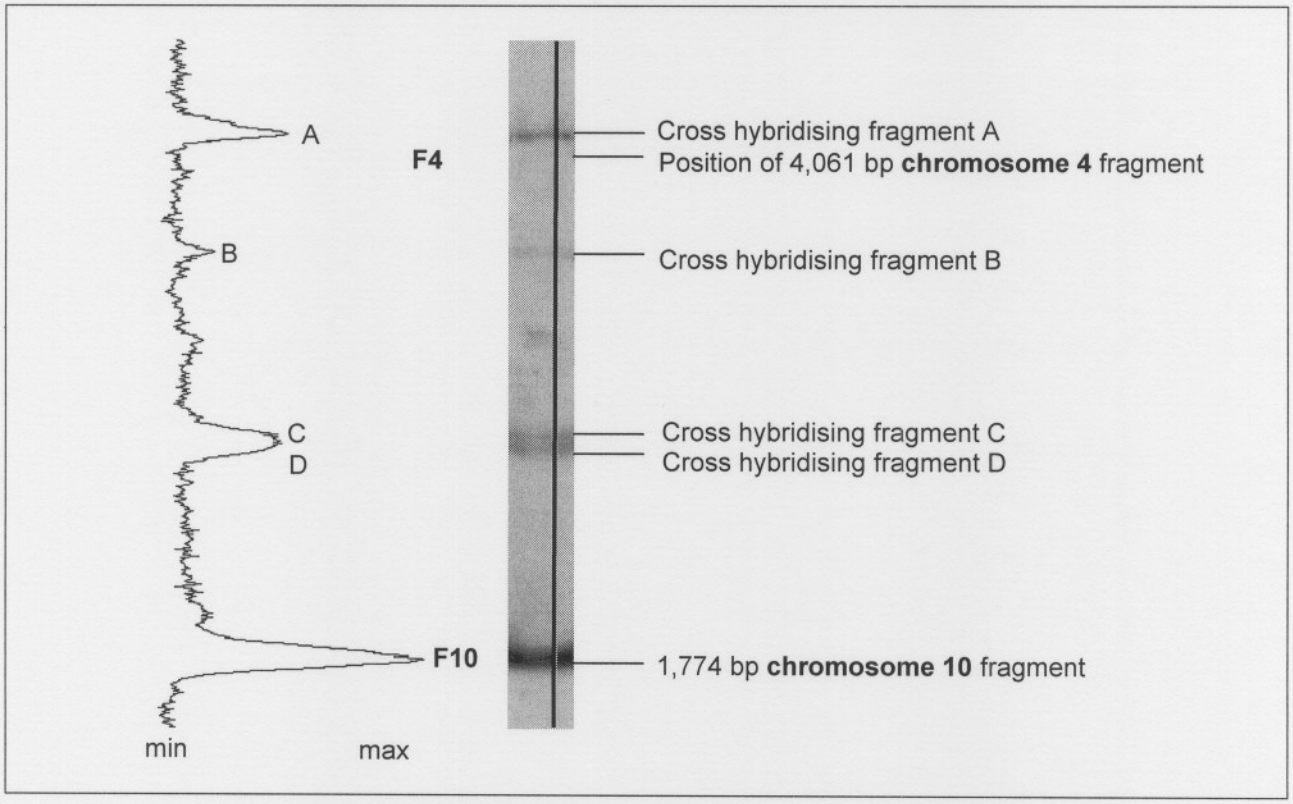


Image presented as exported from Quantity One® version 4.4.1.

Peaks F4 and F10, as illustrated in Figure 6.9, characterised the chromosome 4 and 10 fragments of this monosomic individual. It is noticeable from Figure 6.9 that peak F10 was much higher than peak F4, signifying the greater intensity of this fragment. Although the intensity could not be quantified with the density trace, it was evident that the chromosome 10 fragment was much more intense than the chromosome 4 fragment.

The cross hybridising fragments A, B, C and D was characterised by peaks A, B, C and D respectively. In Figure 6.9 fragment B was seen to be faint due to the surrounding regions' under exposure. However, peak B corresponded to this fragment and was determined via the analysis software. Peak E on the density trace identified the background present in the specific region on the X-ray film.

Figure 6.9: Representative density trace of a monosomic individual

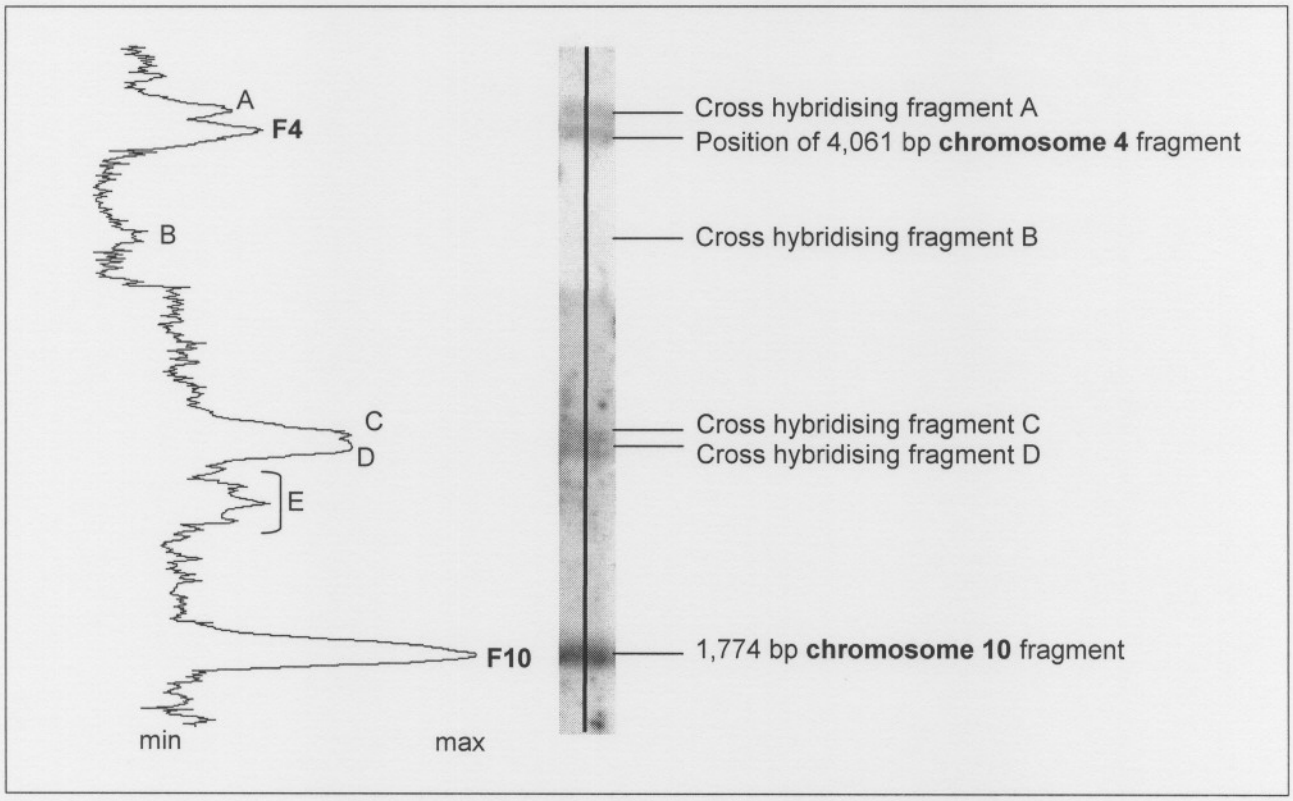


Image presented as exported from Quantity One® version 4.4.1.

A representative density trace of a trisomic individual is presented in Figure 6.10. Cross hybridising fragments A, B, C and D was characterised by peaks A, B, C and D respectively. Peak E corresponded to an artefact on the X-ray film. The intensities of the chromosome 4 and 10 fragments were represented by peaks F4 and F10 respectively. It is evident from the X-ray image in Figure 6.10 that the chromosome 4 fragment was more intense than the chromosome 10 fragment. This greater intensity was indicated by the increased peak height of peak F10.

Figure 6.10: Representative density trace of a trisomic individual

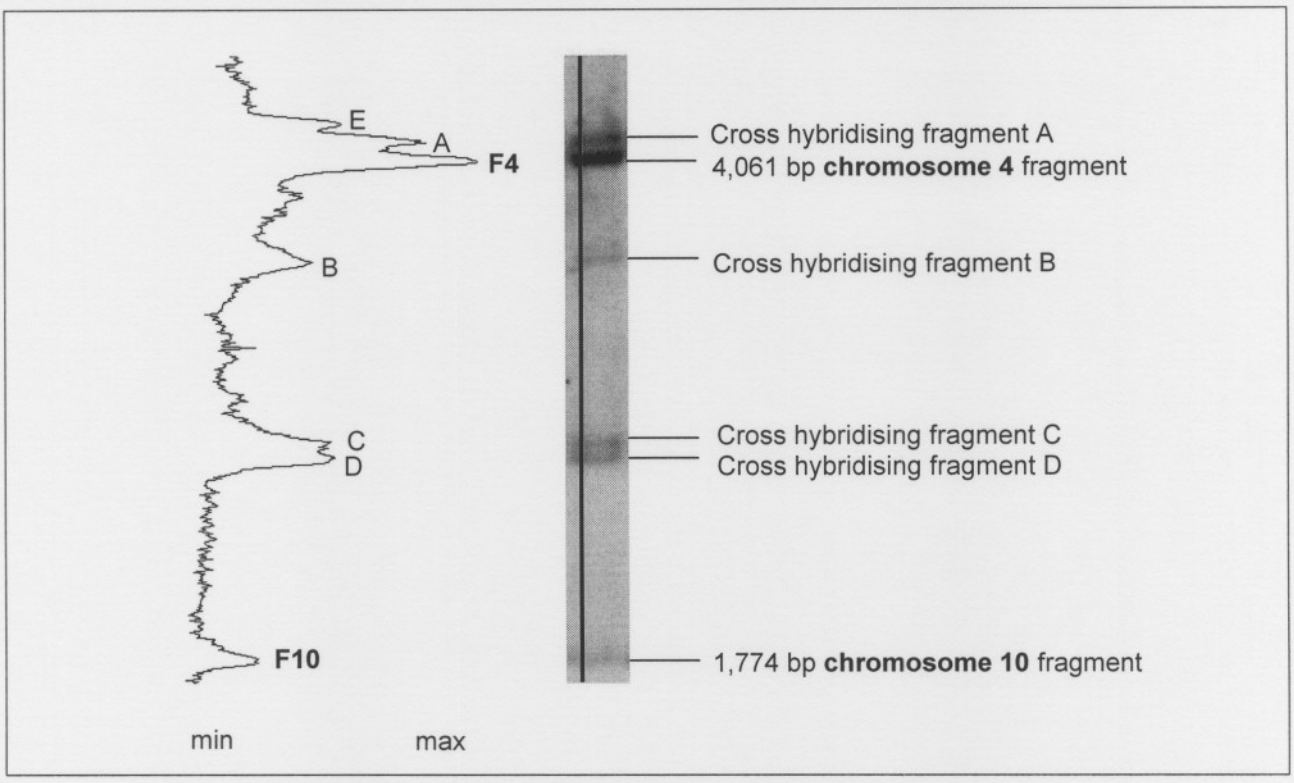


Image presented as exported from Quantity One® version 4.4.1.

In contrast to the density trace of the nullisomic individual illustrated in Figure 6.8, the density trace presented in Figure 6.11 characterises an individual quatrosomic for the chromosome 4 fragment. Peak F4 corresponded to the chromosome 4 fragment intensity, while no peak for the chromosome 10 fragment intensity, indicated at the peak F10 position, was visible. This individual therefore harboured four chromosome 4-type fragments. Peaks A, B, C and D illustrated cross hybridising fragments A, B, C and D. Only one peak was present for cross hybridising fragments C and D, as they did not separate sufficiently and were presented as one fragment on the X-ray film.

Figure 6.11: Representative density trace of a quatrosomic individual

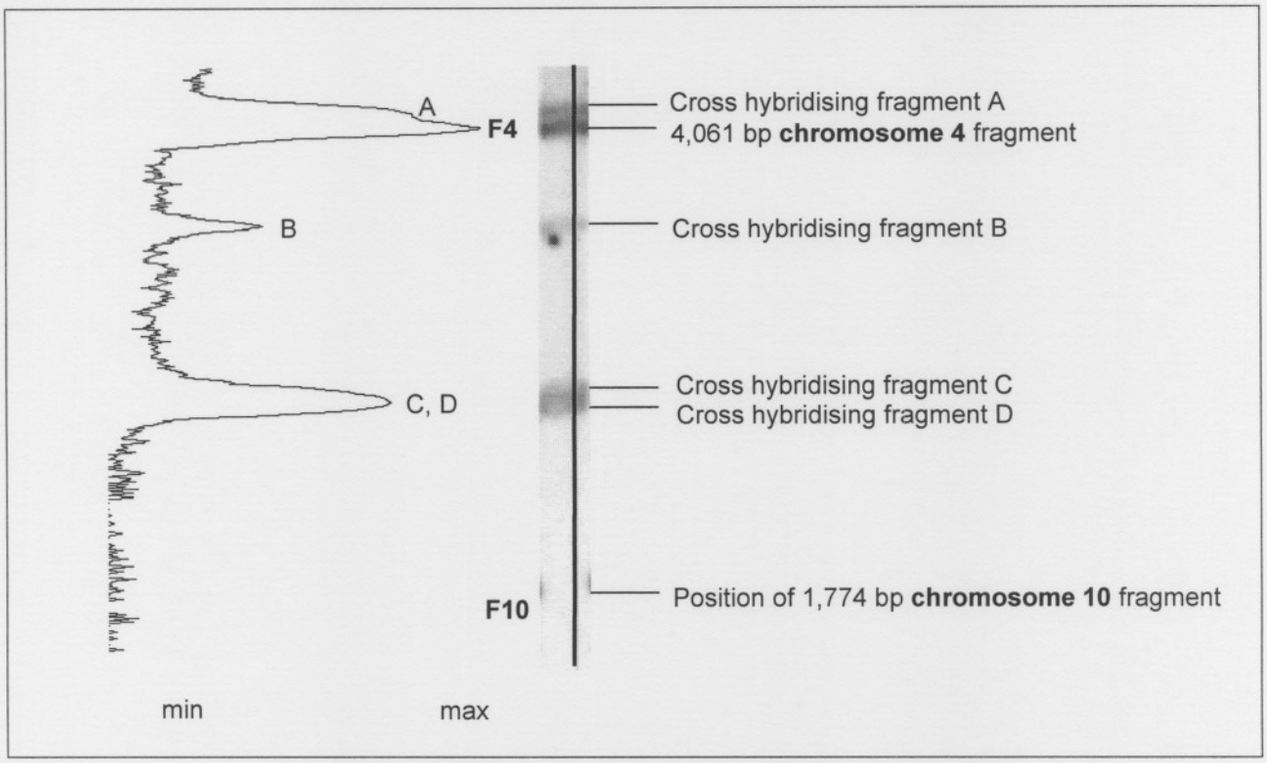


Image presented as exported from Quantity One® version 4.4.1.

6.1.6 CLASSIFICATION OF TRANSLOCATION PROFILES

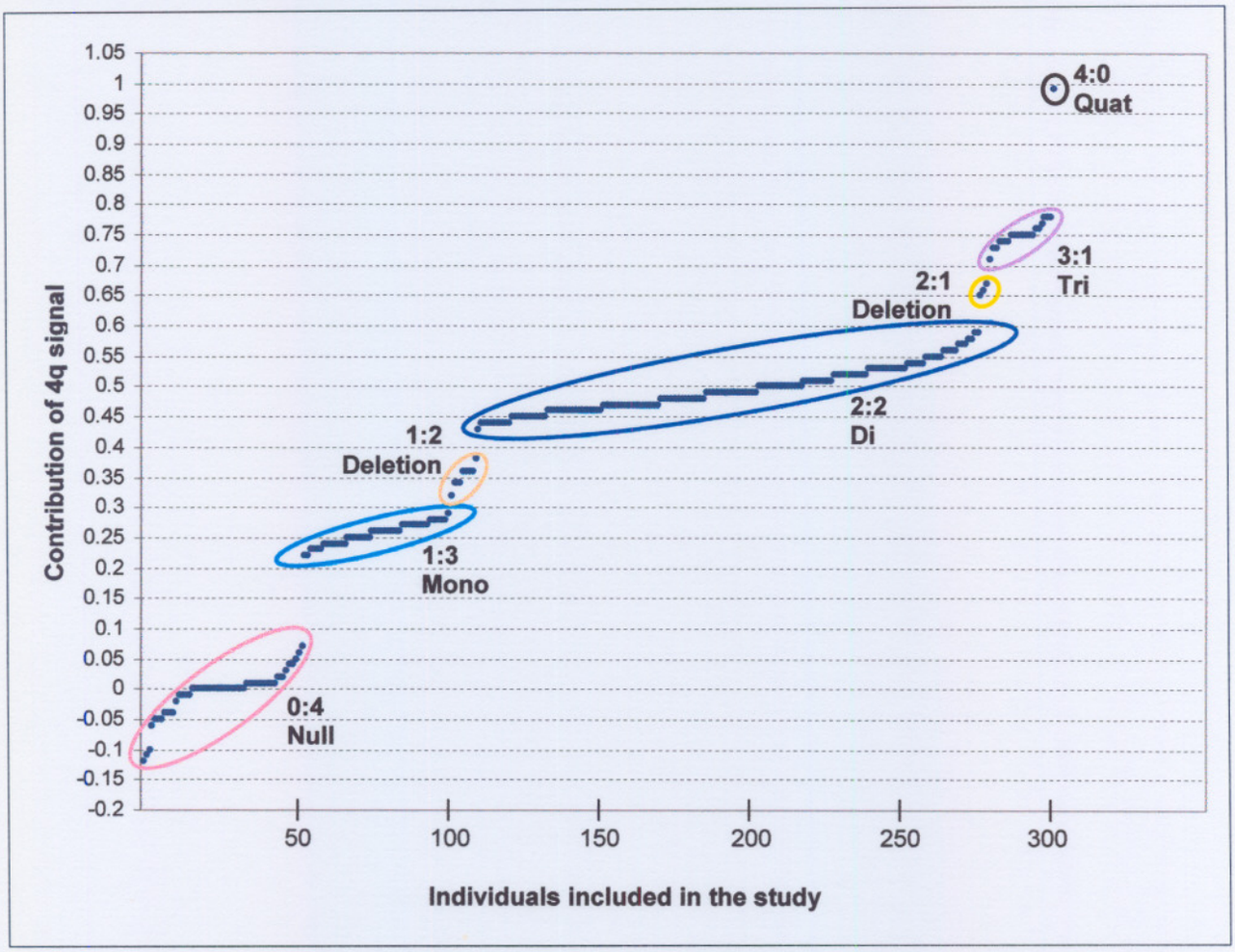
The contribution of the chromosome 4 specific signal values of the 300 individuals included in this study, as presented in Table F.1, Appendix F, were utilised for the classification of the samples into different translocation profiles. Table 5.3 was utilised as a guideline for the differentiation between the different profiles.

6.1.6.1 Translocation profile distribution

The contribution values were plotted as illustrated in Graph 6.1. Seven different translocation profiles, [nullisomy (0:4); monosomy (1:3); deletion of p13E-11 hybridisation site on chromosome 4q35 (1:2); disomy (2:2); deletion of p13E-11 hybridisation site on chromosome 10q26 (2:1); trisomy (3:1) and quatrosomy (4:0)], were identified in the 300 individuals, as indicated by the seven coloured circles drawn around the plotted values. The disomic profile was also classified as a translocation profile in this study for easier classification purposes although no translocation events could be observed in these individuals. The presence of double exchanges, i.e. a 4-type fragment on chromosome 10, and a 10-type fragment on chromosome 4, can not be differentiated from a standard

disomic profile. Samples with a contribution of a 4q signal value of less than zero, as indicated in Graph 6.1, were observed as a result of a higher background value that were due to artefacts.

Graph 6.1: Distribution of translocation profiles



0:4 = null = nullisomy (n = 53); 1:3 = mono = monosomy (n = 48); 1:2 = deletion of p13E-11 hybridisation site on chromosome 4q35 (n = 9); 2:2 = di = disomy (n= 165); 2:1 = deletion of p13E-11 hybridisation site on chromosome 10q26 (n = 3); 3:1 = tri = trisomy (n = 21); 4:0 = quat = quatrosomy (n = 1).

All of the contribution values could be categorised into one of the seven profiles via comparison to the expected values as listed in Table 5.3. However, to determine whether the values could be differentiated without any overlap between the different profiles the confidence intervals for the contribution of the 4q signal of the different profiles were determined.

6.1.6.2 Calculation of the confidence interval distribution

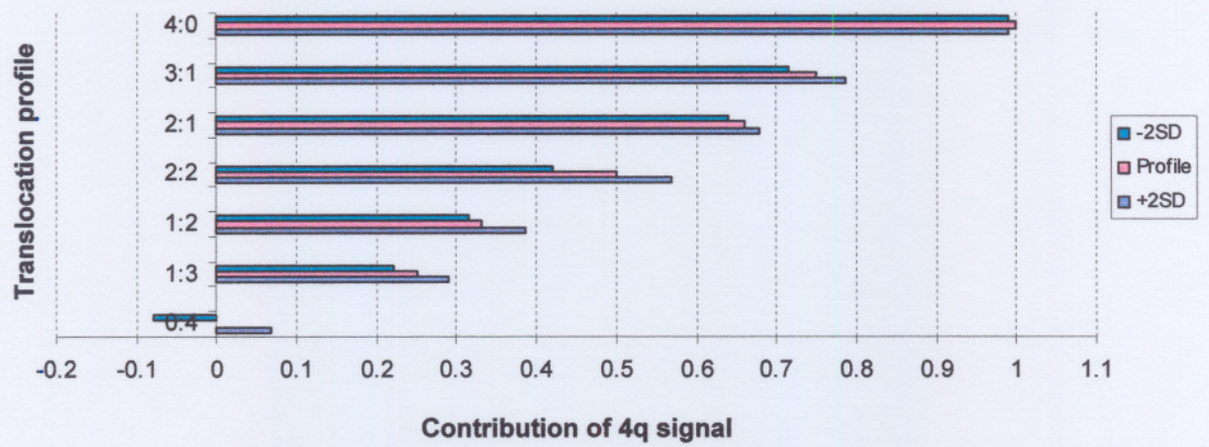
The confidence interval distribution (± 2 standard deviations) for the contribution of the chromosome 4q signal for each of the translocation profiles is presented in Graph 6.2 (a).

The average of each of the different translocation profiles was calculated and utilised to determine the standard deviation for each profile, as depicted in the table in Graph 6.2 (b). The data presented in the table in Graph 6.2 (c) were utilised to compose the confidence interval distribution in Graph 6.2 (a).

The pink bars in Graph 6.2 (a) represent the expected contribution of the chromosome 4q signal value, as depicted in Table 5.3. The purple and green bars represent the ± 2 standard deviations from the mean respectively, as illustrated in the table in Graph 6.2 (c). It is evident from Graph 6.2 (a) that none of the confidence intervals for any of the translocation profiles overlap. The seven different translocation profiles therefore did not overlap and could be differentiated from each other.

Graph 6.2: Confidence interval distribution for the contribution of the chromosome 4q signal for each of the translocation profiles

(a)



(b)

	10-on-4					4-on-10	
Translocation profile	**0:4**	**1:3**	**1:2**	**2:2**	**2:1**	**3:1**	**4:0**
Average	-0.01	0.26	0.35	0.49	0.66	0.75	0.99
Standard deviation	0.04	0.02	0.02	0.04	0.01	0.02	0.00

(c)

	10-on-4					4-on-10	
Translocation profile	**0:4**	**1:3**	**1:2**	**2:2**	**2:1**	**3:1**	**4:0**
+2SD	0.07	0.29	0.39	0.57	0.68	0.79	0.99
Profile[1]	0.00	0.25	0.33	0.50	0.66	0.75	1.00
-2SD	-0.08	0.22	0.32	0.42	0.64	0.72	0.99

1 = The different profiles are as listed in Table 5.3. SD = Standard deviation. The data was utilised in all the analyses without rounding. The data presented in the tables were rounded off to the second decimal.

6.1.7 TRANSLOCATION FREQUENCY

Subsequent to the classification of the 300 individuals included in this study into the different translocation profiles, the translocation frequency for each of the profiles was determined. The translocation frequencies for the two populations included in this study are presented in paragraphs 6.1.7.1 and 6.1.7.2.

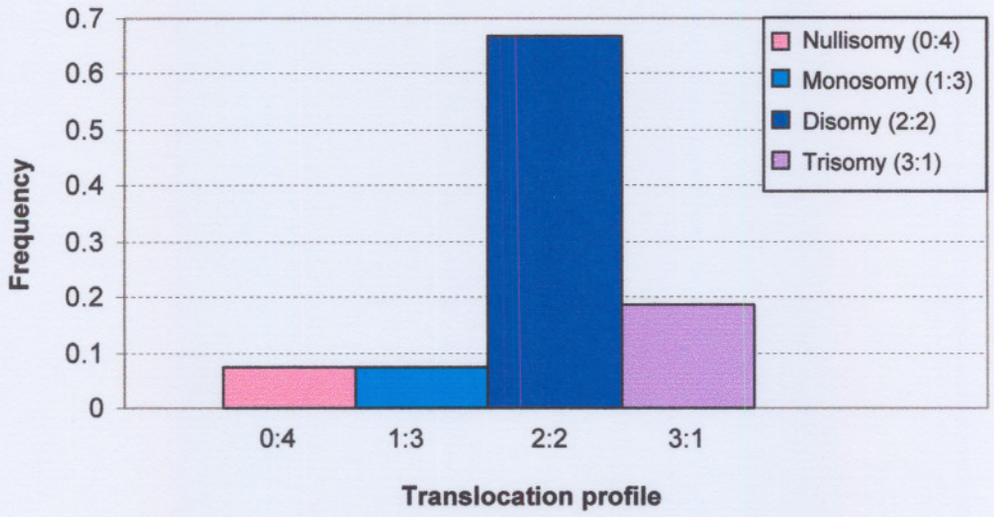
6.1.7.1 Frequency of the translocation profiles in the Khoi-San population

The frequency distribution of the translocation profiles observed in the Khoi-San population is presented in Graph 6.3 (a) and the data utilised to compose the graph is depicted in the table in Graph 6.3 (b). Only four of the seven translocation profiles that were observed in the larger cohort were present in the 27 individuals of the Khoi-San population included in this study. No quatrosomy or deletions of the p13E-11 hybridisation site on either chromosome 4 or 10 were observed in the Khoi-San population. Two thirds (0.67) of this population consisted of a standard disomic profile. After the grouping of the nullisomy and monosomy profiles into the excess of 10-on-4 profile and the trisomy and quatrosomy profiles into the excess of 4-on-10 profile, it was evident that these two grouped profiles had comparable translocation frequencies of 0.15 vs 0.19 respectively. A total translocation frequency of 0.34 was therefore observed for the Khoi-San population. This frequency is comparable to that observed in the reported Caucasian and Asian populations, as discussed in 4.3.2.3. The Caucasian and Asian populations will be collectively indicated as the "Eurasian population" for reference purposes in this study. Both these populations diverged from macrohaplogroup L ca. 50,000 years ago and cluster in haplogroups M and N, as depicted in Figure 5.3.

However, the small sample size of the Khoi-San population, almost ten times smaller than the sample size of the Black South African population included in this study, should be taken into consideration. The absence of the quatrosomic profile as well as the deletion of the p13E-11 hybridisation site on either chromosome 4 or 10 may be due to the insufficient sample size.

Graph 6.3: Frequency distribution of the translocation profiles of the Khoi-San population

(a)



(b)

	10-on-4			4-on-10	
Translocation profile	**0:4**	**1:3**	**2:2**	**3:1**	**4:0**
Total (27)	2	2	18	5	0
Frequency	0.07	0.07	0.67	0.19	0
Combined	0.15		0.67	0.19	

The data was utilised in all the analyses without rounding. The frequencies presented in the tables were rounded off to the second decimal.

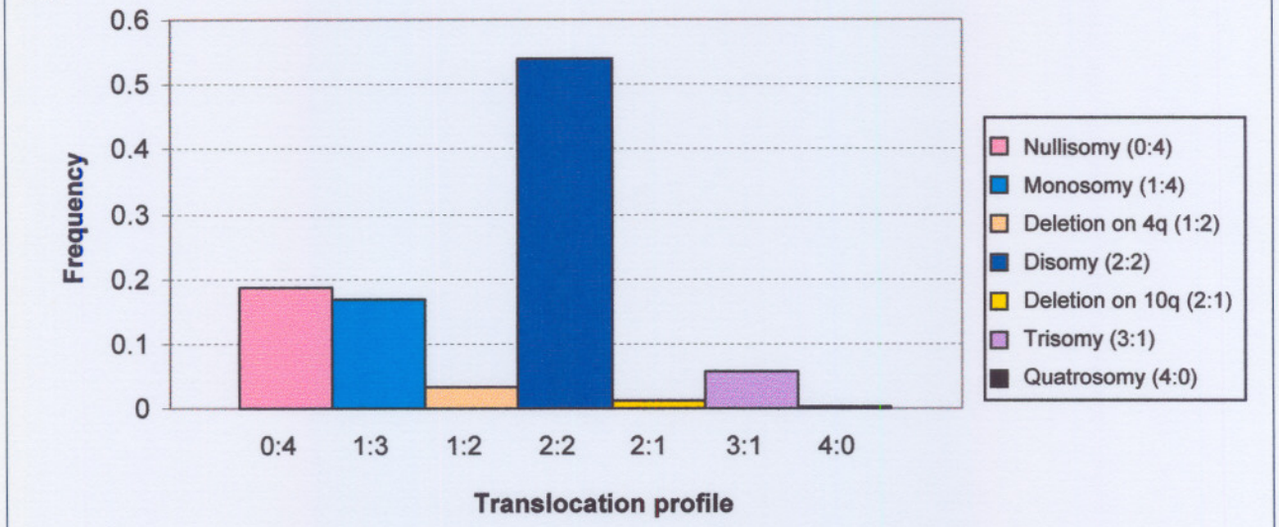
6.1.7.2 Frequency of the translocation profiles in the Black South African population

Seven distinct translocation profiles were observed in the 273 individuals from the Black South African population included in this study. The table depicted in Graph 6.4 (b) was utilised to construct the graph presented in Graph 6.4 (a) illustrating the frequency distribution of these seven profiles. A total translocation frequency of 0.42 was observed for this population. The nullisomy, monosomy, trisomy and quatrosomy profiles were utilised to calculate this frequency. The nullisomy and monosomy profiles (excess of 10-on-4) therefore contributed ca. 85% (0.355/0.42) of the total translocation frequency, indicating that an excess of 10-on-4 fragments in the Black South African population. Less than 5% of the individuals included in the study harboured a contribution value of 0.33 or 0.66. These values, as presented in Table 5.3, indicated a possible deletion of the p13E-11 hybridisation site on chromosomes 4 and 10 respectively. Matsumura *et al.*

(2002) also reported the presence of a deletion of the probe hybridisation site in less than 5% of the individuals included in their study.

Graph 6.4: Frequency distribution of the translocation profiles of the Black South African population

(a)



(b)

	10-on-4					4-on-10	
Translocation profile	**0:4**	**1:3**	**1:2**	**2:2**	**2:1**	**3:1**	**4:0**
Total (273)	51	46	9	147	3	16	1
Frequency	0.187	0.169	0.033	0.538	0.011	0.0596	0.004
Combined	0.355		0.033	0.538	0.011	0.062	

The data was utilised in all the analyses without rounding. The data presented in the tables were rounded off to the third decimal.

6.1.7.3 Chi-square analysis

The χ^2 test was subsequently utilised to determine whether the differences observed between the Khoi-San and the Black South African populations were statistically significant. The values for the Black South African population was utilised as the observed (O) values, while that of the Khoi-San was utilised as the expected (E) values. The χ^2 value was determined utilising Equation 5.2 (page 89).

The E values, as presented in Table 6.1 below, were calculated utilising the frequency of the specific translocation profile for the Khoi-San population as depicted in Graph 6.3 (b). This frequency was multiplied by the total sample size (273) of the Black South African population to obtain the E value. The total of individuals present within a specific

translocation profile, as presented in Graph 6.4 (b), were utilised as the O value. Since the two populations were compared, they were categorised as two different classes. The significance of the differences was therefore determined at a degree of freedom of one. The χ^2 values of the comparison between the nullisomy, monosomy, 10-on-4 (nullisomy and monosomy), disomy, trisomy and 4-on-10 (trisomy and quatrosomy) profiles of the two populations are presented in Table 6.1. Since no quatrosomy or deletions of the p13E-11 site on either chromosome 4 or 10 were observed in the Khoi-San population, these profiles could not be included in the χ^2 analyses. P-values of less than 0.0001 were observed for all the profiles, except for the disomic profile where the p-value is less than 0.01 but bigger than 0.001. The null hypothesis was therefore rejected, indicating that there was a significant difference between these profiles in the two populations included in this study. Evidence was therefore generated in support of the alternative hypothesis. The frequencies of the different translocation profiles therefore differed significantly between the Khoi-San and the Black South African population.

Table 6.1: Chi-square analysis

	Nullisomy	Monosomy	10-on-4	Disomy	Trisomy	4-on-10
O[1]	51	46	97	147	16	17
K-S frequency[2]	0.07	0.07	0.14	0.67	0.19	0.19
E	20.222	20.222	41.333	182.000	50.556	51.667
O-E	30.778	25.778	55.667	-35.000	-34.556	-34.667
χ^2	46.843	32.860	74.970	6.731	23.619	23.260
df = 1	p<0.0001	p<0.0001	p<0.0001	0.001<p<0.01	p<0.0001	p<0.0001

1 = Observed number of Black South African individuals per translocation profile, as presented in Graph 6.4; 2 = Frequency of Khoi-San individuals per translocation profile, as presented in Graph 6.3; K-S = Khoi-San population; E = expected value calculated utilising the frequency of the specific translocation profile for the Khoi-San population multiplied by the total sample size (n = 273) of the Black South African population; O = observed value; X^2 = chi-square; df = degrees of freedom.

Although a significant difference between the populations was observed, the differences in the sample sizes of these populations should be taken into consideration. However, only a limited number of samples were available from the Khoi-San population. None the less, since this is arguably the most unique population on a genetic level, the limited number of available samples was included in this study. The study reported here represents the largest sample size investigated for a particular population in FSHD to date, as indicated in Table 6.2.

6.1.8 META-ANALYSIS OF THE TRANSLOCATION FREQUENCY IN DIFFERENT POPULATIONS

A meta-analysis of the translocation frequency in several different populations is presented in Table 6.2. Both FSHD and control individuals, indicated in sections 1 and 2 of the table, were investigated by several of the authors, whereas some of the authors only investigated the presence of translocations in control individuals within their respective populations. Several of the authors did, however, not differentiate between the different translocation profiles. The comparisons made between the different populations and translocation profiles were thus performed after the combining of the nullisomy and monosomy profiles into the 10-on-4 profile and the trisomy and quatrosomy profiles into the 4-on-10 profile. A summary of the data is presented in section 3 of Table 6.2. This summary contains the averages of the FSHD and control data of all the populations, excluding the populations investigated in this study, which is presented separately.

Via comparison of the different translocation profiles within the different populations, presented in section 3 of Table 6.2, excluding the populations investigated in this study, it was evident that there is 1.7 times more 4-on-10 than 10-on-4 (12.55 vs 7.32) in the control individuals. An excess of 3.7 times that of 4-on-10 to 10-on-4 (19.95 vs 5.38) was further observed in the FSHD patients. There was therefore a trend of more of 4-on-10 in all the populations, FSHD and controls, but the excess of 4-on-10 was much more in the FSHD patients than in the Eurasian controls.

It was further indicated, via the comparison of the populations included in this study, that the Black South African population has 5.7 times more 10-on-4 than 4-on-10 (35.53 vs 6.23), the Khoi-San population has 1.2 times more 4-on-10 than 10-on-4 (18.52 vs 14.82). The Khoi-San population thus have ca. an equal frequency of 4-on-10 and 10-on-4.

Subsequent to comparison of the populations included in this study to the controls included in the populations already reported, it was illustrated that the Khoi-San population has 2 times more 10-on-4 and the Black South African population has 4.9 times more 10-on-4. It was therefore evident that the Black South African population had a much higher frequency of 10-on-4 than was reported for any of the other populations. This was especially evident from the much higher frequency (0.18) of nullisomy in the populations

included in this study is, compared to the control individuals included in the Eurasian studies.

Table 6.2: Meta-analysis of the translocation events between chromosomes 4q and 10q

Population		10-on-4[j]		Standard	4-on-10[j]		n	Exchange ratio (%)
		Null	Mono	Di	Tri	Quat		
1) FSHD patients								
Japanese[a]	FSHD	0.00	8.90	67.86	23.20	0.00	56	32.10
Chinese[b]	Sporadic FSHD	0.00	0.00	56.25	37.50	6.25	16	43.75
Chinese[b]	Familial FSHD	3.45	3.45	93.00	0.00	0.00	29	6.89
Chinese[c]	FSHD	5.71		81.43	12.86		70	18.57
2) Control individuals								
Japanese[a]	Control	0.65	7.19	72.50	19.00	0.65	153	27.50
Korean[a]	Control	0.00	4.03	70.16	25.80	0.00	124	29.80
Chinese[a]	Control	0.00	5.26	80.70	14.04	0.00	114	19.30
Dutch[d]	Control	0.00	10.00	80.00	10.00	0.00	50	20.00
Dutch[e]	Control	0.00	2.88	75.96	8.65	0.96	208	12.49[f]
Chinese[b]	Control	8.00		84.00	8.00		50	16.00
Chinese[c]	Control	9.62		80.77	9.62		52	19.23
Chinese[c]	Relatives of FSHD	10.91		85.45	3.64		55	14.55
Black SA[g]	Controls	18.68	16.85	53.85	5.86	0.37	273	41.76[h]
Khoi-San[i]	Controls	7.41	7.41	66.67	18.52	0.00	27	33.34
3) Summary								
Average	FSHD	5.38		74.64	19.95		171	25.33
Average	Controls	7.32		78.69	12.55		806	19.87
Black SA	Controls	35.53		53.85	6.23		273	41.76
Khoi-San	Controls	14.82		66.67	18.52		27	33.34

Null = nullisomy; Mono = monosomy; Di = disomy; Tri = trisomy; Quat = quatrosomy. a= Matsumura *et al.* (2002). b = Wang *et al.* (2003). c = Su *et al.* (2003). d = Van Deutekom *et al.* (1996). e = Van Overveld *et al.* (2000). f = exchange ratio excludes 11.5% of complex translocation profiles. g = As presented in this study. h = The exchange ratio excludes 4.4% due to complex translocation profiles. i = As presented in this study. j = The authors of reference c did not differentiate between the different translocation profiles. The authors of reference b did not differentiate between the translocation profiles for the control group investigated.

The comparison of the translocation frequencies between all the populations therefore indicated an excess of 4-on-10 in the FSHD and Eurasian control populations, with an excess of 10-on-4 in the Black South African population. The Khoi-San population was observed to be more comparable to the Caucasian populations than to the Black South African population. The frequencies of the Khoi-San and the Caucasian populations were comparable, but that of the Black South African population differ significantly.

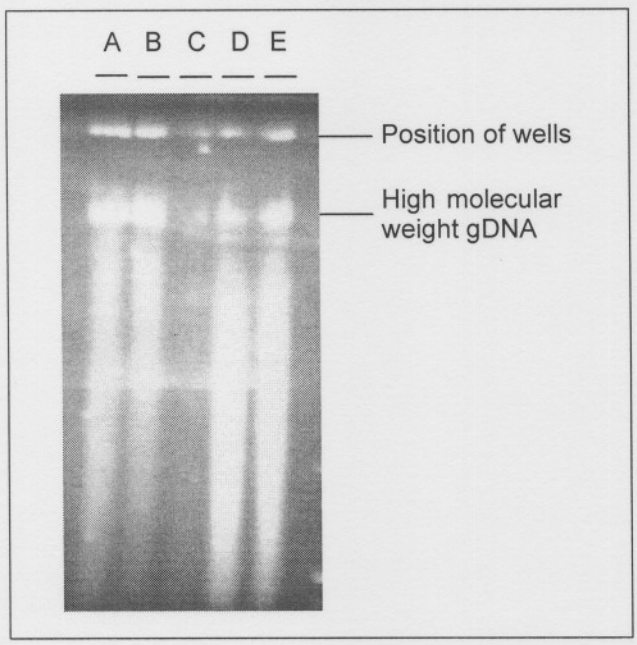
6.2 PULSED FIELD GEL ELECTROPHORESIS ANALYSIS

One of the specific objectives of this study, as discussed in paragraph 4.7.1, was to determine the difference in the allele sizes at the D4Z4 locus between the Black South African, the Khoi-San and other reported populations. Representative individuals of each of the translocation profiles were also included for PFGE analysis for the comparison and verification of the dosage test analyses. However, these objectives could not be fulfilled in their totality due to several experimental, technical and resource limitation difficulties that were encountered during the PFGE analysis.

6.2.1 PREPARATION AND RFLP ANALYSIS OF AGAROSE EMBEDDED gDNA

gDNA was successfully embedded in agarose from cultured cells as well as whole blood, as described in paragraphs 5.1.1.1 and 5.1.2 respectively. A representative agarose gel indicating the RFLP analysis of agarose embedded gDNA, performed as described in paragraph 5.2.1, is presented in Figure 6.12.

Figure 6.12: RFLP analysis of agarose embedded gDNA



Fragments were separated on a 0.5% agarose gel, which was electrophoresed at 1.6 $V.cm^{-1}$ for 28 hours in 1 x TBE buffer.

The agarose gel presented in Figure 6.12 was utilised to monitor the efficiency of the process of embedding the gDNA into agarose as well as the RFLP analysis of the gDNA within the agarose. Even after an electrophoresis time of 28 hours, unresolved high molecular weight gDNA was visible, as presented in Figure 6.12. This is due to the limited mobility of these large molecules in standard agarose electrophoresis (Schwartz and Cantor, 1984; Finney, 2000).

It is evident from the RFLP analyses of the samples included in Figure 6.12, that the gDNA was successfully embedded in the agarose. Furthermore the RFLP reactions on the gDNA were also successful. The amount of gDNA embedded within the plugs of the

sample presented in lane C in Figure 6.12 was less than that of the remaining samples. This may be due to the inadequate resuspension of the lymphocyte pellet resulting in a non homogeneous suspension, one of the difficulties to over come when utilising frozen blood samples for the preparation of agarose embedded gDNA.

6.2.2 AGAROSE GEL ELECTROPHORESIS FOR PULSED FIELD ANALYSIS

The electrophoresis of the agarose embedded gDNA was performed as described in paragraph 5.2.2. The Hoefer, HG 1000 Hula gel pulsed-field electrophoresis unit utilises a rotating circular gel plate in a homogeneous electric field for electrophoresis.

A total pulse time of 22 hours was utilised for the separation of the gDNA. Since the voltage was turned off during the reorientation of the gel the actual time necessary to provide a voltage of 200 V to the gel for 22 hours implied a total elapsed time of 35 hours. This elapsed time included a 30 minute run in time at the beginning of the run during which voltage was applied to the gel without any rotation. This was necessary to ensure that the liquid molecular weight marker that was loaded on the gel was not washed out of the wells via the rotation of the gel.

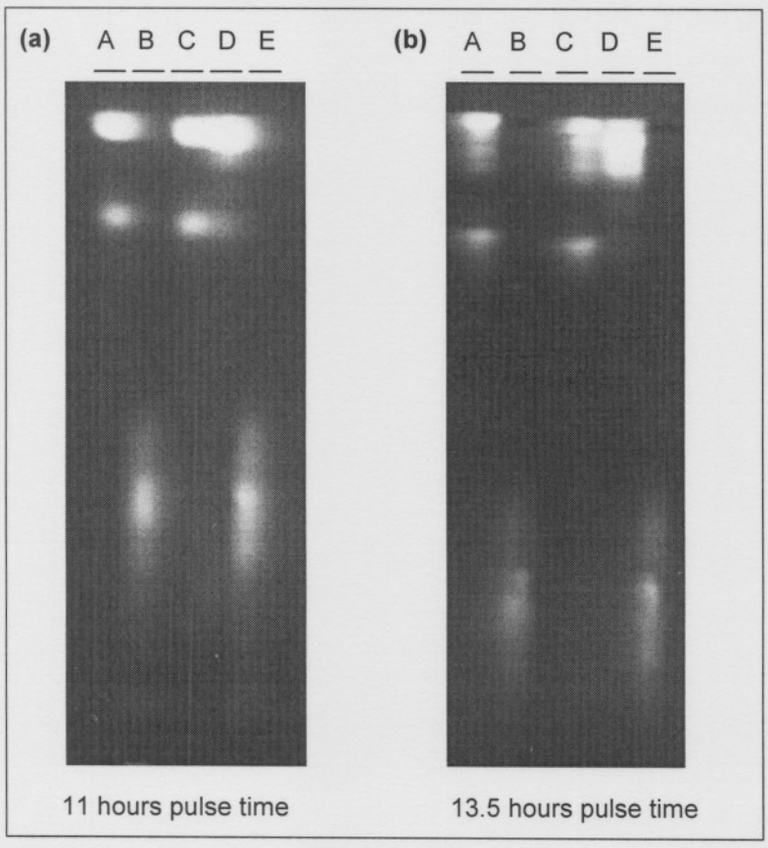
The PFGE system had to be connected to a refrigerated circulation water bath for the cooling of the 0.5 X TBE buffer within the system. However, since this resource was not available within the Centre, a non-refrigerated circulation water bath was utilised. The temperature of the water bath was regulated via the constant addition of ice. The temperature of the buffer was monitored after being pumped across the heat exchanger between the PFGE unit and the water bath, and displayed on the LCD screen of the unit.

Initially EtBr was not included in the gel as the intercalation of the EtBr molecule slows the reorientation of gDNA. The separation of the fragments could therefore not be monitored throughout the run. After the staining of the first PFGE gel with EtBr, subsequent to the 22 hour pulse time, no fragments were visible. It was therefore evident that the gDNA was run off the gel. Since this was a new technique optimised for the first time in the Centre, the initial experiment was repeated to ensure that the absence of gDNA on the gel was not due to any experimental error. However, no gDNA was visible on this gel upon staining with EtBr after the 22 hour pulse time. Subsequently, a PFGE analysis was performed utilising only half of the run time, therefore only including 2 cycles with a total pulse time of 11 hours. Upon EtBr staining, no gDNA could be visualised.

For the optimisation of this analysis, it was crucial to eliminate each of the factors that could influence the migration of the gDNA. The most obvious factor was the temperature control of the system, since this was performed manually via the addition of ice. The temperature of the buffer, as displayed on the LCD screen was therefore logged throughout the runs. However, the temperature on the screen indicated that the temperature of the buffer was between 18°C and 25°C during the entire run. Subsequently the actual temperature inside the buffer chamber was measured via a thermometer. Although the reading on the screen was 21°C a temperature of 40°C was measured inside the buffer chamber. It was therefore evident that this excess temperature increased the current resulting in the higher mobility of the gDNA. Upon further analysis it was noted that the pump of the system was defective. The heat generated throughout the run could therefore not be exchanged. The pump was subsequently repaired.

To enable constant monitoring of the migration of the gDNA throughout the run, as presented in Figure 6.13, EtBr was added to the subsequent gels. A PFGE analysis after 11 hours pulse time is presented in Figure 6.13 (a). The 8-48 kb ladder, as depicted in lanes B and E, was already separating after 11 hours, while the lambda ladder was only starting to separate after 13.5 hours, as illustrated in Figure 6.13 (b). The high molecular weight gDNA sample was also starting to separate after 13.5 hours.

Figure 6.13: Representative PFGE analysis



Fragments were separated on a 0.8% PFGE grade agarose gel, which was electrophoresed at 8.5 $V.cm^{-1}$ in 0.5 x TBE buffer, for the time as indicated below each image. A = Lambda ladder; B = 8-48 kb ladder; C = Lambda ladder, D = Agarose embedded gDNA of control sample; E = 8-48 kb ladder.

However, during the run it was noted that the temperature of the buffer inside the buffer chamber was increasing although the temperature of the water bath was kept constant.

Despite several attempts to overcome this technical hurdle it was eventually not possible to do so. Thus, due to the technical difficulties encountered during the PFGE analysis, it was concluded that it would not be possible to complete these analyses with the available resources.

The advantage of including PFGE in this study would have been that all four alleles, as well as the presence of translocations would have been visualised. It would thus have been possible to compare the allele sizes between the individuals included in this study to those already reported. However, the PFGE analysis would not have been able to detect the 3% of individuals included in this study who had a deletion of the p13E-11 hybridisation site since probe 9B6A was not available for further analysis.

All of the studies reported to date, utilising PFGE and dosage test analyses, obtained comparable results (Van der Maarel *et al.*, 1999; Matsumura *et al.*, 2002) and also the latest reports only utilised the dosage test to detect the translocation events in the specific populations (Matsumura *et al.*, 2002; Su *et al.*, 2003; Wang *et al.*, 2003). The dosage test was therefore utilised to determine the translocation frequency in the populations investigated.

The dosage test would not have detected the translocation event if exchange between 4q and 10q occurred distal to the first polymorphic *Bln* I site or if an individual carried a complex rearrangement resulting in a hybrid repeat array (consisting of both chromosome 4 and 10 repeat units). Van der Maarel *et al.* (1999), however, observed via PFGE that in the majority of cases the entire repeat array was translocated to the non-homologous chromosome and that only a small group (11%) of the translocations resulted in hybrid repeat units. Van Overveld *et al.* (2000) further indicated that only 4.3% of individuals with a hybrid repeat array have translocated repeats distal to the first polymorphic *Bln* I site. The dosage test will therefore rather underestimate the translocation frequency in a particular population. The translocation frequencies illustrated in the populations included in this study are thus viewed as representative of these populations, although it may actually be an underestimation of the true frequencies. At the least, the observed frequencies represent the minimum frequencies in these population.

6.3 HAPLOGROUP ANALYSES

The mtDNA has a high mutation rate (Wallace *et al.*, 1999), and over time specific mtDNA polymorphisms became fixed in different ancestral lineages. These polymorphisms are

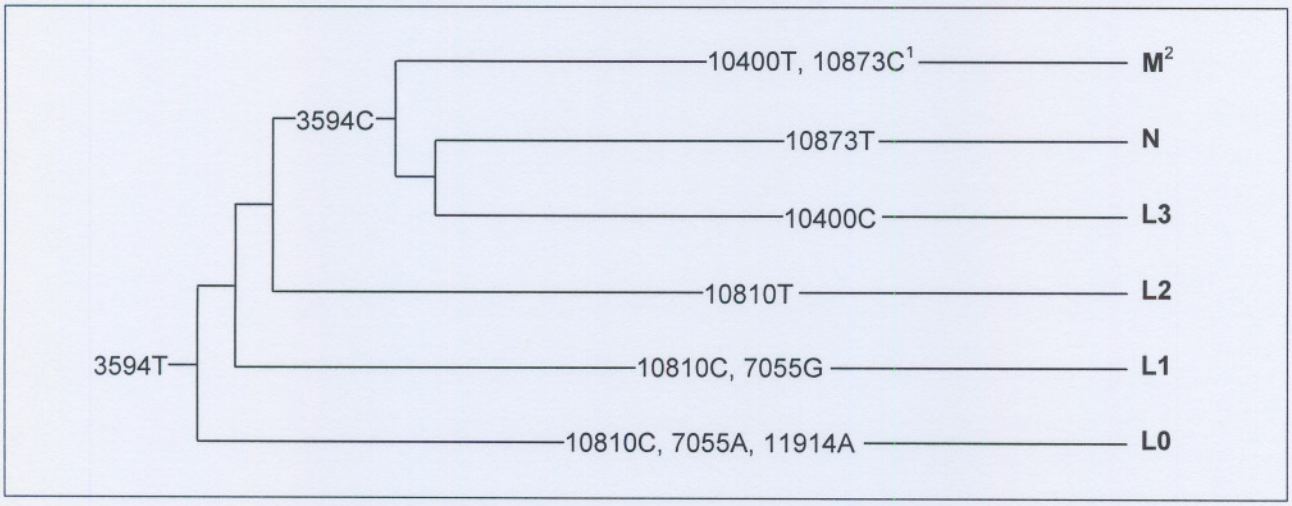
today utilised to delineate specific haplogroups for the reconstruction of the genetic history of humans over time. Evidence within the global human phylogeny revealed that older populations present with greater genetic variation (Wallace, 1994; Wallace, 1995).

The African-specific macro-haplogroup L is divided into three sublineages, namely L0, L1 and L2, as presented in Figure 6.14 (Mishmar *et al.*, 2003). These three sublineages represent the most ancient lineages of mtDNA haplogroups, with L0 being the most ancient, as depicted in Figure 6.14. In the same figure it is also illustrated that three haplogroups diverged from haplogroup L, the African-specific haplogroup L3 and the Eurasian haplogroups M and N.

Haplogrouping of selected individuals was not a specific objective of this study. However, upon dosage test analysis, it was noted that translocation events were an old phenomenon, and already present within the Khoi-San population. Several individuals of the Khoi-San population were included in previous haplogroup studies, and it was observed that these individuals all clustered in haplogroup L0 (Chen *et al.*, 2000; Mishmar *et al.*, 2003). However, no haplogroup data was available for any of the individuals from the Black South African population included in this study. It was therefore of interest to determine the haplogroups of a randomly selected group of these individuals to indicate whether the translocations were only present within specific haplogroups, i.e. L0, or if they were also present in the other haplogroups, L1, L2 and L3, into which the majority of Africans cluster. In addition the delineation of haplogroups would provide an indication of the sampling distribution of the Black South African individuals included in this study.

The L0, L1 and L2 haplogroups are defined by a thymine nucleotide at position 3594 in the mtDNA, while the L3, M and N haplogroups harbour a cytosine nucleotide at this position. These haplogroups are further delineated by specific polymorphisms as illustrated in Figure 6.14. The L0 haplogroup thus contains the following profile: 3594T, 10810C, 7055A and 11914A. Five SNPs within the mtDNA were analysed to determine the haplogroups of selected individuals within the Black South African population as described in paragraph 5.4. Approximately ten samples from each of the different translocation profile classes were randomly selected for haplogroup analysis, except for the quatrosomy class in which there was only one individual. A total of 41 individuals were therefore included. No individuals from the Khoi-San population were selected for haplogroup analysis, since previous studies in this population indicated that these individuals clustered in haplogroup L0 (Chen *et al.*, 2000; Mishmar *et al.*, 2003).

Figure 6.14: Evolutionary relationship and delineation of selected haplogroups

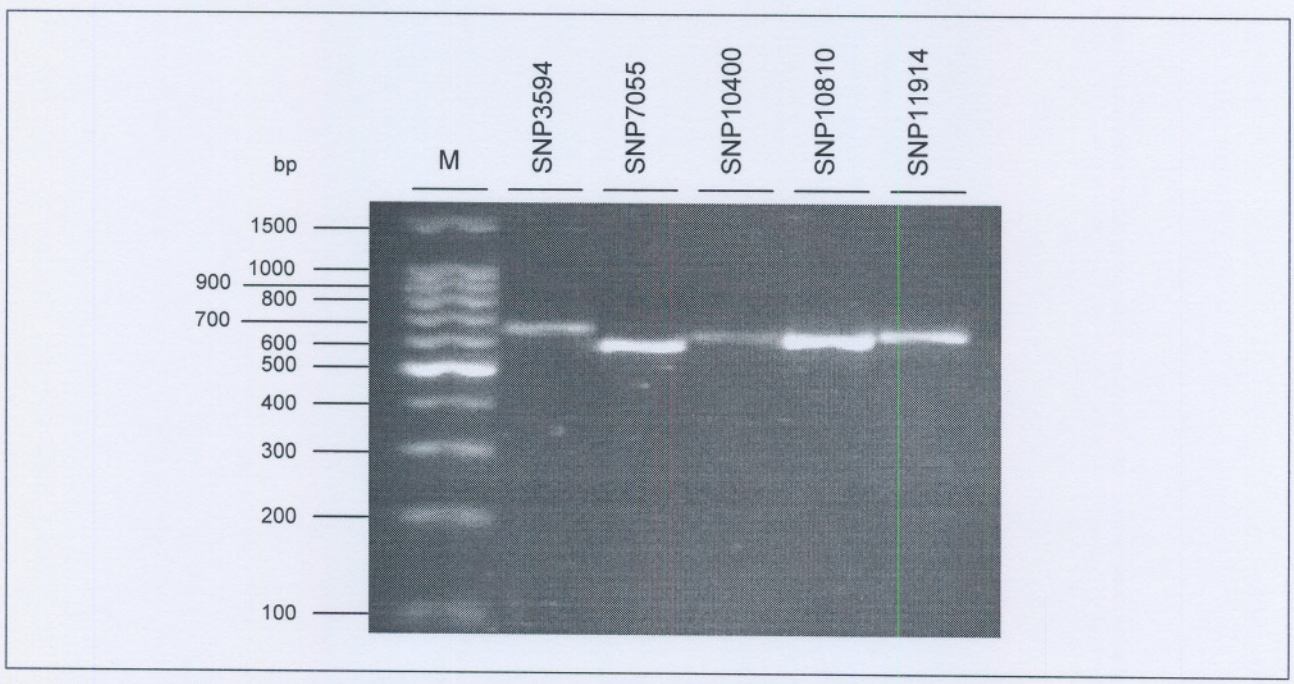


The lines are not drawn to scale. 1 = The numbers / letter combinations indicate the specific nucleotide and nucleotide number that should be present to characterise a particular haplogroup. 2 = The numbers / letter combinations in bold indicate the specific haplogroups.

6.3.1 AMPLIFICATION OF mtDNA FOR HAPLOGROUP ANALYSIS

The regions of the mtDNA encompassing the five SNPs were successfully amplified utilising the primers listed in Table 5.3 and the conditions described in paragraph 5.4. The amplicons of ca. 600 bp, with exact sizes as indicated in Table 5.4, of the five regions are presented in Figure 6.15.

Figure 6.15: Agarose gel electrophoresis of the amplified mtDNA for haplogroup analyses

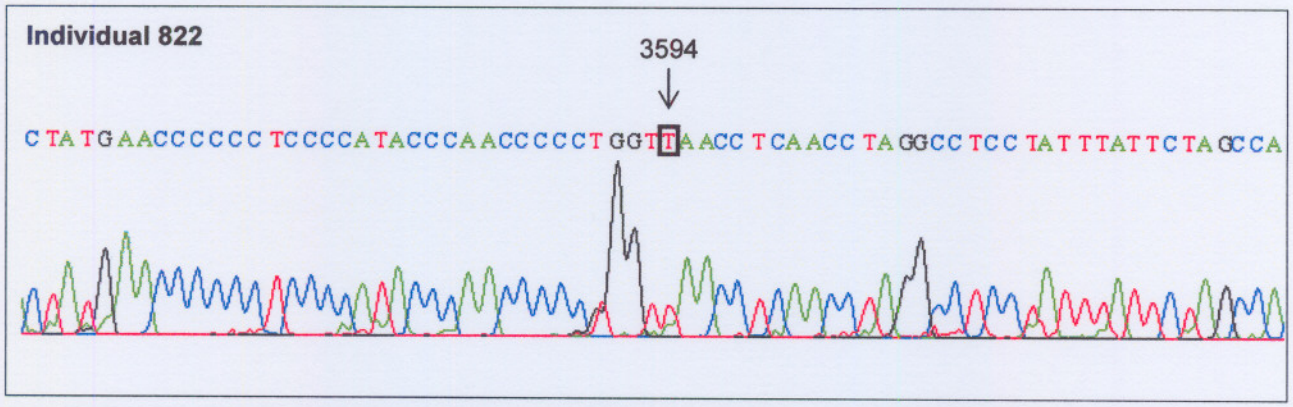


M = Molecular weight marker. Fragments were separated on a 2% agarose gel, which was electrophoresed at 5 $V.cm^{-1}$ for 45 minutes in 1 x TBE buffer.

6.3.2 RFLP ANALYSIS FOR SNP3594

As discussed in paragraph 5.4.2.1 an RFLP strategy was utilised to detect SNP3594. However, the PCR product of one individual (sample 822) was sequenced, as illustrated in Figure 6.16, to confirm that the correct region of the mtDNA was amplified prior to further analyses. The obtained sequence was aligned to the sequence presented in Table 5.4 utilising the BioEdit Sequence Alignment Editor version 5.0.9 programme.

Figure 6.16: Representative electropherogram of the mtDNA sequence encompassing SNP3594



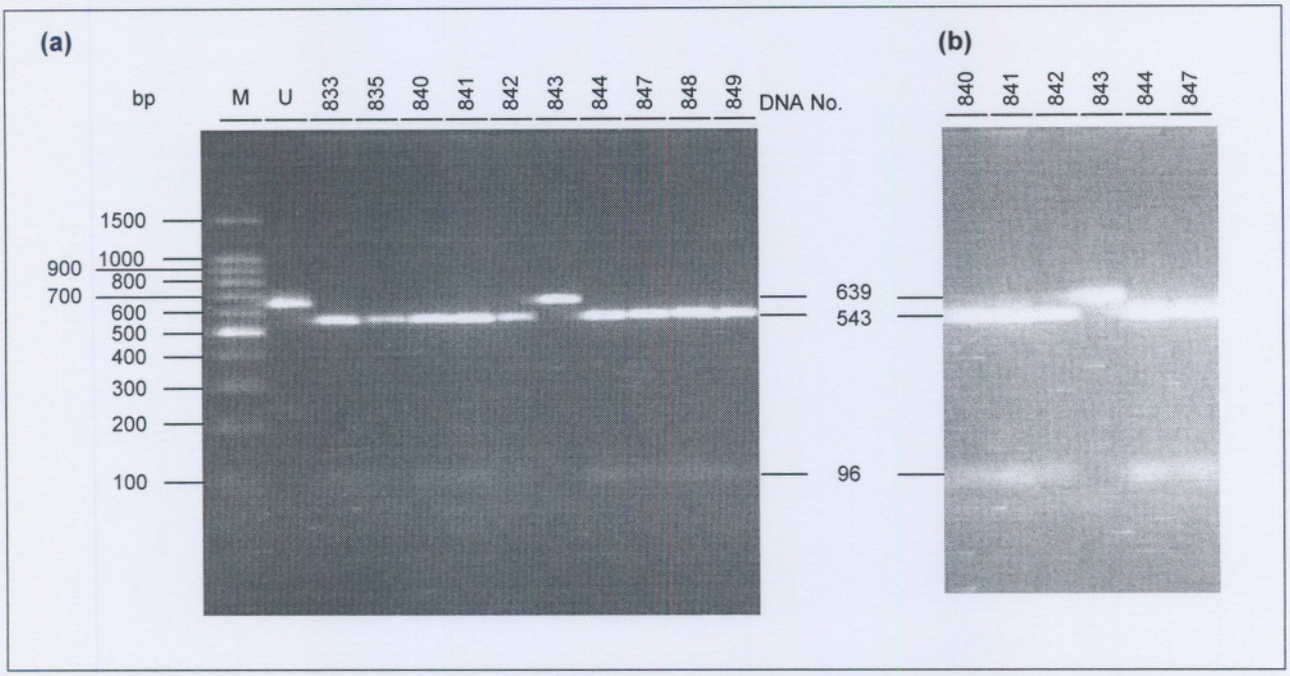
Nucleotide numbering on the sequence is according to MITOMAP (2003). The nucleotide indicated by a box characterises the position of SNP3594.

Subsequent to confirmation that the correct region was amplified, the amplified products of the control individuals were subjected to restriction enzyme digestion with *Ksp* AI harbouring the GTT↓AAC recognition sequence. The presence of a thymine at position 3594 characterises haplogroups L0, L1 and L2, as illustrated in Figure 6.16, while a cytosine at this position signifies haplogroups L3, M and N. The digestion of the 639 bp amplified product into a 543 bp and a 96 bp product therefore signified the presence of a thymine at this position. Haplogroups L3, M and N, represented by a C nucleotide at position 3594 resulted in the 639 bp amplified product remaining undigested. The samples that grouped in haplogroups L0, L1 and L2 were further analysed for SNP10810 as described in paragraph 6.3.3.3. SNP10400 was subsequently analysed for the samples that grouped in haplogroups L3, M and N, as presented in paragraph 6.3.3.2.

A representative agarose gel illustrating the RFLP analyses of selected individuals are presented in Figure 6.17. Sample 843, as presented in Figure 6.17, harboured a 639 bp undigested product signifying that this individual grouped into either haplogroup L3, M or N. The remaining individuals all harboured a 543 bp and a 96 bp product, characteristic of haplogroups L0, L1 and L2. However, the 96 bp product was faint upon initial analysis, as

illustrated in Figure 6.17 (a). The image intensity was subsequently over exposed, as depicted in Figure 6.17 (b), allowing the visualisation of the 96 bp product. The fluorescent signal generated by the 96 bp product was less intense than the 543 bp product due to less ethidium bromide intercalated into the smaller product.

Figure 6.17: Representative agarose gel of the RFLP analysis for SNP3594



M = Molecular weight marker. U = Uncut PCR product. Fragments were separated on a 2% agarose gel, which was electrophoresed at 5 $V.cm^{-1}$ for 90 minutes in 1 x TBE buffer.

6.3.3 AUTOMATED CYCLE SEQUENCE ANALYSIS

The remaining four SNPs were detected via an automated cycle sequencing strategy as discussed in paragraph 5.4.2.2. After the successful amplification of the four regions of the mtDNA encompassing the four SNPs of interest, the amplified products were purified to remove the excess primers, nucleotides, polymerases, salts and mineral oil from the reactions prior to sequencing. These entities were removed, since they could negatively interfere with the sequencing reaction. The average concentration of the purified products for each of the SNPs is presented in Table 6.3. The concentration of all the samples was sufficient for sequencing.

Table 6.3: Average concentration of purified products

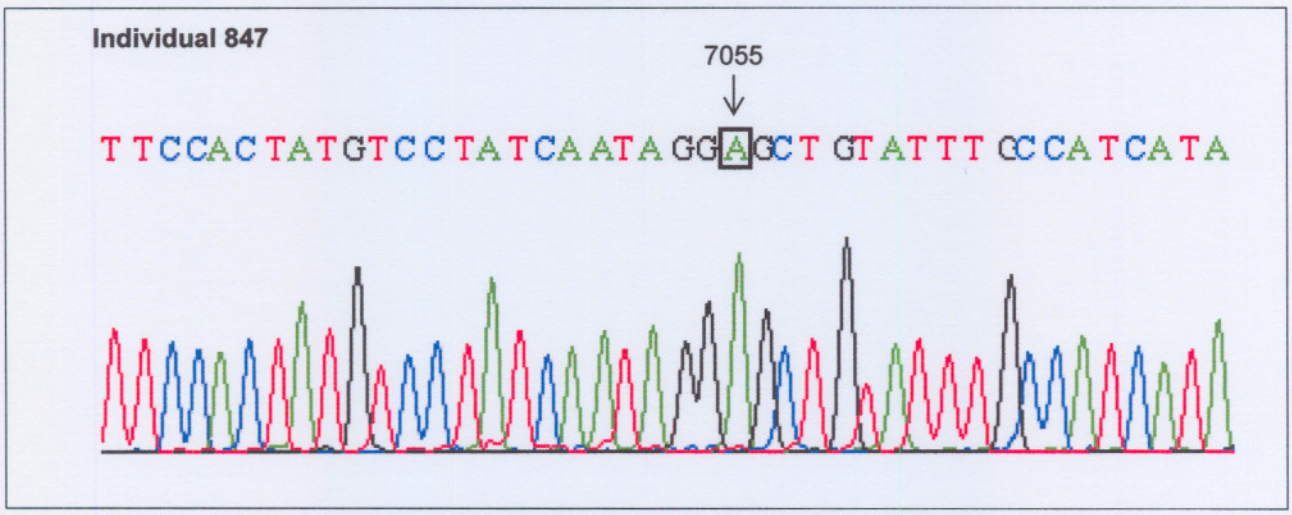
SNP	Average concentration
7055	21.33 $ng.\mu l^{-1}$
10400	15.62 $ng.\mu l^{-1}$
10403	24.66 $ng.\mu l^{-1}$
11914	21.35 $ng.\mu l^{-1}$

6.3.3.1 Single nucleotide polymorphism 7055

Subsequent to the analysis of SNP3594, SNP7055 was utilised to delineate between haplogroups L0 and L1. Haplogroup L1 was characterised by the presence of a G nucleotide at this position, while an A nucleotide at this position signified the L0 haplogroup. All of the samples analysed for this SNP harboured an adenine at this position, as presented in Figure 6.18. These samples were therefore all grouped into haplogroup L0. Thus, none of the samples analysed in this study represented haplogroup L1.

Since haplogroup L0 is further characterised by an adenine at nucleotide 11914, this SNP was subsequently analysed in all the samples analysed for SNP7055 as described in paragraph 6.3.3.4.

Figure 6.18: Representative electropherogram of the mtDNA sequence encompassing SNP7055



Nucleotide numbering on the sequence is according to MITOMAP (2003). The nucleotide indicated by a box characterises the position of SNP7055.

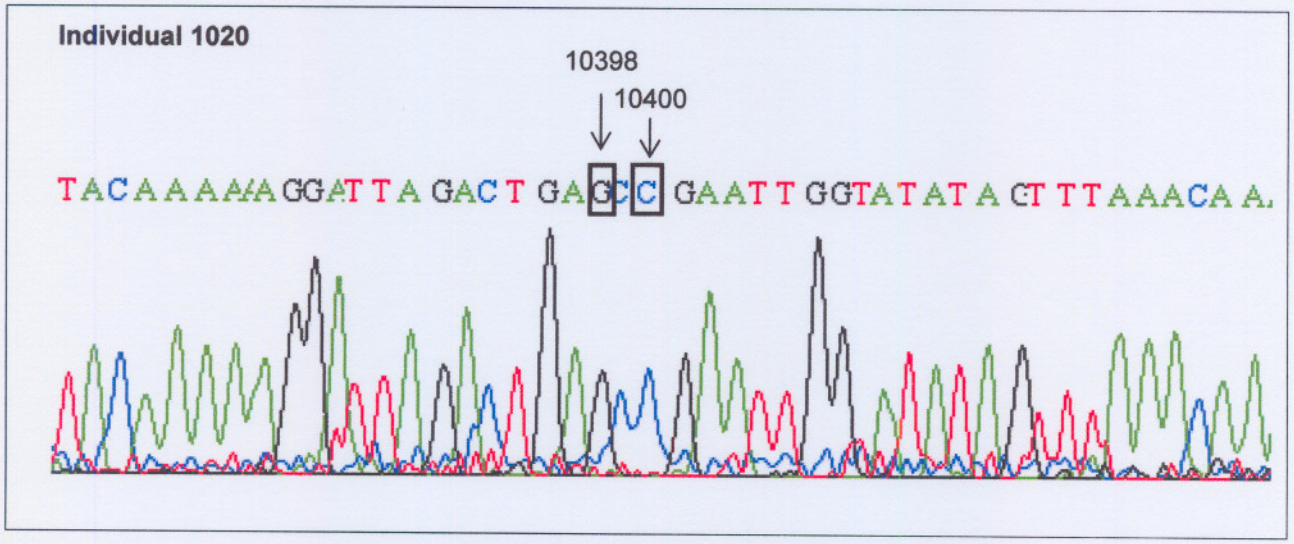
6.3.3.2 Single nucleotide polymorphism 10400

The SNP at position 10400 was utilised to differentiate between haplogroups M and L3. A T nucleotide at this position was characteristic of haplogroup M, while a C nucleotide was indicative of haplogroup L3, as presented in Figure 6.19. Individual 1020 was therefore grouped into haplogroup L3. Higher levels of background were observed for the sequence presented in Figure 6.19 than the sequence illustrated in Figure 6.18. However, the signal

intensity of the background peaks was less than that of the true peaks, and did therefore not influence the base calling of the sequence.

Via comparison to the mtDNA sequence presented in Table 5.6 it was observed that all the sequences analysed for this SNP in this study, harboured a nonsynonymous A to G nucleotide change at position 10398. This change was previously reported to be common to all mtDNA sequences of the African haplogroups (Wallace, 1999; Herrnstadt *et al.*, 2002). The presence of the 10398G polymorphism in this study therefore confirms the presence of this polymorphism in the African population and therefore in the L haplogroups.

Figure 6.19: Representative electropherogram of the mtDNA sequence encompassing SNP10400

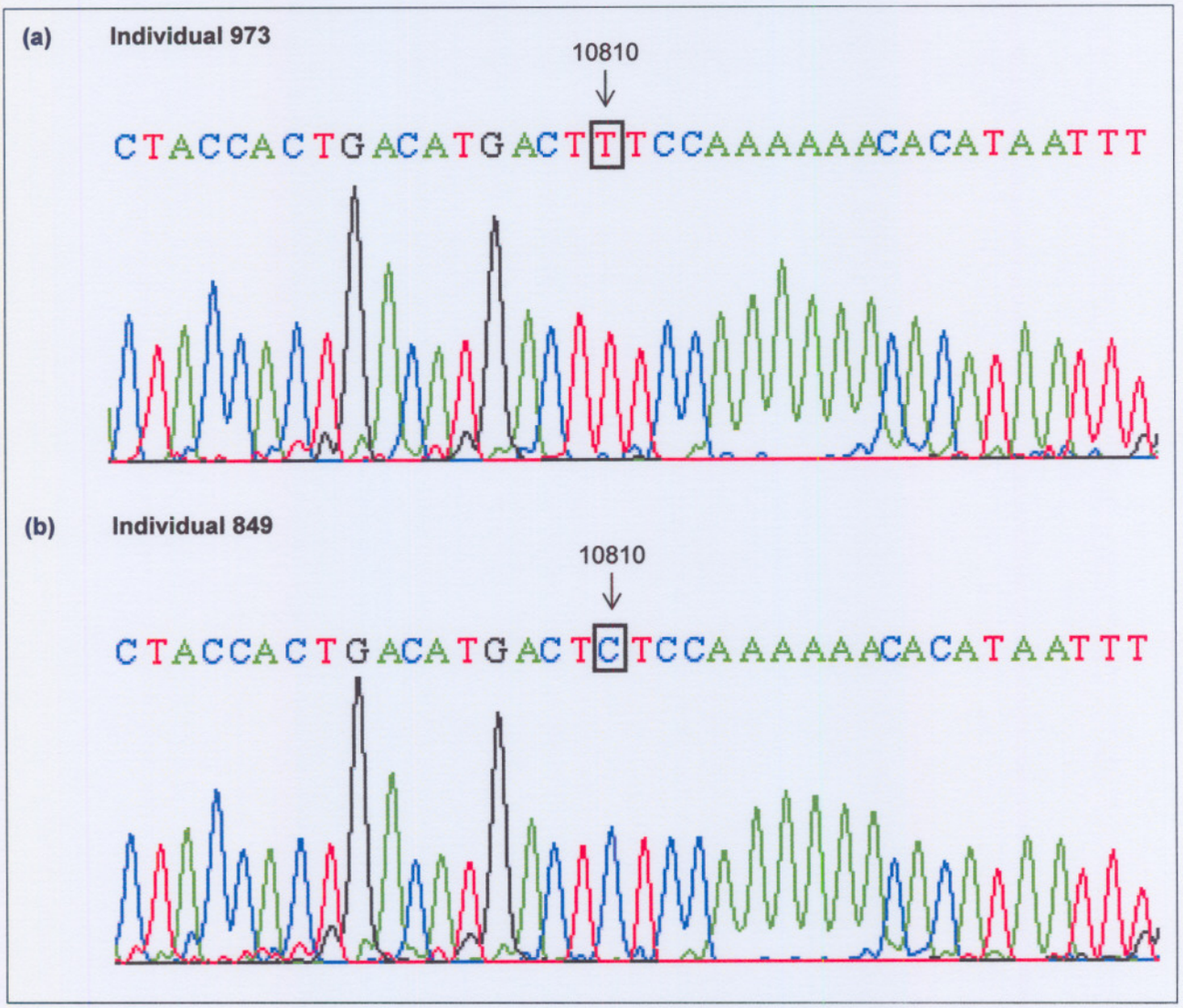


Nucleotide numbering on the sequence is according to MITOMAP (2003). The nucleotide indicated by a box characterises the position of SNP10400.

6.3.3.3 Single nucleotide polymorphism 10810

The samples grouped into haplogroups L0, L1 and L2, after the analysis of SNP3594, were further analysed for SNP10810 to enable delineation of haplogroup L2 from haplogroups L0 and L1. Representative sequences differentiating between these haplogroups are presented in Figure 6.20. As depicted in Figure 6.20 (a), individual 973 harboured a thymine at nucleotide 10810, which is characteristic of haplogroup L2. Individual 849, illustrated in Figure 6.20 (b) was classified into haplogroups L0 and L1, since this individual harboured a cytosine at this position.

Figure 6.20: Representative electropherogram of the mtDNA sequence encompassing SNP10810

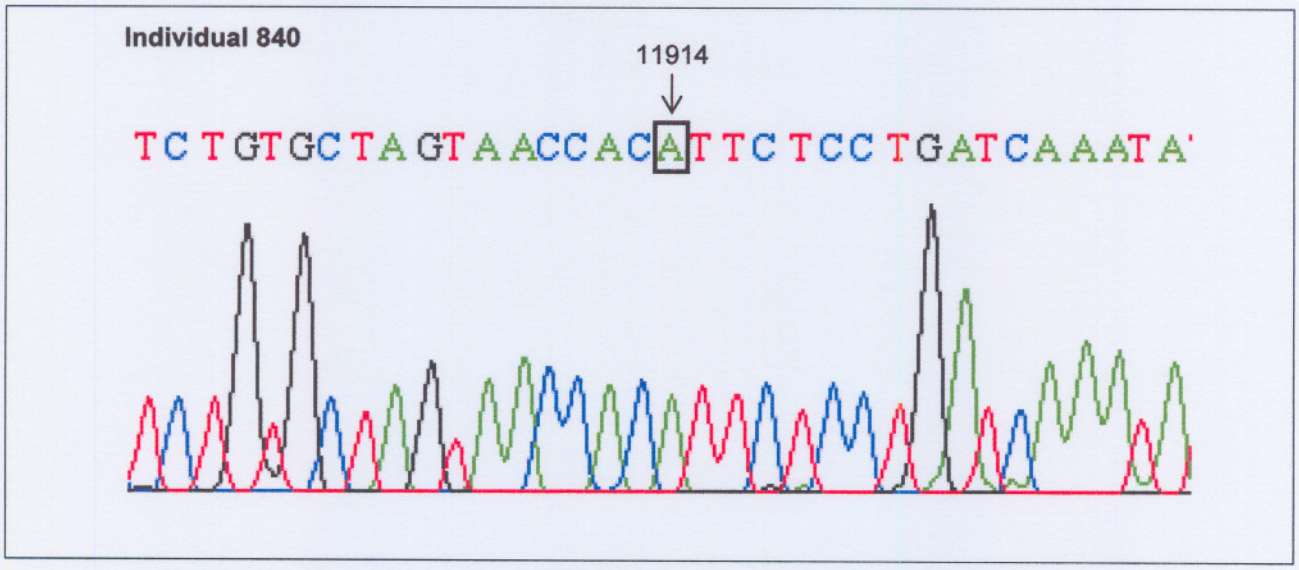


Nucleotide numbering on the sequence is according to MITOMAP (2003). The nucleotide indicated by a box characterises the position of SNP10810.

6.3.3.4 Single nucleotide polymorphism 11914

The SNP at position 11914 was utilised to confirm the inclusion of a particular sample into haplogroup L0, subsequent to the analysis of SNP7055. The presence of an adenine at this position signified the confirmation of the inclusion of a sample into haplogroup L0. Individual 840, as presented in Figure 6.21, was therefore grouped into haplogroup L0, since an adenine nucleotide was present at position 11914. All of the individuals who harboured an adenine at position 7055 (as described in paragraph 6.3.3.1), was confirmed to also harbour an adenine at position 11914, therefore confirming their inclusion into haplogroup L0.

Figure 6.21: Representative electropherogram of the mtDNA sequence encompassing SNP11914



Nucleotide numbering on the sequence is according to MITOMAP (2003). The nucleotide indicated by a box characterises the position of SNP11914.

6.3.4 DELINEATION OF HAPLOGROUPS

The delineated haplogroups of the 41 individuals selected for haplogroup analyses are presented in Table G.1, Appendix G. It is evident from Table G.1 that translocation was not only present within specific haplogroups, i.e. L0, since the individuals included for haplogroup analyses clustered in three different haplogroups, L0, L2 and L3. None of the individuals clustered in haplogroup L1. The individuals included from the Black South African population were, however, sampled at selected geographical regions, which may explain the absence of this specific haplogroup within this sample. The majority of samples grouped into haplogroup L0 (61%), with 27% and 12% clustering into haplogroups L2 and L3 respectively, as presented in Table 6.4. Within haplogroup L0 four different translocation profiles, nullisomy, monosomy, disomy and trisomy, were observed with relatively high frequencies of each. The quatrosomic individual grouped into haplogroup L2. This can, however, not be seen as a representative of the quatrosomic profile, since only one individual was observed to harbour this profile in the cohort investigated. Nullisomic, monosomic, disomic and quatrosomic individuals were grouped into haplogroup L2. The frequencies of each of these profiles were less than that observed in haplogroup L0, except for the nullisomy profile which remained the same, as presented in Table 6.4. Only three profiles, monosomy, disomy and trisomy, were grouped into haplogroup L3. The frequency of the disomic profile for haplogroups L2 and L3 was observed to be 0.18 for both, while the combined frequency of the trisomic and

monosomic profile for haplogroups L2 and L3 is comparable (0.32 vs. 0.31). Thus, no individuals with a nullisomic profile were observed in haplogroup L3.

If the randomly selected individuals included in the haplogroup analyses is a true representation of the Black South African population it is evident that the frequency of an excess of 10-on-4 (nullisomy and monosomy) is higher in haplogroups L0 and L2 than an excess of 4-on-10 (trisomy and quatrosomy), if the quatrosomic profile is not taken into consideration. The frequency of an excess of 10-on-4 and an excess of 4-on-10 is more comparable in haplogroup L3. It can also be noted that the frequencies of both an excess of 10-on-4 and an excess of 4-on-10 decreases from haplogroup L0 to haplogroup L3. The frequencies of the translocation profiles observed in haplogroup L3 is more comparable to that observed in the other populations, where frequencies of 20 to 32% were reported, as discussed in paragraph 4.3.2.3.

Table 6.4: Distribution of haplogroups versus translocation profiles

		L0	L1	L2	L3	n
0:4	Nullisomy	5	0	5	0	10
	Frequency	0.50	0	0.50	0	
3:1	Monosomy	7	0	1	2	10
	Frequency	0.70	0	0.10	0.20	
2:2	Disomy	7	0	2	2	11
	Frequency	0.64	0	0.18	0.18	
1:3	Trisomy	6	0	2	1	9
	Frequency	0.67	0	0.22	0.11	
4:0	Quatrosomy	0	0	1	0	1
	Frequency	0	0	1.0	0	
Total (n)		25	0	11	5	41
Total frequency		0.61	0	0.27	0.12	1.0

The presence of translocation events in the ancient Khoi-San population, clustered in haplogroup L0, as well as some individuals from the Black South African population also clustered in haplogroup L0, indicated that the origin of the translocation event between chromosome 4q and 10q was much older than originally anticipated. The age of the L0 haplogroup was determined to be 142 ± 17 x 10^3 years before present (Mishmar *et al.*, 2003). It is therefore evident that this age is also indicative of the relative evolutionary age of the translocation event.

CHAPTER SEVEN

CONCLUSIONS

The FSHD associated deletion at the D4Z4 locus was detected in 1993 by Van Deutekom *et al.* Despite several years of investigation in several populations, the exact mechanism by which the deletion occurred has not yet been elucidated. Many studies have been performed on different levels to characterise the aetiology of FSHD. Subsequent to one such study, Van der Maarel *et al.* (2000), concluded that the Caucasian population has an increased risk for FSHD, as an excess of 4-on-10 fragments were observed in this population. This emphasised the need to investigate other populations, apart from the Caucasian and Asian population, in order to determine whether this excess was only present in the Dutch population. It was envisaged that these investigations would reveal the reason for FSHD being more prevalent in certain populations.

FSHD cases have been reported in several populations, including, but not limited to the Brazilian, British, Chinese, Dutch, French, German, Italian, Japanese, Polish, South African Caucasian and Turkish populations. The FSHD phenotype has, however, not been reported in the Black South African population, even after the first reported cases in the Caucasian South African population several decades ago. The absence of reported cases in the Black South African population could be due to the fact that previous studies did not include individuals from this population. It can also be due to a low prevalence of FSHD in this population, decreasing the chance of individuals being diagnosed accurately. Alternatively, it is of course possible that there are no individuals affected with FSHD in this population.

In this study the D4Z4 locus was investigated for the first time on a molecular level in the Black South African and the Khoi-San populations. The results obtained were compared between the two populations under investigation, as well as Eurasian populations reported in the literature. It was envisaged that the analysis of the D4Z4 locus over evolutionary time would reveal characteristics that could explain the role of the translocations between chromosomes 4q35 and 10q26 in the pathogenesis of FSHD.

The complexity of the region in the human genome encompassing the FSHD locus was discussed in Chapter four. The presence of interchromosomal exchanges between the repeats on chromosome 4 and chromosome 10, indicating the plasticity of these regions, was a significant finding that contributed towards the understanding of the molecular basis of FSHD (Van Deutekom *et al.*, 1996a). Since the above finding has had significant implications for both the specificity and sensitivity of diagnostic DNA testing in FSHD, several studies have been undertaken in different populations to investigate the translocation events between chromosomes 4 and 10.

7.1 TRANSLOCATION PROFILES

Seven translocation profiles, nullisomy, monosomy, disomy, trisomy, quatrosomy and deletions of the p13E-11 hybridisation site on either chromosome 4 or 10, were observed in the study reported here, as discussed in paragraph 6.1.6. It is significant that each translocation profile could be uniquely identified, as no overlapping of any of the confidence intervals of the different profiles was observed.

All seven of the profiles were observed in the Black South African population. The nullisomic profile was observed in 19%, the monosomic profile in 17%, the disomic profile in 54%, the trisomic profile in 6%, the deletion of the p13E-11 hybridisation site on chromosomes 4 and 10 in 3% and 1% respectively and the quatrosomic profile was only observed in one (0.4%) of the 273 individuals included in this study. It is thus evident that plasticity exists at the D4Z4 locus in the Black South African population, which is evolutionary older than the Eurasian population. The 273 individuals from the Black South African population that were included in this study represent the largest sample size investigated for a particular population to date.

The Khoi-San population harboured only four (nullisomy, monosomy, disomy and trisomy), of the seven profiles observed in this study. No quatrosomic individuals, for the chromosome 4 type fragment, were observed. It can, however, not be definitively stated that in the entire Khoi-San population does not harbour any quatrosomic profiles, as only a small cohort of this population was included in this study. Similarly, the absence of the two profiles that are due to a deletion of the p13E-11 hybridisation site on either chromosome 4 or 10, may also be attributable to the small sample size. The nullisomic and monosomic profiles were each observed in 7%, the disomic profile in 67% and the trisomic profile in 19% of the 27 individuals included in this study. The translocations between

chromosomes 4 and 10 indicated the presence of plasticity at the D4Z4 locus in the Khoi-San population. The plasticity at the D4Z4 locus is thus not a new phenomenon in the human genome. It occurs in the most ancient lineage, and has survived over time. The implications of the presence of translocation events in the two populations included in this study are discussed in paragraph 7.3.

7.2 TRANSLOCATION FREQUENCY

The translocation frequencies of the different profiles of the Black South African and the Khoi-San populations were compared to each other as well as to those already reported. A summary of this comparison is presented in Table 7.1. Since all the reported studies did not differentiate between all the translocation profiles, it was necessary to combine the nullisomic and monosomic profiles into the 10-on-4 profile, and the trisomic and quatrosomic profiles into the 4-on-10 profile to enable the comparison between the three populations. The differentiated as well as the combined profiles for the populations included in this study are listed in Table 7.1.

A significant difference in the translocation frequency between the Khoi-San and the Black South African populations were observed in this study. The Black South African population harboured an excess of 10-on-4 in 36% and an excess of 4-on-10 in 6.4% of the individuals included in this study, whereas 15% and 19% of the Khoi-San population displayed these two translocation profiles respectively. A total translocation frequency of 42% was thus observed in the Black South African population. This was higher than the 34% observed in the Khoi-San population. Although the total translocation frequency of the two populations did not differ significantly, it was evident that the frequencies of the individual translocation profiles were different.

The frequency of the translocation profiles was also different between the Black South African and the reported Eurasian populations, with regard to the 10-on-4 and the 4-on-10 profiles. The Black South African population displayed an excess of 10-on-4 in 36% of the individuals, whereas only 5% and 6% were observed in the FSHD and control individuals in the reported Eurasian populations respectively. The 4-on-10 profile was only observed in 6% of the Black South African population, with 20% and 13% of the FSHD and control individuals in the reported populations displaying this profile respectively. A significant difference was also observed for the total translocation frequency of 42% in the Black

South African population, when compared to the 25% reported in the Eurasian FSHD population, and the 20% reported in the Eurasian controls.

The translocation frequency of the Khoi-San population was more comparable to that of the Eurasian populations, than to the translocation frequency of the Black South African population. This was especially true for the 4-on-10 profile. The implications for the differences observed in the translocation frequencies for the various translocation profiles in the different populations are discussed in paragraph 7.3.

Table 7.1: Translocation frequencies between chromosomes 4q and 10q in the Black South African, Khoi-San and reported populations

Population		Profile frequencies					n	Exchange ratio
		10-on-4		Standard	4-on-10			
		Null	Mono	Di	Tri	Quat		
Black SA	Controls	0.187	0.169	0.539	0.059	0.004	273	0.419
		0.356		0.539	0.063		273	0.419
Khoi-San	Controls	0.074	0.074	0.667	0.185	0.00	27	0.333
		0.148		0.667	0.185		27	0.333
Average for Eurasian populations	FSHD	0.054		0.746	0.200		171	0.254
	Controls	0.073		0.787	0.126		806	0.199

n = total number of individuals investigated. The frequencies presented in the table were rounded off to the third decimal.

An excess of 4-on-10 fragments was observed in both the Eurasian FSHD (20%) and control individuals (13%), as presented in Table 7.1. Van der Maarel *et al.* (2000) postulated that the excess of 4-on-10 fragments observed in the Dutch population is a significant, if not the major predisposing factor, that gives rise to the FSHD-type deletion which is associated with hypomethylation. The presence of fully homologous repeat arrays on non-homologous chromosomes, therefore 10-on-4 and 4-on-10, was hypothesised to result in interchromosomal gene conversion via non-homologous recombination between these two chromosomes. These recombination events are therefore suspected to cause the FSHD associated deletions. FSHD was thus proposed to be a recombination-based disorder and could be classified as a genomic disorder (Stankiewicz and Lupski, 2002).

It was postulated that populations with an excess 4-on-10 are therefore predisposed to FSHD (Van der Maarel *et al.*, 2000). However, Van der Maarel *et al.* (2000), stated that the presence of both 10-on-4 and 4-on-10 result in non-homologous recombination between these two chromosomes, and that these recombination events cause the FSHD associated deletions. Thus, according to Van der Maarel *et al.* (2000) both 10-on-4 and

4-on-10 should predispose a population to FSHD. A significant difference in the frequency of the two profiles (10-on-4 and 4-on-10) was observed in both the Eurasian FSHD and the Eurasian control populations. The Eurasian control population harbours an excess of 4-on-10, therefore predisposing this population to FSHD. The presence of FSHD cases in this Eurasian population thus confirms that these individuals were predisposed via the excess of 4-on-10 in this specific population.

However, to date only the predisposing effect of an excess of 4-on-10 could be illustrated in the Eurasian populations, since no excess of 10-on-4 were observed in any of the Eurasian populations investigated. The hypothesis that an excess of 10-on-4 is also a predisposing factor for FSHD, as stated by Van der Maarel *et al.* (2000), could therefore not be confirmed. As depicted in Table 7.1, a reverse frequency distribution of the 10-on-4 and 4-on-10 profiles was observed in the Black South African population. An excess of 10-on-4 was observed in the Black South African population compared to the excess of 4-on-10 in the Eurasian populations. If an excess of 10-on-4 is also a predisposing factor for FSHD, it would follow that the Black South African population is predisposed to FSHD. However, no FSHD case has to date been reported in this population. Despite several attempts in the past, not a single Black FSHD patient has been identified. Even if such a patient is identified in the future it remains that FSHD is not present in the Black South African population at the same level as observed in the Eurasian population to date. It is therefore proposed that the enrichment of 10-on-4 in the Black South African population is an epigenetic protective factor for FSHD. It is consequently hypothesised that the absence of reported FSHD cases in this population is due to this enrichment.

The translocation frequency of the Khoi-San population was more comparable to that of the Eurasian populations, than to the translocation frequency of the Black South African population. However, as presented in Table 7.1 the Khoi-San population has a higher frequency of 10-on-4, ca. twice as high as in the Eurasian FSHD population and ca. three times higher than in the Eurasian control population. An excess of 10-on-4 fragments was therefore also observed in this small sample from the Khoi-San population. Similar to clinical observations in the Black South African population to date, not a single Khoi-San FSHD individual has yet been reported.

7.3 PLASTICITY AT THE D4Z4 LOCUS IN A PHYLOGENETIC CONTEXT

Haplogroup analysis was not an objective of this study. However, subsequent to detection of translocations within the Khoi-San population, the haplogroups of approximately ten

individuals for each of the translocation profiles were determined in the Black South African population. This pilot study was performed to investigate whether the translocations were only present within specific haplogroups, i.e. L0, or if they were also present in haplogroups L1, L2 and L3, into which the majority of Africans cluster. Since the different haplogroups can be utilised to determine the relative time divergence between different populations clustering in these haplogroups, it was anticipated that the relative evolutionary age of the different profiles could be established.

The Khoi-San is the most diverse and ancient lineage in the world (Chen *et al.*, 2000). The Black South African population represents an intermediate lineage between the Khoi-San population and the Eurasian population. Initially, it was anticipated that the translocation frequency accumulated over time. Since ca. 20% of the Eurasian population was observed to harbour translocation events, it was postulated that a lower frequency of translocation events would be observed in the Black South African population, with even less, or none, in the Khoi-San population. A gradual decrease in the translocation frequency was therefore anticipated. This scenario was, however, not observed in the populations included in this study. The presence of translocation events in the ancient Khoi-San population indicated that the origin of the translocation event between chromosome 4q and 10q was much older than originally anticipated. The age of the L0 haplogroup was determined to be 142 ± 17 x 10^3 years before present (Mishmar *et al.*, 2003). The exact time divergence of the different translocation profiles could not be determined, however, as translocation events were already present in the Khoi-San population, which clustered in the L0 haplogroup, it is evident that the translocation event has a relative evolutionary age of approximately 142 ± 17 x 10^3 years.

The sample size of the Khoi-San population included in this study was small. However, as this is one of the most unique populations on a genetic level and only limited samples were available, the samples were none the less included in this study and did enable the investigation of the D4Z4 locus over evolutionary time. Comparison of the translocation frequencies of the different profiles between the two populations included in this study indicated that the Khoi-San population have translocation frequencies that are comparable to that of the Eurasian populations reported world wide, as illustrated in Table 7.1.

The translocation frequency of the profiles observed in the Black South African population was, however, observed to differ considerably from the Khoi-San and the reported Eurasian populations. The frequencies of the Khoi-San and the Eurasian populations were comparable, but that of the Black South African population differed significantly.

Given the lack of FSHD patients in the Black South African population to date, it would appear that this molecular difference, the enrichment of 10-on-4, as observed in the Black South African population thus had an evolutionary benefit for this population.

7.4 MODEL OF PATHOGENESIS IN FSHD

The deletion of an integral number of 3.3 kb repeat elements at the D4Z4 locus was originally reported as the causative factor in FSHD (Van Deutekom *et al.*, 1993). However, the complexity of the region of the human genome encompassing the FSHD locus was demonstrated through the interaction of several other genetic and epigenetic mechanisms implicated in FSHD. A model illustrating the complex genetic and epigenetic factors that interact to culminate in the FSHD phenotype is presented in Figure 7.1 (a) to (g), emphasising the multi-factorial aetiology of FSHD.

As discussed in paragraph 4.3.2.2, Lemmers *et al.* (2002) indicated that the deletion fragment present in FSHD affected individuals was exclusively associated with the 4qA allele. The deletion fragment alone can thus not cause FSHD, as depicted in Figure 7.1 (a.i), unless it is present on the 4qA variant of chromosome 4, as indicated in Figure 7.1 (a.ii). Similarly, as illustrated in Figure 7.1 (b.i), the 4qA allele alone can not cause FSHD, but has to be associated with a deletion fragment in order to lead to the pathology observed in FSHD, as presented in Figure 7.1 (b.ii).

The presence of a repression complex bound to each of the 3.3 kb repeat elements at the D4Z4 locus, as discussed in paragraph 4.6.4, was demonstrated by Gabellini *et al.* (2002). It was hypothesised that the repression complex can cause over expression of the genes associated with FSHD. This is caused by a reduction of bound complexes due to the decrease in the number of 3.3 kb repeat elements. The loss of repression thus results in over expression. Therefore, as presented in Figure 7.1 (c.i), the presence of the repression complex alone can not cause FSHD. However, when linked with a deletion fragment localised on a 4qA allele, as illustrated in Figure 7.1 (c.ii), the repression complex causes FSHD.

The long distance looping hypothesis, as discussed in paragraph 4.6.3, was proposed after the rejection of the PEV hypothesis by Jiang *et al.* (2003). This model indicated the mechanism by which the deletion at the D4Z4 locus could result in the FSHD phenotype

via the interaction of the repeats with a distant transcription control region, via a looping mechanism. As illustrated in Figure 7.1 (d.ii), the interaction of the repeats with this control region can only occur after the deletion of an integral number of 3.3 kb repeat elements. Furthermore, the deletion should be on the 4qA allele, as depicted in Figure 7.1 (a.i) and (b.ii). The interaction brought about by the long distance looping can not cause FSHD, as illustrated in Figure 7.1 (d.i). The interaction between these two regions may also require the binding of the repression complex to facilitate appropriate activation. This may seem contradictory to the repressor complex hypothesis, as discussed in paragraph 4.6.4. However, Jiang *et al.* (2003) did not indicate the binding of the repressor complex, but described the binding of the separate proteins, constituting the repressor complex. The looping hypothesis could therefore be described as a hybrid between the PEV and the repressor complex hypotheses. The main difference is however, that according to the PEV hypothesis, as described in paragraph 4.6.2, it is assumed that the normal length D4Z4 arrays are heterochromatic. The long distance looping hypothesis does, however, not make provision for the interaction of heterochromatic regions, and for this reason long distance can not be classified as a position effect.

Position effect variegation was considered to be the most likely cause of FSHD for several years (Winokur *et al.*, 1993; Hewitt *et al.*, 1994), as discussed in paragraph 4.6.2. It was originally hypothesised that the normal heterochromatinisation at the D4Z4 locus is lost in FSHD affected individuals resulting in the increase in expression of genes located proximal to the D4Z4 locus. Jiang *et al.* (2003) re-evaluated this hypothesis and indicated that the histone H4 acetylation levels were similar to unexpressed euchromatin and not to constitutive heterochromatin as previously believed. The authors further indicated that there was no position-dependent increase in the expression levels for the genes localised proximal to the D4Z4 locus. As a result, the PEV hypothesis was recently rejected as a possible genetic mechanism for FSHD, as indicated in Figure 7.1 (e.i).

As discussed in paragraph 4.5, a large degree of variation was observed in the gene expression data reported by several authors. Whereas Jiang *et al.* (2003) did not observe any over expression of the genes located proximal to the D4Z4 locus, Gabellini *et al.* (2002) indicated a position-dependent increase in the expression levels for the genes localised proximal to the D4Z4 locus. No model can, however, be excluded until the exact genetic aetiology of FSHD, via the predisposition or protection of individuals to specific mechanisms, has been defined. It may, therefore, even be possible that an unknown

factor, indicated by the question mark in Figure 7.1 (e.ii), interacts with the deletion at the D4Z4 locus, resulting in a PEV. However, given the current data, it seems unlikely that PEV, even in the presence of an unknown factor, can cause FSHD.

Van Overveld *et al.* (2003) illustrated that the D4Z4 locus was hypomethylated in FSHD affected individuals. This hypomethylation was also observed in individuals clinically diagnosed with FSHD but without a 4q35 deletion fragment. These authors thus demonstrated that the hypomethylation of the FSHD locus alone could result in FSHD, as depicted in Figure 7.1 (f.i). Contrary to several earlier hypotheses the deletion associated with the FSHD phenotype is therefore not a requirement for FSHD.

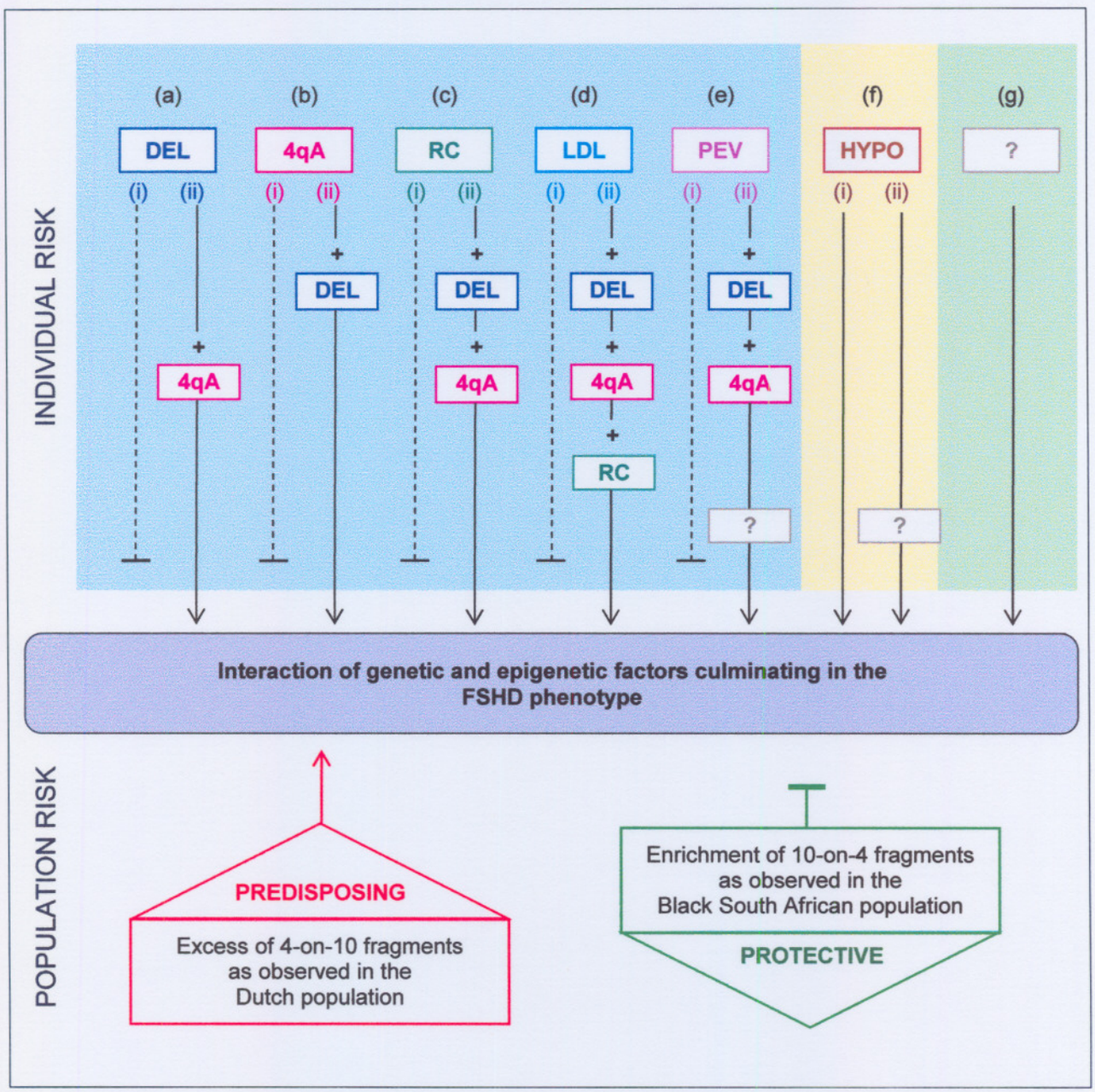
FSHD can thus be caused by different mechanisms, as indicated in the blue, yellow and green shaded areas in Figure 7.1. The mechanisms in the blue shaded area share a common aetiology in that a deletion fragment associated with a 4qA allele is always required in order to express the FSHD phenotype. FSHD can also be caused by the hypomethylation of the D4Z4 locus without a deletion being present, as indicated in the yellow shaded area. It is not yet known whether other genetic or epigenetic factors, indicated by the question mark in Figure 7.1 (f.ii), interact with the hypomethylation at the D4Z4 locus to result in the FSHD phenotype.

However, despite all the genetic and epigenetic factors already reported, there may be additional factors that contribute to the FSHD phenotype, as illustrated in Figure 7.1 (g). Interactions between all the above factors may even explain the clinical heterogeneity observed in FSHD, as discussed in paragraph 3.2.

A new dimension to the complex interactions that result in the FSHD phenotype was added via the data generated in this study. A significant difference was observed in the translocation frequency between the Black South African population and the reported Eurasian populations. The observed differences were therefore on a population level. Van der Maarel *et al.* (2000) postulated that the excess of 4-on-10 observed in the Dutch population, was a significant, if not the major predisposing factor that gives rise to the FSHD-type deletion that is associated with hypomethylation. The predisposed population thus have individuals that are more susceptible to FSHD, such as the 20% of the Eurasian population harbouring an excess of 4-on-10. The presence of multiple individuals affected with FSHD was indeed reported in the Eurasian population. Individuals within a predisposed population are therefore at a higher risk of developing FSHD.

An enrichment of 10-on-4 was observed in the Black South African population. It was postulated that this enrichment is an epigenetic protective factor for FSHD in this population. The individuals within this population are therefore less susceptible to FSHD. As a consequence, the excess and enrichment of specific translocation profiles in different populations is an additional factor that affects the aetiology of FSHD within specific populations.

Figure 7.1: Model of the complex genetic and epigenetic factors that interact to culminate in the FSHD phenotype



DEL = deletion of 3.3 kb repeat elements at D4Z4 locus. 4qA = variant A of the long arm of chromosome 4. RC = repressor complex. PEV = position effect variegation. LDL = long distance looping. HYPO = hypomethylation at D4Z4 locus. ? = unknown factor.

7.5 FUTURE DIRECTIONS OF RESEARCH IN FSHD

Despite the investigation of FSHD on both the molecular and clinical levels for several years and in various populations, the exact aetiology of FSHD has not yet been elucidated. A new dimension was added to the understanding of the complexity of FSHD via the data generated in this study. However, several aspects were identified during the cause of the study that need to be addressed in the future to enable a better understanding of FSHD thus paving the way for elucidating the aetiology of this complex disorder.

The inclusion of the Khoi-San population as well as the haplogroup analyses enabled the estimation of a relative evolutionary age for the translocation event. A significant difference between the Khoi-San and the Black South African populations was observed in this study. It is therefore evident that other populations from macrohaplogroup L should be included in future studies, to determine when the difference between the Khoi-San and Black South African populations occurred. Populations such as the Mbuti Pygmies and the Senegalese, but especially the Biaka Pygmies, should be considered. The Biaka Pygmies is also an ancient lineages but "younger" than the Khoi-San population in evolutionary terms, as illustrated in Appendix H (Chen *et al.*, 2000). The Mbuti Pygmies and Senegalese diverged more recently from the ancestral lineage than the Khoi-San (Chen *et al.*, 2000). Analysis of these populations could explain the differences observed between the Khoi-San and the Black South African populations via elucidation of the internal phylogenetic architecture of macrohaplogroup L.

Since the individuals from the Black South African population were only sampled at selected geographical regions it may explain the absence of haplogroup L1 within this sample group. It will thus be of interest to include individuals clustered within haplogroup L1 in future studies to enable the comparison of the translocation profiles and frequencies among all the L haplogroups.

Van Geel *et al.* (2002) demonstrated the presence of 4qA and 4qB alleles in the Dutch population. The association of the 4qA allele with the deletion fragment, present in the FSHD affected individuals, indicated the functional difference between the two alleles. The South African Caucasian FSHD population was previously investigated on a molecular level (Van der Merwe, 2002), to determine the presence of deletion fragments. Analysis to confirm the presence of only the 4qA allele in affected individuals thus still needs to be performed. It will also be of interest to perform this type of analysis in the

Khoi-San and Black South African populations as it may shed light on when the difference between the 4qA and 4qB alleles occurred.

The hypomethylation of the D4Z4 locus culminating in the FSHD phenotype without the presence of a deletion fragment indicated a novel mechanism for the aetiology of FSHD (Van Overveld *et al.*, 2003). It will therefore be necessary to investigate other disorders, indicated to have methylation deficiencies. For instance, in centromeric instability and facial anomalies (ICF) syndrome, where the D4Z4 locus was also observed to be hypomethylated (Kondo *et al.*, 2000; Bickmore and Van der Maarel, 2003; Van Overveld *et al.*, 2003). ICF differ from FSHD in that the hypomethylation was also observed on chromosome 10q in ICF. Individuals affected with ICF have severe recurrent respiratory infections, facial anomalies, mental retardation and developmental delay (Kondo *et al.*, 2000). However, it is significant that muscular abnormalities were not observed in individuals affected with ICF. The presence of neurological abnormalities in some individuals with ICF and FSHD need to be investigated further in order to determine if this common phenotypic characteristic is due to a shared, or partially shared aetiology. The inclusion of other disorders with methylation deficiencies will therefore enable differentiation between epigenetic factors that are specific to FSHD and those resulting in other disorders. It may also enable the elucidation of the mechanism of how the hypomethylation of the D4Z4 locus can result in the FSHD phenotype.

Since FSHD can be caused by the hypomethylation of the D4Z4 locus alone, as postulated by Van Overveld, (2003), it will be necessary to investigate the methylation status of the locus in different populations, such the Black South African population, in which no FSHD cases have been reported to date.

As discussed earlier in paragraph 6.2, the PFGE analyses could not be performed due to several difficulties that were encountered. The advantage of including PFGE in this study would have been that all four alleles, as well as the presence of translocations would have been visualised. It would thus have been possible to compare the allele sizes between the individuals included in this study to those already reported. PFGE analysis therefore remains a goal of the extended FSHD programme.

The new dimension added via this study does not influence the complex interactions that result in the FSHD phenotype, as presented in Figure 7.1 (a) to (g), since these factors result in FSHD on the individual level. However, in view of the data presented in this study

it is evident that several factors interact on both the individual and population levels. Firstly to predispose specific populations to FSHD, therefore increasing the risk of individuals within these populations of developing FSHD. Understanding of the complex interactions and elucidation of all the factors resulting in FSHD may also clarify the extreme clinical heterogeneity observed in FSHD.

In this era of genomics it is evident that trans-disciplinary analysis will be vital to elucidate the complexity of the region in the human genome encompassing the FSHD locus. The study reported here represents the first effort in FSHD to combine translocation data, via analysis of the Khoi-San and Black South African populations, as well as evolutionary data to open new avenues in order to understand the complexity of the aetiology of FSHD. The findings of this study support the statement by Lupski, (2003), that the...

" ... human genome has evolved an architecture
that may make us as a species more susceptible to rearrangements
causing genomic disorders."

Significant advances in FSHD research has been made since identification of the FSHD associated DNA rearrangements at the D4Z4 locus in 1992. The study reported here revealed the complex phylogenetic plasticity of the human genome architecture at the proximal subtelomeric domains of chromosomes 4q and 10q. This confirms that ...

"... genome architecture has profound implications
for how we as a species evolved and continue to evolve,
as well as ramifications for common traits and human disease."

Lupski (2003)

CHAPTER EIGHT

REFERENCES

8.1 GENERAL REFERENCES

Abbas A.K., Murphy K.M. and Sher A. Functional diversity of helper T lymphocytes. *Nature*, **383**, 787-793, 1996.

Adrian G.S., McCammon M.T., Montgomery D.L. and Douglas M.G. Sequences required for delivery and localisation of the ADP/ATP translocator to the mitochondrial inner membrane. *Mol. Cell. Biol.*, **6**, 626-634, 1986.

Altherr M.R., Bengtsson U., Markovich R. and Winokur S.T. Efforts towards understanding the molecular basis of facioscapulohumeral muscular dystrophy. *Muscle Nerve*, **Suppl 2**, S32-S38, 1995.

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

Bailey 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.

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.

Beckmann J.S., Brown R.H., Muntoni F., Urtizberea A., Bonneman C. and Bushby K.M.D. 66th/67th ENMC sponsored international workshop: The limb-girdle muscular dystrophies, 26-28 March 1999, Naarden, The Netherlands. *Neuromuscul. Disord.*, **9**, 436-445, 1999.

Bickmore W.A. and Van der Maarel S.M. Perturbations of chromatin structure in human genetic disease: recent advances. *Hum. Mol. Genet.*, **12** (2), R1-R7, 2003.

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. *Neuromuscul. Disord.*, **8**, 574-579, 1998.

Bushby K.M.D. The limb-girdle muscular dystrophies – multiple genes, multiple mechanisms. *Hum. Mol. Genet.*, **8** (10), 1875-1882, 1999.

Bushby K.M.D. and Beckmann J.S. The 105th ENMC sponsored workshop: pathogenesis in the non-sarcoglycan limb-girdle muscular dystrophies, Naarden, April 12-14, 2002. *Neuromuscul. Disord.*, **13**, 80-90, 2003.

Bushby K., Muntoni F. and Bourke J.P. 107th ENMC Internationa Workshop: the management of cardiac involvement in muscular dystrophy and myotonic dystrophy. 7th – 9th June 2002, Naarden, the Netherlands. *Neuromuscul. Disord.*, **13**, 166-172, 2003.

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). *Neuromuscul. Disord.*, **10**, 178-181, 2000.

Butz M., Koch M.C., Müller-Felber W., Lemmers R.J.L.F., Van der Maarel S.M. and Schreiber H. Facioscapulohumeral muscular dystrophy: Phenotype-genotype correlation in patients with borderline D4Z4 repeat numbers. *J. Neurol.*, **250**, 932-937, 2003.

Buxton J., Shelbourne P., Davies J., Jones C., Van Tongeren T., Aslanidis C., De Jong P., Jansen G., Anvret M., Riley B., Williamson R. and Johnson K. Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. *Nature*, **355**, 547-548, 1992.

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 *KpnI* repeat units implicated in Facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.*, **63**, 181-190, 1998.

Chen Y-S., Olckers A., Schurr T.G., Kogelnik A.M., Huoponen K. and Wallace D.C. mtDNA variation in the South African Kung and Khwe – and their genetic relationships to other African populations. *Am. J. Hum. Genet.*, **66**, 1362-1383, 2000.

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.

Clark L.N., Koehler U., Ward D.C., Wienberg J. and Hewitt J.E. Analysis of the organisation and localisation of the FSHD-associated tandem array in primates: implications for the origin and evolution of the 3.3 kb repeat family. *Chromosoma*, **105** (3), 180-189, 1996.

Cohn R.D. and Campbell K.P. Molecular basis of muscular dystrophies. *Muscle Nerve,* **23**, 1456-1471, 2000.

Cummings C.J. and Zoghbi H.Y. Trinucleotide repeats: Mechanisms and Pathophysiology. *Annu. Rev. Genomics Hum. Genet.*, **1**: 281-328, 2000.

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.

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.

Driss A., Amouri R., Ben Hamida C., Souilem S., Gouider-Khouja N., Ben Hamida M. and Hentati F. A new locus for autosomal recessive limb-girdle muscular dystrophy in a large consanguineous Tunisian family maps to chromosome 19q13.3. *Neuromuscul. Disord.*, **10**, 240-246, 2000.

Eagle M. Report on the muscular dystrophy campaign workshop: Exercise in Neuromuscular diseases, Newcastle, January 2002. *Neuromuscul. Disord.*, **12**, 975-983, 2002.

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. Fortnightly review: The muscular dystrophies. *BMJ.,* **317**, 991-995, 1998.

Emery A.E.H. Muscular dystrophy: The facts, 2nd edition. Oxford University Press, 2000.

Emery A.E.H. Muscular dystrophy into the new millennium. *Neuromuscul. Disord.*, **12**, 343-349, 2002a.

Emery A.E.H. The muscular dystrophies. *Lancet*, **359**, 687-695, 2002b.

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.

Fanin M., Melacini P., Boito C., Pegoraro E. and Angelini C. LGMD2E patients risk developing dilated cardiomyopathy. *Neuromuscul. Disord.*, **13**, 303-309, 2003.

Faustmann P.M., Farahati J., Rupilius B., Dux R., Koch M.C. and Reiners C. Cardiac involvement in facio-scapulo-humeral muscular dystrophy: a family study using Thallium-201 single-photon-emission-computed tomography. *J. Neurol. Sci.,* **144**, 59-63, 1996.

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.

Finney M. Pulsed-field gel electrophoresis. *Current Protocols in Molecular Biology*, **Unit 2.5B**, 2.5B.1-2.5B.9, 2000.

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). *Neuromuscul. Disord.*, **7**, 55-62, 1997.

Fischbeck K.H. and Garbern J.Y. Facioscapulohumeral muscular dystrophy defect identified. *Nature Genet.*, **2**, 3-4, 1992.

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.

Fitzsimons R.B. Facioscapulohumeral muscular dystrophy. *Curr. Opin. Neurol.*, **12**, 501-511, 1999.

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. *Neuromuscul. 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.

Funnel C.L. and George N.D.L. Epiphora as a presenting sign of facioscapulohumeral muscular dystrophy. *J. Pediatr. Ophthalmol. Strabismus*, **40** (2), 113-114, 2003.

Gabellini D., Green M.R. and Tupler R. Inappropriate gene activation in FSHD: A repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell*, **110**, 339-348, 2002.

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 correct genetic advice and a reliable prenatal diagnosis of the disorder. *Neuromuscul. 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.

Grewal 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.

Harper P.S., Harley H.G., Reardon W. and Shaw D.J. Anticipation in myotonic dystrophy: new light on an old problem. *Am. J. Hum. Genet.*, **51**, 10-16, 1992.

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.

Hsu Y-D., Kao M-C., Shyu W-C., Lin J-C., Huang N-E., Sun H-F., Yang K.D. and Tsao W-L. Application of chromosome 4q35-qter marker (pFR-1) for DNA rearrangement of facioscapulohumeral muscular dystrophy patients in Taiwan. *J. Neurol. Sci.*, **149**, 73-79, 1997.

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. *Neuromuscul. Disord.*, **4**, 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.

Jiang G., Yang F., Van Overveld P.G.M., Vedanarayanan V., Van der Maarel S. and Ehrlich M. Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. *Hum. Mol. Genet.*, **12** (22), 1-13, 2003.

Kazakov V.M., Bogorodinsky D.K., Znoyko Z.V and Skorometz A.A. The Facio-Scapulo-Limb (or facioscapulohumeral) type of muscular dystrophy. *Europ. Neurol.* **11**, 236-260, 1974.

Kazakov V.M. History of the recognition and description of the facioscapulohumeral muscular dystrophy (FSHD) and on the priorities of Duchenne, Erb, Landouzy and Dejerine. *Acta Cardiol.*, **VII**, 79-94, 1995.

Kazakov V.M. On the name of FSHD and the famous discussion between Erb and Landouzy-Dejerine once again. *Acta Myol.*, **II**, 72-74, 1998.

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.

Kondo T., Bobek M.P., Kuick R., Lamb B., Zhu X., Narayan A., Bourc'his D., Viegas-Péquignot E., Ehrlich M. and Hanash S.M. Whole-genome methylation scan in ICF syndrome: hypomethylation of non-satellite DNA repeats D4Z4 and NBL2. *Hum. Mol. Genet.*, **9** (4), 597-604, 2000.

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.

Kotzin B.L. Twins and T-cell responses. *Nature*, **364**, 187-188, 1993.

Krasnianski M., Eger K., Neudecker S., Jakubiczka S. and Zierz S. Atypical phenotype in patients with facioscapulohumeral muscular dystrophy 4q35 deletion. *Arch. Neurol.*, **60**, 1421-1425, 2003a.

Krasnianski M., Neudecker S., Eger K., Jakubiczka S. and Zierz S. Typical facioscapulohumeral dystrophy phenotype in patients without FSHD 4q35 deletion. *J. Neurol.*, **250**, 1084-1087, 2003b.

Kunkel L.M., Monaco A.P., Middlesworth W., Ochs H.D., Latt S.A. Specific cloning of DNA fragments absent from the DNA of a male patient with an X chromosome deletion. *Proc. Natl. Acad. Sci. USA.*, **82**, 4778-4782, 1985.

La Spada A.R., Paulson H.L., Fischbeck K.H. Trinucleotide repeat expansion in neurological disease. *Ann Neurol.*, **36**, 814-822, 1994.

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.

Laoudj-Chenivesse D., Bisbal C., Carnac G., Vassetzky Y., Bouillot S., Hugon G. and Fernandez A. Protein profiling in FSHD muscle: Evidence for altered patterns of fiber-specific muscle proteins expression and mitochondrial dysfunction. Submitted for publication, 2003.

Leclercq I., Mattéotti C., Deneubourg G., Van der Maarel S,.M., Frants R.R., Padberg G., Coppée F. and Belayew A. Presence of a double homeobox gene in the D4Z4 locus of a patient with FSHD despite a subtelomeric exchange between chromosome 4q35 and 10q26. FSHD International Consortium Research Meeting, San Diego, U.S.A., 2001.

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. Complete allele information in the diagnosis of facioscapulohumeral muscular dystrophy by triple DNA analysis. *Ann. Neurol.*, **50**, 816-819, 2001.

Lemmers R.J.L.F., De Kievit P., Sandkuijl L., Padberg G.W., Van Ommen G-J.B., Frants R.R. and Van der Maarel S.M. Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. *Nature Genet.*, **32** (2), 235-235, 2002.

Lemmers R.J.L.F., Osborn M., Haaf T., Rogers M., Frants R.R., Padberg G.W., Cooper D.N., Van der Maarel S.M. and Upadhyaya M. D4F104S1 deletion in facioscapulohumeral muscular dystrophy: Phenotype, size and detection. *Neurology*, **61**, 178-183, 2003.

Lemmers R.J.L.F. Human and Clinical Genetics, Leiden University Medical Centre, Leiden, the Netherlands. Personal communication 2004.

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.

Liao D. Concerted Evolution: molecular mechanism and biological implications. *Am. J. Hum. Genet.*, **64**, 24-30, 1999.

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. *Neuromuscul. Disord.*, **8**, 126-130, 1998.

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

Lupski J.R. Genomic disorders: Recombination-based disease resulting from genome architecture. *Am. J. Hum. Genet.*, **72**, 246-252, 2003.

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.

Mackenzie W.G., Riddle E.C., Earley J.L. and Sawatzky B.J. A neurovascular complication after scapulothoracic arthrodesis. *Clin. Orthop.*, **408**, 157-161, 2003.

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

Marshall J., Leigh D., McArthur S. and De Boer K. Preimplantation genetic diagnosis (PGD) for facioscapulohumeral muscular dystrophy (FSHD) – a reproductive option using genetic linkage analysis. FSHD International Consortium Research Meeting, Los Angeles, U.S.A., 2003.

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.

Matsumura T., Goto K., Yamanaka G., Lee J.H., Zhang C., Hayashi Y.K. and Arahata K. Chromosome 4q:10q translocations: comparison with different ethnic populations and FSHD patients. *BMC Neurol.*, **2** (1), 7-11, 2002.

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

McGonigal A., Thoma A.M. and Petty R.K.H. Facioscapulohumeral muscular dystrophy and myasthenia gravis co-existing in the same patient: a case report. *J. Neurol.*, **249**, 219-220, 2002.

McNally E.M., De Sá Moreira E., Duggan D.J., Bönnemann C.G., Lisanti M.P., Lidov H.G.W., Vainzof M., Passos-Bueno M.R., Hoffman E.P., Zatz M. and Kunke L.M. Caveolin- in muscular dystrophy. *Hum. Mol. Genet.*, **7** (3), 871-877, 1998.

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.

Minetti C., Sotgia F., Bruno C., Scartezzini P., Broda P., Bado M., Masetti E., Mazzocco M., Egeo A., Donati M.A., Volonté D., Galbiati F., Cordone G., Bricarelli F.D., Lisanti M.P. and Zara F. Mutations in the caveolin-3 gene cause autosomal dominant limb-girdle muscular dystrophy. *Nat. Genet.*, **18**, 365-368, 1998.

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.

Mishmar D., Ruiz-Pesini E., Golik P., Macaulay V., Clark A.G., Hosseini S., Brandon M., Easley K., Chen E., Brown M.D., Sukernik R.I., Olckers A. and Wallace D.C. Natural selection shaped regional mtDNA variation in humans. *Proc. Natl. Acad. Sci. USA.*, **100**, 171-176, 2003.

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.

Muchir A., Bonne G., Van der Kooi A.J., Van Meegen M., Baas F., Bolhuis P.A., De Visser M. and Schwartz K. Identification of mutations in the gene encoding lamins A/C in autosomal dominant limb girdle muscular dystrophy with atrioventricular conduction disturbances (LGMD1B). *Hum. Mol. Genet.*, **9** (9), 1453-1459, 2000.

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.

Muntoni F. and Guicheney P. 85th ENMC International Workshop on Congenital Muscular Dystrophy 6th International CMD Workshop, 1st Workshop of the Myo-Cluster Project 'GENRE', 27-28th October 2000, Naarden, The Netherlands, *Neuromuscul. Disord.*, **12**, 69-78, 2002.

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. *Neuromuscul. Disord.*, **13**, 428-439, 2003.

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.

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). *Neuromuscul. Disord.*, **1**, 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.

Ray P.N., Belfall B., Duff C., Logan C., Kean V., Thompson M.W., Sylvester J.E., Gorski J.L., Schmickel R.D. and Worton R.G. Cloning of the breakpoint of an X;21 translocation associated with Duchenne muscular dystrophy. *Nature*, **318** (19), 672-675, 1985.

Persky A.M. and Brazeau G.A. Clinical pharmacology of the dietary supplement creatine monohydrate. *Pharmacol. Rev.*, **53**, 161-176, 2001.

Reilly M.M. and Hanna M.G. Genetic neuromuscular disease. *J. Neurol. Neurosurg. Psychiatry*, **73** (suppl II), ii12-ii21, 2002.

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.

Samuels M.L. Statistics for the life sciences. Prentice-Hall, Inc, 329-334, 1989.

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.

Schwartz D.C. and Cantor C.R. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell*, **37**, 67-75, 1984.

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.

Speer M.C., Vance J.M., Grubber J.M., Graham F.L., Stajich J.M., Viles K.D., Rogala A., McMichael R., Chutkow J., Goldsmith C., Tim R.W. and Pericak-Vance A. Identification of a new autosomal dominant limb-girdle muscular dystrophy locus on chromosome 7. *Am. J. Hum. Genet.*, **64**, 556-562, 1999.

Stankiewicz P. and Lupski J.R. Genome architecture, rearrangements and genomic disorders. *Trends. Genet.*, **18** (2), 74-82, 2002.

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.

Su Q.X., Zhang C., Zeng Y., Lu X.L., Liu X.R., Wang Z.H. and Zhu Y.Z. Translocation between chromosomes 4q35 and 10q26 in facioscapulohumeral muscular dystrophy. *Zhonggua Yi Xue Ke Xue Yuan Xue Bao*, **25** (5), 581-584, 2003. (Article in Chinese)

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., Myers G.J., Weiffenbach B. and Griggs R.C. Scapuloperoneal Syndromes: Absence of linkage to the 4q35 FSHD locus. *Arch. Neurol.*, **52**, 1069-1072, 1995.

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.

Tonini M.M.O., Passos-Bueno M.R., Cerqueira A., Matioli S.R., Pavanello R. and Zatz M. Asymptomatic carriers and gender differences in facioscapulohumeral muscular dystrophy (FSHD). *Neuromuscul. Disord.*, **14,** 33-38, 2004.

Tupler 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.

Tupler R., Perini G., Pellegrino M.A. and Green M.R. Profound misregulation of muscle-specific gene expression in facioscapulohumeral muscular dystrophy. *Proc. Natl. Acad. Sci. USA.*, **96** (22), 12650-12654, 1999.

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.

Uncini A., Galluzzi G., Di Muzio A., De Angelis M.V., Ricci E., Scoppetta C. and Servidei S. Facioscapulohumeral muscular dystrophy presenting isolated monomelic lower limb atrophy. Report of two patients with and without 4q35 rearrangement. *Neuromuscul. Disord.*, **12**, 874-877, 2002.

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.

Utz U., Biddison W.E., McFarland H.F., McFarlin D.E., Flerlage M. and Martin R. Skewed T-cell receptor repertoire in genetically identical twins correlates with multiple sclerosis. *Nature*, **364**, 243-247, 1993.

Vainzof M., Moreira E.S., Suzuki O.T., Faulker G., Valle G., Beggs A.H., Carpen O., Ribeiro A.F., Zanoteli E., Gurgel-Gianneti J., Tsanaclis A.M., Silva H.C.A., Passos-Bueno M.R. and Zatz M. Telethonin protein expression in Neuromuscular disorders. *Biochim. Biophys. Acta*, **1588**, 33-40, 2002.

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 Merwe A. Genetic heterogeneity in South African facioscapulohumeral muscular dystrophy (FSHD) families. M.Sc. thesis, University of Pretoria, South Africa, 2002.

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 Geel M., Dickson M.C., Beck A.F., Bolland D.J., Frants R.R., Van der Maarel S.M., De Jong P.J. and Hewitt J.E. Genomic analysis of human chromosome 10q and 4q telomeres suggests a common origin. *Genomics*, **79** (2), 210-217, 2002.

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.

Van Overveld P.G.M., Lemmers R.J.F.L., Sandkuijl L.A., Enthoven L., Winokur S.T., Bakels F., Padberg G.W., Van Ommen G.J.B., Frants R.R. and Van der Maarel S.M., Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. *Nat. Genet.*, **35**, 315-317, 2003.

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

Wagner K.R. Genetic diseases of muscle. *Neurol. Clin. N. Am.*, **20**, 645-678, 2002.

Wallace D.C. Mitochondrial DNA sequence variation in human evolution and disease. *Proc. Natl. Acad. Sci. USA.*, **91**, 8739-8746, 1994.

Wallace D.C. Mitochondrial DNA variation in human evolution, degenerative disease and aging. *Am. J. Hum. Genet.*, **57**, 201.223, 1995.

Wallace D.C., Brown M.D. and Lott M.T. Mitochondrial DNA variation in human evolution and disease. *Gene*, **238**, 211-230, 1999.

Wallace D.C. Center for Molecular and Mitochondrial Medicine and Genetics, University of California, Irvine, U.S.A. Personal communication, 2004.

Walter M.C., Lochmüller H., Reilich P., Klopstock T., Huber R., Hartard M., Hennig M., Pongratz D., Müller-Felber W. Creatine monohydrate in muscular dystrophies: A double-blind, placebo-controlled clinical study. *Neurology*, **54**, 1848-1850, 2000.

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.

Wang N., Wu Z.Y., Wang C.D., Wang Z.Q., Lin M.T., Fang L. and Murong S.X. Mechanism of translocation between chromosomes 4q and 10q in facioscapulohumeral muscular dystrophy. *Zhonghua Yi Xue Za Zhi*, **83** (8), 650-653, 2003. (Article in Chinese)

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.

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. *Neuromuscul. Disord.*, **3**, 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.

Wilson S.B., Kent S.C., Patton K.T., Orban T., Jackson R.A., Exley M., Porcelli S., Schatz D.A., Atkinson M.A., Balk S.P., Strominger J.L. and Hafler D.A. Extreme Th1 bias of invariant Vα24JαQT cells in type 1 diabetes. *Nature*, **391**, 177-180, 1998.

Wilson S.B., Kent S.C., Horton H.F., Hill A.A., Bollyky P.L., Hafler D.A., Strominger J.L and Byrne M.C. Multiple differences in gene expression in regulatory Vα24JαQT cells from identical twins discordant for type I diabetes. *Proc. Natl. Acad. Sci. USA.*, **97** (13), 7411-7416, 2000.

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.

Winokur S.T., Barrett K., Martin J.H., Forrester J.R., Simon M., Tawil R., Chung S.A., Masny P.S. and Figlewicz D.A. Facioscapulohumeral muscular dystrophy (FSHD) myoblasts demonstrate increased susceptibility to oxidative stress. *Neuromuscul. Disord.*, **13**, 322-333, 2003a.

Winokur S.T., Chen Y-W., Masny P.S., Martin J.H., Ehmsen J.T., Tapscott S.J., Van der Maarel S.M., Hayashi Y. and Flanigan K.M. Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation. *Hum. Mol. Genet.*, **12** (22), 2895-2907, 2003b.

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.

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.

Zatz M., Vainzof M. and Passos-Bueno M.R. Limb-girdle muscular dystrophy: one gene with different phenotypes, one phenotype with different genes. *Curr. Opin. Neurol.*, **13**, 511-517, 2000.

Zatz M., De Paula F., Starling A. and Vainzof M. The 10 autosomal recessive limb-girdle muscular dystrophies. *Neuromuscul. Disord.*, **13**, 532-544, 2003.

8.2 ELECTRONIC REFERENCES

BioEdit, Biological Sequence Alignment Editor, http://www.mbio.ncsu.edu/BioEdit/bioedit/html, 2003.

Maca-Meyer N., González A.M., Larruga J.M., Flores C. and Cabrera V.M. Major genomic mitochondrial lineages delineate early human expansions. BMC genetics, **2,** http://www.biomedcentral.com/1471-2156/2/13, 2001.

NCBI, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, (Bethesda, MD), http://www.ncbi.nlm.nih.gov, 2003.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 310200 (DMD): date last edited 2/12/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003a.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 300377 (Dystrophin): date last edited 9/12/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003b.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 300376 (BMD): date last edited 2/4/2002. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003c.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 310300 (EDMD X-linked): date last edited 2/4/2002. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003d.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 181350 (EDMD-AD): date last edited 8/1/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003e.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 300384 (Emerin): date last edited 19/5/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003f.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 150330 (LMNA): date last edited 11/11/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003g.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 159000 (LGMD1A): date last edited 11/6/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003h.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 159001 (LGMD1B): date last edited 28/7/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003i.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 607801 (LGMD1C): date last edited 6/6/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003j.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 601253 (CAV3): date last edited 6/6/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003k.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 603511 (LGMD1D): date last edited 28/7/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003l.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 253600 (LGMD2A): date last edited 23/9/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003m.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 114240 (CAPN3): date last edited 20/10/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003n.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 253601 (LGMD2B): date last edited 11/11/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003o.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 603009 (DYSF): date last edited 9/9/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003p.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 253700 (LGMD2C): date last edited 27/2/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003q.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 608099 (LGMD2D): date last edited 11/10/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003r.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 600119 (SGCA): date last edited 3/10/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003s.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 604286 (LGMD2E): date last edited 23/9/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003t.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 600900 (SGCB): date last edited 10/11/2003 World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003u.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 601287 (LGMD2F): date last edited 23/9/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003v.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 601411 (SGCD): date last edited 19/11/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003w.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 601954 (LGMD2G): date last edited 1/2/2000. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003x.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 254110 (LGMD2H): date last edited 27/12/2002. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003y.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 607155 (LGMD2I): date last edited 30/7/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003z.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 606596 (FKRP): date last edited 10/11/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003aa.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 600334 (TMD): date last edited 29/12/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003bb.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 254130 (MM): date last edited 12/8/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003cc.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 164300 (OPMD): date last edited 9/8/2002. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003dd.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 602279 (PABP2): date last edited 3/2/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003ee.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 156225 (LAMA2): date last edited 10/6/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003ff.

OMIM™, Online Mendelian Inheritance in Man™, Johns Hopkins University, Baltimore, MD. MIM Number: 158900 (FSHD): date last edited 2/12/2003. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim, 2003gg.

APPENDIX A

TERMS PREVIOUSLY USED TO DESCRIBE FACIOSCAPULOHUMERAL MUSCULAR DYSTROPHY

In the past several terms were used to describe FSHD. A list of these terms is presented in this appendix.

Table A.1: Names under which FSHD was described from 1848 to 1996

	Name	Year	Described
1.	Progressive muscular atrophy with fatty degeneration, *(Atrophie musculaire avec transformation graisseuse)*	1848	Duchenne
2.	Progressive hereditary muscle atrophy with fatty degeneration beginning with a face in adults or in youth or in children, *(Atrophie musculaire progressive avec transformation graisseuse débutant de la face des adultes où adolescence où enfance, héréditaire)*	1855	Duchenne
3.	Progressive hereditary fatty muscular atrophy with the affection of some muscles of face, trunk and limbs in adults, children and congenital period, *(Atrophie musculaire graisseuse progressive siégeant dans quelques muscles de la face, du tronc et des membres des adultes, enfants et congénitale, héréditaire)*	1861 1862	Duchenne
4.	Progressive fatty muscular atrophy of childhood, *(Atrophie musculaire graisseuse progressive de l'enfance)*	1868	Duchenne
5.	Progressive hereditary muscular atrophy of childhood (or youth and adult), (Atrophie musculaire progressive de l'enfance (où de l'adolescence et de l'adulte), hétéditaire	1872	Duchenne
6.	Progressive muscular atrophy of childhood *(Atrophie musculaire progressive de l'enfance)*	1874	Landouzy
7.	Hereditary myopathic form of protopathic progressive muscular atrophy (including Duchenne childhood form, in Roth's opinion)	1876 1880	Roth
8.	Progressive muscular atrophy, myopathic form (involving facial muscles) including Duchenne infantile form, in Lichtheim's opinion, *(Progressive Muskelatrophie, myopathische Form (mit Betheiligung der Gesichtsmuskulatur)*	1878	Lichtheim
9.	Progressive atrophic myopathy the same as Duchenne progressive muscular atrophy of childhood, in the opinion of Landouzy and Dejerine, *(Myopathie atrophique progressive)*	1884	Landouzy and Dejerine

continued ...

Table A.1: continued ...

	Name	Year	Described
10.	Juvenile form of progressive muscular atrophy involving facial muscles the same as Duchenne progressive hereditary muscular atrophy of childhood, in Remak's opinion, *(Juveniler progressiver Muskelatrophie mit Betheiligung der Gesichtsmuskulatur)*	1884	Remak
11.	Basic (central) myopathic form of the protopathic progressive muscular atrophy (including Duchenne hereditary and sporadic childhood form, in Roth's opinion)	1884	Roth
12.	The same as the facio-scapulo-humeral type of Landouzy and Dejerine, in Charcot's opinion, *(Atrophie musculaire progressive de l'enfance de Duchenne (de Boulogne)*	1885	Charcot
13.	Hereditary atrophy of childhood, *(Atrophie infantilis hereditaria de Duchenne)*	1885	Marie and Guinon
14.	Progressive muscle atrophy *(Myopathie atrophique progressive)*: 1. Progressive muscular atrophy of childhood of Duchenne, *(Atrophie musculaire progressive de l'enfance (de Duchenne)* 2. Facio-scapulo-humeral type *(de Landouzy et Dejerine)*	1885	Landouzy and Dejerine
15.	Juvenile form of progressive muscular dystrophy the same as Duchenne progressive muscular atrophy of childhood, the facio-scapulo-humeral type of Landouzy-Dejerine, in Erb's opinion, *(Juvenilen Form der Dystrophia muscularis progressive)*	1886	Erb
16.	Myopathic form of progressive muscular atrophy involving facial muscles the same as Duchenne progressive muscular atrophy of childhood, Landouzy-Dejerine facioscapulohumeral type, in Krecke's opinion, *(Myopathische Form der progressiven Muskelatrophie mit Betheiligung der Gesichtsmuskeln)*	1886	Krecke,
17.	Progressive muscle atrophy, facio-scapulo-humeral type of Landouzy and Dejerine, *(Myopathie atrophique progressive, type facio-scapulo-humeral de Landouzy et Dejerine)*	1886	Ladame
18.	Progressive muscle atrophy (the same as the progressive muscle atrophy in childhood of Duchenne in the opinion of Landouzy and Dejerine) 1. type facio-scapulo-humeral 2. type scapulo-humeral	1886	Landouzy and Dejerine
19.	Juvenile form of progressive muscular atrophy involving facial muscles the same as Duchenne progressive muscular atrophy of childhood, facio-scapulo-humeral type of Landouzy-Dejerine, in Bernhard's opinion *(Juveniler progressiver Muskelatrophie mit Betheiligung der Gesichtsmuskulatur)*	1887	Bernhardt
20.	Basic (central) form of the muscular tabes (including Duchenne hereditary childhood form the same as the facio-scapulo-humeral type of Landouzy and Dejerine, in Roth's opinion)	1887	Roth

continued ...

Table A.1: continued ...

	Name	Year	Described
21.	Facio-scapulo-humeral type of Landouzy and Dejerine (the same as the infantile progressive muscular atrophy of Duchenne, in Sach's opinion)	1888	Sachs
22.	Simple idiopathic muscular atrophy, involving the face and scapulohumeral muscles (the same as the facioscapulohumeral type Duchenne and Landouzy and Dejerine, in Osler's opinion)	1889	Osler
23.	Progressive muscular dystrophy (facio-scapulo-humeral type), the same as Duchenne progressive muscular atrophy of childhood , in Sperling's opinion *(Dystrophia muscularis progressive [facio-scapulo-humeral Typus], Duchenne progressive muscular atrophie de l'enfance)*	1889	Sperling
24.	Muscular atrophy of the youth Duchenne or Landouzy-Dejerine (facio-scapulo-humeral type) *(Atrophiae muscularis juvenilis Duchenne or Landouzy-Dejerine)*	1889	Rybalkin
25.	Progressive muscular dystrophy (facio-scapulo-humeral type), the same as the facioscapulohumeral type of Duchenne and Landouzy-Dejerine, in Bielschowski's opinion *(Dystrophia muscularis progressive [Typus facio-scapulo-humeralis])*	1890	Bielschowski
26.	I. Infantile progressive muscular dystrophy, B. atrophic form: I. with primary facial involvement (Duchenne's infantile form) II. Progressive muscular dystrophy, juvenile and adult (Juvenile form)	1891	Erb
27.	Simple idiopathic muscular atrophy involving the face and shoulder girdle (the same as Landouzy and Dejerine facioscapulohumeral type, in Gowers' opinion)	1891	Gowers
28.	Progressive myopathy of Landouzy-Dejerine, the same as the hereditary infantile form of Duchenne, the facio-scapulo-humeral type of Landouzy and Dejerine in opinion of Blocq and Marinesco	1893	Blocq and Marinesco
29.	Muscular atrophy of Erb form, the same as Duchenne infantile type of muscular atrophy, Landouzy-Dejerine, in Romberg's opinion	1893	Romberg
30.	Descending (basic) form of the muscular tabes in children and adults (the same as Duchenne hereditary progressive muscular atrophy of childhood, the facioscapulohumeral type of Landouzy and Dejerine, in Roth's opinion)	1895	Roth
31.	Muscle atrophy (facio-scapulo-humeral type) *(Myopathie atrophiques [du type facio-scapulo-humeral])*	1909	Landouzy and Lortat-Jacob
32.	Progressive muscle atrophy in childhood, the same as the facio-scapulo-humeral type (Duchenne-Landouzy-Dejerine) in the opinion of Eulenburg and Cohn *(Infantile progressive Muskeldystrophie, forme facioscapulohumeral muscular dystrophy facio-humero-scapulare)*	1911	Eulenburg and Cohn
33.	Infantile atrophic form of progressive muscular dystrophy without or with affected facial muscles (Duchenne, Landouzy and Dejerine) in Strumpell's opinion	1911	Strumpell
34.	Facio-scapulo-humeral type of Landouzy-Dejerine (Duchenne type in terminology of Charcot)	1925	Davidenkov

continued ...

Table A.1: continued ...

	Name	Year	Described
35.	I. Facio-scapulo-humeral type myopathy of Landouzy and Dejerine *(myopathie type facio-scapulo-humeral – Landouzy et Dejerine)* – II. Progressive muscular atrophy of childhood of Duchenne de Boulogne *(atrophie musculaire progressive de l'enfance – Duchenne de Boulogne)* III. Facio-scapulo-humeral type muscular atrophy corresponding to hereditary muscular atrophy of Duchenne (de Boulogne), *(Le type facio-scapulo-humeral de la myopathie atrophique progressive correspond a l'atrophie musculaire héréditaire de Duchenne (de Boulogne)* in Dejerin's opinion)	1926	Dejerine
36.	Landouzy-Dejerine myopathy, facio-scapulo-humeral type (the same as Duchenne progressive muscular atrophy of childhood)	1930 1932	Davidenkov
37.	Autosomal dominant cases beginning with facial affection (including facio-scapulo-humeral muscular dystrophy, in Bell's opinion)	1942	Bell
38.	Facio-scapulo-humeral type (Landouzy and Dejerine) progressive muscular dystrophy	1950	Tyler and Wintrobe
39.	Facioscapulohumeral type of progressive muscular dystrophy	1950	Tyler and Stephens
40.	Facioscapulohumeral type (Landouzy and Dejerine) juvenile or infantile with or without involvement of lower extremities	1951	Levison
41.	Autosomal recessive limb-girdle type (with or without affected face)	1953	Stevenson
42.	Autosomal dominant descending shoulder-girdle form (involving face or shoulder girdle) of muscular dystrophy	1952 1953 1957	Becker
43.	Mild restricted form of muscular dystrophy with primary involvement of shoulder girdle and often the face muscles (including Duchenne progressive muscular atrophy of childhood the same as the facio-scapulo-humeral type of Landouzy and Dejerine, in opinion of Adams, Denny-Brown and Pearson)	1953 1962	Adams, Denny-Brown & Pearson
44.	Facioscapulohumeral type of Landouzy and Dejerine	1954 1955 1956 1963 1964	Walton and Nattrass, Walton
45.	Shoulder-Girdle form of Erb muscular dystrophy	1957	Seitz
46.	Benign descending form of muscular dystrophy	1958	Brugsch and Brockmann-Rohne
47.	Autosomal dominant facioscapulohumeral type (with incipient affection of the face or shoulder girdle)	1959	Morton and Chung
48.	Autosomal dominant descending shoulder-girdle form of muscular dystrophy (the same as the facioscapulohumeral type of Duchenne (1868), Erb (1884), Landouzy and Dejerine (1885), in Becker's opinion	1964	Becker

continued ...

Table A.1: continued ...

	Name	Year	Described
49.	Facioscapulohumeral form (Erb)	1967	Weingarten
50.	Disease of Landouzy and Dejerine *(La maladie de Landouzy et Dejerine)*	1968	Schmitt and Duc
51.	Scapulohumeral variant (Erb) of Type I (Landouzy and Dejerine) muscular dystrophy	1968	Zellweger and McCormick
52.	Facio-scapulo-humeral (Landouzy and Dejerine) (Becker descending type), autosomal dominant	1968 1969 1974	Research group on neuromuscular disorders Walton Walton & Gardner-Medwin
53.	Autosomal dominant facio-scapulo-limb muscular dystrophy (Duchenne, Landouzy and Dejerine): 1. Descending variety with a jump from the face, shoulder girdle to the peroneal group of shin muscles (frequently, a mild course) 2. Gradually descending variety (rarely, a less mild course)	1971 1974-1976 1972	Kazakov Becker
54.	Facio-scapulo-humeral type (Erb-Landouzy-Dejerine)	1982	Sorrel-Dejerine and Fardeau
55.	Progressive muscle atrophy of Landouzy and Dejerine *(Myopathie atrophique progressive de Landouzy et Dejerine)*	1982	Sorrel-Dejerine and Fardeau
56.	Facioscapulohumeral disease	1982	Padberg
57.	Congenital facioscapulohumeral dystrophy (the same as Duchenne progressive fatty muscular atrophy of childhood, in Furukawa's opinion)	1984	Furukawa
58.	Landouzy-Dejerine disease	1985	Rossi
59.	Facioscapulohumeral (Landouzy and Dejerine), autosomal dominant involving face, scapulohumeral and anterior tibial muscles	1988 1988	Walton Research group on neuromuscular disorders
60.	Facioscapulolimb muscular dystrophy (autosomal dominant, gradual descending) Facioscapuloperoneal muscular dystrophy [autosomal dominant, descending with a jump from the face, shoulder girdle to the peroneal group (anterior tibial) muscles]	1992 1995	Kazakov
61.	Facioscapulolimb muscular dystrophy type 1 (FSLD1): a gradually descending with initial facioscapulohumeral (FSH) phenotype, autosomal dominant muscles (Duchenne de Boulogne) Facioscapulolimb muscular dystrophy type 2 (FSLD2): a descending with a jump with initial facioscapuloperoneal (FSP) phenotype, autosomal dominant (Erb, Landouzy and Dejerine)	1996	Kazakov

Adapted from Kazakov (1995); Kazakov (1998).

APPENDIX B

SYNOPSIS OF MOLECULAR INFORMATION FOR SELECTED MUSCULAR AND NEUROMUSCULAR DISORDERS

This appendix contains information on the mode of inheritance, gene loci, gene symbols, and gene products for the muscular dystrophies and other selected neuromuscular disorders.

Table B.1: Mode of inheritance, gene loci, gene symbols and gene products of different types of neuromuscular disorders

Disorder	Inheritance	Gene locus	Gene symbol	Gene product
Muscular dystrophies:				
Duchenne/Becker	XR	Xp21.2	DMD	*Dystrophin*
Emery Dreifuss, X-linked	XR	Xq28	EMD	*Emerin*
Emery Dreifuss, autosomal dominant	AD	1q11-q23	EDMD-AD	*LaminA/C* (LMNA)
Facioscapulohumeral	AD	4q35	FSHD	---
Limb-girdle muscular dystrophies (LGMD)				
LGMD type 1A	AD	5q22-34	LGMD1A	*Myotilin*
LGMD type 1B	AD	1q11-21	LGMD1B	*LaminA/C* (LMNA)
LGMD type 1C	AD	3p25	LGMD1C	*Caveolin-3* (CAV3)
LGMD type 1D	AD	6q23	LGMD1D	---
LGMD type 1E	AD	7q	LGMD1E	---
LGMD type 2A	AR	15q15.1	LGMD2A	*Calpain-3* (CAPN3)
LGMD type 2B	AR	2p13	LGMD2B	*Dysferlin*
LGMD type 2C	AR	13q12	LGMD2C	*γ-sarcoglycan* (SGCG)
LGMD type 2D	AR	17q12-q21.33	LGMD2D	*α-sarcoglycan* (SGCA)
LGMD type 2E	AR	4q12	LGMD2E	*β-sarcoglycan* (SGCB)
LGMD type 2F	AR	5q33-q34	LGMD2F	*δ-sarcoglycan* (SGCD)
LGMD type 2H	AR	9q31-q34	LGMD2H	Tripartite-motif-containing protein 32 (TRIM32)

continued ...

Table B.1: continued ...

Disorder	Inheritance	Gene locus	Gene symbol	Gene product
LGMD type 2G	AR	17q11-q12	LGMD2G	*Telethonin*
LGMD type 2I	AR	19q13.3	LGMD2I	*Fukutin related protein* (FKRP)
LGMD type 2J	AR	2q31	LGMD2J	*Titin*
Distal myopathy:				
Autosomal dominant distal myopathy	AD	14q11.2	MPD1	---
Distal myopathy with rimmed vacuoles	AR	9p12	DMRV (IBM2)	---
Hereditary inclusion body myopathy	AR	9p12	HIBM (IBM2)	acetylglucosamine-2-epimerase *(GNE)*
Distal recessive myopathy (Miyoshi myopathy)	AR	2p12-14	MM	*Dysferlin*
Tibial muscular dystrophy	AD	2q31	TMD	*Titin* (TNN)
Other myopathies:				
Autosomal dominant myopathy with proximal muscle weakness and early respiratory muscle involvement (Edström) (type 1)	AD	2q24-31	MPRM1	---
Autosomal dominant myopathy with proximal muscle weakness and early respiratory muscle involvement (type 2)	AD	2q21	MPRM2	---
Bethlem myopathy (type 1 and 2)	AD	21q22.3	COL6A1-COL6A2	*Collagen type VI subunit α1 or α2*
Bethlem myopathy (type 3)	AD	2q37	COL6A3	*Collagen type VI subunit α3*
Desmin-related myopathy	AD	11q22	DRM	*αB-crystallin* (CRYAB)
Desmin-related myopathy	AD / AR	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	*Poly (A) binding protein 2* (PABP2)
Myofibrillar Myopathy with arrhythmogenic right ventricular cardiomyopathy	AD	10q22	MFM / ARVC	---
Myopathy with joint contractures, ophthalmoplegia and rimmed vacuoles (Inclusion body myopathy 3)	AD	17p13	IBM3	*Myosin heavy chain IIa* (MYHC2A)
HyperCKemia	AD	3p25	CAV3	*Calveolin-3*
Congenital muscular dystrophies (CMD):				
Merosin deficiency CMD	AR	6q2	MDC1A	*Laminin α2 chain of merosin* (LAMA2)

continued ...

Table B.1: continued ...

Disorder	Inheritance	Gene locus	Gene symbol	Gene product
Fukuyama CMD	AR	9q31-q33	FCMD	*Fukutin*
Integrin α7 deficiency	AR	12q	ITGA7	*Integrin α7*
Muscle-eye-brain disease	AR	1p3	MEB	*O-mannoseβ-1,2-N-acetylglucosaminyl transferase* (POMGnT1)
Rigid spine syndrome	AR	1p3	RSMD-1	*Selenoprotein N1* (SEPN1)
CMD with secondary merosin deficiency 1	AR	1q42	MDC1B	---
CMD with secondary merosin deficiency 2	AR	19q1	MDC1C	*Fukutin related protein* (FKRP)
Ullrich syndrome (type 2)	AR	21q2	UCMD	Collagen type VI subunit α2 (COL6A2)
Ullrich syndrome (type 3)	AR	21q3	UCMD	Collagen type VI subunit α3 (COL6A3)
Congenital myopathies:				
Central core disease, autosomal dominant	AD	19q13.1	CCD	*Ryanodine receptor* (RYR1)
Central core disease, autosomal recessive	AR	19q13.1	CCD	*Ryanodine receptor* (RYR1)
Myotubular myopathy	XR	Xq28	MTM1	*Myotubularin*
Nemaline myopathy (type 1)	AD	1q21-q23	NEM1	*α tropomyosin 3* (TPM3)
Nemaline myopathy (type 2)	AR	2q21.1-q22	NEM2	*Nebulin*
Nemaline myopathy (type 3, autosomal dominant)	AD	1q42.1	NEM3	*α-actin skeletal* (ACTA1)
Nemaline myopathy (type 3 autosomal recessive)	AR	1q21-q23	TPM3	*α tropomyosin 3*
Nemaline myopathy (type 4)	AD	9p13	NEM4	*β-tropomyosin* (TPM2)
Nemaline myopathy (type 5)	AR	19q13	NEM5	*Troponin T1* (TNNT1)
Multiminicore disease	AR	1p36	SEPN1	*Selenoprotein N1*
Myotonic syndromes:				
Brody disease	AR	16p12	SERCA1	*Sarcoplasmic reticulum Ca^{2+} ATPase*
Myotonic dystrophy (Steinert)	AD	19q13	DM	*Myotonin-protein kinase*
Myotonic dystrophy type 2	AD	3q21	DM2 PROMM	Zinc finger protein 9 (*ZNF9*)
Myotonia, dominant (Thomsen)	AD	7q35	CLC-1	*Muscle chloride channel*
Myotonia, recessive (Becker)	AR	7q35	CLC-1	*Muscle chloride channel*
Rippling muscle disease (type 1)	AD	1q41	RMD1	---
Rippling muscle disease (type 2)	AD	3p25	CAV3	*Caveolin-3*
Schwartz-Jampel syndrome	AR	1p34-p36.1	SJS	*Perlecan*

XR = X-linked recessive; AR = autosomal recessive; AD = autosomal dominant, --- = no information available. Adapted from Neuromuscular disorders: gene location, (2003).

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. (Padberg *et al.*, 1997).

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. (Padberg *et al.*, 1997).

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 influenced 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 an alternative diagnosis.
1.9	**Hearing loss**
C	Hearing loss is part of the disorder; it starts with high tone perceptive deafness and may 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 years 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 the age of 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.

Adapted from Padberg *et al.* (1997).

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 ...

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 tcatttttt 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	acccctttccc 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 ctggggatct
3901	atacagcact catcaaatct agcatttaaa gactggctca gtaaaggggg ataccgacag
3961	caatagtcca ggctatagat tctaggtgct tcacaaacac attcccaaat atattttctc
4021	tggacttcgc tgtgtttcca agttaaagag aatttttttct 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	cattttctca 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 ccccgggggа gctggaggcc tcggaagagg ccgcctcgct
5461	ggaagcaccc ctcagcgagg aagaataccg ggctctgctg gaggagcttt aggacgcggg
5521	gttgggacgg ggtcgggtgg ttcggggcag ggcggtggcc tctctttcgc ggggaacacc
5581	tggctggcta cggaggggcg tgtctccgcc ccgccccctc 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 ggacccccgc 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 ccccccacc ccccccccc accaccacca
7261	ccaccacccc gccggccggc cccaggcctc gacgccctgg ggtcccttcc ggggtggggc
7321	gggctgtccc agggggctc accgccattc atgaaggggt ggagcctgcc tgcctgtggg
7381	cctgtacaag 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 cacaggggcc 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 gccggggtcc ccaggtcgcc ggggcggcgt gggaacccca agccggggca
8521	gctccacctc cccagcccgc gccccccggac 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 accgcccccc ccccaccccc caccccccac
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 cctcccccct ctcccccttc ctcttcgtct ctccggcccc
10501	accaccacca ccgccaccac gccctccccc cccccccccc cccccacca ccaccaccac

continued ...

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 cgaagggcca ggcacccggg
11281	acagggtggc agggcgcccg cgcaggcagg cggcctgtgc agcgcggccc ccggcggggg
11341	tcaccctgct ccctcgtggg tcgccttcgc ccacaccggc gcgtggggaa cggggcttcc
11401	cgcaccccac gtgccctgcg cgcctggggc tctcccacag ggggctttcg tgagccaggc
11461	agcgagggcc gcccccgcgc 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 cccgcccccc 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 gccccctccc cacccccccac cccccacccc cggaaaacgc gtcgtccccct

continued ...

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.

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 cgcccctcc accgggctga
	gggaacgcct ggctggctac ggaggggcgt gtctccgccc cgcccctcc 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 ...

Table E.1: continued ...

Chr 4:	ggcagagctc tcctgcctct ccaccagccc accccgccgc ctgaccgccc cctccccacc
Chr 10:	ggcagagctc tcctgcctct ccaccagccc accccgccgc ctgaccgccc cctccccacc
	ccccaccccc caccccccgga 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
	ggga*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 ccggggccca 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 ...

Table E.1: continued ...

Chr 4: Chr 10:	ccctgctgcc gacgcgtggg agcccagaga gcggcttccc gttcccgcgg gatccctgga ccctgctgcc gacgcgtggg agcccagaga gcggcttccc gttcccgcgg gatccctgga gaggtccgga gagccggccc ccgaaacgcg ccccccctccc ccctcccccc tctccccctt gaggtccgga gagccggccc ccgaaacgcg ccccccctccc ccctcccccc tctccccctt cctcttcgtc tctccggccc caccaccacc accgccacca cgccctcccc ccccaccccc cctcttcgtc tctccggccc caccaccacc accgccacca cgccctcccc caccaccccc ccc*cc**** ********** ccaccaccac caccaccacc ccgccggccg gccccaggcc cccaccccac accaccacca ccaccaccac caccaccacc ccgccggccg gccccaggcc tcgacgccct ggggtccctt 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: Chr 10:	tcgcctacgc cgccccggct cctccggacg gggcgctctc ccaccctcag gctcctcggt 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

SIGNAL INTENSITY CALCULATIONS

The signal intensities of the chromosome 4 and chromosome 10 fragments were determined as described in paragraphs 5.33 and 5.3.4. The intensity for each fragment was corrected for the background around that particular fragment to obtain a normalised value. For each individual the contribution of chromosome 4 specific signal to the sum of the chromosomes 4 (S4) and 10 (S10) signals was calculated via Equation 5.1 utilising this normalised value. The signal intensity, normalised and contribution values for the fragments of interest for the individuals investigated in this study are presented in Table F.1.

Table F.1: Signal intensity calculations

DNA No.	Chromosome 4			Chromosome 10			Contribution of chr 4
	Signal intensity	Background	Normalised signal intensity	Signal intensity	Background	Normalised signal intensity	
806	2.923	3.090	-0.167	3.808	2.861	0.946	0.00
807	1.610	1.397	0.213	1.839	1.235	0.604	0.26
808	1.695	1.366	0.330	1.959	1.121	0.838	0.28
809	2.111	1.715	0.396	2.544	1.360	1.183	0.25
810	4.204	3.182	1.022	5.240	2.556	2.685	0.27
811	0.896	0.686	0.210	1.176	0.565	0.611	0.25
812	2.811	2.100	0.711	2.969	2.077	0.892	0.44
814	3.571	2.533	1.037	3.799	2.524	1.275	0.44
815	1.368	1.349	0.019	4.081	2.370	1.711	0.01
816	1.298	0.306	0.991	1.964	0.879	1.085	0.51
817	1.310	0.896	0.415	1.017	0.720	0.297	0.58
818	9.435	5.457	3.978	9.786	8.531	1.255	0.78
819	2.424	2.269	0.155	2.532	2.088	0.444	0.28
820	1.305	0.941	0.364	1.125	0.810	0.315	0.53
821	2.202	1.758	0.444	3.144	1.741	1.403	0.26
822	1.583	0.353	1.230	1.819	1.423	0.396	0.75
823	1.506	1.328	0.178	1.466	0.940	0.526	0.27
824	5.532	4.612	0.920	5.520	3.032	2.488	0.27
826	5.088	5.080	0.008	5.014	2.742	2.272	0.00
827	1.311	0.635	0.676	2.370	0.296	2.074	0.24
828	1.075	0.046	1.029	1.160	0.002	1.157	0.47

continued ...

Table F.1: continued ...

DNA No.	Chromosome 4			Chromosome 10			Contribution of chr 4
	Signal intensity	Background	Normalised signal intensity	Signal intensity	Background	Normalised signal intensity	
829	6.387	5.358	1.029	6.673	3.627	3.046	0.25
830	9.291	5.964	3.327	10.315	7.196	3.118	0.51
831	11.484	6.347	5.137	12.430	7.347	5.083	0.49
832	3.607	2.756	0.851	4.749	2.288	2.461	0.27
833	2.376	2.625	-0.248	4.399	1.699	2.699	-0.11
834	2.587	2.244	0.343	2.521	2.047	0.474	0.45
835	2.471	2.738	-0.267	4.329	1.616	2.713	-0.12
837	5.249	1.150	4.100	8.461	1.566	6.895	0.36
838	1.948	1.338	0.611	1.457	0.872	0.585	0.52
839	2.968	1.290	1.678	3.197	1.242	1.955	0.45
840	1.632	1.069	0.563	1.549	1.328	0.221	0.75
841	9.388	2.589	6.798	14.381	7.678	6.702	0.48
842	3.699	2.204	1.495	3.383	1.594	1.789	0.47
843	3.857	2.088	1.769	3.368	1.381	1.986	0.48
844	4.860	2.392	2.469	2.222	1.345	0.877	0.75
845	5.793	2.718	3.075	3.694	2.113	1.581	0.67
847	3.421	1.842	1.579	2.781	1.337	1.444	0.54
848	3.875	1.764	2.111	2.117	1.381	7.364	0.75
849	5.574	2.256	3.318	3.476	2.159	1.317	0.73
851	1.626	1.167	4.593	1.321	1.175	1.465	0.78
852	5.139	2.486	2.653	2.896	2.095	0.801	0.78
853	6.656	2.561	4.095	3.615	2.272	1.343	0.76
854	10.581	1.357	9.224	4.986	1.393	3.593	0.71
855	10.380	7.948	2.432	14.514	7.649	6.865	0.28
856	4.306	2.141	2.165	5.374	3.212	2.162	0.52
883	2.510	0.911	1.599	5.068	0.720	4.348	0.26
885	2.753	1.292	1.461	3.404	1.676	1.728	0.48
887	2.197	1.241	0.956	2.790	1.599	1.191	0.46
888	4.148	1.843	2.305	4.917	2.489	2.427	0.51
894	0.884	0.855	0.030	6.431	2.415	4.016	0.01
902	3.828	3.848	-0.020	8.315	3.885	4.430	-0.01
904	3.228	3.248	-0.020	7.046	2.720	4.326	-0.01
918	5.669	1.544	4.125	5.318	0.328	4.991	0.44
920	3.205	3.208	-0.003	7.800	3.907	3.893	-0.00
921	3.670	1.144	2.526	4.345	1.506	2.839	0.46
922	1.950	0.190	1.760	1.754	0.348	1.406	0.55
926	3.053	3.061	-0.008	8.078	3.911	4.167	-0.00
929	0.517	0.054	0.463	1.384	0.004	1.380	0.26
907	1.109	0.613	0.496	1.447	0.971	0.476	0.53
908	3.242	0.888	2.354	1.457	0.326	1.131	0.66
911	2.865	0.632	2.233	2.590	0.842	1.748	0.57
912	4.638	1.224	3.413	3.938	0.136	3.802	0.46
930	3.713	3.148	0.565	5.528	3.864	1.664	0.25
932	1.699	1.683	0.015	4.646	2.318	2.328	0.01

continued ...

Table F.1: continued ...

DNA No.	Chromosome 4			Chromosome 10			Contribution of chr 4
	Signal intensity	Background	Normalised signal intensity	Signal intensity	Background	Normalised signal intensity	
933	2.524	0.795	1.729	2.326	0.622	1.704	0.51
934	3.800	1.177	2.623	4.512	1.530	2.982	0.46
936	3.026	0.959	2.068	3.640	1.120	2.520	0.44
939	1.263	0.507	0.757	2.466	0.454	2.012	0.28
942	5.080	3.591	1.490	5.731	4.252	1.479	0.51
943	1.274	1.162	0.112	4.961	0.195	4.766	0.02
953	3.409	1.186	2.222	4.529	2.106	2.423	0.47
959	1.241	1.262	-0.021	5.221	2.518	2.702	-0.01
960	4.651	3.143	1.508	5.002	3.660	1.342	0.52
961	1.680	0.480	1.201	1.962	0.817	1.145	0.50
962	1.778	0.984	0.794	2.704	1.784	0.920	0.45
963	4.476	3.031	1.445	7.471	5.858	1.613	0.49
964	2.849	1.683	1.166	3.734	2.355	1.379	0.48
965	5.851	4.862	0.989	10.638	7.781	2.857	0.27
966	2.510	2.147	0.363	2.587	2.083	0.505	0.44
967	6.012	4.354	1.658	8.229	6.552	1.677	0.52
968	14.037	8.052	5.985	15.333	8.849	6.484	0.50
969	4.299	2.282	2.017	5.542	3.102	2.440	0.47
970	3.962	1.659	2.303	5.439	2.862	2.577	0.49
972	4.977	1.922	3.055	5.521	2.853	2.668	0.55
973	5.153	2.942	2.211	5.545	3.139	2.406	0.50
974	2.188	2.187	0.001	2.572	2.039	0.532	0.00
975	2.775	2.726	0.048	6.293	3.651	2.642	0.02
976	1.497	1.145	0.352	1.851	1.729	0.122	0.74
993	2.343	2.108	0.235	2.331	2.036	0.295	0.47
994	3.895	2.504	1.391	4.964	3.288	1.676	0.46
995	2.119	0.592	1.527	1.695	0.463	1.233	0.54
996	0.598	0.194	0.404	0.513	0.096	0.416	0.48
997	0.629	0.621	0.008	3.966	0.108	3.857	0.00
998	6.328	2.091	4.237	8.912	1.118	7.793	0.34
999	3.814	0.887	2.927	5.037	0.584	4.453	0.38
1000	1.974	0.428	1.546	1.210	0.062	1.148	0.56
1005	2.503	1.933	0.569	2.927	2.340	0.587	0.48
1009	9.487	7.621	1.865	11.042	5.337	5.705	0.25
1010	3.600	1.186	2.414	5.026	0.591	4.434	0.34
1012	4.905	1.128	3.777	5.032	0.898	4.134	0.47
1013	1.995	0.435	1.560	3.137	0.472	2.665	0.36
1015	0.160	0.056	0.104	2.907	0.929	1.979	0.05
1017	4.409	2.571	1.838	4.549	2.930	1.619	0.53
1020	0.930	0.301	0.628	2.278	0.691	1.587	0.28
978	7.956	5.001	2.955	11.615	6.490	5.125	0.36
979	6.163	3.990	2.173	7.045	4.679	2.366	0.47
980	1.065	0.395	0.669	1.234	0.504	0.730	0.47
982	1.772	0.633	1.139	1.400	0.248	1.152	0.49

continued ...

Table F.1: continued ...

DNA No.	Chromosome 4			Chromosome 10			Contribution of chr 4
	Signal intensity	Background	Normalised signal intensity	Signal intensity	Background	Normalised signal intensity	
983	1.708	0.443	1.264	2.670	1.268	1.402	0.47
984	4.570	2.168	2.402	5.387	2.863	2.524	0.51
985	4.573	2.418	2.154	6.556	2.826	3.730	0.36
987	2.306	2.309	-0.003	5.292	2.244	3.048	-0.00
988	3.309	1.365	1.944	3.828	1.402	2.426	0.47
989	1.630	1.679	-0.049	2.768	1.703	1.065	-0.05
990	3.939	2.870	1.069	5.243	4.241	1.002	0.52
991	1.411	1.455	-0.044	3.037	1.969	1.068	-0.04
1024	2.656	0.473	2.183	2.690	1.116	1.575	0.57
1025	1.082	0.702	0.380	1.198	0.764	0.434	0.47
1029	1.722	1.813	-0.091	2.686	1.175	1.512	-0.06
1030	2.274	1.241	1.033	4.367	3.249	1.118	0.50
1032	0.974	0.539	0.436	1.045	0.515	0.530	0.45
1034	1.375	1.228	0.147	1.592	1.187	0.404	0.29
1035	1.211	1.207	0.004	1.627	1.333	0.294	0.01
1038	1.291	1.280	0.011	1.584	1.435	0.149	0.07
1087	1.934	1.953	-0.019	9.677	2.690	6.987	-0.00
1089	5.493	2.669	2.824	5.376	2.780	2.596	0.53
1091	1.454	1.139	0.316	2.462	1.340	1.122	0.22
1092	8.793	8.212	0.581	7.943	7.285	0.658	0.50
1093	0.074	0.036	0.038	3.429	0.413	3.016	0.01
1094	1.431	0.511	0.921	2.959	0.544	2.416	0.27
1095	4.170	2.032	2.138	5.447	3.057	2.390	0.47
1096	7.505	3.980	3.525	7.608	4.496	3.112	0.53
1097	3.592	0.224	3.368	1.823	0.666	1.157	0.74
1098	2.773	2.567	0.206	2.656	2.006	0.650	0.26
1100	1.303	0.052	1.251	1.346	0.041	1.305	0.52
1101	9.311	5.145	4.166	9.216	4.777	4.439	0.49
1102	2.738	0.399	2.338	0.046	0.020	0.025	0.99
1104	1.189	1.293	-0.104	2.101	1.012	1.089	-0.10
1106	9.155	7.865	1.290	6.985	6.487	0.498	0.75
1107	4.710	3.396	1.314	6.897	3.550	3.347	0.28
1040	1.574	0.524	1.049	3.596	0.421	3.175	0.24
1042	1.003	0.604	0.399	1.186	0.798	0.388	0.51
1044	1.718	0.031	1.687	2.251	0.345	1.906	0.46
1047	0.497	0.014	0.483	1.709	0.223	1.486	0.24
1050	1.624	1.434	0.190	1.674	1.482	0.192	0.52
1053	7.925	6.678	1.248	7.235	5.885	1.349	0.48
1054	1.555	1.334	0.220	1.683	1.414	0.2695	0.48
1057	2.367	1.330	1.037	3.673	0.886	2.787	0.26
1061	2.625	2.622	0.002	7.541	1.952	5.589	0.00
1064	8.822	7.763	1.059	10.428	7.226	3.203	0.24
1065	2.408	0.732	1.676	2.277	0.515	1.762	0.50
1066	2.361	0.609	1.752	2.185	0.340	1.845	0.49

continued ...

Table F.1: continued ...

DNA No.	Chromosome 4			Chromosome 10			Contribution of chr 4
	Signal intensity	Background	Normalised signal intensity	Signal intensity	Background	Normalised signal intensity	
1067	7.348	6.422	0.926	7.631	6.507	1.124	0.48
1070	2.672	1.210	1.463	4.006	2.258	1.748	0.44
1071	3.818	3.247	0.571	3.798	3.118	0.681	0.48
1072	1.329	1.401	-0.072	3.690	1.687	2.003	-0.04
1074	5.927	3.258	2.669	7.752	4.323	3.429	0.44
1076	2.469	0.496	1.973	3.271	1.277	1.994	0.49
1079	9.976	3.904	6.072	12.023	5.213	6.810	0.46
1085	6.229	1.591	4.638	6.662	2.009	4.653	0.49
1108	10.092	4.722	5.370	6.603	4.945	1.658	0.76
1109	5.584	5.403	0.181	9.932	5.462	4.470	0.04
1113	0.985	0.398	0.588	0.846	0.384	0.462	0.55
1117	2.602	2.433	0.169	2.476	1.808	0.667	0.22
1126	2.199	1.434	0.765	2.490	1.545	0.945	0.45
1128	3.936	1.546	2.390	4.372	1.853	2.519	0.47
1129	1.247	0.540	0.707	1.140	0.537	0.603	0.53
1131	5.443	2.590	2.853	5.173	2.885	2.288	0.56
1132	2.534	2.522	0.012	7.270	3.451	3.819	0.00
1133	8.669	2.754	5.914	10.682	4.206	6.476	0.47
1134	1.872	0.965	0.907	1.740	0.744	0.995	0.46
1137	2.304	0.077	2.226	3.500	0.695	2.805	0.44
1138	2.360	2.326	0.034	6.225	2.977	3.248	0.01
1140	2.959	0.085	2.874	3.149	0.881	2.268	0.55
1141	0.603	0.569	0.034	3.000	0.609	2.391	0.01
1142	8.135	4.088	4.047	10.919	5.978	4.941	0.45
1143	2.275	0.060	2.214	3.741	1.294	2.447	0.47
1144	2.812	0.028	2.784	3.406	0.594	2.812	0.49
1145	2.658	2.368	0.290	2.463	2.155	0.308	0.52
1146	1.166	0.161	1.005	2.332	0.281	2.051	0.32
1147	1.686	0.855	0.830	1.997	1.167	0.830	0.50
1148	2.930	1.860	1.071	3.810	2.598	1.212	0.46
1149	4.482	2.618	1.864	5.205	3.485	1.719	0.52
1151	6.601	1.948	4.653	6.984	2.078	4.906	0.47
1152	0.426	0.430	-0.004	4.522	1.521	3.000	-0.00
1153	5.712	3.883	1.829	4.640	2.844	1.796	0.49
1154	9.869	3.408	6.461	4.149	1.904	2.244	0.73
1155	2.639	2.208	0.431	2.537	2.112	0.425	0.54
1156	2.193	2.199	-0.006	2.513	2.098	0.415	-0.02
1157	7.523	2.235	5.287	6.664	1.572	5.092	0.50
1158	2.371	2.283	0.088	7.065	3.319	3.746	0.02
1159	2.261	1.305	0.956	2.777	1.662	1.115	0.46
1163	3.127	2.001	1.126	2.948	1.667	1.281	0.48
1169	1.727	1.745	-0.018	5.775	1.591	4.184	-0.00
1170	1.564	0.074	1.489	2.055	0.356	1.699	0.47
1171	5.692	2.638	3.054	6.461	2.821	3.640	0.46

continued ...

Table F.1: continued ...

DNA No.	Chromosome 4			Chromosome 10			Contribution of chr 4
	Signal intensity	Background	Normalised signal intensity	Signal intensity	Background	Normalised signal intensity	
1172	7.640	5.450	2.189	15.281	7.728	7.553	0.24
1173	1.039	0.136	0.902	1.186	0.287	0.899	0.50
1174	1.906	1.021	0.885	2.643	2.348	0.295	0.77
1177	2.648	0.782	1.866	3.506	1.291	2.215	0.45
1178	8.421	4.492	3.929	11.810	7.369	4.441	0.49
1179	2.167	1.834	0.333	1.970	1.723	0.247	0.56
1180	1.066	0.577	0.488	1.903	1.422	0.481	0.53
1183	4.303	3.138	1.165	8.124	4.497	3.627	0.26
1184	0.073	0.070	0.004	1.588	0.425	1.164	0.00
1187	2.948	2.988	-0.040	8.546	4.475	4.071	-0.01
1188	1.248	1.238	0.010	3.767	1.689	2.079	0.00
1189	1.145	0.141	1.003	1.174	0.218	0.955	0.51
1190	1.656	0.905	0.751	2.256	1.263	0.992	0.45
1191	3.755	2.873	0.882	6.732	3.945	2.787	0.26
1192	6.019	3.954	2.066	11.732	7.295	4.438	0.34
1193	2.294	0.687	1.607	2.899	0.735	2.164	0.43
1194	1.442	0.433	1.009	2.209	0.915	1.294	0.46
1195	1.432	0.346	1.086	2.255	0.811	1.444	0.45
1196	4.515	2.164	2.351	10.752	3.830	6.922	0.27
1197	1.891	0.883	1.008	2.137	0.999	1.139	0.49
1198	2.704	0.879	1.825	3.536	1.297	2.239	0.46
1199	1.267	0.963	0.304	1.468	1.242	0.227	0.56
1200	8.245	4.486	3.759	10.115	6.042	4.073	0.49
1201	4.472	1.355	3.117	4.877	1.520	3.358	0.47
1202	2.262	0.435	1.827	1.789	0.279	1.510	0.54
1203	1.679	0.474	1.204	4.062	0.792	3.270	0.26
1204	2.438	0.625	1.813	6.323	0.844	5.479	0.24
1205	1.532	0.429	1.102	4.114	0.825	3.288	0.24
1206	1.532	0.402	1.130	4.160	0.806	3.353	0.24
1207	1.614	0.342	1.272	4.605	0.452	4.154	0.23
1208	1.812	1.635	0.177	2.269	2.150	0.119	0.59
1210	1.289	0.238	1.052	9.276	6.000	3.277	0.23
1211	0.171	0.160	0.011	2.128	0.264	1.864	0.01
1212	2.829	1.333	1.496	3.124	1.810	1.314	0.53
1213	5.574	1.125	4.449	6.380	1.966	4.413	0.49
1214	4.165	1.679	2.486	6.878	0.517	6.361	0.27
1215	2.867	1.942	0.926	2.504	1.868	0.636	0.51
1216	4.214	4.204	0.010	8.429	4.193	4.237	0.00
1218	2.825	1.943	0.882	2.502	1.820	0.682	0.48
1219	2.461	1.874	0.587	2.345	1.773	0.572	0.48
1220	0.173	0.112	0.061	1.872	0.511	1.362	0.04
1224	0.217	0.256	-0.040	1.465	0.374	1.091	-0.04
1227	1.469	1.526	-0.057	2.869	1.572	1.298	-0.05
1228	3.131	2.714	0.417	4.177	3.687	0.490	0.45

continued ...

Table F.1: continued ...

DNA No.	Chromosome 4			Chromosome 10			Contribution of chr 4
	Signal intensity	Background	Normalised signal intensity	Signal intensity	Background	Normalised signal intensity	
1230	2.526	1.725	0.801	2.253	1.562	0.690	0.51
1237	2.284	1.819	0.465	2.691	2.256	0.435	0.50
1240	6.234	2.819	3.416	4.646	0.990	3.656	0.47
1241	1.507	1.501	0.007	3.020	2.469	0.552	0.01
1242	2.745	0.421	2.324	2.619	0.552	2.066	0.54
1243	12.935	7.567	5.369	17.476	11.247	6.229	0.46
1244	2.519	1.008	1.511	3.661	2.730	0.931	0.59
1246	4.998	3.608	1.390	11.272	7.311	3.961	0.23
1247	0.908	0.176	0.733	3.569	1.157	2.412	0.23
1248	3.283	2.324	0.959	7.072	4.334	2.738	0.26
1249	2.582	1.218	1.364	3.075	2.074	1.002	0.58
1250	2.165	1.494	0.671	4.071	2.341	1.730	0.25
1256	1.536	1.546	-0.010	5.897	3.342	2.555	-0.00
1257	2.717	1.090	1.627	3.261	2.134	1.127	0.56
1258	3.251	0.379	2.872	2.034	0.456	1.577	0.65
1259	1.526	0.894	0.632	2.523	1.936	0.587	0.48
1260	7.041	6.436	0.606	11.389	9.696	1.693	0.25
1263	2.447	2.068	0.379	2.112	1.819	0.293	0.53
1264	7.157	5.746	1.411	5.846	4.272	1.573	0.44
1268	2.697	2.194	0.503	1.709	1.247	0.462	0.49
1270	2.828	2.844	-0.016	4.248	3.946	0.302	-0.05
1271	1.798	1.765	0.034	2.303	1.818	0.485	0.06
1272	1.687	1.711	-0.024	2.267	1.710	0.557	-0.04
1274	3.807	2.923	0.885	4.395	3.394	1.001	0.46
1275	4.142	2.802	1.340	3.387	1.880	1.506	0.46
1276	2.198	0.569	1.629	2.555	1.071	1.484	0.52
1279	3.757	2.346	1.411	4.786	3.112	1.674	0.45
1282	2.268	1.769	0.499	2.046	1.570	0.476	0.50
1283	0.873	0.807	0.066	3.147	0.789	2.359	0.03
1284	1.877	1.494	0.383	1.983	1.583	0.400	0.48
1298	2.898	0.593	2.305	2.770	0.645	2.124	0.53
1299	3.680	3.668	0.013	6.903	5.301	1.602	0.01
EUH2044	4.455	0.944	3.511	3.745	0.885	2.860	0.52
EUH2062	8.727	5.343	3.384	7.535	3.627	3.908	0.45
EUH2064	13.747	2.737	11.010	4.697	1.121	3.576	0.75
EUH2073	5.209	4.516	0.693	2.411	1.441	0.970	0.44
EUH2087	7.201	2.913	4.288	11.385	6.975	4.409	0.46
EUH2092	3.833	0.985	2.848	3.672	1.406	2.266	0.53
EUH2095	4.485	0.429	4.056	3.234	0.360	2.874	0.57
EUH2097	3.963	1.137	2.826	4.250	1.733	2.517	0.50
EUH2099	0.285	0.274	0.011	2.971	0.415	2.555	0.00
EUH2100	2.633	0.771	1.862	1.765	1.213	0.552	0.75
EUH2107	4.781	0.355	4.426	4.606	0.336	4.270	0.50
EUH2112	4.934	0.940	3.994	2.687	1.419	1.268	0.74

continued ...

Table F.1: continued ...

DNA No.	Chromosome 4			Chromosome 10			Contribution of chr 4
	Signal intensity	Background	Normalised signal intensity	Signal intensity	Background	Normalised signal intensity	
EUH2114	4.323	0.850	3.473	3.648	0.977	2.671	0.54
EUH2117	1.753	0.352	1.401	1.235	0.143	1.093	0.55
EUH2118	2.889	0.615	2.274	2.375	0.570	1.804	0.53
EUH2119	4.017	0.630	3.387	3.622	0.564	3.058	0.50
EUH2125	8.721	0.884	7.837	10.611	2.237	8.374	0.46
EUH2126	2.643	0.260	2.383	2.150	0.276	1.874	0.53
EUH2127	1.303	0.187	1.116	0.842	0.028	0.814	0.55
EUH2128	0.695	0.065	0.630	2.001	0.194	1.807	0.25
EUH2130	4.043	0.287	3.756	4.017	0.337	3.679	0.49
EUH2132	2.734	0.073	2.661	1.298	0.430	0.869	0.74
EUH2135	2.385	0.052	2.333	1.330	0.582	0.747	0.75
EUH2136	2.163	0.708	1.455	4.617	3.210	1.407	0.50
EUH2137	4.413	1.150	3.262	7.136	3.849	3.286	0.49
EUH2139	1.010	1.009	0.001	8.037	3.866	4.171	0.00
EUH2145	2.301	0.385	1.916	5.483	0.629	4.854	0.27

The Khoi-San samples can be differentiated from the Black South African samples via a 'EUH' preceding the DNA number. The data utilised for all the analyses were without rounding. The data presented in the table were rounded off to the third decimal.

APPENDIX G

HAPLOGROUP ANALYSES

The haplogroups of 41 randomly selected individuals were determined as discussed in paragraph 5.4. The specific nucleotide present at each of the five SNPs analysed is presented in Table G.1. The haplogroups deduced from the nucleotide combinations at the SNPs are also depicted in Table G.1.

Table G.1: Haplogroup analyses of selected individuals

DNA No.	Translocation profile	SNP3594	SNP10810	SNP7055	SNP11914	SNP10400	Haplogroup
807	monosomy	T	T	~	~	~	L2
819	monosomy	T	C	A	A	~	L0
821	monosomy	C	~	~	~	C	L3
822	trisomy	T	T	~	~	~	L2
832	monosomy	T	C	A	A	~	L0
833	nullisomy	T	C	A	A	~	L0
835	nullisomy	T	C	A	A	~	L0
840	trisomy	T	C	A	A	~	L0
841	disomy	T	C	A	A	~	L0
842	disomy	T	C	A	A	~	L0
843	disomy	C	~	~	~	C	L3
844	trisomy	T	C	A	A	~	L0
847	disomy	T	C	A	A	~	L0
848	trisomy	T	C	A	A	~	L0
849	trisomy	T	C	A	A	~	L0
852	trisomy	T	T	~	~	~	L2
853	trisomy	T	C	A	A	~	L0
930	monosomy	T	C	A	A	~	L0
960	disomy	T	C	A	A	~	L0
973	disomy	T	T	~	~	~	L2
987	nullisomy	T	C	A	A	~	L0
988	disomy	T	C	A	A	~	L0
989	nullisomy	T	C	A	A	~	L0
1009	monosomy	T	C	A	A	~	L0
1020	monosomy	C	~	~	~	C	L3
1040	monosomy	T	C	A	A	~	L0
1064	monosomy	T	C	A	A	~	L0
1093	nullisomy	T	T	~	~	~	L2
1096	disomy	T	T	~	~	~	L2
1098	monosomy	T	C	A	A	~	L0
1101	disomy	C	~	~	~	C	L3
1102	quatrosomy	T	T	~	~	~	L2
1106	trisomy	T	C	A	A	~	L0
1108	trisomy	C	~	~	~	C	L3

continued ...

Table G.1: continued ...

1109	nullisomy	T	T	~	~	~	L2
1131	disomy	T	C	A	A	~	L0
1132	nullisomy	T	T	~	~	~	L2
1138	nullisomy	T	T	~	~	~	L2
1158	nullisomy	T	T	~	~	~	L2
1171	disomy	T	C	A	A	~	L0
1187	nullisomy	T	C	A	A	~	L0

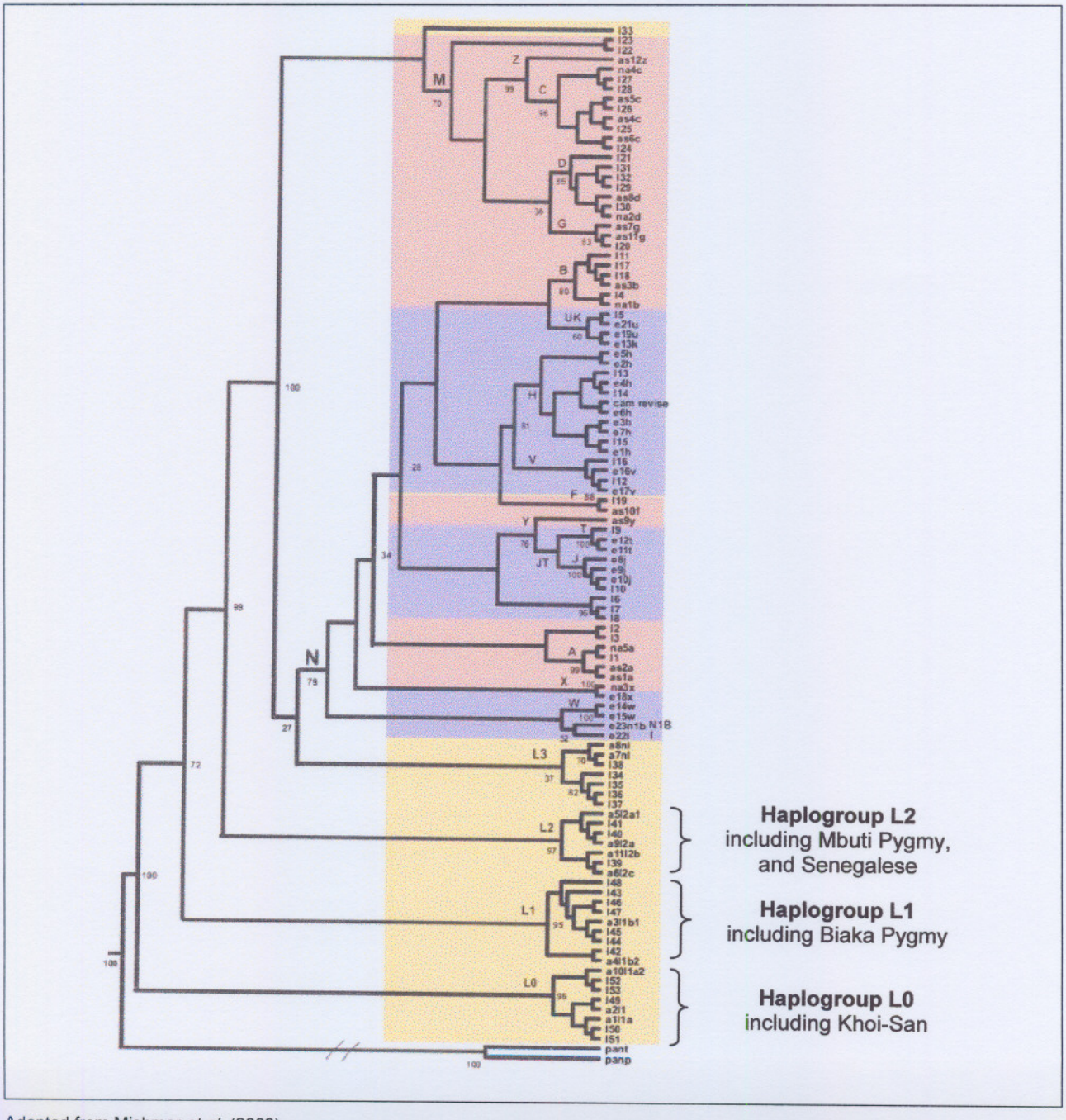
~ = particular SNP was not analysed. L0, L2 and L3 as presented in Figures 5.3 and 6.14.

APPENDIX H

HUMAN mtDNA EVOLUTION

Figure H.1 is a consensus neighbour-joining tree constructed with 104 human mtDNA full genome sequences (Mishmar *et al.*, 2003).

Figure H.1: Consensus neighbour-joining tree of human mtDNA evolution



Adapted from Mishmar *et al.* (2003).

APPENDIX I

CONFERENCES AND MEETINGS AT WHICH RESEARCH WAS PRESENTED DURING THIS STUDY

Research was presented at the following national and international meetings during the period of this study. The presenting author's name is underlined in each case.

I.1 PRESENTATIONS AT INTERNATIONAL CONFERENCES

I.1.1 **53rd Annual meeting of the American Society of Human Genetics,** Los Angeles, U.S.A., November 2003.
Alessandrini M., Van der Merwe A. Schutte C-M and Olckers A. Assessment of the translocation frequency between 4q and 10q in the South African FSHD population: a pilot study (poster presentation).

I.1.2 **FSHD International Consortium Research Meeting**, Los Angeles, U.S.A., November 2003.
Olckers A., Alessandrini M. and Van der Merwe A. Non-radioactive detection of the translocation frequency between 4q and 10q in the South African FSHD population (poster presentation).

I1.3 **Xth International Workshop on Malignant hyperthermia and 22nd Annual Meeting of the European MH Group**, Brunnen, Switzerland, June 2003.
Olckers A., Van der Merwe A., Gericke G.S. and Webber JL. Novel MH locus on chromosome 2q suggested via genome wide linkage screening of South African MH Family SA102 (platform presentation).

I.1.4 **52th Annual meeting of the American Society of Human Genetics**, Baltimore, U.S.A., October 2002.
Van der Merwe A., Alessandrini M., Maree F.F., Schutte C-M., Honey E., Frants R.R., Van der Maarel S.M. and Olckers A. The first compound heterozygote individual identified in the South African FSHD population (poster presentation).

I.1.5 **52th Annual meeting of the American Society of Human Genetics**, Baltimore, U.S.A., October 2002

Alessandrini M., Van der Merwe A., Maree F.F., Schutte C-M., Frants R.R., Van der Maarel S.M. and Olckers A. Molecular analysis suggests the first sporadic case of FSHD in the South African population (poster presentation).

I.2 PRESENTATIONS AT NATIONAL CONFERENCES

I.2.1 **10th Biennial Congress of the Southern African Society of Human Genetics**, Durban, May 2003.

Van der Merwe A., Alessandrini M., Schutte C-M., Honey E., Frants R.R., Van der Maarel S.M. and Olckers A. Molecular analysis identified the first FSHD compound heterozygote in the South African population (platform presentation).

I.2.2 **42nd Annual Congress of the Federation of South African Societies of Pathology**, University of the Free State, Bloemfontein, July 2002.

Alessandrini M., Van der Merwe A., Schutte C-M., Frants R.R., Van der Maarel S.M., and Olckers A. Molecular Diagnosis of Facioscapulohumeral muscular dystrophy (platform presentation).

I.2.3 **Annual congress of the Neurological Association of South Africa**, Sandton, Johannesburg, March 2002.

Van der Merwe A., Alessandrini M., Schutte C-M., Honey E., Frants R.R., Van der Maarel S.M. and Olckers A. A Dual Founder effect in the South African FSHD population is suggested via molecular analysis (platform presentation).

I.2.4 **Annual congress of the Neurological Association of South Africa**, Sandton, Johannesburg, March 2002.

Alessandrini M., Van der Merwe A., Schutte C-M., Honey E., Frants R.R., Van der Maarel S.M. and Olckers A. The relevance of chromosome 10 haplotype analysis in the molecular diagnosis of FSHD (poster presentation).

I.3 RESEARCH PRESENTED AT FACULTY DAY OF THE FACULTY OF HEALTH SCIENCES, POTCHEFSTROOM UNIVERSITY FOR CHE

I.3.1 **Faculty Day 2003:** Faculty of Health Sciences, Potchefstroom University for CHE, South Africa, October 2003.
Van der Merwe A., Alessandrini M., Schutte C-M., Honey E. Frants R.R., Van der Maarel S.M. and Olckers A. Molecular analysis identified first compound heterozygote individual in the South African FSHD population (poster presentation).

I.3.2 **Faculty Day 2003:** Faculty of Health Sciences, Potchefstroom University for CHE, South Africa, October 2003.
Alessandrini M., Van der Merwe A., Schutte C-M., Honey E. Frants R.R., Van der Maarel S.M. and Olckers A. Molecular analysis suggests the first sporadic case of FSHD in the South African FSHD population (poster presentation).

I.4 PUBLISHED ABSTRACTS IN INTERNATIONAL PEER-REVIEWED JOURNALS

I.4.1 Alessandrini M., Van der Merwe A., Schutte C-M. and Olckers A. Assessment of the translocation frequency between 4q an 10q in the South African FSHD population: a pilot study (abstract), *Am. J. Hum. Genet.*,[1] **73(5)**, 270, 2003.

I.4.2 Van der Merwe A., Alessandrini M., Maree F.F., Schutte C-M., Honey E., Frants R.R., Van der Maarel S.M. and Olckers A. The first compound heterozygote individual identified in the South African FSHD population (abstract), *Am. J. Hum. Genet.*,[1] **71(4)**, 523, 2002.

I.4.3 Alessandrini M., Van der Merwe A., Maree F.F., Schutte C-M., Frants R.R., Van der Maarel S.M. and Olckers A. Molecular analysis suggests the first sporadic case of FSHD in the South African population (abstract), *Am. J. Hum. Genet.*,[1] **71(4)**, 523, 2002.

[1] *Am. J. Hum. Genet* : impact factor 10.869