Molecular analysis of the mitochondrial genome in South African patients with suspected mitochondrial disorders

BY

ANNA CATHARINA VAN BRUMMELEN, B.Pharm., M.Sc.

Dissertation submitted for the degree Magister Scientiae in Biochemistry at the Potchefstroomse Universiteit vir Christelike Hoër Onderwys

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

CO-SUPERVISOR: Doctor Izelle Smuts
Department of Paediatrics, Faculty of Health Sciences, University of Pretoria

December 2003

Molekulêre analise van die mitochondriale genoom in Suid-Afrikaanse pasiënte met moontlike mitochondriale siektetoestande

DEUR

ANNA CATHARINA VAN BRUMMELEN, B.Pharm., M.Sc.

Verhandeling ingedien vir die graad Magister Scientiae in Biochemie by die
Potchefstroomse Universiteit vir Christelike Hoër Onderwys

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

MEDESTUDIELEIER: Dokter Izelle Smuts
Departement Pediatrie, Fakulteit Gesondheidswetenskappe, Universiteit van Pretoria

Desember 2003

To my husband, Roy

ABSTRACT

Human mitochondrial DNA (mtDNA) contains 37 genes, which encode 13 proteins (all subunits of the respiratory chain), 22 transfer ribonucleic acids (tRNA), and two ribosomal RNAs. The mtDNA mutation rate is approximately 10 times higher than that of nuclear DNA and mutations therefore accumulate more rapidly. mtDNA damage may result in mitochondrial dysfunction and consequently disease, especially in those tissues most reliant on energy. Therefore, these disorders are often associated with neuromuscular syndromes, are characterised by extensive clinical variation, and are difficult to diagnose.

During this investigation 42 samples from 34 South African paediatric patients with suspected mitochondrial disorders were amplified, sequenced and screened for 24 pathogenic mtDNA mutations in the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and ATPase 6 mitochondrial genes. Whole mtDNA genome sequencing and screening were performed on four patients with mitochondrial disease criteria scores of eight. The nucleotide sequences were compared to the 2001 revised Cambridge reference sequence for any discrepancies. DNA isolated from whole blood was analysed, except for seven patients for whom DNA could be isolated from both blood and muscle.

A total of 103 different reported polymorphisms, 44 different novel synonymous alterations and 17 different potentially pathogenic mutations were detected. None of the detected alterations were reported pathogenic mutations but the 956-965insCCCCC, T2416C, C3254T, G7979A and A13276G alterations should certainly be investigated further. Haplogroup analysis was performed for the four patients with whole mtDNA genome sequence data, as haplogroups can influence disease expression.

Heteroplasmy was not detected for any of the alterations. However, it was demonstrated that low levels of heteroplasmy, detectable via restriction fragment length polymorphism analysis, remain undetected by cycle sequencing. Possible explanations for not detecting reported mutations could be that the pathogenic mutations are nuclear encoded, present in other tissues or that a novel aetiology accounts for mitochondrial disorders in the South African population.

OPSOMMING

Menslike mitochondriale deoksieribonukleïensuur (mtDNS) beslaan 37 gene wat 13 proteïene (subeenhede van die respiratoriese ketting), 22 oordrag-ribonukleïensure (tRNS) en twee ribosomale RNS kodeer. Die mtDNS mutasietempo is ongeveer 10 keer meer as die van die kern DNS en daarom akkumuleer mutasies vinniger. mtDNS skade mag mitochondriale disfunksie en gevolglik siekte veroorsaak, veral in weefsel wat baie energie-afhanklik is. Gevolglik word hierdie siektetoestande dikwels geassosieer met neuromuskulêre sindrome, gekenmerk deur kliniese variasie en moeilike diagnose.

Gedurende hierdie studie is 42 monsters van 34 Suid-Afrikaanse pediatriese pasiënte met moontlike mitochondriale siektetoestande geamplifiseer, die DNS-volgorde is bepaal en die resultate gesif vir 24 patogeniese mtDNS mutasies in die $tRNS^{Leu(UUR)}$, $tRNS^{Lys}$ en ATPase 6 mitochondriale gene. Volledige mtDNS genoom-volgordebepaling en sifting is uitgevoer vir vier pasiënte met 'n mitochondriale siektekriteria-telling van agt. Die nukleotiedvolgordes is vergelyk met die 2001 hersiende Cambridge-verwysingsvolgorde vir enige afwyking. DNS vanuit bloed is geanaliseer, behalwe vir sewe pasiënte vir wie DNS vanuit bloed en spier geïsoleer kon word.

'n Totaal van 103 verskillende gerapporteerde polimorfismes, 44 verskillende nuwe sinonieme veranderinge en 17 verskillende potensiële patogeniese mutasies is waargeneem. Nie een van die waargenome veranderinge was gerapporteerde patogeniese mutasies nie, maar die 956-965insCCCCC, T2416C, C3254T, G7979A en A13276G veranderinge moet definitief opgevolg word. Die haplogroepe van die vier pasiënte van wie volledige mtDNS genoomdata beskikbaar was, is bepaal, aangesien die haplogroep die aard van die siekte kan beïnvloed.

Heteroplasmie is nie waargeneem vir enige van die veranderinge nie, behalwe moontlik die 956-965insCCCCC verandering. Daar is egter bewys dat lae vlakke van heteroplasmie, waarneembaar met restriksie fragment lengte-polimorfisme, nie met DNS volgordebepaling geïdentifiseer kan word nie. Moontlike redes waarom gerapporteerde mutasies nie waargeneem is nie, is dat die patogeniese mutasies nukleêr geleë is, in ander weefsel teenwoordig is, of dat mitochondriale siektetoestande 'n unieke oorsprong het in die Suid-Afrikaanse bevolking.

TABLE OF CONTENTS

CHAPTER THREE

CHAPTER FOUR

LIST OF ABBREVIATIONS AND SYMBOLS

LIST OF SYMBOLS

#	number
%	percent
α	alpha
αβ	three heterodimers that comprise the knob of the F_1 moiety of ATP synthase
$\alpha_3\beta_3\gamma\delta\varepsilon$	stoichiometric ratio of the five different subunit types of the F_1 moiety of ATP synthase
β	beta
γ	gamma
δ	delta
ε	epsilon
ψ	pseudouridine (5-ribosyl uracil)
TψC	thymine-pseudouridine-cytosine
I	complex I
II	complex II
III	complex III
IV	complex IV
V	complex V
12S rRNA	12S ribosomal RNA
16S rRNA	16S ribosomal RNA
28S ribosomal subunit	small subunit of mitochondrial ribosomes (contains 12S rRNA)
39S ribosomal subunit	large subunit of mitochondrial ribosomes (contains 16S rRNA)
55S ribosomes	mitochondrial ribosomes consisting of a large 39S and smaller 28S subunit

LIST OF ABBREVIATIONS

A	alanine
A or a	adenine
A_{260}/ A_{280}	ratio of absorbance measured at 260 nm and 280 nm
AD	Alzheimer's Disease
ADP	adenosine diphosphate
Ala	alanine
ART	antiretroviral therapy
Asn	asparagine
Asp	aspartic acid
ATP	adenosine triphosphate
ATPase 6	gene encoding adenosine triphosphate synthetase subunit 6
ATPase 8	gene encoding adenosine triphosphate synthetase subunit 8
b_{562}	one of two subunits of cytochrome *b* of complex III with specific spectral characteristics at a wavelength of 562 nm
b_{566}	one of two subunits of cytochrome *b* of complex III with specific spectral characteristics at a wavelength of 566 nm
bp	base pair
C	cysteine
C or c	cytosine
°C	degrees centigrade
ca.	circa: approximately
cm	centimetre
CNS	central nervous system
CO I	gene encoding cytochrome *c* oxidase subunit I
CO II	gene encoding cytochrome *c* oxidase subunit II

CO III	gene encoding cytochrome *c* oxidase subunit III
CoQ	coenzyme Q
COX	cytochrome *c* oxidase
CRS	Cambridge reference sequence
CSB	conserved sequence block
CSF	cerebrospinal fluid
cyt *b*	cytochrome *b*
Cyt *b*	gene encoding cytochrome *b*
cyt bc_1	cytochrome bc_1 complex
cyt *c*	cytochrome *c*
cyt c_1	cytochrome c_1
ΔmtDNA	deletions in mitochondrial DNA
D	aspartic acid
dATP	2'-deoxyadenosine-5'-triphosphate
dCTP	2'-deoxycytidine-5'-triphosphate
ddH_2O	double distilled water
DEAF	deafness
DHU	dihydrouridine
D-loop	displacement loop
DNA	deoxyribonucleic acid
[DNA]	DNA concentration
dNTP	2'-deoxynucleotide-5-triphosphate
dTTP	2'-deoxythymidine-5'-triphosphate
E	glutamic acid
EDTA	ethylenediamine tetra-acetic acid
e.g.	*exempli gratia*: for example
et al.	*et altera*: and others
EtOH	ethanol
F	phenylalanine (amino acid) or forward primer
F_0	proton conduction moiety of ATP synthase, embedded in the mitochondrial inner membrane
F_1	catalytic moiety of ATP synthase, which protrudes into the matrix
FeS	Reiske iron-sulphur protein
g	gram
G	glycine
G or g	guanine
gDNA	genomic DNA
H	histidine
H^+	proton/s
H_2O	water
Hae III	restriction endonuclease obtained from *Hemophilus aegyptus* with GGCC as recognition sequence
HeLa	cervical cancer cells from Henrietta Lacks
HIV	human immunodeficiency virus
H_n	protons taken up from electrochemical negative mitochondrial matrix
H_p	protons delivered at electrochemical positive inter-membrane space
HSP	heavy strand promoter
H-strand	heavy strand
I	isoleucine
Ile	isoleucine
IT_{H1}	upstream, more active, heavy strand transcription initiation site
IT_{H2}	downstream, less active, heavy strand transcription initiation site
IT_L	light strand transcription initiation site
k	kilo: 10^3
K	lysine
KSS	Kearns-Sayre syndrome
L	leucine
$L^{(CUN)}$	tRNA leucine recognising codon CUN
$L^{(UUR)}$	tRNA leucine recognising codon UUR
LHON	Leber's hereditary optic neuropathy
LS	Leigh's syndrome
LSP	light strand promoter
L-strand	light strand
μ	micro: 10^{-6}

µg	micrograms
µl	microlitres
µM	micromolar
M	methionine (amino acid) or molar (moles per litre)
m	milli: 10^{-3}
MDC	mitochondrial disease criteria
MELAS	mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes
MERRF	myoclonic epilepsy and ragged-red muscle fibres
Met	methionine
mg	milligram
$MgCl_2$	magnesium chloride
$[MgCl_2]$	magnesium chloride concentration
MIDD	maternally inherited diabetes and deafness
min	minutes
ml	millilitre
mM	millimolar
MM	mitochondrial myopathy
MMC	mitochondrial myopathy and cardiomyopathy
mRNA	messenger RNA
Msp I	restriction endonuclease obtained from *Moraxella species* with CCGG as recognition sequence
mt	mitochondrial DNA fragment (whole genome sequencing)
mtDNA	mitochondrial DNA
mtRNA	mitochondrial RNA
mtTERM	mitochondrial DNA transcription termination protein
mtTFA	former abbreviation of the mitochondrial transcription factor A, now called TFAM
n	nano: 10^{-9}
N	asparagine
Na_2EDTA	di-sodium ethylenediamine tetra-acetic acid
NAD^+	nicotinamide adenine dinucleotide (oxidised)
NADH	nicotinamide adenine dinucleotide (reduced)
NaOAc	sodium acetate
NARP	neuropathy, ataxia and retinitis pigmentosa
NCBI	National Center for Biotechnology Information
ND1	gene encoding NADH dehydrogenase subunit 1
ND2	gene encoding NADH dehydrogenase subunit 2
ND3	gene encoding NADH dehydrogenase subunit 3
ND4	one of two genes encoding NADH dehydrogenase subunit 4
ND4L	one of two genes encoding NADH dehydrogenase subunit 4
ND6	gene encoding NADH dehydrogenase subunit 6
ND7	gene encoding NADH dehydrogenase subunit 7
NEG	negative control
ng	nanograms
NIDDM	non-insulin dependent diabetes mellitus
nm	nanometres
nmol	nanomoles
NRTI	nucleoside reverse transcriptase inhibitor
nt	nucleotide
O_2	oxygen
O_H	heavy strand origin of replication
O_L	light strand origin of replication
OXPHOS	oxidative phosphorylation
p	pico: 10^{-12}
P	proline
PCR	polymerase chain reaction
PD	Parkinson's Disease
PEM	progressive encephalopathy
PEO	progressive external ophthalmoplegia
pH	indicates acidity, numerically equal to the negative logarithm of the H^+ concentration expressed in molarity
P_i	inorganic orthophosphate
pmol	picomoles
POLG	DNA polymerase gamma
POLG2	accessory subunit of POLG

POS	positive control
PS	Pearson syndrome
PUCHE	Potchefstroom University for Christian Higher Education
Q	ubiquinone (oxidised) or glutamine (amino acid)
QH_2	ubiquinone (reduced)
R	arginine (amino acid) or reverse primer
RC	respiratory chain
RCRS	revised Cambridge reference sequence
redox	oxidation/reduction
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
RRF	ragged-red fibres
rRNA	ribosomal RNA
S	serine (amino acid) or Svedberg units (indicating sedimentation velocity)
s	seconds
$S^{(AGY)}$	tRNA serine recognising codon AGY
$S^{(UCN)}$	tRNA serine recognising codon UCN
SDH	succinate dehydrogenase
SSB	mitochondrial single stranded DNA binding protein
T	threonine
T or t	thymine
T_a	estimated annealing temperature
Taq	DNA polymerase from *Thermus aquaticus*
TBE	89.15 mM Tris® (pH 8.1), 88.95 mM boric acid, 2.5 mM di-sodium ethylenediamine tetra-acetic acid
TFAM	transcription factor of mitochondria, formerly called mtTFA
Thr	threonine
T_m	calculated melting temperature
Tris®[1]	tris(hydroxymethyl)aminomethane
Tris®-HCl	Tris®-hydrochloride
Triton®[2] X-100	octylphenolpoly(ethylene-glycolether)$_n$, for n = 10
tRNA	transfer ribonucleic acid
$tRNA^{Ala}$	tRNA alanine
$tRNA^{Asp}$	tRNA aspartic acid
$tRNA^{His}$	tRNA histidine
$tRNA^{Ile}$	tRNA isoleucine
$tRNA^{Leu(UUR)}$	tRNA leucine (specifically recognising the codon UUR)
$tRNA^{Lys}$	tRNA lysine
$tRNA^{Phe}$	tRNA phenylalanine
$tRNA^{Pro}$	tRNA proline
$tRNA^{Ser(AGY)}$	tRNA serine (specifically recognising the codon AGY)
$tRNA^{Thr}$	tRNA threonine
$tRNA^{Val}$	tRNA valine
U	uracil (nucleotide) or unit
U.S.A.	United States of America
UV	ultraviolet
V	valine (amino acid) or volt (electrophoresis)
Val	valine
W	tryptophan
x g	gravitational acceleration
Y	tyrosine

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

LIST OF EQUATIONS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The completion of this dissertation was made possible by the following people whom I would like to thank with all my heart:

The patients and their families who participated in this study without whom this research could not have been conducted.

My supervisor, Prof. Antonel Olckers, for taking a chance with a pharmacist in the field of Human Genetics, for creating opportunities and for having my career at heart. My co-supervisor, Dr. Izelle Smuts, for the clinical evaluation of the patients, as well as for her kindness, support and passion to find the answer to these conditions. Dr. Francois van der Westhuizen for his advice and for performing the biochemical analyses referred to in this study.

My mentor, Wayne Towers, for his faith in the "pharmacist trying to learn Genetics", his encouragement and support. My stand-in mentor, Marco Alessandrini, for helping me to optimise the primers utilised for whole mtDNA genome sequencing and for reading through the chapters of this dissertation. Annelize van der Merwe for all her help, advice and patience in the laboratory. The other members of our team at the Centre for Genome Research, and especially my fellow Masters students, for their support and for teaching me Genetics.

My parents and family for their support and prayers. My husband, Roy, for all his love, encouragement and support and for running the household while his wife was sequencing or dissertation writing. The Lord for giving me the opportunity, strength and ability to complete this degree.

CHAPTER ONE

INTRODUCTION

Mitochondrial deoxyribonucleic acid (mtDNA) was discovered in 1964 by Schatz *et al.* The complete human mtDNA sequence was determined and published in 1981 by Anderson *et al.* and is referred to as the Cambridge reference sequence (CRS). Anderson *et al.* (1981) identified 37 genes that encodes 13 proteins, 22 transfer ribonucleic acids (tRNA), and two ribosomal RNAs (rRNA) within the 16,569 base pair (bp), closed, circular mtDNA molecule (Clayton, 1982).

The most important function of mitochondria is oxidative phosphorylation or OXPHOS (Scholte, 1988), the system that couples cell respiration to the generation of adenosine triphosphate (ATP), the energy intermediate (Mayes, 1993). OXPHOS is mediated by the enzyme complexes of the respiratory chain (RC) and ATP synthase (Adams and Turnbull, 1996).

The mtDNA mutation rate is approximately 10 times higher than that of nuclear DNA and therefore the accumulation of somatic mutations during life is much more rapid in mtDNA (Richter *et al.*, 1988). The 13 proteins encoded by mtDNA are all subunits of the RC (Anderson *et al.*, 1981) and dysfunction of these subunits due to mutations could impair RC function (Larsson and Clayton, 1995) and cause disease.

As discussed in Chapter 2, mtDNA rearrangement mutations and missense mutations can involve the protein coding genes, tRNAs, rRNAs or non-coding regions of the mtDNA genome. The first point mutation reported in mtDNA was in a protein coding gene, the G11778A alteration in the ND4 gene of patients with Leber's hereditary optic neuropathy (LHON), which is a form of maternally inherited blindness (Wallace *et al.*, 1988a). Other common mtDNA disorders are myoclonic epilepsy and ragged-red muscle fibres (MERRF) and mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes (MELAS). Shoffner *et al.*, (1990) identified the genetic cause of MERRF as the A8344G point mutation in the tRNA lysine ($tRNA^{Lys}$) gene, whereas the most common MELAS mutation is the A3243G point mutation in the tRNA leucine ($tRNA^{Leu(UUR)}$) gene (Goto *et al.*, 1990). Prezant *et al.* (1993) were the first to report a pathogenic mitochondrial rRNA

mutation, associated with non-syndromic deafness, the A1555G mutation in the 12S rRNA gene.

However, mitochondria are under dual genetic control, namely mtDNA and nuclear DNA. Mutation in nuclear genes, encoding RC subunits or proteins that are important for mitochondrial biogenesis and maintenance, can cause mitochondrial disorders that are inherited in a Mendelian fashion (Larsson and Clayton, 1995).

The prevalence of mtDNA disorders as a group is comparable with that of Huntington's disease, which affects 6.4 per 100,000 individuals and is more common than Duchenne's dystrophy, which affects 3.2 per 100,000 individuals (Chinnery *et al.*, 2000). These disorders are maternally inherited and the clinical manifestations depend on the energy requirements of the tissue involved and the level of heteroplasmy (Hirano and Pavlakis, 1994). Even if patients harbour the same mutation, the clinical phenotypes often vary, because of differences in overall mutation load between patients (Yasukawa *et al.*, 2002). Therefore, these disorders are clinically, biochemically and molecularly heterogeneous and difficult to diagnose. To establish definitive diagnosis, evidence from at least two relatively independent types of investigation (i.e. clinical, histological, biochemical or molecular) is required (Bernier *et al.*, 2002).

The strict maternal inheritance of mtDNA (Giles *et al.*, 1980) and the lack of recombination make mtDNA a powerful tool for measuring the genetic distance between species and within species. Important conclusions about the origin of modern humans have been determined on the basis of the evolution of the mtDNA (Saccone *et al.*, 1993). Most human mtDNA sequence variation has accumulated sequentially along maternal lineages from sets of mtDNA during and after the process of human colonisation of different geographical regions. These groups of related mtDNAs sharing ancient mutations by descent, are called haplogroups and are often found to be geographically or ethnically specific (Torroni, 2000). Haplogroup data are important from a medical point of view in that it can influence disease expression (Torroni, 2000), for example, the expression of LHON in three haplogroup J families with the mild T10663C mutation and without other primary LHON mutations. It was proposed that a haplogroup J background has an important role in the clinical manifestation of certain LHON mutations (Brown *et al.*, 2002).

The investigation presented here involved the molecular analysis of the mtDNA of South African patients with suspected mitochondrial disorders. The long-term objective of the

research programme is to investigate the genetic aetiology of mitochondrial disorders in the South African population.

In Chapter 2 the mitochondria, OXPHOS, the respiratory chain and the genetic nature of mitochondrial disorders are discussed. Chapter 3 contains the materials and methodology utilised to perform the investigation, and the results obtained are presented and discussed in Chapter 4. The conclusions based on the interpretation of the results are presented in Chapter 5. The literature and electronic sources utilised for background purposes and data analysis are listed under references.

CHAPTER TWO

AETIOLOGY AND PATHOGENESIS OF MITOCHONDRIAL DISORDERS

Mitochondria are vitally important organelles as their matrices are the location of fatty acid β-oxidation and the citric acid (Krebs) cycle, whereas the RC, also known as the electron transport chain, and the OXPHOS system are situated in the inner membrane (Mayes, 1993). The most important function of mitochondria is the OXPHOS pathway (Scholte, 1988), the system that couples respiration to the generation of ATP, the high energy intermediate (Mayes, 1993).

2.1 THE MITOCHONDRION

Mitochondria were discovered and named by Benda (1898), while investigating spermatogenesis in vertebrates and invertebrates. The word "mitochondrion" was derived from the Greek words *mitos* (threads) and *chondrion* (granule). In view of similarities between these bacterium-sized organelles and prokaryotes, Margulis (1970) proposed that mitochondria originated from free-living respiring bacteria that were ingested by eukaryotes early in evolution, about 1.5 to 2 billion years ago (Wallace *et al.*, 1999). Over evolutionary time the initial symbiosis resulted in the host cell becoming totally dependant on the aerobic-based metabolism of the mitochondrion for its viability. In return, the protomitochondrion had a constant nutrient supply, thus rendering its housekeeping functions unnecessary, which resulted in the loss of more than 99% of its genes to the cell's nucleus. The model that is currently accepted is that these bacteria were converted into the protomitochondria and eventually into the modern mitochondria (Schon, 1993).

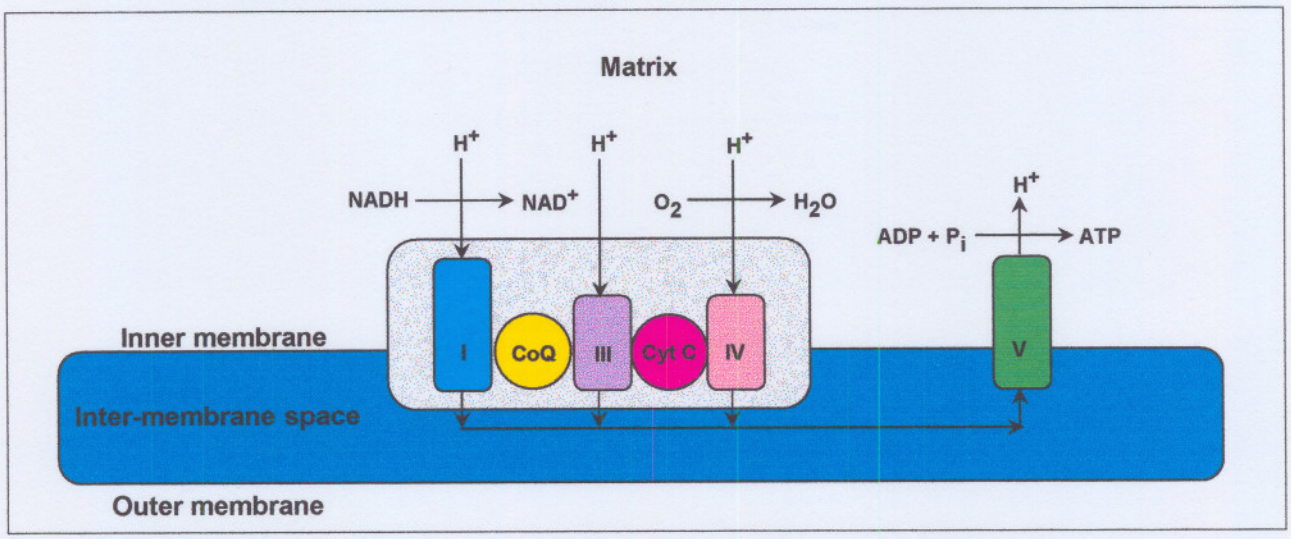
2.2 STRUCTURE AND BIOCHEMISTRY OF THE MITOCHONDRION

Mitchell's chemiosmotic hypothesis (1961), which coupled respiration to ATP synthesis via translocation of protons (H^+) across a semipermeable membrane, explains the unique architecture of the mitochondrion. It is a topologically closed bilayered system with an outer membrane constituting the exterior of the organelle and an invaginated inner mitochondrial membrane, referred to as cristae, surrounding the interior matrix. This

topology is crucial for ATP synthesis through OXPHOS, as it is performed in parallel with the vectorial transport of H^+ across the inner membrane into the matrix (Schon, 1993).

The RC is a functional concept, which means that oxidation of nicotinamide adenine dinucleotide (NADH) dehydrogenase by oxygen (O_2) occurs in a sequential manner as catalysed by the three protein complexes. These complexes are NADH:ubiquinone oxidoreductase (complex I), ubiquinol:ferricytochrome *c* oxidoreductase (complex III or the cytochrome *bc1* complex) and ferrocytochrome *c*:O_2 oxidoreductase [complex IV, cytochrome *c* oxidase (COX) or cytochrome aa_3]. All three complexes are bound to the inner mitochondrial membrane (Wikström, 2003), as presented in Figure 2.1, and are arranged according to increasing operating redox (oxidation/reduction) potential from NADH to O_2 (Bauer *et al.*, 1999).

Figure 2.1: Schematic illustration of the structure and function of the mitochondrial respiratory chain with complex V



I = complex I, III = complex III, IV = complex IV, V = complex V, coenzyme Q (CoQ) and cytochrome *c* (cyt *c*). H^+= hydrogen protons, NADH = nicotinamide adenine dinucleotide (reduced), NAD^+ = nicotinamide adenine dinucleotide (oxidised), O_2 = oxygen, H_2O = water, ADP = adenosine diphosphate, P_i = inorganic orthophosphate and ATP = adenosine triphosphate. Adapted and modified from Larsson and Clayton (1995).

Two additional redox carriers that complete the chain between complexes I and III, and III and IV respectively are ubiquinone (coenzyme Q) and cytochrome *c*. Ubiquinone is a hydrophobic benzoquinone within the inner membrane (Wikström, 2003) and cytochrome *c* is a small haem protein situated in the inter-membrane space and loosely associated with the inner membrane cytosolic side (Adams and Turnbull, 1996). Reducing equivalents such as electrons or hydrogen atoms are transferred through the RC until they reach O_2 at the active site of complex IV, where O_2 is reduced to water (H_2O). Thus, cell respiration is the continuous flux of redox equivalents from substrates to oxygen (Wikström, 2003). The

energy gained along this cascade is coupled to intrinsic proton pumps (Bauer *et al.*, 1999). As electrons are transported along the RC, protons are pumped from the mitochondrial matrix into the inter-membrane space by complexes I, III and IV. This generates an electrical potential as well as a pH gradient across the inner mitochondrial membrane. These effects represent the proton motive force. As protons re-enter the matrix through complex V, the energy from the proton gradient is utilised to produce ATP from adenosine diphosphate (ADP) and inorganic orthophosphate or P_i (Adams and Turnbull, 1996).

Therefore, the respiratory complexes not only function as oxidoreductases, but also have the ability to conserve the free energy of the redox reaction for ATP synthesis. However, complex II (succinate:ubiquinone oxidoreductase) differs from complex I, III and IV in this respect. Complex II was considered to be an integral part of the RC in the past owing to its association with the inner membrane (Wikström, 2003), but it cannot conserve energy. However, complex II functions in the citric acid cycle where it catalyses the oxidation of succinate to fumarate, with transfer of the reducing equivalents to ubiquinone in the inner membrane (Lancaster and Kröger, 2000). Metabolites oxidised by complex II bypass complex I by delivering the reducing equivalents directly to ubiquinone (Wikström, 2003). The subunits of the three RC complexes as well as those of complex II and V are listed in Table 2.1. Thirteen of these subunits are encoded by mtDNA and the balance by the nucleus (Larsson and Clayton, 1995).

Table 2.1: Mitochondrial-encoded subunits of complexes I to V

Complex	Enzyme	Mitochondrial genes	Number of subunits	Reference
I	Nicotinamide adenine dinucleotide (NADH) dehydrogenase	ND1, ND2, ND3, ND4, ND4L, ND5 and ND6.	46	1
II	Succinate dehydrogenase	None	4	2
III	Ubiquinol:ferricytochrome *c* oxidoreductase	Cyt *b*	11	2
IV	Cytochrome *c* oxidase (COX)	CO I, CO II, CO III	13	3
V	F_0-F_1-ATP synthase	ATPase 6, ATPase 8	16	4

ND1, ND2, ND3, ND4, ND4L, ND5 and ND6 = genes encoding subunits 1 to 6 of NADH dehydrogenase, Cyt *b* = gene encoding cytochrome *b*, CO I, CO II, CO III = genes encoding subunits I, II and III of cytochrome *c* oxidase, ATPase 6 and 8 = genes encoding subunits 6 and 8 of ATP synthase, F_0 = proton conduction moiety of ATP synthase and F_1 = catalytic moiety of ATP synthase. Constructed from data listed in Attardi (1993) and 1 = Carroll *et al.* (2002), 2 = Adams and Turnbull (1996), 3 = Campbell and Smith (1993) and 4 = Walker *et al.* (1991).

2.2.1 Complex I (NADH: ubiquinone oxidoreductase/NADH dehydrogenase)

Carroll *et al.* (2002) recently identified three more subunits of complex I and it is, therefore, composed of 46 different subunits with total molecular mass of 980 kilodalton and not only 43 as believed for the last 12 years (Walker, 1992). Complex I is the largest and first complex mediating electron transfer in the RC (Triepels *et al.*, 2001). The overall function of the complex is to serve as an electron acceptor for several NADH-producing reactions and to transfer the electrons from NADH to ubiquinone (Bauer *et al.*, 1999). This electron transfer is associated with proton translocation across the inner membrane as illustrated in Equation 2.1:

Equation 2.1: Proton translocation across the mitochondrial inner membrane

$$NADH + Q + 5H_n^+ \rightarrow NAD^+ + QH_2 + 4H_p^+$$

NADH = nicotinamide adenine dinucleotide (reduced), Q = ubiquinone (oxidised), NAD = nicotinamide adenine dinucleotide (oxidised), QH_2 = ubiquinol (reduced), H_n = protons taken up from electrochemical negative matrix and H_p = protons delivered to electrochemical positive inter-membrane space. Adapted from Adams and Turnbull (1996).

Seven of the 46 subunits are mitochondrially encoded, the genes of which comprise 40% of human mitochondrial DNA (Walker, 1995). The subunits are arranged in an L-shaped structure (Hofhaus *et al.*, 1991). One region is embedded in the mitochondrial membrane while the other protrudes into the matrix. This peripheral section forms the functional NADH dehydrogenase. The mtDNA encoded subunits are all located in the membrane associated structure, which also contains the ubiquinone dehydrogenase (Adams and Turnbull, 1996).

The mechanism by which electron transfer is coupled to proton translocation is poorly understood and mostly speculation. Substrate-induced conformational changes throughout the catalytic part of this complex could be a possible mechanism of action, resulting in proton uptake and release on the opposite side of the membrane (Belogrudov and Hatefi, 1994; Triepels *et al.*, 2001).

2.2.2 Complex II (succinate ubiquinone oxidoreductase/succinate dehydrogenase)

Complex II is also an electron acceptor, receiving electrons from succinate via reduced flavin adenine dinucleotide and transferring these electrons to ubiquinone. Ubiquinone also receives electrons from the flavoprotein-linked steps of fatty acid β-oxidation and sn-glycerophosphate dehydrogenase (Bauer *et al.*, 1999). Complex II contains four

polypeptides and is the only electron-transferring complex of which none of the subunits are encoded by mtDNA (Adams and Turnbull, 1996). As mentioned earlier, complex II is no longer regarded as an integral part of the RC, as it is not situated in the inner mitochondrial membrane and functionally belongs to the citric acid cycle enzymes (Wikström, 2003)

2.2.3 Complex III (ubiquinol:ferricytochrome *c* oxidoreductase)

Complex III or cytochrome *bc1* complex (cyt bc_1) mediates the transfer of electrons from ubiquinol to cytochrome c. Although cytochrome *c* has not been isolated as a component of a complex (Bauer *et al.*, 1999), it acts as an intermediate carrier for the transfer of electrons from complex III to complex IV. Binding sites for cytochrome *c* have been localised on both complexes (Adams and Turnbull, 1996). As a peripheral protein, cytochrome *c* may be readily released from the outer surface of the inner membrane into the inter-membrane space (Bauer *et al.*, 1999).

Complex III is composed of 11 subunits of which only one (Adams and Turnbull, 1996) cytochrome *b* (cyt *b*), is encoded by mtDNA (Anderson *et al.*, 1981). It actually contains two *b* haem groups of cytochrome *b*, designated b_{562} and b_{566} owing to their spectral characteristics (Campbell and Smith, 1993). Two other important subunits of complex III are cytochrome c_1 (cyt c_1) and the Reiske iron-sulphur (FeS) protein (Adams and Turnbull, 1996).

Complex III translocates four protons from the matrix across the inner mitochondrial membrane for each pair of electrons that are transferred from ubiquinol to cytochrome *c* (Adams and Turnbull, 1996). The Q-cycle has been proposed as an explanation for this mechanism. According to this mechanism, electrons received in pairs by ubiquinone from either complex I or complex II can be passed on singly to cyt c_1 via FeS. Simultaneously protons are released into the inter-membrane space (Campbell and Smith, 1993).

2.2.4 Complex IV (ferrocytochrome *c*:O_2 oxidoreductase/cytochrome *c* oxidase)

Electrons from complex III are transferred to complex IV via cytochrome *c* and are associated with proton translocation across the inner membrane. The final step in the RC is the COX catalysed sequential transfer of four electrons from the reduced cytochrome *c* to O_2, forming two H_2O molecules (Bauer *et al.*, 1999). COX comprises 13 subunits.

Subunits I, II and III are encoded and synthesised in the mitochondrion and form the catalytic core of the complex, while the remaining 10 subunits are nuclear-encoded (Campbell and Smith, 1993).

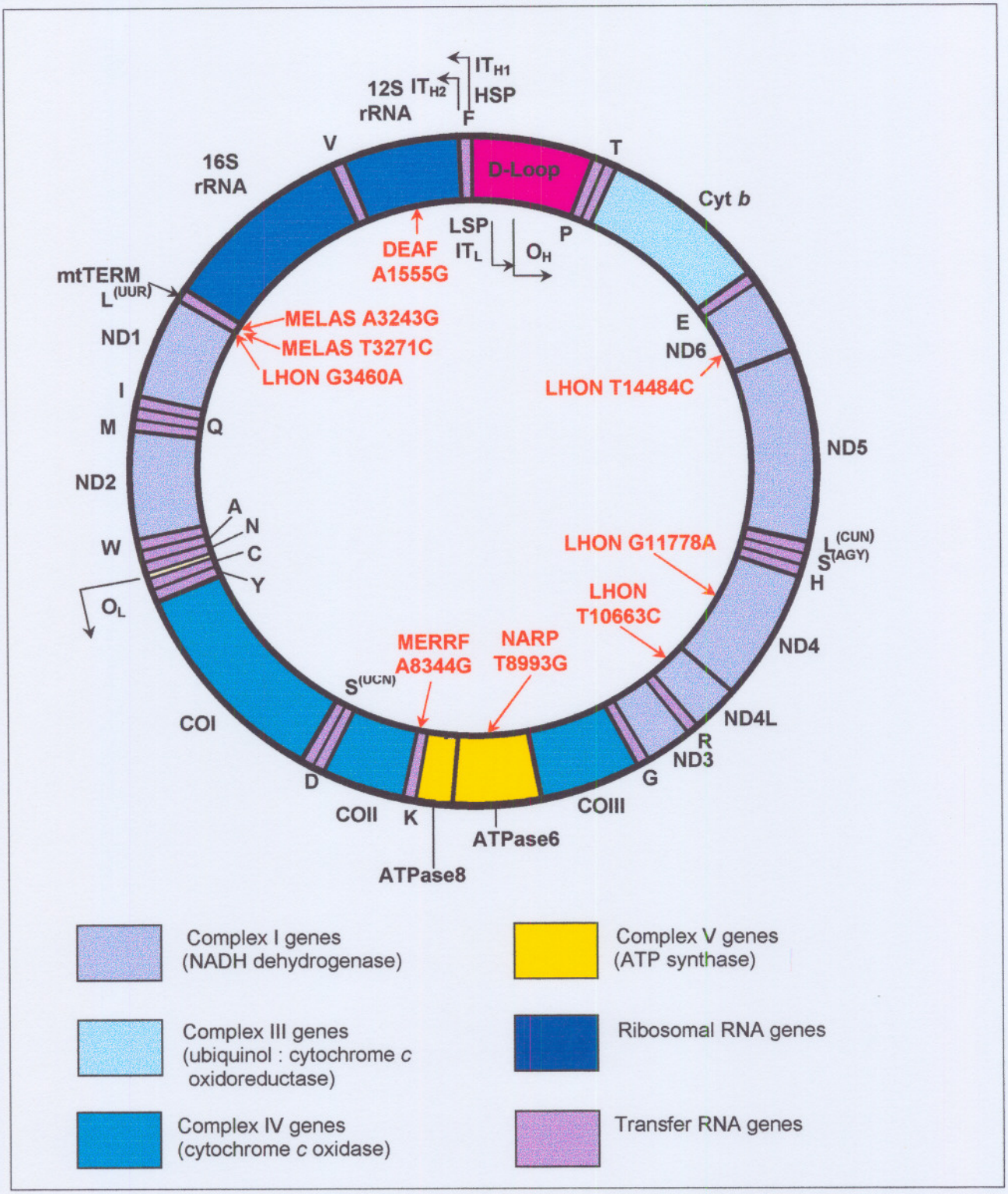
2.2.5 Complex V (F_0-F_1-ATP synthase)

The electrochemical gradient generated across the inner membrane by complexes I, III and IV during OXPHOS provides a proton motive force, which drives ATP synthesis by ATP synthase (Bauer *et al.*, 1999). Bovine ATP synthase is composed of 16 different polypeptides (Walker, 1991) of which two membrane components (ATPase 6 and 8) are encoded by mtDNA in overlapping genes (Anderson, 1981). The ATP synthase consists of F_0 and F_1 moieties, which are responsible for the proton conduction and the catalytic functions of the enzyme, respectively. The catalytic F_1 region protrudes into the matrix and is connected to the F_0 region (embedded in the inner membrane) by a short stalk-like structure (Senior, 1988). F_1 is a soluble protein consisting of five different subunit types, namely alpha (α), beta (β), gamma (γ), delta (δ) and epsilon (ε) in a stoichiometric ratio of $\alpha_3\beta_3\gamma\delta\varepsilon$ (Adams and Turnbull, 1996). The knob of F_1 is composed of three heterodimers of $\alpha\beta$ subunits, whereas γ, δ and ε subunits comprise the stalk (Campbell and Smith, 1993). The F_0 moiety is theoretically composed of three subunits, a (encoded by the mitochondrial ATPase 6 gene), b and c (Adams and Turnbull, 1996). Under certain conditions F_1 can function as an ATPase, thus hydrolysing ATP, formerly known as the mitochondrial ATPase (Campbell and Smith, 1993).

2.3 THE HUMAN MITOCHONDRIAL GENOME

The mitochondrion is the only animal cellular organelle to contain its own DNA other than the nucleus (Schon, 1993). Human cells contain multiple copies (10^3 to 10^4) of a 16,569 bp closed, circular, DNA molecule that is replicated and expressed within the mitochondrial matrix (Clayton, 1982). Mitochondrial DNA was discovered in 1964 by Schatz *et al.* The complete human mtDNA sequence was determined and published in 1981 by Anderson *et al.* and is referred to as the CRS. Anderson *et al.* (1981) identified 37 genes, encoding 13 proteins (all of which are subunits of the RC), 22 tRNAs and two ribosomal rRNAs within the 16,569 bp human mitochondrial genome as presented in Figure 2.2.

Figure 2.2: Morbid and functional map of the human mitochondrial genome



Outer circle = H-strand, inner circle = L-strand, O_H = origin of H-strand replication, O_L = origin of L-strand replication, IT_{H1} and IT_{H2} = H-strand initiation of transcription sites 1 and 2, IT_L, = L-strand initiation of transcription site, HSP = H-strand promoter, LSP =L-strand promoter, rRNA = ribosomal RNA, mtTERM = mitochondrial DNA transcription termination protein, ND1, ND2, ND3, ND4, ND4L, ND5 and ND6 = genes encoding subunits 1, 2, 3, 4, 4L, 5 and 6 of NADH dehydrogenase, CO I, CO II, CO III = genes encoding subunits I, II and III of cytochrome *c* oxidase, ATPase 6 and 8 = genes encoding subunits 6 and 8 of ATP synthase, Cyt *b* = gene encoding cytochrome *b*, D-loop = displacement loop, DEAF = deafness, MELAS = mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes, LHON = Leber's hereditary optic neuropathy, MERRF = myoclonic epilepsy and ragged-red muscle fibres and NARP = neuropathy, ataxia and retinitis pigmentosa. The following letter symbols of amino acids actually indicate the tRNA of that amino acid: F = phenylalanine, V = valine, L = leucine, I = isoleucine, Q = glutamine, M = methionine, W = tryptophan, A = alanine, N = asparagine, C = cysteine, Y = tyrosine, S = serine, D = aspartic acid, K = lysine, G = glycine, R = arginine, H = histidine, E = glutamic acid, T = threonine and P = proline. Note that there are two tRNA genes for leucine, which are differentiated as L$^{(UUR)}$ and L$^{(CUN)}$ and two tRNA genes for serine, differentiated as S$^{(UCN)}$ and S$^{(AGY)}$. Adapted and modified from Taanman (1999) and MITOMAP (2003).

The mtDNA leading strand has been termed the heavy strand (H-strand) because of its greater buoyant density in alkaline cesium chloride gradients as a consequence of a positive guanine (G) and thymine (T) bias in its base composition (Clayton, 1991). Correspondingly, the opposite strand (lagging) of the DNA helix has been termed the light strand (L-strand) because of a relatively high cytosine (C) content. The CRS presents the sequence of the L-strand (Anderson *et al.*, 1981), which is the main coding strand, containing the sense sequence of the rRNAs as well as most of the tRNAs and the messenger RNAs (mRNA).

2.3.1 The Cambridge reference sequence

The CRS has been indispensable for studies of human evolution, population genetics and mitochondrial disorders in the past and present (Andrews *et al.*, 1999). However, it has been recognised that the CRS differs at several sites from mtDNA sequences obtained from other studies. These discrepancies were due to errors in the initial sequencing analysis as well as rare polymorphism in the mtDNA from which the CRS was determined. Another complication of the CRS is that it was principally derived from a single European individual with haplogroup H (Andrews *et al.*, 1999), but has been widely utilised as a mitochondrial genome reference sequence for other haplogroups as well. Furthermore, the CRS sequence originally obtained was not based on human mtDNA alone. It contained some sequences from both Henrietta Lacks (HeLa) cervical cancer cells and bovine mtDNA (Anderson *et al*, 1981).

Andrews *et al.* (1999) reanalysed the original placental mtDNA samples of the CRS investigation and found 11 sequencing errors and seven rare polymorphic alleles. Correction of these errors led to the 2001 revised Cambridge reference sequence (RCRS) with 16,568 bp, due to a single C residue at nucleotide (nt) 3106 instead of the incorrect CC doublet in the CRS presented in 1981. This deleted position at nt 3106 (3106del) was maintained in the RCRS as a gap to retain the historical nucleotide numbers and to prevent confusion (MITOMAP, 2003), as was the rare polymorphism of adenine (A) at nt 750. However, the mtDNA sequence is revised regularly on the website of the National Center for Biotechnology Information (NCBI) and the September 2002 version (NC_001807.4 GI17981852) is indicated to contain 16571 bp. The three extra nucleotides are a T and C at nt 311 and nt 312 respectively, and a C at 16195, without the 3106del (NCBI, 2003). This NCBI version does not retain the original numbering owing to the two extra bases at the beginning of the sequence. To prevent confusion a modified version of

the 2001 RCRS (MITOMAP, 2003), with the original numbering, was utilised for reference purposes in the investigation presented here.

2.3.2 mtDNA replication

Replication of mammalian mtDNA is under relaxed control and takes place independent of the cell cycle phase (Clayton, 1982). The replication machinery and two different replication models will be discussed.

2.3.2.1 mtDNA replication machinery

The machinery required for mammalian mtDNA replication is poorly defined. Only one DNA polymerase, namely DNA polymerase gamma (POLG), has been identified in mammalian mitochondria and is believed to be the replicative enzyme (Spelbrink, 2001). Mammalian mtDNA is dependent on nuclear-encoded proteins for maintenance and faithful propagation. Apart from POLG, three other protein components involved in mammalian mtDNA replication have been well characterised. These are the transcription factor of mitochondria (TFAM, formerly referred to as mtTFA), the accessory subunit of POLG (POLG2) and mitochondrial single-stranded DNA binding (SSB) protein (Spelbrink, 2003).

TFAM is required for accurate and efficient promoter recognition by mammalian mitochondrial RNA polymerase. It activates transcription downstream of its binding site, by unwinding and bending duplex DNA, thereby facilitating access of RNA polymerase to the template (Fisher *et al.*, 1992). POLG2, the β-subunit of POLG, acts as a processivity factor to stimulate the catalytic subunit of human POLG (Carrodeguas and Bogenhagen, 2000). SSB proteins are required to stabilise-single stranded mtDNA regions i.e. in the displacement loop (D-loop) and in replicative intermediates (Zeviani *et al.*, 1995).

Many proteins involved in mammalian mtDNA replication have not been fully characterised (Spelbrink, 2003) e.g. a novel protein, Twinkle, with structural similarity to the phage T7 gene 4 primase/helicase has been identified. Twinkle is apparently critical for lifetime maintenance of mtDNA integrity (Spelbrink *et al.*, 2001)

2.3.2.2 mtDNA replication models

Two models for mammalian replication have been proposed, namely that of Clayton (1982) and the newer model by Holt *et al.* (2000). The Clayton model is an "asynchronous" model of mammalian mtDNA replication (Spelbrink, 2003). Clayton (1982) postulated that the daughter H-strand, or leading strand, is synthesised from the origin of H-strand replication (O_H) on the parental L-strand. The short daughter H-strand of seven Svedberg (S) units, stably associates with the parental closed circle and forms a triplex called the D-loop. H-strand synthesis continues unidirectionally from the D-loop until completed. However, when H-strand synthesis is 67% completed, the origin of replication of the L-stand (O_L) is exposed as single-stranded and initiation of L-strand, or lagging strand, synthesis begins in the opposite direction (Clayton, 1982).

Recently Holt *et al.* (2000) proposed a more conventional "synchronous" model of mammalian mtDNA replication with simultaneous leading and lagging strand synthesis. According to this model mtDNA replication starts from a single origin (at or near O_H) and proceeds around the molecule in one direction. This is in contrast with nuclear DNA where leading and lagging strand synthesis is bidirectional. As DNA synthesis always proceeds in the 5' to 3' direction, short Okazaki fragments are formed on the lagging L-strand (Spelbrink, 2003).

The Clayton (1982) and Holt *et al.* (2000) models of mtDNA replication can be combined into a single model, if a variety of replicative intermediates exist with different numbers of lagging strand start sites. The "synchronous" and "asynchronous" models of replication may represent the extremes of a spectrum in which the frequency of lagging strand initiation varies. Both models of replication apply to mammalian mitochondria, but the ratio of the two types of replicative intermediates is highly variable (Holt *et al.*, 2000).

There is sometimes confusion when the term D-loop is utilised. Many authors, e.g. Spelbrink (2003), regard the D-loop as the third strand, complementary to the L-strand, of approximately 500 nucleotides that arises from O_H and causes displacement of the parental H-strand. However, other authors, e.g. Taanman (1999), regard the D-loop as the region between the genes for the tRNA of phenylalanine ($tRNA^{Phe}$) and the tRNA of proline ($tRNA^{Pro}$) and utilise the term synonymously with control region. In the investigation presented here the D-loop is also regarded as the control region between $tRNA^{Phe}$ and $tRNA^{Pro}$, thus, the 1122 nucleotides from nt 16024 to 576 (MITOMAP, 2003).

2.3.3 mtDNA transcription

The D-loop is the region of mtDNA where protein-DNA interactions are speculated to occur, directing both mtDNA replication (Clayton, 1982) and transcription (Clayton, 1984). There are two major transcription initiation sites in the D-loop. These are the initiation for H-strand transcription site 1 (IT_{H1}) and the initiation for L-strand transcription site (IT_L). The two major transcription sites are situated within 150 bp of each other (Taanman,1999).

As presented in Figure 2.2, IT_L is located within the L-strand promoter (LSP), from which a single large transcript originates, containing the mRNA for the ND6 subunit of complex I and the eight tRNAs encoded by this strand (Attardi, 1986). By contrast, transcription of the H-strand is initiated from two closely located initiation sites (Attardi, 1986). From the upstream, more active site (IT_{H1}), located in the H-strand promoter (HSP), the rRNAs and two tRNAs [$tRNA^{Phe}$ and tRNA valine ($tRNA^{Val}$)] are synthesised as one entity, which terminates at the 16S rRNA/tRNA leucine$^{(UUR)}$ ($tRNA^{Leu(UUR)}$) boundary and consequently yields two rRNA species and two tRNAs. From the downstream, less active, site (IT_{H2}), located near the 5'-end of the 12S rRNA gene, all other tRNAs and the mRNAs encoded within the H-strand are synthesised, in the form of a single polycistronic transcript (Montoya *et al.*, 1983) to be processed to near mature products via endonucleolytic cleavage, by a mitochondrial ribonuclease P. The polycistronic transcript is cleaved before and after a tRNA sequence. Thus, the loci of the tRNA also function as post-transcriptional processing signals (Attardi, 1986).

The mtDNA shows extreme economy of organisation with almost no introns and only the D-loop region, also known as the control region, having a non-coding function (Anderson *et al.*, 1981). The L-strand and H-strand promoters do not overlap and thus function as independent entities (Clayton, 1991). As mentioned above, transcription of the H-strand is performed via two overlapping transcription units (IT_{H1} and IT_{H2}) and underlies the mechanism whereby the rRNA species (as well as $tRNA^{Phe}$ and $tRNA^{Val}$) are synthesised at a rate that is 15 to 60 times that of the mRNAs encoded in the H-strand (Gelfand and Attardi, 1981). In this manner, sufficient amounts of 12S and 16S rRNAs are provided for protein translation. Transcription initiated at IT_{H1}, is terminated via binding of a termination protein (mtTERM) that binds within the $tRNA^{Leu(UUR)}$ gene, immediately downstream of the 16S rRNA and blocks the RNA polymerase (Kruse *et al.*, 1989).

Transcription of mtDNA requires mitochondrial RNA polymerase (mtRNA) and TFAM. TFAM binds at a region upstream of both IT_H and IT_L and activates transcription as a result of DNA binding (Shadel and Clayton, 1993).

Short transcripts initiated at IT_L function as primers for the initiation of replication of the H-strands (Chang and Clayton, 1985). The initiation of L-strand transcription and the initiation of RNA primer formation for mtDNA replication occur through the same mechanism (Clayton, 1991). Apparently mammalian mtDNA replication is intimately linked with mitochondrial transcription. The transition from RNA to DNA synthesis takes place at the conserved sequence blocks (CSB) I-III (Taanman, 1999). The CSB sequences are situated between IT_L and O_H on the H-strand (Larson and Clayton, 1995) and are the most conserved portions of the D-loop (Clayton, 1982). It has been postulated that CSB I-III direct the cleavage of primary transcripts to create the correct primer species for replication (Clayton, 1991).

2.3.4 mtDNA translation

The human mitochondrial translation apparatus consists of mitochondria-specific 55S ribosomes (Attardi and Ojala, 1971), the 22 mtDNA encoded tRNAs, specific nuclear-encoded aminoacyl tRNA synthetases as well as initiation and elongation factors (Attardi, 1993). The 55S ribosomes consist of a large 39S subunit and a smaller 28S subunit, containing the mtDNA encoded 16S rRNA and 12S rRNA species respectively (Attardi and Ojala, 1971).

The differences between the mitochondrial genetic code of mammals and the universal code are presented in Table 2.2 and indicate that the latter is in fact not truly universal. The most striking differences are the use of UGA as a tryptophan recognition codon instead of a stop codon and the use of AGA and/or AGG as stop codons instead of codons that encode arginine. Another interesting difference of the mitochondrial genetic system is the unusual codon recognition pattern, in that it involves a "two out of three" base interaction between codon and anticodon in the four-codon family boxes. Therefore, in the eight family boxes with four codons for one amino acid, there is only one specific mitochondrial tRNA, instead of two (Attardi, 1985).

Table 2.2: Differences between the mitochondrial genetic code of mammals and the universal code

Codon	Mitochondrial code	Universal code
UGA	Tryptophan	STOP
AUA	Methionine	Isoleucine
AGA	STOP	Arginine
AGG	STOP	Arginine

A = adenine, G = guanine, U = uracil and STOP = stop codon. Adapted from Attardi (1993).

2.3.5 mtDNA in evolutionary studies

The mtDNA is strictly maternally inherited (Giles *et al.*, 1980), as the cytoplasm of the fertilised zygote is contributed by the oocyte. The sperm makes no genetic contribution to the mtDNA (Wallace *et al.*, 1999), as its mitochondria are destroyed upon penetration of the oocyte (Sutovsky *et al.*, 1999). These mechanisms of paternal mitochondria degradation can fail, but fortunately this happens rarely. One such case has been reported where a mitochondrial myopathy was paternally inherited (Schwartz and Vissing, 2002; Williams, 2002). However, in general it is accepted that mtDNA missense mutations are either maternally inherited or have arisen as *de novo* mutations in the germline. As discussed in paragraph 2.5.2, a woman carrying a homoplasmic mtDNA point mutation will transmit it to all her offspring (males as well as females), but only the daughters will transmit it to their progeny (Shanske *et al.*, 2001).

In contrast with the high degree of conservation within the rest of the genome, the D-loop region shows great variability in length and base composition among mammals, except in a central region of approximately 250 nucleotides (Saccone *et al.*, 1993). The reason for this high level of variability is that the D-loop contains no genes and is subject to less stringent selective pressure, compared with the rest of the mitochondrial genome. As a consequence, the D-loop retains the fastest evolutionary rate and the highest intraspecific variability within the mitochondrial genome (Clayton, 1982).

Due to its maternal inheritance and the lack of recombination, mtDNA is applied as a powerful tool for measuring the genetic distance between species, but also within species. Important conclusions about the origin of modern humans have been reached on the basis of the evolution of the mtDNA (Saccone *et al.*, 1993). Most human mtDNA sequence variation has accumulated sequentially along maternal lineages from sets of mtDNA founders during and after the process of human colonisation of different geographical

regions. These groups of related mtDNAs sharing ancient mutations by descent, are called haplogroups and are often found to be geographically or ethnically specific (Torroni, 2000).

However, as with questions regarding the strict maternal inheritance of mitochondria, the view that mtDNA is inherited in a clonal fashion and does not undergo recombination has also been challenged (Hagelberg, 1999 and 2003; Awadalla *et al.*, 1999; Eyre-Walker *et al.*, 1999 and 2001). Nuclear DNA undergoes recombination during meiosis I when the maternal and paternal homologs of each chromosome pair form a bivalent, after which crossing-over occurs (Strachan and Read, 1998). On the other hand Hagelberg (2003) stated that "there is now a large body of literature for and against recombination ... of the mitochondrial genome". Anomalies in mtDNA datasets, e.g. the discrepancy between the mtDNA mutation rates observed in different evolutionary timescales (dating the divergence between two species versus those measured within family pedigrees) and a high frequency of homoplasies (a character state shared by different taxa owing to convergence, parallelism or reversal, but not inheritance from a common ancestor), among others, caused some geneticists to question whether mtDNA does not perhaps recombine. The discrepancies in the mtDNA molecular clock are often attributed to rate heterogeneity between sites (Hagelberg, 2003). Hagelberg (2003) further argues that regions currently viewed as hypervariable sites within mtDNA may not be "mutation hotspots" but ancient mutations that were distributed among unrelated lineages worldwide through recombination.

If recombination does occur it will have far-reaching implications for many theories on human evolution, currently based upon mtDNA genetic evidence. An example would be the "out-of-Africa with total replacement hypothesis", which postulates that anatomically modern humans developed in Africa and totally replaced the archaic populations outside Africa (e.g. the Neanderthals) without interbreeding with them (Stringer and Andrews, 1988). Based on mtDNA data the Neanderthals are classified as a separate biological species distinct from modern humans (Krings *et al.*, 1997). If mtDNA recombination occurs, male mtDNA lineages could contribute to offspring. The calculated age of the Mitochondrial Eve will then be an underestimate. This will decrease the value of mtDNA evidence for the out-of-Africa hypothesis and mtDNA data alone will not be sufficient to consign the Neanderthal to a separate biological species (Hagelberg, 2003). The question of human mitochondrial recombination is far from being resolved (Hagelberg, 2003; Eyre-Walker and Awadalla, 2001). However, its occurrence is highly unlikely.

2.4 NUCLEAR DNA MUTATIONS

It is important to emphasise that most mitochondrial disorders are a result of mutations in nuclear encoded genes. For example, COX deficiency presenting as Leigh's syndrome (LS) in infancy is known to be an autosomal recessive disorder (Shanske *et al.*, 2001). The nucleus can cause mitochondrial disease due to defective transcription or translation of the mitochondrial proteins encoded by nuclear genes or alternatively due to mutations of nuclear genes that control mtDNA gene expression (Zeviani *et al.*, 1990). Loss or impaired function of a nuclear encoded RC subunit will lead to a deficiency of the corresponding enzyme complex of the RC. However, the nucleus also encodes other proteins that are important for mitochondrial biogenesis and maintenance, e.g. all the proteins required for mtDNA replication, transcription, processing and translation of mtDNA transcripts, as well as proteins required for mitochondrial protein import. Loss of mtDNA polymerase, mtRNA polymerase or TFAM can cause loss of mtDNA, which can be lethal in early embryonic development. Milder mutations of these proteins may cause mtDNA depletion (Larsson and Clayton, 1995). The dual genetic control of the mitochondria makes mitochondrial disorders unique from a genetic point of view (Shanske *et al.*, 2001).

2.5 MITOCHONDRIAL DNA MUTATIONS AND DISEASE

Since the initial reports of mtDNA deletions (Holt *et al.*, 1988) and missense mutations (Wallace *et al.*, 1988a) that linked mtDNA mutations to disease, there has been an explosion of information on pathogenic mtDNA alterations. The maternal transmission and high copy number of mtDNA make the inheritance of mutations within this genome fundamentally different from the Mendelian inheritance of nuclear mutations (Larsson and Clayton, 1995).

2.5.1 Genetic aetiology

The mtDNA mutation rate is estimated to be around 10 times higher than that of nuclear DNA and therefore the accumulation of somatic mutations during life is much more rapid in mtDNA. As mitochondria consume more than 90% of the O_2 that enters the cell, free oxygen radicals may preferentially cause damage to mtDNA. Furthermore, mitochondria do not contain histones and lack the sophisticated DNA repair mechanisms present in the nucleus (Richter *et al.*, 1988). POLG has 3' to 5' exonuclease proofreading activity (Ropp and Copeland, 1996), but expression of a mutant form of the protein, without 3' to 5'

exonuclease activity, results in the accumulation of mtDNA missense mutations. Over expression of POLG variants deficient in 5' to 3' polymerase activity results in loss of mtDNA (Spelbrink, 2003).

Mutations in mtDNA regulatory regions can cause defective replication, transcription or processing of mitochondrial transcripts. Mutations within the sites responsible for mtDNA replication (e.g. loss of LSP or O_H) would probably be selected against and the mutated mtDNA copy number will decrease. Loss or dysfunction of mtDNA-encoded RC subunits could impair RC function. However, amino acid substitutes in mtDNA-encoded RC subunits may also have pathogenic effects other than those involving RC function. The immune system may recognise mtDNA-encoded proteins specifically, and amino-acid substitutes of the mitochondrially encoded RC subunits may create novel antigens, causing an autoimmune response (Larsson and Clayton, 1995).

With the exception of complex II, each respiratory complex has one or more of its subunits encoded by the mtDNA. Therefore, these enzymes can be affected by mtDNA mutations, either inherited from maternal mitochondria or arising via a sporadic somatic mutation. As the seven genes encoding the subunits of complex I comprise about 40% of human mitochondrial DNA, it is the respiratory enzyme most likely to be affected by mitochondrial mutations and to be involved in human diseases (Walker, 1995).

2.5.2 Heteroplasmy and threshold effect

In contrast to nuclear loci that consist of one maternal and one paternal allele per cell, there are two to 10 copies of mtDNA per mitochondrion and thousands of mitochondria per cell. If termed *homoplasmic* a specific alteration is determined to be present in all mtDNA molecules, but if termed *heteroplasmic* only a certain fraction of the mtDNA copies are altered. In heteroplasmy the fraction of mutated mtDNA (mutation load) may vary significantly between different tissues or between cells of the same tissue (Larsson and Clayton, 1995). A critical number of mutated mtDNAs must be present before tissue dysfunction and expression of the clinical phenotype occurs. This is referred to as the threshold effect. Tissues with high OXPHOS requirements, such as muscle, heart, eye and brain, have relatively low thresholds and are particularly vulnerable to mtDNA mutations. It is therefore not surprising that most mtDNA disorders are encephalomyopathies, affecting primarily brain and muscle tissue (Shanske *et al.*, 2001).

An individual with a homoplasmic mtDNA mutation will transmit the mutation to all offspring, but a mother with a heteroplasmic mtDNA mutation will transmit varying levels of mutated mtDNA to her children. It is also possible that the mutation is not transmitted. The risk of transmission depends on the levels of mutated mtDNA in the mother. Once this reaches a threshold level of 35% to 40%, mutated mtDNA is likely to be transmitted to all children (Larsson *et al.*, 1992).

2.5.3 Mitotic segregation

During cell division the proportion of mutant mtDNAs in daughter cells can shift, and this phenomenon is called mitotic segregation (Shanske *et al.*, 2001). In some tissues and organs, selection may lead to a decrease in the amount of mutant mtDNA, whereas the level may increase in others. This increase occurs through selective proliferation of defective mitochondria, containing a high mutation load, in response to the RC defect (Yoneda *et al.*, 1992). Over time, this will lead to an increase in the amount of mutant mtDNA (positive selection) within postmitotic tissue such as skeletal muscle (Chinnery and Samuels, 1999). If the pathogenic threshold for a particular tissue is exceeded, the phenotype may change, as discussed in paragraph 2.5.4. Thus, in a patient who is heteroplasmic for a pathogenic mutation, the clinical phenotype may change over the course of time (Shanske *et al.*, 2001).

2.5.4 Late onset of mitochondrial disorders and threshold effect

One typical feature of Leber's hereditary optic neuropathy (LHON), myoclonic epilepsy and ragged-red muscle fibres (MERRF) and maturity-onset diabetes and deafness is the late onset and the progressive course of these disorders. Furthermore, they are all due to maternally inherited mutations, thus the genetic defect was present at birth, but the symptoms do not appear until midlife. Apparently the partial OXPHOS defect caused by the inherited mtDNA mutation is augmented by an age-related decline in OXPHOS. Ultimately the combination of both factors results in the energy production of the tissue falling below the expression threshold, resulting in clinical manifestations (Wallace and Lott, 1993).

The age-related OXPHOS decline is probably caused by the progressive accumulation of somatic mtDNA mutations. Inhibition of the RC causes electrons to accumulate at the beginning of the pathway and to be lost from the flavin dehydrogenases, ubiquinone and

cytochrome *b* to produce superoxides and their derivatives. These oxygen radicals cause further damage to the mtDNA and create a self-perpetuating decline in mitochondrial energy production, which ultimately exceeds expression thresholds and results in organ failure (Wallace and Lott, 1993).

2.6 CLINICAL PHENOTYPES CAUSED BY MITOCHONDRIAL DNA MUTATIONS

As mentioned earlier, mitochondrial disorders are unique for several reasons. They are maternally inherited, those associated with mtDNA mutations can be heteroplasmic and the clinical manifestations depend on the energy requirements of the tissue involved (Hirano and Pavlakis, 1994). The clinical phenotype of the organs affected by mitochondrial disorders usually includes those tissues most reliant on mitochondrial energy, such as the occipital and cerebral cortexes, type I skeletal muscle fibres, the auditory pathway, the cerebellum, the brain stem and the heart, in that order. Therefore these disorders are usually characterised by extensive clinical variation (Wallace *et al.*, 1988b). Even if patients harbour the same mutation, the phenotypes and severity of the diseases often vary, because of differences in overall mutation load between patients. Not only can the mutation load be heterogeneous within a tissue, but each patient has his or her own nuclear background which can contribute to phenotypic variation (Yasukawa *et al.*, 2002).

Ragged-red fibres (RRF) are often associated with mtDNA defects and were first noticed by Engel and Cunningham (1963), owing to their tinctorial characteristics, determined when treating skeletal muscle biopsies with a modification of the Gomori trichrome stain. Later, electron microscopic observations indicated that the RRF were actually the proliferation of mitochondrial membranes. It seems that RRF appear when the mtDNA defect interferes with mitochondrial protein synthesis as with mutations of mtDNA encoded tRNAs and rRNAs. Disorders such as Kearns-Sayre syndrome (KSS), MERRF, and mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes (MELAS) can all present with RRF. On the other hand, disorders due to missense mutations of the protein encoding genes, such as LHON and neuropathy, ataxia and retinitis pigmentosa (NARP) syndrome, do not present with RRF (De Vivo, 1993).

2.6.1 mtDNA rearrangements

Large-scale deletions in mtDNA (ΔmtDNA) were first reported by Holt *et al.* in 1988 and subsequently demonstrated to be present in almost all patients with KSS and in ca. 50% of patients with progressive external ophthalmoplegia (PEO) and RRF (Moraes *et al.*, 1989). Most cases were sporadic and presented with the same deletion in various tissues, the disease probably being the result of a single clonally amplified somatic mutation (Johns *et al.*, 1989).

These single deletions should be distinguished from multiple deletions associated with intergenomic signalling defects as discussed in paragraph 2.6.3.1. Individual patients with single deletions usually harbour only one type of deletion per patient and the number of ΔmtDNA varies between different patients and in different tissues from the same patient (heteroplasmy). Virtually all patients with deletions in mtDNA have been sporadic cases. Women with affected offspring and children of affected women are clinically normal, and when tested, do not present with deletions. This suggests that deletions arise *de novo* early in embryogenesis or in the ovum (Shanske *et al.*, 2001).

Single deletions of mtDNA are usually large and often include the loss of several tRNA and protein coding genes. Tissue with high levels of ΔmtDNA is usually affected. Tissue distribution of ΔmtDNA as well as segregation during development and adult life will be important in determining the phenotype (Larsson and Clayton, 1995).

2.6.1.1 Kearns-Sayre syndrome and Pearson syndrome

KSS was the first multisystem mitochondrial disorder to be clinically defined. Kearns and Sayre first described the disorder in 1958 when the authors reported a patient with retinitis pigmentosa, external ophthalmoplegia and complete heart block. A post mortem revealed spongy degeneration throughout the brain (Kearns and Sayre, 1958).

KSS is caused by large-scale deletions of between two and seven kilobase pairs in mtDNA. The deletions are not localised to any single region of the mitochondrial genome and do not seem to be maternally transmitted, but rather arise *de novo*. The clinical syndrome is characterised by early onset (before the age of 20), ophthalmoplegia,

pigmentary retinopathy and at least one of the following conditions i.e. high cerebrospinal fluid (CSF) protein content, heart block or ataxia (Zeviani *et al.*, 1988).

The phenotypic expression of mtDNA deletions is different in children versus adults. Children often have a rapidly progressing disorder with multiple organ involvement e.g. diabetes mellitus, anaemia, tubulopathy, hypoparathyroidism, exocrine pancreatic dysfunction and neurological manifestations. In contrast, adults usually have a slowly progressing neuromuscular disorder with milder symptoms. Adult patients almost always have PEO either alone, or as part of KSS (Larsson and Clayton, 1995). These differences in clinical phenotypes may be explained by the variant tissue distribution of ΔmtDNAs. Widespread distribution with high levels of ΔmtDNA in most tissue will lead to early onset of the disorder, affecting multiple organs, whereas low levels of ΔmtDNA may accumulate with time in non dividing tissues and lead to late onset of the disorder, mainly affecting brain and muscle tissue (Larsson and Clayton, 1995).

In infants deletions of mtDNA cause PS, characterised by severe sideroblastic anaemia, thrombocytopenia, neutropenia and exocrine pancreatic dysfunction (Rötig *et al.*, 1990). Most infants with PS die within their first year, but spontaneous recovery occurs occasionally. Infants who survive PS often develop KSS later during childhood (Larsson *et al.*, 1990). Thus, selection against ΔmtDNA may occur in rapidly dividing tissues such as bone marrow, whereas ΔmtDNA accumulate in non-dividing tissues such as brain and muscle (Larsson *et al.*, 1990). The dynamic differences in ΔmtDNA distribution could explain the different phenotypes associated with a specific mtDNA deletion over time (Larsson and Clayton, 1995).

2.6.2 Missense mutations

The first mtDNA missense mutation was described by Wallace *et al.* (1988a) and was associated with LHON. Missense or point mutations are characterised by single base changes that can affect either protein coding genes or genes involved in protein synthesis, such as tRNA or rRNA genes (Shanske *et al.*, 2001).

2.6.2.1 tRNA missense mutations

Most of the missense mutations reported to date have been in tRNA genes. Since tRNAs are required for translation of the mtDNA encoded proteins, all 13 mtDNA-encoded proteins can be affected, thus explaining the mechanism by which these mutations can affect multiple enzyme activities (Shanske *et al.*, 2001). There have been extensive investigations to determine the method by which these missense mutations result in defective respiration at the molecular level. Mutant MERRF cybrid cells were found to have severely impaired mitochondrial protein synthesis resulting in the production of abortive polypeptides, presumably due to tRNA lysine ($tRNA^{Lys}$) reduction (Enriquez *et al.*, 1995). Yasukawa *et al.* proposed in 2002 that rather than a decrease in the amount of aminoacyl-tRNA, the primary cause of the translational defect resulting in the MERFF and MELAS syndromes is a qualitative abnormality of the mutant tRNAs. Through the development of an *in vitro* mitochondrial translation system it was determined that $tRNA^{Lys}$ harbouring the A8344G mutation as well as $tRNAs^{Leu(UUR)}$ with either the A3243G or U3271C mutations specifically lack post-transcriptional modification of uridine at the anticodon first position. As modification at this position generally controls precise and efficient discrimination of the codons, it implies that mutant tRNAs without the wobble modification are unable to be specific acceptors in the translation process (Yasukawa *et al.*, 2002).

The modification defect in mutant $tRNA^{Lys}$ and $tRNA^{Leu(UUR)}$ was found to be directly due to the missense mutations and not to a secondary effect of mitochondrial dysfunction. Both the wild-type $tRNA^{Lys}$ and $tRNA^{Leu(UUR)}$ have a novel taurine-containing side chain at the 5' position of the uracil base of the wobble nucleotide, while the $tRNA^{Lys}$ also has a 2-thiouridine derivative of this modified uridine. Although putative enzyme(s) of the wobble uridine modification have not been identified to date, the fact that three mutations located in different regions of the two tRNAs cause the modification defect at the same position, implies that this enzyme might specifically recognise the tRNA structure as a whole. The precise tRNA structure could be crucial for enzymatic binding of tRNA as well as the modification reaction (Yasukawa *et al.*, 2002).

Mitochondrial tRNA missense mutations are transmitted maternally and are usually associated with multisystemic disorders, lactic acidosis and massive mitochondrial proliferation in muscle, resulting in RRF (Shanske *et al.*, 2001). The structure and sites of some pathological mutations (including missense mutations) of the $tRNA^{Leu(UUR)}$ molecule are presented in Figure 2.3 (page 27).

2.6.2.1.1 Myoclonic epilepsy and ragged-red muscle fibres

MERRF was the second multisystem mitochondrial disorder to be clinically defined (Fukuhara *et al.*, 1980). It is a rare disorder of the central nervous system (CNS) and skeletal muscle (DiMauro *et al.*, 1985). This clinical syndrome may include myoclonic epilepsy, mitochondrial myopathy, deafness, dementia, ataxia, hypoventilation and mild cardiomyopathy (Rosing *et al.*, 1985). It is associated with defects of OXPHOS complexes I and IV, but its severity depends on the segregation of a heteroplasmic mtDNA mutation. As the OXPHOS capacity declines, tissues of the CNS, skeletal muscle and heart are progressively affected, indicating that tissue-specific energetic thresholds are reached (Wallace *et al.*, 1988b).

Shoffner *et al.*, (1990) identified the genetic cause of MERRF as the A8344G mutation in mtDNA. The mutation alters the thymine-pseudouridine-cytosine (TψC) loop of the lysine tRNA ($tRNA^{Lys}$) gene. Although the mutation is deleterious, it probably does not completely eliminate $tRNA^{Lys}$ function. This is apparent from one patient who was essentially homoplasmic, but only manifested definite symptoms in her early teenage years (Shoffner *et al.*, 1990). The $tRNA^{Lys}$ gene stretches over 70 nucleotides from nt 8295 to 8364 (MITOMAP, 2003). Two other missense mutations T8356C (Silvestri *et al.*, 1992) and G8363A (Ozawa *et al.*, 1997) were also identified to affect the $tRNA^{Lys}$ gene.

2.6.2.1.2 Mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes

The MELAS syndrome was first described in 1975 by Shapira *et al.*, defined as "familial poliodystrophy, mitochondrial myopathy and lactate acidemia". MELAS is a multisystem disorder clinically characterised by encephalopathy, often manifesting as dementia or seizures, stroke-like episodes at a relatively young age (typically by the age of 40 years), and mitochondrial dysfunction in the form of lactic acidosis or RRF (Hirano and Pavlakis, 1994). Recurrent migraine-like headaches with nausea and vomiting are common, as well as hearing loss, short stature, learning difficulties, hemiparesis, hemianopia and limb weakness, whereas some patients present with PEO (De Vivo, 1993)

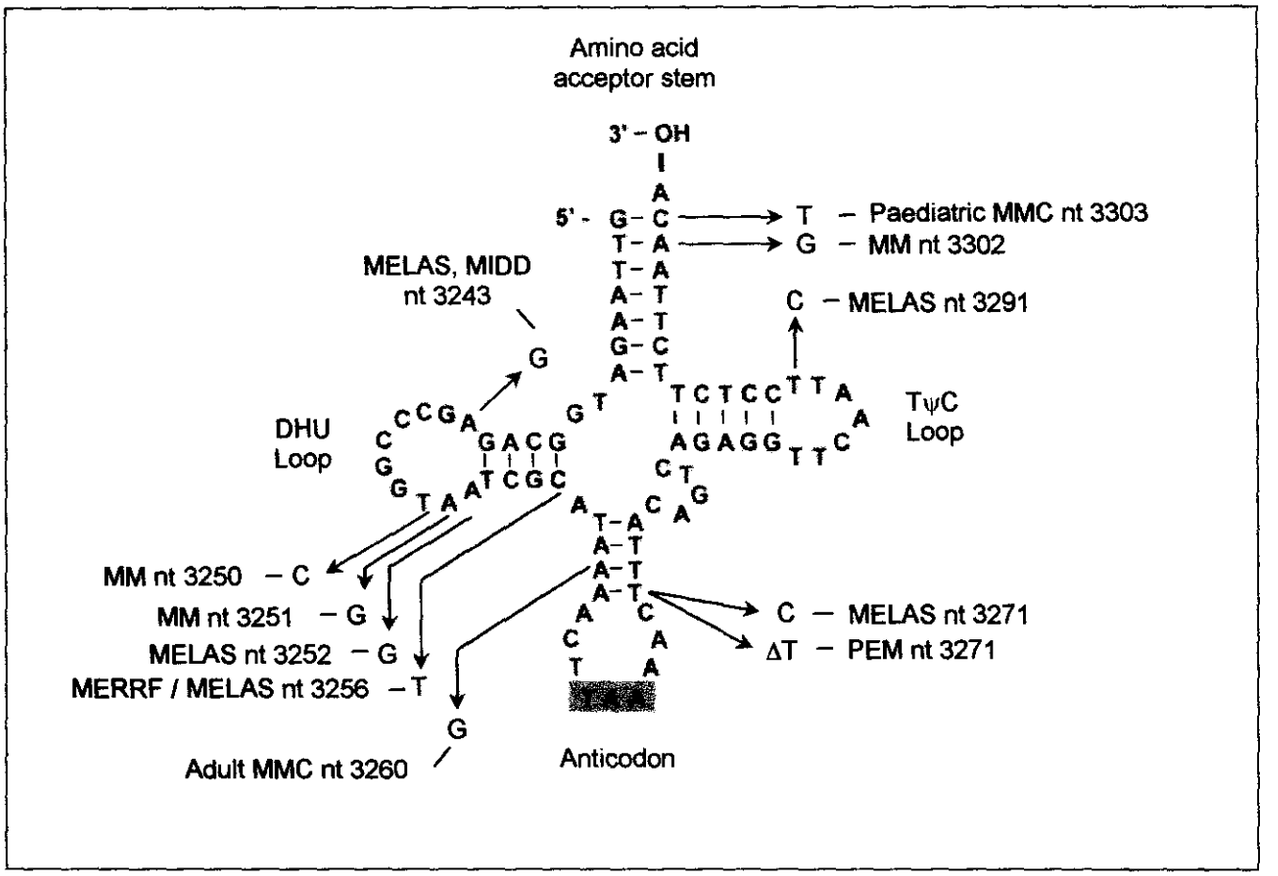
The most common MELAS mutation is the A3243G point mutation in the $tRNA^{Leu(UUR)}$ gene, originally described by Goto *et al.* (1990). This A residue is the first nucleotide in the dihydrouridine loop of the $tRNA^{Leu(UUR)}$ and forms a tertiary hydrogen bond with the T residue of the amino-acid acceptor arm (Goto *et al.*, 1990). Apart from the structural and

functional effect of the A3243G mutation, it was demonstrated that a mutation at this position impairs mitochondrial transcriptional regulation, as the site is embedded in the sequence necessary for transcription termination (Hess *et al.*, 1991). However, the data could not exclude the structural and functional effect of the mutation on the tRNA molecule (Goto *et al.*, 1991). The A3243G point mutation in the $tRNA^{Leu(UUR)}$ gene is the most frequently encountered mtDNA mutation and was the third pathogenic mtDNA mutation to be reported (Shanske *et al.*, 2001). The A3243G mutation is also associated with maternally inherited diabetes and deafness (MIDD), a distinct subtype of non-insulin dependent diabetes mellitus (NIDDM) with sensorineural hearing loss (Van den Ouweland *et al.*, 1994).

In 1991 Goto *et al.* added nine MELAS patients to their original Japanese study group of 31 patients from 1990. Of this MELAS population 80% had the A3243G point mutation and three unrelated patients had a novel T3271C mutation. This T residue is positioned in the anticodon stem of the $tRNA^{Leu(UUR)}$ gene. The detection of this mutation located outside the binding region of mtTERM strengthened the possibility that the abnormal tRNA in itself may be pathogenic (Goto *et al.*, 1990). Besides the Japanese population, both the A3243G (Chinnery *et al.*, 2000) and T3271C (Goto *et al.*, 1994) mutations were also detected in the Caucasian population.

A transition of T to C at nt 3291 in the L-strand of the $tRNA^{Leu(UUR)}$ gene, was identified in one Japanese MELAS patient. This site is the last nucleotide of the TψC loop in the $tRNA^{Leu(UUR)}$ gene. The fact that all three mutations occur within the same tRNA gene supports the idea that there is a close relationship between the mitochondrial $tRNA^{Leu(UUR)}$ gene and MELAS (Goto *et al.*, 1994). The $tRNA^{Leu(UUR)}$ gene stretches over 75 nucleotides from mtDNA 3,230 to 3,304 (MITOMAP, 2003). However, point mutations associated with MELAS, in other mitochondrial tRNA genes as well as protein-encoding genes, have been identified (MITOMAP, 2003).

Figure 2.3: The structure and morbid map of the tRNA$^{Leu(UUR)}$ molecule



MMC = maternal myopathy and cardiomyopathy, MM = mitochondrial myopathy, MELAS = mitochondrial encephalomyopathy *with* lactic acidosis and stroke-like episodes, PEM = progressive encephalopathy, MERRF = myoclonic epilepsy and ragged-red muscle fibres, MIDD = maternally inherited diabetes and deafness. DHU = dihydrouridine, ψ = pseudouridine (5-ribosyl uracil) and nt = nucleotide. Adapted from MITOMAP (2003).

2.6.2.2 rRNA missense mutations

Prezant *et al.* (1993) were the first to report a pathogenic mitochondrial rRNA mutation, namely A1555G, in a highly conserved region of the 12S rRNA gene. Three unrelated patients with maternally transmitted aminoglycoside-induced deafness and a five-generation Arab-Israeli pedigree, with 55 maternal line family members with severe hearing loss, were included in the investigation (Prezant *et al.*, 1993). The use of any of the aminoglycoside antibiotics (e.g. streptomycin, neomycin, gentamycin, kanamycin and amikacin) can cause both vestibular and auditory dysfunction. This is due to the progressive accumulation of these drugs in the perilymph and endolymph of the inner ear, especially when plasma concentrations are high. Ototoxicity is largely irreversible and results from progressive destruction of vestibular or cochlear sensory cells (Chambers and Sande, 1996). In the Pezant *et al.* (1993) investigation, the rare A1555G mutation was the only mutation present in all four families and absent in controls. This mutation is situated in the region known to bind aminoglycosides (Moazed and Noller, 1987) and in which

aminoglycoside resistance mutations have been found in other species (Gravel *et al.*, 1987). The development of non-syndromic deafness in these pedigrees was postulated to be due to the underlying presence of the A1555G mutation alone (Arab-Israeli pedigree) or to the mutation in combination with exogenous administration of aminoglycosides, which accumulate in the cochlea. The A1555G mutation impairs the translational ability of the mitochondrial ribosome (Pezant *et al.*, 1993).

Thyagarajan *et al.* (2000) identified another 12S rRNA missense mutation, T1095C, in a pedigree with maternally inherited sensorineural deafness, levodopa-responsive parkinsonism and neuropathy. They also concluded that the mutation is pathogenic owing to an impairment of OXPHOS by interference with mitochondrial protein synthesis (Thyagarajan *et al.*, 2000).

2.6.2.3 Mutations in protein-coding mtDNA

Cases of patients with mutations in protein-coding genes have mostly been sporadic. Furthermore, these mutations were present only in muscle and were not detectable in blood or fibroblasts. It is believed that these are somatic mutations i.e. spontaneous events that occurred during embryogenesis in muscle after germ layer differentiation and thus do not affect germ line cells (Shanske *et al.*, 2001).

2.6.2.3.1 Leber's hereditary optic neuropathy

The first point mutation reported in mtDNA was G11778A in patients with LHON, a form of maternally inherited blindness, characterised by acute or subacute onset of bilateral vision loss in young adults (Wallace *et al.*, 1988a). LHON is caused by central optic nerve degeneration and there is a bias toward males to present with ophthalmological complications, although cardiac dysrhythmia is also common. The age of onset ranges from adolescence to late adulthood, but the median age of onset is 20 to 24 years. (Wallace *et al.*, 1988a).

The G11778A transition converts a highly conserved arginine to a histidine at codon 340 in the complex I ND4 gene (Wallace *et al.*, 1988a). However, LHON is also associated with two other primary mutations in genes encoding subunits of complex I, namely G3460A and T14484C (Newman, 1993), as well as approximately 30 other missense mutations within the coding and control region of the mitochondrial genome (MITOMAP, 2003). Primary

mutations such as G3460A, G11778A and T14484C account for roughly 90% of all molecularly verified LHON cases. They are designated "primary" mutations because of the high risk of LHON expression associated with them and have been detected as homoplasmic or heteroplasmic mutations (Brown *et al.*, 2002). These primary mutations alter complex I polypeptides and typically induce a complex I functional deficiency. They have been detected in many LHON families, often with multiple affected family members. However, in general there is no co-occurrence of these mutations in the same individual and, to date, these alterations have not been detected in control individuals. Other primary mutations such as the G14459A and A14495G mutations have been detected in LHON families but are rare (Chinnery *et al.*, 2001).

However, some multigenerational LHON families have been reported which lack a known primary mutation. Recently a novel LHON causative mutation, namely the T10663C transition, was identified in the ND4L gene. The T10663C mutation was detected in three LHON families associated with the Western Eurasian (European) mtDNA haplogroup J, with all three lacking other known primary LHON mutations. Haplogroup J is one of nine mtDNA lineages comprising all Western Eurasian mtDNAs. Both the G11778A and T14484C mutations have also been shown to be associated with haplogroup J, much more frequently than would be expected by chance. It is proposed that a haplogroup J background has an important role in the clinical manifestation of certain LHON mutations. Whereas the G3460A and G14459A mutations are sufficiently pathogenic to cause LHON regardless of the patient's mtDNA haplogroup, the functionally "milder" T10663C, G11778A and T14484C mutations require additional genetic or environmental factors, such as haplogroup J, to facilitate expression (Brown *et al.*, 2002).

The LHON missense mutations are usually homoplasmic and RRF are not observed in muscle biopsies (Shanske *et al.*, 2001). Therefore, in contrast with MERRF, the variable symptoms cannot be the result of replicative segregation of a mixed mutant and wild type mtDNA population (Wallace *et al.*, 1988a).

2.6.2.3.2 Leigh's syndrome and neuropathy, ataxia and retinitis pigmentosa

LS is a distinctive neuropathological syndrome characterised by its regional distribution and histopathological features (Leigh, 1951). The lesions are symmetrically distributed in subcortical structures with preferential involvement of the midbrain, pons, basal ganglia, thalamus and optic nerves. Elevated blood and CSF lactate values are a recognised

metabolic correlate to LS and indicate a disturbance in mitochondrial function (Van Coster *et al.*, 1991). Two mitochondrial enzymes, pyruvate dehydrogenase complex (De Vivo *et al.*, 1979) and COX (Willems *et al.*, 1977), have been associated with LS.

The clinical progression of LS can be divided into three phases. Most patients develop normally in the first eight to 12 months of life, learning to sit, stand or even walk. However, they are not without physical symptoms in this period and suffer from gastrointestinal problems such as chronic diarrhoea, recurrent vomiting, anorexia and a decrease in body and head growth (Van Coster *et al.*, 1991). The second phase begins between eight and 12 months of age, when motor regression becomes evident and the patients become increasingly hypotonic. Regression of motor milestones takes place. After the age of one year, CNS dysfunction is noticed with numerous combinations of neurological symptoms. Peripheral neuropathy and abnormal breathing patterns are frequently present (Van Coster *et al.*, 1991). The terminal phase of the disorder is characterised by extreme hypotonia, severe muscle atrophy, dysphagia and dysarthria. Mental and social capacities remain relatively well preserved. The general physical condition of the patient deteriorates, leading to death from respiratory insufficiency around five to six years of age (Van Coster *et al.*, 1991).

Holt *et al.* (1990) described a multisystem mitochondrial disorder with mtDNA heteroplasmy and maternal inheritance, characterised by sensory neuropathy, ataxia and retinitis pigmentosa. Other clinical features were developmental delay, dementia, seizures and proximal limb weakness, but no RRF are present in muscle (Holt *et al.*, 1990).

NARP is invariably heteroplasmic and the clinical phenotype varies markedly as the proportion of mutant mtDNAs fluctuates along the maternal lineage. Mildly affected patients have retinitis pigmentosa and mild neurological symptoms such as migraine. More severely affected patients experience weakness, ataxia, neurological problems and in severe cases, subacute necrotising encephalopathy or LS (Tatuch *et al.*, 1992).

A T8993G mutation in the ATPase 6 gene was identified in the NARP patients and converts a highly conserved leucine to an arginine (Holt *et al.*, 1990). The arginine causes an additional charge group in the proton channel of the ATP synthase, inhibiting ATP synthesis (Tatuch *et al.*, 1992). The ATPase 6 gene stretches over 681 nucleotides from nt 8527 to 9207 (MITOMAP, 2003).

2.6.3 Intergenomic signaling defects

Due to the Mendelian mode of inheritance, these multiple mtDNA deletion disorders have been classified as defects of intergenomic signalling. Therefore, they are mutations in nuclear genes that facilitate an intrinsic tendency of mtDNA to undergo rearrangements or that impair the recognition and elimination of spontaneously occurring rearrangements (Shanske *et al.*, 2001).

2.6.3.1 Dominantly inherited mitochondrial myopathy with multiple deletions of mtDNA

The first family to be described with mtDNA deletions had late onset mitochondrial myopathy, PEO, progressive proximal weakness, bilateral cataracts and precocious deaths as clinical syndromes, which were transmitted as a dominant autosomal trait (Zeviani *et al.*, 1989). Four members of the family were investigated and multiple deletions were detected in all the patients involving the same portion of muscle mtDNA. All the deletions started within a 12-nucleotide stretch at the 5'-end of the D-loop region. All the breakpoints were consistently determined to be associated with candidate sequences for specific interactions with POLG or associated replication factors. This indicated an etiologic correlation between the mechanism generating the lesions and the functions of the D-loop. As the D-loop is a site of active communication between the nucleus and the mtDNA, it was concluded that a mutation of a nuclear-encoded protein can destroy the integrity of mtDNA in a specific, heritable way. Maternal inheritance was ruled out by the recurrent transmission of the clinical and molecular traits to patrilinear descendents (Zeviani *et al.*, 1989). The inheritance of these multiple deletions is in contrast to the described sporadic cases of individual deletions as discussed previously. Zeviani *et al.* (1990) extended their investigation to include more members of the initial family as well as two additional families with similar clinical features and obtained similar results.

2.6.3.2 Mitochondrial DNA depletion syndrome

Another defect of intergenomic signaling is the mtDNA depletion syndrome, a fatal disorder of infancy or childhood in which the number of mtDNAs in mitochondria is severely reduced in different tissues. The number of mtDNAs per mitochondrion has been estimated to vary between two and 10 in different tissues (Bogenhagen and Clayton, 1974). However, two patients investigated by Moraes *et al.* (1991), had less than one

mitochondrial genome per mitochondrion. Many organelles must therefore have been completely devoid of mtDNA. This was in contrast with patients with clinically defined KSS, MERRF and LS in whom the content of mtDNA in muscle was actually increased. COX activity correlated well with the degree of mtDNA depletion and the activity was decreased in tissues with decreased mtDNA levels (Moraes *et al.*, 1991).

The clinical symptoms of this disorder usually arise soon after birth with functional failure of the affected tissue (i.e. muscle, liver, kidney and heart) and early death, because the independent aerobic metabolism is severely impaired. The phenotype may include fatal mitochondrial myopathy, muscle lipid storage and RRF, hepatic failure, renal dysfunction, congestive heart failure, seizures, progressive generalised hypotonia, PEO, respiratory insufficiency and lactic acidosis. Histopathological examination presents with abnormal mitochondrial morphology and proliferation with deficient COX activity (Moraes *et al.*, 1991). The numerous mitochondria demonstrate that mitochondrial proliferation may occur independently of mtDNA replication *in vivo* (Larsson and Clayton, 1995). The non-maternal mode of inheritance suggests a mutation in nuclear DNA, impairing mtDNA replication, i.e. affecting DNA or RNA polymerases, topoisomerases or additional replication factors, as the primary cause of mtDNA depletion (Moraes *et al.*, 1991).

2.6.4 Mitochondrial damage associated with long-term therapy

Human immunodeficiency virus (HIV) infection combination therapy includes at least three antiretroviral drugs chosen from three families of agents, namely nucleoside reverse transcriptase inhibitors (NRTIs), non-nucleoside reverse transcriptase inhibitors and protease inhibitors. Combination therapy of antiretroviral drugs has dramatically improved the prognosis of HIV infection, even though many adverse effects have been associated with these drugs (Vittecoq *et al.*, 2002). In 1995 Lewis and Dalakas were the first to propose that NRTIs inhibit POLG and that this inhibition may be the cause of many tissue-specific adverse effects. NRTI drugs are nucleoside substrates of HIV reverse transcriptase and act as DNA chain terminators through modification of the sugar 3'-hydroxyl groups, hence preventing nucleoside attachment. However, in addition to targeting viral polymerase, these compounds also have the capacity to inhibit POLG, which unlike nuclear DNA polymerases, is unable to discriminate effectively in favour of endogenous nucleosides in the presence of NRTIs (Longley *et al.*, 1998). Zalcitabine has the greatest potency of POLG inhibition, followed by didanosine, stavudine and zidovudine and finally to an equal degree lamivudine, abacavir and tenofovir (Kakuda, 2000).

As POLG is the only mtDNA polymerase, it is critical for the synthesis and repair of the mitochondrial genome. Inhibition of POLG by long-term use of NRTIs can lead to mtDNA depletion (Dalakas *et al.*, 1990; Arnaudo *et al.*, 1991) and mtDNA missense mutations (Martin *et al.*, 2003), possibly causing severe adverse effects. Vitecoq *et al.* (2002) included HIV infected patients in their investigation who complained of unexplained adverse effects and who had received antiretroviral therapy (ART) for at least five years. These patients presented with at least two chronic adverse effects, which were most often neuromuscular in nature and included exercise intolerance and/or amyotrophia, peripheral neuropathy or mental deterioration. Lipodystrophy was almost constantly present and liver steatosis was frequent. Three patients had a severe progressive disease of the CNS with cerebellar and pyramidal symptoms as well as deterioration of cognitive function and seizures. Lactate levels in CSF were elevated in these three patients, although it was normal in blood. Although HIV infection alone can theoretically cause mitochondrial damage (Ferri *et al.*, 2000), a control group of 10 HIV-infected, ART-naïve individuals did not present with any mtDNA alterations compared with five (from a group of 16) HIV-infected, NRTI-treated patients who presented with mtDNA mutations after treatment (Martin *et al.*, 2003).

Recently it was demonstrated that radiotherapy and chemotherapy utilised in the treatment of cancer could induce changes in the mitochondrial genome, including point mutations and deletions. These alterations can be the cause of important adverse effects of cancer treatment (Wardell *et al.*, 2003). It may be helpful to develop agents that specifically protect the mitochondrial genome during radiotherapy and chemotherapy of malignant disease, as well as treatment of HIV infection.

2.7 DIAGNOSIS OF MITOCHONDRIAL DISORDERS

Mitochondrial DNA disorders were found to affect 6.57 per 100,000 individuals in the adult population of working age, and 7.59 per 100,000 unaffected adults and children are at risk of developing mtDNA disease. As a group, the prevalence of mtDNA disorders is comparable with Huntington's disease, which affects 6.4 per 100,000 individuals and is more common than Duchenne's dystrophy, which affects 3.2 per 100,000 individuals (Chinnery *et al.*, 2000). However, RC disorders are clinically, biochemically and molecularly heterogeneous (Bernier *et al.*, 2002) and difficult to diagnose. Especially in children (0-16 years), the clinical presentation and course have enormous variations (Wolf and Smeitink, 2002). Although these disorders may present with various neurological

features (e.g. encephalopathy, myopathy and hearing loss), non-neurologic presentations occur in over 30% of paediatric patients (Munnich *et al.*, 1996).

Evidence from at least two relatively independent types of investigation (i.e. clinical, histological, biochemical or molecular) is usually required to establish a definitive diagnosis. These investigations may include determination of cell redox status (e.g. lactate and lactate/pyruvate ratio), numerical or structural abnormalities of mitochondria in tissue biopsies, molecular studies, enzyme histochemistry and measurement of RC function, including spectrophotometric assays of the individual RC complexes. However, lack of standardised diagnostic criteria complicates the evaluation of diagnostic methodologies and interpretation of published reports. The confirmation or exclusion of an RC disorder is therefore often problematic for clinicians (Bernier *et al.*, 2002).

In 1996 diagnostic criteria were proposed to classify patients' likelihood of having an RC disorder as "definite", "probable", "possible" or "unlikely" based on the presence of certain major and minor criteria for clinical features and laboratory markers (Walker *et al.*, 1996). However, these criteria were suggested on the basis of experience at an adult neurology clinic and were therefore modified to allow for paediatric clinical and histological features and for more sensitive coding of RC enzyme and functional studies (Bernier *et al.*, 2002). These modified adult criteria contained biochemical parameters, but did not indicate optimal biochemical assays and cut-off values, posing difficulties in interpretation of results (Wolf and Smeitink, 2002).

Wolf and Smeitink (2002) developed the mitochondrial disease criteria (MDC), comprising clinical symptoms, metabolic as well as imaging findings and morphological and biochemical investigations of skeletal muscle. In order to avoid a disproportionate contribution from a multitude of parameters in one of these categories, each category (clinical presentation, metabolic investigation and imaging, and muscle histopathology) can contribute a maximum of four points; thus a total of 12 points. Clinical criteria are divided into three main groups: muscular presentation (maximum two points), CNS (maximum two points) and multisystem involvement (maximum three points) with an overall maximum of four points (NEUROLOGY, 2003). Patients are allocated to one of these three groups according to their predominant clinical features. Oxidation rates, carbon-14 labelled substrates (pyruvate, malate and succinate), ATP and phosphocreatine production rate and single enzyme activities (complex I to IV) are considered biochemical criteria and the results of all of these are combined to determine the biochemical likelihood of developing

an RC disorder (Wolf and Smeitink, 2002). All available information for a given patient is collected and scored and results in assignment to one of four levels similar to the adult criteria of Walker *et al.* (1996), as presented in Table 2.3.

Table 2.3: Evaluation of the mitochondrial disease criteria score

MDC score	RC disorder evaluation
1	Unlikely
2 – 4	Possible
5 – 7	Probable
8 – 12	Definite

MDC = mitochondrial disease criteria and RC = respiratory chain. Compiled from NEUROLOGY (2003).

2.8 MOLECULAR INVESTIGATIONS OF MITOCHONDRIAL DISORDERS

In an investigation by Prosser (2001) of 25 clinically well characterised mitochondrial myopathy patients, only one individual (patient 411) harboured the reported MELAS A3243G mutation and only one patient was observed with a novel G3277A alteration in the $tRNA^{Leu(UUR)}$ gene, even though the clinical phenotype was indicative of LS. None of the 23 other reported mutations in these three regions were present in any of the patients investigated. Prosser (2001) concluded that the genetic aetiology of mitochondrial myopathies in the South African population is different from those reported in other international studies, although the clinical phenotypes are often similar or the same. Therefore, whole mitochondrial genome screening was suggested to prevent false negative diagnoses of mitochondrial myopathy patients. However, Prosser (2001) investigated mtDNA from whole blood or fibroblast samples. As result of mitotic segregation these samples might have been negative for mtDNA mutations, even though mutations could have been present in other tissues. Although any organ or tissue may be affected, neuromuscular symptoms are most commonly associated with RC disorders (Wolf and Smeitink, 2002). Therefore, if mutant mtDNA is present, muscle samples are probably more likely to harbour the mtDNA mutations and are preferred for mtDNA investigations.

Prosser (2001) demonstrated that direct sequencing could replace restriction fragment length polymorphism (RFLP) analyses as a more sensitive and specific strategy to screen for mutations in the mitochondrial genome. With RFLP analyses of the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and ATPase 6 mitochondrial genes, 10 mtDNA mutations could be detected. However, sequencing of these three genes facilitated the investigation of the 10 mtDNA

mutations originally screened for, as well as an additional 14 mtDNA mutations detectable by the sequencing strategy for a total of 24 mtDNA mutations under investigation (Prosser, 2001). Especially for screening of entire genomes as performed for four patients in the investigation reported here, automated cycle sequencing is indispensable. Therefore, instead of RFLP analyses, cycle sequencing was performed as detection strategy in this investigation.

2.9 AIMS OF THE INVESTIGATION

The long-term objective of the project is to investigate the genetic aetiology of mitochondrial disorders in the South African population. Furthermore, it is designed to assist in the diagnosis of these patients for informed genetic counselling of their family members and the development of possible future therapeutic strategies.

2.9.1 Specific aims

The objective of this investigation was to screen at least 35 samples from South African patients of different populations, who had been clinically diagnosed with suspected mitochondrial disorders, for the presence of the mtDNA mutations most commonly associated with MELAS, MERRF and LS. Through screening of the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and ATPase 6 mitochondrial genes, the presence of 24 pathogenic mutations was to be investigated. Preferentially muscle samples were analysed; however, if not available, mtDNA isolated from whole blood was investigated.

For a minimum of four additional patients with MDC scores above eight (NEUROLOGY, 2003) and previously investigated by Prosser (2001), outstanding mtDNA regions were to be screened for a complete mitochondrial genome analysis. These patients were negative for the 24 reported mutations in the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and ATPase 6 mitochondrial genes, although they had severe clinical syndromes.

CHAPTER THREE

MATERIALS AND METHODS

Many of the patients examined at the Paediatric Neurology Research Clinic (at Pretoria Academic Hospital) presented with clinical phenotypes suggestive of MELAS, MERRF and LS. In this study three genes often associated with these syndromes, namely the mitochondrial $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and the ATPase 6 genes, were screened for alterations. Four severely affected patients, with MDC scores of eight, were previously investigated without observing pathogenic alterations in these three mitochondrial genes. Sequence analysis of the full mitochondrial genome was performed for these four patients.

The protocols followed in this investigation were performed according to published methods or according to those acquired with the respective kits utilised. Modifications to the published protocols are indicated throughout.

3.1 ETHICAL APPROVAL

Ethical approval has been obtained for this project from the Ethics Committee of the Potchefstroom University for Christian Higher Education (PUCHE) under the title "Mitochondrial DNA (mtDNA) mutations in patients with a suspected mitochondrial disorder in the South African context" with approval number 02M02. Legal informed consent has been obtained from either the individuals participating in this research project, or from their legal guardians.

3.2 PATIENT POPULATION

Samples were collected continuously from patients referred to the Centre for Genome Research. These patients were clinically diagnosed with a possible mitochondrial disorder by Dr. I. Smuts, Department of Paediatrics, Faculty of Health Sciences, University of Pretoria. The patient population investigated had different ethnic origins and included mainly individuals younger than 16 years of age. Muscle samples were analysed preferentially. However, since obtaining samples of such a nature requires invasive procedures, it was not feasible to collect a muscle sample from each patient. In these

cases DNA was isolated from whole blood. If available, muscle tissue and whole blood from the same patient were analysed to enable comparative analysis.

3.3 ISOLATION OF GENOMIC DNA

Genomic DNA (gDNA) was extracted from whole blood utilising the Wizard®[1] Genomic DNA Purification Kit and from muscle with the QIAamp®[2] DNA Mini Kit (See 3.3.1 and 3.3.2). The concentration of the isolated DNA was estimated by measuring the absorbance/optical density at a wavelength of 260 nanometres (nm) by ultraviolet (UV) spectrophotometry with an Eppendorf[3] BioPhotometer. For double-stranded DNA the reading obtained is multiplied with one absorbancy unit, which equals 50 $ng.\mu l^{-1}$, and with the dilution factor (Sambrook and Russell, 2001), as presented in Equation 3.1. The absorbance at 280 nm (A_{280}) was also recorded to determine the concentration of proteins in the isolated DNA. The degree of protein contamination is indicated by the A_{260}/A_{280} ratio, which is acceptable for a value equal to or above 1.8.

Equation 3.1: Calculation of the DNA concentration from the absorbance at 260 nm

$$[\text{double-stranded DNA}] = A_{260} \times (50\ ng.\mu l^{-1} \times \text{dilution factor})$$

[DNA] = DNA concentration and A_{260} = absorbance of samples at 260 nm. Adapted from Sambrook and Russell (2001).

3.3.1 Isolation of genomic DNA from whole blood

The Wizard® Genomic DNA Purification Kit was utilised to extract gDNA from blood samples, according to the manufacturer's instructions. The blood was collected in ethylenediamine tetra-acetic acid (EDTA) tubes and the gDNA was extracted as discussed below.

Three hundred microlitres (µl) of blood was lysed with 900 µl cell lysis solution. The reaction was incubated at room temperature for 10 minutes (min), while mixing by inversion. The sample was centrifuged at 13,000 x gravity (x g) for 30 seconds (s), the supernatant discarded and the white pellet vortexed until the white blood cells were completely resuspended in the remaining supernatant.

[1] Wizard® is a registered trademark of Promega Corporation, Madison, WI, U.S.A.
[2] QIAamp® is a registered trademark of QIAGEN, Clifton Hill, Victoria, Australia.
[3] Eppendorf® is a registered trademark of Eppendorf AG, Hamburg, Germany.

The white cells were lysed by the addition of 300 µl nuclei lysis solution followed by incubation at 37 degrees centigrade (°C) for 30 min. Subsequently the sample was allowed to cool down to room temperature. Protein precipitation was achieved by the addition of 100 µl protein precipitation solution followed by vigorous vortexing. Afterwards the sample was centrifuged at 13,000 x g for 3 min.

DNA was precipitated by transferring the supernatant to a 1.5 ml microcentrifuge tube containing 300 µl 100% isopropanol, followed by inversion of the tube until the white thread-like strands of DNA became visible. The gDNA was recovered via centrifugation for 1 min at 13,000 x g and all remaining salts were washed out with 300 µl 70% ethanol (EtOH). Following another centrifugation step at 13,000 x g for 1 min, the pellet was air-dried for 10 to 15 min. Finally the DNA pellet was rehydrated in 100 µl DNA rehydration solution, followed by incubation overnight at 4°C. The concentration of the isolated DNA was estimated by spectrophotometry as discussed earlier in paragraph 3.3. The 50 $ng.\mu l^{-1}$ DNA working solution was stored at 4°C and the stock solution at -20°C.

3.3.2 DNA isolation from muscle tissue

It is essential for biochemical OXPHOS analyses that muscle samples be transported and homogenised in SETH-medium consisting of 0.25 M (molar) sucrose, 2 millimolar (mM) potassium ethylenediamine tetra-acetic acid (K_2EDTA), 10 mM tris-(hydroxymethyl)-aminomethan (Tris®) and double distilled water (ddH_2O), set at pH 7.4. Samples of 100 milligram (mg) muscle were homogenised in 1 millilitre (ml) SETH-medium (10% weight per volume) in a tight-fitting Potter-Elvehjam homogeniser. The homogenised solution was centrifuged at 600 x g for 10 min, the supernatant was aspirated and sent for biochemical analysis performed by Dr. F.H. van der Westhuizen, Biochemistry, PUCHE. The homogenised pellet was stored at -70°C until DNA extraction could be performed.

The tissue protocol of the QIAamp® DNA Mini Kit was utilised, according to the manufacturer's instructions, for extraction of gDNA from homogenised as well as non-homogenised muscle tissue. Initially 180 µl lysis buffer ATL was added to 100 mg of homogenised pellet or 50 mg of non-homogenised muscle in a 1.5 ml microcentrifuge tube. The non-homogenised muscle was divided into small pieces prior to the addition of lysis buffer. Subsequently, 20 µl of 20 $mg.ml^{-1}$ Proteinase K solution was added, the

sample vortexed and incubated overnight in a waterbath at 56°C, with gentle agitation to allow complete digestion of cellular constituents.

Following the incubation step, 200 µl of lysis buffer AL was added and the samples were vortexed for 15 s and incubated for 10 min at 70°C. Thereafter, 400 µl EtOH was added and the samples were vortexed for 15 s. If samples of only 50 mg homogenised pellet or 25 mg non-homogenised muscle were available, only 200 µl EtOH was added.

After mixing, the homogeneous solution was applied to the QIAamp® spin column (in the 2 ml collection tube provided) and centrifuged at 6,000 x g for 1 min. The spin column was placed into a clean 2 ml collection tube and the filtrate discarded. Subsequently, 500 µl of wash buffer AW1 was pipetted onto the column and it was again centrifuged at 6,000 x g for 1 min. Again the QIAamp® spin column was placed into a clean 2 ml collection tube and the filtrate was discarded. The second wash step followed with the addition of 500 µl of buffer AW2 into the spin column and centrifugation at 20,000 x g for 3 min. To eliminate possible buffer AW2 carryover, the spin column was once again placed into a new collection tube, the filtrate was discarded and the centrifugation step at 20,000 x g repeated. Subsequently the QIAamp® spin column was placed into a clean 1.5 ml microcentrifuge tube and the filtrate discarded.

Finally for the elution of the gDNA, 200 µl of elution buffer AE was pipetted into the spin column. It was incubated for 5 min at room temperature and then centrifuged at 6,000 x g for 1 min. This final step was repeated, yielding a final volume of 400 µl. As with DNA isolated from blood, the concentration of the isolated DNA was estimated by spectrophotometry. The 50 ng.μl^{-1} DNA working solution was stored at 4°C and the stock solution at -20°C.

3.4 POLYMERASE CHAIN REACTION (PCR)

Regions of the mitochondrial genome under investigation were amplified through PCR, utilising a modified version of the protocol described by Mullis *et al.* (1986). The amplification of the fragments was performed in a Thermo Hybaid®[1] Multiblock System 0.5S thermocycler. The standard PCR reaction mixture contained the following in a final reaction volume of 12.5 µl: 1 x PCR buffer [50 mM potassium chloride, 10 mM

[1] Thermo Hybaid® is a registered trademark of Hybaid Limited, Ashford, Middlesex, United Kingdom.

Tris®-hydrochloride (HCl) of pH 9.0, 1.5 mM magnesium chloride ($MgCl_2$) and 0.1% Triton® X-100], 200 micromolar (μM) of each 2'-deoxynucleotide-5-triphosphate [2'-deoxyadenosine-5'-triphosphate (dATP), 2'-deoxycytidine-5'-triphosphate (dCTP), 2'-deoxyguanosine-5'-triphospate and 2'-deoxythymidine-5'-triphosphate], 10 picomoles (pmol) each of the respective forward and reverse primers, 1 unit (U) of Taq DNA polymerase, 100 nanograms (ng) of gDNA and ddH_2O.

3.4.1 Amplification of the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and the ATPase 6 mitochondrial genes

In general PCR conditions used for amplification are optimised from a temperature 2°C below the calculated mean melting temperature (T_m) of the primer set for a specific reaction. This lower temperature, the estimated annealing temperature (T_a) of the primer set, serves as a starting point from where the annealing temperature can be increased or decreased to achieve optimal amplification. The equation utilised to calculate T_a is presented in Equation 3.2.

Equation 3.2: Calculation of the estimated annealing temperature of primer sets

$$T_a = \text{Mean } T_m - 2°C$$

T_a = estimated annealing temperature of primer set in °C and Mean T_m = the mean melting temperature of the primer set in °C.

Oligonucleotide Properties Calculator software (OLIGONUCLEOTIDE PROPERTIES CALCULATOR, 2003) was utilised to determine the T_m of each primer set through nearest neighbour and thermodynamic calculations by using the values published by Sugimoto *et al.* (1996). Another equation, often applied for practical purposes, is that of Thein and Wallace (1986) as presented in Equation 3.3. This equation utilises the adenine/thymine (A + T) and guanine/cytosine (G + C) content to determine the T_m and does not include complex thermodynamic parameters. The Thein and Wallace (1986) equation is, therefore, an easy rule-of-thumb that can be applied to determine the starting point from which optimisation should be initiated.

Equation 3.3: Calculation of the primer melting temperature

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

T_m = calculated melting temperature, A + T = adenine and thymine content of primer sequence and G + C = guanine and cytosine content of primer sequence. Adapted form Thein and Wallace (1986).

The primers utilised for amplification of the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and the ATPase 6 mitochondrial genes, as conveyed by Wallace (1999), are presented in Table 3.1.

Table 3.1: Sequence of PCR primers utilised for amplification of the mitochondrial $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and ATPase 6 genes

Primer	Primer sequence	T_m (°C)	Mean T_m (°C)	T_a (°C)	Size (bp)
$tRNA^{Leu(UUR)}$					
F: ND1-FOR-B	5′-ccc gat ggt gca gcc gc-3′	57	55.5	53.5	364
R: 3.3-REV	5′-gca tta gga atg cca ttg cg-3′	54			
$tRNA^{Lys}$					
F: 8155FOR	5′-ggt ata cta cgg tca atg ctc t-3′	52	52.5	50.5	212
R: 8380REV2	5′-att tag ttg ggg cat ttc act cta-3′	53			
ATPase 6					
F: 952x10BF	5′-cct agc cat ggc cat cc-3′	50	49	47	1,031
R: 9859-REV	5′-cag ata gtg agg aaa gtt ga-3′	48			

T_m = calculated melting temperature in °C according to OLIGONUCLEOTIDE PROPERTIES CALCULATOR (2003), Mean T_m = the mean melting temperature in °C for the primer set, T_a = estimated annealing temperature in °C for the primer set, Size = amplified fragment size in bp, F = forward primer and R = reverse primer. Primers designed in the laboratory of Wallace (1999).

3.4.1.1 Amplification of the $tRNA^{Leu(UUR)}$ gene

The region from nt 3007 to 3370 was amplified, including the $tRNA^{Leu(UUR)}$ gene from nt 3230 to 3304 (MITOMAP, 2003) as presented in Table 3.2. The ND1-FOR-B forward and 3.3-REV reverse primers (Wallace, 1999), as presented in Table 3.1, were utilised. In the sequence presented in Table 3.2 the position of the annealing sites of the primers and the 13 most common mutation loci associated with the $tRNA^{Leu(UUR)}$ gene are indicated.

Table 3.2: Partial sequence of the mitochondrial $tRNA^{Leu(UUR)}$ gene, including 13 common mutation loci and positions of PCR primers

Nucleotide	Sequence
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	ag**A**gcccgg**T** **AA**t**C**g**C**ata**A** aac**T**taaaac **T**ttacagtca gaggttc**A**at **T**cctcttctt
3301	a**AC**aacatac ccatggccaa cctcctactc ctcattgtac ccattctaat cgcaatggca
3361	ttcctaatgc ttaccgaacg aaaaattcta ggctatatac aactacgcaa aggccccaac

The position of the forward primer is indicated by the underlined text (xxx) and the position of the reverse primer is indicated by double underlined text (yyy). The 13 mutation loci within this region are indicated in bold capital letters. The $tRNA^{Leu(UUR)}$ gene is coloured blue. Adapted from MITOMAP (2003).

3.4.1.2 Amplification of the $tRNA^{Lys}$ gene

The $tRNA^{Lys}$ gene stretches from nt 8295 to 8364 (MITOMAP, 2003). However, only the region from nt 8155 to 8380 was amplified, as presented in Table 3.3. The 8155FOR forward and 8380REV2 reverse primers (Wallace, 1999), as presented in Table 3.1, were utilised. In the sequence presented in Table 3.3 the annealing site positions of the primers and the six most common mutation loci associated with the $tRNA^{Lys}$ gene are indicated.

Table 3.3: Partial sequence of the mitochondrial $tRNA^{Lys}$ gene, including six common mutation loci and positions of the PCR primers

Nucleotide	Sequence
8101	agatgcaatt cccggacgtc taaaccaaac cactttcacc gctacacgac cgggggtata
8161	ctacggtcaa tgctctgaaa tctgtggagc aaaccacagt ttcatgccca tcgtcctaga
8221	attaattccc ctaaaaatct ttgaaatagg gcccgtattt accctatagc acccccctcta
8281	ccccctctag agccc**A**ctgt aaagctaact ta**G**cattaac cttttaa**G**tt aaagattaag
8341	a**G**a**A**ccaaca cctct**T**taca gtgaaatgcc ccaactaaat actaccgtat ggcccaccat

The position of the forward primer is indicated by the underlined text (xxx) and the position of the reverse primer is indicated by double underlined text (yyy). The six mutation loci within this region are indicated in bold capital letters. The $tRNA^{Lys}$ gene is coloured blue. Adapted from MITOMAP (2003).

3.4.1.3 Amplification of the ATPase 6 gene

The region surrounding the ATPase 6 gene from nt 8829 to 9859 was amplified, although the gene stretches from nt 8527 to 9207 (MITOMAP, 2003) as presented in Table 3.4. The 952x10BF forward and 9859-REV reverse primers (Wallace, 1999), as presented in Table 3.1, were utilised. In the sequence presented in Table 3.4 the position of the annealing sites of the primers and the three most common mutation loci associated with the ATPase 6 gene are indicated.

Table 3.4: Partial sequence of the mitochondrial ATPase 6 gene, including three common mutation loci and positions of the PCR primers

Nucleotide	Sequence
8821	tctataaacc tagccatggc catcccctta tgagcgggca cagtgattat aggctttcgc
8881	tctaagatta aaaatgccct agcccacttc ttaccacaag gcacacctac accccttatc
8941	cccatactag ttattatcga aaccatcagc ctactcattc aaccaatagc ccTggccgta
9001	cgcctaaccg ctaacattac tgcaggccac ctactcatgc acctaattgg aagcgccacc
9061	ctagcaatat caaccattaa ccttccctct acacttatca Tcttcacaat tctaattcta
9121	ctgactatcc tagaaatcgc tgtcgcctta atccaagcct acgttttcac acttcTagta
9181	agcctctacc tgcacgacaa cacataatga cccaccaatc acatgcctat catatagtaa
9241	aacccagccc atgaccccta acaggggccc tctcagccct cctaatgacc tccggcctag
9301	ccatgtgatt tcacttccac tccataacgc tcctcatact aggcctacta accaacacac
9361	taaccatata ccaatgatgg cgcgatgtaa cacgagaaag cacataccaa ggccaccaca
9421	caccacctgt ccaaaaaggc cttcgatacg ggataatcct atttattacc tcagaagttt
9481	ttttcttcgc aggatttttc tgagcctttt accactccag cctagcccct acccccaat
9541	taggagggca ctggcccccа acaggcatca ccccgctaaa tcccctagaa gtcccactcc
9601	taaacacatc cgtattactc gcatcaggag tatcaatcac ctgagctcac catagtctaa
9661	tagaaaacaa ccgaaaccaa ataattcaag cactgcttat tacaatttta ctgggtctct
9721	attttaccct cctacaagcc tcagagtact tcgagtctcc cttcaccatt tccgacggca
9781	tctacggctc aacatttttt gtagccacag gcttccacgg acttcacgtc attattggct
9841	caactttcct cactatctgc ttcatccgcc aactaatatt tcactttaca tccaaacatc

The position of the forward primer is indicated by the underlined text (xxx) and the position of the reverse primer is indicated by double underlined text (yyy). The three mutation loci within this region are indicated in bold capital letters. The region of the ATPase 6 gene investigated, is coloured blue. Adapted from MITOMAP (2003).

The PCR reaction conditions for amplification of the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and ATPase 6 genes were previously optimised by Prosser (2001) at an annealing temperature of 55 °C with 1.5 mM $MgCl_2$ in the reaction mixture. The cycling parameters for denaturation, annealing and extension of the PCR products are presented in Table 3.5.

Table 3.5: PCR conditions for amplification of mitochondrial DNA

PCR step	# of cycles	Action	Temperature	Duration
1	1	Denaturation	94°C	10 min
2	30	Denaturation	94°C	30 s
		Annealing	x°C	30 s
		Extension	72°C	30 s
3	1	Extension	72°C	7 min

= number of cycles, x = 55°C for amplification $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and ATPase 6, but had different values for whole mitochondrial genome amplification as presented in Table 4.1 (page 54).

3.4.2 Amplification of the full-length mitochondrial genome

The primers published by Maca-Meyer *et al.* (2001) for sequencing the full mitochondrial genome of 42 individuals from different haplogroups, including sub-haplogroup L1-3, were

utilised. The genome was amplified in 32 overlapping fragments of ca. 600 nucleotides each. The primer sets, their calculated and mean melting temperatures, estimated annealing temperatures and amplified fragment lengths are presented in Table 3.6.

Table 3.6: Sequence of PCR primers for amplification of the whole mitochondrial genome

Primer	Primer Sequence	T_m	Mean T_m	T_a	Size
F1:mtL16340	5′-agc cat tta ccg tac ata gca ca-3′	55	54	52	681
R1:mtH408	5′-tgt taa aag tgc ata ccg cca-3′	53			
F2:mtL382	5′-caa aga acc cta aca cca gcc-3′	53	51.5	49.5	603
R2:mtH945	5′-ggg agg ggg tga tct aaa ac-3′	50			
F3:mtL923	5′-gtc aca cga tta acc caa gtc a-3′	52	53	51	607
R3:mtH1487	5′-gta tac ttg agg agg gtg acg g-3′	54			
F4:mtL1466	5′-gag tgc tta gtt gaa cag ggc c-3′	56	54.5	52.5	629
R4:mtH2053	5′-tta gag ggt tct gtg ggc aaa-3′	53			
F5:mtL2025	5′-gcc tgg tga tag ctg gtt gtc c-3′	57	54	52	609
R5:mtH2591	5′-gga aca agt gat tat gct acc t-3′	51			
F6:mtL2559	5′-cac cgc ctg ccc agt gac aca t-3′	60	55.5	53.5	591
R6:mtH3108	5′-tcg tac agg gag gaa ttt gaa-3′	51			
F7:mtL3073	5′-aaa gtc cta cgt gat ctg agt tc-3′	53	53	51	640
R7:mtH3670	5′-ggc gta gtt tga gtt tga tgc-3′	53			
F8:mtL3644	5′-gcc acc tct agc cta gcc gt-3′	58	54.5	52.5	623
R8:mtH4227	5′-atg ctg gag att gta atg ggt-3′	51			
F9:mtL4210	5′-cca ctc acc cta gca tta ctt a-3′	52	53.5	51.5	625
R9:mtH4792	5′-act cag aag tga aag ggg gct a-3′	55			
F10:mtL4750	5′-cca ata cta cca atc aat act c-3′	47	51	49	599
R10:mtH5306	5′-ggt gat ggt ggc tat gat ggt g-3′	55			
F11:mtL5278	5′-tgg gcc att atc gaa gaa tt-3′	50	51	49	593
R11:mtH5832	5′-gac agg ggt tag gcc tct tt-3′	52			
F12:mtL5781	5′-agc ccc ggc agg ttt gaa gc-3′	59	56.5	54.5	626
R12:mtH6367	5′-tgg ccc cta aga tag agg aga-3′	54			
F13:mtL6337	5′-cct gga gcc tcc gta gac ct-3′	57	55	53	601
R13:mtH6899	5′-gca ctg cag cag atc att tc -3′	53			
F14:mtL6869	5′-ccg gcg tca aag tat tta gc-3′	53	53	51	578
R14:mtH7406	5′-ggg ttc ttc gaa tgt gtg gta g-3′	53			
F15:mtL7379	5′-aga aga acc ctc cat aaa cct g-3′	52	53.5	51.5	580
R15:mtH7918	5′-aga tta gtc cgc cgt agt cg-3′	55			
F16:mtL7882	5′-tcc ctc cct tac cat caa atc a-3′	52	51.5	49.5	506
R16:H8345	5′-ttt cac tgt aaa gag gtg ttg g-3′	51			

Table 3.6: continued...

Primer	Primer Sequence	T_m	Mean T_m	T_a	Size
F17:mtL8299	5′-acc ccc tct aga gcc cac tg-3′	56	55	53	603
R17:mtH8861	5′-gag cga aag cct ata atc act g-3′	54			
F18:mtL8799	5′-ctc gga ctc ctg cct cac tca-3′	58	55	53	638
R18:mtH9397	5′-gtg gcc ttg gta tgt gct tt-3′	52			
F19:mtL9362	5′-ggc cta cta acc aac aca cta-3′	51	52	50	609
R19:H9928	5′-aac cac atc tac aaa atg cca gt-3′	53			
F20:mtL9886	5′-tcc gcc aac taa tat ttc act t-3′	51	50	48	617
R20:mtH10462	5′-aat gag ggg cat ttg gta aa-3′	49			
F21:mtL10403	5′-aaa gga tta gac tga acc gaa-3′	50	50.5	48.5	612
R21:mtH10975	5′-cca tga ttg tga ggg gta gg-3′	51			
F22:mtL10949	5′-ctc cga ccc cct aac aac cc-3′	54	53	51	617
R22:mtH11527	5′-caa gga agg ggt agg cta tg-3′	52			
F23:mtL11486	5′-aaa act agg cgg cta tgg ta-3′	51	51	49	629
R23:mtH12076	5′-gga gaa tgg ggg ata ggt gt-3′	51			
F24:mtL12028	5′-ggc tca ctc acc cac cac att-3′	55	54	52	615
R24:mtH12603	5′-acg aac aat gct aca ggg atg-3′	53			
F25:mtL12572	5′-aca acc cag ctc tcc cta ag-3′	53	52.5	50.5	591
R25:mtH13124	5′-att ttc tgc tag ggg gtg ga-3′	52			
F26:mtL13088	5′-agc cct act cca ctc aag cac-3′	56	53	51	618
R26:mtH13666	5′-agg gtg ggg tta ttt tcg tt-3′	50			
F27:mtL13612	5′-aag cgc cta tag cac tcg aa-3′	56	52.5	50.5	614
R27:mtH14186	5′-tgg ttg aac att gtt tgt tgg-3′	49			
F28:mtL14125	5′-tct ttc ttc ttc cca ctc atc c-3′	52	52	50	602
R28:mtH14685	5′-cat tgg tcg tgg ttg tag tcc-3′	52			
F29:mtL14650	5′-ccc cat tac taa acc cac act c-3′	52	52	50	604
R29:mtH15211	5′-ttg aac tag gtc tgt ccc aat g-3′	52			
F30:mtL15162	5′-ctc ccg tga ggc caa ata tc-3′	53	54	52	597
R30:mtH15720	5′-gtc tgc ggc tag gag tca at-3′	55			
F31:mtL15676	5′-tcc cca tcc tcc ata tat cc-3′	49	50	48	524
R31:mtH16157	5′-tga tgt gga ttg ggt ttt tat gta-3′	51			
F32:mtL15996	5′-ctc cac cat tag cac cca aag c-3′	56	54	52	446
R32:mtH16401	5′-tga ttt cac gga gga tgg tg-3′	52			

T_m = calculated melting temperature in °C according to OLIGONUCLEOTIDE PROPERTIES CALCULATOR (2003), Mean T_m = the mean melting temperature in °C for the primer set, T_a = estimated annealing temperature in °C for the primer set, Size = amplified fragment size in bp, F = forward primer and R = reverse primer. Primer sequences were obtained from Maca-Meyer *et al.* (2001).

The PCR reaction conditions for amplification of the 32 overlapping fragments for whole mitochondrial genome sequencing had to be optimised. This was achieved by altering the

annealing temperature and $MgCl_2$ concentration until the PCR product obtained was optimal, as discussed in paragraph 4.1.2.

3.5 AGAROSE GEL ELECTROPHORESIS

To confirm the success of the PCR reaction, the amplified products (amplicons) were separated via agarose gel electrophoresis. The lengths of the fragments were estimated by comparison to fragments of the 100 bp DNA Ladder from Promega®[1]. The PCR products varied between 100 and 2,000 nucleotides (i.e. 364, 212, ca. 600 and 1,031 nucleotides) in size. Therefore, a 2% agarose gel was utilised, containing 1 x TBE buffer [89.15 mM Tris® (pH 8.1), 88.95 mM boric acid, 2.5 mM di-sodium ethylenediamine tetra-acetic acid (Na_2EDTA)] and 0.5 $\mu g.ml^{-1}$ ethidium bromide. A volume of 3 µl was loaded onto the gel containing 2 µl PCR amplified sample and ca. 1 µl gel-loading buffer (0.02% orange G and 50% glycerol in water). The applied voltage [ca. 10 volt (V) per centimetre (cm)] was adjusted according to the size of the gel. In general a voltage of 100 V was applied for mini and midi gels. The amplified fragments were visualised through UV fluorescence at 312 nm and were photographed utilising a photo documentation system.

3.6 PCR PRODUCT PURIFICATION

Subsequent to gel electrophoresis, the PCR products were purified from the primers, nucleotides, polymerases, salts and mineral oil in the PCR reaction mixture, utilising the QIAquick®[2] PCR Purification Kit according to the manufacturer's instructions. To ensure efficient binding of the double stranded PCR product to the silica gel membrane, five volumes of binding or PB buffer was mixed with one volume of PCR sample. The sample was loaded onto the QIAquick® column, followed by centrifugation for 1 min at 13,000 x g. After centrifugation, the elute was discarded. Subsequently the column was washed with 750 µl wash or PE buffer and centrifuged for 1 min at 13,000 x g to remove contaminants binding to the column. The second elute was also discarded and the sample was again centrifuged for another min at 13,000 x g, to remove the residual EtOH at the top of the column. After centrifugation, the column was transferred to a clean 1.5 ml microcentrifuge tube and the PCR product was eluted with 50 µl elution or EB buffer (10 mM Tris®-HCl, pH 8.5) added to the centre of the membrane. This elution took approximately 1 min, after which the sample was centrifuged for an additional 1 min at 13,000 x g. The concentration

[1] Promega® is a registered trademark of Promega Corporation, Madison, WI, U.S.A.
[2] QIAquick® is a registered trademark of QIAGEN Pty. Ltd., Australia.

of the ca. 50 µl final volume was determined spectrophotometrically as discussed in paragraph 3.3.

3.7 CYCLE SEQUENCING

Sequencing of the mtDNA genome was performed utilising the ABI Prism®[1] BigDye™[2] Terminator version 3.1 (v3.1) Ready Reaction Cycle Sequencing Kit, according to the principles originally described by Sanger *et al.* (1977). As the four 2',3'-dideoxynucleotide triphosphates are labelled with different fluorophores, the terminating base can be distinguished and detected. Amplitaq®[3] DNA Polymerase FS was utilised as the sequencing enzyme as it contains a point mutation in the active site leading to less discrimination against dideoxynucleotides than ordinary *Thermus aquaticus* (Taq) DNA polymerase. Amplitaq® DNA Polymerase FS also has a second mutation in the amino terminal domain that virtually eliminates the 5' to 3' nuclease activity.

In order to perform cost-efficient cycle sequencing of the amplified fragments, quarter reactions were utilised. Two microlitres of Big Dye™ 5 x Sequencing Buffer, 2 µl of the Big Dye™ Terminator Ready Reaction Mix [A-, C-, G- and T-BigDye™ Terminators v3.1, deoxynucleoside triphosphates (dATP, dCTP, 2'-deoxyinosine-5'-triphosphate and '2-deoxyuridine-5'-triphosphate), Amplitaq® DNA Polymerase FS, Tris-HCl buffer (pH 9.0) and $MgCl_2$], 3.2 pmol of the primer and a pre-determined amount of the purified PCR product, were mixed for a final reaction volume of 10 µl. The quantity of PCR product added to the cycle sequencing reaction mixture depended on the concentration of the purified PCR product and the fragment size, as presented in Table 3.7. The volume allowed for up to 5 µl of DNA template and 1 µl of primer (3.2 pmol.μl^{-1}) to be utilised per reaction.

[1] ABIPrism® 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.

Table 3.7: Relationship between the quantity of the PCR amplicon utilised in cycle sequencing and the size of the amplicon

PCR Product	Quantity
100 –200 bp	1 – 3 ng
200 – 500 bp	3 – 10 ng
500 –1,000 bp	5 – 20 ng
1000 – 2,000 bp	10 – 40 ng
> 2,000 bp	40 – 100 ng

Adapted from the protocol of the ABI Prism® BigDye™ Terminator v3.0 Ready Reaction Cycle Sequencing Kit.

If the PCR product was not sufficiently concentrated, a more concentrated solution of primer could be utilised to compensate for the additional volume up to a final volume of 10 µl. Cycle sequencing of the PCR product was performed according to the strategy presented in Table 3.8. The first three steps were repeated for 25 cycles.

Table 3.8: Strategy for cycle sequencing

Cycle step	Action	Temperature
1	rapid thermal ramp (1°C.s^{-1})	96°C
	denaturation for 10 s	96°C
2	rapid thermal ramp	50°C
	annealing for 5 s	50°C
3	rapid thermal ramp	60°C
	elongation for 4 min	60°C
4	rapid thermal ramp	4°C
	hold indefinitely	4°C

PCR products were sequenced in the forward direction, but if significant alterations were detected (excluding reported polymorphisms), sequencing was performed in both directions, utilising forward and reverse primers in separate PCR and sequencing reactions. Following the thermal cycling step, the extension products were purified from the unincorporated dye terminators, enzymes and primers, as these could interfere with the base calling during electrophoresis. This was achieved by precipitation of the unincorporated dyes with 95% EtOH and sodium acetate (NaOAc). The precipitation solution was prepared by combining 62.5 µl of 99.8% EtOH, 3.0 µl of 3 M NaOAc (pH 4.6) and 14.5 µl deionised water. The final volume of 80 µl of EtOH/NaOAc was added to the 10 µl reaction mixture and vortexed. The samples were centrifuged for 20 min at 13,000 x g. Immediately after centrifugation the supernatant was removed completely and discarded. Subsequently 250 µl 70% EtOH was added to wash the pellet and the samples

were centrifuged again at 13,000 x g for 10 min. The remaining supernatant was removed and discarded, and the pellet was air-dried at room temperature for 30 min.

The precipitated sequenced samples were electrophoresed and detected by Inqaba Biotechnical Industries (Pty) Ltd on contract. The pellet was resuspended in 6 µl Hi-Di™[1] deionised formamide and 3 µl was injected into a SpectruMedix™[2](SCE2410) Genetic Analysis System sequencer for capillary gel electrophoresis and detection. Depending on the length of the sequenced products, the electrophoresis run time was adjusted. Subsequently the sequences were investigated and aligned, utilising the BioEdit Sequence Alignment Editor (BIOEDIT, 2003) and CLUSTALW computer software (CLUSTALW, 2003). The dye/base relationships and the colours of the BigDye™ terminators v3.1 as they appear on the gel image of the instrument utilised are presented in Table 3.9.

Table 3.9: Colours of bases called on a SpectruMedix™ (SCE2410) Genetic Analysis System sequencer

DNA base	Terminator	Colour of fragment/peak on instrument
A	V3 Dye 2	Green
C	V3 Dye 4	Blue
G	V3 Dye 1	Black
T	V3 Dye 3	Red

A = adenine, C = cytosine, G = guanine and T = thymine.

3.8 MUTATION SCREENING

The nucleotide sequences of the amplified mitochondrial fragments were screened and compared to the RCRS for any alterations, which included polymorphisms and mutations. In analogy to the investigation by Prosser (2001), at least 14 mutations were screened for in the $tRNA^{Leu(UUR)}$ gene, six mutations in the $tRNA^{Lys}$ gene and four mutations in the ATPase 6 gene. Thus, the presence of at least 24 reported mutations were investigated in the proposed investigation, excluding the possibility of novel changes. These 24 reported mutations are presented in Table 3.10.

[1] HiDi™ 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 3.10: Mutations to be investigated in the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and ATPase 6 mitochondrial genes

$tRNA^{Leu(UUR)}$	$tRNA^{Leu(UUR)}$ continued	$tRNA^{Lys}$	ATPase 6
A3243G[1]	T3264C[8]	A8296G[15]	T8993C[21]
T3250C[2]	T3271C[9]	G8313A[16]	T8993G[22]
A3251G[3]	ΔTnt-3271[10]	G8328A[17]	T9101C[23]
A3252G[4]	A3288G[11]	G8342A[18]	T9176C[24]
C3254G[5]	T3291C[12]	A8344G[19]	
C3256T[6]	A3302G[13]	T8356C[20]	
A3260G[7]	C3303T[14]		

1 = Goto *et al.* (1990), 2 = Goto *et al.* (1992), 3 = Sweeney *et al.* (1993), 4 = Morten *et al.* (1993), 5 = Kawarai *et al.* (1997), 6 = Moraes *et al.* (1993), 7 = Zeviani *et al.* (1991), 8 = Suzuki *et al.* (1997), 9 = Goto *et al.* (1991), 10 = Shoffner *et al.* (1995), 11 = Hadjigeorgiou *et al.* (1999), 12 = Goto *et al.* (1994), 13 = Bindoff *et al.* (1993), 14 = Silvestri *et al.* (1994), 15 = Kameoka *et al.* (1998), 16 = Verma *et al.* (1997), 17 = Houshmand *et al.* (1999), 18 = Tiranti *et al.* (1999), 19 = Shoffner *et al.* (1990), 20 = Silvestri *et al.* (1992), 21 = de Vries *et al.* (1993), 22 = Holt *et al.* (1990), 23 = Lamminen *et al.* (1995) and 24 = Thyagarajan *et al.* (1995). Adapted and modified from Prosser (2001).

For the four patients on whom whole mitochondrial genome sequencing was performed the mtDNA was screened for any deviations from the RCRS. Detected alterations were evaluated against the data on MITOMAP (2003), to verify whether they had been reported, whether they resulted in synonymous or non-synonymous changes of proteins or whether they were known polymorphisms. Significant alterations such as reported pathogenic mutations, novel non-synonymous changes or alterations where the functional effect of the change was not known, were confirmed through sequencing of the specific amplicon (derived from separate PCR reactions) in both directions. Reported polymorphisms and synonymous changes detected were not confirmed.

CHAPTER FOUR

RESULTS AND DISCUSSION

The optimisation of experimental protocols are discussed, followed by the results obtained for screening of the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and the ATPase 6 from blood and muscle. Forty-two samples consisting of 18 muscle and 24 blood samples from 34 South African patients were analysed for mutations in the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and ATPase 6 mitochondrial genes. Eight of these muscle and blood samples were obtained from the same individuals. Results of the whole mitochondrial genome sequencing of patients 386, 504, 525 and 1301 are subsequently presented and discussed.

4.1 OPTIMISATION AND APPLICATION OF EXPERIMENTAL PROCEDURES

Commercially available kits were utilised for most of the protocols. However, for DNA isolation from muscle and PCR of fragments for whole mtDNA genome sequencing and cycle sequencing thereof, methods had to be optimised as discussed below.

4.1.1 Isolation of genomic DNA

As discussed in paragraphs 2.8, 2.9 and 3.2, muscle samples were preferentially analysed. Although any organ or tissue may be affected by mtDNA mutations, neuromuscular symptoms are most commonly associated with RC disorders (Wolf and Smeitink, 2002). If heteroplasmy is present in an affected individual, muscle samples are probably more likely to harbour the mtDNA mutations and are, therefore, preferred for mtDNA investigations. However, obtaining a tissue sample requires invasive procedures and for this reason muscle samples were often not available for use in this investigation. In these cases DNA was isolated from whole blood.

Isolation of gDNA from whole blood with the Wizard® Genomic DNA Purification Kit, as described in paragraph 3.3.1, yielded DNA concentrations between 49.5 and 174.3 $ng.\mu l^{-1}$ from 300 µl of whole blood. The degree of protein contamination, indicated by the A_{260}/A_{280} ratio, ranged between 1.5 and 1.7. The A_{260}/A_{280} ratio should preferably have a value of 1.8 or greater. However, the volume of blood samples available was limited as samples were

obtained from paediatric patients. Although not ideal, it was possible to amplify the DNA through PCR, even though the A_{260}/A_{280} values were below 1.8 for all the DNA samples isolated from whole blood. Working dilutions of 50 ng.μl^{-1} were prepared from the stock solution to prevent contamination of the stocks. The stock solutions were stored at -20°C. The biggest risk of using the Wizard® Genomic DNA Purification Kit is decanting the DNA pellet along with the 70% EtOH washing solution after precipitation of the pellet. The QIAamp® DNA Blood Kit is an attractive option here as it uses spin columns, which bind the DNA until eluted into a collection tube, reducing the risk of losing the DNA. However, the kits with spin columns are more expensive than conventional DNA extraction kits. For isolation of DNA from samples that are difficult to obtain, e.g. muscle tissue, the risk of losing the DNA outweighs the cost. Therefore, the QIAamp® DNA Mini Kit, as discussed in paragraph 3.3.2, was utilised.

Different methods of isolating DNA from tissue were applied and evaluated for optimisation purposes. Three different kits, namely the Wizard® Genomic DNA Purification Kit, the Puregene®[1] Kit and QIAamp® DNA Mini Kit (tissue protocol), were investigated utilising bovine muscle tissue. The QIAamp® DNA Mini Kit was determined to be the most reliable for isolation of gDNA from muscle and yielded between 67.6 and 113.9 ng.μl^{-1} from 100 mg of homogenised pellet, or 50 mg of non-homogenised muscle. The A_{260}/A_{280} ratio obtained with the QIAamp® DNA Mini Kit was between 1.7 and 2.0, which was better than with the previously described method. Similarly, working dilutions of 50 ng.μl^{-1} were prepared from the stock gDNA solution to prevent contamination, and the stock solutions were stored at -20°C.

4.1.2 Polymerase chain reaction

A high copy number of mitochondrial genes exist in a cell, as discussed in paragraph 2.3, compared to generally only two copies of nuclear genes per cell. For this reason mtDNA generally amplifies easily via PCR. The $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and ATPase 6 as well as their flanking regions were amplified to ensure a full-length product. The forward and reverse primers utilised, had been optimised previously (Prosser, 2001). Despite the different estimated annealing temperatures of the different primer sets, as presented in Table 3.1, sufficient amplification was achieved at an annealing temperature of 55°C for all three primer sets, along with a $MgCl_2$ concentration of 1.5 mM.

[1] Puregene® is a registered trademark of Gentra Systems Inc., Minneapolis, MN, U.S.A.

For whole mtDNA genome sequencing this genome was amplified in 32 overlapping fragments utilising the primers described by Maca-Meyer *et al.* (2001), as discussed in paragraph 3.4.2. Maca-Meyer *et al.* (2001) utilised these primers to amplify and sequence 42 individuals from different European as well as African haplogroups. As manufacturing of the 32 primer sets was an expensive endeavour, the primers were evaluated beforehand with NET PRIMER (2003) software for possible secondary structures e.g. dimers, hairpins and palindromes. Some of the primers formed dimers, but none of them formed hairpins or palindromes.

In contrast to the amplification of the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and ATPase 6 genes, PCR with the 32 primer sets, as described by Maca-Meyer *et al.* (2001), could not all be performed under the same conditions. The different primer sets had to be optimised and the annealing temperatures and $MgCl_2$ concentrations are listed in Table 4.1.

Table 4.1: Optimised PCR conditions for whole mtDNA genome amplification

Primer	Annealing Temp.	[$MgCl_2$]	Primer[1]	Annealing Temp.	[$MgCl_2$]
F1:mtL16340	55	1.5	F17:mtL8299	55	1.5
R1:mtH408			R17:mtH8861		
F2:mtL382	55	1.5	F18:mtL8799	55	1.5
R2:mtH945			R18:mtH9397		
F3:mtL923	55	1.5	F19:mtL9362	55	1.3
R3:mtH1487			R19:H9928		
F4:mtL1466	55	1.5	F20:mtL9886	55	1.3
R4:mtH2053			R20:mtH10462		
F5:mtL2025	55	1.5	F21:mtL10403	56	1.5
R5:mtH2591			R21:mtH10975		
F6:mtL2559	56	1.5	F22:mtL10949	58	1.5
R6:mtH3108			R22:mtH11527		
F7:mtL3073	55	1.5	F23:mtL11486	56	1.5
R7:mtH3670			R23:mtH12076		
F8:mtL3644	55	1.5	F24:mtL12028	55	1.0
R8:mtH4227			R24:mtH12603		
F9:mtL4210	55	1.5	F25:mtL12572	56	1.5
R9:mtH4792			R25:mtH13124		
F10:mtL4750	55	1.5	F26:mtL13088	55	1.3
R10:mtH5306			R26:mtH13666		

[1] Primer sets 19 to 32 were optimised in collaboration with M. Alessandrini.

Table 4.1: continued...

Primer	Annealing Temp.	[$MgCl_2$]	Primer	Annealing Temp.	[$MgCl_2$]
F11:mtL5278	55	1.5	F27:mtL13612	55	1.0
R11:mtH5832			R27:mtH14186		
F12:mtL5781	58	1.5	F28:mtL14125	53	1.5
R12:mtH6367			R28:mtH14685		
F13:mtL6337	58	1.5	F29:mtL14650	53	1.5
R13:mtH6899			R29:mtH15211		
F14:mtL6869	58	1.5	F30:mtL15162	55	1.0
R14:mtH7406			R30:mtH15720		
F15:mtL7379	55	1.5	F31:mtL15676	53	1.5
R15:mtH7918			R31:mtH16157		
F16:mtL7882	55	1.5	F32:mtL15996	55	1.0
R16:H8345			R32:mtH16401		

Annealing Temp. = experimental annealing temperature utilised for the primer set, [$MgCl_2$] = magnesium chloride concentration in the PCR reaction mixture in mM, F = forward primer and R = reverse primer. Primer sequences were obtained from Maca-Meyer *et al.* (2001).

Positive and negative controls were included in all PCR amplifications. Positive controls were utilised to identify samples where the patient's DNA caused the reaction not to work i.e. due to polymorphisms. Positive controls were also utilised as a wild type reference along with the RCRS for comparison purposes and sequence analysis. Negative controls, containing no DNA, were utilised to test the possibility of reagent and vessel contamination. The addition of dimethyl sulfoxide or formamide was not required in any of the amplification reactions as mtDNA does not form secondary structures that interfere with PCR amplification. The increased temperature, as listed in Table 3.5 (page 44), was sufficient to denature the mtDNA for the PCR reaction. The PCR conditions in Table 3.5 were applied for amplification of the three specific tRNA genes as well as for the whole mitochondrial genome.

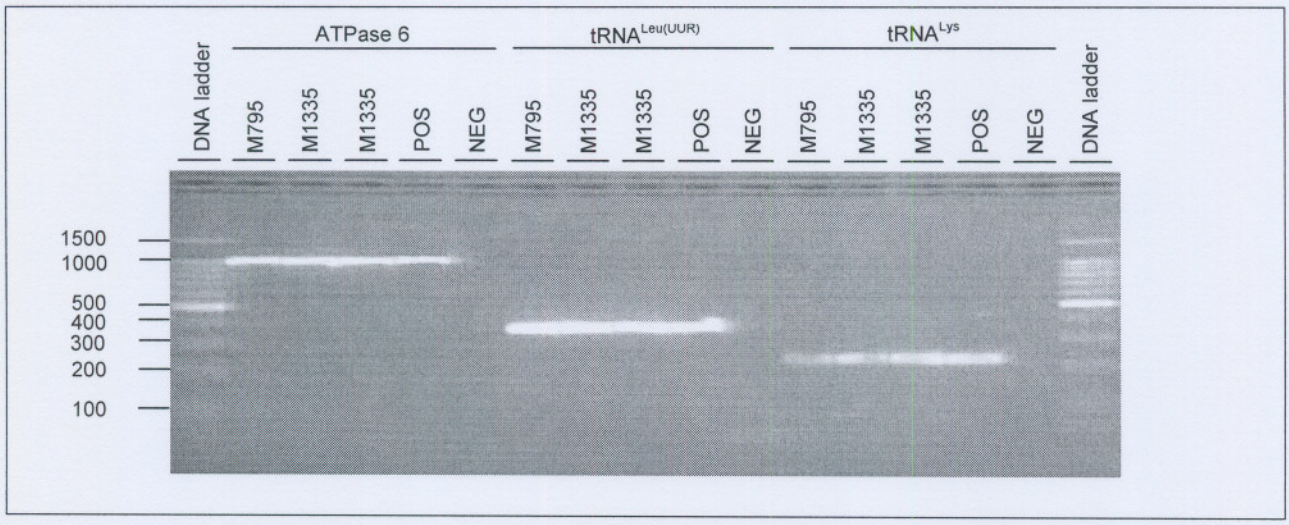
The primers utilised for amplification of the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and the ATPase 6 genes were initially designed for mutation detection via restriction fragment length polymorphism (RFLP) analysis. Primers utilised in RFLP analyses often contain single nucleotide changes to create certain restriction sites. Therefore, the 8380REV2 reverse primer, as presented in Table 3.1, is not completely complementary to the recognition sequence. 8380REV2 contains a C instead of a G at the third last nucleotide. Although RFLP analyses were not utilised in the investigation presented here, it was not economical to synthesise new primers and the primers from the RFLP investigation were utilised unaltered. Therefore, all the $tRNA^{Lys}$ amplicons had a G instead of a C at nt 8359.

However, if whole mtDNA genome sequencing of these patients is performed in future, nt 8359 will have to be resequenced as the nucleotide at that position is not known.

4.1.3 Electrophoresis and PCR product purification

As discussed in paragraph 3.5, successful amplification of the correct fragment size was determined by agarose gel electrophoresis before purification and sequencing. Amplified $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and ATPase 6 fragments from muscle DNA of three patients (795, 1335 and 1336) are presented in Figure 4.1. As was expected, the larger ATPase 6 fragments migrated slower than the shorter $tRNA^{Leu(UUR)}$, and especially $tRNA^{Lys}$ fragments under the influence of the applied current. The three different gene fragments could easily be distinguished based on size and gel mobility.

Figure 4.1: Photographic representation of PCR products for the ATPase 6, $tRNA^{Leu(UUR)}$ and $tRNA^{Lys}$ mitochondrial regions



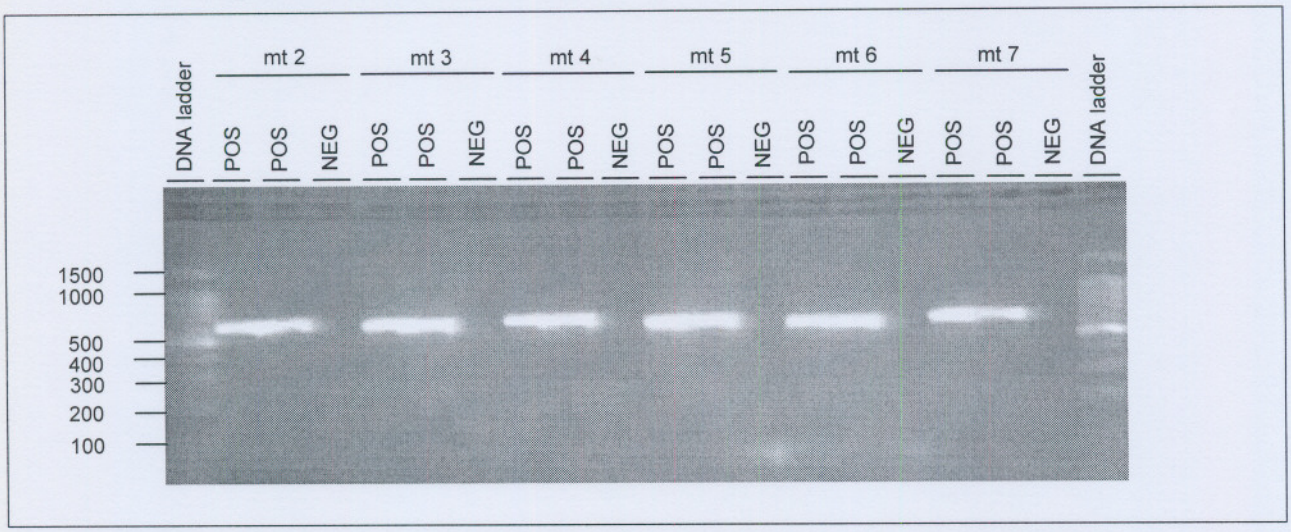
2% Agarose gel electrophoresed at 100 V for 30 min in 1 x TBE buffer. The numbers M795, M1335 and M1336 indicate the patient muscle samples. ATPase 6 = amplicons of the gene encoding ATP synthase subunit 6, $tRNA^{Leu(UUR)}$ = amplicons of the gene encoding tRNA leucine (specifically recognising the codon UUR), $tRNA^{Lys}$ = amplicons of the gene encoding tRNA lysine, DNA ladder = 100 bp from Promega, POS = positive control and NEG = negative control.

All the fragments amplified with the Maca-Meyer *et al.* (2001) primers were ca. 600 bp in length and could not be distinguished according to size or gel mobility as in the case of the amplicons of $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and ATPase 6. The sizes of the fragments amplified for whole mitochondrial genome sequencing were presented in Table 3.6 (page 45).

A photographic representation of six of the 32 overlapping PCR fragments amplified for whole mitochondrial genome sequencing is presented in Figure 4.2. PCR fragments 2, 3, 4, 5, 6 and 7 (Maca-Meyer *et al.*, 2001) were amplified from control DNA for optimisation purposes. The respective negative controls are indicated. Electrophoresis is an

uncomplicated and cost-effective method to confirm that the correct fragment length was amplified successfully before continuing with PCR product purification and sequencing.

Figure 4.2: Photographic representation of PCR products for mitochondrial fragments 2, 3, 4, 5, 6 and 7



2% Agarose gel electrophoresed at 100 V for 30 min in 1 x TBE buffer. DNA ladder = 100 bp from Promega, mt = mitochondrial DNA fragment 2 – 7 of 32 fragments, POS = positive control and NEG = negative control.

After gel electrophoresis, the successful PCR products were purified from the primers, nucleotides, polymerases, salts and mineral oil in the PCR reaction mixture, utilising spin columns of the QIAquick® PCR Purification Kit. The protocol is efficient in terms of cost and time and can be recommended. However, the spin columns are expensive.

The concentrations of the purified PCR products were measured spectrophotometrically, as discussed in paragraph 3.3. Concentrations varied between 10 and 40 ng.μl^{-1}, on average, for the different regions amplified.

4.1.4 Cycle sequencing

The quantity of DNA used for cycle sequencing depended on the purified PCR product concentration and fragment length as presented in Table 3.7 (page 49). However, ca. 8-10 ng of PCR product was utilised for the sequencing of the $tRNA^{Leu(UUR)}$ and $tRNA^{Lys}$ fragments and ca. 15-20 ng for the ATPase 6 fragment. When sequencing the overlapping fragments for whole mitochondrial genome sequencing, ca. 10-20 ng was utilised.

The concentration and quality of the PCR product, as well as precipitation of the sequenced product, are critical factors that determined the success of the sequencing reaction. Excess template will lead to broad peaks of which the individual bases cannot be

distinguished. Insufficient template will lead to peaks with low amplitude overshadowed by background peaks. The quality of the template also plays a large role. If the PCR reaction is not optimal and non-specific secondary amplification occurs, excessive background will be observed. For this reason, time and effort were invested in optimising the PCR conditions for amplification with the 32 primer sets of Maca-Meyer *et al.* (2001). Finally, the precipitation of the sequenced product is critical. Care should be taken not to disturb the pellet when withdrawing the precipitation or wash solution, as it is not visible and can easily be removed as well. It is also important that the pellet is completely air-dried afterwards, as moisture can cause the pellet to be unstable.

The automated cycle sequencing technique worked extremely well for mutation detection, as a large region could be screened for alterations from the RCRS. Compared to RFLP analysis, as performed previously by Prosser (2001), much more data are collected with a single sequence than with one RFLP experiment. However, as discussed in paragraph 4.5, the cycle sequencing technique is less sensitive and low levels of heteroplasmy cannot be detected, whereas it is detectable to a certain extent by RFLP analysis.

4.2 SCREENING OF THE MITOCHONDRIAL $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ AND THE ATPase 6 GENES

Whole blood and/or muscle samples of selected patients with suspected mitochondrial disorders were screened for alterations of the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and the ATPase 6 genes. For eight patients (800, 878, 1233, 1301, 1302, 1316, 1335 and 1336) blood and muscle samples were available and the results were compared.

4.2.1 Sequence analysis of whole blood samples

Twenty four whole blood samples were screened for sequence variations in and flanking the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and the ATPase 6 mitochondrial genes. Twelve different neutral alterations, including six novel synonymous alterations as presented in Table 4.2, were detected. Heteroplasmy could not be detected for any of these alterations by sequencing.

Silent or synonymous mutations do not alter the encoded protein and do not result in a disease phenotype. They involve bp substitutions that do not result in the substitution of a different amino acid into the encoded protein, owing to degeneracy of the genetic code (Winter *et al.*, 2002). Non-synonymous mutations, on the other hand, result in the

substitution of a different amino acid into the encoded protein. Over time, silent mutations tend to accumulate in DNA and are commonly referred to as polymorphisms (Winter *et al.*, 2002). However, according to other authors polymorphisms are neutral alterations that occur at a frequency between 0.01 and 0.99 (Falconer, 1981). The alterations depicted as "polymorphisms" in the investigation presented here were reported as such, but do not necessarily occur at the defined frequency range and could represent rare variants. However, they are all neutral alterations that do not result in a disease phenotype.

Table 4.2: Neutral alterations detected in the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and the ATPase 6 mitochondrial genes isolated from whole blood

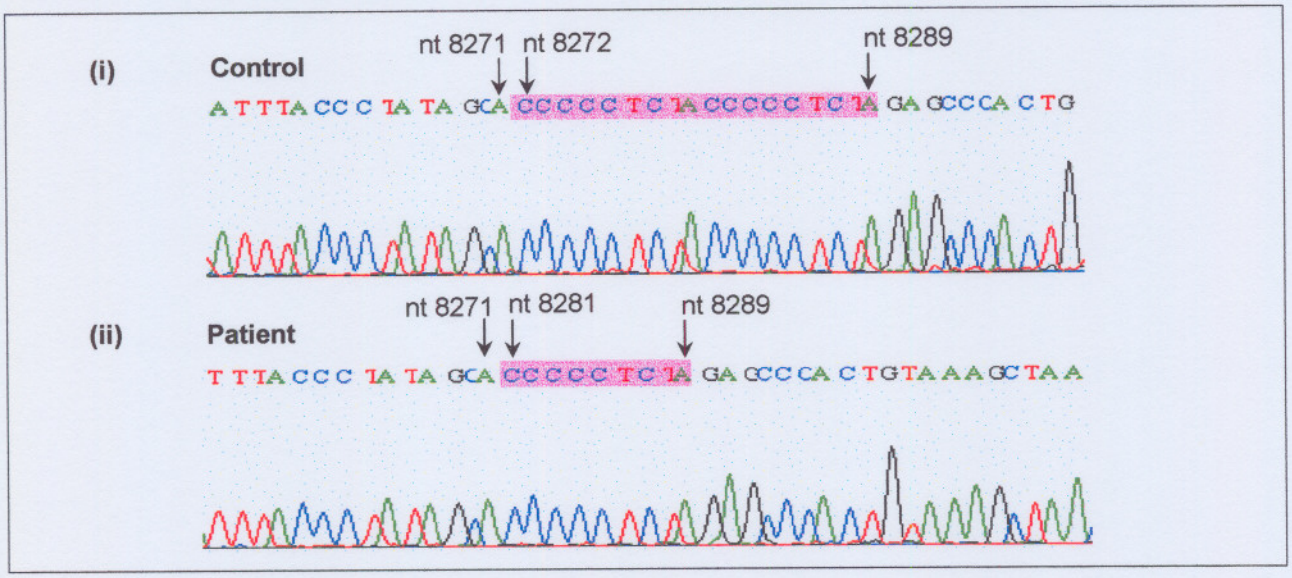
Patient #	Alteration	Locus	Alteration Type	Reference
799	T3336C	ND1	Polymorphism	Sternberg *et al.* (1998)
767	G8206A	COII	Polymorphism	Ozawa *et al.* (1995)
759 870	G8251A	COII	Polymorphism	Ruvolo *et al.* (1993)
789	8271_8281 delCCCCCTCTA	$tRNA^{Lys}$, COII intergenic	Polymorphism	Cann and Wilson (1983)
1314	G8269A	COII	Polymorphism	Spagnalo *et al.* (2001)
799	A8901G	ATPase 6	Synonymous	Not reported
804	T8928C	ATPase 6	Synonymous	Not reported
759 858 870 1233	C9042T	ATPase 6	Synonymous	Not reported
1302	A9072G	ATPase 6	Synonymous	Not reported
759	T9111C	ATPase 6	Synonymous	Not reported
767 1301 1316 1335 1336	A9221G	COIII	Polymorphism	Silva *et al.* (2003)
759 858 870 1233	A9347G	COIII	Synonymous	Not reported

Blue = alteration also found in DNA from muscle of this patient, # = number, ND1 = gene encoding subunit 1 of NADH dehydrogenase, $tRNA^{Lys}$ = gene encoding tRNA lysine, CO II = gene encoding subunit II of cytochrome c oxidase, ATPase 6 = gene encoding ATP synthase subunit 6 and COIII = gene encoding subunit III of cytochrome c oxidase.

The intergenic 9 bp deletion between the COII and $tRNA^{Lys}$ genes is a length polymorphism first reported by Cann and Wilson (1983). The same alteration was subsequently reported in different populations, namely the American Indians (Torroni *et al.*, 1992), Asians (Ballinger *et al.*, 1992) and African Pygmies from sub-haplogroup L1

(Chen *et al.*, 1995). The intergenic 9 bp deletion detected in patient 789 (nt 8271-8281) is depicted in Figure 4.3.

Figure 4.3: Electropherogram of the 9 bp deletion (nt 8271-8281) in patient 789



The highlighted sequence indicates the area of deletion. nt = nucleotide.

With respect to length changes, there are different sequence classes in the COII/tRNA^{Lys} intergenic region. Most humans have two copies of the 9 bp repeated sequence, as presented in Figure 4.3(i), and this is probably the ancestral state, as evidenced by phylogenetic analysis and comparison with non-human primates (Horai *et al.*, 1993). However, two individuals with a triplication of the 9 bp repeated sequence have been reported, one Chuckchi from Siberia (Shields *et al.*, 1992) and a Tharu from Nepal (Passarino *et al.*, 1993). It was originally proposed that the 9 bp deletion had a single origin in Asia (Wrischnik *et al.*, 1987). However, it has been demonstrated that the deletion has arisen at least three times: once in Africa and probably two separate times in Asia. The most likely origin of the deletion is slipped mispairing during replication (Albertini *et al.*, 1982). Statistical analyses support the independent origin of the deletion in Africa and in Asia (Redd *et al.*, 1995). The 9 bp deletion was also detected by Prosser (2001) in two black African individuals who were sisters.

Six novel synonymous alterations were detected in DNA from blood, namely the A8901G, T8928C, C9042T, A9072G and T9111C alteration in the ATPase 6 gene and the A9347G alteration in the COIII gene. Four patients (759, 858, 870 and 1233) had both the C9042T and the A9347G alteration. All four patients are black African individuals and the combination is probably indicative of a specific haplogroup or sub-haplogroup. However, these alterations have not been reported and are not included in the haplogroup exclusion

criteria listed in Appendix B, Table B.1. As phylogenetics is unravelled, new trees are compiled and it is possible that these alterations may be included in future. Haplogroup analysis will be discussed in paragraph 4.3.

One potentially pathogenic mutation was detected in DNA from blood of patient 1314, as presented in Table 4.3. This alteration has been reported (Ozawa *et al.*, 1991a).

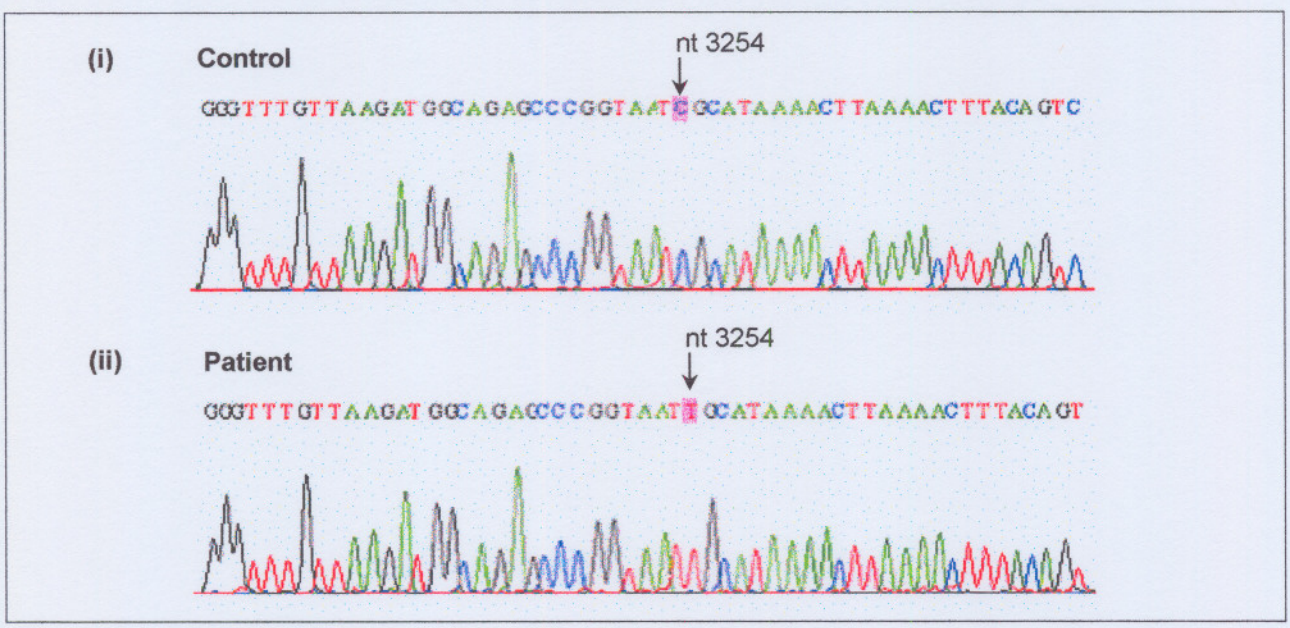
Table 4.3: Potentially pathogenic mutation detected in the mitochondrial $tRNA^{Leu(UUR)}$ isolated from whole blood of patient 1314

Alteration	Locus	Alteration Type	Reference
C3254T	$tRNA^{Leu(UUR)}$	Functional effect unknown	Ozawa *et al.* (1991a)

$tRNA^{Leu(UUR)}$ = gene encoding tRNA leucine (specifically recognising the codon UUR)

An electropherogram of the C3254T alteration is depicted in Figure 4.4. Heteroplasmy could not be detected by sequencing.

Figure 4.4: Electropherogram of the C3254T alteration in patient 1314



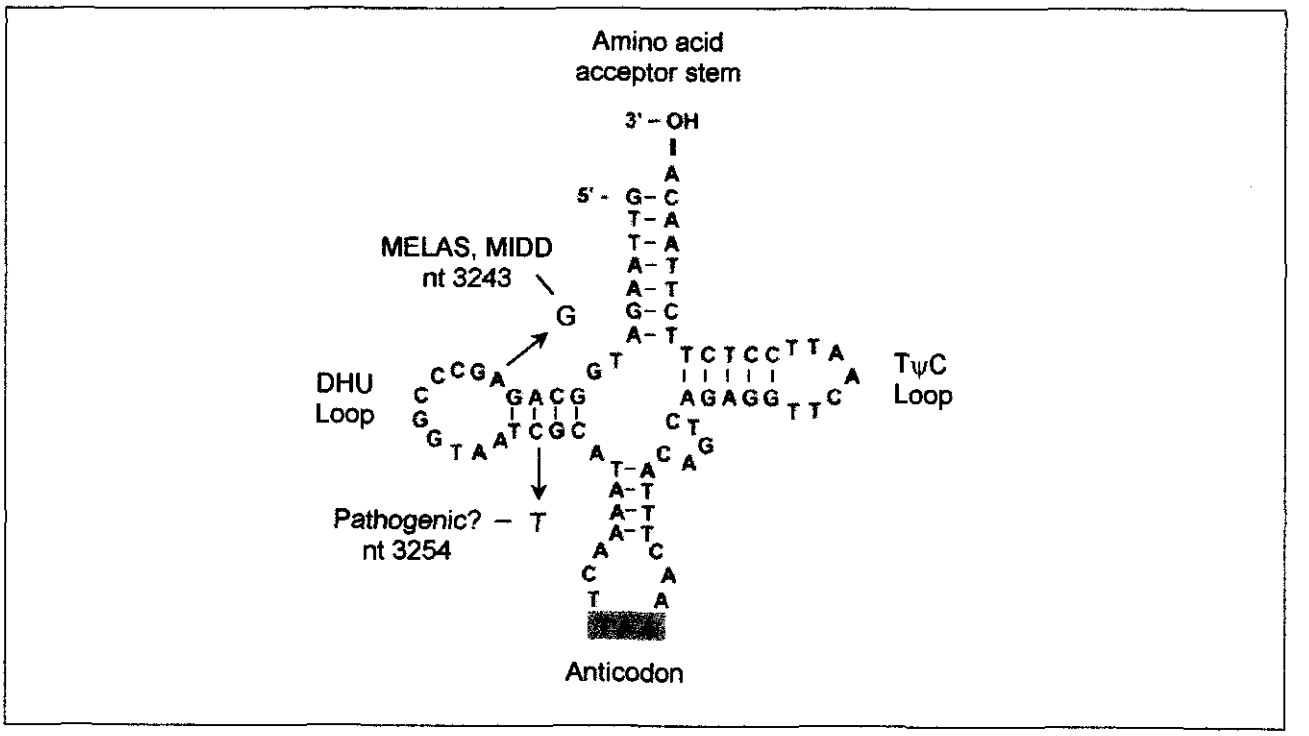
The highlighted sequence indicates the position of the C3254T alteration. nt = nucleotide.

The C3254T alteration detected in the dihydrouridine loop of $tRNA^{Leu(UUR)}$ (as presented in Figures 4.4 and 4.5) was first reported as one of many point mutations in a patient with fatal infantile cardiomyopathy (Ozawa *et al.*,1991a, Ozawa *et al.*,1991b, Obayashi *et al.*, 1992). This same individual also had missense mutations in $tRNA^{Ile}$ and $tRNA^{Trp}$, namely A4317G and C5554T, respectively (Ozawa *et al.*, 1991a, Ozawa *et al.*, 1991b, Obayashi *et al.*, 1992). According to MITOMAP (2003) the C3254T alteration is a polymorphism, but the C3254G alteration is regarded as a mutation associated with myopathy (Kawarai *et al.*,

1997). However, according to the literature (Ozawa *et al.*, 1991a, Ozawa *et al.*, 1991b, Obayashi *et al.*, 1992) no functional studies have been reported to determine the pathogenicity of the C3254T alteration, nor the C3254G alteration. Neither of the two changes affects base pairing in the $tRNA^{Leu(UUR)}$ molecule, as this position is one of three non-standard base pairs (Obayashi *et al.*, 1992).

The secondary structure of the $tRNA^{Leu(UUR)}$ molecule and the position of the C3254T alteration vs. the A3243G mutation are presented in Figure 4.5. The morbid map of the $tRNA^{Leu(UUR)}$ molecule was presented in Figure 2.3 (page 27).

Figure 4.5: The A3243G and C3254T alterations on the $tRNA^{Leu(UUR)}$ molecule



MELAS = mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes, MIDD = maternally inherited diabetes and deafness, DHU = dihydrouridine, ψ = pseudouridine (5-ribosyl uracil) and nt = nucleotide. Adapted and modified from MITOMAP (2003).

At first glance, the C3254T substitution in the $tRNA^{Leu(UUR)}$ molecule seems to be beneficial and non-pathogenic by causing the molecule to be more stable with an A – T bond instead of the A – C bond. However, this more stable bond may cause rigidity in the molecule, limiting its flexibility and function. Furthermore, this alteration site is positioned within the binding region of the mitochondrial termination factor, mtTERM (Kruse *et al.*, 1989) and it was suggested that alterations of this position may interfere with transcription regulation (Kawarai *et al.*, 1997).

Other mutations of the dihydrouridine loop, as presented in Figure 2.3 (page 27), such as the well-known A3243G mutation (Goto *et al.* 1990), cause devastating phenotypes. However, the adenine at nt 3243 (indicated in red) is evolutionarily conserved as presented in Table 4.4, whereas the cytosine at nt 3254 (indicated in blue) is not. In fact, 14 of the 16 species listed in Table 4.4 have a thymine at nt 3254, suggesting that the C3254T alteration is probably not pathogenic but may rather be an evolutionary adaptation. Alternatively, 15 of the 16 species in Table 4.4 have an adenine at nt 3243. Only one of the species listed, namely *Canis familiaris*, has a guanine at nt 3243, illustrating the evolutionary conservation of adenine at this position.

Table 4.4: The conservation of adenine at nt 3243 and cytosine at nt 3254 in the mitochondrial tRNA$^{Leu(UUR)}$ across 16 species

Species	Sequence from nt 3234 - 3256
Alligator mississipiensis	GGTTGGCAG **A**GCCTGGCTTAAT**T**GC
Balaenoptera musculus	AGGTGGCAG **A**GTTCGG-TAAT**T**GC
Balaenoptera physalus	AGGTGGCAG **A**GTTCGG-TAAT**T**GC
Bos Taurus	AGGTGGCAG **A**GCCCGG-TAAT**T**GC
Canis familiaris	GGGTG-CAG **G**GCCCGG-TAAC**T**GC
Cebus albifrons	AGATGGCAG **A**GCCCGG-CAAT**T**GC
Cyprinus carpio	GGGTGGCAG **A**GCATGGTAAAT**T**GC
Drosophila melanogaster	ATATGGCAG **A**-------TTAG**T**GC
Equus asinus	GGGTGGCAG **A**GCC-GG-AAAT**T**GC
Equus caballus	GGGTGGCAG **A**GCCCGG-AAAT**T**GC
Erinaceus europaeus	GTGTGGCAG **A**GCCCGG-TAAT**T**GC
Felis catus	GGGTGGCAG **A**GCCCGG-TAAC**T**GC
Gallus gallus	GCGTGGCAG **A**GCTCGG-CAAA**T**GC
Gorilla gorilla	AGATGGCAG **A**GCCCGG-TAAT**C**GC
Hippoppotamus amphibious	CGATGGCAG **A**GCCCGG-TAAT**T**GC
Homo sapiens sapiens	AGATGGCAG **A**GCCCGG-TAAT**C**GC

Red = nucleotide at nt 3243 and blue = nucleotide at nt 3254. Obtained from Procaccio (2003).

Patient 1314 harbours the C3254T alteration. This patient is a Caucasian female who presented with resistant epilepsy and developmental delay. Biochemically she has an abnormal profile of organic acids, amino acids and carnitine in the urine.

Functional studies will have to be conducted to determine whether the C3254T alteration culminates in a functional change of the tRNA$^{Leu(UUR)}$ molecule and whether it plays a role in the phenotype of this patient. However, the first step will be to test healthy maternal relatives of the patient for the alteration to exclude it being a neutral inherited mutation. If

the healthy maternal relatives are negative for the alteration, functional studies should definitely be initiated.

4.2.2 Sequence analysis of muscle samples

Eighteen muscle samples were also screened for sequence variations of the mitochondrial $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and the ATPase 6 genes and surrounding regions. Initially DNA isolation from muscle was optimised on bovine muscle tissue, as muscle biopsies are invasive, expensive procedures and excess human muscle samples were not available. For eight patients (800, 878, 1233, 1301, 1302, 1316, 1335 and 1336) both muscle and blood samples were analysed and the results were compared. Three of the novel synonymous alterations detected in DNA from blood, namely the C9042T, A9072G and A9347G alterations, were also detected in DNA from muscle. The neutral alterations detected in DNA from muscle are presented in Table 4.5. Heteroplasmy could not be detected for any of these alterations.

Table 4.5: Neutral alterations detected in the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and the ATPase 6 mitochondrial genes isolated from muscle

Patient #	Alteration	Locus	Alteration Type	Reference
757	8271_8281 delCCCCCTCTA	$tRNA^{Lys}$, COII intergenic	Polymorphism	Cann and Wilson (1983)
758 795 1290	G8251A	COII	Polymorphism	Ruvolo *et al.* (1993)
757 758 795 1233 1290	C9042T	ATPase 6	Synonymous	Not reported
1302	A9072G	ATPase 6	Synonymous	Not reported
1301 1316 1335 1336	A9221G	COIII	Polymorphism	Silva *et al.* (2003)
757 758 795 1233 1290	A9347G	COIII	Synonymous	Not reported

Blue = alteration also found in DNA from whole blood of this patient, # = number, $tRNA^{Lys}$ = gene encoding tRNA lysine, CO II = gene encoding subunit II of cytochrome c oxidase, ATPase 6 = gene encoding ATP synthase subunit 6 and COIII = gene encoding subunit III of cytochrome c oxidase.

As was expected, in the patients of whom blood and muscle tissue were analysed, synonymous alterations and polymorphisms detected in DNA from whole blood were also detected in DNA from muscle. The C9042T, A9072G, A9221G and A9347G alterations were present in both tissues of the patients indicated in blue in Tables 4.2 and 4.5.

The intergenic 9 bp deletion at nt 8271-8281 between the COII and $tRNA^{Lys}$, here detected in the muscle of patient 757, was discussed in detail in 4.2.1.1. As in DNA from blood, the C9042T and A9347G combination was very common in DNA from muscle and was found in four more patients (757, 758, 795 and 1290). The C9042T and A9347G synonymous alterations are probably novel polymorphisms and haplogroup-related as discussed in paragraph 4.2.1. This combination was detected in the DNA from blood and muscle of patient 1233.

One non-synonymous, potentially pathogenic mutation, the T9098C alteration, was detected in DNA isolated from muscle of two individuals. The alteration has not been reported and heteroplasmy could not be detected for either of the patients. Information regarding this alteration is depicted in Table 4.6.

Table 4.6: Potentially pathogenic mutation detected in the ATPase 6 mitochondrial gene isolated from muscle

Patient #	Alteration	Locus	Amino Acid Change	Alteration Type
476 753	T9098C	ATPase 6	Ile → Thr	Non-synonymous alteration

= number, ATPase 6 = gene encoding ATP synthase subunit 6, Ile = isoleucine and Thr = threonine.

The T9098C alteration, causing an isoleucine to threonine change at position 191 of ATPase 6 (MITOANALYZER, 2003), was detected in two Caucasian patients (476 and 753). Isoleucine is a non-polar, hydrophobic amino acid, whereas threonine is a polar, hydrophilic amino acid (Winter *et al.*, 2002). The two patients have distinct phenotypes ranging from failure to thrive and lactic acidosis in patient 476 to CNS symptoms and organic acids in the urine in patient 753.

The 39 species evaluated to determine evolutionary amino acid conservation are listed in Table 4.7. The substituted isoleucine is not conserved across species, although the six species in Table 4.7 that have isoleucine at position 191 are all primates. The other 33 species investigated have a threonine at that position. It seems as if after evolving from

threonine into isoleucine in the primates, this amino acid evolved back into threonine in these two patients. This is based on the assumption that isoleucine at position 191 is conserved in humans. If an amino acid is described as conserved in the investigation presented here, it means conservation across species. The fact that the two patients have the same non-conserved non-synonymous alteration and two different phenotypes, limits the likelihood of it being pathogenic. However, even patients with the same pathogenic mutation can have heterogeneous phenotypes i.e. the A3243G mutation can cause MELAS (Goto *et al.*, 1990) and MIDD (Van den Ouweland *et al.*, 1994). Therefore, future analyses will have to be undertaken to definitely exclude the T9098C alteration as being pathogenic.

Table 4.7: Species evaluated to determine amino-acid conservation

Species	Species
Alligator mississippiensis	*Macropus robustus*
Balaenoptera musculus	*Mus musculus*
Balaenoptera physalus	***Nycticebus coucang***
Bos Taurus	*Ornitorhynchus anatinus*
Canis familiaris	*Oryctolagus cuniculus*
Cebus albifrons	*Ovis aries*
Cyprinus carpio	***Pan paniscus***
Drosophyla melanogaster	***Pan troglodytes***
Equus asinus	***Papio hamadryas***
Equus caballus	*Phoca vitulina*
Erinaceus europaeus	***Pongo pygmaeus***
Felis catus	***Pongo pygmaeus abelii***
Gallus gallus	*Rattus norvegicus*
Gorilla gorilla	*Rhinoceros unicornis*
Hippopotamus amphibious	*Sciurus vulgaris*
Homo sapiens sapiens	*Sus scrofa*
Hylobates lar	*Talpa europaea*
Lama pacos	***Tarsius bancanus***
Loxodonta africana	*Xenopus laevis*
Macaca sylvanus	

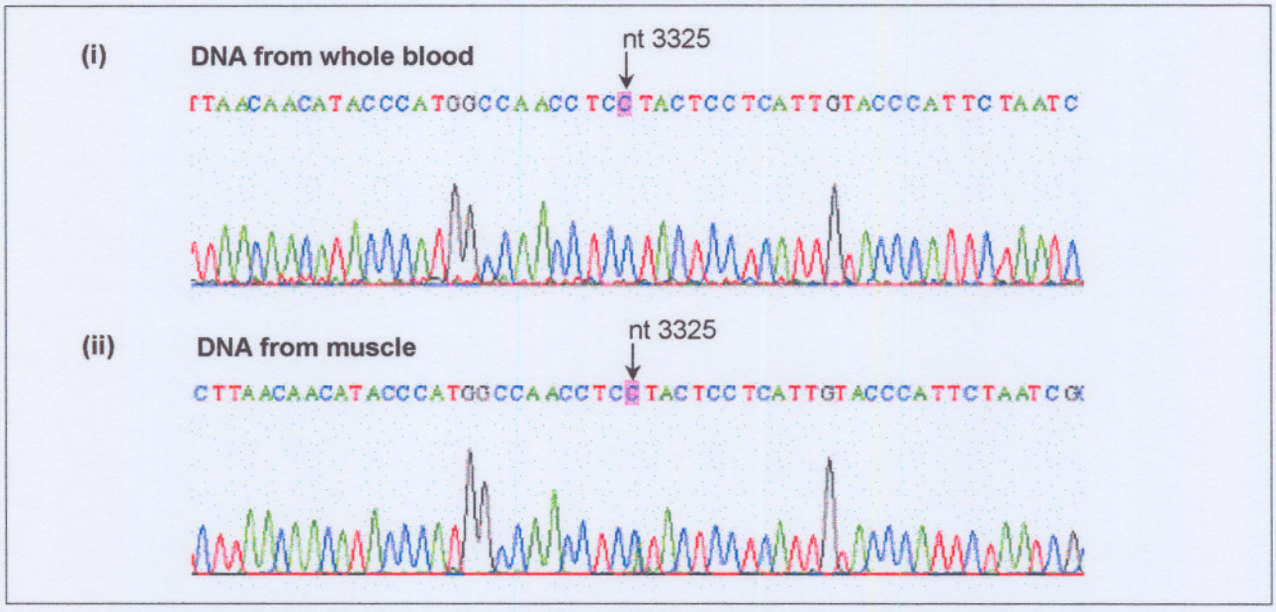
The primates are indicated in bold letters. Compiled from Ruiz-Pesini (2004).

4.2.3 Alteration confirmation

As mentioned earlier, significant alterations, such as reported pathogenic mutations, novel non-synonymous changes or alterations where the functional effect of the change was not

known, were confirmed through sequencing of the specific amplicon (derived from separate PCR reactions) in both directions. The importance of confirmation of significant alterations was accentuated by the detection of a "heteroplasmic" C3325A change in DNA from muscle of patient 1336, but not in the blood. DNA from blood of this patient was homoplasmic for cytosine at nt 3325, whereas DNA from muscle was "heteroplasmic" and both cytosine and adenine peaks were visible on the sequence at this position. There are no reports of the C3325A alteration in the literature. The "heteroplasmy" is presented in Figure 4.6.

Figure 4.6: Electropherogram of the C3325A alteration in patient 1336



The highlighted sequence indicates the position of the "heteroplasmic" C3325A alteration. nt = nucleotide.

However, when attempting to confirm the presence of the C3325A alteration in ND1 of patient 1336 by sequencing of another $tRNA^{Leu(UUR)}$ PCR product in the reverse direction, the C3325A alteration could not be detected. As the initial adenine peak was distinct, and the C3325A alteration causes a highly conserved leucine to change to a methionine, its disappearance was investigated. The first PCR product in which the C3325A alteration was detected, was resequenced in the forward direction to determine whether the C3325A alteration was an error in the PCR reaction, during which an adenine was built into the amplicon at that position. However, the C3325A alteration was not detected in the second sequencing attempt. To exclude the slim but possible chance of sample exchange, the whole batch of $tRNA^{Leu(UUR)}$ PCR products sequenced together with 1336 on the day when the C3325A alteration was detected, was resequenced. The C3325A alteration was not detected in any of these samples.

According to Preisig (2003) the adenine peak noticed in the initial electropherogram could be the result of contamination at the stage of preparing the sequencing reaction or when resuspending it for injection into the sequencer. He postulated that a contaminating dye could have been detected at that position or otherwise it was simply an artefact. This case highlights the necessity for sequencing in both directions in order to confirm sequence alterations.

4.3 HAPLOGROUP ANALYSIS

Haplogroup analysis was not one of the aims of the investigation presented here, but haplogroups can influence disease expression (Torroni, 2000). This was discussed in paragraph 2.6.2.3.1 with regard to haplogroup J and the development of LHON. Therefore, haplogroup analysis was performed on the four patients of whom whole mtDNA genome sequences were generated in this investigation. For the other patients screened only for alterations of $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and ATPase 6, haplogroup analysis could not be performed, as the required polymorphic data were not available. The inclusion and exclusion criteria utilised for haplogroup analysis are presented in Appendix B. The results of the haplogroup analysis will be discussed separately for each patient in paragraph 4.4 below.

4.4 WHOLE MITOCHONDRIAL GENOME SCREENING

Whole mitochondrial genome screening was performed for patients 386, 504, 525 and 1301. These patients have MDC scores of eight and no pathogenic alterations in $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ or ATPase 6 genes have yet been determined. DNA isolated from whole blood was analysed for patients 386, 504 and 525, although it would have been preferable to utilise muscle. Unfortunately DNA from muscle was only available for patient 1301. The alterations detected with cycle sequencing in the four patients are depicted in Appendix A. The potentially pathogenic mutations detected were confirmed with sequencing in both directions and will be discussed.

4.4.1 Patient 386

This patient is a Caucasian female who presented with symptoms at birth and died at the age of six weeks. Her clinical phenotype included intrauterine growth retardation, intractable seizures, apnoea and lactic acidosis.

The 13 alterations detected in the mtDNA genome of patient 386 are presented in Appendix A, Table A.1. The alterations were evaluated and compared to the data available on MITOMAP (2003). Twelve of the 13 alterations detected, were reported polymorphisms. Only one alteration was a novel non-synonymous potentially pathogenic mutation, namely the G7979A in the COII gene. Heteroplasmy could not be detected, but the G7979A alteration results in an amino acid substitution, namely aspartic acid to asparagine (MITOANALYZER, 2003). Both amino acids are polar, but aspartic acid is acidic and asparagine is basic (Winter *et al.*, 2002). Aspartic acid is not conserved at position 132 of COII and was only present in 12 of the 39 species listed in Table 4.7. However, 10 of these are primates, which could indicate the importance of the amino acid for this group of mammals. Therefore, the investigation of the pathogenicity of G7979A alteration should be pursued further.

Table 4.8: Potentially pathogenic mutation detected in patient 386

Alteration	Locus	Amino Acid Change	Alteration Type
G7979A	COII	Asp → Asn	Non-synonymous alteration

COII = gene encoding cytochrome *c* oxidase subunit II, Asp = aspartic acid and Asn = asparagine.

Haplogroup analysis was performed as explained in Appendix B. As patient 386 has polymorphisms 1 to 5 in Table B.1, the patient belongs to the same sub-haplogroup of haplogroup H as the individual from whom the CRS was compiled. Therefore, it is not surprising that patient 386 had very little deviation from the RCRS.

4.4.2 Patient 504

Patient 504 is a black African female who presented with symptoms of a mitochondrial disorder at the age of six months. Clinically she presented with stunted growth, muscle wasting, myopathy, developmental delay, hepatomegaly and brain atrophy. Biochemical analyses revealed lactic acidosis, and raised amino acids were detected in the plasma and urine.

From the outset the DNA of patient 504 behaved differently during amplification, as many of the PCR reactions failed and had to be repeated. This can probably be explained by alterations being present within primer annealing sites, as 88 alterations were eventually detected in her mtDNA. Temperature in the thermocycler is not constant across all 96 wells and it is possible that owing to the additional mtDNA alterations the conditions were

sometimes suboptimal, resulting in a success rate of approximately two out of three for amplification of DNA from this patient. The alterations that were detected are presented in Appendix A, Table A.2. Five novel non-synonymous changes, as well as a five nucleotide insertion of cytosines in the 12S rRNA gene, were detected. Heteroplasmy could not be detected for 87 of the 88 alterations identified in patient 504. However, it appears as if the 956-965insCCCCC is heteroplasmic, although this should be confirmed, as discussed later in this section. The potentially pathogenic mutations are depicted in Table 4.9 and discussed below.

Table 4.9: Potentially pathogenic mutations detected in patient 504

Alteration	Locus	Amino Acid Change	Alteration Type
956-965insCCCCC	12S rRNA	~	Shoffner *et al.* (1993)
C5911T	COI	Ala → Val	Non-synonymous alteration
A7146G	COI	Thr → Ala	Non-synonymous alteration
A8566G	ATPase6 ATPase8	Ile → Val	Non-synonymous alteration Synonymous alteration
A13276G	ND5	Met → Val	Non-synonymous alteration
G15431A	Cyt *b*	Ala → Thr	Non-synonymous alteration

12S rRNA = gene encoding 12S ribosomal RNA, ~ = mutation not found in protein-coding gene, COI = gene encoding cytochrome *c* oxidase subunit I, ATPase 6 and 8 = genes encoding subunits 6 and 8 of ATP synthase, ND5 = gene encoding subunit 5 of NADH dehydrogenase, Cyt *b* = gene encoding cytochrome *b*, Ala = alanine, Val = valine, Ile = isoleucine, Met = methionine and Thr = threonine.

Two novel alterations result in amino acid substitutions in the COI subunit of cytochrome oxidase. C5911T causes a substitution of alanine with valine (MITOANALYZER, 2003), which are both polar (Winter *et al.*, 2002), at position three of COI (MITOANALYZER, 2003). However, valine has two methyl groups (Rodwell, 1993) that can cause sterical problems in the polypeptide. Alanine at position three is not conserved and occurs in only five species listed in Table 4.7, of which four are primates. As alanine is not conserved, the C5911T alteration is probably not pathogenic, but future analysis should be undertaken to exclude it.

Another amino acid substitution is induced by the A7146G alteration at position 415 of COI. This alteration has also not been reported in literature and initially it was thought that it is potentially pathogenic, as it causes a polar threonine to change to a non-polar alanine (Winter *et al.*, 2002; MITOANALYZER, 2003). Threonine is conserved in 13 species in Table 4.7, of which only four are primates. However, when the haplogroup data depicted in Appendix B were analysed, it was observed that the A7146G alteration is the 27th

polymorphism utilised to define sub-haplogroup L0 and L1 in Table B.1. Therefore, this alteration was regarded as a polymorphism. This alteration was also detected in patient 525.

The ATPase 6 and 8 genes overlap in mtDNA (Anderson, 1981) from position 8527 to 8572 (MITOMAP, 2003). That is the reason why the novel A8566G alteration affects both genes. However, A8566G causes a synonymous alteration in the ATPase 8 reading frame, whereas it causes an isoleucine to valine substitution at position 14 of ATPase 6 (MITOANALYZER, 2003). Both isoleucine and valine are non-polar (Winter *et al.*, 2002). However, the lack of one methyl group in valine compared to isoleucine (Rodwell, 1993), could affect its stereochemistry and interaction with other amino acids in the ATPase 6 polypeptide. The substitution could thus affect the physical properties of the ATPase 6. Isoleucine at position 14 is not conserved and is found in 12 of the species listed in Table 4.7, of which seven are primates. Therefore, it is unlikely that the A8566G alteration is pathogenic, but further analysis is required to confirm this conclusion.

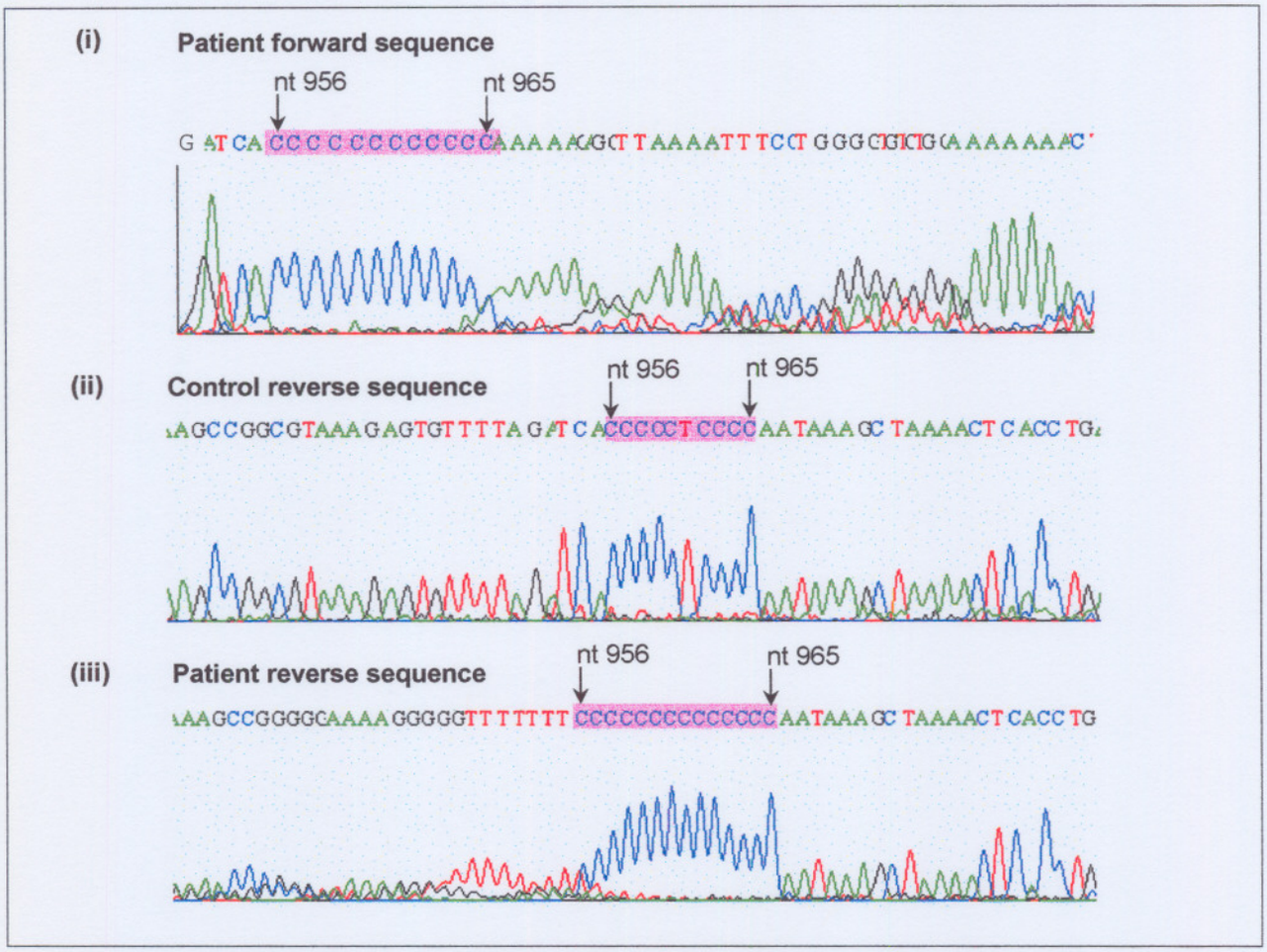
The A13276G alteration results in a methionine to valine substitution at position 314 of the ND5 subunit of NADH dehydrogenase (MITOANALYZER, 2003). Both amino acids are non-polar (Winter *et al.*, 2002), but methionine contains a sulphur atom (Rodwell, 1993). Methionine at position 314 is highly conserved and is present in 31 of the 39 species listed in Table 4.7. However, three of the 12 primates in Table 4.7 do not have methionine at that position, and two (*Gorilla gorilla* and *Hylobates lar*) of the three have valine at position 314. *Pan troglodytes* (chimpanzee) also has a methionine at this position and the chimpanzee is the closest relative to *Homo sapiens sapiens* (Horai *et al.*, 1992). The A13276G alteration was also detected in patient 525. There are similarities in the clinical phenotypes of patients 504 and 525 and they also belong to the same haplogroup. Therefore, the alteration could be a polymorphism, but it could also be an alteration that becomes pathogenic when expressed in the L0 sub-haplogroup. Further analyses should definitely be performed to determine the pathogenicity of the A13276G alteration.

The G15431A alteration in Cyt *b* of complex III results in an alanine to threonine change at position 229 of the polypeptide (MITOANALYZER, 2003). Alanine is non-polar, whereas threonine is polar (Winter *et al.*, 2002). Threonine is also a larger molecule than alanine (Rodwell, 1993) and steric hindrance may occur in the Cyt *b* polypeptide. The substitution can, therefore, influence the physical properties of the polypeptide. Alanine at position 229 of Cyt *b* is not at all conserved and is present in this position in only three species listed in

Table 4.7. Apart from *Homo sapiens sapiens*, alanine is not found in this position in primates. Therefore, despite the change in physical properties, the substitution is probably not pathogenic.

The alteration of most interest was the insertion of approximately five cytosines in 12S rRNA between nt 956 and 965. The electropherograms of the forward and reverse sequence of the insertion are presented in (i) and (iii) in Figure 4.7.

Figure 4.7: Electropherogram of the 956-965insCCCCC insertion in patient 504



The highlighted sequence indicates the position of the five cytosine insertion. nt = nucleotide.

When the first electropherogram of the forward sequence was obtained, as presented in Figure 4.7 (i), contamination was suspected owing to the image of a mixed template downstream of the cytosine repeat sequence. The sequencing reaction of this particular region was successful in all the other individuals and the positive control. The forward sequencing reaction was then repeated three times with PCR amplicons obtained from three separate reactions to exclude contamination. The PCR and sequencing primers utilised were made up new from the stock solutions. However, after the third time of

obtaining an electropherogram as depicted in Figure 4.7 (i), it was clear that contamination was not the problem and alterations at the forward primer annealing site were suspected. The region under investigation was then amplified with flanking forward and reverse primers and these amplicons, along with the original amplicon, were sequenced in the reverse direction. Electropherograms similar to Figure 4.7 (iii) were obtained for all three reverse reactions. Note that in the reverse sequence upstream of the cytosine repeat it also appears to have a mixed template. The sequence downstream of the cytosine repeat could be interpreted and was compared to the RCRS.

Shoffner *et al.* (1993) reported an insertion of approximately five cytosines within the 5'-CCCCCTCCCC-3' between nt 956 and 965 in 12S rRNA. However, Shoffner *et al.* (1993) detected it in a late-onset Alzheimer's Disease (AD) and Parkinson's Disease (PD) patient, whereas the patient included in this investigation is paediatric. The link between mitochondrial dysfunction and oxidative damage and the pathogenesis of AD and PD is still not resolved (Beal, 2003; Carrieri *et al.*, 2001), but the alteration may be pathogenic. Unfortunately the alteration was only briefly discussed by Shoffner *et al.* (1993) and the exact sequence obtained was not provided.

The forward and reverse sequences differ in the amount of cytosines inserted. In the forward sequence only 12 cytosines are visible, compared to 14 in the reverse sequence. Therefore, from the forward sequence it appears as if only two cytosines are inserted between nt 956 and 965 along with a T961C (Ozawa *et al.*, 1995) alteration. From the reverse sequence it appears as if five cytosines are inserted between nt 956 and 965.

The discrepancy in the number of inserted cytosines between the forward and reverse sequences is difficult to explain, but it is possible that it is due to slippage of the sequencing enzyme when sequencing in the forward direction. It is, however, unlikely that the enzyme would slip to the same extent in all three the forward sequencing reactions. Alternatively, it can be due to the cytosine repeat being close to the annealing site of the forward primer. Further analyses with primers amplifying the region with the cytosine repeat more to the middle are required. However, for this reason the sequences in the reverse direction were regarded as more reliable. If that is the case then the 14 cytosines can be explained by an insertion of five cytosines in the position of the thymine at nt 961.

The appearance of the mixed template downstream of the cytosine repeat in the forward sequence and upstream of the repeat in the reverse sequence may be explained by

heteroplasmy. If the insertion is not present in all the mtDNA genomes it will result in two amplicons of different length. Sequencing these simultaneously will appear as a mixed template. This image of heteroplasmy differs from the classical presentation as depicted in Figure 4.11 (ii). The reason is that the alteration under discussion is a heteroplasmic *insertion* and it leads to amplicons of different length being superimposed on one another. The image in Figure 4.11 (ii) depicts the A3243G *substitution* in which only one base is superimposed onto another. However, the 956-965insCCCCC alteration and an alternative method of analysis should definitely be pursued further. Heteroplasmy should be confirmed by RFLP analysis or high resolution electrophoresis. The maternal relatives of patient 504 should be tested to determine if they have the same alteration, as this will give an indication of the pathogenic nature of the insertion.

According to haplogroup analysis, utilising Appendix B, patient 504 is part of sub-haplogroup L0. This is consistent with the high number of alterations detected in this patient. Sub-haplogroup L0 is regarded as the most ancient haplogroup and has the greatest variation (Chen *et al.*, 1995; Mishmar *et al.*, 2003).

4.4.3 Patient 525

Patient 525 is a person of mixed ancestry[1], formerly referred to as "coloured". The patient is female and presented with symptoms of a mitochondrial disorder at the age of seven months. Her clinical phenotype included stunted growth, muscle wasting, microcephaly, and developmental delay. Lactic acidosis and increased organic acids in the urine were detected by biochemical analyses.

PCR amplification of patient 525 was often not successful and this could also be due to the high number of alterations, as discussed for patient 504. Ninety-one alterations, presented in Table A.3, were detected in patient 525. Nine novel alterations were regarded as potentially pathogenic mutations and are presented in Table 4.10. The other 82 alterations are listed in Appendix A, Table A.3. Heteroplasmy could not be detected for any of these alterations.

[1] Mixed ancestry is difficult to define, however, in this case it may include African, European Caucasian, Khoi-San and Malay ancestry.

Table 4.10: Potential pathogenic mutations detected in patient 525

Alteration	Locus	Amino Acid Change	Alteration Type
498delC	Control region	~	Functional effect unknown
C597T	$tRNA^{Phe}$	~	Functional effect unknown
A4225G	ND1	Met → Val	Non-synonymous alteration
5899insC	Non-coding region	~	Functional effect unknown
A7146G	COI	Thr → Ala	Non-synonymous alteration
A12172G	$tRNA^{His}$	~	Functional effect unknown
A13276G	ND5	Met → Val	Non-synonymous alteration
G15930A	$tRNA^{Thr}$	~	Functional effect unknown
T15941C	$tRNA^{Thr}$	~	Functional effect unknown

~ = mutation not found in protein-coding gene, $tRNA^{Phe}$ = gene encoding tRNA phenylalanine, ND1 = gene encoding subunit 1 of NADH dehydrogenase, COI = gene encoding cytochrome *c* oxidase subunit I, $tRNA^{His}$ = gene encoding tRNA histidine, ND5 = gene encoding subunit 5 of NADH dehydrogenase, $tRNA^{Thr}$ = gene encoding tRNA threonine, Met = methionine, Val = valine, Thr = threonine and Ala = alanine.

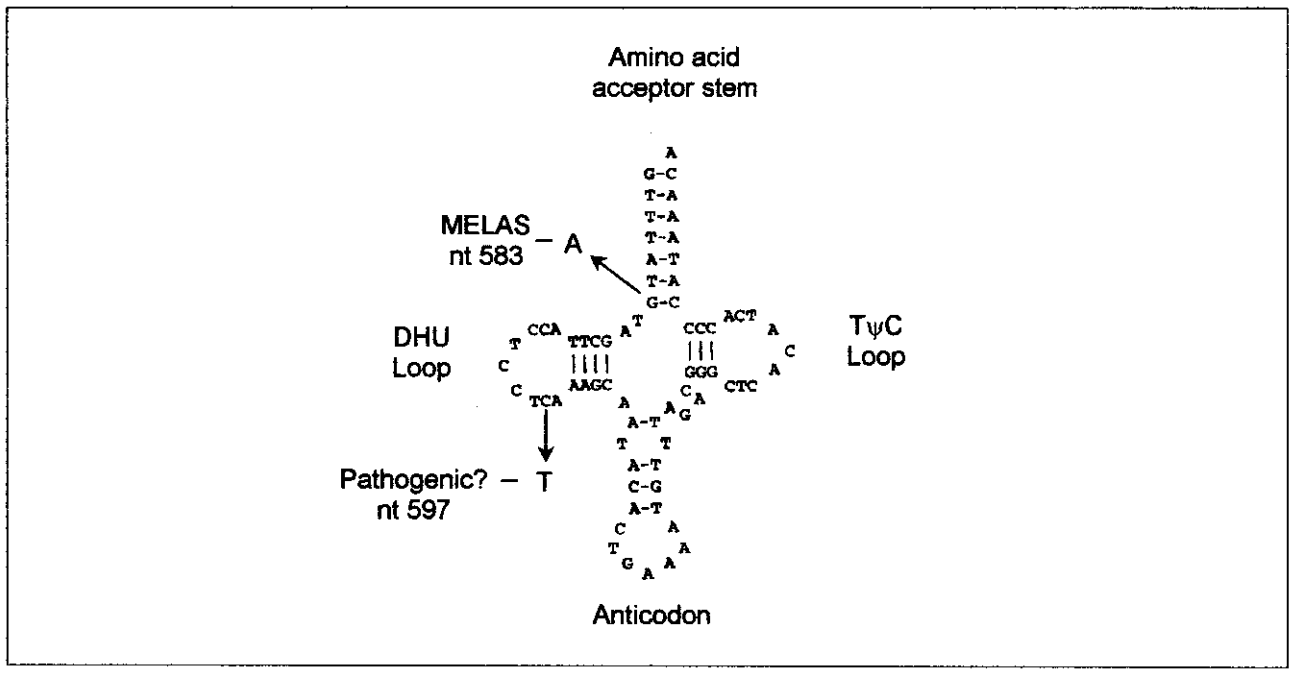
The cytosine deletion at nt 498 and the cytosine insertion at nt 5899 are both in non-coding regions. However, as the functional effects of alterations in these regions are not known, they were confirmed with sequencing in the reverse direction and are regarded as potentially pathogenic. Especially 498delC is important, as the control region is the site of regulation of replication and transcription (Clayton, 1982; Clayton, 1984). IT_{H1} and IT_L, located in HSP and LSP respectively, are situated within 150 bp of each other in the control region (Attardi, 1986). Furthermore, POLG, TFAM, SSB, RNA primers and other proteins not characterised yet, bind at regular intervals throughout the control region (Spelbrink, 2003). If 498delC and 5899insC independently are absent in healthy maternal relatives of patient 525, functional studies will have to be performed to determine their pathogenicity.

Four alterations were detected in tRNA genes of patient 525, namely the C597T alteration in $tRNA^{Phe}$, A12172G in $tRNA^{His}$, and G15930A and T15941C in $tRNA^{Thr}$. As with the non-coding regions it is difficult to determine the functional effect of tRNA gene alterations without performing functional translation studies. Sternberg *et al.* (2001) came to the same conclusion after screening for tRNA sequence variations in 166 patients with suspected mitochondrial disorders. Phylogenetic conservation and heteroplasmic state are generally utilised for functional assessment. However, the Caucasian polymorphic T4336C alteration (Leory and Norby, 1994) was found to be heteroplasmic in muscle of one patient (Sternberg *et al.*, 2001). Some nearly homoplasmic pathogenic mutations may be detected as homoplasmic owing to lack of sensitivity of the detection method (Sternberg *et al.*, 2001). This was also experienced in the investigation presented here, as discussed in

paragraph 4.5. Taking phylogenetic conservation as the determining criteria is also not necessarily ideal. For example the A8344G alteration most commonly associated with MERFF (Shoffner *et al.*, 1990) alters a nucleotide which is only mildly conserved (Sternberg *et al.*, 2001). Therefore, functional assessment remains challenging.

The C597T alteration is located in the dihydrouridine loop of $tRNA^{Phe}$. Another alteration, G583A, in the amino acid acceptor stem of $tRNA^{Phe}$ is associated with MELAS (Hanna *et al.*, 1998). Both alterations are presented in Figure 4.8.

Figure 4.8: The G583A and C597T alterations on the $tRNA^{Phe}$ molecule



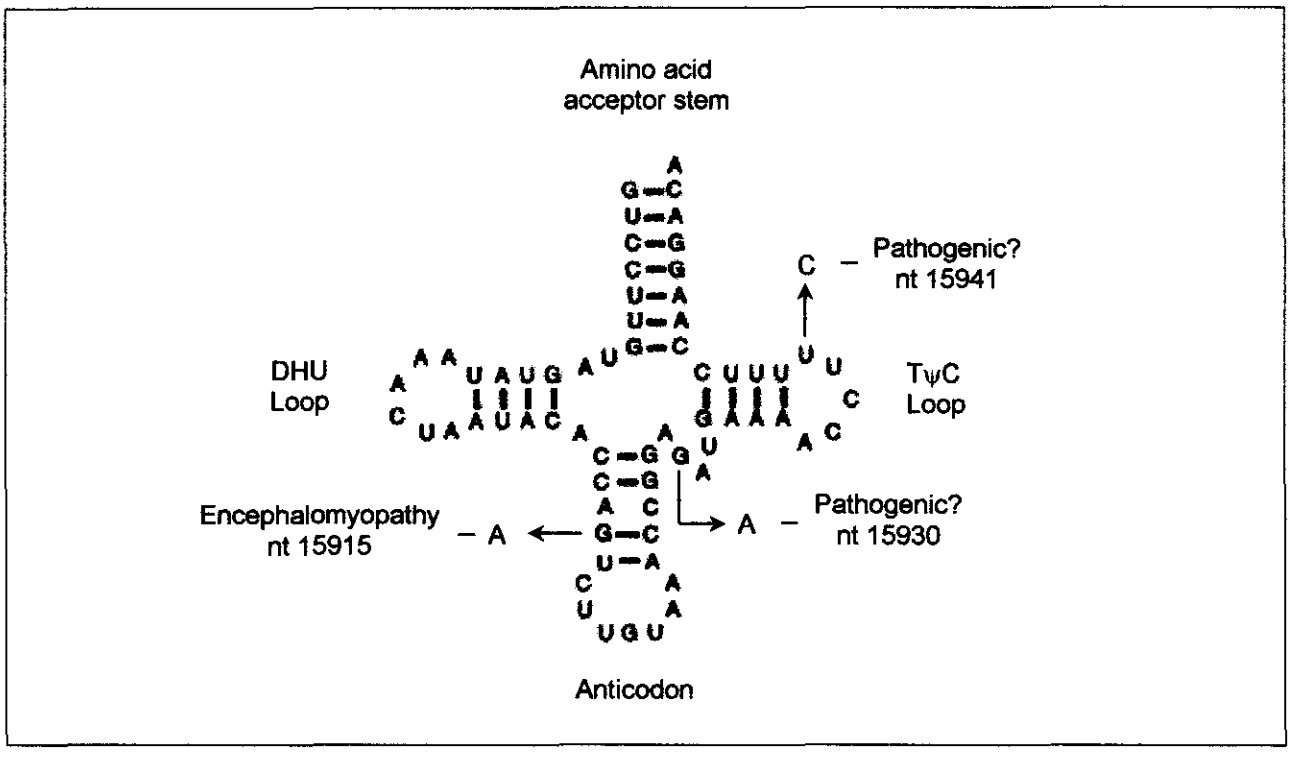
MELAS = mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes, DHU = dihydrouridine, ψ = pseudouridine (5-ribosyl uracil) and nt = nucleotide. Adapted and modified from Kleinle *et al.* (1998).

The G15930A alteration is located in the variable loop of $tRNA^{Thr}$ and the T15941C is situated in the TψC loop, as presented in Figure 4.9. The G15915A alteration in $tRNA^{Thr}$ has been reported to be associated with mitochondrial encephalomyopathy (Nishino *et al.*, 1996).

Three non-synonymous alterations were detected that could possibly be pathogenic. The A4225G alteration results in a methionine to valine substitution at position 307 of the ND1 subunit of NADH dehydrogenase (MITOANALYZER, 2003). As mentioned earlier when the A13276G alteration was discussed, methionine and valine are both non-polar (Winter *et al.*, 2002) but methionine contains a sulphur atom (Rodwell, 1993) and the three dimensional properties are not the same. Methionine is not conserved at position 307 and

is present at this position in only 11 species listed in Table 4.7. Twenty-three species have a leucine at position 307. Nine of the 11 species with methionine at this position are primates, possibly indicating its importance for this group. However, because of the low conservation of methionine, it is not likely that the A4225G alteration is pathogenic.

Figure 4.9: The G15915A, G15930A and T15941C alterations on the tRNA^{Thr} molecule



DHU = dihydrouridine, ψ = pseudouridine (5-ribosyl uracil) and nt = nucleotide. Adapted and modified from Nishino *et al.* (1996).

Both the A7146G alteration in COI and the A13276G alteration in ND5 were also detected in patient 504 and will not be discussed again. However, it is of interest that there are similarities in the clinical phenotypes of patients 504 and 525.

Through haplogroup analysis, utilising Appendix B, it was determined that patient 525 belongs to sub-haplogroup L0. Patient 525 does not have the fourth polymorphism listed in Table B.1, namely A1438G, which defines the sub-haplogroup of haplogroup H which represents the CRS. However, patient 504, who is in sub-haplogroup L0, does have the A1438G alteration. According to Ruiz-Pesini (2003) the A1438G alteration occurred in the ancestral lineage of patient 504 in parallel with the lineage defining haplogroup H, but not in the lineage of patient 525. Patients 504 and 525 have many of the same alterations and it is, therefore, not surprising that they are classified in the same sub-haplogroup.

4.4.4 Patient 1301

Patient 1301 is a black African female who presented with the clinical phenotype of a mitochondrial disorder at birth. She had muscle wasting and weakness, myopathy, developmental delay, external ophthalmoplegia, ptosis and hypermobile joints. Her biochemical profile included lactic acidosis, organic acids and amino acids in the urine.

Fifty-five alterations were detected in the mtDNA from muscle of patient 1301 as presented in Table A.4. However, 49 of these were reported polymorphisms and five were novel synonymous alterations. Only one novel T2416C alteration in 16S rRNA, as presented in Table 4.11, is potentially pathogenic. The T2416C alteration was also detected in DNA isolated from blood of patient 1301, but heteroplasmy could not be detected.

Table 4.11: Potentially pathogenic mutation detected in patient 1301

Alteration	Locus	Amino Acid Change	Alteration Type
T2416C	16S rRNA	~	Functional effect unknown

16S rRNA = gene encoding 16S ribosomal RNA and ~ = mutation not found in protein-coding gene.

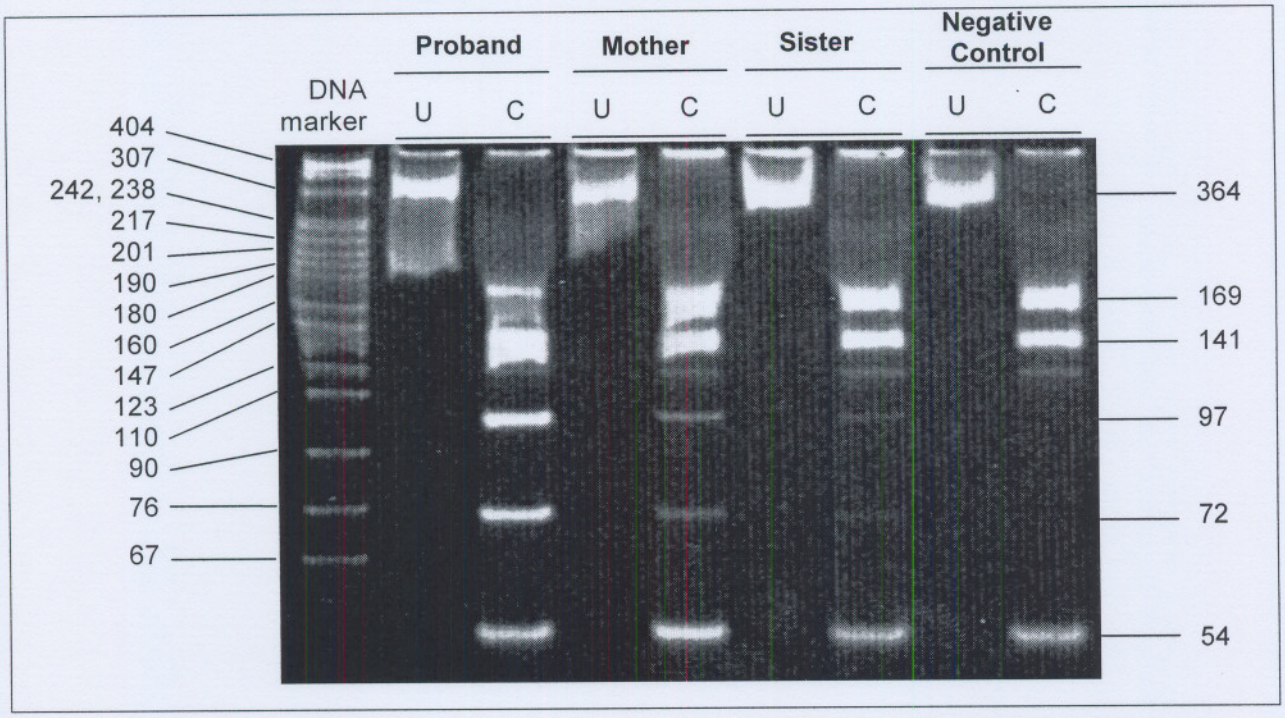
As for alterations in tRNAs and in non-coding regions, it is difficult to determine the functional effect of alterations in rRNA without performing functional studies. However, as discussed in paragraph 2.3.4, 16S rRNA forms part of the large 39S subunit of the mitochondrial 55S ribosomes (Attardi and Ojala, 1971) and synthesis of the 13 mtDNA proteins (Anderson *et al.*, 1981) are dependent on this translation apparatus. Pathogenic mutations of rRNA can, therefore, influence all proteins synthesised. If healthy maternal relatives of the patient do not have the T2416C alteration, it should definitely be investigated further. Haplogroup analysis, as discussed in Appendix B, revealed that patient 1301 belongs to sub-haplogroup L2.

4.5 HETEROPLASMY

A major limitation of the cycle sequencing technique is the inability to detect low levels of heteroplasmy. Prosser (2001) detected the MELAS A3243G mutation through RFLP analysis in a single proband and subsequently in two maternal relatives with different levels of heteroplasmy. In the photograph presented in Figure 4.10, the 97 bp and 72 bp RFLP fragments are visible after *Hae* III digestion of the $tRNA^{Leu(UUR)}$ amplicon of the proband (411), the mother (431) and the sister (432). The two determining RFLP

fragments had a high intensity in the severely affected proband (411), intermediate intensity in the mildly affected mother (431) and low intensity in the asymptomatic sister (432).

Figure 4.10: Different levels of heteroplasmy of the A3243G mutation in one family analysed through RFLP analysis



12% polyacrylamide gel electrophoresed at 250 V for 2 hours and 30 minutes in 1 X TBE buffer. Stained with SYBR® Gold for 30 min. DNA marker = pBR322 digested with *Msp* I, U = uncut, C = cut and NEG = negative control. Adapted from Prosser (2001).

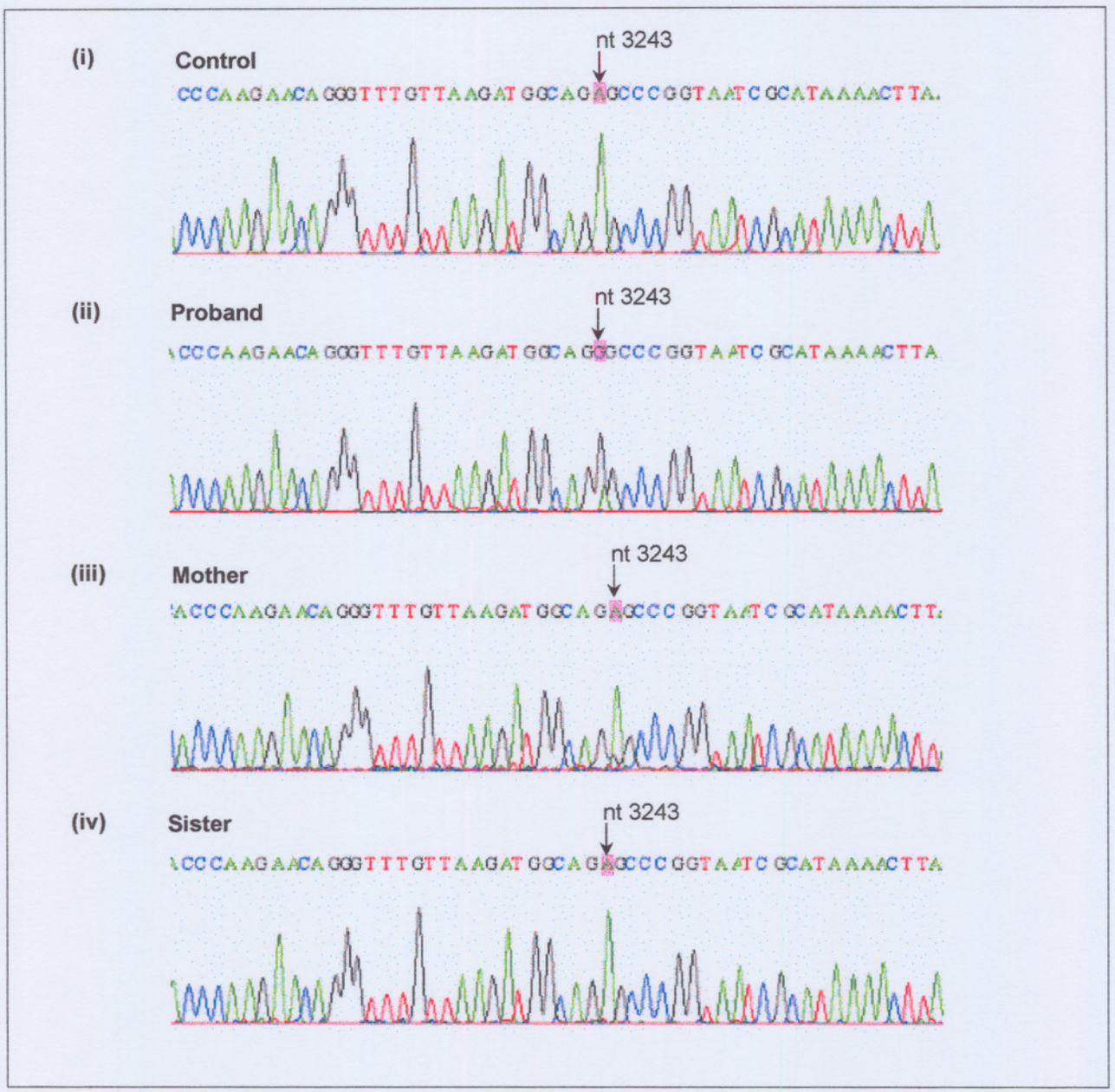
Subsequent to the initial investigation, DNA isolated from blood of the above three individuals were sequenced to visualise the heteroplasmy on an electropherogram for sequence data analysis purposes. The electropherogram obtained is presented in Figure 4.11.

As presented in Figure 4.11 (iv), the low level of the A3243G mutation in the sister (432) was not visible with cycle sequencing. Although the sister did not present with any symptoms suggestive of a mitochondrial disorder, she may develop them later if the percentage of abnormal mtDNA in relevant tissues increases to threshold levels. Individual 432 is also at risk of transmitting the mutation to her offspring.

From Figures 4.10 and 4.11 it is clear that there is a minimum level below which heteroplasmy cannot be detected by cycle sequencing. This threshold will be even higher if the electropherogram contains some background noise. This minimum level can be determined by preparing a series of samples consisting of different ratios of wild type and

A3243G mutant amplicon. In this manner the limitation of the cycle sequencing technique will be quantified in the next phase of this investigation.

Figure 4.11: Different levels of heteroplasmy of the A3243G mutation in one family analysed by cycle sequencing



The highlighted sequence indicates the position of the A3243G mutation.

Heteroplasmy could not be detected (and confirmed) for the alterations discussed in Chapter 4. However, heteroplasmy was concluded to be present, based on the mixed template sequences observed in the forward and reverse directions of the 956-965insCCCCC alteration detected in patient 504. This conclusion should be confirmed, as discussed in paragraph 4.4.2. As the cycle sequencing technique was demonstrated not to be sensitive enough to detect low levels of heteroplasmy, it could in

fact have been present in some of the detected alterations. It is surprising that higher levels of heteroplasmy, as presented in Figure 4.11 (ii), were not detected in even a single sample. In general homoplasmic alterations are regarded as not pathogenic, as they would not be viable had they been pathogenic. However, it is speculated that a mild pathogenic mutation that alters the structure of a protein, tRNA or rRNA enough to cause disease but not death, can be homoplasmic.

It is unfortunate that heteroplasmy could not be determined owing to the limitation of the sequencing technique. However, now that this limitation is known, the minimum detectable level of heteroplasmy should be quantified and more sensitive techniques can be applied in future where heteroplasmy is suspected.

No reported pathogenic mutation could be detected in the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ or the ATPase 6 genes of 34 patients, nor in the entire mtDNA of patients 386, 504, 525 and 1301. However, a total of 103 different reported polymorphisms, 44 different novel synonymous alterations and 17 different potentially pathogenic mutations were detected. The 17 potentially pathogenic mutations included eight novel non-synonymous alterations and nine alterations in tRNA, rRNA and non-coding regions, of which the functional effect was unknown. All 17 potentially pathogenic mutations should be investigated further. However, it is essential that the maternal relatives of the following five alterations should be tested, as these alterations were identified as more likely to be pathogenic. The G7979A non-synonymous alteration detected in COII of patient 386, the 956-965insCCCCC in 12S rRNA of patient 504, the non-synonymous A13276G alteration detected in the ND5 subunit of patients 504 and 525, the T2416C alteration detected in 16S rRNA of patient 1301 and the C3254T alteration detected in the $tRNA^{Leu(UUR)}$ of patient 1314. If the healthy maternal relatives of these patients are negative for the specific alterations, functional studies should be conducted to characterise the pathogenicity of these alterations.

CHAPTER FIVE

CONCLUSION

In the investigation presented here 42 samples from 34 South African paediatric patients, clinically diagnosed with mitochondrial disorders, were analysed for 24 pathogenic mtDNA mutations in the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and ATPase 6 mitochondrial genes. Whole mtDNA genome sequencing and screening were performed on one patient from this group who was negative for the 24 reported mutations and who had an MDC score of eight. Three other paediatric patients with MDC scores of eight, previously screened and determined to be negative for the 24 pathogenic mutations (Prosser, 2001), were also included in the whole mtDNA genome screening programme. The whole mtDNA genome screening strategy was applied to determine whether the phenotype of these patients is determined or influenced by an alteration in their mtDNA or not. In addition to the original aims of the investigation, haplogroup analysis was performed on these four patients, as the haplogroup of an individual can influence disease expression (Torroni, 2000).

Blood and muscle samples were available for eight patients, and DNA from both tissues was analysed. In one patient (1336) a non-synonymous alteration substituting a conserved amino acid was detected in DNA from muscle, but not from blood. As discussed in paragraph 4.2.3, this alteration could not be confirmed after sequencing in the reverse direction, indicating an artefact in the original experiment. This experience accentuated the critical importance of sequence confirmation. Alteration differences between the two types of tissue could not be confirmed in the mtDNA of any of the above seven patients.

In all samples analysed a total of 103 different reported polymorphisms, 44 different novel synonymous alterations and 17 different potentially pathogenic mutations were detected. None of the detected alterations was reported pathogenic mutations. The 17 potentially pathogenic mutations included eight novel non-synonymous alterations and nine alterations in tRNA, rRNA and non-coding regions, of which the functional effect was unknown. The non-synonymous alterations were evaluated according to the evolutionary conservation of the substituted amino acid and possible alteration in physical properties of the polypeptide. Heteroplasmic status was also considered, but heteroplasmy could not be excluded owing to a limitation of the detection technique. For this reason it is possible that

low levels of heteroplasmy could have been present but remained undetected. Homoplasmic alterations are generally regarded as non-pathogenic, thus polymorphic or neutral, because they would not have been viable had they been pathogenic. However, a mild pathogenic mutation that alters the structure of a protein, tRNA or rRNA enough to cause disease but not death, can possibly be homoplasmic. As discussed in paragraph 2.6.2.3.1, the G3460A, G11778A and T14484C mutations, which are prominent risk factors for LHON, have been detected as homoplasmic and heteroplasmic mutations (Brown *et al.*, 2002). For this reason the homoplasmic nature of the alterations detected in this study should not be the sole criteria upon which the decision is made not to investigate alterations further.

If the functional effect of an alteration was not known during this investigation, it was regarded as potentially pathogenic in that it was confirmed with sequencing in the reverse direction and reported to be investigated further. These alterations with unknown functional effect included alterations of tRNA, rRNA and non-coding regions that have not been reported previously. The alterations in non-coding regions included the cytosine deletion and insertion detected at nt 498 and nt 5899 respectively of patient 525. As the mtDNA has extreme economy of organisation with almost no introns (Anderson *et al.*, 1981) it was argued that the non-coding regions that remained in the mtDNA may have an important function which is not yet known, such as regulation. The 498delC alteration could be especially important, as the control region is the site of regulation of replication and transcription (Clayton, 1982; Clayton, 1984).

The investigation presented here forms part of an extended research programme in which the mtDNA of 94 patients has been screened to date. Most were screened for alterations in the mitochondrial $tRNA^{Leu(UUR)}$ (including the A3243G mutation), $tRNA^{Lys}$ and ATPase 6 genes. Whole mtDNA genome sequencing was performed in a previous investigation for six patients (Prosser *et al.*, 2001) and, therefore, whole mtDNA genome data from 10 patients are currently available. In this entire group a single patient harboured a reported mutation (Prosser, 2001), namely the A3243G mutation in the $tRNA^{Leu(UUR)}$ gene, associated with MELAS (Goto *et al.*,1990). Prosser (2001) also detected the A3243G mutation in two maternal relatives of the proband. Other pathogenic mutations reported in other populations have not, to date, been detected in this cohort of South African patients. This is in contrast with the investigation by Sternberg *et al.* (2001) performed in France where 21 patients with the MELAS A3243G alteration and two patients with the MERRF A8344G alteration were identified from a cohort of 166 patients with suspected

mitochondrial disorders. There is a striking difference between the results obtained by Sternberg *et al.* (2001) and those of the extended research programme, of which this study is part. Sternberg *et al.* (2001) detected the A3243G alteration in 12.65% of the patients included in their study, whereas this alteration was present in only 1.06% of the cohort of 94 South African patients.

There could be several explanations why these reported mutations were not observed in the group of South African patients. In the patients of whom only the $tRNA^{Leu(UUR)}$, $tRNA^{Lys}$ and ATPase 6 mitochondrial genes were screened, it is possible that the pathogenic mutations could be located in areas of the mtDNA genome that, as yet, have not been screened. Whole mtDNA genome screening will have to be performed on these patients before alterations in the mtDNA can be excluded from the aetiology of their disorders.

Another possibility is that the mutations causing the mitochondrial dysfunction are nuclear-encoded. For 12 years it was thought that complex I comprised only 43 subunits (Walker, 1992) and the RC 90 subunits (Larsson and Clayton, 1995). However, recently Carroll *et al.* (2002) identified three additional subunits of complex I and currently the RC is believed to consist of 93 different subunits. Only 13 of the approximately 93 RC subunits (Larsson and Clayton, 1995; Carroll *et al.*, 2002) are encoded by mtDNA. The balance of the subunits is encoded by the nucleus (Larsson and Clayton, 1995), and most mitochondrial disorders are thus as a result of mutations in nuclear-encoded genes (Shanske *et al.*, 2001). However, the nuclear genome is more complex than the mtDNA genome. Furthermore, screening for pathogenic mutations in nuclear DNA would not involve only the regions coding for RC subunits. As discussed in paragraph 2.3.2.1, mtDNA is dependent on nuclear-encoded proteins for maintenance and faithful propagation (Spelbrink, 2003). However, many proteins involved in mammalian mtDNA replication have not been fully characterised (Spelbrink, 2003). Apart from the replication machinery a nuclear pathogenic mutation can involve a large number of loci, for example the nuclear-encoded rRNAs of the mitochondria-specific 55S ribosomes, proteins involved in the mitochondrial import of nuclear encoded RC subunits, replication factors as well as transcription factors and proteins involved in post-transcriptional and post-translational processing of nuclear as well as mtDNA encoded transcripts, to name a few. For this reason, alterations in the mtDNA are generally first excluded prior to screening nuclear DNA for pathogenic mutations. However, it is also possible that after screening all the nuclear genes involved in the mitochondria the pathogenic mutations would still not be detected. Interaction and communication between the nucleus and mitochondria are still

not fully understood. However, disruption of these networks will also result in pathology, and may not necessarily be evident from DNA sequencing analysis.

For most of the patients included in the investigation only blood samples could be obtained owing to the invasive nature of muscle biopsies. It is therefore possible that the mutations are not present in blood because of mitotic segregation of mtDNA genomes throughout cellular differentiation. However, DNA from the muscle of 18 patients was analysed and whole mtDNA genome sequencing was performed for one of these individuals. None of the reported mutations was found in any of these samples. It is possible that other tissues such as liver, brain or heart muscle harboured the mutation in some of these patients, but these tissues were unfortunately not available. Although none of the reported mutations was detected during this investigation, 17 different potentially pathogenic mutations were identified in mtDNA from either blood or muscle.

It is also possible that a novel genetic aetiology may account for the majority of mitochondrial phenotypes observed in the South African population (Olckers *et al.*, 2001). Most of the patients included in this investigation were of African descent. However, most of the reported mutations were originally detected in non-African haplogroups. Haplogroup L mtDNAs have significantly lower frequencies of non-synonymous versus synonymous alterations compared to temperate and arctic haplogroups (Ruiz-Pesini *et al.*, 2004). As people migrated out of Africa into Europe and Siberia some haplogroups were exposed to severe cold, which led to adaptive selection of mtDNA genotypes with partially uncoupled OXPHOS. OXPHOS uncoupling implies that less ATP is available, but more heat is produced. Uncoupling mutations thus decrease the production of ATP and reactive oxygen species, but increase the chances of energetic failure. Therefore, these haplogroups will have an increased susceptibility to diseases of energy deficiency, but in turn serve as a protective factor from neurodegenerative diseases and aging (Ruiz-Pesini *et al.*, 2004). It is possible that alterations regarded as polymorphisms in other populations may become pathogenic when expressed upon an African or Southern African genetic background. An example of this phenomenon is the expression of LHON in three haplogroup J families that harboured only the mild T10663C mutation. It was proposed that the haplogroup J background played an important role in the clinical manifestation of certain LHON mutations (Brown *et al.*, 2002). The hypothesis of a novel genetic aetiology is arguably the main reason for not detecting the reported pathogenic mutations in the South African population.

Another possible explanation is the inability of the sequencing strategy utilised to detect low levels of heteroplasmy. In the investigation by Prosser (2001) it was demonstrated that cycle sequencing could replace RFLP analysis, thus enabling screening for many mutations in the same sequence. It was argued that the sequencing strategy provides a more reliable molecular diagnosis (Prosser, 2001). However, in the investigation presented here it was demonstrated that low levels of heteroplasmy that were detectable in RFLP analysis could not be detected by cycle sequencing, as discussed in paragraph 4.5. Cycle sequencing is more specific, but less sensitive than RFLP analysis. Excluding the 956-965insCCCCC alteration identified in patient 504, as discussed in paragraph 4.4.2, all the alterations detected in this investigation appeared to be homoplasmic. However, in view of the limitation of the sequencing strategy, it is possible that low levels of heteroplasmy could have been present in some cases. Unfortunately, large-scale screening such as whole mtDNA genome analysis, will currently be too time-consuming and costly utilising a technique such as RFLP. An alternative detection strategy providing the same amount of data as cycle sequencing, along with the ability to detect heteroplasmy at every nucleotide position, would be ideal. Cycle sequencing is a qualitative method, but if the quantity of every nucleotide could be determined, and the ratio calculated, heteroplasmy would be detectable at each position. However, such a quantitative sequencing strategy is unfortunately not yet available. Detection methods such as denaturing gradient gel electrophoresis or denaturing high performance liquid chromatography differentiate between two species of mtDNA, but do not indicate the number of differences between the species. It is, therefore, not ideal for the analysis of mtDNA, which often harbours many differences, as presented in paragraphs 4.4.2 and 4.4.3 for patients 504 and 525 respectively.

Further analysis will have to be conducted for the 17 different potentially pathogenic mutations detected in this investigation. The presence of these mutations in maternal relatives of the probands will have to be tested in order to exclude the possibility of these alterations being inherited neutral mutations. If these alterations are not present in the healthy family members the alterations arose *de novo* and could possibly be pathogenic. However, functional studies will have to be performed to confirm their pathogenic nature. Of the 17 potentially pathogenic mutations, five were identified as more likely to be pathogenic, as discussed for each patient in Chapter 4. The G7979A non-synonymous alteration detected in COII of patient 386, the 956-965insCCCCC in 12S rRNA of patient 504, the non-synonymous A13276G alteration detected in the ND5 subunit of both patients 504 and 525, the T2416C alteration detected in 16S rRNA of patient 1301 and the C3254T

alteration detected in the $tRNA^{Leu(UUR)}$ of patient 1314, should definitely be investigated further.

During interpretation of the data generated in this investigation it was difficult to determine whether detected alterations were reported or novel pathogenic mutations and polymorphisms. Worldwide the MITOMAP (2003) database is utilised to verify whether alterations have been reported and whether they are pathogenic or known polymorphisms. However, certain reported alterations are not reflected on MITOMAP (2003), such as the Caucasian polymorphic T4336C alteration (Leroy and Norby, 1994) referred to in paragraph 4.4.3. Furthermore, certain alterations are indicated as polymorphisms or pathogenic mutations for which functional studies have not been performed, such as the "polymorphic" C3254T and "pathogenic" C3254G alteration, discussed in paragraph 4.2.1. Furthermore, all the publications cited on MITOMAP (2003) do not always contain the alterations associated with it on the database, but are often only related literature. This hampered identification of the original publication, as it was not necessarily the earliest reference cited. In the investigation presented here, MITOMAP (2003) was often utilised for evaluation of detected alterations and to determine whether they had been reported or not. Other databases, such as PubMed Central (2004), were also searched for detected alterations, but no reports could be found. Polymorphisms in particular are usually not included in publication abstracts and can, therefore, not be traced. All the above complicated the ascertainment of the novel status of alterations detected in this investigation.

It is clear that an official reference system is required to which genetic data can be submitted for classification. The system should be self-regulatory in that a list of strict criteria is adhered to with each submission in order to classify alterations as "pathogenic mutations", "polymorphisms" or "functional effect unknown". An editor with outstanding knowledge of the field should evaluate the data on a regular basis. Proof of functional analyses with approved detection strategies should be available on request. Criteria such as amino acid conservation and heteroplasmic state could be regarded as additional motivation for a particular classification. However, as discussed in paragraph 4.4.3, the high conservation or heteroplasmic state is unfortunately not always indicative of a pathogenic mutation. Therefore, functional studies should be a requirement prior to alterations being classified as pathogenic. If certain analyses are still outstanding, alterations should rather be classified as "functional effect unknown" to ensure that no misconceptions are created.

Mitochondrial alterations can be regarded as pathological when they are absent in a large number of controls from the same ethnic group. The presence of an OXPHOS defect and segregation with the disease strengthen the possibility of the alteration being pathogenic (Holt, 2003). In the investigation presented here the data were compared with only one control and the RCRS. It is unfortunate that the RCRS had to be utilised as a reference sequence for patients from haplogroup L. The CRS was principally derived from a single European individual with haplogroup H (Andrews *et al.*, 1999). As there is no African reference sequence available, any unreported deviation from the RCRS of which the functional effect was not known had to be regarded as potentially pathogenic. However, many of these "potentially pathogenic" alterations are probably polymorphisms. If an African reference sequence had been available, true polymorphisms could have been excluded. In turn this implies that unnecessary analyses could have been avoided and more patients screened. The reason for so few alterations being reported for patient 386 was that this patient was of the same haplogroup as the individual from which the CRS was compiled. Therefore, the single potential pathogenic mutation detected in this patient, G7979A in the COII gene, is more likely to be pathogenic than many of the alterations detected in the patients belonging to haplogroup L. An African reference sequence is definitely a necessity for effective evaluation of alterations detected in this population.

As none of the reported pathogenic mutations were detected, none of the patients screened could definitely be diagnosed on a molecular level. However, 17 novel potentially pathogenic mutations were detected during this investigation. Their pathogenic nature should be confirmed in the future through functional analyses.

Despite the lack of an optimal detection strategy, applicable reference sequence and alteration evaluation system, important limitations of data interpretation and techniques that are utilised worldwide were identified. If these are addressed, mitochondrial disorders could be investigated more efficiently.

Elucidating the aetiology of mitochondrial disease in the South African population will create the opportunity to develop specific drugs for patients suffering from these disorders. In this manner patients who are currently only supplemented with substrates of the RC could be treated more efficiently in future. However, for this to become a reality it is vital that the genetic basis of these disorders be established in this population.

It is evident that the South African population is unique on the genomic level. In particular the results generated during this study support the hypothesis that the aetiology of mitochondrial disorders is unique in this population (Olckers *et al.*, 2001). It is conceivable that this uniqueness may be attributed to the fact that alterations could interact with haplogroup-specific polymorphisms to create the clinical phenotypes of mitochondrial disorders that are observed in the South African population.

REFERENCES

6.1 GENERAL REFERENCES

Adams P.L. and Turnbull D.M. Disorders of the electron transport chain. *J. Inherit. Metab. Dis.*, **19**, 463-469, 1996.

Albertini A.M., Hofer M., Calos M.P. and Miller J.H. On the formation of spontaneous deletions: the importance of short sequence homologies in the generation of large deletions. *Cell*, **29**, 319-328, 1982.

Anderson S., Bankier A.T., Barrell B.G., De Bruijn M.H.L., Coulson, A.R., Drouin J., Eperon I.C., Nierlich D.P., Roe B.A., Sanger F., Schreier P.H., Smith A.J.H., Staden R. and Young I.G. Sequence and organization of the human mitochondrial genome. *Nature*, **290**, 457-465, 1981.

Andreu A.L., Bruno C., Hadjigeorgiou G.M., Shanske S. and DiMauro S. Polymorphic variants in the human mitochondrial cytochrome *b* gene. *Mol. Genet. Metab.*, **67**, 49-52, 1999.

Andrews R.M., Kubacka I., Chinnery P.F., Lightowlers R.N., Turnbull D.M. and Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.*, **23**, 147, 1999.

Aquadro C.F. and Greenberg B.D. Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. *Genetics*, **103**, 287-312, 1983.

Arnaudo E., Dalakas M., Shanske S., Moraes C.T., DiMauro S. and Schon E.A. Depletion of muscle mitochondrial DNA in AIDS patients with zidovudine-induced myopathy. *Lancet*, **337**, 508-510, 1991.

Attardi G. and Ojala D. Mitochondrial ribosomes in HeLa cells. *Nature New Biol.*, **229**, 133-136, 1971.

Attardi G. Animal mitochondrial DNA: an extreme example of genetic economy. *Int. Rev. Cytol.*, **93**, 93-145, 1985.

Attardi G. The elucidation of the human mitochondrial genome: a historical perspective. *Bioessays*, **5** (1), 34-39, 1986.

Attardi G. The human mitochondrial genetic system. In: Mitochondrial DNA in Human Pathology. DiMauro S. and Wallace D.C, editors. New York: Raven Press Ltd., 9-25, 1993.

Awadalla P. Eyre-Walker A. and Maynard Smith J. Linkage disequilibrium and recombination in hominid mitochondrial DNA. *Science*, **286**, 2524-2525, 1999.

Ballinger S.W., Schurr T.G., Torroni A., Gan Y.Y., Hodge J.A., Hassan K., Chen K.H. and Wallace D.C. Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient mongoloid migrations. *Genetics*, **130** (1), 139-152, 1992.

Bauer M.F., Gempel K., Hofmann S., Jaksch M., Philbrook C. and Gerbitz K.D. Mitochondrial Disorders. A diagnostic challenge in clinical chemistry. *Clin. Chem. Lab. Med.*, **37**, 855-876, 1999.

Beal M.F. Bioenergetic approaches for neuroprotection in Parkinson's disease. *Ann. Neurol.* **53**, S39-S48, 2003.

Belogrudov G. and Hatefi Y. Catalytic sector of complex I (NADH:ubiquinone oxidoreductase): subunit stoichiometry and substrate-induced conformation changes. *Biochemistry*, **33**, 4571-4576, 1994.

Benda C. Ueber die Spermatogenese der Vertebraten und höherer Ervertebraten. II. Theil: Die Histiogenese der Spermien (The spermatogenesis of vertebrates and higher invertebrates. The histology of sperm). *Arch. Anal. Physiol., (Physiol. Abt.)*, 393-398, 1898.

Bendall K.E., Macaulay V.A., Baker, J.R. and Sykes B.C., Heteroplasmic point mutations in the human control region. *Am. J. Hum. Genet.*, **59**, 1276-1287, 1996.

Bernier F.P., Boneh A., Dennett X., Chow C.W., Cleary M.A. and Thorburn D.R. Diagnostic criteria for respiratory chain disorders in adults and children. *Neurology*, **59**, 1406-1411, 2002.

Bindoff L.A., Howell N., Poulton J., McCullough D.A., Morten K.J., Lightowlers R.N., Turnbull D.M. and Weber K. Abnormal RNA processing associated with a novel tRNA mutation in mitochondrial DNA. *J. Biol. Chem.*, **268** (26), 19559-19564, 1993.

Bogenhagen D. and Clayton D.A. The number of mitochondrial deoxyribonucleic acid genomes in mouse L and human HeLa cells. *J Biol Chem.*, **249**, 7991-7995, 1974.

Brown M.D., Starikovskaya E., Derbeneva O., Hosseini S., Allen J.C., Mikhailovskaya I.E., Sukernik R.I. and Wallace D.C. The role of mtDNA background in disease expression: a new primary LHON mutation associated with Western Eurasian haplogroup *J. Hum. Genet.*, **110**, 130-138, 2002.

Brown M.D., Zhadanov S., Allen J.C., Hosseini S., Newman N.J., Atamonov V.V., Mikhailovskaya I.E., Sukernik R.I. and Wallace D.C. Novel mtDNA mutations and oxidative phosphorylation dysfunction in Russian LHON families. *Hum. Genet.*, **109**, 33-39, 2001.

Campbell P.N. and Smith A.D. Biochemistry illustrated, 3rd ed. London: Churchill Livingstone, 183-188, 1993.

Cann R.L. and Wilson A.C. Length mutations in human mitochondrial DNA. *Genetics*, **184**, 699-711, 1983.

Carrieri G., Bonafé M., De Luca M., Rose G., Varcasia O., Bruni A., Maletta R., Nacmias B., Sorbi S., Corsonello F., Feraco E., Andreev K.F., Yashin A.I., Franceschi C. and De Benedictis G. Mitochondrial DNA haplogroups and APOE4 allele are non-independent variables in sporadic Alzheimer's disease. *Hum. Genet.*, **108**, 194-198, 2001.

Carrodeguas J.A. and Bogenhagen D.F. Protein sequences conserved in prokaryotic aminoacyl-tRNA synthetases are important for the activity of the processivity factor of human mitochondrial DNA polymerase. *Nucleic Acids Res.*, **28**, 1237-1244, 2000.

Carroll J., Shannon R.J., Fearnley I.M., Walker J.E. and Hirst J. Definition of the nuclear encoded protein composition of bovine heart mitochondrial complex I. *J. Biol. Chem.*, **277**, 50311-50317, 2002.

Chambers H.F. and Sande M.A. Antimicrobial agents (continued). The Aminoglycosides. In: Goodman and Gilman's The Pharmacological Basis of Therapeutics, 9th ed. Hardman J.G. and Limbird L.E., editors. New York: The McGraw-Hill Companies, 1103-1121, 1996.

Chang D.P. and Clayton D.A. Priming of human mitochondrial DNA replication occurs at the light-strand promoter. *Proc. Natl. Acad. Sci. U.S.A.*, **82**, 351-355, 1985.

Chen Y.S., Torroni A., Excoffier L., Santachiara-Benerecetti A.S. and Wallace D.C. Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. *Am. J. Hum. Genet.*, **57**, 133-149, 1995.

Chinnery P.F. and Samuels D.C. Relaxed replication of mtDNA: a model with implications for the expression of disease. *Am. J. Hum. Genet.*, **64**, 1158-1165, 1999.

Chinnery P.F., Brown D.T., Andrews R.M., Singh-Kler R., Riordan-Eva P., Lindley J., Applegarth D.A., Turnbull D.M. and Howell N. The mitochondrial ND6 gene is a hot spot for mutations that cause Leber's hereditary optic neuropathy. *Brain*, **124**, 209-218, 2001.

Chinnery P.F., Johnson M.A., Wardell T.M., Singh-Kler R., Hayes C., Brown D.T., Taylor R.W., Bindoff L.A. and Turnbull D.M. The epidemiology of pathogenic mitochondrial DNA mutations. *Ann. Neurol.*, **48**, 188-193, 2000.

Clayton D.A. Replication and transcription of vertebrate mitochondrial DNA. *Annu. Rev. Cell Biol.*, **7**, 453-478, 1991.

Clayton D.A. Replication of animal mitochondrial DNA. *Cell*, **28**, 693-705, 1982.

Clayton D.A. Transcription of the mammalian mitochondrial genome. *Annu. Rev. Biochem.*, **53**, 573-594, 1984.

Dalakas M.C., Illa I., Pezeshkpour G.H., Laukaitis J.P., Cohen B. and Griffin J.L. Mitochondrial myopathy caused by long-term zidovudine therapy. *N. Engl. J. Med.*, **322**, 1098-1105, 1990.

De Vivo D.C. Mitochondrial DNA Defects. In: Mitochondrial DNA in Human Pathology. DiMauro S. and Wallace D.C., editors. New York: Raven Press Ltd., 39-52, 1993.

De Vivo D.C., Haymond M.W., Obert K.A., Nelson J.S. and Pagliara A.S. Defective activation of the pyruvate dehydrogenase complex deficiency in subacute necrotizing encephalomyelopathy (Leigh disease). *Ann. Neurol.*, **6**, 483-484, 1979.

De Vries D.D., Van Engelen B.G.M., Gabreëls F.J.M., Ruitenbeek W. and Van Oost B.A. A second missense mutation in the mitochondrial ATPase 6 gene in Leigh's syndrome. *Ann. Neurol.*, **34**, 410-412, 1993.

DiMauro S., Bonilla E., Zeviani M., Nakagawa M. and De Vivo D.C. Mitochondrial myopathies. *Ann. Neurol.*, **17**, 521-538, 1985.

DiRienzo A. and Wilson A.C. Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 1597-1601, (1991).

Doda J.N., Wright C.T and Clayton D.A. Elongation of displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. *Proc. Natl. Acad. Sci. U.S.A.*, **75**, 677-681, 1981.

Engel W.K. and Cunningham C.G. Rapid examination of muscle tissue: an improved trichrome stain method for fresh-frozen biopsy section. *Neurology*, **13**, 919-923,1963.

Enriquez J.A., Chomyn A. and Attardi G. MtDNA mutation in MERRF syndrome causes defective aminoacylation of $tRNA^{Lys}$ and premature translation termination. *Nat. Genet.*, **10**, 47-55, 1995.

Eyre-Walker A. and Awadalla P. Does human mtDNA recombine? *J. Mol. Evol.*, **53**, 430-435, 2001.

Eyre-Walker A., Smith N.H. and Maynard-Smith J. How clonal are human mitochondria? *Proc. R. Soc. Lond. B. Biol. Sci.*, **266**, 477-483, 1999.

Falconer D.S. Introduction to quantitative genetics, 2nd ed. Essex: Longman Scientific & Technical, 22-45, 1981.

Fan L., Sanschagrin P.C., Kaguni L.S. and Kuhn L.A. The accessory subunit of mtDNA polymerase shares structural homology with aminoacyl-tRNA synthetases: implication for a dual role as a primer recognition factor and processivity clamp. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 9527-9532, 1999.

Ferri K.F., Jacotot E., Blanco J., Esté J.A. and Kroemer G. Mitochondrial control of cell death induced by HIV-encoded proteins. *Ann. N.Y. Acad. Sci.*, **926**, 149-164, 2000.

Fisher R.P., Lisowky T., Parisi M.A. and Clayton D.A. DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein. *J. Biol.Chem.*, **267**, 3358-3367, 1992.

Fukuhara N., Tokigushi S., Shirakawa K. and Tsubaki T. Myoclonus epilepsy associated with ragged-red fibres (mitochondrial abnormalities): Disease entity or syndrome? Light and electron microscopic studies of two cases and review of literature. *J. Neurol. Sci.*, **47**, 117-133, 1980.

Gelfand R. and Attardi G. Synthesis and turnover of mitochondrial ribonucleic acid in HeLa cells: the mature ribosomal and messenger ribonucleic acid species are metabolically unstable. *Mol. Cell. Biol.*, **1**, 497-511, 1981.

Giles R.E., Blanc H., Cann H.M. and Wallace D.C. Maternal inheritance of human mitochondrial DNA. *Proc. Natl. Acad. Sci. U.S.A*, **77**, 6715-6719, 1980.

Goto Y.I., Nonaka I. and Horai S. A mutation in the $tRNA^{Leu(UUR)}$ gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature*, **348**, 651-653, 1990.

Goto Y.I., Nonaka I. and Horai S. A new mtDNA mutation associated with mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes (MELAS). *Biochim. Biophys. Acta.*, **1097**, 238-240, 1991.

Goto Y.I., Tojo M., Tohyama J., Horai S., and Nonaka I. A novel point mutation in the mitochondrial $tRNA^{Leu(UUR)}$ gene in a family with mitochondrial myopathy. *Ann. Neurol.*, **31**, 672-675,1992.

Goto Y.I., Tsugane K., Tanabe Y., Nonaka I. and Horai S. A new point mutation at nucleotide pair 3291 of the mitochondrial $tRNA^{Leu(UUR)}$ gene in a patient with mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes (MELAS). *Biochem. Biophys. Res. Commun.*, **202**, 1624-1630, 1994.

Gravel M., Melancon P. and Brakier-Gingras L. Cross-linking of streptomycin to the 16S ribosomal RNA of *Escherichia coli*. *Biochemistry*, **26**, 6227-6232, 1987.

Hadjigeorgiou G.M., Kim S.H., Fischbeck K.H., Andreu A.L., Berry G.T., Bingham P., Shanske S., Bonilla E. and DiMauro S. A new mitochondrial DNA mutation (A3288G) in the $tRNA^{Leu(UUR)}$ gene associated with familial myopathy. *J. Neurol. Sci.*, **164**, 153-157, 1999.

Hagelberg E. Recombination or mutation rate heterogeneity? Implication for mitochondrial Eve. *Trends Genet.*, **19**, 84-90, 2003.

Hagelberg E., Goldman N., Lió P., Whelan S., Schiefenhövel W., Clegg J.B. and Bowden D.K. Evidence for mitochondrial DNA recombination in a human population of island Melanesia. *Proc. R. Soc. Lond. B. Biol. Sci.*, **266**, 485-492, 1999.

Hanna M.G., Nelson I.P., Morgan-Hughes J.A. and Wood N.W. MELAS: a new disease associated mitochondrial DNA mutation and evidence for further genetic heterogeneity. *J. Neurol. Neurosurg. Psychiatry.* **65**, 512-517, 1998.

Hess J.F., Parisi M.A., Bennett J.L. and Clayton D.A. Impairment of mitochondrial transcription termination by a point mutation associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature*, **351**, 236-239, 1991.

Hirano M. and Pavlakis S.G. Mitochondrial myopathy, encephalopathy, lactic acidosis and strokelike episodes (MELAS): Current concepts. *J. Child Neurol.*, **9**, 4-13, 1994.

Hofhaus G., Weiss H. and Leonard K. Electron microscopic analysis of the peripheral and membrane parts of mitochondrial NADH dehydrogenase (complex I). *J. Mol. Biol.*, **221**, 1027-1043, 1991.

Holt I.J. Genetics of Mitochondrial Diseases. Holt I.J., editor. Oxford: Oxford University Press, xxi-xxiii, 2003.

Holt I.J., Harding A.E. and Morgan-Hughes J.A. Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. *Nature*, **331**, 717-719, 1988.

Holt I.J., Harding A.E., Petty R.K.H. and Morgan-Hughes J.A. A new mitochondrial disease associated with mitochondrial DNA heteroplasmy. *Am. J. Hum. Genet.*, **46**, 428-433, 1990.

Holt I.J., Lorimer H.E. and Jacobs H.T. Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA. *Cell*, **100**, 515-524, 2000.

Horai S. and Hayasaka K. Intraspecific nucleotide sequence differences in the major noncoding region of human mitochondrial DNA. *Am. J. Hum. Genet.*, **46**, 828-842, 1990.

Horai S., Satta Y., Hayasaka K., Kondo R., Inoue T., Ishida T., Hayashi S. and Takahata N. Man's place in *Hominoidea* revealed by mitochondrial DNA genealogy. *J. Mol. Biol.*, **35**, 32-43, 1992.

Horai S., Kondo R., Nakagawa-Hattori Y., Hayashi S., Sonoda S. and Tajima K. Peopling of the Americas, founded by four major lineages of mitochondrial DNA. *Mol. Biol. Evol.*, **10**, 23-47, 1993.

Houshmand M., Lindberg C., Moslemi A.R., Oldfors A. and Holme E. A novel heteroplasmic point mutation in the mitochondrial $tRNA^{Lys}$ gene in a sporadic case of mitochondrial encephalomyopathy: de novo mutation and no transmission to the offspring. *Hum. Mutat.*, **13**, 203-209, 1999.

Howell N., Kubacka I., Halvorson S., Howell B., McCullough D.A. and Mackey D. Phylogenetic analysis of the mitochondrial genomes from Leber hereditary optic neuropathy pedigrees. *Genetics*, **140**, 285-302, 1995.

Huoponen K., Lamminen T., Juvonen V., Aula P., Nikoskelainen E. and Savontaus J.L. The spectrum of mitochondrial DNA mutations in families with Leber hereditary optic neuroretinopathy. *Hum. Genet.*, **92**, 379-384, 1993.

Johns D.R., Drachman D.B. and Hurko O. Identical mitochondrial DNA deletion in blood and muscle. *Lancet*, **1**, 393-394, 1989.

Jorde L.B., Bamshad M.J., Watkins W.S., Zenger R., Fraley A.E., Krakowiak P.A., Carpenter K.D., Soodyall H., Jenkins T. and Rogers A.R. Origins and affinities of modern humans: a comparison of mitochondrial and nuclear data. *Am. J. Hum. Genet.*, **57**, 523-538, 1995.

Kakuda T.N. Pharmacology of nucleoside and nucleotide reverse transcriptase inhibitor-induced mitochondrial toxicity. *Clin. Ther.*, **22**, 685-708, 2000.

Kameoka T., Isotani H., Tanaka K., Azukari K., Fujimura Y., Shiota Y., Sasaki E., Majima M., Furukawa K., Haginomori S., Kitaoka H. and Ohsawa N. Novel mitochondrial DNA mutation in $tRNA^{Lys}$ (8296A→G) associated with diabetes. *Biochem. Biophys. Res. Commun.*, **245**, 523-527, 1998.

Kawarai T., Kawakami H., Kozuka K., Izumi Y., Matsuyama Z., Watanabe C., Kohriyama T. and Nakamura S. A new mitochondrial DNA mutation associated with mitochondrial myopathy $tRNA^{Leu(UUR)}$ 3254C-to-G. *Neurology*, **49**, 598-600, 1997.

Kearns T.P. and Sayre G.P. Retinitis pigmentosa, external ophthalmoplegia and complete heart block. *Arch. Ophthalmol.*, **60**, 280-289, 1958.

Kleinle S., Schneider V., Moosmann P., Brandner S., Krahenbuhl S. and Liechti-Gallati S. A novel mitochondrial $tRNA^{Phe}$ mutation inhibiting anticodon stem formation associated with a muscle disease. *Biochem. Biophys. Res. Commun.*, **247**, 112-115, 1998.

Kösel S., Egensperger R., Mehraein P. and Graeber M.B. No association of mutations at nucleotide 5460 of mitochondrial NADH dehydrogenase with Alzheimer's disease. *Biochem. Biophys. Res. Commun.*, **203**, 745-749, 1994.

Krings M., Stone A., Schmitz R.W., Krainitzki H., Stoneking M. and Pääbo S. Neanderthal DNA sequences and the origin of modern humans. *Cell*, **90**, 19-30, 1997.

Kruse B., Narasimhan N. and Attardi G. Termination of transcription in human mitochondria: identification and purification of a DNA binding protein factor that promotes termination. *Cell*, **58**, 391-397, 1989.

Lamminen T., Majander A., Juvonen V., Wikström M., Aula P., Nikoskelainen E. and Savontaus M.L. A mitochondrial mutation at nt 9101 in the ATP synthase 6 gene associated with deficient oxidative phosphorylation in a family with Leber hereditary optic neuroretinopathy. *Am. J. Hum. Genet.*, **56**, 1238-1240, 1995.

Lancaster C.R.D. and Kröger A. Succinate:quinone oxidoreductases: new insights from X-ray crystal structures. *Biochim. Biophys. Acta.*, **1459**, 422-431, 2000.

Larsson N.G. and Clayton D.A. Molecular genetic aspects of human mitochondrial disorders. *Annu. Rev. Genet.*, **29**, 151-178, 1995.

Larsson N.G., Holme E., Kristiansson B., Oldfors A. and Tulinius M. Progressive increase of the mutated mitochondrial DNA fraction in Kearns-Sayre syndrome. *Pediatr. Res.*, **28**, 131-136, 1990.

Larsson N.G., Tulinius M.H., Holme E., Oldfors A., Andersen O., Wahlström J. and Aasly J. Segregation and manifestations of the mtDNA $tRNA^{Lys}$ $A→G^{(8344)}$ mutation of myoclonus epilepsy and ragged-red fibers (MERRF) syndrome. *Am. J. Hum. Genet.*, **51**, 1201-1212, 1992.

Leigh D. Subacute necrotizing encephalomyelopathy in an infant. *J. Neurol. Neurosurg. Psychiatry*, **14**, 216-221, 1951.

Leory D.and Norby S. A new human mtDNA polymorphism: tRNA(Gln)/4336 (T→C). *Clin. Genet.*, **45**, 109-110, 1994.

Lewis W. and Dalakas M.C. Mitochondrial toxicity of antiviral drugs. *Nat. Med.*, **1**, 417-422, 1995.

Longley M.J., Ropp P.A., Lim S.E. and Copeland W.C. Characterisation of the native and recombinant catalytic subunit of human DNA polymerase γ: identification of residues critical for exonuclease activity and dideoxynucleotide sensitivity. *Biochemistry*, **37**, 10529-10539, 1998.

Margulis L. Origin of eukaryotic cells. New Haven: Yale University Press, 1970.

Martin A.M., Hammond E., Nolan D., Pace C., Den Boer M., Taylor L., Moore H., Martinez O.P., Christiansen F.T. and Mallal S. Accumulation of mitochondrial DNA mutations in human immunodeficiency virus-infected patients treated with nucleoside-analogue reverse-transcriptase inhibitors. *Am. J. Hum. Genet.*, **72**, 549-560. 2003.

Mayes P.A. The respiratory chain and oxidative phosphorylation. In: Harper's Biochemistry, 23rd ed. Murray R.K., Granner D.K., Mayes P.A. and Rodwell V.W., editors. Connecticut: Appleton & Lange, 119-130, 1993.

Mishmar D., Ruiz-Pesini E., Golik P., Macaylay 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. U.S.A.*, **100**, 171-176, 2003.

Mitchell P. Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature*, **191**, 144-148, 1961.

Moazed D. and Noller H.F. Interaction of antibiotics with functional sites in 16S ribosomal RNA. *Nature*, **327**, 389-394, 1987.

Montoya J., Gaines G. and Attardi G. The pattern of transcription of the human mitochondrial rRNA genes reveals two overlapping transcription units. *Cell*, **34**, 151-159, 1983.

Moraes C.T, DiMauro S., Zeviani M., Lombes A., Shanske S., Miranda A.F., Nakase H., Bonilla E., Werneck L.C., Servidei S., Nonaka I., Koga Y., Spiro A.J., Brownell K.W., Schmidt B., Schotland D.L.,

Zupanc M., De Vivo D.C., Schon E.A. and Rowland L.P. Mitochondrial DNA deletions in progressive ophthalmoplegia and Kearns-Sayre syndrome. *N. Engl. J. Med.*, **320**, 1293-1299, 1989.

Moraes C.T., Ciacci F., Bonilla E., Jansen C., Hirano M., Rao N., Lovelace R.E., Rowland L.P., Schon E.A. and DiMauro S. Two novel pathogenic mitochondrial DNA mutations affecting organelle number and protein synthesis. *J. Clin. Invest.*, **92**, 2906-2915, 1993.

Moraes C.T., Shanske S., Tritschler H.J., Aprille J.R., Andreetta F., Bonilla E., Schon E.A. and DiMauro S. mtDNA Depletion with variable tissue expression: a novel genetic abnormality in mitochondrial diseases. *Am. J. Hum. Genet.*, **48**, 492-501, 1991.

Morten K.J., Cooper J.M., Brown G.K., Lake B.D., Pike D. and Poulton J. A new point mutation associated with mitochondrial encephalomyopathy. *Hum. Mol. Genet.*, **2** (12), 2081-2087, 1993.

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

Munnich A., Rotig A., Chretien D., Cormier V., Bourgeron T., Bonnnefont, J.-P., Saudubray J.-M. and Rustin P. Clinical presentation of mitochondrial disorders in childhood. *J. Inherit. Metab. Dis.*, **19**, 521-527, 1996.

Newman N.J. Leber's hereditary optic neuropathy: new genetic considerations. *Arch. Neurol.*, **50**, 540-548, 1993.

Nishino I., Seki A., Maegaki Y., Takeshita K., Horai S., Nonaka I. and Goto Y. A novel mutation in the mitochondrial $tRNA^{Thr}$ gene associated with a mitochondrial encephalomyopathy. *Biochem. Biophys. Res. Commun.*, **225**, 180-185, 1996.

Obayashi T., Hattori K., Sugiyama S., Tanaka M., Tanaka T., Itoyama S., Deguchi H., Kawamura K., Koga Y., Toshima H., Takeda N., Nagano M., Ito T. and Ozawa T. Point mutations in mitochondrial DNA in patients with hypertrophic cardiomyopathy. *Am. Heart J.*, **124**, 1263-1269, 1992.

Olckers A., Prosser D., Wallace D.C., Brown M.D. and Smuts I. Mitochondrial myopathies: a South African perspective. 51st Annual meeting of the American Society of Human Genetics, San Diego, U.S.A., 2001.

Ozawa M., Nishino I., Horai S., Nonaka I. and Goto Y. Myoclonus epilepsy associated with ragged-red fibers: A G-to-A mutation at nucleotide pair 8363 in mitochoncrial $tRNA^{Lys}$ in two families. *Muscle Nerve*, **20**, 271-278, 1997.

Ozawa T., Katsumata K., Hayakawa M., Tanaka M., Sugiyama S., Tanaka T., Itoyama S., Nunoda S. and Sekiguchi M. Genotype and phenotype of severe mitochondrial cardiomyopathy: a recipient of heart transplantation and the genetic control. *Biochem. Biophys. Res. Commun.*, **207**, 613-620, 1995.

Ozawa T., Tanaka M., Sugiyama S., Ino H., Ohno K., Hattori K., Ohbayashi T., Ito T., Deguchi H., Kawamura K., Nakane Y. and Hashiba K. Patients with idiopatic cardiomyopathy belong to the same mitochondrial DNA gene family of Parkinson's disease and mitochondrial encephalomyopathy. *Biochem. Biophys. Res. Commun,* **177**, 518-525, 1991a.

Ozawa T., Tanaka M., Ino H., Ohno K., Sano T., Wada Y., Yoneda M., Tanno Y., Miyatake T., Tanaka T., Itoyama S., Ikebe S., Hattori N. and Mizuno Y. Distinct clustering of point mutations in mitochondrial DNA among patients with mitochondrial encephalomyopathies and with Parkinson's disease. *Biochem. Biophys. Res. Commun,* **176**, 938-946, 1991b.

Passarino G., Semino O., Modiano G. and Santa-Chiara-Benerecetti A.S. COII/$tRNA^{Lys}$ intergenic 9-bp deletion and other mtDNA markers clearly reveal that the Tharus (southern Nepals) have Oriental affinities. *Am. J. Hum. Genet.*, **53**, 609-518, 1993.

Preisig O. Inqaba Biotechnical Industries (Pty) Ltd. Pretoria, South Africa. Personal communication, 2003.

Prezant T.R., Agapian J.V., Bohlman M.C., Bu X., Oztas S., Qiu W.Q., Arnos K.S., Cortopassi G.A., Jaber L., Rotter J.I., Shohat M. and Fischel-Ghodsian N. Mitochondrial ribosomal RNA mutation associated with both antibiotic-induced and non-syndromic deafness. *Nat. Genet.*, **4**, 289-294, 1993.

Procaccio V. Centre for Molecular and Mitochondrial Medicine and Genetics, University of California at Irvine, Irvine, U.S.A. Personal communication, 2003.

Prosser D.O. Mitochondrial DNA (mtDNA) mutations in patients with suspected myoclonic epilepsy and ragged red muscle fibres (MERRF), Leigh syndrome (LS), and mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes (MELAS) [dissertation]. Pretoria: University of Pretoria, 109-145, 2001.

Prosser D.O., Smuts I., Brown M.D., Wallace D.C. and Olckers A. Molecular analysis of mitochondrial myopathies within South Africa. Annual Congress of the Neurological Association of South Africa, Durban, South Africa, 2001.

Redd A.J., Takezaki N., Sherry S.T., McGarvey S.T., Sofro A.S.M. and Stoneking M. Evolutionary History of the COII/tRNALys intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific. *Mol. Biol. Evol.*, **12**, 604-615, 1995.

Richter C., Park J.W. and Ames B.N. Normal oxidative damage to mitochondrial and nuclear DNA is extensive. *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 6465-6467, 1988.

Rodwell V.W. Amino Acids. In: Harper's Biochemistry, 23rd ed. Murray R.K., Granner D.K., Mayes P.A. and Rodwell V.W., editors. Connecticut: Appleton & Lange, 23-32, 1993.

Ropp P.A. and Copeland W.C. Cloning and characterization of the human mitochondrial DNA polymerase, DNA polymerase gamma. *Genomics*, **36**, 449-458, 1996.

Rose G., Passarino G., Carrieri G., Altomare K., Greco V., Bertolini S., Bonafe M., Franceschi C., De Benedictis G. Paradoxes in longevity: sequence analysis of mtDNA haplogroup J in centenarians. *Eur. J. Hum. Genet.*, **9**, 701-707, 2001.

Rosing H.S., Hopkins L.C., Wallace D.C., Epstein C.M. and Weidenheim K. Maternally inherited mitochondrial myopathy and myoclonic epilepsy. *Ann. Neurol.*, **17**, 228-237, 1985.

Rötig A., Cormier V., Blanche S., Bonnefont J.-P., Ledeist F., Romero N., Schmitz J., Rustin P., Fischer A., Saudubray J.-M. and Munnich A. Pearson's marrow-pancreas syndrome. A multisystem mitochondrial disorder in infancy. *J. Clin. Invest.*, **86**, 1601-1608, 1990.

Ruiz-Pesini E. Centre for Molecular and Mitochondrial Medicine and Genetics, University of California at Irvine, Irvine, U.S.A. Personal communication, 2003.

Ruiz-Pesini E., Mishmar D., Brandon M., Procaccio V. and Wallace D.C. Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science*, **303**, 223-226, 2004.

Ruvolo M., Zehr S., von Dornum M., Pan D., Chang B. and Lin J. Mitochondrial COII sequences and modern human origins. *Mol. Biol. Evol.*, **10**, 1115-1135, 1993.

Saccone C., Lanave C., Pesole G. and Sbisà E. Peculiar features and evolution of mitochondrial genome in mammals. In: Mitochondrial DNA in Human Pathology. DiMauro S. and Wallace D.C, editors. New York: Raven Press Ltd., 27-37, 1993.

Sambrook J. and Russel D.W. Molecular cloning, a laboratory manual, 3rd ed. New York: Cold Spring Harbor Laboratory Press, A8.20-A8.21, 2001.

Sanger F., Nicklen S. and Coulson A.R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 5463-5467, 1977.

Schatz G., Haslbrunner E. and Tuppy H. Deoxyribonucleic acid associated with yeast mitochondria. *Biochem. Biophys. Res. Commun.*, **15**, 127-135, 1964.

Scholte H.R. The biochemical basis of mitochondrial diseases. *J. Bioenerg. Biomembr.*, **20**, 161-191, 1988.

Schon E.A. Mitochondria. In: Mitochondrial DNA in Human Pathology. DiMauro S. and Wallace D.C, editors. New York: Raven Press Ltd., 1-7, 1993.

Schwartz M. and Vissing J. Paternal inheritance of mitochondrial DNA. *N. Engl. J. Med.*, **347**, 576-580, 2002.

Senior A.E. ATP synthesis by oxidative phosphorylation. *Physiol. Rev.*, **147**, 417-430, 1988.

Shadel G.S. and Clayton D.A. Mitochondrial transcription initiation, variation and conservation. *J. Biol. Chem.*, **268**, 16083-16086, 1993.

Shanske A.L., Shanske S. and DiMauro S. The other human genome. *Arch. Pediatr. Adolesc. Med.*, **155**, 1210-1216, 2001.

Shapira Y., Cederbaum S.D., Cancilla P.A., Nielsen D. and Lippe B.M. Familial poliodystrophy, mitochondrial myopathy and lactate acidemia. *Neurology*, **25**, 614-621, 1975.

Shields G.F., Hecker K., Voevoda M.I. and Reed J.K. Absence of the Asian-specific region V mitochondrial marker in Native-Beringians. *Am. J. Hum. Genet.*, **50**, 758-765, 1992.

Shoffner J.M., Bialer M.G., Pavlakis S.G., Lott M., Kaufman A., Dixon J., Teichberg S. and Wallace D.C. Mitochondrial encephalomyopathy associated with a single nucleotide pair deletion in the mitochondrial $tRNA^{Leu(UUR)}$ gene. *Neurology*, **45**, 286-292, 1995.

Shoffner J.M., Brown M.D., Torroni A., Lott M.T., Cabell M.F., Mirra S.S., Beal M.F., Yang C.C., Gearing M., Salvo R., Watts R.L., Juncos J.L., Hansen L.A., Crain B.J., Fayad M., Reckord C.L. and Wallace D.C. Mitochondrial DNA variants observed in Alzheimer Disease and Parkinson Disease Patients. Genomics, **17**, 171-184, 1993.

Shoffner J.M., Lott M.T., Lezza A.M.S., Seibel P., Ballinger S.W. and Wallace D.C. Myoclonic epilepsy and ragged-red fibre disease (MERRF) is associated with a mitochondrial DNA $tRNA^{Lys}$ mutation. *Cell*, **61**, 931-937, 1990.

Silva W.A. Jr., Bonatto S.L., Holanda A.J., Ribeiro-Dos-Santos, A.K., Paixao B.M., Goldman G.H., Abe-Sandes K., Rodriguez-Delfin L., Barbosa M., Paco-Larson M.L., Petzl-Erler M.L., Valente V., Santos S.E. and Zago, M.A. Correction: mitochondrial DNA variation in Ameridians. *Am. J. Hum. Genet.*, **72**, 1346-1348, 2003.

Silvestri G., Moraes C.T., Shanske S., Oh S.J. and DiMauro S. A new mtDNA mutation in the $tRNA^{Lys}$ gene associated with myoclonic epilepsy and ragged-red fibres (MERRF). *Am. J. Hum. Genet.*, **51**, 1213-1217, 1992.

Silvestri G., Santorelli F.M., Shanske S., Whitley C.B., Schimmenti L.A., Smith S.A. and DiMauro S. A new mtDNA mutation in the $tRNA^{Leu(UUR)}$ gene associated with maternally inherited cardiomyopathy. *Hum. Mutat.*, **3**, 37-43, 1994.

Spagnalo M., Tomelleri G., Vattemi G., Filosto M., Rizzuto N. and Tonin P. A new mutation in the mitochondrial $tRNA^{Ala}$ gene in a patient with opthalmoplegia and dysphagia. *Neuromuscul. Disord.*, **11**, 481-484, 2001.

Spelbrink J.N. Replication, repair, and recombination of mitochondrial DNA. In: Genetics of Mitochondrial Diseases. Holt I.J., editor. Oxford: Oxford University Press, 3-25, 2003.

Spelbrink J.N., Li F.Y., Tiranti V., Nikali K., Yuan Q.P., Tariq M., Wanrooij S., Garrido N., Comi G., Morandi L., Santoro L., Toscano A., Fabrizi G.M., Somer H., Croxen R., Beeson D., Poulton J., Suomalainen A., Jacobs H.T., Zeviani M. and Larsson C. Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localised in mitochondria. *Nat. Genet.*, **28**, 223-231, 2001.

Sternberg D., Chatzoglou E., Laforêt P., Fayet G., Jardel C., Blondy P., Fardeau M., Amselem S., Eymard B. and Lombès A.. Mitochondrial DNA transfer RNA gene sequence variations in patients with mitochondrial disorders. *Brain*, **124**, 984-994, 2001.

Sternberg D., Danan C., Lombes A., Laforet P., Girodon E., Goossens M. and Amselem S. Exhaustive scanning approach to screen all the mitochondrial tRNA genes for mutations and its application to the investigation of 35 independent patients with mitochondrial disorders. *Hum.Mol.Genet.*, **7**, 33-42, 1998.

Strachan, T. and Read A.P. Chromosome structure and function. In: Human Molecular Genetics. BIOS Scientific Publishers Limited, 33-59, 1998.

Stringer C.B. and Andrews P. Genetic and fossil evidence for the origin of modern humans. *Science*, **239**, 1263-1268, 1988.

Sudoyo H., Suryadi H., Lertrit P., Pramoonjago P., Lyrawati D. and Marzuki S. Asian-specific mtDNA backgrounds associated with the primary G11778A mutation of Leber's hereditary optic neuropathy. *J. Hum. Genet.*, **47**, 594-604, 2002.

Sugimoto N., Nakano S., Yoneyama M. and Honda K. Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes. *Nucleic Acids Res.*, **24**, 4501-4505, 1996.

Sutovsky P., Moreno R.D., Ramalho-Santos J., Dominko T., Simerly C. and Schatten G. Ubiquitin tag for sperm mitochondria. *Nature*, **402**, 371-372, 1999.

Suzuki Y., Suzuki S., Hinokio Y., Chiba M., Atsumi Y., Hosokawa K., Shimada A., Asahina T. and Matsuoka K. Diabetes associated with a novel 3264 mitochondrial $tRNA^{Leu(UUR)}$ mutation. *Diabetes Care*, **20** (7), 1138-1140, 1997.

Sweeney M.G., Bundey S., Brockington M., Poulton K.R., Winer J.B. and Harding A.E. Mitochondrial myopathy associated with sudden death in young adults and a novel mutation in the mitochondrial DNA leucine transfer $RNA^{(UUR)}$ gene. *Q. J. Med.*, **86**, 709-713, 1993.

Taanman J. The mitochondrial genome: structure, transcription, translation and replication. *Biochim. Biophys. Acta.*, **1410**, 103-123, 1999.

Tatuch Y., Christodoulou J., Feigenbaum A., Clarke J.T.R., Wherret J., Smith C., Rudd N., Petrova-Benedict R. and Robinson B.H. Heteroplasmic mtDNA mutation (T-G) at 8993 can cause Leigh disease when the percentage of abnormal mtDNA is high. *Am. J. Hum. Genet.*, **50**, 852-858, 1992.

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

Thyagarajan D., Bressman S., Bruno C., Przedborski S., Shanske S., Lynch T., Fahn S. and DiMauro S. A novel mitochondrial 12S rRNA point mutation in parkinsonism, deafness and neuropathy. *Ann. Neurol.*, **48** (5), 730-736, 2000.

Thyagarajan D., Shanske S., Vazquez-Memije M., De Vivo D., and DiMauro S. A novel mitochondrial ATPase 6 point mutation in familial bilateral striatal necrosis. *Ann. Neurol.*, **38**, 468-472,1995.

Tiranti V., Carrara F., Confalonieri P., Mora M., Maffei R.M., Lamantea E. and Zeviani M. A novel mutation (8342G→A) in the mitochondrial $tRNA^{Lys}$ gene associated with progressive external ophthalmoplegia and myoclonus. *Neuromuscul. Disord.*, **9**, 66-71, 1999.

Torroni A. mtDNA haplogroups in human populations and disease studies, *J. Cult. Heritage*, **1**, 33-34, 2000.

Torroni A., Bandelt H.J., D'Urbano L., Lahermo P., Moral P., Sellitto D., Rengo C., Forster P., Savontaus M.L., Bonne-Tamir B., and Scozzari R. mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. *Am. J. Hum. Genet.*, **62**, 1137-1152, 1998.

Torroni A., Schurr T.G., Yang C.C., Szathmary E.J., Williams R.C., Schanfield M.S., Troup G.A., Knowler W.C., Lawrence D.N., Weiss K.M. and Wallace D.C. Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. *Genetics*, **130** (1), 153-162, 1992.

Triepels R.H., Van den Heuvel L.P., Trijbels J.M. and Smeitink J.A. Respiratory Chain Complex I Deficiency. *Am. J. Med. Genet.*, **106**, 37-45, 2001.

Van Coster R., Lombes A., De Vivo D.C., Chi T.L., Dodson W.E., Rothman S., Orrechio E.J., Grover W., Berry G.T., Schwartz J.F., Habib A. and DiMauro S. Cytochrome *c* oxidase-associated Leigh syndrome: phenotypic features and pathogenetic speculations. *J. Neurol. Sci.*, **104**, 97-111, 1991.

Van den Ouweland J.M.W., Lemkes H.H.P.J., Trembath R.C., Ross R., Velho G., Cohen D., Froguel P. and Maassen J.A. Maternally inherited diabetes and deafness is a distinct subtype of diabetes and associates with a single point mutation in the mitochondrial $tRNA^{Leu(UUR)}$ gene. *Diabetes*, **43**, 746-751, 1994.

Verma A., Piccoli D.A., Bonilla E., Berry G.T., DiMauro S. and Moraes C.T. A novel mitochondria G8313A mutation associated with prominent initial gastrointestinal symptoms and progressive encephaloneuropathy. *Pediatr. Res.*, **42**, 448-454, 1997.

Vittecoq D., Jardel C., Barthélémy C., Escaut L., Cheminot N., Chapin S., Sternberg D., Maisonobe T. and Lombès A. Mitochondrial damage associated with long-term antiretroviral treatment: associated alteration or causal disorder? *J. Acquir. Immune Defic. Syndr.*, **31**, 299-308, 2002.

Walker J.E. Determiniation of the structures of respiratory enzyme complexes from mammalian mitochondria. *Biochim. Biophys. Acta.*, **1271**, 221-227, 1995.

Walker J.E. The NADH: ubiquinone oxidoreductase (complex I) of respiratory chains. *Q. Rev. Biophys.*, **25**, 253-324, 1992.

Walker J.E., Lutter R., Dupuis A. and Runswick M.J. Identification of the subunits of F_1F_0-ATPase from bovine heart mitochondria. *Biochemistry*, **30**, 5369-5378, 1991.

Walker U.A., Collins S. and Byrne E. Respiratory chain encephalomyopathies: a diagnostic classification. *Eur. Neurol.*, **36**, 260-267, 1996.

Wallace D.C. and Lott M.T. Maternally inherited diseases. In: Mitochondrial DNA in Human Pathology. DiMauro S. and Wallace D.C, editors. New York: Raven Press Ltd., 63-83, 1993.

Wallace D.C. Centre for Molecular Medicine, Emory University School of Medicine, Atlanta, U.S.A. Personal communication, 1999.

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

Wallace D.C., Lott M.T., Torroni A. and Shoffner J.M. Report of the committee on human mitochondrial DNA. *Cytogenet. Cell. Genet.*, **58**, 1103-1123, 1991.

Wallace D.C., Singh G., Lott M.T., Hodge J.A., Schurr T.G., Lezza A.M.S., Elsas L.J. and Nikoskelainen E.K. Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. *Science*, **242**, 1427-1430, 1988a.

Wallace D.C., Zheng X., Lott M.T., Shoffner J.M., Hodge J.A., Kelley R.I., Epstein C.M. and Hopkins L.C. Familial mitochondrial encephalomyopathy (MERRF): Genetic, pathophysiological and biochemical characterisation of a mitochondrial DNA disease. *Cell*, **56**, 601-610, 1988b.

Wardell T.M., Ferguson E., Chinnery P.F., Borthwick G.M., Taylor R.W., Jackson G., Craft A., Lightowlers R.N., Howell N. and Turnbull D.M. Changes in the human mitochondrial genome after treatment of malignant disease. *Mutat. Res.*, **525**, 19-27, 2003.

Wikström M. Oxidative phosphorylation: An overview. In: Genetics of Mitochondrial Diseases. Holt I.J., editor. Oxford: Oxford University Press, 69-84, 2003.

Willems J.L., Monnens L.A.H., Trijbels J.M.F., Veerkamp J.H., Meyer A.E.F.H., Van Dam K. and Van Haelst U. Leigh's encephalomyelopathy in a patient with cytochrome *c* oxidase deficiency in muscle tissue. *Pediatrics*, **60**, 850-857, 1977.

Williams R.S. Another surprise from the mitochondrial genome. *N. Engl. J. Med.*, **347**, 609-612, 2002.

Winter P.C., Hickey G.I. and Fletcher H.L. Instant Notes in Genetics. 2nd ed. Hames B.D., editor. Oxford: BIOS Scientific Publishers Limited, 88-89, 293, 2002.

Wolf N.I. and Smeitink J.A.M. Mitochondrial disorders. A proposal for consensus diagnostic criteria in infants and children. *Neurology*, **59**, 1402-1405, 2002.

Wrischnik L.S., Higuichi R.G., Stoneking M., Erlich H.A., Arnheim N. and Wilson A.C. Length mutations in human mitochondrial DNA: direct sequencing of enzymatically amplified DNA. *Nucleic Acids Res.*, **15**, 529-542, 1987.

Yasukawa T., Suzuki T., Ohta S. and Watanabe K. Wobble modification defect suppresses translational activity of tRNAs with MERRF and MELAS mutations. *Mitochondrion*, **2**, 129-141, 2002.

Yoneda M., Chomyn A., Martinuzzi A., Hurko O. and Attardi G. Marked replicative advantage of human mtDNA carrying a point mutation that causes the MELAS encephalomyopathy. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 11164-11168, 1992.

Zeviani M., Amati P., Comi G., Fratta G., Mariotti C. and Tiranti V. Searching for genes affecting the structural integrity of the mitochondrial genome. *Biochim. Biophys. Acta*, **1271**, 153-158, 1995.

Zeviani M., Bresolin N., Gellera C., Bordoni A., Pannacci M., Amati P., Moggio M., Servidei S., Scarlato G. and DiDonato S. Nucleus-driven multiple large-scale deletions of the human mitochondrial genome: a new autosomal dominant disease. *Am. J. Hum. Genet.*, **47**, 1904-1914, 1990.

Zeviani M., Gellera C., Antozzi C., Rimoldi M., Morandi L., Villani F., Tiranti V. and DiDonato S. Maternally inherited myopathy and cardiomyopathy: association with mutation in mitochondrial DNA $tRNA^{Leu(UUR)}$. *Lancet*, **338**, 143-147, 1991.

Zeviani M., Moraes C.T., DiMauro S., Nakase H., Bonilla E., Schon E.A. and Rowland L.P. Deletions of mitochondrial DNA in Kearns-Sayre syndrome. *Neurology*, **38**, 1339-1346, 1988.

Zeviani M., Servidei S., Gellera C., Bertini E., DiMauro S., and DiDonato S. An autosomal dominant disorder with multiple deletions of mitochondrial DNA starting at the D-loop region. *Nature*, **339**, 309-311, 1989.

6.2 ELECTRONIC REFERENCES

BIOEDIT Sequence Alignment Editor, http://www.mbio.ncsu.edu/BioEdit/bioedit.html, 2003.
CLUSTALW, Multiple Sequence Alignments, http://www.cmbi.kun.nl./cgi-bin/clustalw.pl, 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.
MITOANALYZER, National Institute of Standards and Technology, Gaithersburg, MD, U.S.A. http://www.cstl.nist.gov/biotech/strbase/mitoanalyzer.html, 2003.
MITOMAP, A Human Mitochondrial Genome Database, http://www.mitomap.org, 2003.
NCBI, National Center for Biotechnology Information, National Institutes of Health, http://www.ncbi.nlm.nih.gov., 2003.
NET PRIMER, PREMIER Biosoft International, http://www.PremierBiosoft.com, 2003.
NEUROLOGY, http://neurology.org/cgt/content/full/59/9/1402/DCI, 2003.
OLIGONUCLEOTIDE PROPERTIES CALCULATOR, http://www.basic.nwu.edu/biotools/oligocalc.html, 2003.
PubMed Central, U.S. National Library of Medicine's digital archive of life sciences journal literature. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Pmc, 2004.

APPENDIX A

WHOLE mtDNA GENOME SEQUENCING ALTERATIONS

In Tables A.1 to A.4 all the alterations detected for patients 386, 504, 525 and 1301 are presented. These patients were screened for any deviations from the RCRS. The potentially pathogenic mutations, discussed in paragraph 4.4., are depicted in red, whereas the alterations utilised for haplogroup analyses are in blue. The alterations defined as "polymorphisms" were reported as such, although they do not necessarily occur at a frequency range of 0.01 to 0.99 (Falconer, 1981). However, they are neutral and do not result in a disease phenotype (Winter *et al.*, 2002).

Thirteen alterations were detected in the mtDNA from whole blood of patient 386 and are depicted in Table A.1. Heteroplasmy could not be detected for any of these alterations.

Table A.1: Alterations detected in patient 386

Alteration	Locus	Alteration Type	Reference
A263G	Control region	Polymorphism	Aquadro and Greenberg (1983)
315insC	Control region	Polymorphism	Andrews *et al.* (1999)
A750G	12S rRNA	Polymorphism	Ozawa *et al.* (1991a)
A1438G	12S rRNA	Polymorphism	Ozawa *et al.* (1991a)
G3010A	16S rRNA	Polymorphism	Ozawa *et al.* (1991a)
A4769G	ND2	Polymorphism	Ozawa *et al.* (1991a)
G7337A	COI	Polymorphism	Nishino *et al.* (1996)
G7979A	COII	Non-synonymous (Asp → Asn)	Not reported
A8860G	ATPase 6	Polymorphism	Silva *et al.* (2003)
A15326G	Cyt *b*	Polymorphism	Ozawa *et al.* (1991a)
A15758G	Cyt *b*	Polymorphism	Andreu *et al.* (1999)
G16319A	Control region	Polymorphism	Ozawa *et al.* (1991a)
T16519C	Control region	Polymorphism	Aquadro and Greenberg (1983)

A reference is provided for reported alterations (MITOMAP, 2003). Novel alterations were classified as synonymous or non-synonymous by MITOANALYSER (2003). Blue = alterations utilised for haplogroup analysis, red = potential pathogenic mutations, rRNA = ribosomal RNA, ND2 = gene encoding subunit 2 of NADH dehydrogenase, CO I and CO II = genes encoding subunits I and II of cytochrome *c* oxidase, ATPase 6 = gene encoding subunit 6 of ATP synthase, Cyt *b* = gene encoding cytochrome *b*, Asp = aspartic acid and Asn = asparagine.

Eighty-eight alterations were detected in mtDNA from whole blood of patient 504 and are depicted in Table A.2. Heteroplasmy could not be detected for any of the alterations.

Table A.2: Alterations detected in patient 504

Alteration	Locus	Alteration Type	Reference
A93G	Control region	Polymorphism	Jorde *et al.* (1995)
A95C	Control region	Polymorphism	Jorde *et al.* (1995)
T236C	Control region	Polymorphism	Aquadro and Greenberg (1983)
G247A	Control region	Polymorphism	Aquadro and Greenberg (1983)
A263G	Control region	Polymorphism	Aquadro and Greenberg (1983)
309insCC	Control region	Polymorphism	Torroni *et al.* (1998)
315insC	Control region	Polymorphism	Andrews *et al.* (1999)
522-523delCA	Control region	Polymorphism	Kleinle *et al.* (1998)
A750G	12S rRNA	Polymorphism	Ozawa *et al.* (1991a)
G769A	12S rRNA	Polymorphism	Moraes *et al.* (1993)
T825A	12S rRNA	Polymorphism	Prezant *et al.* (1993)
956-965insCCCCC	12S rRNA	Functional effect unknown	Shoffner *et al.* (1993)
G1018A	12S rRNA	Polymorphism	Moraes *et al.* (1993)
C1048T	12S rRNA	Polymorphism	Moraes *et al.* (1993)
A1438G	12S rRNA	Polymorphism	Ozawa *et al.* (1991a)
A2245G	16S rRNA	Polymorphism	Moraes *et al.* (1993)
A2706G	16S rRNA	Polymorphism	Ozawa *et al.* (1991a)
G2758A	16S rRNA	Polymorphism	Moraes *et al.* (1993)
T2885C	16S rRNA	Polymorphism	Moraes *et al.* (1993)
C3516A	ND1	Synonymous	Not reported
C3594T	ND1	Synonymous	Not reported
A4104G	ND1	Synonymous	Not reported
C4312T	$tRNA^{Ile}$	Polymorphism	Moraes *et al.* (1993)
T4586C	ND2	Synonymous	Not reported
A4769G	ND2	Polymorphism	Ozawa *et al.* (1991a)
T5096C	ND2	Synonymous	Not reported
G5231A	ND2	Polymorphism	Ozawa *et al.* (1991b)
T5442C	ND2	Polymorphism	Ozawa *et al.* (1991a)
G5460A	ND2	Polymorphism	Kösel *et al.* (1994)
C5603T	$tRNA^{Ala}$	Polymorphism	Moraes *et al.* (1993)
C5911T	COI	Non-synonymous (Ala → Val)	Not reported
T6185C	COI	Synonymous	Not reported
C7028T	COI	Polymorphism	Nishino *et al.* (1996)
A7146G	COI	Non-synonymous (Thr → Ala)	Not reported
C7256T	COI	Polymorphism	Silva *et al.* (2003)
G7521A	$tRNA^{Asp}$	Polymorphism	Prezant *et al.* (1993)
A8191G	COII	Synonymous	Not reported
T8404C	ATPase 8	Synonymous	Not reported

Table A.2: continued...

Alteration	Locus	Alteration Type	Reference
C8428T	ATPase 8	Polymorphism	Moraes *et al.* (1993)
C8468T	ATPase 8	Polymorphism	Moraes *et al.* (1993)
A8566G	ATPase 6 ATPase 8	Non-synonymous (Ile → Val) Synonymous	Not reported
C8655T	ATPase 6	Synonymous	Not reported
A8701G	ATPase 6	Polymorphism	Ozawa *et al.* (1991b)
A8860G	ATPase 6	Polymorphism	Silva *et al.* (2003)
C9042T	ATPase 6	Synonymous	Not reported
A9347G	COIII	Synonymous	Not reported
T9540C	COIII	Polymorphism	Ozawa *et al.* (1991b)
G9755A	COIII	Synonymous	Not reported
C9818T	COIII	Synonymous	Not reported
A10398G	ND3	Polymorphism	Obayashi *et al.* (1992)
G10589A	ND4L	Polymorphism	Howell *et al.* (1995)
C10664T	ND4L	Synonymous	Not reported
G10688A	ND4L	Synonymous	Not reported
T10810C	ND4	Synonymous	Not reported
T10873C	ND4	Polymorphism	Ozawa *et al.* (1991b)
T10915C	ND4	Polymorphism	Ozawa *et al.* (1991b)
G11176A	ND4	Synonymous	Not reported
A11641G	ND4	Synonymous	Not reported
G11719A	ND4	Polymorphism	Ozawa *et al.* (1991a)
G11914A	ND4	Polymorphism	Sudoyo *et al.* (2002)
G12007A	ND4	Polymorphism	Moraes *et al.* (1993)
G12127A	ND4	Synonymous	Not reported
C12705T	ND5	Polymorphism	Ozawa *et al.* (1991b)
A12720G	ND5	Polymorphism	Brown *et al.* (2001)
A13105G	ND5	Polymorphism	Prezant *et al.* (1993)
A13276G	ND5	Non-synonymous (Met → Val)	Not reported
C13506T	ND5	Synonymous	Not reported
C13650T	ND5	Polymorphism	Silva *et al.* (2003)
A14007G	ND5	Synonymous	Not reported
T14308C	ND6	Polymorphism	Sudoyo *et al.* (2002)
C14766T	Cyt *b*	Polymorphism	Moraes *et al.* (1993)
C15136T	Cyt *b*	Synonymous	Not reported
A15326G	Cyt *b*	Polymorphism	Ozawa *et al.* (1991a)
G15431A	Cyt *b*	Non-synonymous (Ala → Thr)	Not reported
G16129A	Control region	Polymorphism	Ozawa *et al.* (1991a)
C16148T	Control region	Polymorphism	Aquadro and Greenberg (1983)

Table A.2: continued...

Alteration	Locus	Alteration Type	Reference
C16168T	Control region	Polymorphism	DiRienzo and Wilson (1991)
T16172C	Control region	Polymorphism	Ozawa *et al.* (1991b)
C16187T	Control region	Polymorphism	Aquadro and Greenberg (1983)
C16188G	Control region	Polymorphism	Aquadro and Greenberg (1983)
T16189C	Control region	Polymorphism	Aquadro and Greenberg (1983)
C16223T	Control region	Polymorphism	Nishino *et al.* (1996)
A16230G	Control region	Polymorphism	Aquadro and Greenberg (1983)
C16278T	Control region	Polymorphism	Ozawa *et al.* (1991b)
A16293G	Control region	Polymorphism	Aquadro and Greenberg (1983)
T16311C	Control region	Polymorphism	Bendall *et al.* (1996)
C16320T	Control region	Polymorphism	Aquadro and Greenberg (1983)

A reference is provided for reported alterations (MITOMAP, 2003). Novel alterations were classified as synonymous or non-synonymous by MITOANALYSER (2003). Blue = alterations utilised for haplogroup analysis, red = potential pathogenic mutations, rRNA = ribosomal RNA, ND1, ND2, ND3, ND4, ND4L, ND5 and ND6 = genes encoding subunits 1, 2, 3, 4, 4L, 5 and 6 of NADH dehydrogenase, CO I, CO II, CO III = genes encoding subunits I, II and III of cytochrome c oxidase, ATPase 6 and 8 = genes encoding subunits 6 and 8 of ATP synthase, Cyt b = gene encoding cytochrome b, $tRNA^{Ile}$ = gene encoding tRNA isoleucine, $tRNA^{Ala}$ = gene encoding tRNA alanine, $tRNA^{Asp}$ = gene encoding tRNA aspartic acid, Ala = alanine, Val = valine, Ile = isoleucine, Met = methionine and Thr = threonine.

The 91 alterations detected in the mtDNA from whole blood of patient 525 are depicted in Table A.3. Heteroplasmy could not be detected for any of these alterations.

Table A.3: Alterations detected in patient 525

Alteration	Locus	Alteration Type	Reference
A73G	Control region	Polymorphism	Aquadro and Greenberg (1983)
T146C	Control region	Polymorphism	Aquadro and Greenberg (1983)
T152C	Control region	Polymorphism	Aquadro and Greenberg (1983)
T195C	Control region	Polymorphism	Aquadro and Greenberg (1983)
C198T	Control region	Polymorphism	Jorde *et al.* (1995)
G247A	Control region	Polymorphism	Aquadro and Greenberg (1983)
315insC	Control region	Polymorphism	Andrews *et al.* (1999)
498delC	Control region	Functional effect unknown	Not reported
522-523delCA	Control region	Polymorphism	Kleinle *et al.* (1998)
C597T	$tRNA^{Phe}$	Functional effect unknown	Not reported
A750G	12S rRNA	Polymorphism	Ozawa *et al.* (1991a)
G769A	12S rRNA	Polymorphism	Moraes *et al.* (1993)
T825A	12S rRNA	Polymorphism	Prezant *et al.* (1993)
G1018A	12S rRNA	Polymorphism	Moraes *et al.* (1993)
C1048T	12S rRNA	Polymorphism	Moraes *et al.* (1993)
A2706G	16S rRNA	Polymorphism	Ozawa *et al.* (1991a)
G2758A	16S rRNA	Polymorphism	Moraes *et al.* (1993)

Table A.3: continued...

Alteration	Locus	Alteration Type	Reference
T2885C	16S rRNA	Polymorphism	Moraes *et al.* (1993)
C3516A	ND1	Synonymous	Not reported
C3594T	ND1	Synonymous	Not reported
A3756G	ND1	Synonymous	Not reported
A3981G	ND1	Synonymous	Not reported
C4025T	ND1	Polymorphism	Huoponen *et al.* (1993)
A4044G	ND1	Synonymous	Not reported
A4104G	ND1	Synonymous	Not reported
A4225G	ND1	Non-synonymous (Met → Val)	Not reported
C4312T	$tRNA^{Ile}$	Polymorphism	Moraes *et al.* (1993)
A4769G	ND2	Polymorphism	Ozawa *et al.* (1991a)
A5153G	ND2	Synonymous	Not reported
T5442C	ND2	Polymorphism	Ozawa *et al.* (1991a)
G5460A	ND2	Polymorphism	Kösel *et al.* (1994)
5899insC	Non-coding	Functional effect unknown	Not reported
T6185C	COI	Synonymous	Not reported
T6815C	COI	Polymorphism	Not reported
C7028T	COI	Polymorphism	Nishino *et al.* (1996)
A7146G	COI	Non-synonymous (Thr → Ala)	Not reported
A7154G	COI	Synonymous	Not reported
C7256T	COI	Polymorphism	Silva *et al.* (2003)
G7521A	$tRNA^{Asp}$	Polymorphism	Prezant *et al.* (1993)
C8113A	COII	Polymorphism	Ruvolo *et al.* (1993)
G8152A	COII	Polymorphism	Ruvolo *et al.* (1993)
G8251A	COII	Polymorphism	Ruvolo *et al.* (1993)
G8392A	ATPase 8	Polymorphism	Vittecoq *et al.* (2002)
C8468T	ATPase 8	Polymorphism	Moraes *et al.* (1993)
C8655T	ATPase 6	Synonymous	Not reported
A8701G	ATPase 6	Polymorphism	Ozawa *et al.* (1991b)
A8860G	ATPase 6	Polymorphism	Silva *et al.* (2003)
C9042T	ATPase 6	Synonymous	Not reported
A9347G	COIII	Synonymous	Not reported
T9540C	COIII	Polymorphism	Ozawa *et al.* (1991b)
G9755A	COIII	Synonymous	Not reported
A10398G	ND3	Polymorphism	Obayashi *et al.* (1992)
G10589A	ND4	Polymorphism	Howell *et al.* (1995)
C10664T	ND4L	Synonymous	Not reported
G10688A	ND4L	Synonymous	Not reported
T10810C	ND4	Synonymous	Not reported

Table A.3: continued...

Alteration	Locus	Alteration Type	Reference
T10873C	ND4	Polymorphism	Ozawa *et al.* (1991b)
T10915C	ND4	Polymorphism	Ozawa *et al.* (1991b)
G11719A	ND4	Polymorphism	Ozawa *et al.* (1991a)
T11854C	ND4	Synonymous	Not reported
G11914A	ND4	Polymorphism	Sudoyo *et al.* (2002)
G12007A	ND4	Polymorphism	Moraes *et al.* (1993)
T12121C	ND4	Synonymous	Not reported
A12172G	$tRNA^{His}$	Functional effect unknown	Not reported
A12234G	$tRNA^{Ser(AGY)}$	Polymorphism	Sternberg *et al.* (1998)
C12705T	ND5	Polymorphism	Ozawa *et al.* (1991b)
A12720G	ND5	Polymorphism	Brown *et al.* (2001)
A12810G	ND5	Polymorphism	Ozawa *et al.* (1991b)
A13105G	ND5	Polymorphism	Prezant *et al.* (1993)
A13276G	ND5	Non-synonymous (Met → Val)	Not reported
C13506T	ND5	Synonymous	Not reported
C13650T	ND5	Polymorphism	Silva *et al.* (2003)
C13767T	ND5	Synonymous	Not reported
T14221C	ND6	Synonymous	Not reported
T14502C	ND6	Polymorphism	Ozawa *et al.* (1991a)
C14766T	Cyt *b*	Polymorphism	Moraes *et al.* (1993)
A15326G	Cyt *b*	Polymorphism	Ozawa *et al.* (1991a)
G15466A	Cyt *b*	Synonymous	Not reported
A15766G	Cyt *b*	Polymorphism	Andreu *et al.* (1999)
G15930A	$tRNA^{Thr}$	Functional effect unknown	Not reported
T15941C	$tRNA^{Thr}$	Functional effect unknown	Not reported
G16129A	Control region	Polymorphism	Ozawa *et al.* (1991a)
C16187T	Control region	Polymorphism	Aquadro and Greenberg (1983)
T16189C	Control region	Polymorphism	Aquadro and Greenberg (1983)
A16212G	Control region	Polymorphism	Horai and Hayasaka (1990)
C16223T	Control region	Polymorphism	Nishino *et al.* (1996)
A16230G	Control region	Polymorphism	Aquadro and Greenberg (1983)
T16243C	Control region	Polymorphism	Aquadro and Greenberg (1983)
T16311C	Control region	Polymorphism	Bendall *et al.* (1996)
G16390A	Control region	Polymorphism	Chen *et al.* (1995)
T16519C	Control region	Polymorphism	Aquadro and Greenberg (1983)

A reference is provided for reported alterations (MITOMAP, 2003). Novel alterations were classified as synonymous or non-synonymous by MITOANALYSER (2003). Blue = alterations utilised for haplogroup analysis, red = potential pathogenic mutations, rRNA = ribosomal RNA, ND1, ND3, ND4, ND4L and ND5 = genes encoding subunits 1, 2, 3, 4, 4L, 5 and 6 of NADH dehydrogenase, CO I, CO II, CO III = genes encoding subunits I, II and III of cytochrome *c* oxidase, ATPase 6 and 8 = genes encoding subunits 6 and 8 of ATP synthase, Cyt *b* = gene encoding cytochrome *b*, $tRNA^{Phe}$ = gene encoding tRNA phenylalanine, $tRNA^{Ile}$ = gene encoding tRNA isoleucine, $tRNA^{Asp}$ = gene encoding tRNA aspartic acid, $tRNA^{Ser(AGY)}$ = gene encoding tRNA serine (specifically recognising codon AGY), $tRNA^{Thr}$ = gene encoding tRNA threonine, Met = methionine, Val = valine, Thr = threonine and Ala = alanine.

Fifty-five alterations were detected in mtDNA from muscle of patient 1301, as presented in Table A.4. Heteroplasmy was not detected for any of the alterations.

Table A.4: Alterations detected in patient 1301

Alteration	Locus	Alteration Type	Reference
A73G	Control region	Polymorphism	Aquadro and Greenberg (1983)
T146C	Control region	Polymorphism	Aquadro and Greenberg (1983))
T152C	Control region	Polymorphism	Aquadro and Greenberg (1983)
T195C	Control region	Polymorphism	Aquadro and Greenberg (1983)
A263G	Control region	Polymorphism	Aquadro and Greenberg (1983)
309insC	Control region	Polymorphism	Torroni *et al.* (1998)
315insC	Control region	Polymorphism	Andrews *et al.* (1999)
A750G	12S rRNA	Polymorphism	Ozawa *et al.* (1991a)
G769A	12S rRNA	Polymorphism	Moraes *et al.* (1993)
G1018A	12S rRNA	Polymorphism	Moraes *et al.* (1993)
A1438G	12S rRNA	Polymorphism	Ozawa *et al.* (1991a)
T2416C	16S rRNA	Functional effect unknown	Not reported
A2706G	16S rRNA	Polymorphism	Ozawa *et al.* (1991a)
C2789T	16S rRNA	Polymorphism	Rose *et al.* (2001)
C3594T	ND1	Synonymous	Not reported
G3918A	ND1	Synonymous	Not reported
A4104G	ND1	Synonymous	Not reported
A4769G	ND2	Polymorphism	Ozawa *et al.* (1991a)
A5285G	ND2	Synonymous	Not reported
C7028T	COI	Polymorphism	Nishino *et al.* (1996)
T7175C	COI	Polymorphism	Silva *et al.* (2003)
C7256T	COI	Polymorphism	Silva *et al.* (2003)
C7274T	COI	Polymorphism	Silva *et al.* (2003)
G7521A	$tRNA^{Asp}$	Polymorphism	Prezant *et al.* (1993)
A7771G	COII	Polymorphism	Silva *et al.* (2003)
G8206A	COII	Polymorphism	Ozawa *et al.* (1995)
A8701G	ATPase6	Polymorphism	Ozawa *et al.* (1991b)
A8860G	ATPase6	Polymorphism	Silva *et al.* (2003)
A9221G	COIII	Polymorphism	Silva *et al.* (2003)
T9540C	COIII	Polymorphism	Ozawa *et al.* (1991b)
T10115C	ND3	Polymorphism	Silva *et al.* (2003)
A10398G	ND3	Polymorphism	Obayashi *et al.* (1992)
T10873C	ND4	Polymorphism	Ozawa *et al.* (1991b)
G11719A	ND4	Polymorphism	Ozawa *et al.* (1991a)
G11914A	ND4	Polymorphism	Sudoyo *et al.* (2002)

Table A.4: continued...

Alteration	Locus	Alteration Type	Reference
T11944C	ND4	Polymorphism	Ozawa *et al.* (1991b)
A12693G	ND5	Polymorphism	Silva *et al.* (2003)
C12705T	ND5	Polymorphism	Ozawa *et al.* (1991b)
G13590A	ND5	Polymorphism	Ozawa *et al.* (1995)
C13650T	ND5	Polymorphism	Silva *et al.* (2003)
A13803G	ND5	Polymorphism	Silva *et al.* (2003)
A14566G	ND6	Polymorphism	Chinnery *et al.* (2001)
C14766T	Cyt *b*	Polymorphism	Moraes *et al.* (1993)
A15244G	Cyt *b*	Polymorphism	Andreu *et al.* (1999)
G15301A	Cyt *b*	Polymorphism	Ozawa *et al.* (1991b)
A15326G	Cyt *b*	Polymorphism	Ozawa *et al.*(1991a)
T15629C	Cyt *b*	Synonymous	Not reported
T15784C	Cyt *b*	Polymorphism	Silva *et al.* (2003)
C16223T	Control region	Polymorphism	Nishino *et al.* (1996)
C16278T	Control region	Polymorphism	Ozawa *et al.* (1991b)
C16286T	Control region	Polymorphism	Horai and Hayasaka (1990)
C16294T	Control region	Polymorphism	Aquadro and Greenberg (1983)
A16309G	Control region	Polymorphism	Horai and Hayasaka (1990)
G16390A	Control region	Polymorphism	Chen *et al.* (1995)
T16519C	Control region	Polymorphism	Aquadro and Greenberg (1983)

A reference is provided for reported alterations (MITOMAP, 2003). Novel alterations were classified as synonymous or non-synonymous by MITOANALYSER (2003). Blue = alterations utilised for haplogroup analysis, red = potential pathogenic mutations, rRNA = ribosomal RNA, ND1, ND3, ND4, ND4L, ND5 and ND6 = genes encoding subunits 1, 2, 3, 4, 4L, 5 and 6 of NADH dehydrogenase, CO I, CO II, CO III = genes encoding subunits I, II and III of cytochrome *c* oxidase, ATPase 6 and 8 = genes encoding subunits 6 and 8 of ATP synthase, Cyt *b* = gene encoding cytochrome *b* and $tRNA^{Asp}$ = gene encoding tRNA aspartic acid.

APPENDIX B

HAPLOGROUP ANALYSIS

Haplogroup analysis could be performed for patients 386, 504, 525 and 1301, as their whole mtDNA genome sequences were available, including the haplogroup determining nucleotides. Table B.1 below was applied for the analysis, starting at number 1 in the bottom right corner and working upwards through the table until number 30 is reached. For an individual to be in haplogroups L0 or L1, all 30 polymorphisms listed in Table B.1 should be present. However, Table B.1 cannot distinguish between haplogroups L0 and L1. Unpublished data provided by Ruiz-Pesini (2003) was applied for this differentiation.

Table B.1: Exclusion criteria for haplogroup L from other haplogroups

Exclusion Criteria	Polymorphism	#
L0 and L1 from other haplogroups	T825A	30
	G2758A	29
	T2885C	28
	A7146G	27
	C8468T	26
	C8655T	25
	G10688A	24
	T10810C	23
	A13105G	22
	C13506T	21
L0, L1 and L2 from other haplogroups	G769A	20
	G1018A	19
	C3594T	18
	A4104G	17
	C7256T	16
	G7521A	15
	C13650T	14
Haplogroup N	A8701G	13
	T9540C	12
	T10873C	11
Haplogroup R	C12705T	10
Haplogroup HV	G11719A	9
	C14766T	8

Table B.1 continued...

Haplogroup H	A2706G	7
	C7028T	6
Sub-haplogroup of H containing the CRS	A750G	5
	A1438G	4
	A4769G	3
Sub-haplogroup of H containing the RCRS	A8860G	2
	A15326G	1

= number. Obtained from Ruiz-Pesini (2003).

APPENDIX C

CONFERENCES AT WHICH RESEARCH CONTAINED IN THIS DISSERTATION WERE PRESENTED

Parts of the research contained in this dissertation were presented at the following international conference. The name of the presenting author is underlined.

C.1 RESEARCH PRESENTED AT INTERNATIONAL CONFERENCES

53rd Annual meeting of the American Society of Human Genetics: Los Angeles, U.S.A., November 2003.

Van Brummelen A.C., Smuts I., Wallace D.C. and Olckers A. Molecular screening of patients with mitochondrial disorders in the South African population (poster presentation).

Smuts I., van Brummelen A.C., Wallace D.C. and Olckers A. Evaluation of a mitochondrial scoring system in the South African population (poster presentation).