Molecular screening of the G3277A alteration in a Black South African diabetic population

BY

MADELEINE NICOLETTE WESSELS, B.Sc.(Hons)

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 Annamarie Kruger
Research co-ordinator, Preventative and Therapeutic Intervention,
Faculty of Health Sciences,
Potchefstroom University for Christian Higher Education

December 2003

Molekulêre sifting van die G3277A verandering in 'n Swart Suid-Afrikaanse diabetiese populasie

DEUR

MADELEINE NICOLETTE WESSELS, B.Sc.(Hons)

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: Doktor Annamarie Kruger
Navorsings-koördineerder, Terapeutiese en Voorkomende Intervensie,
Fakulteit Gesondheidswetenskappe,
Potchefstroomse Universiteit vir Christelike Hoër Onderwys

Desember 2003

TO RORY

ABSTRACT

The diabetic disorder is a collection of genetic diseases with a common phenotype of impaired glucose homeostasis and affects many different organs, resulting in a clinically heterogeneous phenotype with strong genetic as well as environmental influences. The basis of diabetes is the aberrant production and utilisation of glucose, in conjunction with factors affecting the influence of insulin on the body and its role in glucose absorption. The mitochondrion has a vital role to play in energy production, and therefore warrants investigation.

A mutation at position 3243 in the mitochondrial genome has been associated with the diabetic phenotype when expressed in heteroplasmic levels of approximately 30%. This mutation has been detected in the European diabetic population but is not detected in the Black South African diabetic population. In a previous study to determine the prevalence of the 3243 mitochondrial mutation in a Black South African diabetic cohort, another alteration, G3277A, was observed in ca. 3% of a small cohort investigated. In the present study it was investigated whether this alteration could be associated with the Black South African diabetic population.

Blood samples collected from 222 diabetic and 237 non-diabetic individuals were analysed via RFLP and automated sequencing techniques for the presence of this G3277A alteration as well as the A3243G mutation in the samples that were sequenced. It was observed that seven control individuals and one patient individual harboured the G3277A alteration. None harboured the A3243G mutation but three novel alterations were detected. Via statistical analysis it was determined that the G3277A alteration was not in Hardy-Weinberg equilibrium in either the diabetic or non-diabetic populations, therefore no assumption regarding the association of these alterations with the type 2 diabetic phenotype could be made. It is therefore suggested that in the future larger groups of individuals are screened for the A3277G alteration.

OPSOMMING

Diabetes bestaan uit 'n versameling genetiese afwykings met 'n gemeenskaplike fenotipe, naamlik verminderde glukose-homeostase, en beïnvloed verskillende organe, wat lei tot 'n klinies heterogene fenotipe met sterk genetiese en omgewingsinvloede. Die grondslag van die afwyking is die foutiewe produksie en gebruik van glukose, in samewerking met faktore wat beide die invloed van insulien op die liggaam en die rol van insulien in glucose-opname beïnvloed. Die mitochondrion vervul 'n belangrike rol in energieproduksie, en moet dus ook in ag geneem word.

'n Mutasie by posisie A3243G in die mitochondriale genoom word geassosieer met diabetes wanneer dit uitgedruk word in heteroplasmiese vlakke van ongeveer 30%. Hierdie mutasie word waargeneem in Europese bevolkings, maar is nog nie waargeneem in die Swart Suid-Afrikaanse diabetiese bevolking nie.

In 'n vorige studie waarin die frekwensie van die 3243 mitochondriale mutasie in 'n Swart Suid-Afrikaanse diabetiese groep bepaal is, is 'n ander verandering, G3277A, in ongeveer 3% van die groep wat ondersoek is waargeneem. In die huidige studie is ondersoek ingestel of hierdie verandering geassosieer kan word met die Swart Suid-Afrikaanse diabetiese bevolking.

Bloedmonsters is versamel van 222 diabetiese en 237 nie-diabetiese individue en geanaliseer vir die teenwoordigheid van die G3277A verandering deur middel van RFLP en volgordebepalendegnieke. Monsters waarvan die DNS-volgordes bepaal is, is ook vir die teenwoordigheid van die A3243G-mutasie geanaliseer. Sewe van die kontrole-individue en een pasïent het die G3277A verandering getoon. Geen individu het die A3243G mutasie getoon nie, maar drie nuwe veranderinge is waargeneem. Deur middel van statistiese analise is die afwesigheid van die Hardy-Weinberg ewewig van die G3277A-verandering in beide die diabetiese en nie-diabetiese groepe opgemerk. Daarom kan geen afleidings aangaande die assosiasie van hierdie veranderinge met die tipe 2 diabetiese fenotipe gemaak word nie. Daar word dus annbeveel dat groter groepe van individue in die toekoms gesif word vir die A3277G verandering.

TABLE OF CONTENTS

CHAPTER FIVE

REFERENCES

LIST OF ABBREVIATIONS AND SYMBOLS

LIST OF SYMBOLS

α	alpha
β	beta
δ	delta
γ	gamma
μ	micro
μg	microgram
μl	microlitre
μM	micromolar
%	percent
π	sample proportion
ψ	pseudouridine
χ^2	chi-square

LIST OF ABBREVIATIONS

12S rRNA	12 svedberg units ribosomal RNA
16S rRNA	16 svedberg units ribosomal RNA
A and a	adenine (in DNA sequence)
A_{260}	absorbency at 260 nm
A_{280}	absorbency at 280 nm
A_{260}/A_{280}	ratio of absorbency measured at 260 nm and 280 nm
ADP	adenosine diphosphate
ADP:O	ratio of adenosine diphosphate to oxygen
Ala	alanine
APM1	adiponectin
APS	ammonium persulphate
Arg	arginine
Asn	asparagine
Asp	aspartic acid
ATP	adenosine triphosphate
ATPase	adenosine triphosphatase
bp	base pair
BHE/Cdb	a laboratory bred strain of rat used in research
BMI	body mass index
BSA	bovine serum albumin
C and c	cytosine (in DNA sequence)
C	control
Ca^{2+}	calcium ion
CAD	coronary artery disease
*CAPN*10	calpain 10
cm	centimetre
CO I	cytochrome *c* oxidase subunit I
CO II	cytochrome *c* oxidase subunit II
CO III	cytochrome c oxidase subunit III
CO_2	carbon dioxide
CoA	coenzyme A
CoQ10	coenzyme Q10
COOH	carboxyl group
COX II	subunit II gene of the cytochrome *c* oxidase complex
CPEO	chronic progressive external ophthalmoplegia
CR	control region
CRS	Cambridge reference sequence
Cys	cysteine
Cyt	cytochrome
°C	degrees centigrade

DIAGEN	Genetics of diabetes study
DdH_2O	double distilled water
dATP	2'-deoxyadenosine-5'-triphosphate
dCTP	2'-deoxycytidine-5'-triphosphate
ddNTP	dideoxynucleotide triphosphate
dGTP	2'-deoxyguanosine-5'-triphosphate
D-loop	displacement loop
DM	diabetes mellitus
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
dTTP	2'-deoxythymidine-5'-triphosphate
E	expected
e.g.	for example
e^-	electron
et al	*et altera:* and others
EDTA	ethylenediamine tetra-acetic acid
F	female
EtBr	ethidium bromide
F_o	transmembrane proton channel of complex V
F_1	hydrophobic component of complex V
FADH	flavin adenine dinucleotide
FFA	free fatty acids
G	gram
G	glucose
G and g	guanine (in DNA sequence)
gDNA	genomic DNA
GK	glucokinase
Gln	glutamine
Glu	glutamic acid
Gly	glycine
Glut	glucose transporter
H^+	hydrogen ion
H_2O	water
H_A	alternative hypothesis
HCl	hydrogen chloride
His	histidine
HLA	human leukocyte antigen
HNF	human necrosis factor
H_o	null hypothesis
HSP	heavy strand promoter
HVR	hypervariable region
H-W	Hardy-Weinberg
i	insulin
IDDM	insulin dependent diabetes mellitus
i.e.	that is
IGT	impaired glucose tolerance
Ile	isoleucine
IPF	insulin promoter factor
IR	insulin receptor
IRS	insulin receptor substrate
K^+	potassium
K_{ATP}	ATP-sensitive potassium channels
kB	kilobase
KCl	potassium chloride
kg	kilogram
KOAc	potassium acetate
KOH	potassium hydroxide
l	litre
Leu	leucine
LSP	light strand promoter
Lys	lysine
M	molar
M	male
mg	milligram

MAP	mitogen-activated protein
MBS	multiblock system
Met	methionine
MELAS	mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes
MERRF	myoclonic-epilepsy and ragged red fibres
Mg^{2+}	magnesium ion
MgCl	magnesium chloride
Mg(OAc)	magnesium acetate
$mg.ml^{-1}$	milligram per millilitre
MHC	major histocompatibility complex
MIDD	maternally inherited diabetes mellitus and deafness
min	minutes
ml	millilitre
$mmol.l^{-1}$	millimoles per litre
mM	millimolar
mRNA	messenger RNA
MODY	maturity onset diabetes of the young
mt	mitochondrial
mtDNA	mitochondrial DNA
MTERF	mitochondrial transcription termination factor
N	nano
Na^+	sodium ion
NADH	nicotinamide adenine dinucleotide (reduced form)
Na_2EDTA	ethylene diamine tetra-acetic acid di-sodium salt
NaOAc	sodium acetate
ND	NADH dehydrogenase
NeuroD	neuronal transcription factor
ng	nanogram
NH_2	amino group
NF-κB	nuclear factor κB
NIDDM	non insulin dependent diabetes mellitus
nm	nanometre
nM	nanomolar
nt	nucleotide
N.Sotho	northern Sotho
O	observed
Ob	obesity gene
O_2	molecular oxygen
OD	optical density
O_H	heavy strand origin of replication
O_L	light strand origin of replication
OXPHOS	oxidative phosphorylation
P	patient
p	petite
p_i	inorganic phosphate
PAGE	polyacrylamide gel electrophoresis
PAI-1	plasminogen activator inhibitor
PPAR	peroxisome proliferator-activated receptor
PCR	polymerase chain reaction
PBS	phosphate buffer saline
pH	indicates acidity: numerically equal to the negative logarithm of H^+ concentration expressed in molarity
Phe	phenylalanine
PI3	phosphatidylinositol 3 kinase
pmol	pico mole
pM	picomolar
POWIRS	profiles of obese women with insulin resistance syndrome
PP	pancreatic polypeptide
Pro	proline
PUCHE	Potchefstroom University for Christian Higher Education
q	designation for the long arm of a chromosome
R_a	rate of glucose appearance
R_d	rate of glucose disposal
RFLP	restriction fragment length polymorphism

ROS	reactive oxygen species
RNA	ribonucleic acid
rRNA	ribosomal RNA
S	svedberg unit
SE	standard error of proportion
sec	seconds
Ser	serine
SNP	single nucleotide polymorphisms
S.Sotho	southern Sotho
T and t	thymine (in DNA sequence)
T1DM	type 1 diabetes mellitus
T2DM	type 2 diabetes mellitus
T2D	type 2 diabetic
T_a	optimal annealing temperature
Taa I	restriction endonuclease isolated from *Thermus aquaticus* I with recognition site 5' – ACN↓GT – 3'
Taq polymerase	DNA deoxynucleotidyltransferase from *Thermus aquaticus*: EC 2.7.7.7
TBE	89.15 mM Tris (pH 8.0), 88.95 mM boric acid, 2.498 mM Na_2EDTA
TCA	Tricarboxylic acid cycle
Thr	threonine
T-loop	TψC loop
T_m	calculated annealing temperature
Tris	Tris®[1]: tris (hydroxymethyl)aminomethan:2-amino-2-(hydroxymethyl)-1,3-propanediol: $C_4H_{11}NO_3$
tRNA	transfer RNA
$tRNA^{Ala}$	transfer RNA molecule for adenine
$tRNA^{Arg}$	transfer RNA molecule for arginine
$tRNA^{Asn}$	transfer RNA molecule for asparagine
$tRNA^{Asp}$	transfer RNA molecule for aspartic acid
$tRNA^{Cys}$	transfer RNA molecule for cysteine
$tRNA^{Gln}$	transfer RNA molecule for glutamine
$tRNA^{Glu}$	transfer RNA molecule for glutamic acid
$tRNA^{Gly}$	transfer RNA molecule for glycine
$tRNA^{His}$	transfer RNA molecule for histidine
$tRNA^{Ile}$	transfer RNA molecule for isoleucine
$tRNA^{Leu}$	transfer RNA molecule for leucine
$tRNA^{Leu(CUN)}$	transfer RNA molecule for leucine with anticodon CUN
$tRNA^{Leu(UUR)}$	transfer RNA molecule for leucine with anticodon UUR
$tRNA^{Lys}$	transfer RNA molecule for lysine
$tRNA^{Met}$	transfer RNA molecule for methionine
$tRNA^{Phe}$	transfer RNA molecule for phenylalanine
$tRNA^{Pro}$	transfer RNA molecule for proline
$tRNA^{Ser}$	transfer RNA molecule for serine
$tRNA^{Ser(AGY)}$	transfer RNA molecule for serine with anticodon AGY
$tRNA^{Ser(UCN)}$	transfer RNA molecule for serine with anticodon UCN
$tRNA^{Thr}$	transfer RNA molecule for threonine
$tRNA^{Trp}$	transfer RNA molecule for tryptophan
$tRNA^{Tyr}$	transfer RNA molecule for tyrosine
$tRNA^{Val}$	transfer RNA molecule for valine
Trp	tryptophan
Tyr	tyrosine
U	units
UCP	uncoupling protein
UK	United Kingdom
USA	United States of America
UV	ultraviolet
Val	valine
WHR	waist to hip ratio
x g	gravitational force

[1] Tris® is a registered trademark of the United States Biochemical Corporation, Cleveland, OH, U.S.A.

LIST OF EQUATIONS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank the following individuals and institutions by expressing my sincerest appreciation for all the assistance given to me during my master's study in 2003:

All the individuals, patients and controls, who participated in this study by donating their blood samples, without which this research would not have been possible.

Prof Antonel Olckers, my supervisor, without whose guidance, encouragement and dedication, I would not have been able to accomplish my goals this past year. The many hours spent by her in assessing my work and organising assistance in the interest of furthering my scientific career made this degree possible. I will carry the skills learnt under her guidance with me for the rest of my life in whatever direction I may follow. **Dr Annamarie Kruger**, my co-supervisor, whose expert assistance kept me focused on the subject at hand and whose support I am truly grateful for. My mentor, **Tumi Semete**, whose countless hours of assessing, evaluating and encouraging enabled me to produce a product of which I can be proud. **Wayne Towers**, for the unfailing support as well as the scientific knowledge that was so willingly shared. Thanks to the other members of the team, **Annelize van der Merwe, Marco Alessandrini, Tharina van Brummelen, Christa Mouton, Desiré Hart, Desiré Dalton, Michelle Freeman and Jake Darby**, for a sense of humour that could alleviate the stress that we all encountered, and for their reliability coupled with their willingness to help at all times; each made his or her own invaluable contribution in assisting me on the road to completion of this project.

To **Prof. Peter Schwarz** for his collaboration and support, as well as the rest of the team from Germany, **Uta**, **Astrid** and **Jutta**, who aided in the collection of patient information and blood specimens without which this project truly would not have been able to reach completion.

The **Centre for Genome Research** and **DNAbiotec (Pty) Ltd** for providing the financial support and infrastructure, making this opportunity possible. The **PUCHE** staff for always providing all help that was requested, academic and administrative. **External companies**,

such as BioRad and Inqaba, for their patience in dealing with requests and enquiries and unfailing willingness to assist in finding answers.

I would like to thank my **mother**, extended family (**Gail, Nolene, Koosie, Janet and Rob**) and friends, especially **Danay**, for all their love and kind words of encouragement. But most of all I would like to thank **Rory**, my fiancé, for always being there to love and support me through the good times and the bad, who always pulled me back up when I didn't feel I could carry on. The patience that you have shown towards me through all my studies will always bear witness to your unending love for me. Thank you.

CHAPTER ONE

INTRODUCTION

Diabetes mellitus, a hyperglycaemic disorder, currently affects approximately 5 percent (%) of the global population. It has been estimated that by the year 2025, this figure will have doubled (Simpson *et al.*, 2003).

1.1 BACKGROUND

The main cause for the massive increase in the incidence of diabetes world-wide is the exercise-lenient lifestyle that an increasing number of individuals are adopting. For this reason, it is more likely that any genetic predisposition to the disorder will manifest itself. The lack of exercise prevents the energy ingested in the form of glucose to be utilised to its fullest capacity, resulting in phenotypic complications due to the excessive build-up of this sugar in the body (Marx, 2002).

1.2 PROBLEM STATEMENT

Elucidating a genetic cause underlying the occurrence of type 2 diabetes has been the subject of research over the last two decades and is discussed briefly in Chapter two. Recently, information regarding adipose tissue-released hormones and their influence on the consumption of glucose within the body has led to an improved understanding about the hormone adiponectin (Arita *et al.*, 1999) and other related proteins. These proteins, through the interaction with insulin, can affect the glucose-tolerance level of an individual. Polymorphisms or alterations within the genetic coding regions of these proteins are also being investigated for the effect that they could have on the diabetic phenotype.

Many genetic markers within the nuclear genome as well as the mitochondrial genome have been associated with diabetes (McCarthy and Froguel, 2002; Alcolado and Thomas, 1995). It has been observed that there are higher incidences of individuals with diabetes who have mothers with the disorder than there are affected individuals with affected fathers. Therefore due to the maternal mode of inheritance of mitochondrial DNA, which is

elaborated on further in Chapter two, the mitochondrial genome should be investigated with regard to the type 2 diabetic (T2D) phenotype. A mitochondrial-related form of diabetes has already been identified, known as maternally inherited diabetes mellitus and deafness (Van den Ouweland *et al.*, 1994).

The A3243G mitochondrial mutation (Van den Ouweland *et al.*, 1992), is more commonly associated with the MELAS (mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes) phenotype (De Vries *et al.*, 1994). This alteration has been found in European populations to be a causative mutation for the development of type 2 diabetes when present at a heteroplasmic level of over 30 % (Van den Ouweland *et al.*, 1994). This, as well as other diabetic-associated mutations, is discussed in Chapter 2.5.3.

A growing incidence of type 2 diabetes has been observed within the Black South African populations (Walker *et al.*, 2002). Few research projects have focused on this increasing need to elucidate a genetic predisposition within these populations towards the disorder. In 2001 a project undertaken by Towers *et al.* to determine the prevalence of the A3242G mutation within the Black South African population, discovered an alteration at the 3277 position from guanine to adenine. In a pilot study this alteration was observed in 3 % of the 100 diabetic patients who were investigated.

1.3 AIMS OF THIS STUDY

A control population was, however, not screened in the aforementioned investigation. It was, therefore, the aim of this current research project to screen and determine the prevalence of the G3277A alteration within a further 379 cohort composed of both control and patient individuals. The outlined aim was achieved via direct sequencing of a region of the mitochondrial genome encompassing the 3277 position, as well as restriction fragment length polymorphism (RFLP) analysis. Techniques and procedures that were utilised are described in Chapter three, with the resultant findings being discussed in Chapter four.

The cause of a disorder in one individual is often not identical to the cause of the same disorder in a second individual. For this reason certain therapies are effective in one group of individuals but not in another. Determining population-specific susceptibility alterations, therefore, has significant implications for the future of patient care.

1.4 OUTLINE OF DISSERTATION

Chapter two contains an outline of diabetes and complications of the disorder in general, concluded with the specific aims of this study. This is followed by the methodology and study design in Chapter three of how the investigation was pursued in terms of the aims outlined in Chapter two. Chapter four discusses the results obtained from the research undertaken and in Chapter five a conclusion is reached and recommendations are made for future research. Chapter six contains a list of the references utilised during the accumulation of information throughout the duration of this study.

CHAPTER TWO

THE MOLECULAR AETIOLOGY AND PATHOGENESIS OF TYPE 2 DIABETES MELLITUS

2.1 INTRODUCTION

Diabetes mellitus (DM) is a metabolic disorder due to the incorrect processing of glucose in the blood stream in conjunction with the body's dysfunctional interplay with insulin. It was realised in the 1970's that DM appears to consist of a number of heterogeneous syndromes characterised by a continuum of metabolic changes secondary to insufficient insulin action as well as various tissue changes (Fajans *et al.*, 1978). DM is a complex polygenic disease that was described by Himsworth in 1936 where the differences between the main types of diabetes are clearly outlined as well as their possible causes. This disorder is influenced by both environmental factors such as life-style, as well as genetic factors, with the interplay of many genes causing subtle alterations within the phenotype. It is hypothesised that there are twice as many people with genetic alterations that could predispose them to diabetes. However, these individuals do not express the diabetic phenotype as opposed to the number of individuals that do exhibit diabetic complications (Berdanier, 2001). The resulting phenotype is due to the disruption of cell signalling in the pancreatic beta (β)-cells, compromising insulin production and/or secretion. This perturbation causes alterations in signalling in many other cells and tissues of the body such as the brain, muscle, liver and adipose tissue (Margolis *et al.*, 2002).

2.2 DIABETES MELLITUS

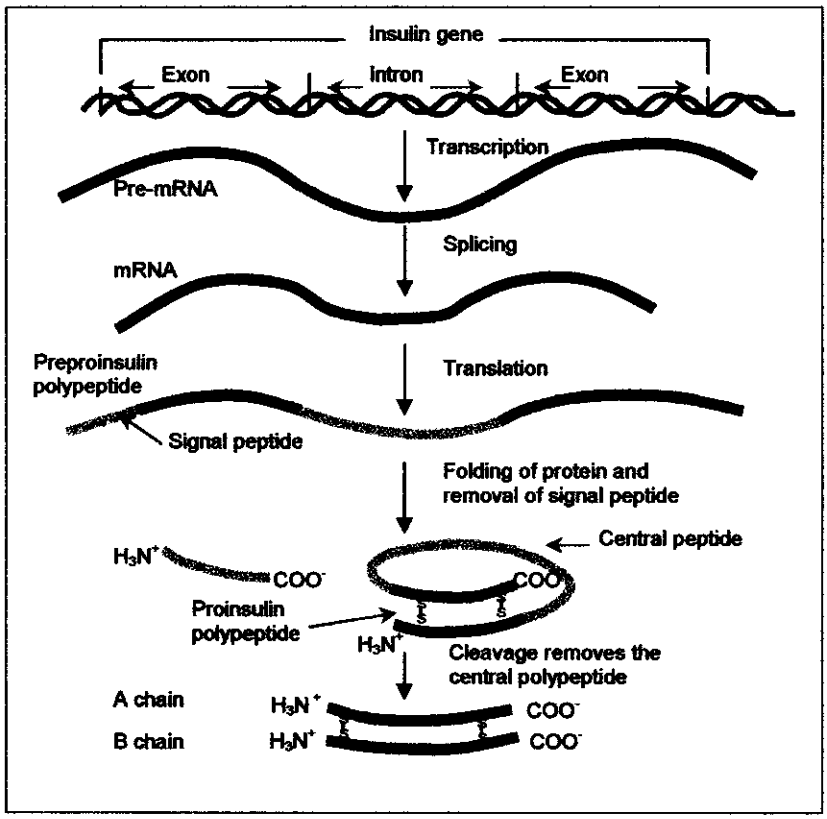
DM is prevalent in about 5-6 % of the general population (King *et al.*, 1998) and can be divided into two main clinical types. Type one diabetes mellitus (T1DM) is typified by the autoimmune destruction of the β-cells resulting in decreased insulin production, while type two diabetes mellitus (T2DM) is characterised by the body's lack of response to insulin due to many contributing factors (Vadheim and Rotter, 1989). In both instances, insulin is not being utilised correctly thereby not promoting the correct metabolic processing of glucose that naturally occurs under glucose-tolerant conditions.

2.2.1 The hormone insulin

Insulin has been in use as a treatment for diabetes for decades and was initially harvested from animals in large amounts to treat individuals with diabetes (Narang *et al.*, 1984). It is currently produced in genetically-modified micro-organisms to meet the demands of the growing diabetic population (Narang *et al.*, 1984).

2.2.1.1 Insulin

Figure 2.1: Synthesis of insulin



Adapted from Campbell, (1995).

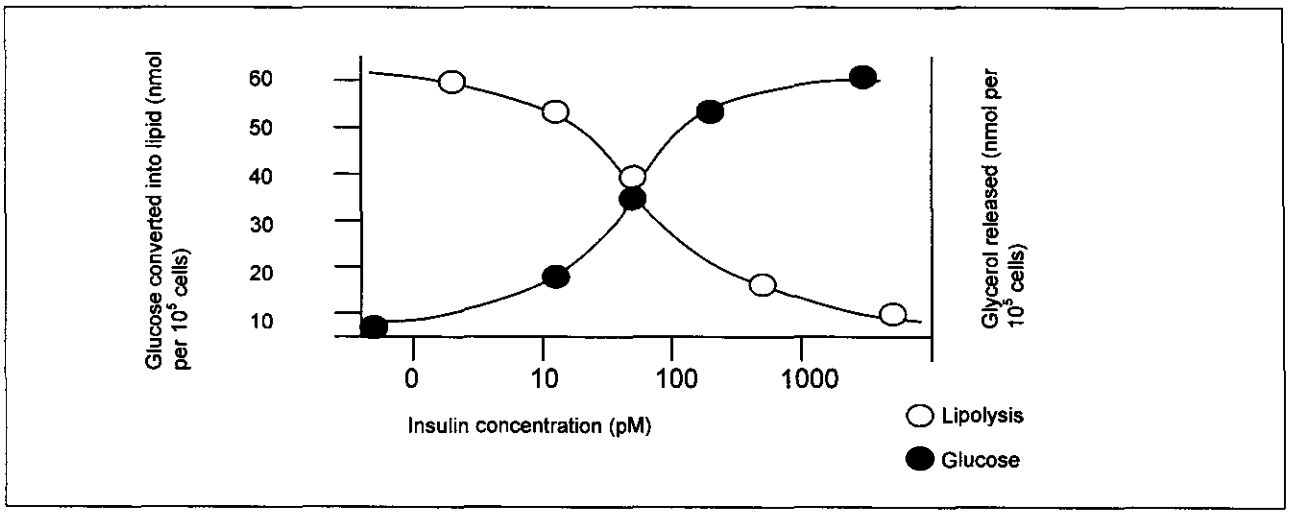
The main organ targeted during diabetes is the pancreas that has both exocrine and endocrine capabilities (Gannon and Wright, 1999). The exocrine tissue produces and secretes substances to aid in digestion while the endocrine tissue, known as the islets of Langerhans, produces and secretes hormones into the blood (Mader, 1996). There are four islet endocrine cells types: alpha (α), which produces glucagon; β, which produces insulin; delta (δ), which produces somatostatin; and pancreatic polypeptide (PP) producing cells (Gannon, 2001).

As depicted in Figure 2.1, the hormone insulin consists of two peptide chains attached via disulphide bonds. It is secreted as a preprohormone that is absorbed by cells. During the process of absorption, the signal peptide is cleaved, resulting in proinsulin. The general structure of proinsulin is NH_2-(B chain)-(C peptide)-(A chain)-COOH (Narang *et al.*, 1984). The proinsulin molecule is finally activated to produce insulin by the cleavage of the C peptide via peptidases and the formation of disulphide bonds (Narang *et al.*, 1984).

Insulin has both excitatory and inhibitory activities (Sonksen and Sonksen, 2000) and acts to control homeostasis between gluconeogenesis and cellular oxidation as well as various other metabolic functions (Rossetti and Giaccari, 1990). It inhibits endogenous glucose production mainly in the liver by reducing gluconeogenesis or glycogenolysis (Brown *et al.*, 1978) and stimulates hepatic and peripheral glucose uptake (Rossetti and Giaccari, 1990).

After one has eaten, glucose enters the blood stream and subsequently enters the cell. Within the cellular interior the glucose molecule is phosphorylated from which this phosphorylated molecule can form part of one of two processes: glucose storage via glycogen synthesis, which is the method by which glucose is stored in the body, or glycolysis where glucose is catabolised to produce energy in the form of adenosine triphosphate (ATP) thus maintaining cellular homeostasis (Rossetti and Giaccari, 1990). As shown in Figure 2.2, an insulin response is usually triggered by an increase of glucose in the blood (Sonksen and Sonksen, 2000). The primary function of insulin is not to stimulate the uptake of glucose by cells but to prevent alternative sources of energy from being utilised while glucose is present in the blood e.g. gluconeogenesis (Brown *et al.*, 1978), lipolysis (Torroni, 2000), ketogenesis (Reusch, 2002) and proteolysis (Sonksen and Sonksen, 2000). Therefore it accelerates the uptake of glucose into the cells from the bloodstream.

Figure 2.2: Insulin action on adipocytes



Adapted from Sonksen and Sonksen, (2000).

During prolonged fasting, glycogen stores become depleted and insulin levels decrease sufficiently to allow the catabolism of amino acids and lipids which are utilised as precursors for gluconeogenesis. Glucogenic amino acids along with the small amount of

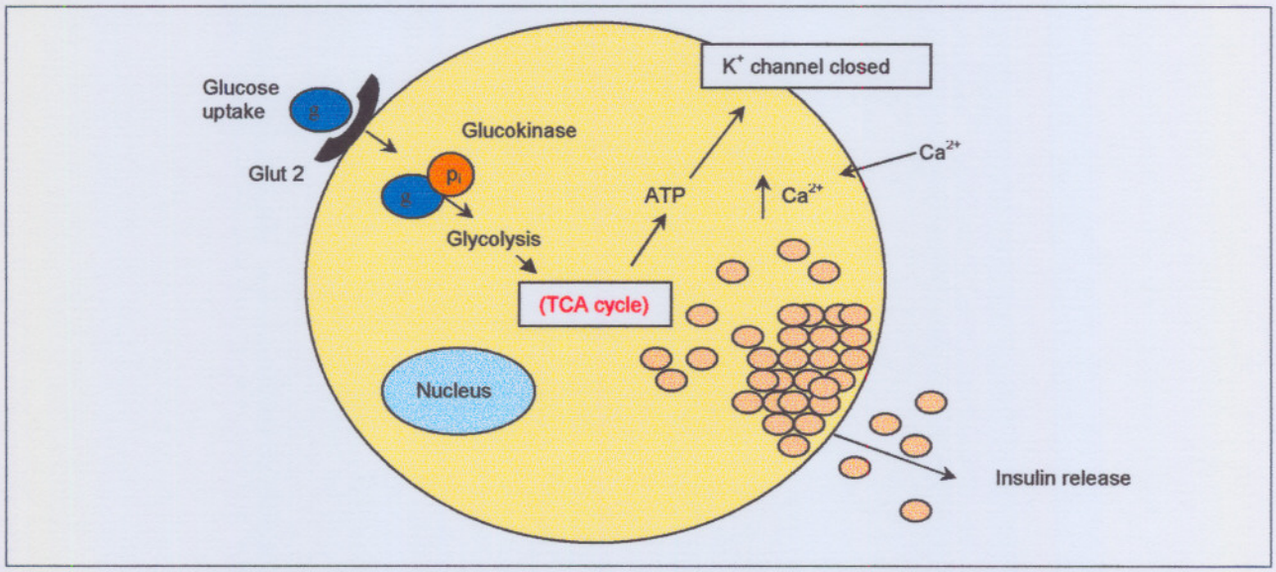
glycerol produced via lipolysis subsequently maintain the blood glucose levels in times of famine. This, however, causes a gradual loss of structural protein and eventually leads to pneumonia and death from prolonged starvation (Sonksen and Sonksen, 2000).

2.2.1.2 The β-cells

β-cells have a vital contribution in the pathogenesis of insulin dysfunction. The β-cells respond to a variety of hormones and nutrients by synthesising and releasing insulin (Margolis *et al.,* 2002). These cells are subject to signals generated by other cells of the islet, and those generated by other tissues (Margolis *et al.*, 2002). Elevated levels of glucose in the blood induce β-cells to respond with a large transient release of insulin. This is followed by a longer, slower, second phase of insulin secretion (Margolis *et al.,* 2002).

As depicted in Figure 2.3, the cascade of events in the β-cells that follow the uptake of glucose and its phosphorylation via glucokinase (GK) involves rapid and modulated changes in the energy-intermediate, ATP. These changes alter the ATP-sensitive potassium channels (K_{ATP}) which regulate the calcium ion (Ca^{2+}) fluxes. Utilising mice as a model in the investigation of diabetes in humans, it was illustrated that those without K_{ATP} do not present with a diabetic phenotype (Margolis *et al.*, 2002). This suggests that in addition to activating the classical K_{ATP}-dependent pathway which leads to insulin secretion, insulin secretogogues can stimulate insulin secretion via a K_{ATP}-independent pathway (Margolis *et al.*, 2002). A decline in glucose concentration, regardless of the cause, is known to be associated with a fall in the serum potassium (K^+) concentration (Brown *et al.*, 1978). The above implies that insulin affects K^+ transport into the cells independently of glucose transport (Brown *et al.*, 1978).

Figure 2.3: Glucose-stimulated insulin pathway



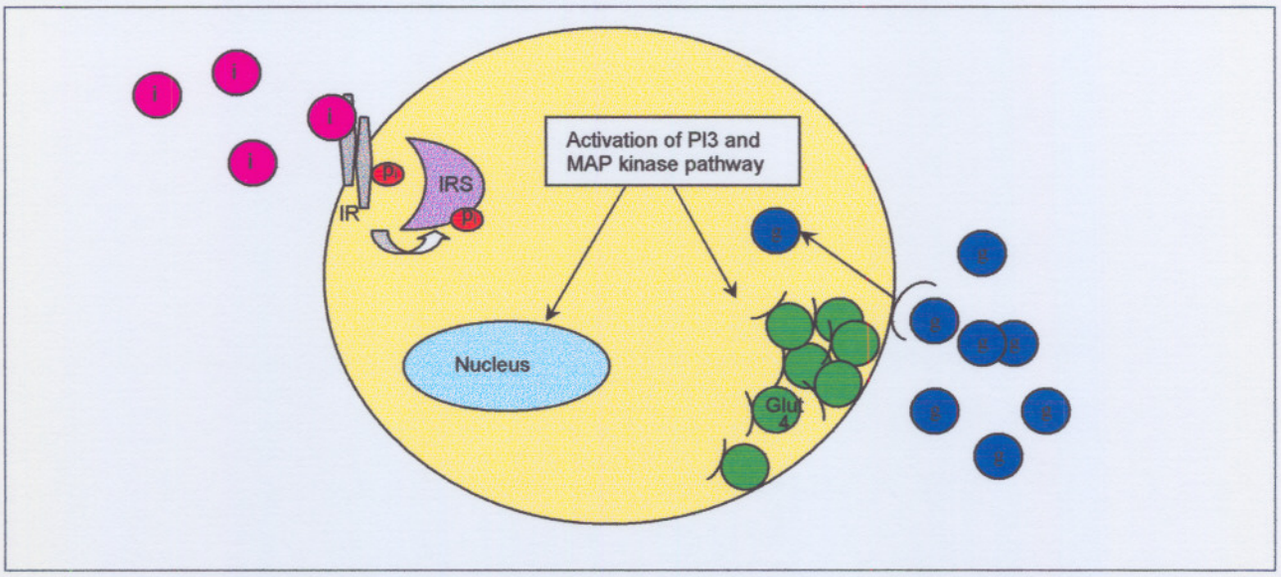
g = Glucose, Glut 2 = Glucose transporter 2, ATP = Adenosine triphosphate, TCA = Tricarboxylic acid. Ca^{2+} = calcium ions, K^+ = potassium ions, p_i = phosphate. Adapted from Gannon, (2001).

Dysfunction of the β-cells, which includes malfunction or even death, involves several factors as well as various other factors that are involved in insulin resistance (Marx, 2002). Fatty acids can trigger apoptosis in β-cells by first being converted into toxic ceramides which may be due to the body being resistant to the effects of leptin, an adipocyte-released hormone (Marx, 2002). There may also be evidence that the destruction of the β-cells during diabetes is as a result of increased generation of reactive oxygen species (ROS) formed in the mitochondria or decreased antioxidant capacity of β-cells (Margolis *et al.*, 2002).

2.2.1.3 Role of insulin in diabetes

As illustrated in Figure 2.4, the cellular response to insulin is mediated by one of two pathways, either the phosphatidylinositol 3-kinase (PI3) or the mitogen-activated protein (MAP) kinase pathway (Cusi *et al.*, 2000). When insulin is released into the blood stream, it binds to the membrane-bound hetero-oligomeric insulin receptor (IR) located on the surface of the cell. The binding of insulin stimulates the autophosphorylation of the receptor. Phosphorylation of the receptor allows for the phosphorylation of the cellular proteins, insulin receptor substrate (IRS)-1 and IRS-2, which are recruited as part of the protein complex. These are subsequently coupled to the MAP and PI3 pathways which lead to the translocation of vesicles containing the glucose-transporter (Glut 4) to the plasma membrane and facilitates glucose uptake into the relevant tissues (Gannon, 2001), where it is processed to produce the required energy.

Figure 2.4: Insulin action in target tissues



i = insulin, IR = insulin receptor, IRS = insulin receptor substrate, MAP = mitogen-activated protein kinase signalling cascade, PI3 = phosphoinositol 3 kinase signalling cascade, Glut 4 = glucose transporter, p_i = phosphate, g = glucose. Adapted from Gannon, (2001).

The MAP pathway is associated with mitogenic effects such as cell growth and proliferation which are maintained during insulin resistance. The PI3 pathway, which is also initiated by the phosphorylation of IRS-1 and -2, is responsible for the metabolic effects of insulin including glucose transport and glycogen synthesis as well as lipid metabolism. During insulin resistance, these processes are affected due to the impaired phosphorylation (Cusi *et al.*, 2000).

Glucose transport is the major determinant of glucose disposal at low insulin concentrations (Rossetti and Giaccari, 1990). As mentioned earlier in this section, glucose is taken up by the cells via glucose transporters i.e. glucose transporter one to glucose transporter six. Glut 4 is present in muscle and adipose tissue and is 'insulin sensitive' (Sonksen and Sonksen, 2000). This implies that in addition to the Glut molecules already in the cell membrane, additional Glut receptors are present in the cell cytoplasm that could be recruited in response to insulin towards the membrane to aid in glucose uptake (Sonksen and Sonksen, 2000). It is postulated that there are two major and independent defects in the skeletal muscle of diabetic individuals. One is a proximal defect i.e. glucose transport or phosphorylation, which is responsible for the impaired total glucose uptake at low plasma insulin concentration. The other is a distal defect i.e. glycogen synthesis, with its primary function, that of intracellular distribution of glucose, being altered (Rossetti and Giaccari, 1990). The presence of prolonged moderate hyperglycaemia and hypoinsulinaemia determines these two distinct cellular defects (Rossetti and Giaccari, 1990). When experiments were performed on mice where the IR was deleted in muscle

utilising tissue-specific deletions it was observed that this muscle-specific deletion did not result in glucose-intolerance or insulin resistance (Bruning *et al.*, 1998). This implies that an unidentified compensatory mechanism must exist, as it has been previously illustrated that muscle tissue is a primary site for insulin resistance in T2DM patients (Gannon, 2001). However, the same deletion in the liver resulted in severe glucose intolerance (Michael *et al.*, 2000).

It has been illustrated that increased levels of insulin does not induce peripheral cells to take up glucose. Therefore the process of the cells obtaining glucose is not a primary concern, as this function is still being fulfilled. The two main metabolic processes that are impaired during hyperglycaemia acting in parallel are:

1) excess glucose in the blood stream resulting in osmotic diuresis, and
2) ketone overproduction causing metabolic acidosis.

Fasting hyperglycaemia is an initial indicator of early stages of diabetes where impaired glucose tolerance (IGT) is observed. There is usually excessive glycosuria as well as thirst, polyuria and weight loss. These conditions are all caused by the excess production of glucose in the liver. Glycosuria develops due to the level of glucose taken in by the peripheral tissues (R_d) being less than the glucose appearance in the blood from liver production (R_a). Generally these two processes need to be maintained in homeostasis and if too much glucose is present in blood circulation, as there is in a hyperglycaemic state, a renal threshold is reached and the excess is secreted in the urine. As a result of high levels of glucose in the blood, an osmotic movement of water out of the cells will be induced. Due to the disposal of excess glucose in the blood, a high level of glucose is observed in the kidneys which the body is unable to reabsorb and that, therefore, require removal. Osmotic diuresis and polyuria develop and the body begins dehydrating despite excess water that is being taken in orally due to an increased thirst. (Sonksen and Sonksen, 2000).

As mentioned earlier (section 2.2.1.3) the cause for the excess glucose in the blood, characteristic of the diabetic phenotype, is the production of glucose from the liver through alternative processes not being inhibited, in addition to decreased glucose absorption from the blood. Under normal glucose-tolerant conditions, alternative processes, such as lipolysis, ketogenesis and proteolysis are inhibited after eating, therefore the only glucose that is being metabolised is that which has been ingested. Gluconeogenesis relies on protein breakdown, as well as small contributions from lipolysis, which results in protein

wasting and eventually structural breakdown. Lipids and ketones are also utilised by several other tissues as alternative sources of energy to minimise protein loss. As these processes are continuing in insulin-insensitive individuals, an accumulation of free-fatty-acids (FFA) and ketones as by-products are observed in the liver. From these molecules the liver produces ketoacids which are secreted in the blood and absorbed into peripheral cells. As the concentration of these ketoacids increases, the higher the acidity of the blood becomes, and acidosis eventually occurs. (Sonksen and Sonksen, 2000).

Similar complications are observed in the two main types of diabetes commonly known (discussed below). Their symptoms are, however, brought about by defects in significantly diverse mechanisms.

2.2.1.3.1 Type 1 diabetes mellitus

The less common type of diabetes known as T1DM or insulin dependent diabetes mellitus (IDDM) is due to the autoimmune destruction of the pancreatic β-cells (Lindgren and Hirschorn, 2001) by lymphocytes, macrophages and neutrophil granulocytes (Gannon, 2001) and is usually found in about 10% of all diabetic cases (Vadheim and Rotter, 1989). Its characteristic phenotype is the loss of insulin in the body (Wucherpfennig, 2001) and the inability to produce insulin due to β-cell destruction (Marx, 2002). In the absence of insulin, cells are not prompted to absorb glucose. Insulin loss occurs following the inability of the pancreas to produce insulin, a condition known as insulitus. The β-cells are targeted as antigens by the T-cells in the blood stream. This destruction prevents insulin production, which could result in unconsciousness, as the brain requires a constant supply of glucose for optimal functioning. To counteract this, the injection of insulin is required to perform the function that the pancreas is no longer capable of (Mader, 1996).

Susceptibility to IDDM is brought about by susceptibility alleles in the major histocompatibility complex (MHC) or human leukocyte antigen (HLA) as it is named in humans, which is located on chromosome 6p21 (Wucherpfennig, 2001). In this condition, IDDM, the liver is not stimulated to store glucose and cells are not able to utilise the circulating glucose for their energy requirements. This defect leads to the alternate production of glucose via gluconeogenesis. Therefore as described in paragraph 2.2.1.3, movement of water out of cells subsequently occurs, which results in osmotic diuresis causing dehydration. Alternative sources of energy are sought out which results in protein breakdown. Therefore, along with gluconeogenesis, there is an increased level of amino

acids in the muscle from proteolysis. Together with proteolysis, lipid breakdown, which is usually largely inhibited after the ingestion of glucose-containing food, causes FFA and glycerol release. An increase of glycerol in the environment makes it hypertonic. The FFA and amino acids, which are insoluble, are transported to the liver where ester bonds are formed within these molecules to make them more soluble. This produces ketoacids which are rapidly taken up by the cells. An accumulation of these in the cellular environment produces metabolic acidosis which lowers the pH of the blood, which in conjunction with the dehydration, causes coma and eventually death (Mader, 1996). These repercussions of insulin loss in the body are duplicated in a T2DM phenotype as this disorder, through alternative processes, eventually also suffers diminished β-cell mass and therefore insulin production.

A small percentage of individuals develop T1DM as a result of complications associated with viral infection. It has been estimated that 12-20% of individuals affected by the rubella virus, which can cause a condition known as congenital rubella syndrome, develop diabetes 5-20 years post-infection (Rewers and Hamman, 1995). Certain viruses express antigens on their surfaces that resemble proteins expressed by the β-cells. The body builds up immunity to these antigens on the surfaces of these pathogens and in turn develop autoimmunity to these proteins found in the β-cells, which ultimately will lead to their destruction and eventual loss. (Szopa *et al.*, 1993)

2.2.1.3.2 Maturity onset diabetes of the young

Maturity onset diabetes of the young (MODY) is a collection of rare inherited forms of diabetes that result from mutations in single genes (Margolis *et al.*, 2002) and are a subtype of T2DM (Gannon, 2001). MODY is distinguished by an autosomal dominant and monogenic mode of inheritance as well as mild hyperglycaemia due to a β-cell secretary defect (Fajans *et al.*, 1978; Tattershall and Fajans, 1975). Unlike T2DM (see section 2.2.1.3.3), it is characterised by low insulin production as well as decreasing glucose-stimulated-insulin-release (Gannon, 2001). Mutations in the genes encoding GK and in five factors involved in regulation of β-cell gene transcription as well as five other genes that encode transcription factors which regulate genes involved in insulin production, have been identified as both necessary and sufficient to cause this form of diabetes (Bell and Polonsky, 2001).

2.2.1.3.3 Type 2 diabetes mellitus

The most prevalent form of DM is T2DM, also known as non insulin dependent diabetes mellitus (NIDDM). T2DM is caused by a complex interaction of multiple susceptibility loci and environmental factors (Margolis *et al.,* 2002) and accounts for about 90% (Vadheim and Rotter, 1989) of the entire diabetic population world-wide. T2DM is defined by hyperglycaemia which could be due to impairment of insulin action and/or secretion (Lindgren and Hirschorn, 2001). It is characterised by cellular resistance to the effects of insulin (Himsworth, 1936) combined with the eventual failure of the β-cells to produce adequate insulin to compensate for the resistance (Alberti and Zimmet, 1998). This group of disorders is a major cause of heart disease, renal failure, blindness, damage to the feet and generally any pathway associated with the necessary processing of glucose (Nathan, 1995; Marx, 2002).

In this investigation, the molecular pathogenesis of T2DM was investigated with specific focus on the mitochondrial (mt) involvement in the development of the disorder. Even though individuals with mt defects cannot be classified exclusively as T2DM, many of the characteristics of mt defects include those of T2DM. A general overview of the mitochondrion, as well as its contribution to the development of diabetes, is discussed in section 2.4.

2.3 THE AETIOLOGY OF NON INSULIN DEPENDENT DIABETES MELLITUS

In any form of diabetes the major complication originates from the increased extracellular glucose concentration, which is aggravated by the utilisation of alternative sources of energy within the intracellular environment. Since cells require glucose for energy, complications arising from this are far-reaching if left untreated. The clinical features of T2DM are commonly associated with hyperinsulinemia, hypertension, hyperlipidemia, obesity, insulin resistance (Gutierrez *et al.,* 1998), atherosclerosis, microangiopathy, neuropathy, complications during pregnancy, overweight babies, cataracts (Fajans *et al.*, 1978), hypercoagulation and fibrinolysis which all increase the risk of cardiovascular diseases such as coronary artery disease or CAD (Kannel *et al.*, 1979). A few of the factors influencing the development of the conditions mentioned above are discussed briefly in the following sections to outline examples of different pathways via which T2DM could arise or how different hormones and other factors can influence the progress of the

diabetes. The aspects discussed, are by no means an overview of all the elements that could affect the development of the disorder.

2.3.1 Physiological symptoms of T2DM

To fully understand the perturbations that occur within any disorder, analysts undertaking research into disorders in general, need to investigate the network of systems within the cell. The intricacy of these networks is demonstrated in DM which involves the interplay of many intracellular and extracellular factors as well as a multitude of cellular functions. Many biochemical, genetic and homeostatic pathways, previously thought to act independently of one another, may in fact have secondary effects on, or interactions with, one another that is beyond current understanding. For instance, triglycerides and their processing have been identified as key elements in the pathogenesis of a vast majority of patients with T2DM. Yet this observation does not provide answers to many questions relating to the onset and complications surrounding diabetes.

2.3.1.1 The role of obesity in diabetes

With an increasing occurrence of obese individuals in the global population over the last few years, a direct increase in the prevalence of T2DM has also been observed. This does not imply, though, that all obese individuals are diabetic, or that all diabetic patients are obese (Marx, 2002). The diabetic phenotype involves more than just obesity as a complication. Yet obesity is defined as an accumulation of excess body fat which frequently accompanies insulin resistance, hypertension, dyslipoproteinaemia and vascular disease (Arita *et al.*, 1999).

In a correspondence written to *Nature* by Danforth in 2000, he suggested a hypothesis that may lend credence to the role of obesity in DM. Adipocytes are terminal cells, meaning that when filled to capacity the cells are extremely insulin resistant. The only way that excess energy taken in can be balanced, is if the adipose organ expands (Danforth, 2000). Danforth suggested that T2DM in obese individuals is a result of the inability of adipose tissue to expand, therefore not being able to accommodate excess calories. The excess energy in the form of FFA is then stored in the liver, muscle, blood and urine, as seen in diabetic individuals. This hypothesis is supported by a study performed by Okuno *et al.*, (1998) where Zucker rats were treated with thiazolidinediones, an anti-hyperglycaemic

agent. Insulin sensitivity improved due to an increasing number of new adipocytes being generated (Okuno *et al.*, 1998). Therefore too few adipocytes could predispose an individual to T2DM.

Until recently, fatty tissue was hypothesised to be a storage depot only. Apart from its function of storing fatty acids, it is now known to have a dynamic role in releasing a variety of hormone-like substances that circulate in the blood and affect other tissues (Marx, 2002).

Some examples of these compounds are leptin, resistin, adiponectin, and even the fatty acids themselves. There are many others, many of which are cytokines and hormones involved in potentially regulating glucose homeostasis. Leptin, the product of the obesity (*ob*) gene (Zhang *et al.*, 1994) is an antisteatotic hormone (Unger and Orci, 2000) known for suppressing the appetite and as a result reduces weight-gain. The human resistin gene is a protein associated with pulmonary inflammation (Smith *et al.*, 2003) in addition to also modulating certain points in the insulin-signalling pathway (Sentinelli *et al.*, 2002). Lastly, adiponectin appears to promote the effects of insulin thereby making the body more sensitive to insulin, but its production has been found to be decreased in certain obese individuals (Arita *et al.*, 1999).

However, insulin itself has an important function to fulfil within adipocytes. Insulin has an anti-lipolytic effect, which enhances the clearance of FFA from the blood stream. Lipids can be important secretogogues of insulin, such as a hormone-sensitive lipase. If these lipases are secreted, it induces insulin to be secreted to prevent the breakdown of triglycerides into fatty acids, which interferes with other pathways (Brown *et al.*, 1978). Fatty acids in obese individuals accumulate in muscle, a prime insulin target. These, therefore, interfere with the pathway that transmits insulin signals into the muscle cell interior and prevents the removal of glucose from the bloodstream, storing it as glycogen. Therefore if insulin sensitivity can be improved, the consequences of dyslipidaemia from insulin resistance i.e. atherogenicity, could be limited (Reusch, 2002).

2.4 MITOCHONDRIAL COMPONENT OF DIABETES

With continuous investigation showing that maternal inheritance plays a larger role in diabetes than previously realised, it is important to study that which is maternally inherited

i.e. the mitochondria, parental imprinting of nuclear genes, or the intrauterine environment (Suzuki *et al.*, 2003). It has also been shown that the insulin secretory ability and blood glucose regulation are more severely impaired in the patients whose mothers are glucose intolerant (Lindgren and Hirschorn, 2001), thereby lending credence to the investigation of the mitochondria in the development of diabetes.

DM linked to mutations in the mitochondrial deoxyribonucleic acid (DNA) is referred to as maternally inherited diabetes mellitus and deafness or MIDD (Van den Ouweland *et al.*, 1994) or mt diabetes mellitus (Katagiri *et al.*, 1994; Gerbitz *et al.*, 1995). MIDD or mitochondrially related diabetes have been investigated for many years (Van den Ouweland *et al.*, 1992; Reardon *et al.*, 1992; Gerbitz *et al.*, 1995) and are not as rare as was previously thought. Many disorders manifest with diabetes as a symptom or side-effect which can be attributed to mt involvement. This subset of diabetes is characterised by maternal inheritance, progressive impairment of insulin secretion leading to an insulin-requiring state and a clinical picture of NIDDM. Manifestation of the IDDM phenotype at the onset is possible (Katagiri *et al.*, 1994). Individuals identified to have this form of diabetes appear to fall somewhere in between that of a type one diabetic and a T2D in terms of management of the clinical symptoms. They are not excessively overweight even though exercise may not be an option, as many mutations in the mitochondria affect the muscle through alterations in ATP production (Berdanier, 2001). This can be explained if one looks at the mitochondria in more detail including its function.

2.4.1 Mitochondrial genetics and diabetes

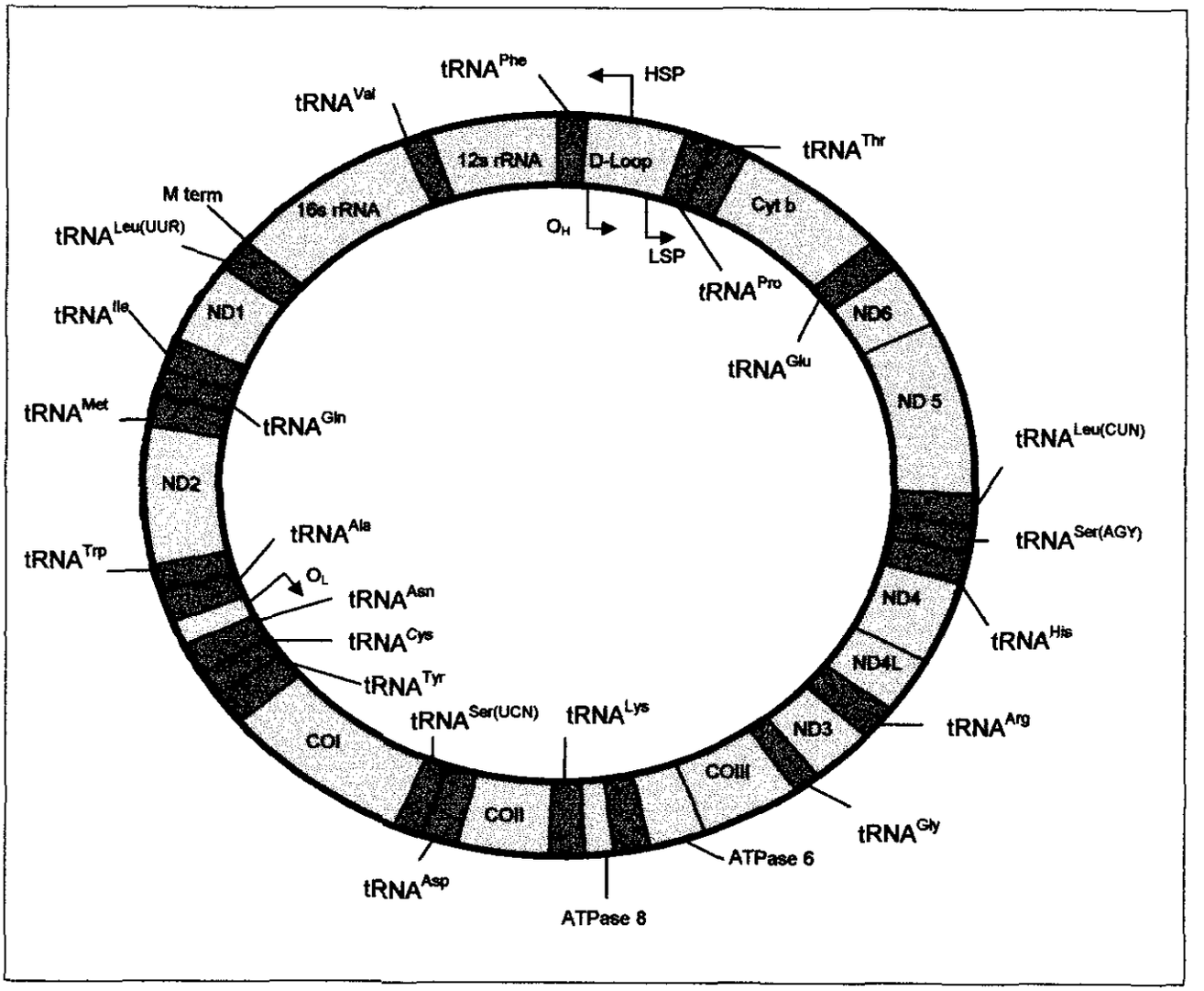
The mitochondria, often referred to as "the power houses of the cell", are the source of more than 90% of the cellular energy (Chance *et al.*, 1979). The energy is produced in the form of ATP, through the oxidative phosphorylation (OXPHOS) pathway involving the five complexes as shown in Figure 2.5 (Lamson and Plaza, 2002). Their function is intimately related to insulin secretion and possibly insulin action (Gerbitz *et al.*, 1995).

2.4.1.1 The mitochondrial genome

The mitochondria are endosymbionts that formed symbiotic relationships with eukaryotes very early in evolution and have evolved into a unique mammalian intracellular organelle (Alcolado and Thomas, 1995). They contain an inner and an outer membrane with the

inner core containing two to ten circular DNA molecules (Van den Ouweland *et al.*, 1994) of 16569 base pairs (bp) in length (Anderson *et al.*, 1981) attached transiently to the inner mt membrane (Shearman and Kalf, 1977). The genome has one heavy strand containing mostly guanine and adenine residues, and a light strand containing mostly cytosine and thymine residues. Both of these strands contain coding sequences (Alcolado and Thomas, 1995). Each human cell contains thousands of mitochondria, which are the sites of the vast majority of intracellular energy (ATP) production (Alcolado and Thomas, 1995).

Figure 2.5: Schematic representation of the mitochondrial genome



The dark shaded areas are the 22 tRNA, the open regions denote the control regions, rRNA genes and the protein-coding genes. D-loop = displacement loop, CO I-III = cytochrome c oxidase subunits I-III. ATP = ATPase. ND = NADH dehydrogenase subunits. Cyt = cytochrome. O_H and O_L = origins of the heavy (H) strand and light (L) strand mtDNA replication respectively. HSP and LSP = promoters for transcription from the H and L template strands respectively. M term = transcription termination site. Adapted from Hess *et al.*, (1991) and MITOMAP, (2003).

The mt genome encodes 13 enzyme subunits (complex I-V, except complex II) involved in the OXPHOS pathway (Alcolado and Thomas, 1995), all of which also require nuclear-derived gene products. It also encodes 22 transfer ribonucleic acids (tRNAs) and 2 ribosomal RNAs (rRNAs) as reported by Alcolado and Thomas in 1995. Due to the fact

that the mt genome contains virtually no introns, lacks protective histones and has no effective DNA repair mechanisms, the mitochondria is more vulnerable to mutation than its nuclear counterpart, which possesses these protective mechanisms (Kunkel *et al.,* 1981; Wallace, 1992; Johns, 1996; Yakes and Van Houten, 1997).

2.4.1.2 Mitochondrial transcription and replication

Mitochondrial DNA (mtDNA) replication appears to be under relaxed control with lack of restriction with regard to cell cycle phase. Figure 2.5 illustrates the origin of heavy-strand (O_H), or leading strand, replication found in the displacement loop (D-loop) while the origin of replication of the light-strand (O_L), or the lagging strand, is found within a cluster of five tRNA genes. O_H replication is initiated before O_L replication which occurs once the leading strand has replicated about two thirds of the way (Larsson and Clayton, 1995).

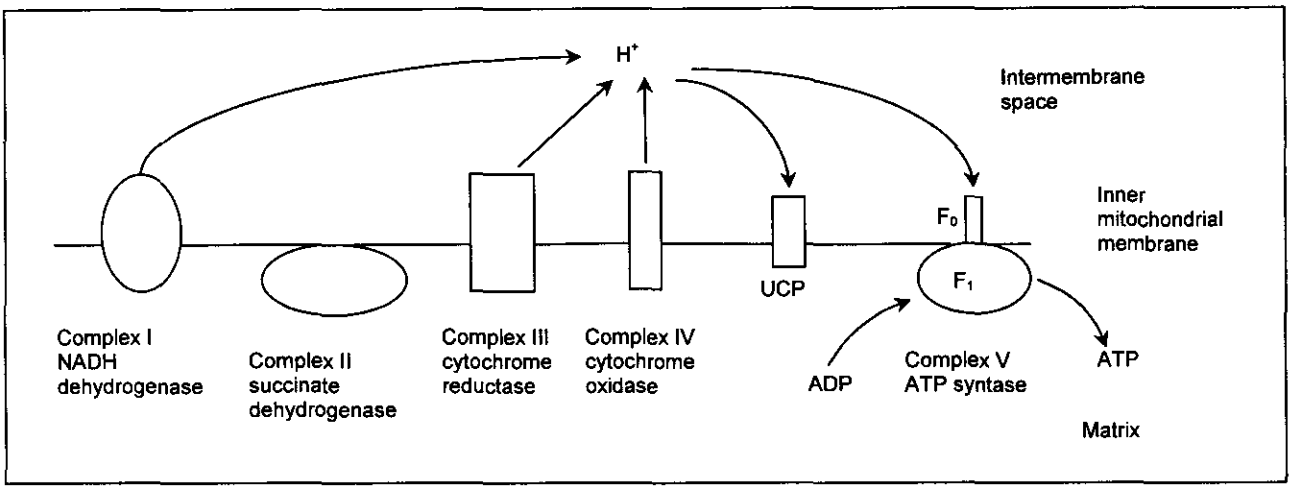
The D-loop, a short triple-stranded structure, is the control site for both transcription and replication (Clayton, 1984), with each strand having its own promoter i.e. heavy-strand promoter (HSP) and light-strand promoter or LSP (Larsson and Clayton, 1995). The human mtDNA transcription termination site is located at the 16s rRNA and $tRNA^{Leu(UUR)}$ gene boundary (Hess *et al.*, 1991). This is a controlling factor for the relative synthesis of rRNA compared to other RNA molecules and it has been observed that transcripts from the HSP are in different relative amounts with genes being more proximal to the promoter i.e. the 12S and 16S rRNA genes being transcribed at a higher frequency than those more distal (Clayton, 1984).

2.4.2 Mitochondria and glucose metabolism

The primary function of the mitochondria is energy production via the OXPHOS pathway. As mentioned in section 2.4.1.1 there are five multiprotein complexes involved in the OXPHOS pathway that all ultimately function to carry hydrogen ions (H^+) into the mt matrix. The transfer of electrons results in a gradient that in effect leads to the formation of ATP through release of the ions back over the mt inner membrane (Larsson and Clayton, 1995). This process is illustrated in Figure 2.6. The five complexes are complex I, the nicotinamide adenine dinucleotide (NADH) dehydrogenase complex; complex II, the succinate dehydrogenase complex; complex III, the cytochrome reductase complex;

complex IV, the cytochrome oxidase complex and complex V, the F_1F_0 ATP syntase complex (Lamson and Plaza, 2002).

Figure 2.6: Schematic representation of the mitochondrial respiratory chain



NADH = nicotinamide adenine dinucleotide, UCP = uncoupling protein, ADP = adenosine diphosphate, ATP = adenosine triphosphate, F_1 and F_0 forms ATPase complex, H^+ = protons. Adapted from Lamson and Plaza, (2002).

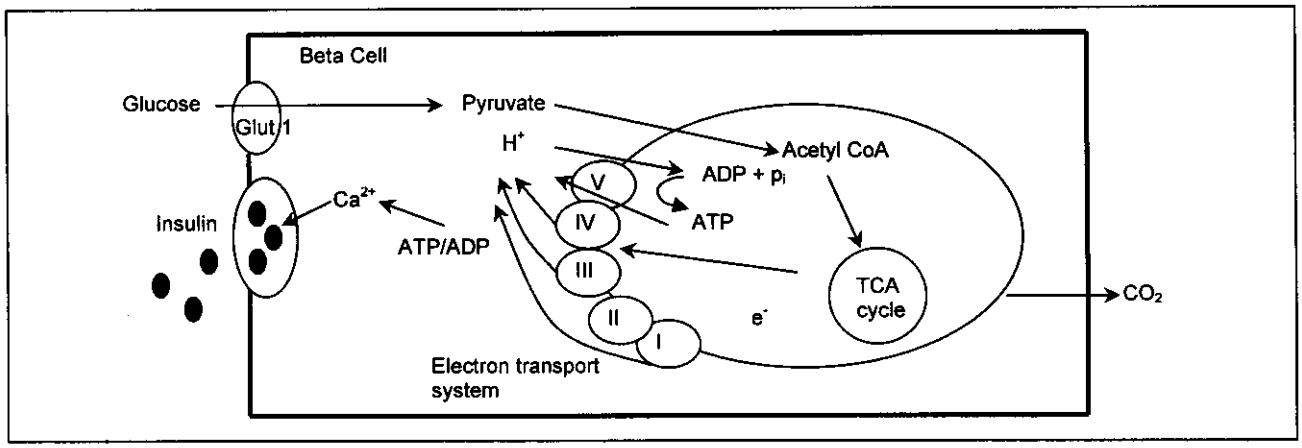
The energy production process begins with the oxidation of NADH and the entry of the H^+ into the respiratory chain. As protons are pumped out of the matrix into the intermembrane space, each complex moves electrons along the chain. The ultimate phosphorylation of adenosine diphosphate (ADP) to ATP occurs as a result of a proton gradient created by the oxidation of various compounds by the first four complexes. The proton gradient creates a transmembrane potential utilised by complex V to drive the synthesis of ATP (Lamson and Plaza, 2002).

ATP syntase consists of two portions F_0 and F_1. The F_0 portion is embedded in the inner mitochondrial membrane while the F_1 portion protrudes out into the inner matrix. The complex moves and rotates as energy from the respiratory chain or H^+ is being captured. This movement is essential for the activity of the complex and the production of the energy-rich ATP (Berdanier, 2001).

The maintenance of homeostasis of glucose and insulin is based upon the ATP generated by glucose in the β-cell. As depicted in Figure 2.7 glucose is transported across the cellular membrane by Glut-1 where it is converted to pyruvate via the glycolytic pathway. The latter molecule is transported into the mitochondria where it is further broken down by the tricarboxylic acid cycle (TCA) producing a small amount of ATP as well as reducing flavin adenine dinucleotide (FADH) and NADH which further produces additional ATP. The ATP/ADP ratio increases as this process continues, which allows the ATP-sensitive K^+

channels to close, depolarising the Ca^{2+} channels. This action causes an influx of Ca^{2+} into the cytosol, triggering the exocytosis of insulin secretary vesicles produced by the Golgi apparatus. Insulin processing, which is ATP and pH dependent, takes place within the vesicles ultimately resulting in insulin release (Lamson and Plaza, 2002).

Figure 2.7: Schematic representation of the release of insulin in the β-cell



Glut-1 = glucose transporter 1, ADP = adenosine diphosphate, ATP = adenosine triphosphate, H^+ = protons, TCA = tricarboxylic acid cycle, e^- = electron, CO_2 = carbon dioxide. I, II, III, IV and V = complex I, II, III, IV or V, Ca^{2+} = calcium ion, p_i = phosphate, CoA = coenzyme A. Adapted from Lamson and Plaza, (2002).

One way in which the development of diabetes can be influenced by the respiratory chain, leading to insufficient insulin production, is through diet, with a classical symptom of the common T2D individual being obese. A saturated-fat-rich diet reduces the inner membrane fluidity and therefore impairs the movement of the adenosine triphosphatase (ATPase), causing inefficient energy generation by the OXPHOS pathway in the form of ATP (Berdanier, 2001). This has been postulated in work performed on BHE/Cdb rats where the animals were fed different forms of dietary lipids and the ADP:O ratio measured to determine the mobility of F_1F_0ATPase (Kim and Berdanier, 1998). A rat strain-dependent temperature difference as well as a decrease in ATP synthesis efficiency was observed that could be explained by the rigidity of the surrounding lipid environment of the ATPase molecule (Kim and Berdanier, 1998).

Another pathway of mt involvement in diabetes is oxidative stress, which can cause a decrease in the number of mitochondria present within the cell. As noted at the beginning of section 2.4, ATP is produced through the OXPHOS pathway. A negative side-effect of this pathway is the generation of ROS through the 'leaking' of electrons to oxygen forming superoxide radicals (Cooke *et al.*, 2003). The accumulation of these reactive species results in harmful, and in many cases, irreparable damage.

2.4.2.1 Reactive oxygen species

In the presence of increased concentrations of glucose, endothelial generation of ROS is shown to be enhanced. This will appear in all cells where the uptake of glucose is not dependent on the presence of insulin. ROS can be formed by several candidate pathways: nicotinamide adenine dinucleotide phosphate oxidase, the mt respiratory chain, xanthine oxidase, the archidonic cascade and microsomal enzymes (Cross and Jones, 1991).

Endothelial cells in a hyperglycaemic environment cannot control their need for glucose, producing an accelerated flux of glucose through glycolysis. This in turn produces excessive pyruvate that enters the TCA cycle, overloading the mitochondria thereby causing excessive ROS generation (Schmidt and Stern, 2000). Hyperglycaemic conditions also activate nuclear factor κB (NF-κB), which is sensitive to the oxidant environment (Schreck *et al.*, 1991; Schreck *et al.*, 1992; Marui *et al.*, 1993) that links hyperglycaemia to the expression of multiple genes related to vascular stress responses (Collins, 1993).

The presence of ROS in the mt environment contributes to the increased occurrence of mutations within the genome via DNA damage. These mutations are transferred to the increasing number of mitochondria as they divide through normal cellular activities (Wei, 1998). With a 10-20 fold higher rate of mutation than the nuclear genome (Richter, 1992), defective electron transport complex subunits would be encoded by the mutated mt genome and could eventually cause increased superoxide production of physiological concentrations of glucose. The five ROS-producing pathways mentioned above would be continually activated despite the absence of hyperglycaemia (Schmidt and Stern, 2000).

2.4.3 Mitochondrial involvement in type 2 diabetes

MtDNA is characterised by high nucleotide (nt) diversity between species and even between different populations within one species (Finnilä, 1999). Most pathogenic mutations occur in a heteroplasmic state (Finnilä, 1999). If all mt genomes in the cell contain the same genotype, they are classified as being homoplasmic. Varying amounts of heteroplasmy and the distribution within the different tissues of the body and even between individuals is one plausible explanation for the wide variation in phenotypes observed in patients with mt disorders. This is due to the fact that different tissues have different dependencies on oxidative phosphorylation for normal function (Larsson and Clayton, 1995).

The mitochondrion has been found to play an integral part in the insulin system. 0.1% to 9% of the diabetic population may be affected due to a mutation in the mtDNA (Berdanier, 2001). Genetic alterations associated with T2DM are often expressed via mutations in the mt genome. They can occur in either the homoplasmic or heteroplasmic state leading to a number of genetic disorders all ultimately expressing T2DM. (Lamson and Plaza, 2002). A number of these mutations have already been investigated and an association between diabetes and particular populations has been established, as illustrated in section 2.5.3.

In an investigation by Katagiri in 1994, four out of 300 diabetic patients were found to have diabetes due to a mt gene mutation showing that this aetiology of diabetes is not particularly rare and merits further investigation. The influence of the mitochondria can be either quantitative or qualitative. The qualitative input refers to genetic mutations in the mtDNA, while the quantitative effect refers to decreases in mtDNA copy number being linked to diabetes development. (Lamson and Plaza, 2002).

It has been noted that peripheral blood mtDNA content is lower in the offspring of T2DM patients than in control families without a history of diabetes, indicating that mtDNA copy number may be inherited (Song *et al.*, 2001). MtDNA is vital for maintaining the mt function and energy demands of the body, and lower copy numbers could influence insulin resistance in a population.

The integrity of mtDNA is essential for maintaining the mt function and the energy demands of the body, but little attention has been paid to the quantitative aspects of mtDNA in diabetes (Song *et al.*, 2001). It has been shown that inhibition of OXPHOS in the pancreatic islets impairs insulin secretion (Yousufzai *et al.*, 1982). Therefore it is very likely that mt gene defects could decrease OXPHOS capacity in pancreatic β-cells resulting in an impaired glucose-induced insulin secretion in IGT or diabetic individuals (Katagiri *et al.*, 1994). It has been suggested that different alterations in the mt genome results in different diabetic phenotypes (van den Ouweland *et al.*, 1992).

2.5 GENETIC ASPECTS OF TYPE 2 DIABETES MELLITUS

Since the early 1960's it has been hypothesised that there is a genetic basis to diabetes in general (Rimoin, 1971; Kobberling, 1971; Simpson, 1962). The rapidly increasing occurrence of diabetes in the population "may represent a 'normal' evolutionary adaptation

to specific environmental conditions rather than a disease caused by a collection of 'bad' genes" (McClain, 2002). Therefore through evolution, populations have retained alleles containing mutations producing certain phenotypes. This is not to say the allele is flawed but just an aberration of the wild-type, a process that occurs in every organism without exception. This aberration can, however, have a negative effect on the organism due to changes in environmental stress factors, resulting in a negative phenotype for the individual.

Clustering of the disease in families indicates genetic influences in susceptibility towards the disease. This is supported in diabetes by high concordance in monozygotic twins, compared with dizygotic twins (Rewers and Hamman, 1995; Newman *et al.*, 1987). There is also wide variation in the prevalence among different ethnic groups i.e. Pima Indians that have the highest incidence in the world with 50% of their adult population being affected (Williams *et al.*, 2000). Information obtained from an investigation conducted on the Pima Indians concluded that diabetic parents are more likely to have diabetic children (Marx, 2002). In addition, conditions in the womb of diabetic mothers raise the chance of the offspring being diabetic (Marx, 2002). To date, a correlation between these two factors has not been found, and whether this has a direct or indirect effect is yet to be determined. A hypothesis known as the "thrifty phenotype" (Hales and Barker, 1992) could explain the observations presented that there is a negative correlation between birth weight and T2DM. The hypothesis outlines that the aforementioned observation could be due to foetal malnutrition as a consequence of an inadequate intrauterine environment. Therefore, as a consequence of insufficient nutrients to the foetus, more high-priority organs such as the brain and heart would receive preference for the utilisation of nutrients at the expense of organs such as the pancreas. This would cause the underdevelopment of the islets of Langerhans leading to glucose intolerance, and subsequently a T2DM phenotype (Hales and Barker, 1992).

2.5.1 The genetic basis of diabetes

Due to the polygenic nature of this disorder, it is vital to identify the candidate genes responsible for the observed diabetic phenotype. Various strategies have been applied to identify disease alleles. One strategy is linkage analysis (Lindgren and Hirschorn, 2001). An advantage of this method is that no prior knowledge is needed about which genes to test. However, this approach has limited power to detect modest genetic effects (Risch and Merikangas, 1996) generally associated with complex diseases such as T2DM.

Candidate gene studies utilise potential biological relevance to the disease to select genes of interest. The frequencies of each allele are compared in populations of cases and controls. An allele that is present at significantly higher frequency in affected individuals is deemed to be associated with the disease (Lindgren and Hirschorn, 2001).

Association studies could also be used. This has a potential advantage for studying multigenic diseases. The disadvantages are firstly, that the case-control studies can give false-positives if multiple ethnic groups with differing risks of disease are combined (Morton and Collins, 1998). Secondly, statistical criteria for significance have not been developed and widely accepted for association studies, especially with regard to multiple hypotheses testing (Altshuler *et al.*, 1998).

In the following section, several candidate nuclear genes have been highlighted as having an association with diabetes. The precise function of a variety of them has not been completely identified, yet linkage studies have demonstrated association between these genes and the different types of diabetic phenotypes.

2.5.2 Nuclear candidate genes associated with diabetes

The expression of a range of genes is regulated by the levels of glucose. There is difficulty in dissecting the contribution of the genetic component towards the disorder, as diabetes has proven to be highly polygenic.

Data are continuously being generated to show heterozygosity at several loci, as depicted in Table 2.1, (including β-cell transcription factors) can predispose an individual to T2DM (Gannon, 2001). Several candidate genes have been determined, although most are involved in either the production of insulin or the body's responses to it (Marx, 2002).

Table 2.1: Candidate genes for type 2 diabetes

Mutated Gene	Function	Effect	Associated with
HNF-4-α, HNF-1-β IPF-1, NeuroD1	Transcription factors	Down regulates insulin secretion	MODY (human)
HNF-1-α	Transcription factor	Down regulates insulin secretion	MODY
Glucokinase	Glucose metabolism	Down regulates insulin secretion	MODY
Calpain-10	Protease	Unknown	T2DM in Mexican and African Americans
PPAR-γ	Transcription factor	Down regulates insulin secretion	T2DM
Insulin receptor	Transmits signals into cell	Down regulates insulin sensitivity and secretion	Human diabetes (rare); mouse models
IRS1 and IRS2	Insulin signalling	Down regulates insulin sensitivity	Mouse models
Akt2	Insulin signalling	Down regulates insulin sensitivity	Mouse models
11- β-HSD	Glucocorticoid synthesis	Increases blood lipids, down regulates insulin sensitivity	Mouse models
UCP2	Down regulates ATP synthesis	Down regulates insulin secretion	Mouse models
Resistin	Fat cell "hormone"	Down regulates insulin sensitivity	Mouse models
Adiponectin	Fat cell "hormone"	Up regulates insulin sensitivity	Mouse, human studies

HNF = human necrosis factor, NeuroD1 = neuronal transcription factor, MODY = maturity onset diabetes of the young, PPAR = peroxisome proliferator-activated receptor, IRS = insulin receptor substrate, UCP = uncoupling protein, 11-β-HSB = glucocorticoid producing enzyme, IPF-1 = insulin promoter factor-1. Adapted from Marx, (2002).

Illustrations of the scope of genetic variation and continuous generation of results linking diabetes to novel sites within the genome make the task of identifying a singular cause for the disorder in a particular diabetic individual impossible. For instance, a gene that has received much attention over the last four years is one that was discovered by Horikawa *et al.*, in 2000 called calpain-10 (*CAPN 10*), that encodes for a cysteine protease, and has been linked to T2DM. The gene is located within a region known as the NIDDM 1 locus on chromosome 2 (Horikawa *et al.*, 2000). The encoded protein has been postulated to play a role in insulin secretion and action by inhibiting insulin's response to glucose. The protein could therefore have a role in glucose homeostasis (Cox, 2001). A variation found to be in the non-coding region (Cox, 2001) has been reported to possibly account for 25% of diabetes in African Americans (Marx, 2002). Variations in the *CAPN 10* gene are associated with significant reduction in the skeletal muscle *CAPN 10* messenger RNA (mRNA) levels as well as measures of insulin resistance in non-diabetic patients in certain populations (Baier *et al.*, 2000). Combinations of variants are more strongly associated with the disease rather than just individual polymorphisms (Cox, 2001).

2.5.3 Mitochondrial genetic component

Until recently, mutations in the mtDNA were thought to be rare and only associated mainly with the neuromuscular system (Wallace, 1992). However, a number of mutations have been associated with T2DM phenotype (Lamson and Plaza, 2002). To date, 42 mutations have been found in the mtDNA that can be associated with diabetes as a primary disorder as depicted in Table 2.2.

Table 2.2: Mitochondrial diabetic mutations

Gene	Position of the alteration
$tRNA^{Leu}$	3243, 3252, 3256, 3271, 3290, 3291, 12308
ND1	3316, 3348, 3394, 3396, 3423, 3434, 3438, 3447, 3480, 3483, 4216
ND2	4917
$tRNA^{Cys}$	5780
$tRNA^{Ser}$	7476
COXII	8245, 8251
$tRNA^{Lys}$	8344
ATPase 6	8993, 8860, 8894 (human)
ND3	10398
ND4	11778
$tRNA^{Glu}$	14709
$tRNA^{Thr}$	15904, 15924, 15927, 15928
D-loop	16069, 16093, 16126

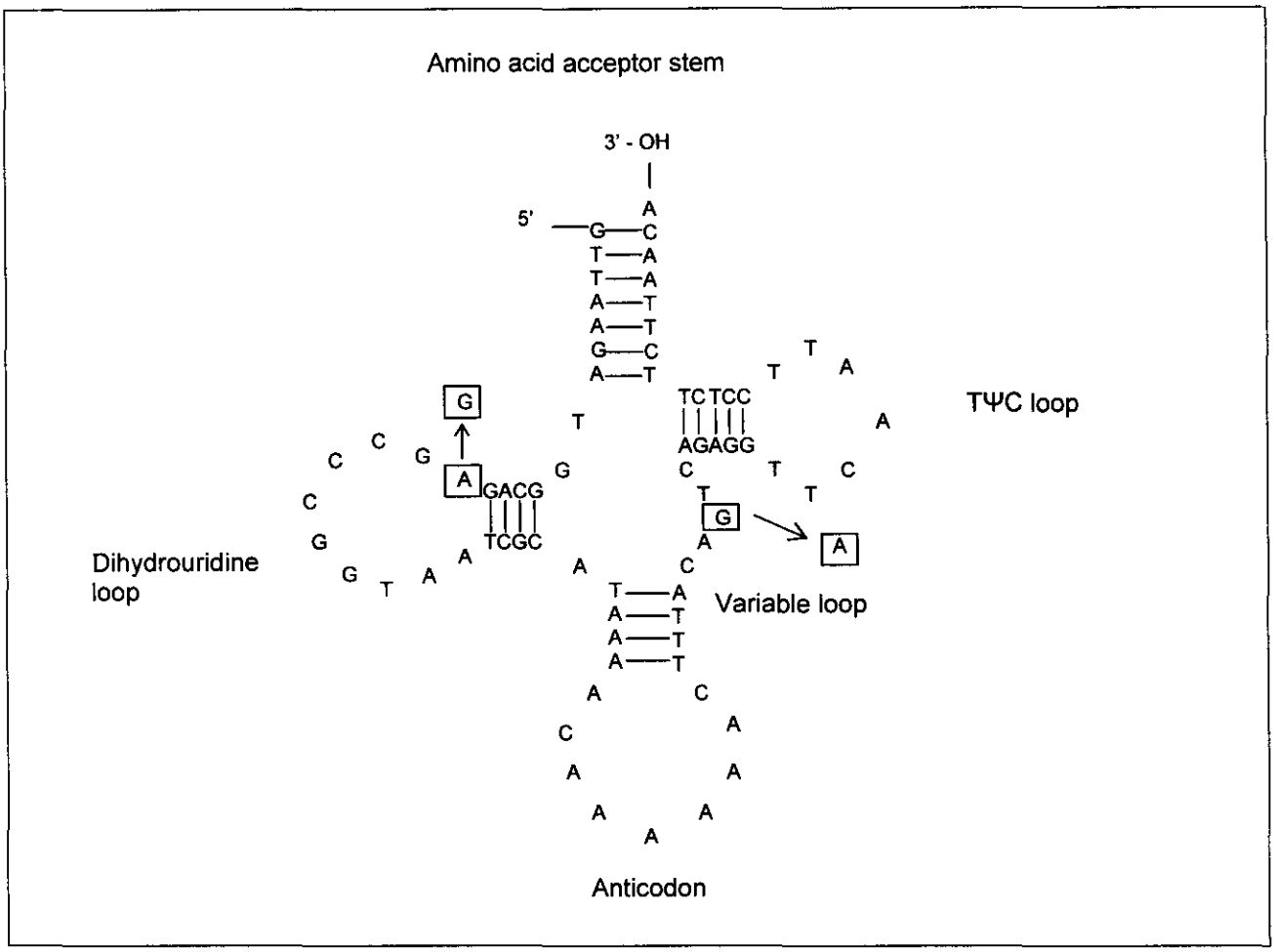
Mitochondrial nucleotide numbering is according to the Cambridge reference sequence. Adapted from Berdanier, (2001).

There are many mt disorders that have impaired glucose metabolism as a secondary complication, and all mt disorders have elevated lactate levels due to the inability of the mitochondria to metabolise pyruvate, the end product of glycolysis, which is in turn converted to lactate (Berdanier, 2001). Mutations in a number of genes i.e. those that regulate gene expression as well as structural genes or tRNAs, can cause mt related disorders, affect mt function and therefore alter tissue function. The most extensively investigated mt DNA mutation with an association to diabetes is the A3243G mutation.

2.5.3.1 A3243G mutation

The tRNA$^{Leu(UUR)}$ gene has been termed an 'etiological hotspot' for mutations due to the number of mutations located within this region as recorded in Table 2.2. The A3243G mutation is a point mutation resulting in a change of an adenine to guanine at position 3243 (Van den Ouweland *et al.*, 1992) within the tRNA$^{Leu(UUR)}$ gene which can be observed in Figure 2.8.

Figure 2.8: Cloverleaf structure of the tRNA$^{Leu(UUR)}$ molecule



Ψ = pseudouridine. Adapted from Van den Ouweland *et al.*, (1994).

This particular aberration renders the template sequence ineffective in transcription termination and shows a dramatic loss of functioning of the termination sequence (Hess *et al.*, 1991). This may cause defects in mt protein synthesis (Hess *et al.*, 1991). Therefore, the mutation causes significant dysfunction of one of the control regions in the mt genome. A major control sequence for the regulation of the relative amount of 16S and 12S rRNA synthesis versus the relative amounts of the other genes encoded by the mtDNA could be

altered as a consequence of the A3243G mutation (Hess *et al.*, 1991). This could impact on the proper processing of the $tRNA^{Leu(UUR)}$ molecule.

The A3243G mutation is also associated with MELAS, MERRF (myoclonic-epilepsy and ragged red fibres), Leigh syndrome and CPEO or chronic progressive external ophthalmoplegia (De Vries *et al.*, 1994). It is also a known cause of one variant of DM (Van den Ouweland *et al.*, 1992; Reardon *et al.*, 1992) when in a heteroplasmic state of 20-30% (Van den Ouweland *et al.*, 1994). This mt gene mutation is not always transmitted to all offspring of affected mothers. However, with successive generations, diabetes tends to manifest in affected individuals at an earlier age. This could be due to the fact that the mt genome with the mutation is selectively amplified in successive generations (Kobayashi *et al.*, 1992) and rapid segregation of mtDNA populations occurs (Hauswirth and Laipis, 1985) with a probability that the polymorphic mtDNA segregate unequally among siblings (Laipis *et al.*, 1988).

The presence of the A3243G mutation results in a subtype of diabetes with mitochondrially related complications (Suzuki *et al.*, 2003) including cardiomyopathy, cardiac conductance disorders, neuromuscular symptoms, macular pattern dystrophy neuropsychiatric disturbance and sensory hearing loss. This mutation accounts for 0.5-2.8% of the diabetic population, is maternally inherited and occurs in non-obese patients (Katagiri *et al.*, 1994). Anti-glutamic acid decarboxylase antibodies are absent, while clinical phenotypes of acute onset IDDM, and slow progressive IDDM or NIDDM are presented as a result of this mutation. Insulin therapy is required due to the progressive insulin secretory defect. Progressive neurosensory deafness (Van den Ouweland *et al.*, 1994) is shown along with advanced microvascular complications.

Late onset MELAS appears to be one of the characteristic complications in DM patients with the A3243G mutation. These mitochondria have impaired oxidative phosphorylation, which leads to reduced ATP synthesis. The cells with higher heteroplasmic mutations depend more strongly on ATP supplied by glycolysis. Therefore, carbohydrate deficiency during illness in diabetic patients with mutant mitochondria might cause ATP deprivation, leading to the stroke-like episodes (Suzuki *et al.*, 2003). If there is a gradual increase in the percentage of mutated mtDNA in the pancreatic β-cells due to the replicative advantage of the mtDNA with the mutation as discussed earlier in the section, there will also be a gradual decrease in insulin secretion capacity contributing to a diabetic phenotype (Yoneda *et al.*, 1992).

2.5.3.2 ATPase mutations

Many mutations that occur in the genes that encode for proteins associated with the OXPHOS pathway, compromise the process as well as the clinical condition of the patient affected. As mentioned earlier, an unhealthy diet containing large amounts of fatty acids can lead to the reduction of efficiency of the ATPase complex. Mutations within the genes that encode for the constituent proteins will also negatively affect the functioning of this highly essential pathway.

In the BHE/Cdb rat discussed in paragraph 2.4.2, a mutation was detected at the 8204 position which resulted in a base substitution of aspartic acid to asparagine, altering the charge ratio in the channel. A second mutation at position 8289 resulted in a threonine to serine substitution which altered the structure of the complex to the extent that it affected the mobility, lending a degree of rigidity to the complex, impairing its free movement and ATP synthesis. (Berdanier, 2001).

2.5.3.3 Other mitochondrial mutations associated with diabetes

An A to G transition at nt 5656 has been reported at a low frequency in patients with DM. It has been demonstrated that it alters a site of mt cleavage which could interfere with the tRNA processing. (Thomas *et al.*, 1996)

Another alteration identified by Towers *et al.,* in a Black South African diabetic patient cohort in 2001, as well as by Sternberg *et al.*, (2001) in a European Caucasian mitochondrial patient cohort, is located at position 3277 in the mt genome. This was detected in approximately 3% (n=100) of the Black South African diabetic patients sampled and only 0.006% (n=166) patients with mt myopathies in the European group. Neither of these investigations included a control group. Mutations are being detected at an accelerated rate due to the increasing accuracy and efficiency of the technology available today. Very little is currently known, however, about this multifactorial disease in Africa, which therefore warrants investigation.

2.6 SPECIFIC AIM OF THIS INVESTIGATION

The significant contribution made by the mitochondria on diagnosis of a vast number of disorders is increasingly being realised. The mitochondrion plays a larger role in diabetes than was previously suggested (Lamson and Plaza, 2002). The broad aim of this project is to elucidate the molecular aetiology and pathogenesis of T2DM in the Black South African population. The study presented in this thesis has focused on determining whether there is an association between mt alterations and T2DM, with particular emphasis on the A3243G and the G3277A alterations. This will determine whether these are alterations that could be associated with the diabetic disorder. The determination of these alterations as diabetic specific in this population could assist in the future with early diagnosis and treatment of diabetes in affected individuals.

The specific aim of this project was to determine if there is an association between the G3277A and A3243G alterations and the T2D phenotype, by screening diabetic and non-diabetic subjects.

CHAPTER THREE

MATERIALS AND METHODS

Ethical approval was obtained for this project from the Ethics Committee of the Potchefstroom University for Christian Higher Education (PUCHE) in 2002 and was assigned the number 02M08 under the title "Molecular analysis of non-insulin dependent diabetes mellitus (NIDDM) in the South African population".

3.1 STUDY DESIGN

This study forms part of a multi-disciplinary research programme, in which diabetes mellitus in the Black South African population is investigated. It incorporates resources and projects being undertaken at the Technical University of Dresden in Germany under the supervision of Prof P. Schwarz; the University of California in Irvine, California, USA under the supervision of Prof D. Wallace as well as the Department of Nutrition at the PUCHE. The molecular component for the study reported here was conducted at the Centre for Genome Research PUCHE.

A previous study was performed wherein a group of 100 female diabetic individuals from Pretoria, South Africa, were screened for the A3243G mutation. Another alteration at position 3277 whereby a guanine nucleotide was altered to an adenine nucleotide was detected in ca. 3% of the patient cohort (Towers, 2001). This current project focussed on elucidating the frequency of the alteration at position 3277 in the diabetic Black South African population in comparison to a non-diabetic control population to determine whether this alteration could be associated with DM in this particular group of individuals. Molecular techniques i.e. sequencing and restriction fragment length polymorphism (RFLP), were utilised to carry out this investigation and protocols were followed according to the manufacturers' specifications. Modifications to the above protocols are indicated throughout the text. To interpret the results obtained, statistical analyses were performed.

3.1.1 Patient and control population

The cohort of individuals included in the investigation consisted of Black South African diabetic and non-diabetic individuals. These individuals who were matched for age and fasting glucose level were screened for the G3277A alteration and, where possible, the A3243G causative mutation mentioned above. Informed consent was obtained from all subjects included in the project. Patients and controls were recruited from clinics around Pretoria, South Africa. In addition, control subjects were also recruited from Brits Community Hospital and from the "Profiles Of Obese Women with Insulin Resistance Syndrome" (POWIRS) study conducted at PUCHE. The control patients were collected from hypertension clinics. Hypertension may also be a complication of T2DM, therefore these control individuals were tested for diabetes. It was thereby established that the control individuals collected were truly non-diabetics.

The inclusion criteria for individuals in the patient group were:

- a fasting blood glucose level exceeding a value of 6.1 millimoles per litre ($mmol.l^{-1}$),
- previous diagnosis of DM, and
- the age of diagnosis with diabetes was above 25 years.

A fasting blood glucose level between 6.1 $mmol.l^{-1}$ and 7.8 $mmol.l^{-1}$ is only an indication of IGT. Patients with a value higher than 7.8 $mmol.l^{-1}$ is an indication of complete DM. However, IGT individuals were also included in the patient group for this study. To ensure a clear distinction between patient and control individuals fasting blood glucose levels were also obtained for the control group. These values were all below the exclusion level of 6.1 $mmol.l^{-1}$.

In addition to the above criteria, additional information was collected via a detailed questionnaire that included:

- Body Mass Index (BMI)
- Waist Hip ratio (WHR)
- Daily lifestyle patterns
- Previous medical treatments
- Family history in relation to diabetes and hypertension.

This information will be of assistance for stratification purposes when a larger sample size of individuals is collected.

3.2 METHODS EMPLOYED FOR THIS STUDY

Various methods were undertaken to achieve the aims outlined in section 2.6. The techniques involved in each of these are discussed in the following paragraphs.

3.2.1 DNA isolation

Approximately 9.0 millilitres (ml) of whole blood were obtained from each patient in ethylenediamine tetra-acetic acid (EDTA) tubes. Collection in these tubes prevented clotting of the blood as well as preserving the integrity of the DNA. Genomic DNA (gDNA) was isolated from 3 ml of each sample utilising the Nucleospin®[1] (paragraph 3.2.1.1) or the Promega Wizard®[2] (paragraph 3.2.1.2) gDNA purification kit. The two different procedures were utilised due to the availability of the kits mentioned. The DNA isolated that was not utilised, was stored at -70°C to protect against degradation.

3.2.1.1 Genomic DNA purification

The protocol for isolating DNA utilising the Nucleospin® kit is as follows: Three ml of lysis buffer was added per 3 ml of whole blood sample in a sterile 15 ml centrifuge tube. Double distilled water (ddH_2O) was added to the blood solution to make up a total volume of 15 ml. To separate the white blood cells from the lysed red blood cells, the solution was centrifuged at 1,300 x gravitational acceleration (x g) in an Eppendorf®[3] 5810 R centrifuge utilising a fixed angle rotor F-45-30-11 for 15 minutes (min) at 4°C. The white blood cell pellet was washed with phosphate buffer saline (PBS) to elute any remaining lysis buffer and re-establish the leukocytes by vortexing the mixture for one min. The mixture was poured into an Eppendorf® tube and 400 microlitres (µl) of the nuclei lysis buffer to lyse the white blood cells was added along with 50 µl Proteinase K to denature any proteins in the solution. The solution was vortexed until the pellet had completely dissolved after which the solution was incubated at 70°C in a water bath for 10 min. The DNA was precipitated by the addition of 420 µl of 100% ethanol and vortexing. To collect the DNA, the sample was placed through a Nucleospin® column and centrifuged at 10,416 x g for two min until the entire solution had been centrifuged through the column and the column had dried.

[1]Nucleospin® is a registered trade mark of Macherey-Nagel, Germany.
[2]Wizard® is a registered trade mark of the Promega Corporation, Madison, WI, USA.
[3]Eppendorf® is a registered trade mark of Eppendorf, Hamburg, Germany.

This allowed the DNA to adhere to the column while the other cellular components were removed in the remaining solution.

The DNA was washed by the addition of 700 µl of 70% ethanol to the spin tube and centrifuged at 10,416 x g for one min. An additional washing step was undertaken by the addition of 300 µl of ethanol after which the DNA was centrifuged at 15,000 x g for two min. To collect the DNA from the column, 200 µl of elution buffer that was incubated at 70°C was added. The column containing the elution buffer was incubated at 70°C for 15 min to allow the DNA to be released from the column. The DNA was collected by centrifugation at 10,416 x g for two min. The elution step was repeated to maximise the yield of DNA obtained from the column. Due to the low yield of DNA obtained for some samples using the above mentioned kit, the Promega Wizard® DNA isolation kit was used as an alternative strategy.

3.2.1.2 Alternative method of genomic DNA purification

Three ml of whole blood was utilised in DNA isolation using the Promega Wizard® kit. Nine ml of cell lysis solution was added into a sterile 15 ml centrifuge tube, followed by the addition of the blood sample. Red blood cell lysis was obtained by mixing the solution. The cellular mixture was incubated at room temperature for 10 min to allow complete cell lysis, after which the leucocytes were collected by centrifugation at 2,000 x g for 10 min. Subsequent to centrifugation the supernatant was discarded and the pellet vortexed to resuspend the cells in the remaining solution. Three ml of nuclei lysis solution was added to the pellet to lyse the white blood cells. Fifteen µl of the protein precipitation solution was subsequently added to denature the proteins. To collect the protein and cellular debris the solution was centrifuged at 2,000 x g for 10 min, after which the supernatant containing the DNA was transferred into 3 ml of isopropanol and mixed until white threads of DNA became visible. These DNA threads were recovered by centrifugation at 2,000 x g for one min. The isopropanol was discarded and 70% ethanol was added to wash the DNA pellet. To facilitate air drying of the DNA, the ethanol was discarded and the tube containing the DNA left on the lab bench for 15 min after which 250 µl DNA rehydration solution [10 millimolar (mM) Tris-HCl (pH 7.4), 1 mM EDTA) was added. In order to ensure efficient rehydration of DNA, the sample was incubated at 65°C for one hour. The DNA-containing solution was transferred into an Eppendorf® tube and stored at 4°C.

3.2.1.3 DNA quantification

The quantity of DNA per isolation was determined spectrophotometrically using the Eppendorf® Biophotometer at an optical density (OD) of 260 nanometers (nm) in combination with Equation 3.1 as described by Sambrook *et al.*, 1989. Working dilutions of 50 ng.μl^{-1} of each sample were prepared subsequent to determining the concentration of the stock solutions. The stock solutions were stored at -20°C and the working dilutions at 4°C.

Equation 3.1: Calculation of the DNA concentration from the optical density value at 260 nm

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

Adapted from Sambrook *et al.*, (1989).

3.2.2 Amplification of DNA template

Two methods of analysis were utilised through the duration of this project i.e. automated sequencing and RFLP, to detect for the G3277A alteration and the A3243G mutation. Before either of these techniques could be carried out, the DNA region under investigation harbouring these changes was amplified, in order to increase the concentration of the aforementioned region.

3.2.2.1 Analysis of the $tRNA^{Leu(UUR)}$ region within the mitochondrial genome

The region of genome that harbours the above-mentioned changes is that of the mitochondrial genome between bases 3007 and 3370, a region of 364 bases in length. The sequence illustrated in Table 3.1 encompasses the region of interest.

Table 3.1: Partial genomic DNA sequence of the mitochondrial tRNA$^{Leu(UUR)}$ gene, indicating nucleotides 3001 to 3600

Nucleotide number	DNA sequence
3001	ggacatcccg atggtgcagc cgctattaaa ggttcgtttg ttcaacgatt aaagtcctac
3061	gtgatctgag ttcagaccgg agtaatccag gtcggtttct atctaccttc aaattcctcc
3121	ctgtacgaaa ggacaagaga ataaggcct acttcacaaa gcgccttccc ccgtaaatga
3181	tatcatctca acttagtatt atacccacac ccacccaaga acagggtttg ttaagatggc
3241	agaggcccggt aatcgcataa aacttaaaac tttacagtca gaggttcaat tcctcttctt
3301	aacaacatac ccatggccaa cctcctactc ctcattgtac ccattctaat cgcaatggca
3361	ttcctaatgc ttaccgaacg aaaaattcta ggctatatac aactacgcaa aggccccaac
3421	gttgtaggcc cctacgggct actacaaccc ttcgctgacg ccataaaact cttcaccaaa
3481	gagcccctaa aacccgccac atctaccatc accctctaca tcaccgcccc gaccttagct
3541	ctcaccatcg ctcttctact atgaaccccc ctccccatac ccaaccccct ggtcaacctc

The forward primer sequence is indicated by the double underlined text (xx), while the reverse primer is indicated by the single underlined text (xx). The 3243 position is indicated by the circle. The 3277 position is indicated by an open block. The yellow region indicates the recognition site for restriction enzyme *Taa* I, discussed in paragraph 3.2.4. Sequence adapted from MITOMAP, (2003).

3.2.2.2 Primers

Before sequencing or RFLP could be performed, the fragment of 364 bp from the tRNA$^{Leu(UUR)}$ gene of the mitochondrial genome was amplified via the polymerase chain reaction (PCR). The primers utilised are listed in Table 3.2. (Wallace, 1999).

Table 3.2: Primer sequences used for the amplification of the mitochondrial nucleotide region 3001 to 3400

Primer	Primer sequence	Product size	Calculated T_m
NDI-FOR	**F: 5'**-ccc gat ggt gca gcc gc-**3'**	364 bp	60°C
3.3-REV	**R: 5'**-gca tta gga atg cca ttg cg-**3'**		60°C

Primer sequences obtained from Wallace, (1999).

Primers utilised for this screening experiment were optimised by changing certain parameters i.e. the annealing temperature and the number of cycles for amplification. The primer annealing temperature (T_a) for the PCR reaction was determined from the melting temperature (T_m) as calculated in Equation 3.2. The annealing temperature was optimised experimentally by utilising gel electrophoresis to observe the quality of product amplified to determine the optimal temperature required to obtain the minimal levels of secondary product as possible.

Equation 3.2: Calculation of the annealing temperature

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

A = adenine; C = cytosine, G = guanine, T = thymine. Adapted from Thein and Wallace, (1986).

3.2.2.3 Amplification of DNA via the polymerase chain reaction

In order to perform the sequencing reaction, a concentration of the DNA template was required that is between three to ten ng for a template length that ranges from 200 to 500 bp. This was obtained via PCR by first amplifying the region of interest. A PCR master mixture consisted of: 1 X PCR buffer [50 mM potassium chloride (KCl), 10mM Tris-HCl (pH 9.0), 0.1% Triton X-100®[1]], 1.5 mM magnesium chloride ($MgCl_2$), 1 U of Taq DNA polymerase (Promega), 10 micromolar (µM) forward and reverse primer (as indicated in Table 3.2), 200 µM each of 2'-deoxyadenosine-5'-triphosphate (dATP), 2'-deoxythymidine-5'-triphosphate (dTTP), 2'-deoxycytidine-5'-triphosphate (dCTP), 2'-deoxyguanosine-5'-triphosphate (dGTP) and ddH_2O. The master mixture was aliquoted into separate PCR tubes to which 100 ng of gDNA was added to make up a total volume of 12.5 µl. The reaction mix was finally overlaid with mineral oil to prevent sample evaporation. A positive and negative control that utilised control DNA and ddH_2O respectively was included in the reaction to make up the final volume indicated. The positive control was included to monitor the success of the reaction. The negative control was utilised to monitor that none of the utilised reagents were contaminated with foreign DNA.

A standard PCR programme was utilised to amplify the target region and was achieved utilising a Thermo Hybaid® MultiBlock System 0.5 µl Satellite (MBS 0.5S) thermal cycler. The programme involved:

- An initial 10 min denaturation step at 94°C;
- 35 cycles of
  - 30 sec denaturation at 94°C;
  - 30 sec annealing at 55°C;
  - 45 sec extension period at 72°C;
- Followed by a 7 min extension step at 72°C;
- A final indefinite hold step at 4°C.

[1] Triton® is a registered trade mark of Rohm & Haas Company, Philadelphia, PA, USA.

The success of the amplification procedure was determined via 2% agarose gel electrophoresis as described in section 3.2.5.1. The images of the fragments visualised under the ultraviolet (UV) transilluminator were captured as electronic copies by utilising a video documentation system.

Following the amplification of the target region to be sequenced, the PCR product was purified to remove any contaminating reagents such as 2'-deoxynucleotide triphosphates (dNTPs), primers, enzymes and salts, from the prior reaction. This was achieved utilising the QIAquick®[1] PCR purification kit according to the manufacturer's instructions.

Five volumes of the binding or PB buffer were added to every 1 volume of PCR product sample to bind the DNA to the silica matrix. The mixture was briefly vortexed and added to the QIAquick column, after which the columns were placed in the centrifuge for 60 sec at 17,900 x g. The supernatant was discarded and 750 µl of wash or PE buffer was added. The column was again centrifuged for 60 sec at 17,900 x g. To ensure the complete removal of the wash buffer, the supernatant was discarded and the column again centrifuged for the same time period and at the same speed as before. To elute the DNA from the column, the column was placed in a 1.5 microcentrifuge tube, 50 µl of elution buffer (10 mM Tris-HCl, pH 8.5) added and the column centrifuged for a final 60 sec at 17,900 x g.

Two µl of the purified product was electrophoresed via agarose gel electrophoresis, as described in paragraph 3.2.5.1, on a 2% agarose gel to ascertain the efficiency of the purification procedure. Ten µl of the sample was further added to 90 µl of ddH_20 and quantified by making use of UV spectrophotometry. This provided an indication of the quantity of sample required to add to the sequencing sample mixture.

3.2.3 Cycle sequencing

The most cost effective strategy was sought to determine the presence of the 3277 and 3243 alterations. It was decided that automated sequencing would be optimal in providing conclusive results. This method of direct sequencing is based on the principle of dideoxy chain termination developed by Sanger *et al.*, (1977). The ABI Prism®[2] BigDye terminator version 3.1 Ready Reaction Cycle Sequencing Kit was utilised. Amplification of the target

[1] QIAquick® is a registered trade mark of QIAGEN Pty. Ltd, Victoria, Australia.
[2] ABI is a registered trade mark of Applera Corporation, Foster City, CA, USA.

sequence was made possible by the use of the sequencing enzyme AmpliTaq®[1] DNA polymerase, FS. Detection of the sequence was facilitated by the differential labelling of the four 2',3'-dideoxynucleotide triphosphates (ddNTPs) with various fluorophores.

3.2.3.1 DNA sequence determination

Once the region of interest had been amplified, purified and quantified, a sequencing reaction was set up. All the required components were supplied in the ABI PRISM® sequencing kit mentioned in paragraph 3.2.3. The cycle sequencing reaction consisted of the following: 2 µl of the BigDye™ 5 X Sequencing Buffer (Tris-HCl, pH 9.0 and $MgCl_2$), 2 µl of the BigDye™ Ready Reaction Premix which contains all the dye terminators, dNTPs, AmpliTaq® DNA polymerase FS and $MgCl_2$, 3.2 picomoles (pmol) of the selected primer, and 10-20 ng of the purified PCR product. A final reaction volume of 10 µl was utilised and the difference in volume was made up utilising ddH_2O. This reaction mixture was subjected to the following programme as depicted in Table 3.3, in order to amplify the sequencing product in the Thermo Hybaid® MultiBlock System.

Table 3.3: Cycle sequencing PCR programme

Step	Cycles	Temperature	Length of time	Purpose
1	25	96°C	10 sec	Denaturation
		50°C	5 sec	Annealing
		60°C	4 min	Extension
2	1	4°C	Indefinite	Hold

3.2.3.2 Product precipitation

Subsequent to the sequencing programme, the fragments obtained during amplification were precipitated, removing the unincorporated BigDye™ terminators from the extended fragments. The strategy was a sodium acetate (NaOAc)/ethanol precipitation and involved the following components: the sequencing product was added to 62.5 µl of 99.8% ethanol, 3 µl of 3 M NaOAc solution and 14.5 µl of ddH_2O. The final volume of 90 µl that included the sequencing product was centrifuged at 13,000 x g for 20 min in the Eppendorf® 5810R centrifuge. The supernatant was discarded and 250 µl of 70% ethanol was added to wash

[1] AmpliTaq DNA Polymerase FS is a registered trade mark of Roche Molecular Systems Inc., Pleasanton, CA, USA.

the sample. The sample was centrifuged at 13,000 x g for 10 min, followed by the removal of the residual ethanol and allowed to air dry for 30 min.

In order to visualise the sequence, the precipitate was resuspended in 6 µl Hi-Di™[1] formamide solution. Three µl was electrophoresed on a SpectruMedix™[2] (SCE2410) Genetic analysis system. Sequence evaluation and alignment was performed via the utilisation of the freeware programme BioEdit Sequence Alignment Editor Version 5.0.9 (BIOEDIT, 2003).

3.2.4 Restriction fragment length polymorphism

An alternative strategy was sought besides that mentioned in section 3.2.3 in order to allow for a direct method of alteration detection. This method is more time-efficient as compared to sequencing. The method of RFLP[3] is not very informative in that it does not provide data about the region under analysis other than about the nature of the alteration.

The restriction enzyme, *Taa* I, which was utilised in these experiments, digests genomic DNA at a specific sequence that is defined as follows: 5' ACN/GT 3'. This sequence matches the sequence surrounding the wild-type sequence of the 3277 position in the mitochondrial genome that is demonstrated in Table 3.1. This, therefore, indicates that the enzyme, *Taa* I would digest the PCR amplified DNA segment into two fragments if the wild-type sequence was present. The sizes of the two fragments to be generated were calculated to be 271 bp and 93 bp. If the G3277A alteration was present at the 3277 position, an A would be present in the place of the G at the 3' end of the digestion site. The enzyme would no longer recognise this sequence and would, therefore, not digest it. The size of the fragment expected to be observed if the alteration is present is 364 bp.

The region of interest that was amplified and purified, as described in paragraph 3.2.2.1, was digested with the *Taa* I restriction enzyme. Ten µl of gDNA was placed in a 1.5 ml reaction tube containing 5 U of enzyme, one tenth of the total reaction volume of Buffer Y$^+$/Tango™[4] [33 mM Tris-acetate, 10 mM magnesium acetate, 66 mM potassium acetate and 0.1 mg.ml^{-1} bovine serum albumin or BSA (pH 7.9 at 37°C)] and 16.5 µl of ddH_2O to a total volume of 30 µl. The reaction mixture was incubated in a water bath at 65°C

[1] Hi-Di™ is a trade mark of the Applera Corporation, Foster City, CA, USA.
[2] SpectruMedix™ is a trade mark of SpectruMedix LLC., State College, PA, USA.
[3] *Taa* I RFLP strategy was developed by B.Semete in 2003.
[4] Y$^+$/Tango is the trade mark of Fermentas, Hanover, MD, USA.

overnight. Fifteen µl of the product was electrophoresed on a 2% agarose gel as described in paragraph 3.2.5.1 to determine the presence of the G3277A alteration which was later analysed under UV light.

3.2.5 Gel electrophoresis

It was often necessary to analyse the size of the PCR product as well as its purity. The most efficient and cost-effective means to achieve this form of analysis was via gel electrophoresis.

3.2.5.1 Agarose gel electrophoresis

DNA fragments obtained after amplification with PCR and RFLP were separated on agarose gels. These gels were utilised to analyse fragments obtained to determine the success of the reaction and whether the expected size fragment was amplified. PCR products were analysed on two % midi or mini gels. The mini gel which had a total volume of 25 ml and was composed of 0.5 grams (g) molecular grade agarose (Promega), 1 X TBE buffer [89.15 mM Tris® (pH 8.1), 88.95 mM boric acid, 2.5 mM Na_2EDTA] and 0.5 $\mu g.ml^{-1}$ ethidium bromide or EtBr (Sigma®). The midi gel was made up to a volume of 100 ml, utilising 2 g of agarose, 1 X TBE buffer and 0.5 $\mu g.ml^{-1}$ EtBr.

Two µl of sample was loaded into the wells with 1 µl of 6 X Blue/Orange Loading Dye (15% Ficoll®[1] 400, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 0.4% orange G, 10 mM Tris-HCl (pH 7.5) and 50 mM EDTA). A 100 bp DNA marker (Promega) was utilised as a reference to which the sizes of products could be compared. The gels were subsequently electrophoresed in 1 X TBE buffer at 12.5 volts per centimetre (cm) for 30 min. To visualise fragments, an UV transilluminator was utilised.

3.2.5.2. Non-denaturing polyacrylamide gel electrophoresis

Polyacrylamide gels are utilised in order to obtain clearer resolution than agarose gels, and are the choice of gel when visualising small fragments. In particular, they allow for smaller fragments i.e. less than 100 bp in size, to be visualised more clearly than on an

[1] Ficoll is as trade mark of Pharmacia Biotech AB, Piscataway, NJ, USA.

agarose gel. Twelve % gels were utilised in experiments that required the use of polyacrylamide gel electrophoresis (PAGE). The reasoning for utilisation of this percentage of gel was due to the fact that they were the gels in use at the time, which were being shared between different experiments done by different individuals. Ideally, a five % gel could have been run for two hours. A 40 ml solution was prepared which consisted of 12 ml of a 40 % stock solution of unpolymerised polyacrylamide [acylamide (Stratagene) and bis-acrylamide (Promega) in a ratio of 19:1], four ml of 10 X TBE and ddH_2O to make up 40 ml. To this 40 µl of TEMED was added, which acts as a catalyst for the polymerising procedure, and 350 µl of 10 % ammonium persulphate (APS) which acts as an oxidising agent for the polymerisation reaction.

After an hour, once the gel had set, the samples were loaded and electrophoresed for six hours in 0.5 X TBE at 250 volts. The gel apparatus was disassembled and the gel stained in SYBR®[1] Gold (five µl SYBR® Gold and 500 µl 1 X TBE) for 30 min. The fragments were visualised under UV light utilising an UV transilluminator, after which an electronic copy of the image was captured for documentation purposes.

3.3 STATISTICAL ANALYSIS

In order to determine whether the G3277A alteration or the A3243G mutation was associated with DM in the Black South African diabetic population, statistical analyses were performed on the generated data. The main aim for utilising this form of analysis was to determine if there was a statistically significant difference between the distributions of the frequencies of the investigated alterations in the diabetic population as compared to the control/non-diabetic population. It would not be possible to make this inference without the use of statistics, as molecular techniques alone cannot demonstrate if an association can be made between two entities.

The ideal sample size that would be required in this study in order to obtain statistically significant results for the prevalence of the alteration within the diabetic versus the non-diabetic Black South African population was determined. The equation that was utilised to calculate the relevant sample size is stated below in Equation 3.3.

[1] SYBR® is a registered trade mark of Molecular, Probes, Inc., Leiden, The Netherlands.

Equation 3.3: Sample size estimation

$$SE_{\pi} = \sqrt{\frac{\pi (1 - \pi)}{n}}$$

SE = standard error of proportion, π = sample proportion, n = population size. Adapted from Samuels, (1989).

The π symbol represented the proportion of the population that were expected to harbour the G3277A alteration. The standard error (SE) was 0.01, representing a confidence interval of 99% against a type I error i.e. rejecting a true hypothesis. The results obtained from the number of individuals that were recruited were analysed according to the following outlined analyses.

3.3.1 Statistical analysis of the G3277A alteration

A null hypothesis (H_o) and alternative hypothesis (H_A) were generated. These are outlined in Table 3.4.

Table 3.4: Null and alternative hypothesis for association between alteration and diabetic phenotype

H_o:	The genotypic frequency of the investigated alteration does not differ between the diabetic and non-diabetic population. Therefore the alteration is not associated with the T2DM phenotype.
H_A:	The genotypic frequency of the investigated alteration does differ between the diabetic and non-diabetic population, suggesting that this alteration is associated with the T2DM phenotype.

In order to begin investigating the above hypotheses, the frequencies of the G3277A alteration and A3243G mutation in the investigated populations needed to be calculated. The results of these calculations were interpreted via utilising the stated hypotheses.

3.3.1.1 Hardy-Weinberg equilibrium

When the frequency of an allele (or alteration for the purpose of this research) is being investigated within two populations i.e. a patient/diabetic population in comparison to a control/non-diabetic population, it is necessary to determine whether the two populations

are in Hardy-Weinberg (H-W) Equilibrium. This allows certain assumptions to be made about these populations as well as allowing them to be compared since the alteration investigated is in the same equilibrium status within both diabetic and non-diabetic populations. The assumptions made are that an infinitely large population is being investigated and that there is no inbreeding, migration or mutation, and the allelic and genotypic frequencies will not change from one generation to the next (Last, 2001). A second H_0 and H_A was therefore generated. These are stipulated in Table 3.5.

Table 3.5: Null and alternative hypothesis for Hardy-Weinberg equilibrium status of diabetic and non-diabetic population

H_o:	The frequency of the specific alteration under investigation in the Black South African diabetic as well as the non-diabetic population is in H-W equilibrium in the two populations separately.
H_A:	The frequency of the specific alteration under investigation in the two aforementioned populations is not in H-W equilibrium within the two separate populations.

In order to apply statistical analyses to test for the above hypotheses, observed and expected genotypic frequencies of the alterations had to be calculated. The allelic frequency of the most common allele was denoted with 'p', and that of the alteration with 'q' i.e.

- The frequency of 'G' at position 3277 = p_{3277},
- The frequency of 'A' alteration at position 3277 = q_{3277},

The allelic frequencies for the wild-type sequence were calculated by utilising Equation 3.4.

Equation 3.4: Allelic frequency of the most common allele

$$p = \frac{[\text{number of homozygotes} + \frac{1}{2}(\text{number of heterozygotes})]}{\text{Total number of individuals}}$$

p = frequency of wild-type sequence in the investigated population. Adapted from Tamarin, (1996).

The value for the frequency of the alteration in the population could have been calculated according to Equation 3.4. However, as the total of the calculated frequencies should always equate to one, the frequency of the sequence containing the alteration was calculated via Equation 3.5.

Equation 3.5: Allelic frequency of alteration of interest

$$q = 1 - p$$

q = frequency of sequence with alterations of interest in the investigated population. Adapted from Tamarin, (1996).

To calculate whether the population was in H-W equilibrium the allelic frequencies were needed to calculate expected genotypic frequencies. The expected genotypic frequencies are the frequencies in which the alterations should have been found, in order for the alterations investigated to be in H-W equilibrium. These expected genotypic frequencies were converted into the whole number of individuals that they would equate to in the analysed population. These numbers were then compared to the actual observed numbers of individuals with and without the alterations.

As the alterations that were being investigated are found within the mitochondrial genome, it is necessary to clarify to what the genotypic frequencies are referring. For the purpose of this investigation, individuals could not be referred to as having separate loci that possess the alteration or not, but rather that they were homoplasmic for the alteration or the wild-type sequence. Those individuals that possessed mitochondrial genomes with both the wild-type and the alteration were referred to as heteroplasmic. Homoplasmic individuals were treated as homozygous individuals for a nuclear genome locus in order to simplify the calculation for the genotypic frequencies, and heteroplasmic individuals were treated as heterozygous individuals at a nuclear genome locus.

The probability of an individual being homoplasmic was denoted as f(1,1) for possessing only the wild-type sequence, f(2,2) for possessing only the alteration and f(1,2) for being heteroplasmic. Therefore, the expected genotypic frequencies were calculated from the values calculated for p and q as presented in Equation 3.6.

Equation 3.6: Expected genotypic frequency calculation

$$f(1,1) = p \times p = p^2$$
$$f(1,2) = 2pq$$
$$f(2,2) = q \times q = q^2$$

p = allelic frequency of wild-type sequence in the investigated population; q = allelic frequency of sequence with alterations of interest in the investigated population; f(1,1) = probability of homoplasmic wild-type genotype; f(2,2) = probability of homoplasmic genotype for the alteration of interest; f(1,2) = probability of a heteroplasmic genotype. Adapted from Tamarin, (1996).

From the calculated expected genotypic frequencies, the expected numbers of individuals with the respective genotypes could be calculated and compared to the observed numbers

of individuals with the same genotypes via the use of contingency table analysis. The method of comparison used was the chi-squared (χ^2) test, illustrated in Equation 3.7, which is one of the most common tests for detecting whether two or more population distributions differ significantly from one another (Last, 2001). In this example, the population distributions are referring to the distributions of the genotypic frequencies. The χ^2 value was obtained by squaring the difference between the observed number of individuals for a particular genotype and the expected number for the same genotype, after which the value was divided by the expected number of individuals for that genotype.

Equation 3.7: Chi square test

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

χ^2 = chi square value; O = frequency of the number of individuals observed with the G3277A alteration; E = calculated frequency of individuals expected to have the G3277A alteration. Adapted from Samuels, (1989).

The value that was obtained represents the significance of the difference between the observed and expected numbers. This value, however, is meaningless unless compared against a standard of chi-square values known as critical values. For this comparison, a confidence interval and degrees of freedom had to be determined. The confidence interval represents the value that determines whether to reject the H_o, and a 95% confidence interval is generally utilised. Therefore, there is a 5% probability of the χ^2 value being greater than the critical value by chance alone. The degrees of freedom were calculated by referring to the contingency table and multiplying the number of row categories minus one by one less than the number of column categories (Last, 2001). This produced a value of two degrees of freedom, therefore the χ^2 value was compared to the critical value of 5.991.

3.3.1.2 Comparison of the G3277A alteration between the patient and control population

Once H-W equilibrium was determined, it was then possible to determine whether the alteration investigated was observed more frequently in the T2DM affected subset of the population or not. The χ^2 equation was again utilised to compare the diabetic and non-diabetic population subsets and determine if there was a significant difference between the frequencies observed in the two populations. This would allow the conclusion

to made about whether or not the alteration could be considered associated with the T2D phenotype, to be made.

3.3.2 Statistical analysis of the A3243G alteration

The A3243G mutation was the second mutation investigated in this project. The results from the analyses for this alteration were subjected to the same statistical analysis as outlined for the G3277A alteration.

CHAPTER FOUR

RESULTS AND DISCUSSION

Owing to the polygenic nature of the diabetic disorder, linking any genetic alteration to a particular population group can significantly increase the diagnostic power of the current methods to analyse the diabetic phenotype. Therefore the type of research that has been undertaken in this molecular investigation is important in order to obtain a clear indication of the various alterations in different populations and whether these could be associated with the development of the diabetic phenotype.

The results obtained were to determine the incidence of the A3243G and G3277A alteration in 379 Black South African individuals. This group consisted of a combination of diabetic and non-diabetic individuals. The results from the detection of the A3243G alteration would further strengthen the investigation performed in 2001 by Towers *et al.*, to detect the presence of the A3243G alteration in the South African Black diabetic population. Since the aforementioned investigation was only carried out on 100 diabetic patients, results from the current investigation would add value to those obtained in 2001. This chapter is therefore devoted to discussing the outcomes of the methods outlined in Chapter three, which were utilised to attain the aims of the study presented here.

4.1 PATIENT AND CONTROL POPULATION

In total, 142 diabetic patients were studied along with 237 control individuals. This makes the total number of patients analysed for the G3277A alteration 222 which includes 80 diabetic patients from the cohort of 100 patients included in the previous investigation (Towers *et al.*, 2001). Of the original 100 individuals examined during the previous study in 2001, 20 were resampled during the second study reported here and for this reason only 80 individuals of the original study could be added to the current investigation.

4.2 RESULTS OBTAINED WITH SPECIFIC METHODS

In the following sections the success of the techniques utilised through the duration of the study was analysed. Optimisation procedures are briefly outlined where instances arose that required optimisation.

4.2.1 DNA isolation

All patient and control DNA samples were obtained and isolated as described in paragraph 3.2.1. When the yields of the samples isolated from the Wizard® kit were compared to those samples isolated via the use of the Nucleospin® kit it was observed that, in general, the Wizard® kit yielded higher DNA concentrations. The procedure of the Wizard® kit also proved to be more time-effective than that of the Nucleospin® kit. Therefore in future, in view of the time-efficiency and higher yield production, if it were necessary to select a kit for use in DNA isolation, the Wizard® kit would be recommended.

As determined by the A_{260}/A_{280} ratio the gDNA obtained utilising either of the kits mentioned in the paragraph above was of the highest purity and the concentrations obtained were sufficient to perform sequence and RFLP analysis. The DNA concentrations of all the samples were determined as described in paragraph 3.2.1.3 and a range of values between 23.3 $ng.\mu l^{-1}$ and 2490.8 $ng.\mu l^{-1}$ was obtained. The average yield of gDNA was calculated to be 250.5 $ng.\mu l^{-1}$.

Two methods of analysis were utilised to determine whether the 222 diabetic and 237 non-diabetic individuals carried the A3243G and the G3277A alterations. The success of each method is outlined in the following sections.

4.2.2 Polymerase chain reaction

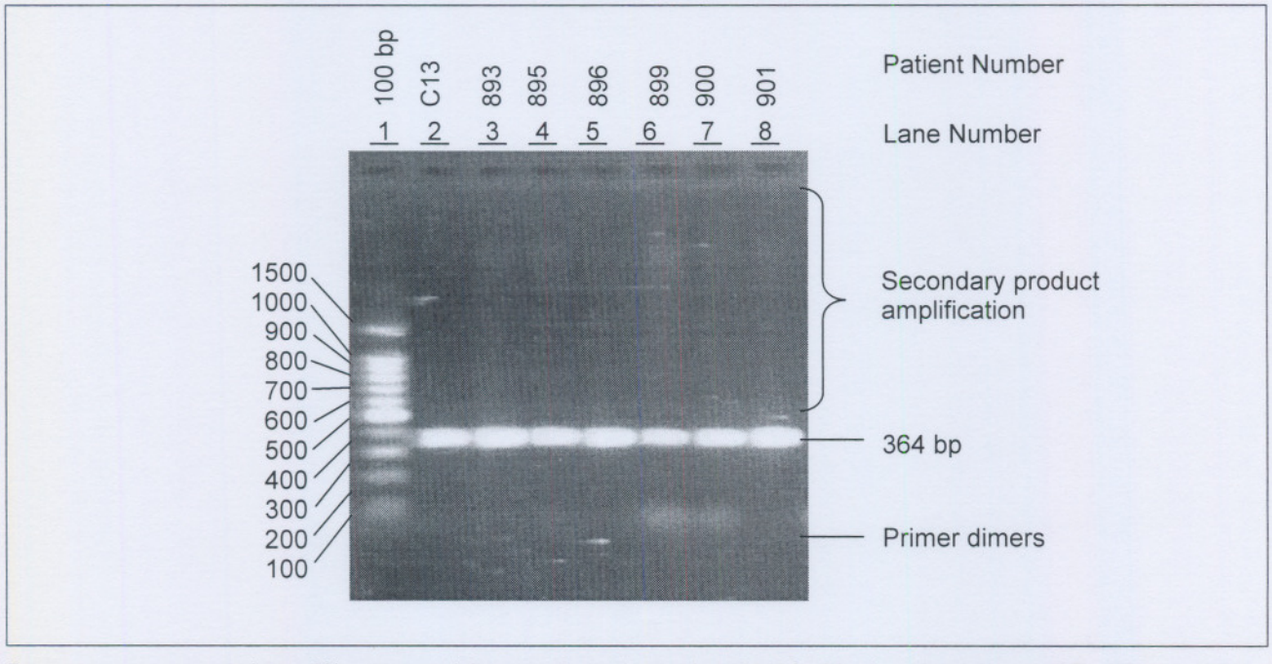
The method utilised to amplify the region of interest was PCR. The PCR product obtained from amplification utilising the Thermo Hybaid® Block was utilised to sequence the amplified region or to digest it during an RFLP procedure.

4.2.2.1 Amplification of template DNA

The mitochondrial DNA region encompassing the positions of interest was amplified with conventional Promega® *Taq* DNA polymerase. Mitochondrial DNA has proved to be ideal for amplification purposes, producing large amounts of product without requiring extensive optimisation. This is due to the high copy number of the mitochondrial genome in comparison to nuclear DNA. However, owing to the nature of the template DNA being amplified without difficulty, it was necessary to optimise the annealing temperature to inhibit secondary product formation, which could influence the downstream sequencing reaction.

The PCR product obtained was sufficient. However, the secondary product that was also formed could interfere with the sequencing reaction. These secondary fragments would not have been eliminated during PCR product purification, and could therefore have produced a signal similar to that of contamination on a sequencing electropherogram. It was therefore imperative that secondary product, depicted in Figure 4.1, be reduced to a minimum. Unfortunately, since a scanned image had to be used here, the quality of the image produced does not reflect the clarity of the original image.

Figure 4.1: Photographic representation of PCR with secondary product amplification from reaction with annealing temperature of 55°C



Fragments were separated on a 2% agarose gel that was electrophoresed at 100 V for 30 min and stained with EtBr. 100 bp = 100 bp marker (Promega), C13 = Positive control, numbers between 893 and 901 = samples investigated.

As illustrated in Figure 4.1 secondary product was formed when utilising an annealing temperature of 55°C. A final annealing temperature was optimised at 59°C, which is illustrated in Figure 4.5 on page 55. A positive and negative control was performed with every reaction. The products from this particular amplification experiment were not utilised further due to the high level of secondary product as well as particular reactions not being successful.

The following factors were taken into consideration when a PCR reaction failed to produce amplified product:

a) That there was not enough template DNA in the reaction tube. In this instance a new value for the concentration of the DNA in the working dilution was obtained spectrophotometrically to verify the amount of working dilution needed per reaction.

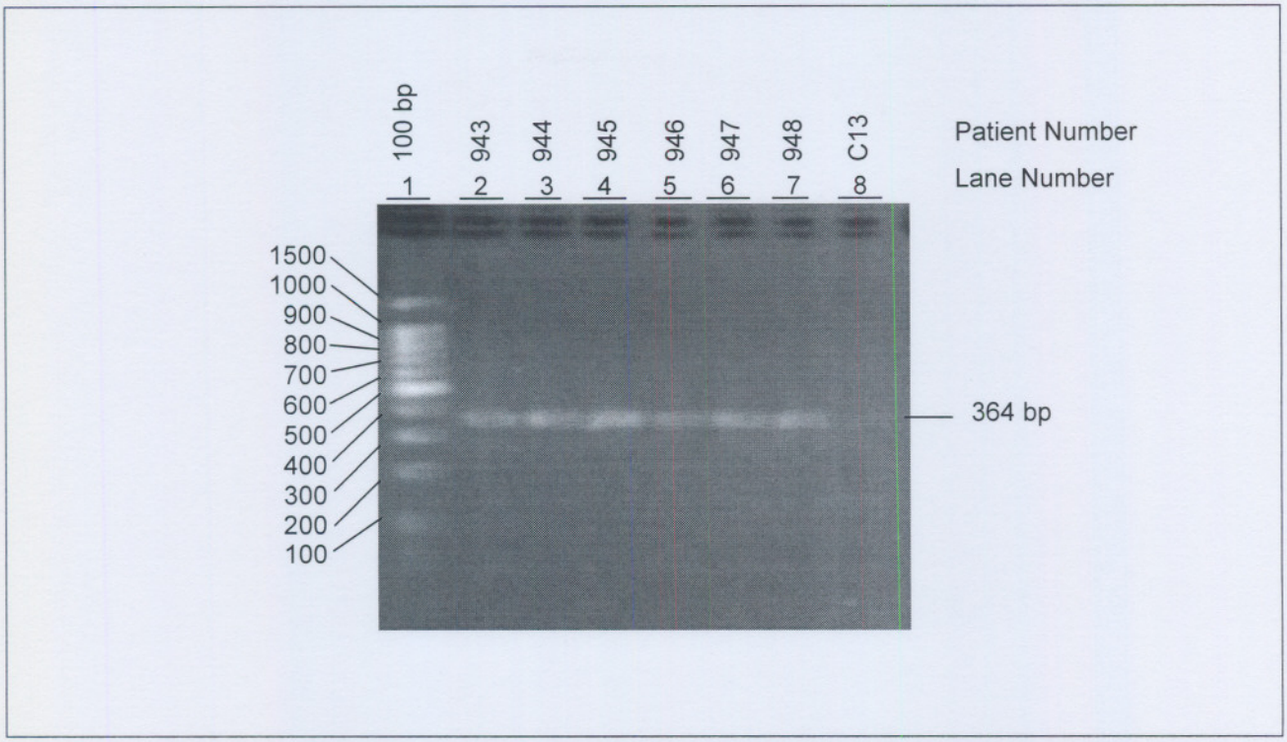
b) One of the reagents in the mixture could have been overlooked and left out of the reaction, in which case the reaction was repeated with caution in order to rectify this.

c) The temperature of the Thermo Hybaid® PCR machine could have failed, in which case the reaction would not have been driven to completion. To resolve this, the reaction was repeated and the programme monitored to ensure that this fault was not a possible cause.

4.2.3 PCR product purification

In order to have pure DNA product from which sequencing could be achieved it was necessary to purify the PCR amplified fragments. The QIAquick purification kit rendered sufficiently high yields that were within a range of 7.3 $ng.\mu l^{-1}$ to 40.2 $ng.\mu l^{-1}$. An example of an agarose gel containing purified PCR product is illustrated in Figure 4.2.

Figure 4.2: Photographic representation of purified amplified fragments



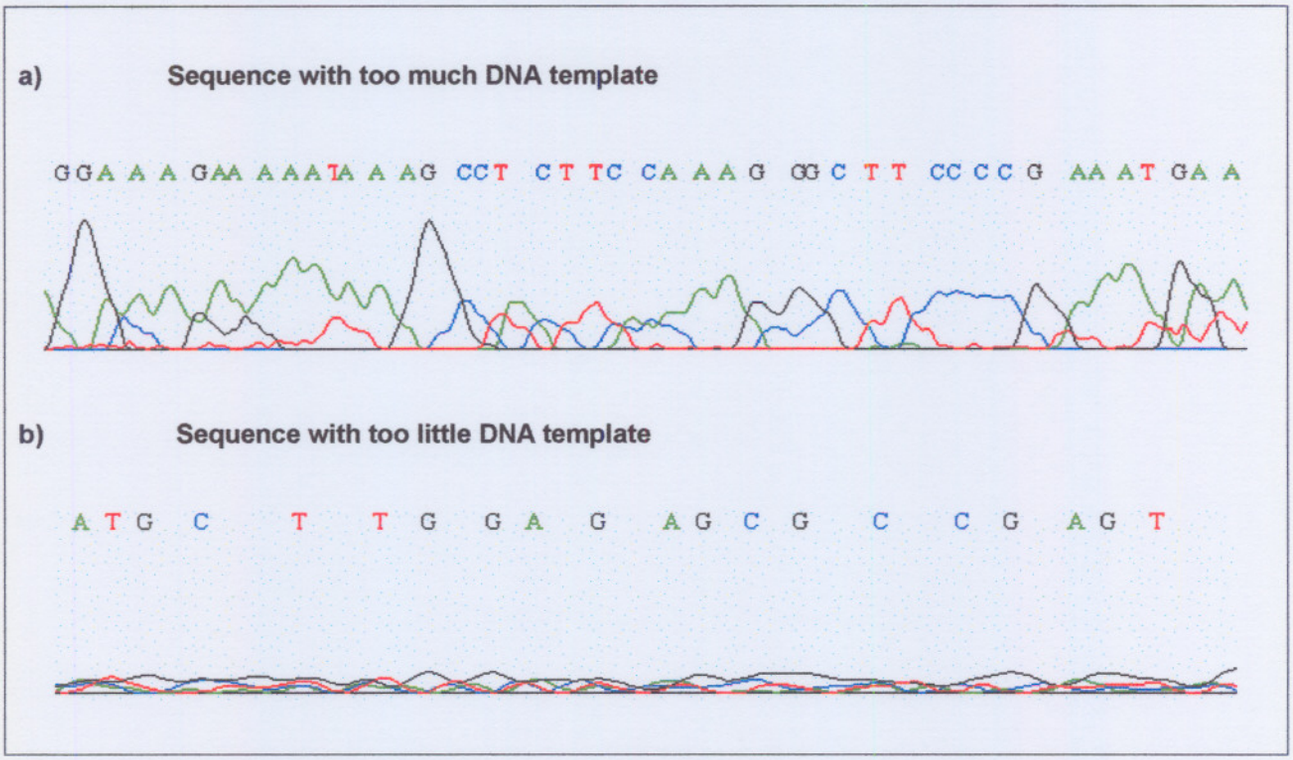
Fragments were separated on a 2% agarose gel that was electrophoresed at 100 V for 30 min and stained with EtBr. 100 bp = 100 bp marker (Promega), C13 = positive control DNA, 943-948 = samples investigated.

4.2.4 Cycle sequencing

This method of automated sequencing was employed as opposed to manual sequencing. The automated form of sequencing is more efficient in terms of time. Results were generated in the form of an electropherogram and the different nucleotide bases are represented by differently coloured peaks.

The sequencing reaction is highly sensitive but it is not possible to distinguish between the template of interest and that of a contaminating template. Therefore, the purity of the template as well as the concentrations of certain reagents can have a significant impact on the outcome of the reaction. DNA concentrations of the template DNA used in the reactions should be measured accurately. For the purpose of these experiments both an agarose gel and spectrophotometric readings were obtained and compared in order to ensure the correct concentrations of DNA were measured before being added to the sequencing reaction mixture. If too much DNA was added broad peaks were visualised as depicted in Figure 4.3 a). If too little DNA was added, no peaks could be distinguished as illustrated in Figure 4.3 b).

Figure 4.3: Representative electropherogram of failed sequence reaction

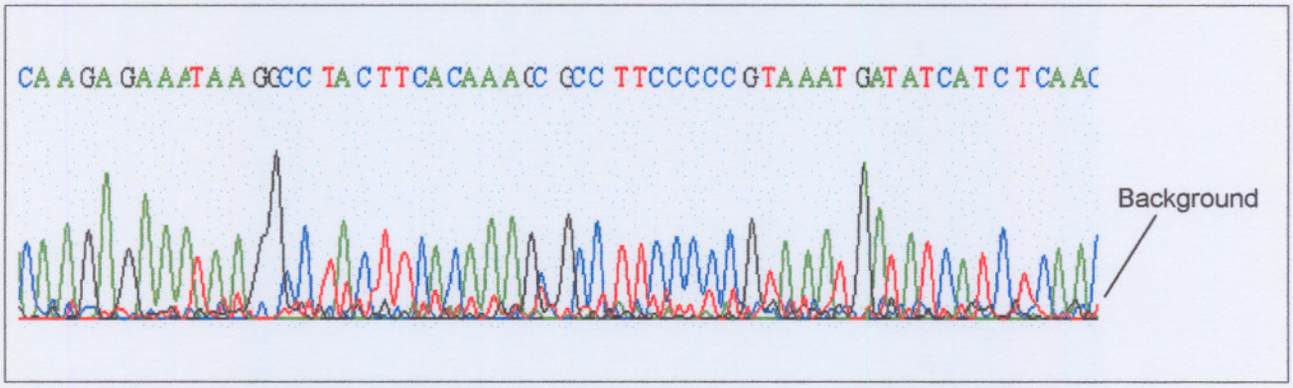


A = adenine; C = cytosine; G = guanine; T = thymine.

An optimised reaction was achieved for this procedure. An example of an ideal sequencing electropherogram is illustrated with Figure 4.6.

Following the sequencing reaction in the Thermo Hybaid® block, it was necessary to precipitate the extended fragments from the reaction mixture of the excess primers, dNTPs, ddNTPs, buffer and enzyme. This particular method was highly sensitive and could play a vital role in the outcome of the final product. During the initial precipitation step of this method, whilst the sequencing product is added to the ethanol mixture, it is essential that the product is not left to stand in the ethanol before centrifugation. It should be centrifuged immediately to prevent any precipitation of unincorporated ddNTPs. This will prevent the manifestation of background in the electropherogram as represented in Figure 4.4. The appearance of background can obscure the result, making it more challenging to identify heteroplasmy, in the case of mitochondrial DNA, or heterozygosity, in the case of nuclear DNA.

Figure 4.4: Representative electropherogram of a sequence reaction with background caused by precipitation of unincorporated dye terminators



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

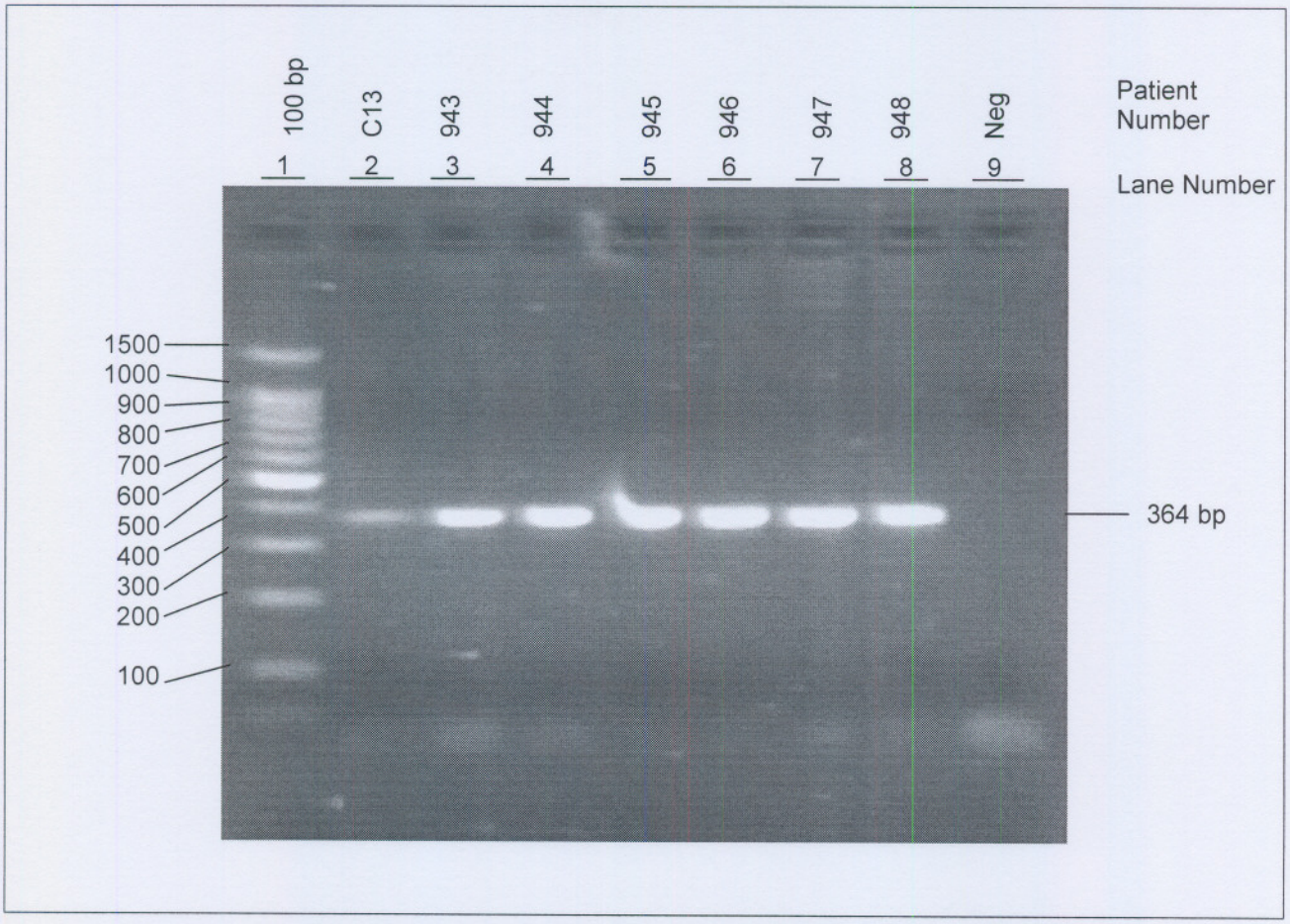
4.3 ANALYSIS OF SEQUENCING RESULTS

Most of the results obtained during this research were obtained utilising automated sequence analysis. The resultant electropherograms were scrutinised for the accuracy of the base-calling, whether the investigated alterations were present or not, as well as if any novel changes could be observed.

4.3.1 Analysis of amplified product

Prior to sequencing, the region of interest first had to be amplified to obtain a PCR product free of secondary product. As presented in Figure 4.5, pure PCR product was obtained for all samples analysed, which was further utilised for sequencing purposes. The region of interest was 364 bp in length. The amplified PCR fragment lies between the marker fragments indicating the 400 bp and 300 bp sizes. It was therefore assumed that the region of interest was successfully amplified. Furthermore the product was free of contamination, as additional fragments could not be observed.

Figure 4.5: Photographic representation of the amplified fragments of region 3107 to 3370



Fragments were separated on a 2% agarose gel that was electrophoresed at 100 V for 30 min and stained with EtBr. 100 bp = 100 bp marker (Promega), C13 = positive control DNA, Neg = negative control, 943-948 = samples investigated.

4.3.2 Analysis of the purification of target region

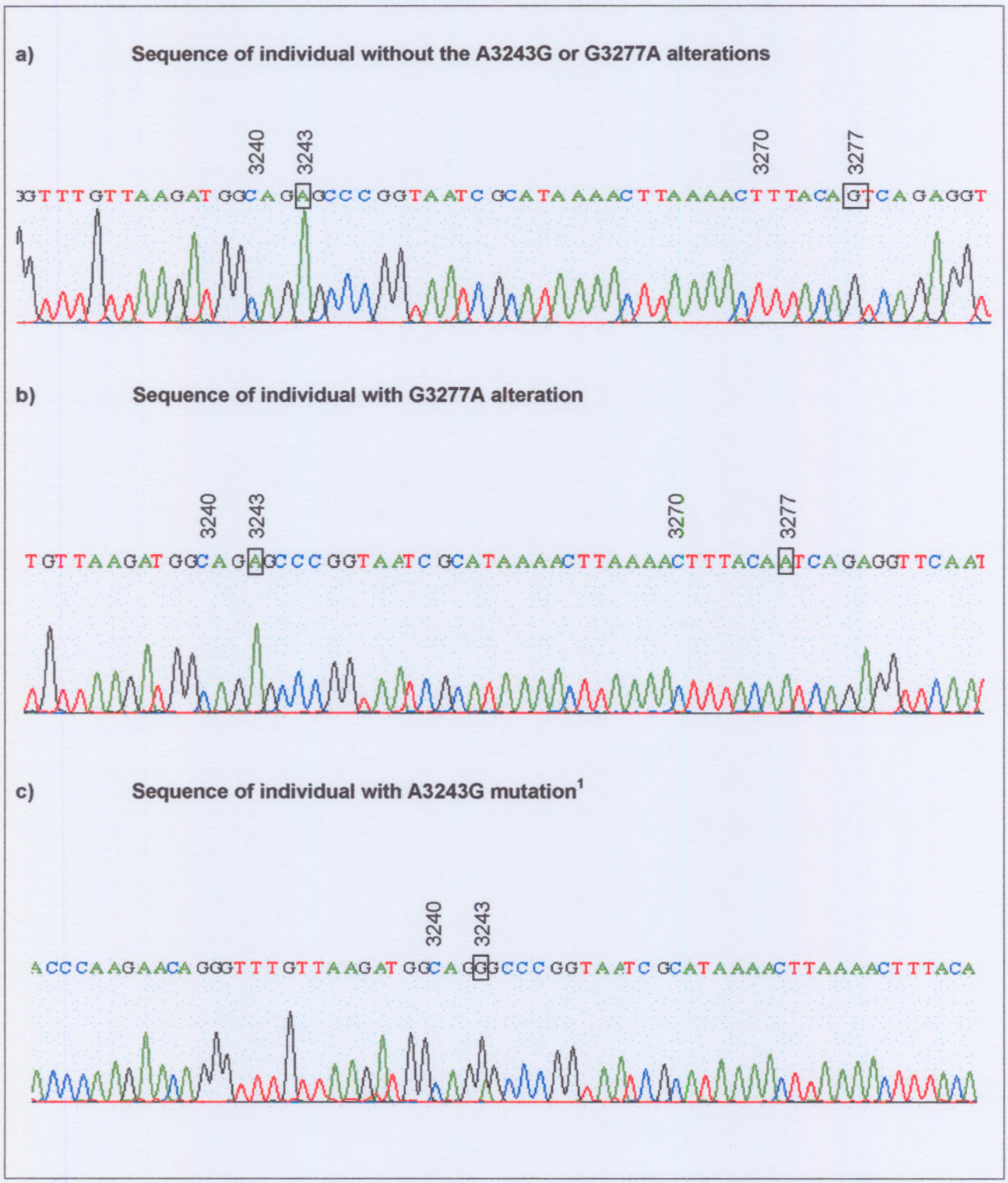
Purification was performed, as described earlier in paragraph 3.2.2.3, subsequent to amplification of the DNA region to remove any contaminating enzymes or reagents. To ensure that the purification procedure had been successful and that there was no significant loss in the amount of product, a sample of the purified product was electrophoresed on a gel. A representation of this is depicted in Figure 4.2 given earlier.

4.3.3 Sequence analysis of target region

Subsequent to amplifying and purifying the region of interest the DNA was subjected to sequencing. Successful sequences clearly illustrated the order of nucleotides in the region. Figure 4.6a) represents an electropherogram that does not harbour either the A3243G or G3277A alterations. Figure 4.6b) is a representation of a patient that harbours the G3277A

alteration and Figure 4.6c) is a patient with the A3243G alteration. These figures are representations of the results obtained from the analysed patients.

Figure 4.6: Representative electropherograms of individuals with and without alterations of interest



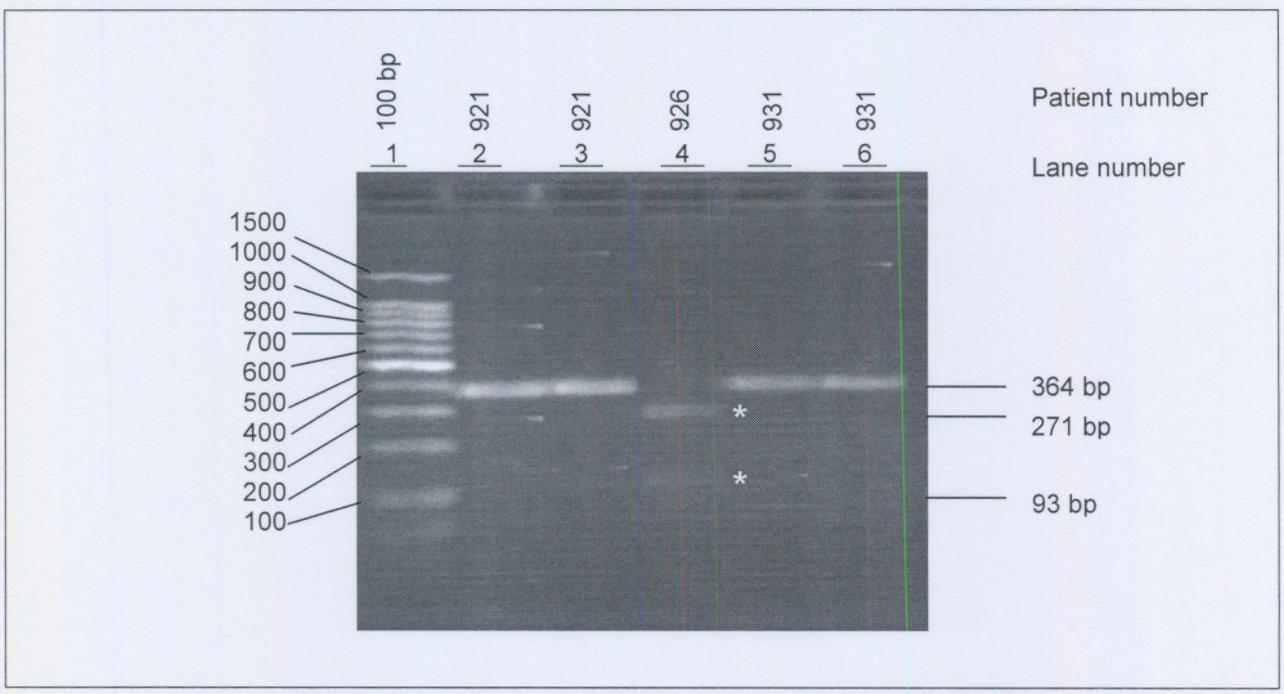
A = adenine; C = cytosine, G = guanine, T = thymine. Square block indicates alteration of interest. Numbers of above sequence indicates the number of the sequence according to MITOMAP, (2003).

[1] Image obtained from A.C van Brummelen.

4.4 RESTRICTION FRAGMENT LENGTH POLYMORPHISM ANALYSIS

Due to the time-consuming procedure of sequencing, a second strategy was attempted to obtain the required results. The RFLP strategy utilised the enzyme *Taa* I to digest the purified PCR product obtained as described in paragraph 4.2.3. If the alteration was present, only one fragment was expected to be observed when the DNA product that had been incubated with the enzyme was electrophoresed on an agarose gel. If the alteration was not present, two fragments were expected to be observed from the product presented on the agarose gel. This is clearly demonstrated in Figure 4.7.

Figure 4.7: Representation of RFLP analysis of the G3277A alteration



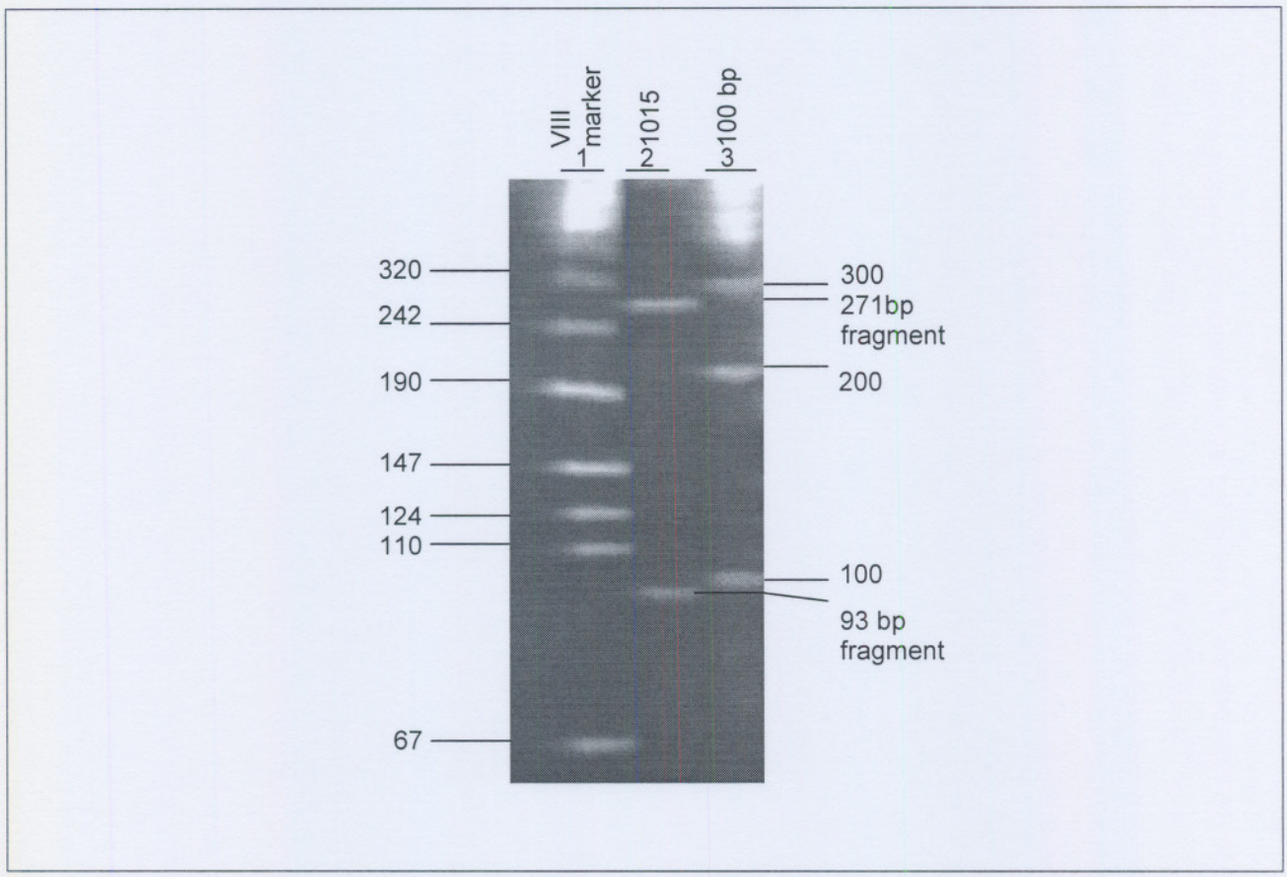
Fragments were separated on a 2% agarose gel that was electrophoresed at 100 V for 30 min and stained with EtBr. 100 bp = 100 bp marker (Promega). 271 bp and 93 bp fragment indicated by white asterisks.

The experiment, for which Figure 4.7 is a representation, was intended to optimise the RFLP procedure utilising the *Taa* I enzyme. The individuals whose DNA was digested had been sequenced prior to the RFLP procedure. These were utilised during this optimisation experiment to verify that the results obtained from RFLP were those that were expected. Therefore, it was known beforehand which individuals had the G3277A alteration.

The two fragments observed in lane four were calculated from Figure 3.1 to be 271 bp for the larger fragment and 93 bp for the smaller fragment if the 364 bp fragment was digested with *Taa* I. It was, however, observed in some of the agarose gels that the smaller fragment did not appear to be 93 bp when sized against a 100 bp marker, as is illustrated

above in Figure 4.7. In order to determine the true size of this fragment, a series of agarose and polyacrylamide gels were electrophoresed to estimate the size of the fragment. A representation of the polyacrylamide gels is demonstrated in Figure 4.8.

Figure 4.8: Representation of polyacrylamide gel to determine size of small fragment



Fragments were separated on a 12% polyacrylamide gel that was electrophoresed at 250 V for 6 hours and stained with Sybr Gold. VIII = DNA molecular weight marker VIII (Roche), 1015 = individual 1015, 100 bp = 100 bp marker (Promega).

From Figure 4.8, it was determined that the size of the smaller fragment was approximately 93 bp. This was in accordance with the calculated size of the fragment. It appears on the agarose gel that the smaller fragment does not migrate at the correct position, but when electrophoresed on a PAGE gel it migrated correctly according to the marker.

4.5 RESULTS OBTAINED FOR THE A3243G ALTERATION

As outlined in the specific aims in Chapter two, where it was possible individuals were screened for the A3243G mutation. No individuals were subjected to RFLP analysis for the A3243G mutation, therefore only those individuals whose DNA was sequenced were screened for this particular mutation. This is due to the primary focus of this particular

research project being the G3277A alteration. None of the 265 individuals analysed harboured the A3243G mutation.

The purpose of the investigation of this mutation in the Black South African diabetic population is the association noted in the Caucasian diabetic population (Van den Ouweland *et al.*, 1992; Reardon *et al.*, 1992). This mutation is more commonly involved in the manifestation of the MERRF phenotype, which can include diabetes as a complication. This was demonstrated in an investigation performed by Sternberg *et al.,* in 2001 where 21 of the 166 patients screened for mitochondrial mutations harboured the A3243G mutation, of which nine manifested with diabetes. It was suggested that a similar association could have been detected in the Black South African diabetic population. Yet, out of all the individuals screened in the present investigation together with those screened by Towers *et al.,* in 2001 none has presented with the mutation in either the diabetic or non-diabetic population. This would indicate that no association can be made with the A3243G mutation and T2DM in this population.

4.6 RESULTS OBTAINED FOR THE G3277A ALTERATION

The main objective of this project was to determine the prevalence of the G3277A alteration in the individuals screened. The G3277A alteration was previously observed with a frequency of 3% in a patient population screened according to the results obtained by Towers *et al.*, (2001). A control population was not screened in the previous project to which the results obtained in the patient population could be compared. Therefore no comparison could be made with the present group of individuals investigated, in terms of the numbers of control individuals expected to harbour the alteration. A total of 379 individuals were screened for this alteration, making use of sequencing and RFLP techniques. Of the 379 individuals, 265 individuals were screened via sequencing and the remaining 114 individuals were screened via RFLP. A total of eight individuals were observed to harbour the G3277A alteration. These results were confirmed by RFLP analysis. Of these eight, seven were control individuals and one was a patient.

The profiles for those individuals with the alteration are provided in Table 4.1. This data were examined in order to determine if there is any similarity between the individuals, so as to ascertain whether the G3277A alteration was associated with any common phenotype. From the information presented in Table 4.1, no common feature was apparent. This does not, however, imply that there is no association. Statistical analyses

were performed in order to examine whether an association could be demonstrated as outlined in section 4.8.

Table 4.1: Profiles of individuals harbouring the G3277A alteration

Individual	Sex	Patient (P) / Control (C)	Ethnic Group			Other Medical conditions
			Individual	Mother	Father	
903	F	C	Xhosa	Zulu	Unknown	Hypertension
921	F	C	Xhosa	Xhosa	Xhosa	None
931	F	C	Ndebele	Ndebele	Ndebele	None
967	F	C	Tswana	S.Sotho	Tswana	None
1032	M	C	Xhosa	Xhosa	Xhosa	None
1108	F	P	Zulu	Zulu	Zulu	Hypertension
1165	F	C	N.Sotho	N.Sotho	N.Sotho	Arthritis
1241	F	C	Tswana	Tswana	Tswana	None

F = female, M = male, C = control, P = patient, S.Sotho = Southern Sotho, N.Sotho = Northern Sotho, numbers between 903 and 1241 = individual investigated.

Subsequent to a species comparison for the 3277 position, as demonstrated in Table 4.2, it was observed that this position is not highly conserved between species. It was, however, apparent that the closest relative to man in this comparison, being the gorilla, shared the same nucleotide at this position. As is observed in Figure 2.8 the 3277 position is situated within the variable loop, indicating that alterations are a commonplace event in this region. From this, it can be postulated that owing to this site not appearing to be highly conserved, it may not have a functionally important position in the $tRNA^{Leu(UUR)}$ gene, and that changes at this location will only have minor effects, if any. Functional studies, however, would need to be performed in future to determine if this is truly so. However, from these observations it appears unlikely that the G3277A alteration would have any association with T2DM. This postulation was examined via statistical analysis.

Table 4.2: Interspecies conservation of the mitochondrial tRNA$^{Leu(UUR)}$ gene

tRNA^{Leu} 3230-3304

```
                1            15 16           30 31           45 46           60 61           75 76
Bovine          GTTA--AGGTGGCAG AGCCCGG-TAATTGC ATAAAACTTAAACTT TTATACCCAGAGATT CAAATCCTCTCCTTA ACA-
Capuchin        GTTA--AGATGGCAG AGCCCGG-CAATTGC ATAAAACTTAAAACT TTACATCAGAGGTT CAACTCCTCTTCTTA ACA-
Human           GTTA--AGATGGCAG AGCCCGG-TAATCGC ATAAAACTTAAAACT TTACATCAGAGGTT CAATTCCTCTTCTTA ACA-
Gorilla         GTTA--AGATGGCAG AGCCCGG-TAATCGC ATAAAACTTAAAACT TTATATCAGAGGTT CAATTCCTCTTCTTA ACA-
Hedgehog        GTTA--GTGTGGCAG AGCCCGG-TAATTGC ATAAAACTTAAAACT TTATACCCAGAGGTT CAAATCCTCTCTCTA ACA-
Blue Whale      GTTG--AGGTGGCAG AGTTCGG-TAATTGC ATAAAACTTAAACCT TTACATCAGAGGTT CAAATCCTCTCCCCA ACA-
Fin Whale       GTTG--AGGTGGCAG AGTTCGG-TAATTGC ATAAAACTTAAACTT TTACACCCAGAGGTT CAAATCCTCTCCCCA ACA-
Wild ass        GTTA--GGGTGGCAG AGCC-GG-AAATTGC GTAAAACTTAAACCT TTACACCCAGAGGTT CAACTCCTCTCCCTA ACA-
Horse           GTTA--GGGTGGCAG AGCCCGG-AAATTGC ATAAAACTTAAACCT TTACATCAGAGGTT CAACTCCTCTCCCTA ACA-
Dog             GTTA--GGGTG-CAG GGCCCGG-TAACTGC GTAAAACTTAAACCT TTACTTCAGAGGTT CAATTCCTCTCCCTA ACA-
Cat             GTTA--GGGTGGCAG AGCCCGG-TAACTGC ATAAAACTTAAGCTT TTATTTCAGAGGTT CAATTCCTCTCCTTA ACA-
Hippo           GTTG--CGATGGCAG AGCCCGG-TAATTGC ATAAAACTTAAACCT TTACACCCAGAGGTT CAAATCCTCTTCACA ACA-
Alligator       GCTA--GGTTGGCAG AGCCTGGCTTAATGC AAAAGGCCTAAACCC TT-TATCAGAGATT CAAATCCTCTACCTA GCA-
Chicken         GCTA--GCGTGGCAG AGCTCGG-CAAATGC AAAAGGCTTAAGCCC TT-TACCAGAGGTT CAAATCCTCTCCCTA GCT-
Carp            --TACTGGGTGGCAG AGCATGGTAAATTGC GAAAGGCCTAAGCCC TTTTACCCAGAGGTT CAAATCCTCTTCCCA GTTT
Drosophila      TCTA--ATATGGCAG A-------TTAGTGC AATAGATTTAAG-CT CTATAATAAAG--T ATTTTACTTTTATTA GAA-
                                1                                2
```

A = adenine; C = cytosine, G = guanine, T = thymine. 1 = position of nt 3243, 2 = position of 3277. Adapted from Procaccio, (2003).

4.7 OTHER ALTERATIONS NOTED WITHIN THE INVESTIGATED POPULATION

While screening for the occurrence of the G3277A alteration and the A3243G mutation in the individuals whose DNA was sequenced, several other alterations were noted. One of the alterations was noted at the 3200 position in the mitochondrial genome in two individuals i.e. 887 and 1096. This alteration is a thymine to adenine change, which is situated 30 bp prior to the tRNA$^{Leu(UUR)}$ gene within the 16S rRNA gene. This alteration has not been reported previously and further investigation would have to be undertaken to determine if this alteration has implications for the functioning of the molecule. A representative electropherogram is depicted in Figure 4.9 for this alteration.

Figure 4.9: Electropherograms of individuals observed with alteration at 3200 position

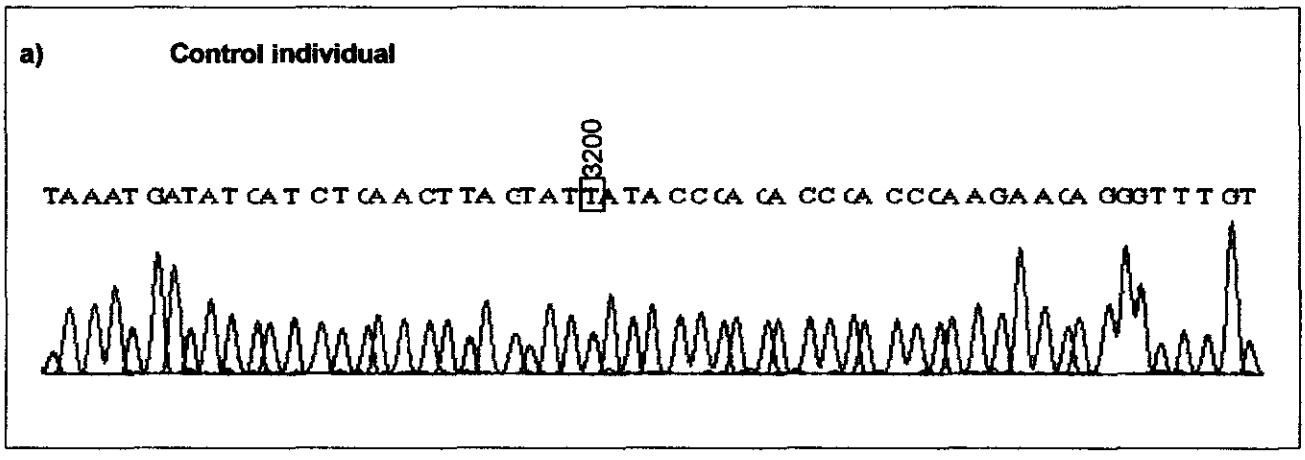
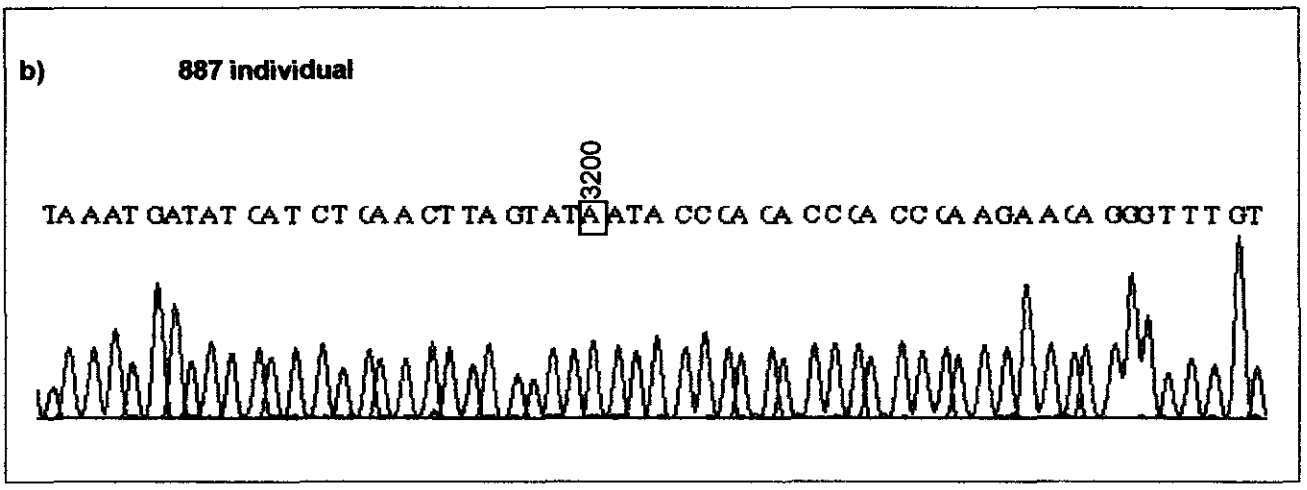


Figure 4.9: continued ...



A = adenine; C = cytosine, G = guanine, T = thymine. 3200 = position of alteration. Square block indicates alteration of interest. Numbers in above sequence indicates the number of the sequence according to MITOMAP, (2003).

Similarities between individuals 887 and 1096 were noted. It was observed that they were both female, diagnosed T2D individuals but from different ethnic backgrounds.

A second alteration was also observed in individual 1091 at position 3308 consisting of a thymine to cytosine change, which is situated 4 bp past the $tRNA^{Leu(UUR)}$ gene in the mitochondrial genome in the NDI gene. Individual 1091 is an obese non-diabetic female diagnosed with hypertension. It has not been observed that any previous investigation has reported this alteration. An electropherogram presenting this finding can be found in Figure 4.10.

Figure 4.10: Electropherogram of individual harbouring an alteration at position 3308

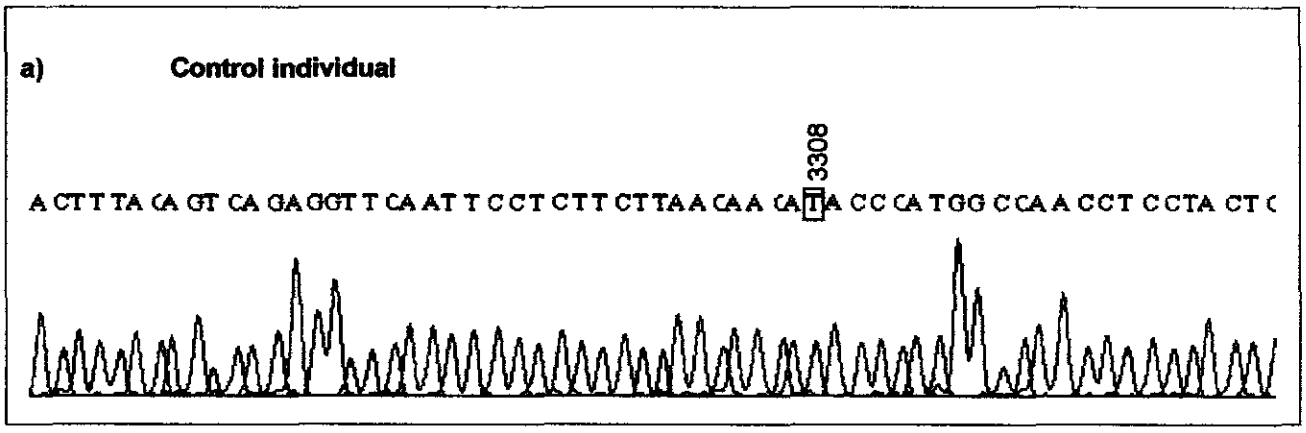
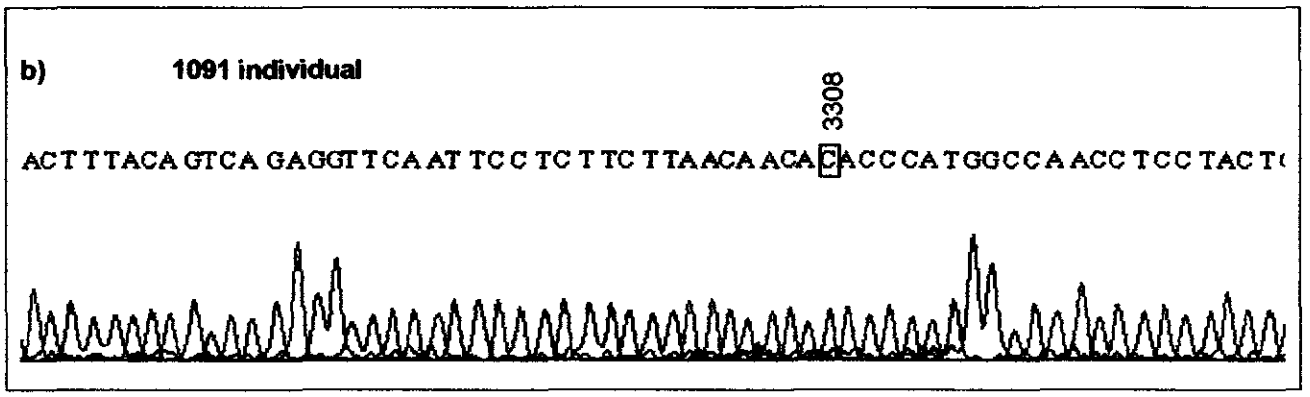


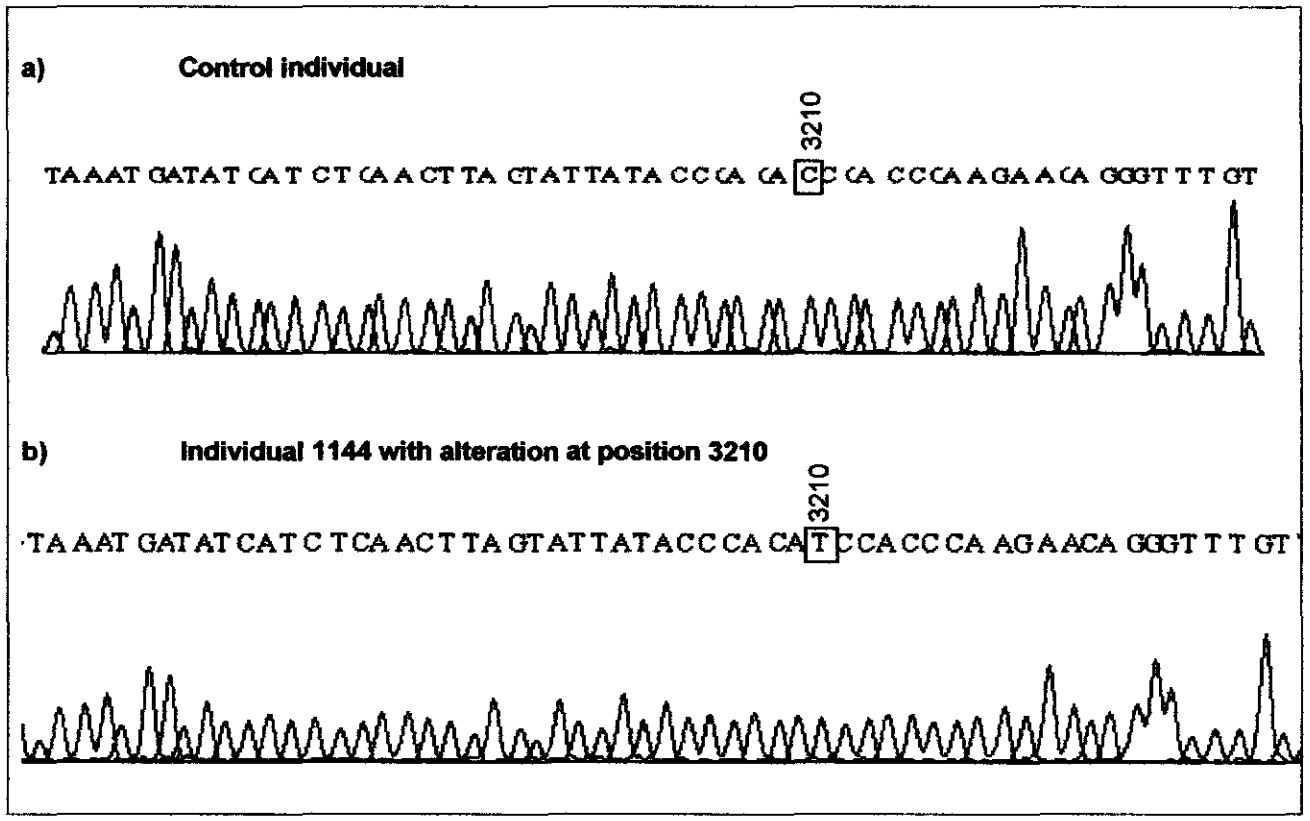
Figure 4.10: continued ...



A = adenine; C = cytosine, G = guanine, T = thymine. 3308 = position of alteration. Square block indicates alteration of interest. Numbers above sequence indicate the number of the sequence according to MITOMAP, (2003).

The third alteration noted at position 3210 within the 16S rRNA gene of the mitochondrial genome is a cytosine to thymine substitution. This change was detected in individual 1144, a non-diabetic female. No other individual has been noted with this alteration in this investigation and it has not been reported previously. The electropherogram depicting this alteration is presented in Figure 4.11.

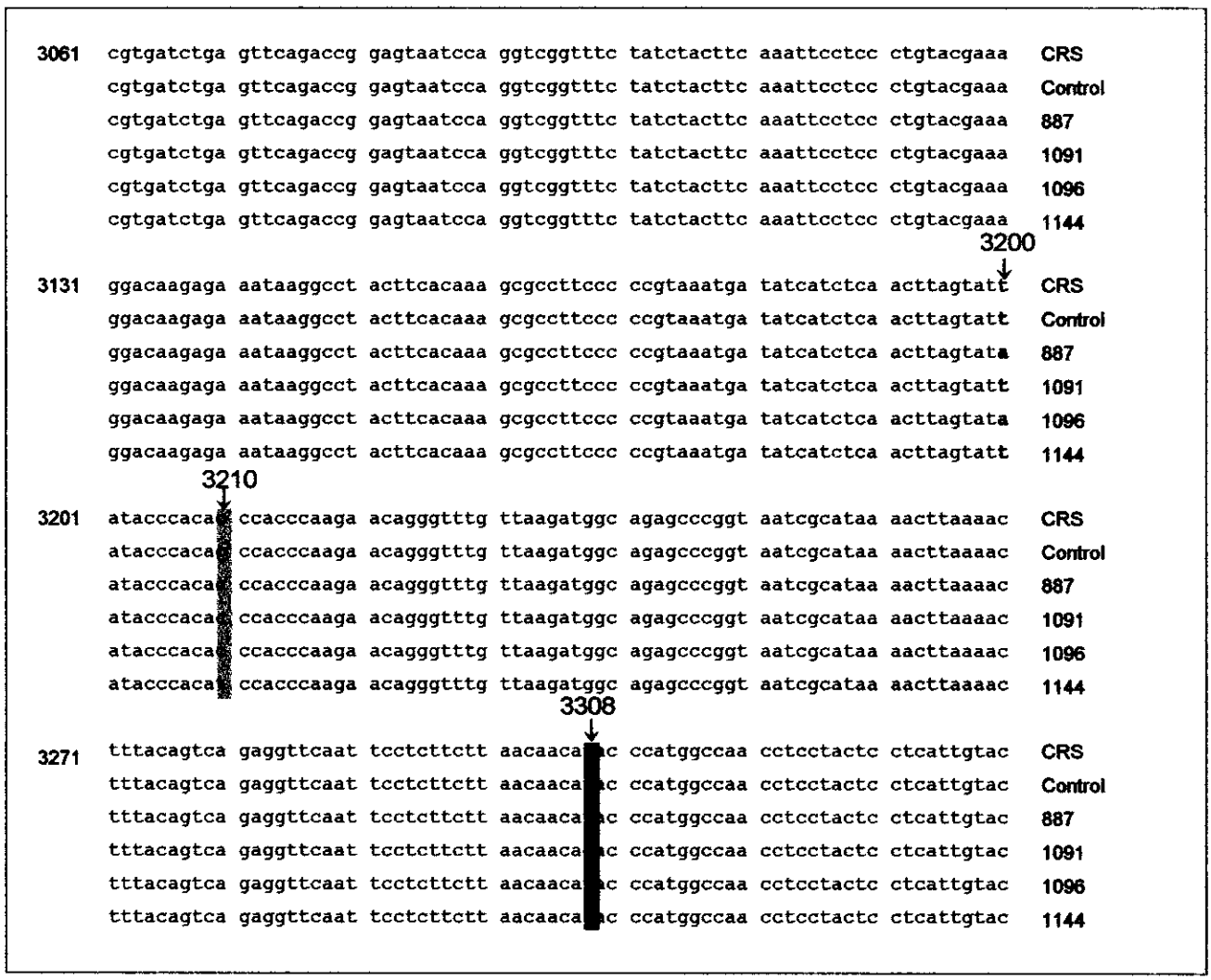
Figure 4.11: Electropherogram of individual harbouring an alteration at 3210 position



A = adenine; C = cytosine, G = guanine, T = thymine. Square block indicates alteration of interest. Numbers above sequence indicate the number of the sequence according to MITOMAP, (2003).

The presence of the above-mentioned alterations was determined via analysing the entire region sequenced, in the 265 individuals whose mitochondrial DNA was sequenced. Figure 4.12 illustrates the sequence alignments of the region sequenced for the four individuals (887, 1091, 1096 and 1144) in which the above additional alterations were observed.

Figure 4.12: gDNA sequence alignment of a portion of the tRNA$^{Leu(UUR)}$ gene



Nucleotide numbering on the alignment is according to the Cambridge Reference Sequence on MITOMAP. CRS = the Cambridge Reference Sequence, control = a control individual, 887-1144 = indicate individual samples noted with alterations in this study. Pink = indicates position 3308 in the mitochondrial genome, yellow = indicates position of 3200 in the mitochondrial genome, blue = position of 3210 in the mitochondrial genome. Numbering of sequence is according to MITOMAP, (2003).

4.8 STATISTICAL DETERMINATION OF ALTERATION FREQUENCY

To determine whether any association exists between the alterations investigated and the T2DM phenotype within the investigated population, statistical analyses were performed. Utilisation of Equation 3.3 generated a value for n of 291 patients that would be required in the study. Control individuals were also required to which the patient samples could be compared. It was indicated that the same number of control individuals would be required

in order to make a statistically valid conclusion (Silman and Macfarlane, 2002). Only 222 diabetic patient individuals and 237 control individuals were collected, leaving the statistics in deficit, making it problematic to draw a final conclusion on the occurrence of the A3243G and G3277A alterations within the population. It was still possible, however, to formulate a calculation based on the current data. This result will lend credence to further analysis of the region in a larger cohort or when sufficient data are collected. Nevertheless, a hypothesis can be generated based on the current information. Out of the 379 individuals that were screened in total, only eight individuals were identified as having the G3277A alteration and none were found to have the A3243G mutation as mentioned in paragraph 4.5. The results outlined in the following paragraphs were calculated utilising the equations presented in section 3.3.

4.8.1 Calculation of Hardy-Weinberg equilibrium for the G3277A alteration

As discussed in Chapter three, it was first determined whether the alteration was in H-W equilibrium within the investigated populations. In order to do this a contingency table was drafted as presented in Table 4.5.

As described in Chapter three, 'p' denotes the symbol for the frequency of the most common allele i.e. 'G' at position 3277 in the mitochondrial genome, while 'q' denotes the frequency of the alteration or the less common allele i.e. A. From Equations 3.4 and 3.5 the allelic frequencies were calculated as depicted in Table 4.3.

Table 4.3: Calculated allelic frequencies G3277A alteration in the diabetic and non-diabetic cohorts

Allele	p	q	Total
Diabetic	0.98	0.02	1
Non-diabetic	0.97	0.03	1

p = frequency of wild-type sequence in the investigated population; q = frequency of sequence with alterations of interest in the investigated population.

The genotypic frequencies of the two alleles i.e. the 'G' and the 'A', for the diabetic and non-diabetic populations are listed in Table 4.4. This is a calculation of the expected proportion of individuals with the respective genotypes in the population.

Table 4.4: Expected genotypic frequencies in the diabetic and non-diabetic cohorts

Genotype	p^2	q^2	2pq	Total
Diabetic	0.96	0	0.04	1
Non-diabetic	0.94	0	0.06	1

p = frequency of wild-type sequence in the investigated population; q = frequency of sequence with G3277A alteration in the investigated population.

The above proportions were then multiplied with the total number of individuals included in the investigation to calculate how many individuals were expected to harbour the respective genotypes in the respective populations. The contingency table for the diabetic population can be observed in Table 4.5.

Table 4.5: Contingency table to determine Hardy-Weinberg equilibrium in the diabetic cohort

Genotypes	1,1	1,2	2,2	Total
Observed numbers (O)	218	0	4	222
Expected proportions	p^2	2pq	q^2	$p^2 + 2pq + q^2$
Expected numbers (E)	214	8	0	222
$\chi^2 = (O-E)^2/E$	0.078	8	Undefined	8.078

χ^2 = chi square value; O = number of individuals observed with the G3277A alteration; E = calculated number of individuals expected to have the G3277A alteration.

The calculated χ^2 value of 8.078 was then compared to the determined critical value of 5.991 described in paragraph 3.3.1.1. This calculated value far exceeded the critical value. Therefore, the null hypothesis is rejected, indicating that the diabetic population is not in H-W equilibrium. A contingency table was also constructed for the non-diabetic population as illustrated in Table 4.6.

Table 4.6: Contingency table to determine Hardy-Weinberg equilibrium in the non-diabetic cohort

Genotypes	1,1	1,2	2,2	Total
Observed numbers (O)	230	0	7	237
Expected proportions	p^2	2pq	q^2	$p^2 + 2pq + q^2$
Expected numbers (E)	223	14	0	237
$\chi^2 = (O-E)^2/E$	0.219	14	undefined	14.219

χ^2 = chi square value; O = number of individuals observed with the G3277A alteration; E = calculated number of individuals expected to have the G3277A alteration.

It was concluded from the value of this χ^2 calculation, which is 14.219 when compared to the critical value, that the non-diabetic population was also not in H-W equilibrium. Therefore, no comparison of the G3277A alteration between the diabetic and non-diabetic cohorts could be made. Only when both are in H-W equilibrium can they be compared. H-W provided a method of normalising the information observed in the diabetic population as compared to the non-diabetic population. The conclusion that can be drawn from these results is that although the samples were collected randomly, the group of individuals collected may be a subgroup within the larger population in which this G3277A alteration is under selection, or that mutation, migration or inbreeding is experienced. A larger sample size would enable researches to confirm this hypothesis.

The most apparent reason for the two populations not being in H-W equilibrium though is the proportion of homozygotes to heterozygotes. As can be observed in Table 4.5 and Table 4.6, no heterozygotes were observed in either of the populations. The decrease in heterozygosity within a population usually implies inbreeding (Tamarin, 1996). If inbreeding is taking place within a population, that population cannot be in H-W equilibrium as this affects the genotypic frequencies, as observed above. The probability that inbreeding is taking place within these groups is unlikely, as these samples were collected from various clinics in and around the Pretoria area in South Africa. It is postulated that the appearance of only homozygous individuals for either the alteration or the wild-type sequence could be due to chance and, therefore, a larger cohort of individuals would be needed to either confirm or refute the findings presented here. If the alteration was under selection, however, it could also have brought about the same result. However, the method of detection utilised, i.e. sequencing, may not have been sensitive enough to detect low levels of heteroplasmy. It is unlikely, though, that this would be a contributing factor, as the results were confirmed for those individuals harbouring the G3277A alteration with RFLP and no heteroplasmy was detected.

4.8.2 Hardy-Weinberg equilibrium calculation for the A3243G mutation

No individuals were observed with this mutation. Therefore no statistical inferences can be made with regard to this mutation in the investigated population. For the reasons outlined in paragraph 4.5, it is postulated that the A3243G mutation is not associated with the T2D phenotype in the Black South African population.

Therefore, with regard to the specific aim of this project, i.e. to determine if there was an association between the G3277A alteration and the T2D phenotype in the Black South African population, no final result could be reached. This, however, does not imply that there is no connection but further investigations will have to be undertaken to explore this concept in greater depth. The second part of the specific aim with regard to the A3243G's association with T2D in the investigated population appears unlikely from the evidence presented. Future screenings of the mitochondria in a larger diabetic cohort would have to be undertaken to attempt to establish an association between a mitochondrial alteration and T2D in the Black South African population. To date, none has been identified.

CHAPTER FIVE

CONCLUSIONS

This investigation focused on the mitochondrial involvement in the development of T2DM in the Black South African population. During the past decade T2DM has been observed to have increased dramatically in the investigated population (Walker *et al.*, 2002). Few genetic factors have been implicated in the occurrence of T2DM in the Black population, yet it is widely understood that T2DM is a highly polygenic disorder with many genetic factors potentially contributing to its manifestation. These factors, however, interact to a large extent with the environment and, depending on environmental stresses i.e. eating or exercise habits, can have a small or severe influence on the individual's health status. Unfortunately, the adoption of a Westernised life-style by this population can have an influence due to the radical change in diet.

The foremost reason, as discussed in section 2.4, for investigating the mitochondrion is that T2DM is a disorder involving the body's interplay with glucose. ATP is produced from the break-down of glucose, and a major component of the ATP-producing pathway occurs within the mitochondrial environment. It has already been demonstrated that a mutation in the mt genome at the 3243 position, from an adenine to guanine, can result in the development of T2DM (Van den Ouweland *et al.*, 1992; Reardon *et al.*, 1992). This mutation has, to date, only been reported to be associated with diabetes in the Caucasian population and not within the Black population. Nonetheless, it provides evidence that alterations in the mitochondrial DNA could have an influence in the development of diabetes.

5.1 RESULTS GENERATED IN THIS STUDY

When it was observed in 2001 by Towers *et al.*, that 3% of the Black South African female diabetic patients that were studied harboured an alteration at the 3277 position whereby a guanine was altered to an adenine, it was questioned whether this alteration could be involved with the development of T2DM in the Black South African population. The present investigation focused on elucidating an answer to this question. An additional 379 blood

samples were collected from a second cohort of Black South African individuals apart from those collected in 2002, and included diabetic and non-diabetic samples. Molecular techniques such as real-time PCR, sequencing and RFLP analysis were attempted to determine the presence of the alteration within both diabetic and non-diabetic individuals. The majority of information was obtained via sequencing, from which it could also be determined if the individuals harboured the A3243G mutation, adding value to the study undertaken in 2001 by Towers *et al.* No individual was found to harbour the A3243G mutation. Therefore, only statistical analyses providing information pertaining to the association of the G3277A alteration with T2DM were calculated.

It was first determined whether the aforementioned alteration was in H-W equilibrium. From the results obtained, as discussed in section 4.8, it was observed that this alteration was not in H-W equilibrium in either the diabetic or non-diabetic population, implicating that an unknown factor is influencing the frequency of this alteration within the groups investigated. It is therefore not possible to formulate assumptions regarding the G3277A alteration's association with the diabetic phenotype, as these assumptions would be meaningless without the two populations having the commonality of both being in H-W equilibrium.

5.2 CONCLUSIONS BASED ON THE GENERATED RESULTS

The reason for the A3277G alteration not being in H-W equilibrium in these cohorts can be deduced from the number of homozygotes observed. As is illustrated in Chapter four, there are no heterozygotes in the population for this alteration. An increase in homozygosity is usually as a result of inbreeding, which alters genotypic frequencies, resulting in the population not being in H-W equilibrium. This could be as a result of the fact that only 222 diabetic patients and 237 non-diabetic individuals were included in this project. Compared to the outcome of the number of individuals needed for this investigation as calculated in section 4.8, page 64 i.e. 291 for each cohort, the sample size was too small. Therefore additional diabetic and non-diabetic individuals should be recruited in the future. These individuals should be from the same population, so as to eliminate as many additional factors as possible that could influence the outcome of the analysis.

An alternative explanation for this observation is that this alteration is a haplogroup specific alteration, not having any influence on the health of an individual. Different ethnic groups have specific alterations that have arisen through evolution and which have no influence on the health of the individual. It would appear, though, that the above mentioned explanation is unlikely as this alteration was observed by Sternberg *et al.*, (2001) in a European Caucasian population. In addition, the few samples identified with the alteration appear to have no common ethnic origin. Stratification of samples in large cohorts would provide more information about the genetic basis of disease. This may provide concrete information as to the true nature of an alteration and whether it could be classified as pathogenic or not.

5.3 IMPACT OF THE STUDY

The information that is available to date suggests that the G3277A alteration is not pathogenic. It has been observed in two populations i.e. Black South African and European Caucasian (Towers *et al.*, 2001; Sternberg *et al.*, 2001), in which neither presented with a high frequency of this alteration. When information regarding the conservation of the alteration within species was examined, the position did appear to be conserved within primates. In addition, the alteration is observed within the variable loop of the $tRNA^{Leu(UUR)}$ molecule. Lastly, in this investigation the alteration was observed with a far higher frequency among the control individuals in comparison to the patient cohort. From the arguments presented here, it is unlikely that the G3277A alteration is associated with diabetes. Functional studies would still have to be performed to confirm this assumption.

The A3243G mutation that was also screened for in this research project produced the same results as those generated in 2001 by Towers *et al.* No association could be found between diabetes in the investigated population and the A3243G mutation as it was not detected in any of the individuals. The reason for screening for this mutation in the current investigation was to determine if the alteration was present. Control non-diabetic individuals were also included in this investigation. The results could have been verified as having a potential association with the investigated phenotype or not, by observing in which cohort (diabetic or control) it was found and comparing it to the frequency of appearance in the other cohort.

5.4 RECOMMENDATIONS FOR FUTURE INVESTIGATIONS

It may be of value in the future to investigate the alteration at the T3200A position, as discussed in section 4.7, detected as one of the three additional alterations observed during this project. This alteration was observed in two diabetic patients but in none of the non-diabetic individuals. To date, there have been no reports regarding an alteration at this position. The other two alterations i.e. T3308C and C3210T were only observed in non-diabetic individuals and thus not associated with the T2DM phenotype.

The most effective way to manage T2DM is to balance diet, exercise and hypoglycaemic agents. In the last few years several companies have convincingly demonstrated the capacity of preventative and therapeutic strategies to reduce hyperglycaemia. New therapeutic and preventative treatments are urgently required and these are most likely to arise out of the rational drug discovery industry based on a thorough molecular understanding of the fundamental processes determining disease pathogenesis. Treatment strategies that focus on decreasing oxidative stress as well as increasing mitochondrial function may present important options.

In order to determine if alterations in the mitochondrion have an influence in the pathogenesis of T2DM the whole mt genome of more diabetic individuals would have to be screened in a future study. The influences of nuclear genes are also to be elucidated for alterations that could contribute to the development of this phenotype. The answers to these studies will provide much needed information for the development of individualised medication, to assist in the prevention of this disorder. In addition it will facilitate the early detection, appropriate treatment and continued management of the complex T2D phenotype.

REFERENCES

6.1 GENERAL REFERENCES

Alberti K.G. and Zimmet P.Z. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet. Med.*, **15**, 539-553, 1998.

Alcolado J.C. and Thomas A.W. Maternally Inherited Diabetes Mellitus: the Role of Mitochondrial DNA Defects. *Diabet. Med.*, **12**, 102-108, 1995.

Altshuler D., Krugluak L. and Lander E. Genetic polymorphisms and disease. *N. Engl. J. Med.*, **338**, 1626, 1998.

Anderson S., Bankier A.T., Barrel B.G., de Bruin M.H.C., Coulson A.R., Drouin J., Eperon I.C., Nierlich D.P., Roe B.A., Sanger F., Schreier A.R., Smith A.J.H., Staden R. and Young I.G. Sequence and organisation of the human mitochondrial genome. *Nature*, **290**, 457-470, 1981.

Arita Y., Kihara S., Ouchi N., Takahashi M., Maeda K., Miyagawa J., Hotta K., Shimomura I., Nakamura T., Miyaoka K., Kuriyama H., Nishida M., Yamashita S., Okubo K., Matsubara K., Muraguchi M., Ohmoto Y., Funahashi T. and Matsuzawa Y. Paradoxical decrease of an adipose-specific protein, adiponectin in obesity. *Biochem. Biophys. Res. Commun.*, **257**, 79-83, 1999.

Baier L.J., Permana P.A., Yang X., Piatley R.E., Hanson R.L., Shen G.Q., Mott D., Knowler W.C., Cox N.J., Horikawa Y., Oda N., Bell G.I. and Bogardus C. A calpain 10 gene polymorphism is associated with reduced muscle mRNA levels and insulin resistance. *J. Clin. Invest.*, **106**, R69-R73, 2000.

Bell G.I. and Polonsky K.S. Diabetes mellitus and genetically programmed defects in β-cell function. *Nature*, **414**, 788-791, 2001.

Berdanier C.D. Mitochondrial gene expression in diabetes mellitus: effect of nutrition. *Nutr. Rev.*, **59**, 61-70, 2001.

Brown P.M., Tompkins C.V., Juul S. and Sonsken P.H. Mechanism of action of insulin in diabetic patients: a dose related effect on glucose production and utilisation. *Br. Med. J.*, **1**, 1239-1242, 1978.

Bruning J.C., Michael M.D., Winnay J.N., Hayashi T., Horsch D., Accili D., Goodyear L.J. and Kahn C.R. A muscle-specific insulin receptor knockout exhibits features of the metabolic syndrome of NIDDM without altering glucose tolerance. *Mol. Cell*, **2**, 559-567, 1998.

Campbell M.K. Biochemistry. 2nd Edition. Sauders College Publishing, Harcourt Brace College Publishers, 236, 1995.

Chance B, Sies H. and Boveris A. Hydroperoxide metabolism in mammalian cells. *J. Cell. Physiol.*, **59**, 527-605, 1979.

Clayton D.A. Transcription of the mammalian mitochondrial genome. *Annu. Rev. Biochem.*, **53**, 557-594, 1984.

Collins T. Biology of Disease: Endothelial nuclear factor κB and the initiation of the atherosclerotic lesion. *Lab. Invest.*, **68**, 449-508, 1993.

Considine R.V., Sinha M.K., Heiman M.L., Kriaudunas A., Stephans T.W., Nyce M.R., Ohannesian J.P., Marco C.C., McKee L.J., Bauer T.L. and Caro J.F. Serum immunoreactive concentrations in normal-weight and obese humans. *N. Engl. J. Med.*, **334**, 292-295, 1996.

Cooke M.S., Evans M.D., Dizdaroglu M. and Lunec J. Oxidative DNA damage: mechanisms, mutation, and disease. *F.A.S.E.B.*, **17**, 1195-1214, 2003.

Cox N. Challenges in identifying genetic variation affecting susceptibility to type 2 diabetes: examples from studies of the calpain-10 gene. *Hum. Mol. Genet.*, **10**, 2301-2305, 2001.

Cross A. and Jones O. Enzymic mechanisms of superoxide production. *Biochim. Biophys. Acta.*, **1057**, 281-298, 1991.

Cusi K., Maezono K., Osman A., Pendergrass M., Patti M.E., Pratipanawatr T., De Fronzo R.A., Kahn C.R. and Mandarino L.J. Insulin resistance differentially affects the PI3 kinase- and MAP kinase-mediated signalling in human muscle. *J. Clin. Invest.*, **105**, 311-320, 2000.

Danforth E. Jr. Failure of adipocyte differentiation causes type II diabetes mellitus? *Nat. Genet.*, **26**, 13, 2000.

De Vries D., de Wijs I., Ruitenbeek W., Begeer J., Smit P., Bentlage H. and van Oost B. Extreme variability of clinical symptoms among sibs in a MELAS family correlated with heteroplasmy for the mitochondrial A3243G mutation. *J. Neurol. Sci.*, **124**, 77-82, 1994.

Fajans S.S., Clouter M.C. and Crowther R.L. The Banting Memorial lecture 1978. Clinical and aetiological heterogeneity of idiopathic diabetes mellitus. *Diabetes*, **27**, 1112-1125, 1978.

Finnilä S., Hassinen I.E. and Majamaa K. Restriction fragment analysis as a source of error in detection of heteroplasmic mtDNA mutations. *Mutation Research Genomics*. **406**, 109-114, 1999.

Gannon M. Molecular genetic analysis of diabetes in mice. *Trends Genet.*, **17**, S23-S28, 2001.

Gannon M. and Wright C.V.E. Endodermal patterning and organogenesis. In: Cell lineage and fate determination. Moody, S.A., ed. Academic Press, 583-615, 1999.

Gerbitz K.D., van den Ouwerland J.M.W., Maassen J.A. and Jaksch M. Mitochondrial diabetes mellitus: a review. *Biochim. Biophys. Acta.*, **1271**, 253-260, 1995.

Gutierrez C., Vendrell J., Broch M., Aguilar C., Llor C., Simon I. and Richart C. GLUT1 gene polymorphism in non-insulin-dependent diabetes mellitus: genetic susceptibility relationship with cardiovascular risk factors and microangiopathic complications in a Mediterranean population. *Diabetes Res. Clin. Pract.*, **41**, 113-120, 1998.

Hales C.N. and Barker D.J.P. Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis. *Diabetologia*, **35**, 595-601, 1992.

Hauswirth W.W. and Laipis P.J. Transmission genetics of mammalian mitochondria: a molecular model and experimental evidence. In: Quagliarello E., Slater E.C., Palmieri F., Sacconne G., Kroon A.M. (eds). Achievements and perspectives of mitochondrial research, vol. 2. Biogenesis, Elsevier Biomedical, Amsterdam, 49-59, 1985.

Hess J.F., Parisi M.A., Bennet 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.

Himsworth H.P. Diabetes mellitus: its differentiation into insulin-sensitive and insulin insensitive types. *Lancet*, **1**, 127-130, 1936.

Horikawa Y., Oda N., Cox N.J., Li X., Orho-Melander M., Hara M., Hinokio Y., Lindner T.H., Mashina H., Swartz P.E.H., del Bosque-Plata L., Horikawa Y., Oda Y., Yoshiuchi I., Colilla S., Polonsky K.S., Wei S., Concannon P., Iwasaki N., Schulze J., Baier L.J., Bogardus C., Groop L., Boerwinkle E., Hanis C.L. and Bell G.I. Genetic variation in the gene encoding calpain-10 is associated with type2 diabetes mellitus. *Nat. Genet.*, **26**, 163-175, 2000.

Johns D.R. The other human genome: mitochondrial DNA and disease. *Nat. Med.*, **2**, 1065-1068, 1996.

Kannel W.B. and McGee D.L. Diabetes and cardiovascular risk factors: the Framingham study. *Circulation*, **59** (1), 8-13, 1979.

Katagiri H., Asano T., Ishihara H., Inukai K., Anai M., Yamanouchi T., Tsukuda K., Kikuchi M., Kitaoka H., Ohsawa N., Yazaki Y. and Oka Y. Mitochondrial diabetes mellitus: prevalence and clinical characterisation of diabetes due to mitochondrial $tRNA^{Leu(UUR)}$ gene mutation in Japanese patients. *Diabetologia*, **37**, 504-510, 1994.

Kim M.J.C. and Berdanier C.D. Nutrient-gene interactions determine mitochondrial function: effect of dietary fat. *F.A.S.E.B. J*, **12**, 243-248, 1998.

King H., Aubert R.E. and Herman W.H. Global burden of diabetes, 1995-2025. Prevalence, numerical estimates, and projections. *Diabetes Care*, **21** (9), 1414-1431, 1998.

Kobayashi Y., Ichihashi K., Ohta S., Nihei K., Kagawa Y., Yanagisawa M. and Momoi M.Y. The mutant mitochondrial genes in the mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes (MELAS) were selectively amplified through generations. *J. Inher. Metab. Dis.*, **15**, 803-808, 1992.

Kobberling J. Studies on the genetic heterogeneity of diabetes mellitus. *Diabetologia*, **7**, 46-49, 1971.

Kunkel T.A. and Loeb L.A. Fidelity of mammalian DNA polymerases. *Science*, **213**, 765-768, 1981.

Lamson D.W. and Plaza S.M. Mitochondrial factors in the pathogenesis of diabetes: a hypothesis for treatment. *Altern. Med. Rev.*, **7**(2), 94-111, 2002.

Laipis P.J., Van de Walle M.J. and Hauswirth W.W. Unequal partitioning of bovine mitochondrial genotypes among siblings. *Proc. Natl. Acad. Sci. USA.*, **85**, 8107-8110, 1988.

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

Last J.M. A dictionary of epidemiology. Fourth Edition. New York, Oxford United Press, Inc., 2001.

Lindgren C.M. and Hirschorn, J.N. The genetics of type 2 diabetes. *The Endocrinologist*, **11**, 178-187, 2001.

Mader S. Biology. Fifth Edition. Wm. C. Brown Publishers, 777-778, 1996.

Margolis R., Laughlin M. and Haft C.R. β-cell biology in the 21st century. *Trends Endocrinol. Metab.*, **13** (5), 185-187, 2002.

Marui N., Offerman M.K., Swerlick R., Kunsch C., Rosen C.A., Ahmad M., Alexander R.W. and Medford R.M. Vascular cell adhesion molecule-1 (VCAM-1) gene transcription and expression are regulated through an antioxidant-sensitive mechanism in human vascular endothelial cells. *J. Clin. Invest.*, **92**, 1866-1874, 1993.

Marx J. Unravelling the causes of diabetes. *Science*, **296**, 686-689, 2002.

McCarthy M.I. and Froguel P. Genetic approaches to the molecular understanding of type 2 diabetes. *Am. J. Physiol. Endocrinol. Met.*, **283**, E217-E225, 2002.

McClain D.A. Hexoamines as mediators of nutrient sensing and regulation in diabetes. *J. Diabetes Complications*, **16**, 72-80, 2002.

Michael M.D., Kulkarni R.N., Postic C., Previs S.F., Shulman G.I., Magnuson M.A. and Kahn C.R. Loss of insulin signalling in hepatocytes leads to severe insulin resistance and progressive hepatic dysfunction. *Mol. Cell.*, **6**, 87-97, 2000.

Morton N.E. and Collins A. Tests and estimates of allelic association in complex inheritance. *Proc. Natl. Acad. Sci. USA*, **95**, 11389-11393, 1998.

Nathan D.M. Prevention of long-term complications of non-insulin-dependent diabetes mellitus. *Clin. Invest. Med.*, **18**, 332-339, 1995.

Narang S.A., Brousseu R., Georges F., Michiewicz J., Prefontaine G., Stawinski J. and Sung W. The human preproinsulin gene: synthesis, cloning, gene manipulation, and expression studies. *Can. J. Biochem. Cell. Biol.*, **62**, 209-216, 1984.

Newman B., Selby J.V., King M.C., Slemenda C., Fabsitz R. and Friedman G.D. Concordance for Type-2 (non-insulin dependent) diabetes mellitus in male twins. *Diabetologia*, **30**, 763-768, 1987.

Okuno A., Tamemoto H., Tobe K., Ueki K., Mori Y., Iwamoto K., Umesono K., Akanuma Y., Fujiwara T., Horikoshi H., Yazaki Y. and Kadowaki T. Troglitazone increases the number of small adipocytes without the change of white adipose tissue mass in obese zucker rats. *J. Clin. Invest.*, **101**, 1354-1361, 1998.

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

Reardon W., Ross R.J., Sweeney M.G., Luxon L.M., Pembrey M.E., Harding A.E. and Trembath R.C. Diabetes mellitus associated with a pathogenic point in mitochondrial DNA. *Lancet*, **340**, 1376-1379, 1992.

Reusch J.E. Current Concepts in Insulin Resistance, Type 2 Diabetes Mellitus, and the Metabolic Syndrome. *Am. J. Cardiol.*, **90**, 2002.

Rewers M. and Hamman R.F. Risk factors for non-insulin-dependent diabetes: In: diabetes in America, Edited by Group NDD, 179-220, Bethesda MD, NIH/NIDDK, 1995.

Richter C. Reactive oxygen and DNA damage in mitochondria. *Mutat. Res.*, **275**, 249-255, 1992.

Rimoin D.L. Inheritance of diabetes mellitus. *Med. Clin. North Am.*, **55**, 807-819, 1971.

Risch N. and Merikangas K. The future of genetic studies of complex human diseases. *Science,* **273**, 1516-1517, 1996.

Rossetti L. and Giaccari A. Relative contribution of Glycogen Synthesis and Glycolysis to Insulin-mediated Glucose Uptake. *J. Clin. Invest.*, **85**, 1785-1791, 1990.

Sambrook J., Fritsch E.F. and Maniatis T. Molecular cloning, a laboratory manual, Second Edition, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York, 1989.

Samuels M.L. Statistics for the life sciences. Prentice Hall Publishers, 329-335, 1989.

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

Schmidt A.M. and Stern D. A radical approach to the pathogenesis of diabetic complications. *TiPS,* **21**, 2000.

Schreck R., Rieber P. and Baeuerle A. Reactive oxygen intermediates as apparently widely used messengers in the action of the NF-κB transcription factor and HIV-1. *E.M.B.O. J.*, **10**, 2247-2258, 1991.

Schreck R., Meier B., Mannek D.N., Droge W. and Baeuerle P.A. Dithiocarbamates as potent inhibitors of NF-κB activation in intact cells. *J. Exp. Med.,* **175**, 1181-1194, 1992.

Semete B. Centre for Genome Research, Potchefstroom University for Christian Higher Education, Pretoria, South Africa, Personal communication, 2003.

Sentinelli F., Romeo S., Arca M., Filippi E., Leonetti F., Banchieri M., Di Mario U. and Baroni M.G. Human resistin gene, obesity, type 2 diabetes: mutation analysis and population study. *Diabetes,* **51**, 860-862, 2002.

Shearman C.W. and Kalf G.F. DNA replication by a membrane-DNA complex from rat liver mitochondria. *Arch. Biochem. Biophys.*, **182**, 573-586, 1977.

Silman A.J. and Macfarlane G.J. Epidemiological studies: A practical guide. Second edition. United Kingdom, Cambridge University Press, 90-92, 2002.

Simpson N.E. Genetics of diabetes: A study of 233 families of juvenile diabetes. *Ann. Hum. Genet.*, **26**, 1-21, 1962.

Simpson R.W., Shaw J.E. and Zimmet P.Z. The prevention of type 2 diabetes – lifestyle change or pharmacotherapy? A challenge for the 21st century. *Diabetes Res. Clin. Pract,* **59**, 165-180, 2003.

Smith S.R., Bai F., Charbonneau C., Janderova L. and Argyropoulos G. A promoter genotype and oxidative stress potentially link resistin to human insulin resistance. *Diabetes*, **52**, 1611-1618, 2003.

Song J., Oh J.J., Sung Y., Pak Y.K., Park K.S. and Lee H.K. Peripheral blood mitochondria DNA content is related to insulin sensitivity in offspring of type 2 diabetic patients. *Diabetes Care*, **24** (5), 865-869, 2001.

Sonksen P. and Sonksen J. Insulin: understanding its action in health and disease. *Br. J. Anaesth.*, **85**, 69-79, 2000.

Steppan C.M. and Lazar M.A. Resistin and obesity-associated insulin resistance. *Trends Endocrinol. Metab.*, **13** (1), 18-23, 2002.

Sternberg D., Chatzoglou E., Laforet P., Fayet G., Jardel C., Blondy P., Fardeau M., Amselem S., Eymard B. and Lombes A. Mitochondrial DNA transfer RNA gene sequence variantions in patients with mitochondrial disorders. *Brain*, **124** (5), 984-994, 2001.

Suzuki S., Oka Y., Kadowaki T., Kanatsuka A., Kuzuya T., Koayashi M., Sanke T., Seino Y. and Nanjo K. Clinical features of diabetes mellitus with the mitochondrial DNA 3243 (A-G) mutation in Japanese: Maternal inheritance and mitochondria-related complications. *Diabetes Res. Clin. Pract.*, **59**, 207-17, 2003.

Szopa T.M., Tichener P.A., Portwood N.D. and Taylor K.W. Diabetes due to viruses - some recent developments. *Diabetologia*, **36**, 687-695, 1993.

Tamarin R.H. Principles of Genetics. Fifth edition, Wm.C.Brown publishers, 533-535, 1996.

Tattershall R. and Fajans S.S. A difference between the inheritance of classic juvenile-onset and maturity-onset type diabetes of young people. *Diabetes*, **24**, 45-53, 1975.

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

Thomas A.W., Edwards A., Sherratt E.J., Majid A., Gagg J. and Alcolado J.C. Molecular scanning of candidate mitochondrial tRNA genes in type 2 (non-insulin dependent) diabetes mellitus. *J. Med. Genet.*, **33**, 253-255, 1996.

Torroni A. MtDNA haplogroups in human populations and disease studies. *J. Cult. Heritage*, **1**, s33-s34. 2000.

Towers W.G., Prosser D., Rheeder P., Wallace D.C. and Olckers A. Molecular Screening for the mitochondrial A3243G mutation in a black NIDDM South African population. The neurological association of South Africa meeting, Sandton Convention Centre, Johannesburg, South Africa, 2001.

Unger R.H. and Orci L. Lipotoxic diseases of nonadipose tissue in obesity. *Int. J. Obes. Relat. Metab. Disord.*, **24** (Suppl.4), s28-s32, 2000.

Vadheim C.M. and Rotter J.I. Genetics of diabetes mellitus. In molecular and cellular biology of diabetes mellitus. Vol 1, 139-148, Draznin B., Melmed S., LeRoith D. and Alan R. USS. New York: Liss, c1989.

Van Brummelen A.C. Centre for Genome Research, Potchefstroom University for Christian Higher Education, Pretoria, South Africa, Personal communication, 2003.

Van den Ouweland J.M., Lemkes H.H.P., Ruitenbeek W., Sandkuijl L.A., de Vijlder M.F., Struyvenberg J.J.P., van de Kamp J.J.P. and Maassen J.A. Mutation in mitochondrial $tRNA^{Leu(UUR)}$ gene in a large pedigree with maternally transmitted type II diabetes mellitus and deafness. *Nat. Genet.*, **1**, 368-371, 1992.

Van den Ouweland J.M., Lemkes H.H., 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.

Walker A.R.P., Walker B.F. and Adam F. Variations in occurrences of nutrition-related disease in Sub-Saharan African in the stages of transition: what of the future? *Nutrition*, **18**, 71-74, 2002.

Wallace D.C. Diseases of mitochondrial DNA. *Annu. Rev. Biochem.*, **61**, 1175-1212, 1992.

Wallace D.C. Personal communication, 1999.

Wei Y.H. Oxidative stress and mitochondrial DNA mutations in human aging. *Proc. Soc. Exp. Biol. Med.*, **217**, 53-63, 1998.

Williams R.C., Long J.C., Hansen R.L., Sievers L. and Knowler W.C. Individual estimates of European genetic admixture associated with lower body-mass index, plasma glucose, and prevalence of type 2 diabetes in Pima Indians. *Am. J. Hum. Genet.*, **46**, 527-538, 2000.

Wucherpfennig K.W. Insights into autoimmunity gained from structural analysis of MHC-peptide complexes. *Curr. Opin. Immun.*, **13**, 650-656, 2001.

Yakes F.M., and van Houten B. Mitochondrial DNA damage is more extensive and persists longer than nuclear DNA damage in human cells following oxidative stress. *Proc. Natl. Acad. Sci. USA.*, **94**, 514-519, 1997.

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 encephalopathy. *Proc. Natl. Acad. Sci. USA.*, **89**, 11164-11168, 1992.

Yousufzai S.Y.K., Bradford M.W., Shrago E. and Ewart R.B.L. Characterisation of the adenine nucleotide translocase in pancreatic islet mitochondria. *F.E.B.S. Lett.*, **137**, 205-208, 1982.

Zhang Y., Picenca R., Maffei M., Barone M., Leopol L. and Friedman J.M. Positional cloning of the mouse obese gene and its human homologue. *Nature*, **372**, 425-432, 1994.

6.2 ELECTRONIC REFERENCES

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

MITOMAP, A Human Mitochondrial Genome Database, Centre for Molecular Medicine, Emory University, Atlanta, GA, USA. http://www.gen.emory.edu/mitomap.html, 2003.