$CoenzymeQ_{10}$-associated gene mutations in South African patients with respiratory chain deficiencies

L Jonck

22122516

Dissertation submitted in partial fulfillment of the requirements for the degree *Magister Scientiae* in Biochemistry at the Potchefstroom Campus of the North-West University

Supervisor: Prof FH van der Westhuizen

Co-supervisor: Prof I Smuts

May 2015

“A four-letter alphabet called DNA.”

— Matt Ridley

Acknowledgements

I would like to express my utmost gratitude to the following people without whom this dissertation and the past five years would not have been possible.

First and foremost, my supervisor, **Prof. Francois van der Westhuizen**. Not only for the success behind my studies with your vast experience guiding me, but for going the extra mile to help support me through the year (Plan A, B and C), for believing in me, and for always making sure that I know what I am capable of. Thank you so, so much.

My co-supervisor, **Prof. Izelle Smuts**, for the endless excitement towards any work performed on your patients. Thank you for fitting me into a very tight schedule and sharing your experiences (and patients) with me!

Dr. Roan Louw, for enzyme analysis made fast and easy. The help and contributions were deeply appreciated.

Dr. Etresia van Dyk and **Maryke Schoonen**, the NGS queen and bioinformatics expert. If not for you, I would still be praying for a miracle. You are lifesavers! Thank you for your willingness to always help and share your knowledge.

Mari van Reenen, for statistical advice and contributions.

Dries Sonnekus, for language editing.

The **National Research Foundation (NRF)** and **North-West University (NWU)** for their research grants.

A special thanks to the **Medical Research Council (MRC)** for making this study possible with their financial support.

My **mito lab family**, a family on a professional-yet-loving level. I will miss you dearly.

Van der Westhuizen family, especially **Venessa**, for making me feel the warmest of welcomes in your home.

My **family** and **friends**, near and far, for supporting me in some way, but special thanks to:

My loving **father** and **mother**, for supporting me for the past 23 years. Daddy cool, your interest in my work always fascinated me; you somehow always knowing more than me. Mammie blue, for the packages sent from home, filled with endless love, you knowing me better than myself. I love you both endlessly.

My **meerkat sisters**, being a part of such a fantastic three is something to be most grateful for. You were the best role models a little sister could ask for. Thank you for always sharing in my joy and always being proud of me.

The **high five, fab five**, for your friendship over the past years, the conversations shared over glasses of wine and not very well planned braais. I am so grateful for meeting you my friends; friendships that will last a lifetime. I will miss you each day, not being able to see you whenever I want. Be blessed, you helped me through so much without even realising it.

And last but not least, **Richart**, for words that can simply not begin to describe. You took the stress, the tears, the moods and the punches. You are my biggest motivator; you are my mainstay, my best friend, and the love of my life.

I thank my ***Heavenly Father*** for each and every one of you. I pray that He will repay you in ways I will never be able to. Trust in Him, “For he shall give his angels charge over thee, to keep thee in all thy ways.”- Psalm 91:11.

Abstract

CoenzymeQ_{10} (CoQ_{10}) functions as an electron carrier in mitochondria transporting electrons from CI and CII to CIII in the respiratory chain (RC) for normal cellular energy (ATP) production. Mutations in genes of a complicated and not yet well understood CoQ_{10} biosynthesis cause primary CoQ_{10} deficiency, a rare autosomal recessive mitochondrial disorder (MD) with diverse heterogeneous clinical phenotypes. Although the major function of CoQ_{10} is to serve as electron transfer molecule it furthermore possesses multiple metabolic functions which can result in secondary CoQ_{10} deficiency. Five main clinical phenotypes are associated with CoQ_{10} deficiency although distinct genotype-phenotype associations are still absent due to the limited molecular genetic diagnoses of most reported CoQ_{10} deficiency cases. A correlation was found between reduced levels of CoQ_{10} in muscle tissue and deficient CII + III RC enzyme activities in a South African patient cohort, the current indicators for potential CoQ_{10} deficiency. The aim of the study was therefore to identify nuclear-encoded mutations in genes associated with CoQ_{10} deficiencies in a cohort of South African patients diagnosed with respiratory chain deficiencies (RCDs). A high throughput target enrichment strategy was performed in order to identify previously reported and/or possible novel CoQ_{10}-assosciated disease-causing variants using Ion Torrent next generation sequencing (NGS) and an in-house developed bioinformatics pipeline. The data obtained were compared to clinical presentations of the patients to interpret the results of the identified variants considered to be possibly pathogenic. Targeted genes associated with primary CoQ_{10}- and secondary CoQ_{10} deficiency was successfully sequenced in 24 patients, identifying 16 possible disease-causing variants. Of these variants three compound heterozygous variants were identified in three patients in genes *ETFDH, COQ6* and *COQ7,* which were considered to be pathogenic according to the available data provided. Further validation of these three variants supported its pathogenicity in at least two of these variants (*ETFDH* and *COQ6*). In conclusion: This study contributed to better understanding the aetiology of a South African cohort of patients diagnosed with MDs. It also highlighted the valuable role of NGS for such investigations, and furthermore identified areas in the biochemical and molecular diagnostic strategy where improvements could be made in the future.

Keywords: CoenzymeQ_{10}; CoQ_{10} deficiency; mitochondrial disorder; respiratory chain deficiencies; next generation sequencing; bioinformatics.

Table of Contents

List of Figures

List of Tables

List of Equations

List of symbols and abbreviations

α	alpha
β	beta
µg	microgram
µl	microliter
µM	micromolar
°C	degrees Celsius
1000G	1000 Genomes project
2Fe-2S	*Rieske centre*
A	adenine
Abs	absorbance
AcCoA	acetyl-coenzymeA
ADCK3	*COQ8*: aarF domain containing kinase 3
ADCK4	aarF domain containing kinase 4
ADP	adenosine diphosphate
AFR	African
Ala	alanine
AOA1	ataxia-oculomotor-aprataxia 1
APOE	apolipoprotein E
APTX	aprataxin
Arg	arginine
Asp	aspartic acid
ATP	adenosine triphosphate
BCA	bicinchoninic acid
bp	base pare
BRAF	v-raf murine sarcoma viral oncogene homolog B
BRENDA	BRaunschweig ENzyme DAtabase
BSA	bovine serum albumin
C	carbon
C	cytosine
cDNA	complementary DNA
CI	complex I: NADH:ubiquinone oxidoreductase
CI + III	complex I + III: NADH:cytochrome *c* reductase
CII	complex II: succinate:ubiquinone oxidoreductase
CII + III	complex II + III: succinate:cytochrome *c* reductase
CIII	complex III: ubiquinol:ferricytochrome *c* oxidoreductase
CIV	complex IV: ferrocytochrome *c*:oxygen oxidoreductase
CK	creatine kinase
CoA	coenzymeA

CoQ	coenzymeQ
CoQ_{10}	coenzymeQ_{10}: ubiquinone
COQ10A	coenzyme Q10 homolog A
COQ10B	coenzyme Q10 homolog B
$CoQ_{10}H$•	semiquinone
$CoQ_{10}H_2$	ubiquinol
COQ2	decraprenyl-4OH-benzoate transferase
COQ2	coenzyme Q2 4-hydroxybenzoate polyprenyltransferase
COQ3	coenzyme Q3 methyltransferase
COQ4	coenzyme Q4
COQ5	coenzyme Q5 homolog, methyltransferase
COQ6	coenzyme Q6 monooxygenase
COQ7	coenzyme Q7 homolog, ubiquinone
COQ8	*ADCK3*: aarF domain containing kinase 3
COQ9	coenzyme Q9
CoV	co-efficient of variance
CRC	clinically referred controls
CS	citrate synthase; µmole/min citrate synthase
CSF	cerebrospinal fluid
CV	complex V: ATP phosphohydrolase
Cys	cysteine
cyt b_H	cytochrome b_H
cyt b_L	cytochrome b_L
cyt *c*	cytochrome *c*
cyt c_1	cytochrome c_1
Da	dalton
dbSNP	SNP database
DCIP	2,6-dichloroindophenol, sodium salt hydrate
ddNTP	2',3'-dideoxynucleotide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	2'-deoxynucleotide
dsDNA	double stranded DNA
DTNB	5,5'-dithio-bis[-2-nitrobenzoic acid]
DTT	dithiothreitol
e^-	electrons
e.g.	for example
EC	Enzyme Commission
EDTA	ethylene diaminetetraacetic acid
EGTA	ethylene glycol tetraacetic acid
ELC	enhanced chemiluminescence

emPCR	emulsion PCR
Eq.	equation
ES	enrichment system
ETF	field-effect transistor
ETFA	electron-transfer-flavoprotein, alpha polypeptide
ETFB	electron-transfer-flavoprotein, beta polypeptide
ETFDH	electron-transferring-flavoprotein dehydrogenase
EUR	European
FAD	flavin adenine dinucleotide
$FADH_2$	flavine adenine dinucleotide
Fe	iron
Fe^{2+}	iron-ferrous oxidation state
Fe^{3+}	iron-ferric oxidation state
FMN	flavin mononucleotide
FPP	farnesyl-pyrophosphate
g	gram
g	gravitational force
G	guanine
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
gDNA	genomic DNA
Glu	glutamic acid
Gly	glycine
GPP	geranyl pyrophosphate
H^+	hydrogen ion; proton
H_2O	water
HCl	hydrochloric acid
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HGVS	Human Genome Variation Society
His	histidine
HMG-CoA	3-hydroxy-3-methylglutaryl-coenzymeA
HPLC	high-performance liquid chromatography
HPLC-MS/MS	high-performance liquid chromatography coupled to tandem mass spectrometry
HS	high sensitivity
ID	identification
IgG	immunoglobulin G
Ile	isoleucine
IMM	inner mitochondrial membrane
IMS	inter-membrane space
In silico	performed via computer simulation
IPP	isopenetyyl pyrophosphate
ISPs	ion sphere particles

K	potassium
K_2HPO_4	dipotassium phosphate
kb	kilobyte
kDa	kilodalton
KH_2PO_4	potassium dihydrogen phosphate
KOH	potassium hydroxide
KPi-buffer	potassium-phosphate buffer
Leu	leucine
LoF	loss of function
M	molar
MADD	multiple acyl-CoA dehydrogenase deficiency
MD	mitochondrial disorder
Met	methionine
mg	milligram
min	minute
ml	millilitre
mM	millimolar
mm	millimetre
MPS	Massively Parallel Sequencing
mtDNA	mitochondrial DNA
NAD^+	nicotinamide adenine dinucleotide
NADH	reduced nicotinamide adenine dinucleotide
NaN_3	sodium azide
NaOH	sodium hydroxide
NCBI	National Center for Biotechnology Information
nDNA	nuclear DNA
ng	nanogram
NGS	next generation sequencing
nM	nanometer
nmol	nanomole
NOX	nicotinamide adenine dinucleotide oxidase
nsSNP	non-synonymous SNP
NWU	North-West University
O	oxygen
OH	hydroxyl
OMIM	Online Mendelian Inheritance in Man
OMM	outer mitochondrial membrane
OXPHOS	oxidative phosphorylation
P	patient
p	protein
PBS	phosphate buffered saline

PCR	polymerase chain reaction
PDSS1	prenyl (decaprenyl) diphosphate synthase, subunit 1
PDSS2	prenyl (decaprenyl) diphosphate synthase, subunit 2
PDVF	polyvinylidenedifluoride
PGM	Personal Genome Machine
Phe	phenylalanine
pM	picomolar
PolyPhen-2	Polymorphism Phenotyping version 2
PPS	PolyPhen-2 prediction score
Pro	proline
Q	coenzymeQ
$Q^{\bullet -}$	semiquinone
QH_2	ubiquinol
R^2	coefficient of determination
RC	respiratory chain
RCD	respiratory chain deficiency
RefSNP	reference SNP
ROS	reactive oxygen species
rRNAs	ribosomal RNAs
rs	reference number
S	sulphur
SC	supercomplex
SD	standard deviation
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel-electrophoresis
sec	seconds
Ser	serine
SIFT	sorting intolerant from tolerant
SNP	single nucleotide polymorphism
SNV	single nucleotide variation
SPS	SIFT prediction score
T	thymine
TCA	tricarboxylic acid
Ter	termination
TGS	triglycine sulfate
TGX	Tris-Glycine eXtended
Thr	threonine
TNB	5-dithio-bis-2-nitrobenzoic acid
Tris.HCl	tris-hydrochloric acid (2-amino-2-hydroxymethyl)-1,3-diol
Triton X-100	octylphenolpoly(ethylene-glycoether)n: C24
tRNAs	transfer RNAs

Trp	tryptophan
Tween 20	polyoxyethylene (20) sorbitan monolaurate
UCII	µmole/min complex II
UCP	uncoupling protein
UV	ultraviolet
v	version
V	volts
v/v	volume per total volume
v_1	initial velocity value
Val	valine
VCF	variant caller file
VEP	variant effect predictor
WB	western blot
w/v	weight per volume

Chapter 1

1 Introduction

CoenzymeQ_{10} (CoQ_{10}) is the predominant form of coenzymeQ (CoQ) in humans. It fulfils the central role of electron carrier in the mitochondrial respiratory chain (RC) transporting electrons between CI and CII to CIII, which leads to the formation of proton motive force resulting in the production of cellular energy (ATP) (Crane *et al.,* 1985; Ernster & Dallner, 1995). CoQ_{10} deficiency is a mitochondrial disorder (MD), a term that refers to a group of deficiencies categorized by RC impairment that constrain ATP production (Chi, 2014). The complex biosynthesis of CoQ_{10}, and the number of functions attributed to $CoQ_{10,}$ is believed to be the cause of the wide variety of clinical syndromes associated with CoQ_{10} deficiency (Quinzii *et al.,* 2007, Bentinger *et al.,* 2010). The clinical representations of CoQ_{10} deficiency results in encephalomyopathy, severe multi-system infantile disease, cerebellar ataxia, Leigh syndrome and isolated myopathies (Ogasahara *et al.,* 1989; Rötig *et al.,* 2000; Musumeci *et al.,* 2001; Van Maldergem *et al.,* 2002; Lamperti *et al.,* 2003; Lalani *et al.,* 2005; Quinzzi *et al.,* 2007; Horvath *et al.,* 2012).

Primary CoQ_{10} deficiency is an autosomal recessive disorder. Mutations in any of the genes involved in CoQ_{10} biosynthesis may cause CoQ_{10} deficiency and, despite the identification of these genes, the number of reported patients with a genetically diagnosed CoQ_{10} deficiency is still relatively low. Since no distinct genotype-phenotype correlation is known, it contributes to the difficulty in making a genetic diagnosis. Decreased CoQ_{10} levels in muscle (Miles *et al.,* 2008) and combined complex deficiency of CI + III and/or CII + III of the RC serve as good indicators (Ogasahara *et al.,* 1989) to pursue genetic testing.

In a study performed on a South African patient cohort diagnosed with RC deficiency (RCD) by Smuts *et al.* (2010) it was found that patients who presented with a predominant myopathic phenotype had a combined CII + CIII enzyme deficiency. Subsequently Wilsenach (2014) performed a study measuring the muscle CoQ_{10} levels in this cohort of patients and, as expected, the majority of patients with a combined CII + III deficiency had reduced levels of CoQ_{10}. The prevalence of nuclear-encoded CoQ_{10}-associated gene mutations in South African RCD patients is unknown and this study was initiated to investigate the genetic basis of CoQ_{10} deficiencies in a cohort of patients with RCDs.

The aim of this study was therefore to identify mutations in nuclear-encoded genes associated with CoQ_{10} deficiencies in a cohort of South African patients diagnosed with RCDs. This was accomplished by firstly re-evaluating CII + III activities in the selected cases with CII + III

deficiencies from the cohort, followed by a high throughput target enrichment sequence strategy to identify known or possible novel CoQ_{10}-assosciated gene variants[1]. The data obtained were then compared to available clinical presentations of these patients to ultimately interpret the results and to identify the cases where a pathogenic variant was considered the likely cause of disease. Additional follow-up validation strategies including Sanger sequence and immunoblotting analysis for protein structure validation was also included in order to support the pathogenicity of identified variants.

This study aims to contribute towards gaining insight into the role of genetic background and variation in MDs and ultimately improve the complex diagnostic process of patients with RCDs in South Africa.

.

[1] See Chapter 2, Section 2.11 for nomenclature on genetic variation.

Chapter 2

Literature Overview

2.1 Introduction

Chapter 1 presented the background and rationale of this study, which also described the correlation between CII + III deficiency and reduced levels of CoQ_{10} in patients with RCDs. This chapter gives an overview of the mitochondrion's structure and properties of which the focus will be on CoQ_{10} and its central role in the RC energy production, CoQ_{10} biosynthesis, the variety of functions of CoQ_{10}, and associated CoQ_{10} deficiencies that have been reported to occur in humans. This overview will form the basis and motivation behind the study including the methods used to accomplish this study. The problem statement, aim, objectives and experimental strategy of this study are also given in this chapter.

2.2 The mitochondrion

Mitochondria are small, highly organised double membrane organelles that produce most of the cellular ATP needed to drive numerous energy-requiring processes (Taanman, 1999). The shape and number of mitochondria per cell differ greatly and depend on the particular tissue and energy requirements of the cell (Henze & Martin, 2003; McBride *et al.,* 2006). The cellular energy production system known as oxidative phosphorylation (OXPHOS) depends on respiratory CI-V (complex I: NADH:ubiquinone oxidoreductase; complex II: succinate:ubiquinone oxidoreductase; complex III: ubiquinol:ferricytochrome *c* oxidoreductase; complex IV: ferrocytochrome *c*:oxygen oxidoreductase) and complex V (ATP phosphohydrolase), all of which are embedded in the inner mitochondrial membrane (IMM) (Anderson *et al.*, 1981; Endo *et al.*, Krauss, 2001; Henze & Martin, 2003; 2011; Marín-García, 2013). The mitochondrion contains about 1000-1500 proteins involved in a plethora of functions (Endo *et al.,* 2011), which could be considered to mainly revolve around energy metabolism to assure normal cellular function, of which four of the five OXPHOS enzyme complexes (CI, CIII, CIV and CV) are encoded by genes from both the mitochondrial DNA (mtDNA) and nuclear DNA (nDNA) (Anderson *et al.,* 1981) as will be further discussed in Section 2.5.

2.2.1 Mitochondrion structure

Mitochondria are enclosed by two membranes which in themselves are involved in the key functions of the mitochondrion. Both these membranes are composed of phospholipid bilayers and vary in permeability, which is likely achieved by the formation of membrane-spanning pores through which the intermembrane space proteins are released (Chipuk *et al.,* 2006). The outer

mitochondrial membrane (OMM) is very permeable, while the IMM is strictly selective. The OMM contains mostly porin proteins which form large pores, making it permeable to molecules smaller than 10,000 Da. The IMM is much more resistant and will only allow the free movement of water, oxygen and carbon dioxide over it. Special transporter proteins are responsible for the movement of ions, substrates and fatty acids over the IMM (Herrmann & Neupert, 2000). The IMM is folded inwards in a highly organised manner and are called cristae. This increases the surface area of the inner membrane carrying the key enzymatic machinery of OXPHOS (Krauss, 2001). Membranes separate the matrix from the cytosolic environment, dividing the mitochondrion into two compartments: i) the inter-membrane space (IMS), which is situated between the two membranes, and ii) the matrix, which is enclosed by the IMM (Mannella, 2006; Marín-García, 2013).

The OXPHOS complexes are arranged throughout the IMM, together with two mobile carriers (coenzymes), namely cytochrome *c* (cyt *c*) and CoQ, which carries electrons and protons during OXPHOS (Hauss *et al.,* 2005). Although the mitochondrial RC complexes are usually described in literature as single working units connected by the two mobile carriers, it is proposed that the different respiratory complexes are assembled into higher order supramolecular structures called supercomplexes (SCs). These SCs are said to contain multiple copies of certain complexes which serve to increase the rate and efficiency of electron transfer (Schägger, 2002; Schägger & Pfeiffer, 2000). Although a number of functions have been discussed to be associated with SCs no definite explanation have been formulated and, therefore, the existence and role of SCs remains undecided and debatable (Lapuente-Brun *et al.,* 2013).

2.2.2 Cellular energy production

The most recognized function of mitochondria is the production of energy needed to drive normal cellular function. Mitochondrial respiration is driven by a chain of chronologically organized redox reactions of carbon substrates obtained from food such as carbohydrates, fatty acids and amino acids (Spinazzola & Zeviani, 2009) into energy releasing ATP.

Energy metabolism starts in the cytosol of the cell with the catabolism (breakdown of complex molecules to less complex forms) of proteins to amino acids, polysaccharides to monosaccharides, and lipids to fatty acids and glycerol. Further metabolisation via the glycolysis pathway, amino acid metabolism and β-oxidation follows respectively. Glucose is oxidised to pyruvate molecules through a series of enzymatic reactions occurring in the glycolysis pathway, entering the mitochondrion through the double membrane. When ATP is needed, the pyruvate molecules are metabolised into acetyl-coenzymeA (AcCoA) molecules, which are oxidized in the tricarboxylic acid (TCA) cycle (Neustadt & Pieczenik, 2008). Fatty acid and amino acid metabolism also provide AcCoA molecules that enter the TCA cycle which, in turn, provides

reduced electron carriers namely reduced nicotinamide adenine dinucleotide (NADH) and reduced flavine adenine dinucleotide ($FADH_2$) molecules needed to drive the electron transport system.

These reduced electron carriers enter CI and CII in the RC respectively and produce ATP with the help of shuttle molecules, CoQ and cyt *c* (DiMauro & Schon, 2003; Neustadt & Pieczenik, 2008; Nicholls & Ferguson, 2002).

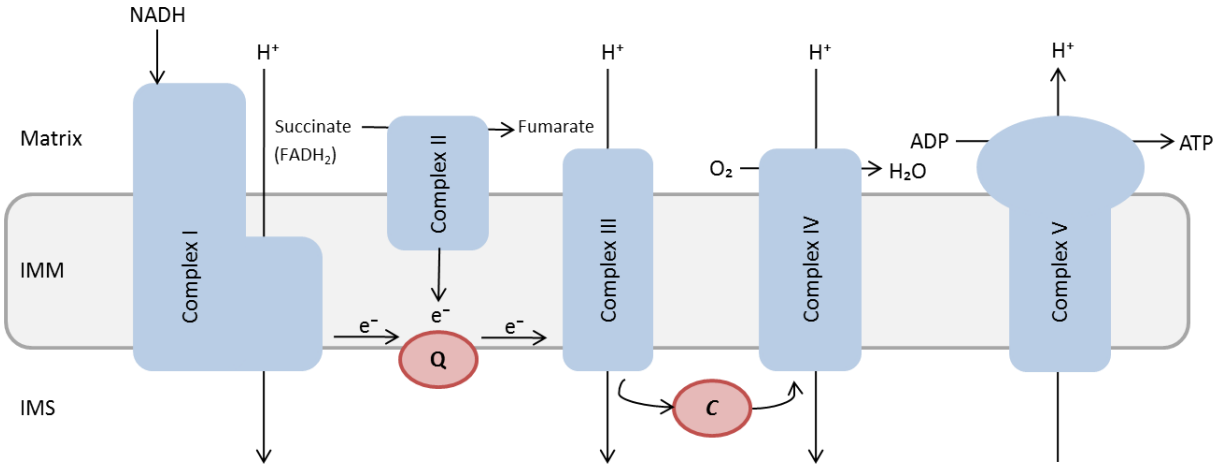

Figure 2-1: Schematic presentation of the respiratory chain, showing the enzymes of the OXPHOS system. Electrons are delivered to the RC by NADH (at CI) and succinate in the form of $FADH_2$ (at CII) and carried to molecular oxygen via RC enzyme complexes and coenzymes. As electrons (e^-) flow along the respiratory chain, protons (H^+) are pumped from the inner mitochondrial membrane (IMM) into intermembrane space (IMS) through complexes I, III, and IV, creating a proton gradient. These protons then flow back into the matrix via complex V, producing ATP from ADP. NADH: reduced nicotinamide adenine dinucleotide; $FADH_2$: flavine adenine dinucleotide; Q: coenzymeQ; *c*: cytochrome *c*; O_2: oxygen; H_2O: water; ADP: adenosine diphosphate; ATP: adenosine triphosphate (Adapted from DiMauro & Schon, 2003).

As is illustrated in Figure 2-1 NADH and $FADH_2$ correspondingly donate electrons to the first two complexes in the RC. During a series of redox reactions the electron-transporting complexes pass electrons derived from NADH and $FADH_2$ onto a final recipient, oxygen, via protein-bound redox centres consequently forming water (Krauss, 2001). The energy generated during these redox reactions is used by CI, CIII and CIV to transport protons (hydrogen ions, H^+) across the IMM into the IMS. This creates a proton and electrochemical gradient across the IMM. The potential energy generated is transformed into free energy during the movement of protons through CV back into the matrix, producing energy during the phosphorylation of ADP to ATP (Mitchell, 1961), as will be further discussed in detail in Section 2.4.

2.3 CoenzymeQ_{10}

CoQ was discovered as a result of a long ongoing investigation into the mechanism and compounds involved in cellular energy production (Crane, 2007). CoQ belongs to a series of molecules, also known as quinones, which are similar in chemical structure but differ in lengths of the isoprenoid side chain (El-Najjar *et al.,* 2011), of which CoQ_{10} is the predominant form of CoQ in humans (Fischer *et al.,* 2011). CoQ_{10}, also known as ubiquinone (oxidised form), owed to its occurrence in every cell in all tissues, forms part of the IMM of the RC (Ernster & Dullner, 1995). In 1958 the complex structure was determined by Wolf *et al.* (1985) and consists of an ubiquinone head group attached to a 10 isoprenoid unit side chain as presented in Figure 2-2 (Turunen *et al.,* 2004; Fischer *et al.,* 2011). Its hydrophilic benzoquinone ring and hydrophobic isoprenoid side chain makes this complex amphipathic, which contributes to the position of CoQ_{10} in the IMM (Lenaz *et al.,* 2007).

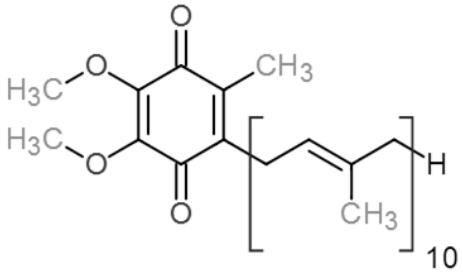

Figure 2-2: The chemical structure of CoQ_{10}. The predominant human form of CoQ consisting of a ubiquinone head attached to a 10 isoprenoid unit side chain (Adapted from Berg *et al.,* 2002).

2.3.1 CoQ_{10} Biosynthesis

The biosynthesis of CoQ_{10} in mammalian cells are extremely complicated and involves the combination of the i) benzoquinone ring, predominantly derived from the amino acid tyrosine (or phenylalanine), and the ii) synthesis of the 10-decraprenyl side chain from AcCoA through the mevalonate pathway (Ernster & Dullner, 1995; Tran & Clarke, 2007), as illustrated in Figure 2-3.

The amino acid, tyrosine, is readily available from dietary intake to form the aromatic benzoquinone structure but can also be derived from phenylalanine. Tyrosine/phenylalanine is converted to 4-OH-benzoate (Ernster & Dullner, 1995) through a series of reactions. The synthesis of the isoprenoid side chain through the mevalonate pathway starts with three molecules AcCoA condensed to form HMG-CoA (3-hydroxy-3-methylglutaryl-coenzymeA). This step is performed by two enzymes namely acetoacetyl CoA thiolase and HMG-CoA synthase. Mevalonate is then formed from HMG-CoA through HMG-CoA reductase, following phosphorylation through two reactions by mevalonate kinase and phosphomevalonate kinase,

respectively. The product, mevalonate pyrophosphate, delivers isopenetyl pyrophosphate (IPP), the key structure of the synthesis of the polyisopropenoid side chain of CoQ_{10}, which also serves as the precursor for farnesyl-pyrophosphate (FPP) (Turunen *et al.,* 2004). FPP is condensed with a number of IPP molecules by trans-prenyl-transferase (COQ1), using geranyl pyrophosphate (GPP) (formed as intermediary) as an enzyme bound intermediate to ultimately deliver the long polyisoprenoid side chain, namely decaprenyl pyrophosphate (Tran & Clarke, 2007; Turunen *et al.,* 2004).

Condensation of the decraphenyl pyrophosphate chain into the 4-OH-benzoate ring through catalisation of decraprenyl-4OH-benzoate transferase (COQ2) produces decarprenyl-4OH-benzoate. After the condensation the benzoate ring undergoes change by C-hydroxylation, decarboxylation and O-methylation to ultimately synthesize CoQ_{10} (Bentinger *et al.,* 2010; Tran & Clarke, 2007). No less than ten nuclear encoded genes (*COQ2-COQ10*, *PDSS1* and *PDSS2*)[2] have been identified to be required in the biosynthesis of CoQ_{10} (Tran & Clarke, 2007; Quinzzi *et al.,* 2008). Even though the biosynthesis of CoQ_{10} is not completely defined the most detailed information of the pathway is provided by the study of yeast and bacteria. In humans, the CoQ_{10} biosynthesis pathway and its necessary enzymes encoded by the required genes remains inconclusive and only defined to an extent (Desbats *et al.,* 2014; Trevisson *et al.,* 2011; Potgieter *et al.,* 2013).

[2] Throughout this dissertation, italic gene symbols refer to genes, while non-italic gene symbols describe the protein encoded by the indicated gene. The official gene symbols were obtained from HUGO Gene Nomenclature Committee (HGNC) (http://www.genenames.org/).

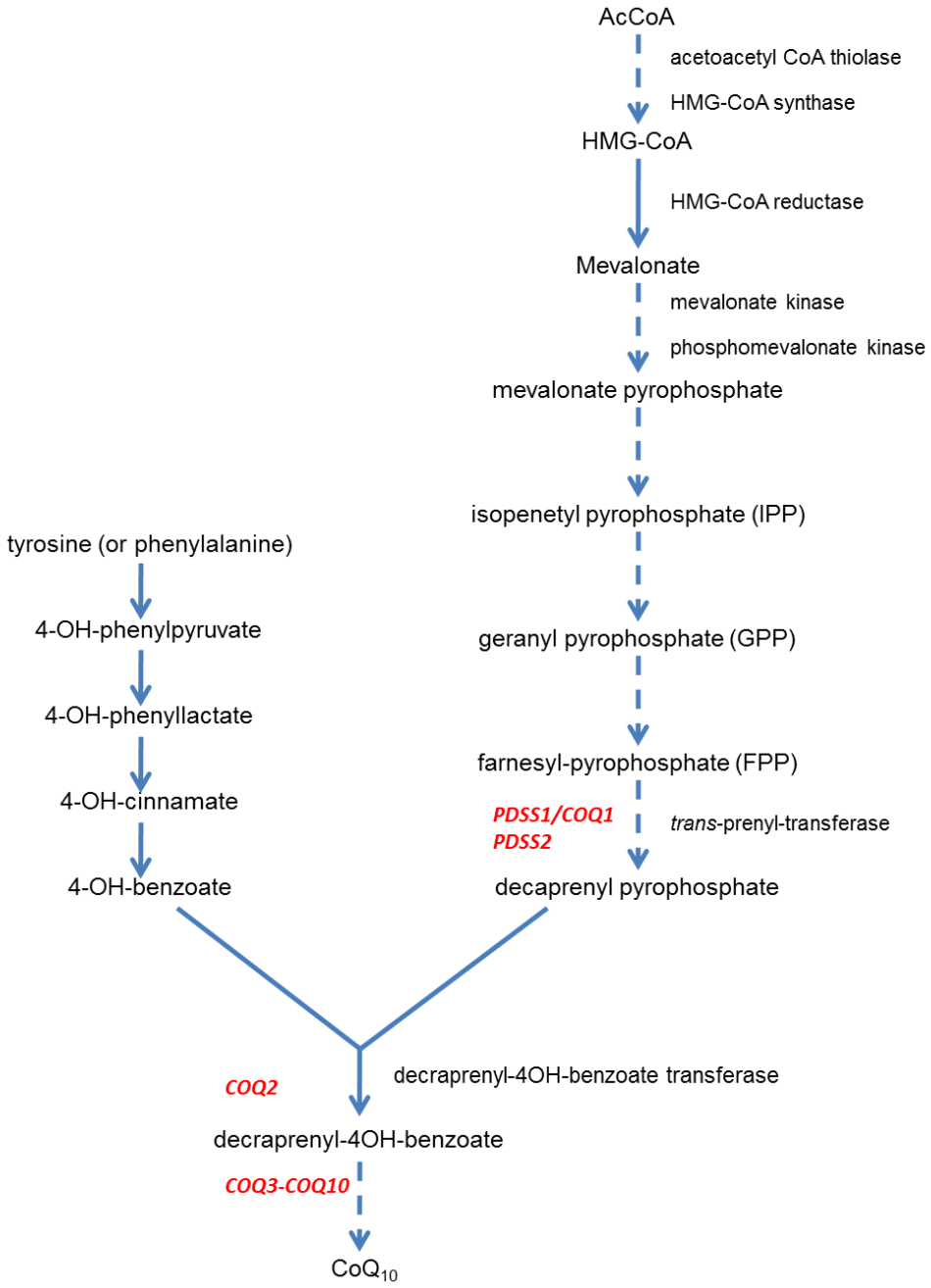

Figure 2-3: Schematic representation of the pathways and enzymes involved in CoQ_{10} biosynthesis. CoQ_{10} is synthesized in the mitochondrial inner membrane where at least 12 genes are involved. The enzymatic reaction pathway starts with the mevalonate pathway, with acetyl-CoenzymeA (AcCoA) as its first substrate, and ends with farnesylpyrophosphate after a series of reactions. The terminal reactions of the CoQ_{10} biosynthesis involve the condensation of 4-OH-benzoate derived from the amino acids tyrosine or phenylalanine, and decaprenyl pyrophosphate from the mevalonate pathway. After condensation, enzymes reactions ring follow to form the functional CoQ_{10} product. Genes involved in the biosynthesis of CoQ_{10} are indicated in red (Adapted from Turunen *et al.,* 2004; Ernster & Dullner, 1995; Quinzzi *et al.,* 2008; Potgieter *et al.,* 2013). CoA: coenzymeA; HMG-CoA: 3-hydroxy-3-methylglutaryl-coenzymeA; OH: hydroxyl.

2.3.2 CoQ_{10} function

CoQ_{10} has many functions in addition to its central role in the mitochondrial RC as electron carrier from CI and CII to CIII (Quinzii *et al.,* 2007). These other functions of CoQ_{10} have been the focus of investigation for the past 20 years (Bentinger *et al.,* 2010) and is summarized in Table 2-1. Other than just experimental observations, it should be noted that some of these functions are still just probably associated with *Homo sapiens*, and are still under further investigation.

Table 2-1: Functions of CoQ_{10} (constructed from Bentinger *et al.,* 2010).

Electron transport in RC (Turunen *et al.,* 2004; Mitchell, 1975)
Mitochondrial regulation of NAD^+/NADH ratio (Gomez-Diaz *et al.,* 1997)
Serves as antioxidant (Ernster & Dallner, 1995)
Regulation of mitochondrial transition pore opening (Papucci *et al.,* 2003)
Essential for regulation of mitochondrial uncoupling proteins (Echtay *et al.,* 2001)
Exerts manifold anti-inflammatory properties (Doring *et al.,* 2007)
Possesses anti-atherosclerotic effects (Turunen *et al.,* 2002)
Modulate endothelial function (Hamilton *et al.,* 2007)
Mediates oxidation of sulphide in yeast (Bentinger *et al.,* 2010; Saiki *et al.,* 2003)

The mitochondrial RC is a principal cellular source of free radicals formed by reactive oxygen species (ROS), generated by the loss of electrons spilling from the RC complexes (Papa & Skulachev, 1997). Studies reveal that reduced CoQ_{10} (ubiquinol; $CoQ_{10}H_2$) acts as an outstanding antioxidant by scavenging free radicals that prevent lipid peroxyl radical formation. An even greater significance regarding $CoQ_{10}H_2$ function is the regeneration of the antioxidant α-tocopherol (Vitamin E), and the control of ascorbate levels (Ernster & Dallner, 1995; Frei *et al.,* 1989). Another function ascribed to CoQ involves NADH oxidase (NOX), a protein involved in electron transfer over the mitochondrial membrane which is situated in the external surface of the plasma membrane (DeHann *et al.,* 1997). The reduced form of NOX is depended on CoQ for regulation of the cytosolic NAD^+/NADH ratio, contributing to the before mentioned ascorbate maintenance, as well as cell growth and differentiation (Gomez-Diaz *et al.,* 1997; Crane *et al.,* 1985).

As mentioned in Section 2.2.1 the OMM contains large pores that make it permeable to molecules smaller than 10,000 Da (Herrmann & Neupert, 2000). The much more resistant IMM can undergo increased permeability by mitochondrial permeability transition pores, causing adverse effects such as depolarisation of the membrane gradient, which ultimately leads to ATP

depletion when uncontrolled. CoQ is known to counter these undesirable effects, which include depolarisation of the proton gradient, liberation of cytochrome *c* and ATP depletion, by affecting these permeability transition pores of the IMM (Papucci *et al.,* 2003). CoQ is also involved in the regulation of uncoupling proteins (UCPs) which are positioned in the IMM. UCPs are in charge of translocating H^+ from the outside to the inside of the mitochondrion matrix and suppresses oxygen radicals by uncoupling the proton gradient from the OXPHOS system by generating heat rather than ATP (Echtay *et al.,* 2001).

It is said that the effectiveness of CoQ as electron carrier in the RC, as well as its function as antioxidant, is not affect by the length of the isoprenoid side chain (Turunen *et al.,* 2004). The activation of UCPs and the regulation of permeability transition pores on the other hand do indeed have requirements when CoQ species with isoprenoid side chain of different lengths are involved (Echtay *et al.,* 2001; Walter *et al.,* 2000).

2.4 CoQ_{10} redox cycle

The redox active benzoquinone ring of CoQ_{10} makes three oxidation states possible to undergo redox cycling. The fully oxidized state, ubiquinone (CoQ_{10}), possesses two keto groups. An intermediate is formed with the addition of an electron and proton, called a semiquinone ($CoQ_{10}H•$). A fully reduced form is reached through the addition of another electron and proton, called ubiquinol ($CoQ_{10}H_2$) (Berg *et al.,* 2002; Mathews *et al.,* 2000). The three oxidation states of CoQ_{10} are illustrated in Figure 2-4.

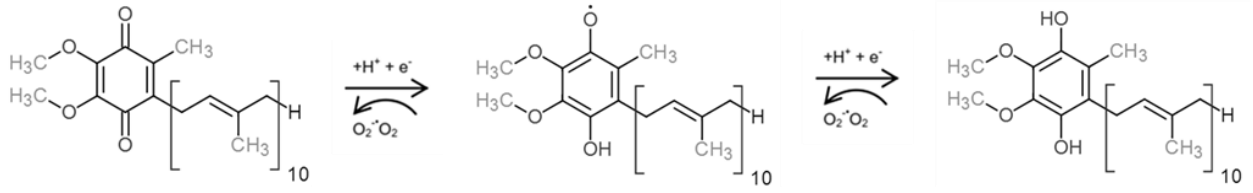

Figure 2-4: Illustration of the CoQ_{10} redox cycle. The reduction of ubiquinone (CoQ_{10}) through a semiquinone intermediate ($CoQ_{10}H•$), to ubiquinol ($CoQ_{10}H_2$). H^+: proton, e^-: electron; O: oxygen (Adapted from Berg *et al.,* 2002).

2.4.1 Complex I: NADH:ubiquinone oxidoreductase (EC 1.6.5.3) [3]

CI serves as the entry point for electrons into the RC as mentioned in Section 2.2.2. NADH transfers two electrons to the catalyser of the reaction, a flavin mononucleotide (FMN) prosthetic group, delivering the reduced form of $FMNH_2$, while being oxidised to NAD^+. The electrons

[3] The enzyme names and Enzyme Commission (EC) numbers were obtained from BRENDA (http://www.brenda-enzymes.info/).

generated are transported from reduced $FMNH_2$ to the iron-sulphur centres of CI (Berg *et al.,* 2002; Garret & Grisham, 2010). Iron ions in the iron-sulphur centres undergo oxidation-reduction reactions between the reduced form (Fe^{2+}) and oxidised (Fe^{3+}) state (Mathews *et al.,* 2000). Two electrons are transferred to CoQ_{10} from the iron-sulphur centres through CI, pumping four final hydrogen ions from the matrix to the cytosol, while reducing CoQ_{10} to $CoQ_{10}H_2$ (Berg *et al.,* 2002; Garret & Grisham, 2010; Mathews *et al.,* 2000). This reaction catalysed by CI is represented in Equation 2-1 (Berg *et al.,* 2002).

$$\mathrm{NADH} + \mathrm{CoQ_{10}} + 5\mathrm{H^+(matrix)} \rightarrow \mathrm{NAD^+} + \mathrm{CoQ_{10}H_2} + 4\mathrm{H^+(cytosol)} \qquad \textbf{(Eq.2-1)}$$

2.4.2 Complex II: succinate: ubiquinone oxidoreductase (EC 1.3.5.1)

Once oxidation of succinate to fumarate occurs in the TCA cycle $FADH_2$ is generated from the flavoprotein subunit (FAD) in CII. $FADH_2$ stays in the complex while transferring electrons to the iron-sulphur centre, and thereafter to CoQ_{10}, following incorporation into the RC. This reaction forms the reduced state of CoQ_{10}, $CoQ_{10}H_2$, as represented in Equation 2-2. No protons are pumped from CII across the IMM (Berg *et al.,* 2002; Garret & Grisham, 2010).

$$\mathrm{succinate} + \mathrm{CoQ_{10}} \rightarrow \mathrm{fumarate} + \mathrm{CoQ_{10}H_2} + 2\mathrm{H^+} \qquad \textbf{(Eq.2-2)}$$

2.4.3 Complex III: ubiquinol: ferricytochrome c oxidoreductase (EC 1.10.2.2)

Cytochrome is a one electron carrier protein that contains a heme prosthetic group. During electron transport the iron ion of the cytochrome undergoes oxidoreduction cycles between reduced ferrous (+2) state and oxidised ferric (+3) state. CIII ultimately pump protons out of the mitochondrial matrix by catalyzing the transfer of electrons from $CoQ_{10}H_2$ to oxidised cytochrome *c,* as illustrated in Equation 2-3 (Berg *et al.,* 2002).

$$\mathrm{CoQ_{10}H_2} + 2\ \mathrm{Cyt}\ c\ (\mathrm{ox})\ 2\mathrm{H^+(matrix)} \rightarrow \mathrm{CoQ_{10}} + 2\ \mathrm{Cyt}\ c\ (\mathrm{red})\ 4\mathrm{H^+(cytosol)} \qquad \textbf{(Eq.2-3)}$$

CIII contains three hemes: a low infinity heme b_L and high infinity heme b_H within the cytochrome *b* subunit, as well as *c*-type heme within the cytochrome c_1 subunit, giving CIII the alternative name cytochrome bc_1. CIII also contain a 2Fe-2S iron-sulphur protein named *Rieske centre,* as well as two binding sites for CoQ_{10} known as Q_0 and Q_i, the latter positioned closer to the inside of the matrix (Berg *et al.,* 2002; Mathews *et al.,* 2000).

2.4.4 The Q cycle

The process in which the coupling of electrons is transferred from CoQ_{10} to cytochrome *c,* to the pumping of electrons across the IMM, is a severely complicated system known as the *Q cycle* (Figure 2-5). $CoQ_{10}H_2$ binds in the Q_0 location transferring one electron to the *Rieske* 2Fe-2S

cluster, which is directly transferred to cytochrome c_1 and thereafter to cytochrome *c*. A second electron is initially transferred to cytochrome b_L before being transferred to cytochrome b_H, and ultimately bound in the Q_0 location to an oxidised ubiquinone. The reduction of oxidised CoQ_{10} leads to $CoQ_{10}H•$ formation, releasing $2H^+$ protons into the IMS. The process is repeated when a second molecule of $CoQ_{10}H_2$ binds to the Q_0 location proceeding the same way as first described, except that, this time, the second electron is transferred through cytochrome b_L and b_H to $CoQ_{10}H•$ bound in the Q_i position. $CoQ_{10}H•$ receives $2H^+$ protons from the mitochondrial matrix and becomes oxidised to $CoQ_{10}H_2$, contribution to the proton gradient formation. The *Q cycle* ultimately produces two molecules CoQ_{10} from the oxidation of two molecules $CoQ_{10}H_2$. It also produces reduced $CoQ_{10}H_2$ from one molecule CoQ_{10} and two reduced molecules cytochrome *c*. Four protons are ultimately released to the cytoplasmic side of the mitochondrion, as well as two protons removed from the mitochondrial matrix (Berg *et al.,* 2002; Mathews *et al.,* 2000).

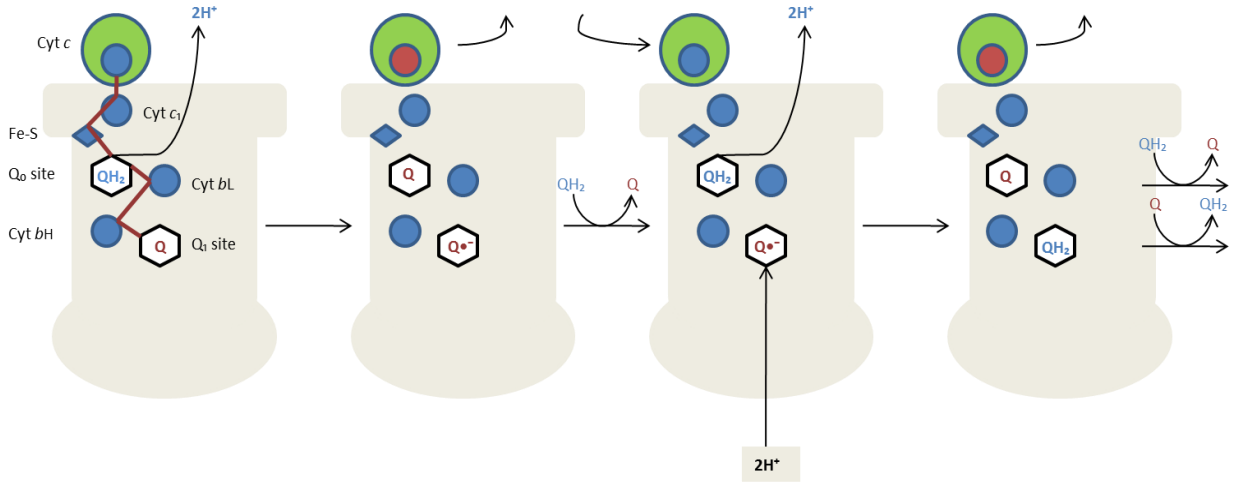

Figure 2-5: An illustration of the Q Cycle. The mechanism involves the connection of electron transfer from CoQ_{10} (Q) to cytochrome *c* (cyt *c),* leading to transmembrane proton export. $CoQ_{10}H•$ (Q•-) is formed when an electron is transferred from $CoQ_{10}H_2$ (QH_2), while a second electron from QH_2 is transferred to cyt *c*. The formed Q is separated and replaced by a second QH_2, donating an electron to a second cyt *c* molecule, as well as the reduction of Q•- to QH_2 via another electron donation, resulting in the uptake of two protons from the matrix. Oxidized forms are indicated in blue font and reduced forms in red font. Fe-S: *Rieske centre*; cyt c_1: cytochrome c_1; cyt b_L: cytochrome b_L; cyt b_H: cytochrome b_H; H^+: protons (Adapted from Berg *et al.,* 2002).

2.4.5 Complex IV: ferrocytochrome c: oxygen oxidoreductase (EC 1.9.3.1)

CIV is the terminal and final electron acceptor in the RC reducing O_2 to two H_2O molecules by oxidation of the reduced cytochrome *c* produced by CIII. When fully oxidised, CIII initiate the catalytic cycle transferring electrons to oxygen which eventually lead to water (H_2O) formation.

The protons needed for this reaction are provided from the matrix and once the H_2O is released from the enzyme, the initial state is once again regenerated (Berg *et al.,* 2002).

2.5 Mitochondrial disorder: a complex genetic disorder

There exist multiple copies of 16,567 base pair circular double-stranded mtDNA within the mitochondrial matrix of each human cell (Chu *et al.,* 2012). mtDNA consists of 37 genes that encode for 13 polypeptides, 22 transfer RNAs (tRNAs) and two ribosomal RNAs (rRNAs) of which the 13 polypeptides form part of the OXPHOS system (seven mtDNA encoded subunits in CI, one in CIII, three in CIV and two in CV). Each of the RC complexes are also encoded by nuclear genes assembled together with the mtDNA-encoded subunits found in the IMM (Anderson *et al.,* 1981; Scarpulla, 1997). Primary MD is the consequence of dysfunctional mitochondrial respiration due to mutations in the mtDNA and nDNA (which includes CoQ_{10} deficiency), while secondary MD is acquired through external mechanisms that influence mitochondrial function (Filler *et al.,* 2014). It is no surprise that the effects are so devastating due to abnormal cellular function in which mitochondria play a crucial part. Disorders from nDNA related mutations are normally inherited in the Mendelian pattern, while mtDNA are inherited maternally (Sue & Sohon, 2000).

A wide variety of clinical and biochemical are associated with MD of which CoQ_{10} has been implicated in mitochondrial involvement and have now been established as being associated with more “common” diseases such as neurodegenerative diseases, cancer, diabetes and ageing (Desbats *et al.,* 2014; Finkel & Holbrook, 2000; Gogvadze *et al.,* 2008; Zsurka & Kunz, 2013).

2.6 CoQ_{10} deficiency as a clinical disorder

Primary CoQ_{10} deficiency is a autosomal recessive disorder affecting the biosynthesis pathway of CoQ_{10} due to mutations in *COQ* genes (Quinzzi *et al*., 2007), while secondary CoQ_{10} deficiency is caused by non-genetic factors such as inadequate dietary consumption, extreme endogenous use of CoQ_{10} (Quinzzi & Hirano, 2011; Emmanuele *et al*., 2012; Potgieter *et al*., 2013) or defects in genes unrelated to CoQ_{10} biosynthesis e.g. other mitochondrial myopathies, mitochondrial DNA depletion syndrome, glutaric aciduria type II, etc (Desbats *et al.,* 2014*)*. It is however, still unclear if CoQ_{10} deficiency is mainly caused by the impairment of the RC, or rather due to the impairment of the other functions attributed to $CoQ_{10,}$ such as serving as anti-oxidant for free radical scavenging (Quinzzi & Hirano, 2010; Horvath, 2012).

CoQ_{10} deficiency is determined by measuring CoQ_{10} in muscle and/or fibroblasts. It was not until 2006 when the first genetic diagnosis was made concerning CoQ_{10} disease even though the

clinical data dates back to Ogasaharas' initial discovery in 1989 (Ogasahara *et al.,* 1989). To date primary CoQ_{10} deficiencies have been identified in eight COQ_{10} biosynthesis genes (*PDSS1, PDSS2, COQ2, COQ4, COQ6, ADCK3, ADCK4,* and *COQ9*) and secondary CoQ_{10} deficiencies are much more common. Although most probably out-dated the main five clinical phenotypes of CoQ_{10} deficiency have been defined as:

i) Encephalomyopathy with exercise intolerance, myopathy, regular myoglobinuria, seizures, ataxia and raged red fibres (Ogasahara *et al.,* 1989);
ii) Severe multi-system infantile disease with encephalopathy, cardiomyopathy, ataxia, optic nerve atrophy, deafness and nephrotic syndrome (Rötig *et al.,* 2000);
iii) Predominant cerebellar ataxia and cerebellar atrophy (Musumeci *et al.,* 2001; Lamperti *et al.,* 2003);
iv) Leigh syndrome with growth retardation, ataxia and deafness (Van Maldergem *et al.,* 2002), and
v) Isolated myopathy (Lalani *et al.,* 2005; Quinzzi *et al.,* 2007; Horvath *et al.,* 2012).

2.6.1 The encephalomyopathy phenotype

The first ever CoQ_{10} deficient patients with the encephalomyopathy phenotype were described by Ogasahara *et al*. (1989). This clinical profile presented with myopathy, myoglubinuria, seizures and mental retardation, together with a biochemical profile of elevated creatine kinase (CK), lactic acidosis and deficient levels of muscle CoQ_{10}, and decreased combined enzyme activities CI + III and CII + III (Ogasahara *et al.,* 1989). A few reports have been associated with this clinical phenotype since 1989 and in a single case was associated with a mutation in the *ADCK3* gene (Aure *et al.,* 2004). Both the cases, reported by Ogashara *et al.* and Aure *et al.,* presented with normal or elevated CI, CII, CIII and CIV enzyme activities, as well as normal levels of CoQ_{10} when measured in fibroblast and serum (Ogasahara *et al.,* 1989; Aure *et al.,* 2004).

2.6.2 The multisystem infantile disease phenotype

Rötig *et al.* (2000) was the first to describe a multi-systemic infantile variant of CoQ_{10} deficiency. Three siblings presented with neurological symptoms such as nystagmus, optic atrophyataxia, dystonia, weakness, sensorineural hearing loss and progressive nephropathy. Quinzzi *et al.* reported two siblings in 2006 with the multi-system infantile disease with severe nephrotic syndrome which is not usually associated with other MDs. A homozygous missense mutation in the *COQ2* gene encoding para-hydroxybenzoate-polyprenyltransferase was detected as the cause of the disease (Quinzzi *et al.,* 2006). Another two siblings harbouring base-pair deletions in the *COQ2* gene with the related phenotype was reported, and a further two siblings from a

family with multysistemic disease including deafness, encephaloneuropathy, obesity, livedo reticularis and cardiac valvulopathy were identified with a mutation in the *PDSS1* gene (Mollet *et al.,* 2007). Further CoQ_{10} deficiencies reported with this specific phenotype have also been identified in the mutated *PSDD2* gene presenting nephrotic syndrome together with Leigh syndrome (López *et al.,* 2006), as well as compound heterozygous *COQ6* gene mutations (Heeringa *et al.,* 2011).

2.6.3 The predominant cerebellar ataxia phenotype

Cerebellar ataxia is the most regular phenotype associated with CoQ_{10} deficiencies with recognition as early as the first CoQ_{10} deficiency (Horvath *et al.,* 2012). Cerebellar ataxia is also mostly accompanied by cerebellar atrophy, seizures, development delay, mental retardation and muscle weakness (Musumeci *et al.,* 2001; Lamperti *et al.,* 2003; Horvath *et al.,* 2012). Deficient muscle and fibroblast CoQ_{10} levels have been reported in a few cerebellar ataxia cases and generally occurred in childhood and adolescences except for a reported case by Gironi *et al.* (2004) describing late-onset cerebellar ataxia with hypogonadism. Ataxia-oculomotor-aprataxia 1 (AOA1) is a disease caused by a homozygous stop codon mutation in the *APTX* gene and have been associated with the cerebellar ataxia phenotype (Quinzzi *et al.,* 2010). As the *APTX* gene is involved in the encoding of the protein aprataxin that repairs DNA strands breaks, and the precise interaction between *APTX* and the biosynthesis of CoQ_{10} is unclear, this disorder is placed in the secondary CoQ_{10} deficiency category (Horvath *et al.,* 2012). Mutations have been identified in the *ADCK3/COQ8* gene in both patients with mild (adulthood) to severe cerebellar ataxia (childhood), presenting with clinical profiles which included spasticity, dystonia, tremor and migraine in these specific gene mutations (Mollet *et al.,* 2007; Horvath *et al.,* 2012).

2.6.4 Leigh syndrome with growth retardation, ataxia and deafness

It has been reported that two sisters represented childhood onset Leigh disease, growth retardation, infantilism, ataxia, deafness and lactic acidosis (Van Maldergem *et al.,* 2002). Another patient who presented Leigh syndrome, including neonatal liver disease, pancreatic insufficiency, tyrosinemia, hyperammonemia, subsequent sensorineural hearing loss and reduced combined complex activity (CI + III; CII + III) indicating a CoQ_{10} deficiency, was reported in 2003 by Leshinsky-Silver and colleagues. The responsible molecular defect in both these cases is, however, not known and can be confused with other phenotypes (Horvath *et al.,* 2012; Leshinsky-Silver *et al.,* 2003).

2.6.5 The isolated myopathy phenotype

Multiple acyl-CoA dehydrogenase deficiency (MADD) is an autosomal recessive disorder of the fatty acid and amino acid metabolism due to an *ETFDH* gene mutation (and in fewer common cases the *ETFA* and *ETFB* genes) that encodes for a component of the electron-transfer system in mitochondria essential for electron transfer from a number of mitochondrial flavin-containing dehydrogenases to the main RC. Mutations in the *ETFDH* gene known to cause MADD is associated with an isolated myopathy clinical phenotype of secondary CoQ_{10} deficiency (Gempel *et al.,* 2007; Olsen *et al.,* 2007). These deficiencies have been reported by Gempel *et al.* (2007) to be homozygous or compound heterozygous deficiencies and expanded on the clinical features first found in the *ETFDH* gene mutation in three patients, which included exercise intolerance, proximal myopathy, increased serum CK and fatigue.

Salviati *et al.* (2012) identified a patient with mental retardation, encephalomyopathy and dysmorphic features caused by haploinsufficiency of the *COQ4* gene that encodes a protein for the biosynthesis of CoQ_{10}. This clinical phenotype was dissimilar from the other phenotypes described, but appears to be most comparable with the original CoQ_{10} deficient patients discovered by Ogasahara *et al*. (1989). It is, however, noteworthy that the majority of reported CoQ_{10} deficient patients still require the precise location and nature of defects in the CoQ_{10} biosynthesis which have thus far not yet been identified (Quinzii *et al.,* 2007).

2.7 CoQ_{10} deficiency treatment

There is currently no effective treatment for OXPHOS deficiencies. However, unlike other RC deficiencies, patients with CoQ_{10} deficiency have developed improved clinical profiles with oral supplementation of exogenous CoQ_{10}, thus being potentially treatable (Quinzii *et al.,* 2007). Individuals with primary or secondary CoQ_{10} deficiency have shown to benefit from CoQ_{10} supplementation (DiMauro *et al.,* 2007). Since dietary uptake of CoQ_{10} is very limited, numerous efforts are ongoing on improving the bioavailability of oral administration, focused mainly towards preparations in forms that are absorbed more effectively, or by substitutes which possess similar functional properties to that of CoQ_{10} (Bentinger *et al.,* 2010; Kapoor & Kapoor, 2013). It is also noticeable that CoQ_{10} treatments tend to not have severe side effects and being low of cost (Itkonen *et al.,* 2013).

Little is known about the absorption and metabolism of CoQ_{10} in different human organs and it is expected that the characterization of CoQ_{10} biosynthesis and regulation will support understanding of CoQ_{10} metabolism and its advantageous use in clinical therapies (Bhagavan & Chopra, 2006; Ernster & Dullner, 1995; Kapoor & Kapoor, 2013; Tran & Clarke, 2007). Short and long chain CoQ_{10} supplements have been available to date for treating patients with the

deferent clinical profiles. Although the oxidized form, CoQ_{10}, was mainly used, it was not until recently that a new and stable form, $CoQ_{10}H_2$, was formulated. This isoform appears to be more stable than CoQ_{10} (Horvath *et al.,* 2012). Salviati *et al.* (2012) reported the first case of a $CoQ_{10}H_2$ treated patient found with the first *COQ4* deficiency. It is, however, necessary to compare the efficiency of $CoQ_{10}H_2$ vs. CoQ_{10} treatments in different CoQ_{10} deficient patients. This will allow treatment optimisation in favour of the patients to benefit from CoQ_{10} supplementation, since different responses to CoQ_{10} supplementation was observed in patients with different genetic mutations for reasons which are not yet known (López *et al.,* 2010).

2.8 Diagnosis of CoQ_{10} deficiency

There is currently no single guiding principle for biochemical and molecular evaluation of suspected MDs including CoQ_{10} deficiency. Multiple biochemical and molecular approaches are necessary to diagnose complex genome MDs (Haas *et al.,* 2008). Clinical assessment is usually the first suspicion for MDs but the verification via biochemical and molecular evaluation is necessary to confirm the diagnosis at hand (Haas *et al.,* 2007).

2.8.1 Biochemical Analysis

There are no clear diagnostic metabolites for RCDs. Analysis of urinary, blood or cerebrospinal fluid (CSF) metabolites such as alanine and other amino acids, lactate, TCA cycle intermediates ethylmalonic acid and 3-methyl glutaconic acid, as well as other organic acids provide persistent indicators for RCDs. Specific biomarkers such as increase in lactate:pyruvate ratio and TCA cycle intermediates have been implicated although it is important to note that there are no specific ranges of abnormal values used to identify MDs (Haas *et al.,* 2008). More recently, metabolomics investigations have provided indications that an urine metabolic “biosignature” (combination of metabolites) may exist for RCDs (Reinecke *et al.,* 2012; Smuts *et al.,* 2013, Venter *et al.,* 2014), but still needs further development.

2.8.2 Enzyme assays

The activities of the different RC complexes measured by spectrophotometric assays are the primary data used for the diagnosis of RCDs. Muscle biopsies are usually used for tissue sample due to the majority of MD patients having skeletal muscle involvements as well as muscle being rich of mitochondria (Rodenburg, 2011). The measurements of complex activities is based on absorbance change of the substrates present in the reaction depending on the complex being assayed (Wong, 2013a) and will be described in detail in Section 3.3.

2.8.3 CoQ_{10} measurements

As mentioned in Section 2.3.2 CoQ_{10} mediate electron transport in the RC contributing to ATP synthesis (Mathews *et al.,* 2000; Lenaz *et al.,* 2007; Ernster *et al.,* 1969). This is why a link between decreased levels of CoQ_{10} and CoQ_{10} deficiency is expected, especially together with a combined CI + III and CII + III deficiency due to impaired electron transfer (Miles *et al.,* 2008).

Miles and colleagues (2008) recognized that measuring the total CoQ_{10} content, rather than the oxidized and reduced levels of CoQ_{10}, presented the best prediction of RC enzyme abnormality. This was indeed recently confirmed by Itkonen, Suomalainen and Turpeinen through a study of mitochondrial CoQ_{10} determination in patients with RCDs (Itkonen *et al.,* 2013). The most used methods of CoQ_{10} determination have been high-performance liquid chromatography (HPLC) methods coupled with electronchemical detection (ECD) (Miles *et al.,* 2008), ultra violet (UV) detection (El-Najjar *et al.,* 2011) and HPLC coupled to tandem mass spectrometry (HPLC-MS/MS) (Itkonen *et al.,* 2013).

Clinical and biochemical characterization is very important in diagnostic procedures, as it contributes to the selection of candidate genes for further genetic investigation to ultimately achieve precise diagnosis.

2.9 Molecular genetic studies

Molecular defects can be detected through various available methods. The traditional approach to molecular analysis of defects caused by nuclear genes is based on identifiable clinical symptoms followed by screening for mutations, sequencing of specific known candidate genes one at a time or (more recently) in combination using next generation sequencing (NGS) technology. The use of exome sequencing has also become widely used in major diagnostic centres. Clinical and biochemical evaluations assist in narrowing down possible candidate genes to a smaller selected group to be sequenced. For instance, if biochemical enzyme activity analyses of mitochondrial RC complexes show deficient CI + III and CII + III activities, genes encoding for CoQ_{10} biosynthesis should be selected for analysis. Different genes may cause related clinical manifestations but the same gene may cause a diverse clinical variety, perplexing the choice of candidate genes for sequence analysis (Wong, 2013b).

As this study will focus largely on molecular genetic investigations a more detailed overview of the sequencing approaches will be given in the following sections to support the experimental strategy that was used in this study.

2.9.1 Sanger sequencing

The gold standard and traditional approach of sequencing is Sanger (also called dideoxy- or chain-termination-) sequencing. The Sanger method involves the synthesis of a complementary DNA template using natural 2'-deoxynucleotides (dNTPs) and termination of synthesis using 2',3'-dideoxynucleotides (ddNTPs) by DNA polymerase. A set of fragments occur as a result of competitive synthesis and termination of synthesis by means of the fragments differing in nucleoside monophosphate units. The DNA sequence is then revealed by separating the fragments by size through high-resolution gel electrophoresis or capillary chromatography. Sanger sequencing detects dye-labelled fragments through laser produced fluorescence emissions by tagging the primer or the terminating ddNTP with specific fluorescent dyes. This results in revealing the DNA sequence, automated of four different colours assigned to a specific base (Metzker, 2005; Sanger *et al.,* 1977). Sanger sequencing is used as the standard sequencing procedure when involvement of specific loci are suspected, when a limited number of samples are to be sequenced, and when "deep sequencing" is not required (i.e. basic allele frequency to determine homo- or heterozygosity).

2.9.2 Next Generation Sequencing

In the recent decade NGS or Massively Parallel Sequencing (MPS) has basically transformed genomics research by undergoing a paradigm shift that allows molecular diagnostics to be performed in fast technical and affordable ways (Voelkerding *et al.,* 2009). Since 2005, when the first NGS system was integrated, NGS has paved the way to clinical diagnosis by generating a large amount of interpretable data to ultimately discover genetic variants to cause rare complex diseases (Pabinger *et al.,* 2014). There are different NGS platforms, every method with its individual advantages and disadvantages, that uses its own distinctive sequencing chemistries and machine hardware configurations (Margulies *et al*., 2005; Rothberg *et al*., 2011; Shendure *et al.,* 2005; Wong, 2013b; Zhang *et al*., 2014). All factors, such as cost per base, coverage fold and simplicity of data analysis, should be taken into account when selecting the most suitable platform for a sequencing project.

NGS relies on basic principles using high throughput parallel sequencing by amplification of single DNA molecules of targeted genes (usually through a polymerase-based clonal replication process) before being divided onto a solid matrix known as library preparation. This enriched target genes are then ultimately sequenced via repeated sequencing chemistries (Meldrum *et al.,* 2011).

Target gene enrichment selectively enrich the coding regions of the gene of interest utilising a method of preference depending on the number and size of the genes or the total size of the

targeted genes in megabases. PCR/multiplex PCR (polymerase chain reaction) methods that make use of gene sequence specific primers are useful for gene coding region enrichment if the total size of the target region is lower than a few hundred kilobyte (kb). In the case of a very large total size of target genes the synthesis of thousands of primers for PCR is impractical due to the fact that PCR conditions are tough to optimise (Meldrum *et al.,* 2011; Wong, 2013b). There are alternatives available to overcome the obstacle involving large scale primer designing for targeted genes, as will be presented in Section 3.6, as used in this study.

Exome sequencing has become a very popular approach in the field of genetic diagnosis in single gene disorders by eliminating high cost and excessive data produced from whole-genome sequencing through only sequencing the ~1% of the genome that codes for specific protein sequences. The amplification of a specific target region from the complete genome can easily be accomplished, which contributes greatly to clinical application in understanding human health (Meldrum *et al.,* 2011; Pabinger *et al.,* 2014).

2.10 Ion Torrent Personal Genome Machine

The Ion Torrent Personal Genome Machine (PGM) for NGS was available, suitable and thus used during this study. This instrument enables high throughput sequencing of targeted nuclear genes in a time- and cost-effective manner. The Ion Torrent platform developed by Life Technologies is the first ever platform to eliminate cost and complexity of optical detection by making use of 'natural' or unmodified dNTPs as a replacement for colour detected dNTPs (Ross & Cronin, 2011).

2.10.1 Technology of the Ion Torrent

Ion Torrent sequencing relies on the basic principle of "sequencing-by-synthesis", through detecting chemical changes within a reaction on a semiconductor chip with each reaction connected to its own sensor or field-effect transistor (ETF). Each well on the Ion Torrent sequencing chip holds about a million copies of DNA fragments. During sequencing the PGM sequencer floods the chip with one dNTP after another. As standard DNA polymerase sequencing occur H^+ is released from the elongating 3'OH end of the complementary DNA strand when dNTPs are incorporated. Each Ion sensor, working as the world's smallest pH-meter, detects this H^+ concentration pH change by creating a recordable voltage change (chemical signal to a digital signal) (Niedringhaus *et al.,* 2011; Merriman *et al.,* 2012). The reaction which takes place on the Ion Torrent sequencing chip is illustrated in Figure 2-6, while Chapter 3.6 will provide a detailed discussion of the sequencing method and materials used during this study.

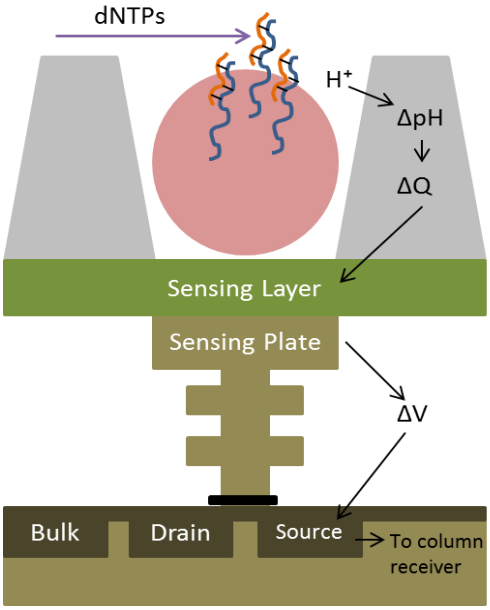

Figure 2-6: A schematic presentation of the Ion Torrent sequencing chip. Illustrated is a single well containing an Ion Sphere particle amplified with a DNA template. When a deoxynucleotides (dNTP) is incorporated, the hydrogen ion (H^+) released changes the pH of the well. This pH change (ΔpH) is detected by the sensing layer/plate and translates the chemical signal to digital information. ΔQ; charge change, ΔV; voltage change (Adapted from Ion Torrent via Life Technologies, Application note number C022802, 2011).

2.11 Classification of variation

The following section gives a brief description of the nomenclature used in this dissertation. The terminology to describe genetic variants can sometimes be confusing because the nomenclature used by different laboratories varies. Before the standardisation of nomenclature systems by the Human Genome Variation Society (HGVS) documentation to support human genetic variation was easily miss-communicated. First and foremost, as suggested by Den Dunnen and Antonarakis (2000), in this dissertation terms such as mutations and polymorphism are not used in order to avoid confusion. In this dissertation the frequently used term 'variant' implies a small genetic variation identified in genetic sequence, either with a benign or deleterious genetic effect. The following standards used for terminology of the variants identified are briefly described.

- Nucleotide substitutions are expressed as e.g. c.41G>A, where the prefix 'c' refers to the sequencing type, complementary DNA (cDNA), or the standard reference referring to genomic (gDNA) as used for sequence in this study, and the symbol greater than (>) indicates the base substitution of the 41st nucleotide G to A.

- Protein substitutions are expressed as e.g. p.Ala49Ser, where the prefix ‘p’ indicates a protein change, followed by the codon number of the alanine to serine amino acid change caused by the variant.
- More than one variant on an allele in the same gene of a patient are called compound heterozygous variants, and are expressed as e.g. c.[41G>A] + [859G>T] and p.[Trp14Ter] + [Ala49Ser].
- The reference SNP (RefSNP) refers to the database of SNPs that have previously been uniquely identified and reported and have an applied reference number (rs). The rs number found in the NCBI dbSNP (database for SNP variants) (http://www.ncbi.nml.nih.gov//snp/) provide the frequency of the SNP detected in specified populations, providing further details if available, such as clinical phenotypes associated with the specific SNP (Ogino *et al.,* 2007; Goodeve *et al.,* 2011).

2.12 Problem statement

In a study conducted by Smuts *et al*. (2010) different ethnical groups of South African MD patients were described of which there were a remarkable diversity in the clinical manifestations regarding the ethnicities of the patients. In this study the group of African patients primarily presented muscular involvement (62.5%), whereas the other ethnic groups which included Caucasian, mixed ancestry and Indian patients, presented encephalomyopathy (30%) as well as central nervous system (7.5%) associations. Imperative to this study was the significant finding of the distinction between the African patients and the other ethnicities concerning the combined RC CI + III and CII + III enzyme deficiencies. A noteworthy 52.5% of the African patients were diagnosed with these specific combined RC complex deficiencies, with combined CII + III deficiency being predominant in the African patient group.

These findings lead to the initiation of an additional study with the assumption that CoQ_{10} depletion could be a main attributing factor of the combined RC complex (CI + III and CII + III) deficiencies in these patients, given CoQ_{10} being the electron carrier from CI and CII to CIII (Lenaz *et al.,* 2007; Quinzii *et al.,* 2007; Smuts *et al.,* 2010). The study (Wilsenach, 2014) prior to this study investigated the levels of CoQ_{10} in muscle tissue of the entire cohort of suspected RCD patients, which included confirmed and unconfirmed RC enzyme deficiencies. Optimal sample preparation and extraction of CoQ_{10} from the quadriceps muscle samples, using an optimized method involving HPLC-MS/MS for the quantification of the total CoQ_{10}, was performed and reported in the dissertation of Wilsenach (2014). A very accurate set of data was obtained and a positive correlation established between combined CII + III deficiencies and reduced levels of CoQ_{10} in the study cohort of South African patients with RCDs. The study

presented here follows up on these findings by investigating the mutations that could be causative in these CoQ_{10} deficiency cases.

Mutations in any of the genes involved in CoQ_{10} biosynthesis may be the cause of CoQ_{10} deficiency in these patients, or alternatively a result of secondary CoQ_{10} deficiency. Before this study no information existed on the prevalence of nuclear-encoded CoQ_{10}-associated gene mutations in South African patients and since the hallmark of CoQ_{10} deficiency is reduced levels of CoQ_{10} in muscle and/or fibroblast together with combined RC complex deficiencies (Montero *et al.,* 2008), it was clear that this needed to be investigated. Based on biochemical data and relevant literature, the question arises whether nuclear-encoded CoQ_{10} biosynthesis genes are most likely affecting South African patients with combined CII + III deficiencies and reduced levels of CoQ_{10}.

2.13 Aim, objectives and experimental strategy

In an effort to confirm if CoQ_{10} quantification could possibly be used to aid in the diagnosis of mitochondrial CoQ_{10} deficiency, and whether the expected reduced levels of CoQ_{10} is the cause or even a main factor of the patients' deficiency, the presence of nuclear-encoded CoQ_{10}-associated gene mutations in the cases with CoQ_{10} deficiency in the South African RCD study cohort needed to be investigated.

Figure 2-7 illustrates the strategy formulated to address the *aim* of this study which was *to identify nuclear-encoded mutations in genes associated with CoQ_{10} deficiencies in a cohort of South African patients diagnosed with RCDs.* The following *objectives* were formulated to achieve this aim:

i) to confirm combined CII + III deficiency via repeated analysis of spectrophotometric enzyme activity assays on selected patient cohort and patient controls (CRC);
ii) generation of DNA sequence data of nuclear-encoded CoQ_{10}-associated genes using Ion Ampliseq custom panel design for targeted gene enrichment and the Ion Torrent PGM platform for NGS;
iii) identification of previously reported and potential novel pathogenic DNA variants by using a bioinformatics workflow;
iv) comparison of clinical profiles of patients found with CoQ_{10}-associated disease-causing variants for genotype-phenotype correlation; and finally
v) Sanger sequence validation and protein structure analysis for validation purposes on selected candidate pathogenic variants using sodium dodecyl

sulphate polyacrylamide gel-electrophoresis (SDS-PAGE) and western blotting analysis.

The strategy designed for these objectives is illustrated in Figure 2-7, on which the main investigations enzymology, molecular genetic testing, bioinformatics, clinical profile comparison and Sanger sequence and protein structure analysis were conducted. The design is based on the overall study concerning extended clinical and molecular genetics data advancing knowledge regarding South African paediatric patients. The selection of the patient case and control groups, as illustrated in Figure 2-7, will be described in detail in Section 3.2. Enzyme activity assays were performed in order to confirm CII + III deficiency on the selected cohort as *first objective*. The highlighted part illustrates the focus of this study which consists of molecular genetic testing of targeted CoQ_{10}-associated genes using NGS as *second objective*, followed by a bioinformatics workflow identifying possible pathogenic variants as *third objective*. Comparing clinical profiles of the patients identified with possible disease-causing variants served as *fourth objective*, while Sanger sequence validation and immunoblotting analysis for protein structure validation on candidate pathogenic variants we performed as *fifth objective*. Detailed methods used in the illustrated strategy (Figure 2-7) are represented in Chapter 3 followed by interpretation of the results and detailed discussion in Chapters 4 and 5, respectively.

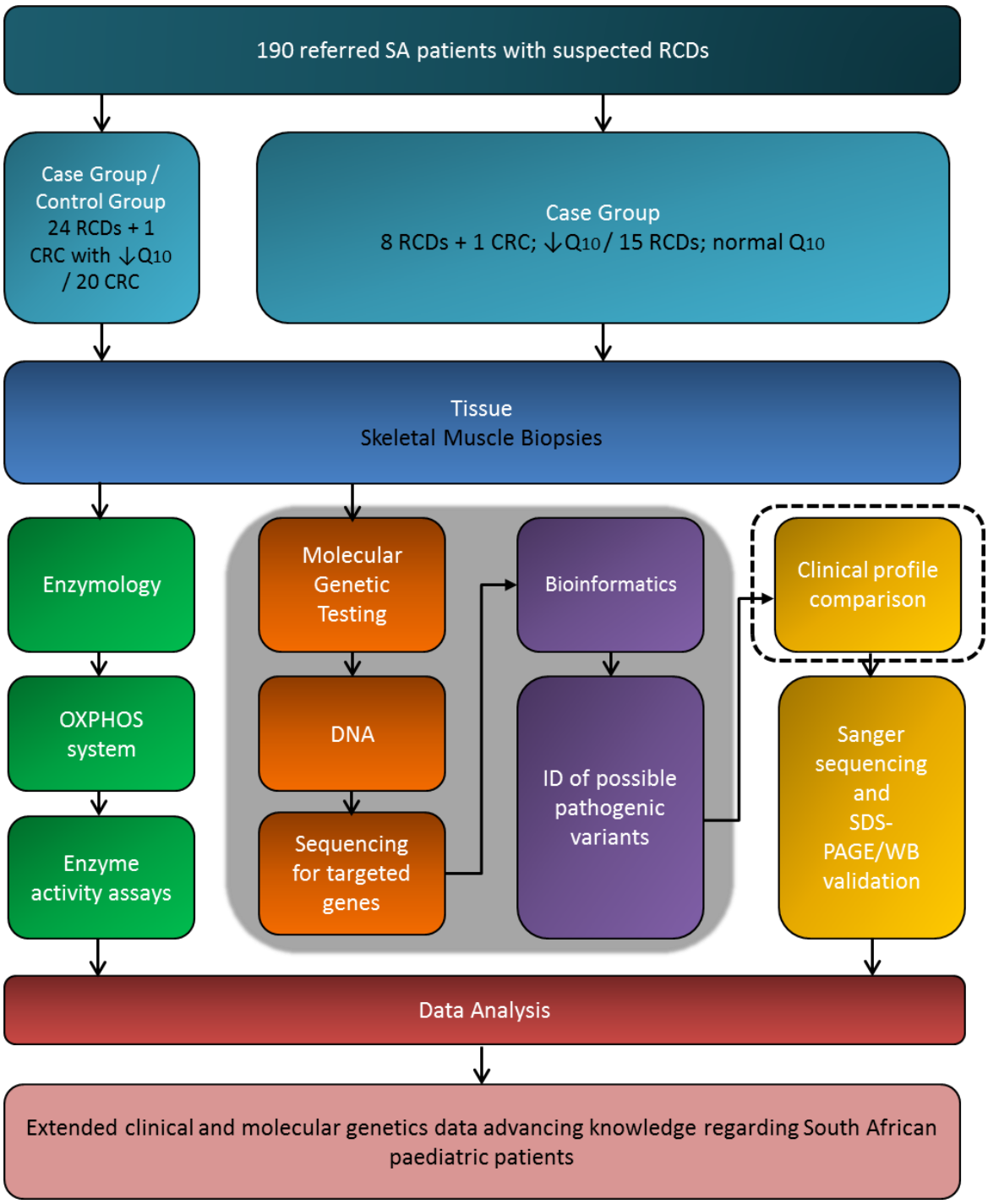

Figure 2-7: Schematic representation of the experimental strategy used to achieve the aims and objectives, forming part of a larger comprehensive investigation. This study consisted of five objectives, of which the two highlighted grey parts present the main focus of the study. The dotted lines resemble the part completed from previous studies that were used to complete the final interpretations of the results obtained in this study. SA: South African; RCDs: respiratory chain deficiencies; CRC: clinically referred controls; ↓: reduced; Q_{10}: coenzymeQ_{10}; ID: identification; SDS-PAGE: sodium dodecyl sulphate polyacrylamide gel-electrophoresis WB: western blot.

Chapter 3

Materials and Methods

3.1 Introduction

This chapter provides further detail of the study design, materials, as well as the methods and technologies that were used to conduct this study. As discussed in Chapter 2 diagnostic confirmation of CoQ_{10} deficiencies requires a variety of investigations including enzyme activity analysis, molecular genetic testing and (to confirm novel pathogenic variants) often also protein structural investigations. The patients selected for this study form part of a comprehensive study of South African patients referred to the Paediatric Neurology Unit at Steve Biko Academic Hospital, Pretoria, South Africa, with suspected MDs over the past 17 years (1997-2013). Since 2006 until 2013 190 muscle biopsies were performed and analysed according to an optimised protocol for the RC enzyme deficiencies. These analyses were performed at the Mitochondria Research Laboratory, Centre for Human Metabonomics, North-West University (NWU). A targeted molecular genetics investigation of a specific group of genes has been used in this study. For this purpose an Ion Torrent PGM NGS platform was available for genetic analyses, followed by a bioinformatics workflow for gene variant identification. Sanger sequencing and immunoblotting protein structure analysis were also included for validation purposes. The methods used to produce the key information of this study are described in full detail in this chapter. As this study is part of a larger multi-disciplinary study the contributions of collaborators are also indicated in this chapter.

3.2 Ethics and patients

Ethical approval was acquired by Prof. I. Smuts (co-supervisor), Department of Paediatrics, Faculty of Health Sciences, University of Pretoria, for the collection of muscle biopsies from all patients used in this study (approval number 91/98 with amendments *Mitochondrial disorders in the South African context*). Ethical approval under project number NW-00170-13-S_1 was also obtained from the Ethics Committee of the NWU for the work described here, which was performed at the NWU.

For this study a total of 25 patients were selected from the larger cohort of 190 patients of whom all but one patient (P54) had been diagnosed enzymatically with a combined CII + III RC deficiency in muscle tissue. Specific attention were given to these combined CII + III deficient patients during the study by Wilsenach (2014) in which a positive effect between combined CII + III deficiency and reduced CoQ_{10} levels were identified. This data were used for the selection of

patients in this study, as mentioned in Section 2.12. Of these 25 selected patients nine patients had reduced CoQ_{10} levels and 16 patients normal CoQ_{10} levels. This selected cohort consisted of four (16%) Caucasian males, eight (32%) African males, three (12%) Caucasian females and 10 (40%) African females. Age at onset of symptoms for these patients ranged from neonatal period to adulthood.

In this study, enzyme activity analysis to confirm a combined complex CII + III deficiency was repeated on muscle tissue homogenates of all 25 patients. Nine patients were therefore cases with reduced CoQ_{10} levels previously diagnosed with CII + III deficiency (except for P54 who had no enzyme deficiency but was included due to reduced CoQ_{10} levels measured) and 16 patients were cases with previous CII + III enzymatic diagnosis but normal CoQ_{10} levels. A patient group that presented strong MD phenotype features, but with no formerly detected enzyme deficiency, was also included and is referred to in this study as *clinically referred controls* (CRC). These CRC samples served as the reference samples to set the enzyme activity reference range against which the 25 case samples could be compared (Section 3.4). The enzymatic analyses thus also included 20 CRC samples consisting of five CRCs of each race and gender.

Nuclear-encoded CoQ_{10}-associated gene mutations were further investigated on 24 of the 25 selected patients (P72 excluded due to insufficient sample). These 24 samples included the nine patients with low levels of CoQ_{10} ($<5^{th}$ percentile of the CRC as reported by Wilsenach, 2014) which can be considered the test group in this study, as well as 15 patients with CoQ_{10} $>5^{th}$ percentile of the CRC (Wilsenach, 2014), which could be considered a control group where no pathogenic variants were expected in the genetic analyses.

For validation purposes candidate pathogenic variants were confirmed using Sanger sequencing and immunoblotting protein structure analyses, using SDS-PAGE and western blot analysis, were conducted. These validation analyses were performed on three patients (P2, P43 and P78) where possible compound heterozygous variants from the sequence data were found. The results conducted from these analyses were evaluated and ultimately interpreted along with the clinical data which were previously compiled by Smuts and colleagues at the Steve Biko Academic Hospital. The results and discussions will be presented in Chapters 4 and 5.

3.3 Citrate synthase and respiratory chain enzyme analyses

These analyses were performed based on previous published methods (Shepherd & Garland, 1969; Rahman *et al.,* 1996; Janssen *et al.,* 2007) using kinetic spectrophotometry assay which was modified for diagnostic use at the Mitochondria Research Laboratory, NWU. These protocols make use of a Synergy HT microplate reader (BioTek Instruments, Winooski, United

States) and Gene5 Data Analysis software (BioTek Instruments, v1.05). The protocols make use of microtiter 96-well plates which allows analyses of 96 samples simultaneously, allowing analyses to be done in triplicate and in an extremely shorter period of time eliminating batch-to-batch variation for increased repeatability.

3.3.1 Materials

The following materials and reagents were used: 5,5'-dithio-bis[-2-nitrobenzoic acid] (DTNB) (Roche Diagnostics, Basel, Switzerland, 104 477[4]); Trisma (Sigma-Aldrich, St. Louis, United States, T8655); Triton X-100 (Sigma-Aldrich, T9284); acetyl-CoA, trilithium salt (Roche Diagnostics, 10 101 907 001); oxaloacetate (Sigma-Aldrich, 17,125-5); potassium dihydrogen phosphate (KH_2PO_4) (Fluka Analytical via Sigma-Aldrich, St. Louis, United States, 60218); dipotassium phosphate (K_2HPO_4) (Fluka, 60353); fatty acid-free bovine serum albumin (BSA) (Roche Diagnostics, 775 835); EDTA.K (Fluka, O3660); ATP, disodium salt (Roche Diagnostics, 10 519 979 001); 2,6-dichloroindophenol, sodium salt hydrate (DCIP) (Sigma-Aldrich, D1878); decylubiquinone (Sigma-Aldrich, D7911); dimethyl sulfoxide (DMSO) (Sigma-Aldrich, C6164); rotenone (Sigma-Aldrich, R8875); ethanol (Merck Chemicals, Darmstadt, Germany, 1.00983.2500); antimycin A (Sigma-Aldrich, A8674); succinate (Sigma-Aldrich, S7903); sodium azide (NaN_3) (Sigma-Aldrich, S8032); cytochrome *c* (Sigma-Aldrich, C7752) ; Milli-Q prepared water.

Instrumentation[5] used throughout enzyme activity assays: Synergy HT Multi-detection microplate reader (BioTek) with Gene5 Data Analysis software (BioTek Instruments, v1.05).

3.3.2 Buffers, solutions and specific reagents preparation

A 0.5 M potassium-phosphate buffer (KPi-buffer, pH 7.4) consisting of 95 mM KH_2PO_4 and 405 mM K_2HPO_4 was used in CII and CII + III analyses. This buffer was diluted to 50 mM with water. For the citrate synthase analysis, DTNB was dissolved in 1 M Tris.HCl (pH 8.0). Rotenone and

[4] The numbers in brackets in this chapter indicate the catalogue numbers of the reagents purchased from the relevant providers.

[5] Throughout this study, unless specified otherwise, the following consumables were used: 8-Strip PCR Tubes and Caps (Roche Diagnostics, 11667009001); Eppendorf Safe-Lock and LoBind microcentrifuge tubes (various volumes, Sigma-Aldrich); Eppendorf epT.I.P.S (various volumes, Sigma-Aldrich), Costar 96-Well polystyrene standard microplates (Cole-Parmer Instrument Company, London, United Kingdom). In this dissertation PBS always refers to 1 x PBS solution (1.06 mM KH_2PO_4, 155 mM NaCl, 2.97 mM Na_2HPO_4-$7H_2O$, pH 7.4).

antimycine A were dissolved in ethanol; BSA was dissolved in 0.5 M KPi-buffer, and dicylubiquinone dissolved in DMSO. Reduced cytochrome *c* and reduced ubiquinol were prepared from the oxidised forms as described by Du Toit (2007) and Luo *et al.* (2008) for routine work at the Mitochondria Research Laboratory, NWU. Where necessary pH values were adjusted using either potassium hydroxide (KOH) or hydrochloric acid (HCl).

3.3.3 Methods

The assay conditions used during the enzyme analyses are indicated below. The biological material that was used, was so-called "600 *g* supernatants" previously prepared from muscle biopsies as follows: In short, the muscle tissue from all the samples were homogenized at 10% (w/v) in an isotonic homogenization buffer (EGTA, 0.1 mM, pH 7.2; HEPES, 5 mM; mannitol, 210 mM; sucrose, 70 mM), by using a motor-driven Potter-Elvehjem glass-teflon homogenizer. Homogenates were centrifuged at 600 x *g* for 10 min and the supernatant used for further analysis. Before storage at -80°C for further use, protein content analyses were done on these samples using the bicinchoninic acid (BCA) method as originally described by Smith *et al.* (1985). In this study this analysis was repeated in selected cases to test its accuracy, especially as accuracy matters a lot for SDS-PAGE analysis. Substrates used, depending on the enzyme being assayed, were added in abundance (typically >10 x Km) to the reactions and either substrate use or product formation monitored by measurement of the absorbance change by the reduction or oxidation reactions of the specific compounds as described in more detail below.

3.3.3.1 *CS Activity*

Citrate synthase serve as mitochondrial marker enzyme occurring at relative constant levels in the mitochondrion (Shepherd & Garland, 1969). Information obtained from a publication by Shepherd and Garland (1969) describes the CS measurements. The reaction, in a final volume of 200 µl, consisted of 3 µl (~ 5 – 10 µg) skeletal muscle 600 *g* supernatant, DTNB (0.1 mM), Triton X-100 (0.04%, v/v), acetyl-CoA (60 µM), and oxaloacetate (0.5 mM), was used to initiate the reaction, measured kinetically at 30 ˚C over 5 min. Triplicate analyses of each sample were done in 96-well microtiter plates which included a reference sample allowing adjustment of batch-to-batch variation.

Table 3-1: Substrates used and products formed during citrate synthase assays.

Enzyme complex	Substrates		Products		Wavelength measured (nm)
CS	Acetyl-CoA **(DTNB)**	oxaloacetate	citrate	CoA **(TNB)**	412

Highlighted DTNB/TNB indicates a coupled reaction, producing (CoA-induced) yellow-coloured TNB from DTNB which could be measured at the specific wavelength indicated.

The DTNB-CoA conjugation forming TNB was followed kinetically in 45 sec intervals for 5 min in a Synergy HT Multi-detection microplate reader at 412 nm. The initial velocity/linear rate increase was determined using linear regression with R^2-values >0.99. These calculations were done using the Gen5 Data Analysis software (v1.05). The calculated initial velocity value (v_1) was expressed as mAbs/min and further used to calculate the specific activity, which was normalised to protein content as indicated in Equation 3.1. A molar extinction coefficient (7465 mM^{-1}) previously determined using a series of product (CoA) and chromogen was used in 200 µl to convert absorbance to molar value.

$$\mu\text{mol/min/mg (CS)} = \left(\frac{v_1}{7465}\right) * 0.2\left(\frac{\mu\text{l protein x}\frac{\mu\text{g}}{\mu\text{l}}}{1000}\right)$$

$$\text{nmol/min/mg} = \text{CS x } 1000 \qquad \textbf{(Eq.3-1)}$$

3.3.3.2 *CII activity*

CII was measured as a potential secondary marker to normalize CII + III on. The CII assay was based on the conditions described by Janssen *et al.* (2007). The reaction, in a final volume of 200 µl, consisted of 10 µl skeletal muscle supernatant, KPi-buffer (50 mM), BSA (0.1%, w/v), EDTA.K (0.08%, w/v), ATP (0.2 mM), DCIP (80 µM), decylubiquinone in DMSO (50 µM), rotenone (2.5 µM), antimycin A (1 µM) and succinate (1.5 mM) was used to initiate the reaction. The reaction was measured kinetically over 5 min, at 30 ˚C. Triplicate analyses of each sample were done in 96-well microtiter plates, which included a reference sample as described for CS.

Table 3-2: Substrates used and products formed during CII enzyme assays.

Enzyme complex	Substrates		Products		Wavelength measured (nm)
CII	Decylubiquinone **(DCIP)**	succinate	Decylubiquinol	fumarate	600

Highlighted DCIP indicate an alternative electron acceptor, resulting in a discolouring of the blue DCIP which could be measured at the specific wavelength.

DCIP was used as a proton acceptor in this assay. The initial velocity/linear rate decrease was determined using linear regression with R^2-values >0.99. These calculations were done using the Gen5 Data Analysis software (v1.05). The (v_1) value was expressed as mAbs/min, and further used to calculate the specific activity, which was normalised to protein content or expressed per unit of CS, as indicated in Equation 3.2. A molar extinction coefficient (12712 mM^{-1}) previously determined using a series of product (DCIP) was used in 200 µl to convert absorbance to molar value.

$$\mu\text{mol/min/mg} = \left(\frac{v_1}{12712}\right) * \ 0.2\left(\mu\text{l protein} * \frac{\frac{\mu g}{\mu l}}{1000}\right)$$

$$\text{nmol/min/CS} \ = \ \mu\text{mol/min/mg x } 1000/\text{CS} \qquad \textbf{(Eq.3-2)}$$

3.3.3.3 *CII + III activity*

CII assays were based on the conditions described by Rahman *et al.* (1996). The reaction, in a final volume of 200 µl, consisted of 6 µl skeletal muscle supernatant, KPi-buffer (50 mM), succinate (10 mM), NaN_3 (20 mM), EDTA.K (0.08 %, w/v), BSA (0.1 % w/v), rotenone (2.5 µM), ATP (2 mM) and cytochrome *c* (oxidised, 0.5 mM) was used to initiate the reaction. The reaction was monitored kinetically over 5 min at 30 ˚C. Triplicate analyses of each sample were done in 96-well microtiter plates, which included a reference sample allowing adjustment of batch-to-batch variation.

Table 3-3: Substrates used and products formed during CII + III enzyme assays.

Enzyme complex	Substrates		Products		Wavelength measured (nm)
CII + CIII	Cytochrome *c* (oxidised)	succinate	**Cytochrome *c* (reduced)**	fumarate	550

Highlighted cytochrome *c* (reduced) indicates the reagent measured during absorption at the specific wavelength.

The initial velocity/linear rate increase was determined using linear regression with R^2-values >0.99. These calculations were done using the Gen5 Data Analysis software (v1.05). The calculated initial velocity value (v_1) were expressed as mAbs/min and further used to calculate the specific activity, which was normalised to protein content or expressed per unit CS, as indicated in Equation 3.3. A molar extinction coefficient (4180 mM^{-1}) previously determined using a series of product (cytochrome *c*) was used in 200 µl to convert absorbance to molar value.

$$\mu\text{mol/min/mg} = \left(\frac{v_1}{4180}\right) * 0.2\left(\mu\text{l protein} * \frac{\frac{\mu g}{\mu l}}{1000}\right)$$

$$\text{nmol/min/CS} = \mu\text{mol/min/mg x } 1000/CS \qquad \textbf{(Eq.3-3)}$$

3.4 Statistical analyses of enzyme activity reference ranges

All parameters measured in triplicate were manually surveyed and evident outliers that most likely occurred due to experimental errors were discarded in an effort to reduce variation in data. For assessment of intra-batch variation the co-efficient of variance (CoV) was determined for the measured repeats with a CoV <10% arbitrarily selected as an acceptable intra-batch variation. The average initial reaction velocity of each sample was subsequently used to normalise against the average initial reaction velocities of CS and CII respectively to generate two values for CII + III activity, i.e. CII + III/CS and CII + III/CII.

Data processing to determine the reference ranges was done using an in-house developed software program by Dr. Gerhard Koekemoer (formerly from the Statistical Consultation Services, NWU) using Transformation Kernel Density Estimation (Sheather & Marron, 1990). This procedure allows density estimation for a limited number reference values used (the CRC in this study) to determine an estimated distribution range (including the 5th and 10th percentiles) of the normalised reference values for the specific diagnostic criteria used in this study. This procedure and specific diagnostic criteria was also followed for determining the distribution of

the reference values of all enzyme analyses previously done on this cohort (Smuts *et al.,* 2010), and will be further discussed in Section 4.2.

Broadly speaking, there are two approaches when calculating percentiles. The first assumes that the data follows a known distribution, usually normal. The second makes no distributional assumptions and instead estimates the distribution from the data provided. For the first approach to be reasonable the assumption that the data follows a known distribution must hold. Here we made use of a formal test for deviation from normality, the Shapiro-Wilk test, to test this assumption. The test was selected as it is powerful and able to detect departures from normality, not just in the mean and variance, but also with regard to skewness and kurtosis (Razali & Wah, 2011). The test did not reveal significant deviations from normality. However, given the small sample size the Shapiro-Wilk test may still not have sufficient power to reliably reject the null hypothesis (Razali & Wah, 2011). Therefore, the Transformation Kernel Density Estimates, which have been shown to be efficient in small samples, were reported.

3.5 DNA isolation and Quantification

The reagents and components used during DNA isolation and quantification were provided in NucleoSpin Tissue kits (Macherey-Nagel, Düren, Germany, 740952.10/10/.50/.250) and Qubit dsDNA HS Assay Kits (Invitrogen, Carlsbad, United States, Q32851) respectively. The use of these are discussed below except were indicated otherwise. Some of the contents used for these analyses (such as buffers) are not revealed and can therefore not be presented in this section.

The instruments used for DNA isolation and quantification: Heraeus Multifuge X3R Centrifuge (Thermo Scientific via Thermo Fisher Scientific, Waltham, United States, 75004515) and Qubit 2.0 Fluorometer (Invitrogen, Q32853).

3.5.1 DNA isolation

Total DNA (nDNA and mtDNA) was isolated from muscle homogenates of which a small volume was initially stored for this purpose during sample preparation using NucleoSpin Tissue kits according to the protocols provided. The homogenates (± 20 µl) were incubated overnight along with 25 µl proteinase K solution (undisclosed concentration) and 180 µl of the provided Buffer T1 at 25°C. Of Buffer B3, 200 µl Buffer was added to the samples and vortexed for 20 sec, and then incubated for 10 min at 70°C. After ensuring lyses of the samples by visible estimation, the adjustment of the DNA binding conditions occurred by adding ethanol (Merck Chemicals, 96-100%, v/v, 1.00983.2500) and thereafter vortexing vigorously. The samples were subsequently placed in NucleoSpin Tissue Columns and centrifuged at 11000 x *g* for 1 min to assure binding

of the DNA to the silica membranes. The silica membranes were washed with 500 µl Buffer BW and 600 µl Buffer B5 for 1 min at 11000 x *g* respectively and dried by centrifuging for 1 min at 15000 x *g*. DNA elution followed by centrifuging twice with 50 µl in 70°C pre-warmed elution buffer for 3 min at 15000 x *g*.

3.5.2 DNA quantification

A final volume of 100 µl was obtained from the DNA isolation and the total DNA quantified using reagents and standards from Qubit dsDNA HS Assay Kit. Fluorescence readings were done using Qubit 2.0 Fluorometer (Invitrogen, Q32853) and software provided. Qubit assay tubes (Invitrogen, Q32856) were set up for two standards necessary for calibration of the Qubit 2.0 Fluorometer, as well as the 24 patient samples. 5400 µl Qubit working solution was prepared by diluting 5373 µl Qubit dsDNA HS reagent in 27 µl Qubit dsDNA HS buffer (in a 1:200 ratio) to a final volume of 200 µl needed for each of the 24 samples and two standards. For each of the two standards, 190 µl Qubit working solution was first loaded into the tubes followed by the addition of 10 µl Qubit DNA standards to each appropriate tube. For the 24 patient samples 199 µl Qubit working solution was loaded into tubes, followed by addition of 1 µl of each DNA sample to each individual assay tube. The final volumes of 200 µl were vortexed for 3 sec, and incubated at room temperature for 2 min before calibration of the Qubit 2.0 Fluorometer with the two standards. After calibration, readings of the samples were performed and displayed by the Qubit 2.0 Fluorometer, and the concentrations of each sample calculated by using Equation 3-4, where QF value presents the value given by the Qubit 2.0 Fluorometer and *x* presents the microliters from each sample added to the individual assay tubes.

$$\text{Concentration of sample} = \text{QF value} \times \left(\frac{200}{\text{x}}\right) \qquad \textbf{(Eq.3-4)}$$

3.6 Next-generation sequencing

This study made use of Ion Torrent PGM (Thermo Scientific, 4462921) for NGS as previously mentioned in Section 2.10. This instrument and associated chemistry was chosen as it is available at the institution where this study was performed and is a suitable approach for use in massively parallel NGS (Millat *et al.,* 2014).

Ion AmpliSeq Custom Panel design (Life Technologies via Thermo Fisher Scientific, Carlsbad, United States, IAD53496_133) was selected for amplicon preparation. Ion AmpliSeq Custom Panels are designed with a free web-based assay design tool (https://www.ampliseq.com) that transforms your selected targeted genes into custom primer pools for amplification of the genomic target regions to construct a library by making use of as little as 10 ng gDNA per primer pool (Yoshimura *et al.,* 2014). The reason for including more than one primer pool is to

improve coverage by overcoming the overlapping of amplicons prone to chemical interactions with each other during the sequencing run. This occurs usually when more than one amplicon is required to cover a specific target region and are therefore split into separate primer pools. The AmpliSeq Custom Panel designed for targeting the coding sequences of the genes for this study, as listed in Table 3-4, allowed analysis of 304 amplicons (combined size of 61.37 kb) with an overall coverage of 98.17% of targeted regions as described in more detail in Appendix A. For sequencing a total of 20 ng genomic DNA was required per sample.

The targeted genes sequenced in the 24 selected patients (nine patient cases with reduced levels of muscle CoQ_{10} and 15 patient cases with normal levels of muscle CoQ_{10}) in this study were selected based on the genes involved in CoQ_{10} biosynthesis, *as well as* frequently reported genes known to cause secondary CoQ_{10} deficiency or have an effect on CoQ_{10} levels. The main focus of this study was, however, on the coding sequences of genes from the biosynthetic pathway, of which functions of each selected gene is presented in Table 3-4.

Table 3-4: Targeted genes selected for sequencing based on enzyme function.

Genes involved in CoQ_{10} Biosynthesis*	
Gene	**Function**
PDSS1 (COQ1)	Prenyl disphosphate synthase
PDSS2	Prenyl disphosphate synthase
COQ2	4HB-prenyl transferase
COQ3	O-methyltransferase
COQ4	Organization of multienzyme complex
COQ5	C-methyltransferase
COQ6	Mono-oxigenase (C5 hydroxylation)
COQ7	Hydroxylase (C6 hydroxylation)
ADCK3 (COQ8)	Atypical kinase phosphorylation of COQ proteins
COQ9	Not known
COQ10A	CoQ chaperone
COQ10B	CoQ chaperone
Genes involved in secondary CoQ_{10} deficiencies	
Gene	**Function**
APTX	Double stranded DNA repair
ETFDH, ETFA, ETFB	Electron-transfer-flavoprotein
BRAF	Regulating of signalling pathway, affecting cell division, differentiation, and secretion
APOE	CoQ_{10} metabolism

*Genes of main focus in the study.

The Ion Torrent PGM sequencing process, as illustrated in Figure 3-1, is a simple workflow that involves DNA library preparation, template preparation of library fragments via emulsion PCR (emPCR), followed by “sequencing-by-synthesis” stages. Some of the contents used for these sequencing kits (such as buffers) are not revealed by the manufacturer and can therefore not be fully described in this section.

Instrumentation used throughout molecular genetic testing: MJ Mini Gradient Thermo Cycler (Bio-Rad Laboratories, Hercules, United States, PTC-1148EDU), Eppendorf Centrifuge (Eppendorf, Hamburg, Germany, 5424).

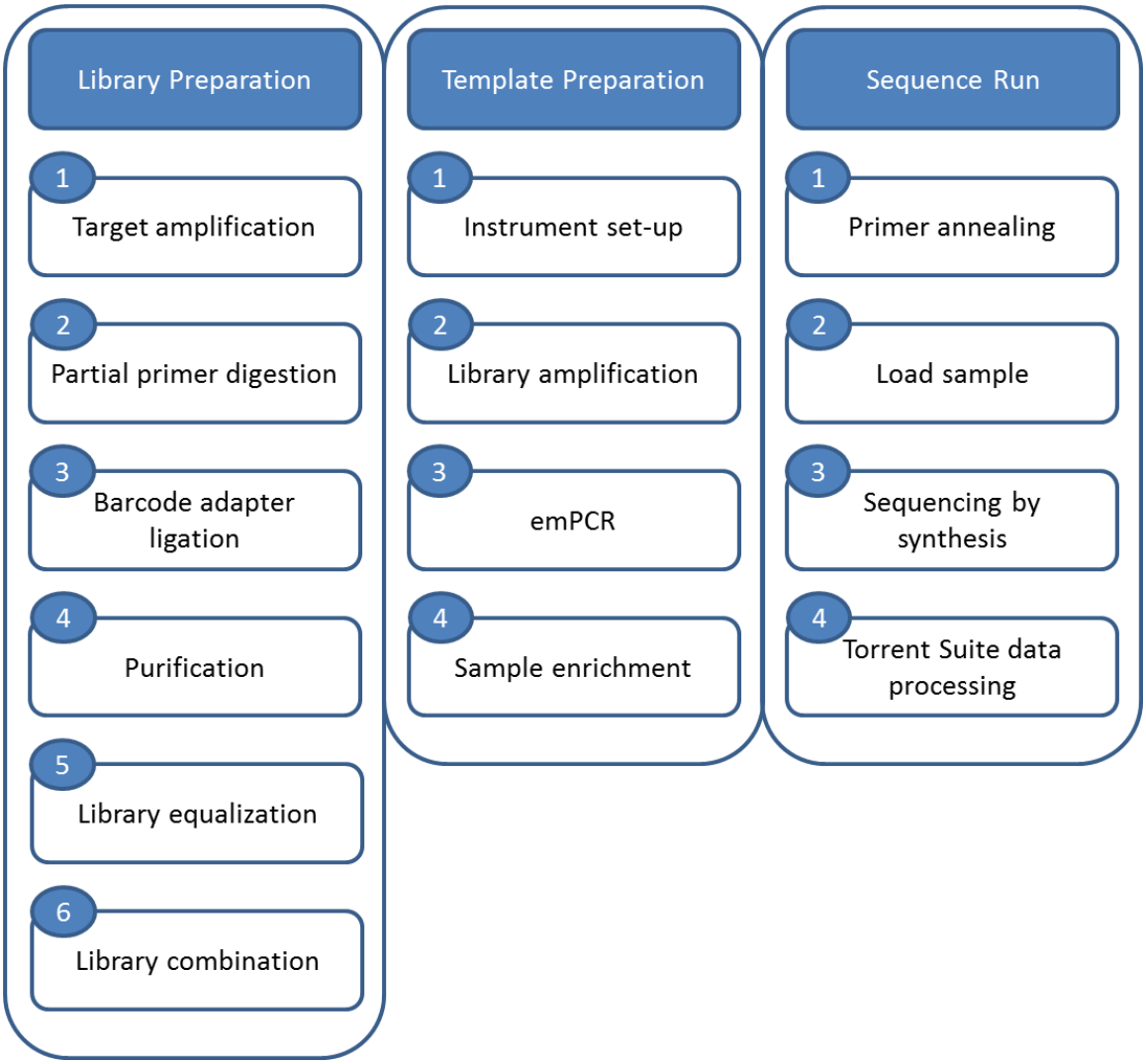

Figure 3-1: Schematic representation of the Ion Torrent PGM sequencing workflow followed. The process involves DNA library preparation, template preparation of library fragments via emulsion PCR (emPCR), and sequencing stages. Consecutive steps involved in each stage are represented in boxes below main stages. In short, generation of the DNA library fragments starts with an amplification round of the target regions, followed by ligation of adapters to the products. The barcoded library fragments are then clonally amplified via emPCR onto Ion sphere particles (ISPs) and these ISPs deposited onto the semiconductor chip into the wells followed by a sequencing run.

3.6.1 DNA Library preparation

Reagents and components used during library preparation were provided in the Ion AmpliSeq Library Kits (Life Technologies, 4480441) and constructed according to the manufacturer's instructions (Publication Number MAN0006735, Publication Part Number 4480441, Revision A.0, Revision Date 10 December 2013), as discussed below unless indicated otherwise.

3.6.1.1 *Amplification of targets*

Each target amplification reaction consisted of 4 µl 5X Ion AmpliSeq HiFi Mix, 10 µl 2X Ion AmpliSeq Primer Pool and 10 ng gDNA adding up to a final volume of 20 µl in a single PCR tube. The amplification of the target genomic regions was done using PCR by loading the PCR tubes into a thermo cycler. The reaction started with an initial holding step at 99°C for 2 min to activated the enzyme, followed by 18 cycles of DNA denaturing (99°C) for 15 sec and primer annealing and DNA extension (60°C) for 4 min. The samples were then held and cooled at 10°C until removal from the thermo cycler.

3.6.1.2 *Partial digestion of primer sequences*

The amplicons in the tubes were modified for effective target assessment by partial digestion of the primers. Each amplified sample was treated with 2 µl FaPa Reagent giving a final volume of 22 µl. The PCR tubes were then loaded in a thermo cycler and set on a program to run at 50°C for 10 min, 55°C for 10 min, 60°C for 20 min and held at 10°C for up to an hour to assure optimal primer digestion.

3.6.1.3 *Ligation of barcode adapters*

A unique barcode from Ion Xpress Barcode Adapters 17-32 Kit (Life Technologies, 4474009) and Ion Xpress Barcode Adapters 33-48 Kit (Life Technologies, 4474518) was added to each sample to enable library identification on a single Ion Chip. The barcode adapters are ~ 10 bp long, 3' end adapters containing the sequencing primer, with the 5' end adapter containing a biotin label contributing to purification and amplification steps. For each diluted barcode adapter mix, 1 µl P1 Adapter, 1 µl Xpress Barcode and 2 µl Nuclease-free Water was added to a final volume of 4 µl. For the barcode ligation reaction 4 µl Switch Solution and 2 µl of the diluted barcode adapter mix was added to each tube containing 22 µl digested sample, making up a total volume of 28 µl. Addition of 2 µl DNA ligase followed and the tubes loaded in the thermo cycler. The samples were set on a reaction program to run at 22°C for 30 min, 72°C for 10 min, and held at 10°C.

3.6.1.4 *Purification of unamplified library*

To each library 45 µl (1.5X sample volume) Agencourt AMPure XP Reagent was added and mixed thoroughly to suspend the beads with the DNA for suitable size DNA fragment selection. Incubation of the tubes at room temperature for 5 min followed before a pellet was obtained via a Dynamag-2 Magnet rack (Life Technologies, 12321D). After discarding the supernatant the beads were washed two times with 150 µl 70% (v/v) ethanol and air dried after removing the supernatant without disturbing the pellet.

3.6.1.5 *Equalize library*

The Ion Library Equalizer kit (Life Technologies, 4482298) was used to normalise the library concentrations at ~100 pM without the necessity of quantification. Reagents and components used were provided in the Ion Library Equalizer kit and performed according to the manufacturer's instructions (Publication Number MAN0006735, Publication Part Number 4482298, Revision A.0, Revision Date 10 December 2013), as discussed below unless indicated otherwise.

3.6.1.5.1 Amplification of the library

To each bead pellet 50 µl Platimum PCR SuperMix High Fidelity and 2 µl Equalizer Primers were added and vortexed. Droplets were collected by spinning down the tubes and placed on the Dynamag-2 Magnet rack for 2 min, followed by the transfer of 50 µl supernatant from each sample to a new PCR tube and loaded in the thermo cycler. The reaction started with an initial holding step at 98°C for 2 min to activated the enzyme, followed by 7 cycles of DNA denaturing (98°C) for 15 sec and DNA extension (64°C) for 1 min. The samples were then held and cooled at 10°C until removal from the thermo cycler (for up to 1 hour).

3.6.1.5.2 Equalizer beads

The Equalizer Beads were mixed and brought to room temperature. To each reaction sample 3 µl beads and 6 µl Equalizer Wash Buffer was added. The tubes were placed on the Dynamag-2 Magnet rack for 3 min until a clear solution was presented. The tubes were removed from the magnet rack and 6 µl Equalizer Wash Buffer added per reaction followed by re-suspension.

3.6.1.5.3 Combining of equalizer beads with amplified library

When the library amplification thermo cycler reaction finished, 10 µl Equalizer Capture was added to each amplified library and vortexed to collect droplets by spinning down the tubes. Incubation for 5 min at room temperature followed. After the Equalizer Beads was gently vortexed 6 µl washed beads were transferred to each sample tube and mixed with the captured reaction by pipetting up and down. Incubation for 5 min at room temperature followed before being placed on the Dynamag-2 Magnet rack for 2 min until the solution was clear. The supernatant was discarded without disturbing the pellet and 150 µl of Equalizer Wash Buffer added to each reaction. The beads were washed before adding another 150 µl of Equalizer Wash Buffer to each reaction for another bead wash.

3.6.1.5.4 Elution of equalized library

The supernatant was discarded without disrupting the pellet and 100 µl Equalizer Elution Buffer added to each pellet and vortexed. The library was eluted by incubation of 5 min in the thermo cycler at 32°C and thereafter incubated for 5 min at room temperature on the Dynamag-2 Magnet rack.

3.6.1.6 *Combination of equalised libraries*

Multiple barcoded libraries were combined to allow loading onto a single Ion Chip for cost and sequencing time reduction. The number of combined libraries in a single sequence run depended on the size of the chip used, as well as the required 30 x coverage. All the libraries were prepared on a single Ion 316 Chip v2 (Life Technologies, 4482261). The separate libraries prepared from the same DNA sample using the primer pools provided by the Ion AmpliSeq Custom Panel was combined by adding 3 µl of each reaction into a single tube. The library was then diluted to a concentration of ~100 pM by adding 2 µl of the combined libraries from the tube with 23 µl Nuclease-Free Water to a final volume of 25 µl.

3.6.2 Template preparation

Reagents and components used during the template preparation, including the components necessary for the instrument setup, were provided by the Ion OneTouch Template OT2 200 Kit (Life Technologies, 4480974) for the use of the Ion OneTouch 2 system (Life Technologies, 4474779). The template preparation was performed according to the manufacturer's instructions (Publication Number MAN0007221, Revision 5.0) and is discussed below (unless indicated otherwise).

3.6.2.1 *Instrument set-up*

The setup and installation of the Ion OneTouch 2 Instrument and reagents which included the OneTouch Recovery Tubes, Ion OneTouch Recovery Router and OneTouch 2 Amplification Plate were done according to the provided protocol.

3.6.2.2 *Preparation of amplification solution*

For the amplification solution 25 µl Nuclease-Free Water, 500 µl Ion PGM Template OT2 200 Reagent Mix, 300 µl Ion PGM Template OT2 200 PCR Reagent B, 50 µl Ion PGM Template OT2 200 Enzyme Mix and 25 µl diluted library was added to a Lo-Bind Tube to a final volume of 900 µl. The Ion PGM Template OT2 200 Ion Sphere Particles (ISPs) were re-suspended by vortexing for 1 min at maximum speed, and 100 µl added to the amplification solution to a final

volume of 1000 µl. The amplification solution was pippeted slowly through the Ion OneTouch Plus Reaction Filter sample port, followed by the addition of 1500 µl Ion OneTouch Reaction Oil which were slowly pippeted through the same port.

3.6.2.3 *Amplification through emPCR*

After the installation and assembly of the Ion OneTouch Plus Reaction Filter on the Ion OneTouch 2 Instrument as described in the provided protocol, the samples were amplified via emPCR during a run for ~6 hours.

3.6.2.4 *Recovery and enrichment of the template-positive ISPs*

For the recovery of the template-positive ISPs all but ~50 µl of each Ion PGM OT2 Recovery Solution was removed from each Recovery Tubes to avoid disturbance of the pellet containing the template-positive ISPs. The template-positive ISPs were re-suspended and transferred into the first well of the provided 8-well strip from the Ion OneTouch enrichment system (ES) provided by the OneTouch 2 system. The enrichment of the template-positive ISPs was performed using the Ion OneTouch ES. A Melt-Off Solution was prepared by combining 280 µl Tween Solution and 40 µl NaOH (1 M) to a final volume of 320 µl. Dynabeads MyOne Streptavidin C1 beads were washed and resuspended for 20 sec and 13 µl transferred to a new 1.5 mL Lo-Bind Tube. The tube was placed on a Dynamag-2 Magnet rack and the supernatant discarded after the solution cleared. 130 µl of MyOne Beads Wash Solution was added to the Dynabeads MyOne Streptavidin C1 beads and vortexed for 30 sec and added into well 2 in the 8-well strip. Wells 3-5 was filled with 300 µl Ion OneTouch Wash Solution, well 7 filled with 300 µl freshly prepared Melt-off solution and wells 5 and 6 left empty. After the ~35 min run was performed on the Ion OneTouch ES according to the protocol, the PCR tube containing the enriched ISPs was gently inverted five times before proceeding to the sequencing run.

3.6.3 Ion Torrent Sequencing

Ion PGM sequencing run, as well as components necessary for instrument cleaning and initiation, were provided in the Ion PGM Sequencing 200 Kit v2 (Life Technologies, 4482006) and discussed below unless indicated otherwise. Ion PGM Sequencing 200 Kit v2 allows sequencing up to 200 bp reads using the Ion 316 Chip v2 (Life Technologies, 4483188). Ion PGM sequencing run was performed according to the manufacturer's instructions (Publication Number MAN0007273, Publication Part Number 4482006, Revision 1.0, Revision Date 26 November 2012).

3.6.3.1 *Sequencing primer annealing*

Enrichment of template-positive ISPs was done by adding 5 µl Control ISPs directly to the entire enriched, template-positive ISPs in a 0.2 ml non-polystyrene PCR tube. The content of the tubes were mixed by pipetting up and down before being centrifuged for 2 min at 15500 x *g*. Without disturbing the pellet, 15 µl supernatant was left in the tube when removal of the supernatant was done. Addition of 12 µl Sequencing Primer followed to a total volume of 27 µl (adding Annealing Buffer where necessary). The pellet was disrupted by pipetting up and down, and loaded in the thermo cycler for a program set for 95°C for 2 min, followed by 37°C for 2 min, by using the heated lid option. The sample remained in the cycler at room temperature during the Chip Check which is to ensure that the Ion 316 Chip v2 is working properly before sample loading occurs.

3.6.3.2 *Binding of sequence polymerase to ISPs and loading of chip*

After annealing the sequencing primer to the ISPs 1 µl Ion PGM Sequencing 200 v2 Polymerase was added to the ISPs to a total volume of 7 µl. The sample was mixed by pipetting up and down and incubated for 5 min at room temperature. The Ion 314 Chip v2 was loaded as described by the protocol by pipetting the 7 µl sample into the loading port. The ISPs were deposited onto the chip at a low rate to avoid bubble formation and placed on the MiniFuge and centrifuged for 30 sec with the chip tab pointing in. After centrifuging, the sample was steadily pipetted up and down three times to mix the sample, and placed on the MiniFuge and centrifuged for 30 sec with the chip tab pointing out. This centrifuging and sample mixing step was repeated before discarding the liquid from the loading port of the chip and placed on the Ion PGM for the sequencing run.

3.6.3.3 *Sequencing run*

The selected planned run was performed by real-time measurements of H^+ ions produced during DNA synthesis when nucleotides in set order (TACG) was run across the chip one at a time with wash steps in between as described in Section 2.10.1.

3.7 Data Analysis

Once the data was generated on the Ion PGM, it was automatically transferred to the standard Ion PGM System Torrent Suite Software (v4.0.2). Signal processing and base calling algorithms of the Torrent Suite software (v4.0.2) produced the DNA sequences for primary data analysis. The primary analyses consisted of basic interpretations of the raw sequence data obtained. This included the total number of amplicon -reads, -lengths, -depths, coverage obtained and quality scores of the data acquired. Quality assessment automatically trimmed and filtered the raw data

to eliminate low quality sequences. The high quality reads for each patient were aligned against the reference human genome (*Homo sapiens,* GrCh37) in order to gain a consensus sequence for every individual. This enabled information on the detected variations, including the variants identified in each individual, as well as estimation of allele frequencies (%) by using Torrent Variant Caller plug-in software (v4.0).

Data mining and identification of possible novel and/or previously reported variants were identified using available open source databases during a secondary data analysis. The first part of this in-house developed pipeline was to run raw BAM data files obtained from the Torrent Suite (v4.0.2) through the Ensembl online VEP runner (variant effect predictor, v73, http://www.ensembl.org) (McLaren et al., 2010). The output files contained the variants of importance in a variant caller file (VCF). GEMINI (v0.6.5, http://gemini.readthedocs.org/en/latest/index.html), a genome mining framework database, explored and filtered the VCF file according to specific arguments and/or criteria with the help of 3rd party databases (Paila *et al.,* 2013). The variants were filtered and classified as either novel (coding variants with or without Loss of Function; LoF) or previously reported (coding variants with or without LoF). Each previously identified variant is present in the NCBI dbSNP (http://www.ncbi.nml.nih.gov//snp/) and thus have a unique SNP rs number (Sherry *et al.,* 2001) as mentioned in Section 2.11. To further investigate the identified variants, variant interpretation software, SIFT (sorting intolerant from tolerant) and PolyPhen-2 (Polymorphism Phenotyping, v2) were used to sort the intolerant from tolerant variants (Kumar *et al.,* 2009; Adzhubei *et al.,* 2010) and predicted the polymorphism impact respectively (Zeng *et al.,* 2014). The PolyPhen-2 prediction score (PPS) is a scaled score that ranges between 0 and 1 to predict the severity of the SNP in the case of a severely affected transcript. A score closer to 1 predicted the amino acid change to be deleterious, while a score closer to 0 indicated a neutral effect. The SIFT prediction score (SPS) presented predicted the effect caused by the SNP. A SIFT score <0.05 predicted the variant to be deleterious, while a score >0.05 predicted the variant to be tolerant (Wei *et al.,* 2011). Deleterious non-synonymous (nsSNPs) can result in dramatic phenotypic consequences due to the amino acid change it causes, while tolerated nsSNPs do not alter protein function (Wang *et al.,* 2009).

The second part of the pipeline indicated whether the newly found variants as well as the previously reported variants had the potential to be pathogenic. The databases used for pathogenic classification were OMIM and CliniVar (http://www.ncbi.nlm.nih.gov/clinvar/) (Landrum et al., 2014), identifying if a known clinical phenotype (clinical significance) was connected to the variants identified. Further mined data observed the variant occurrence in specific populations. This data was retrieved from the 1000 Genomes Project (1000 Genomes Project Consortium, 2012). The 1000 Genomes project (1000G) is an international project

recording human genetic variation by sequencing the whole genomes of over a 1000 anonymous individuals over the world of different ethnical groups to provide insight to genetic variations in the diverse populations. This data was furthermore used to motivate the possible pathogenic state of the identified variants by means of population occurrence.

Tertiary analysis and ultimate interpretation of the results were done by making use of clinical profiles of the patients obtained from previous studies, as well as Sanger sequence validation and SDS-PAGE/western blot analyses for protein structure validation of three patients identified with possible compound heterozygous variants, as mentioned in Section 3.2. Figure 3-2 illustrates the workflow followed during data analysis consisting of the primary, secondary, and tertiary data analysis.

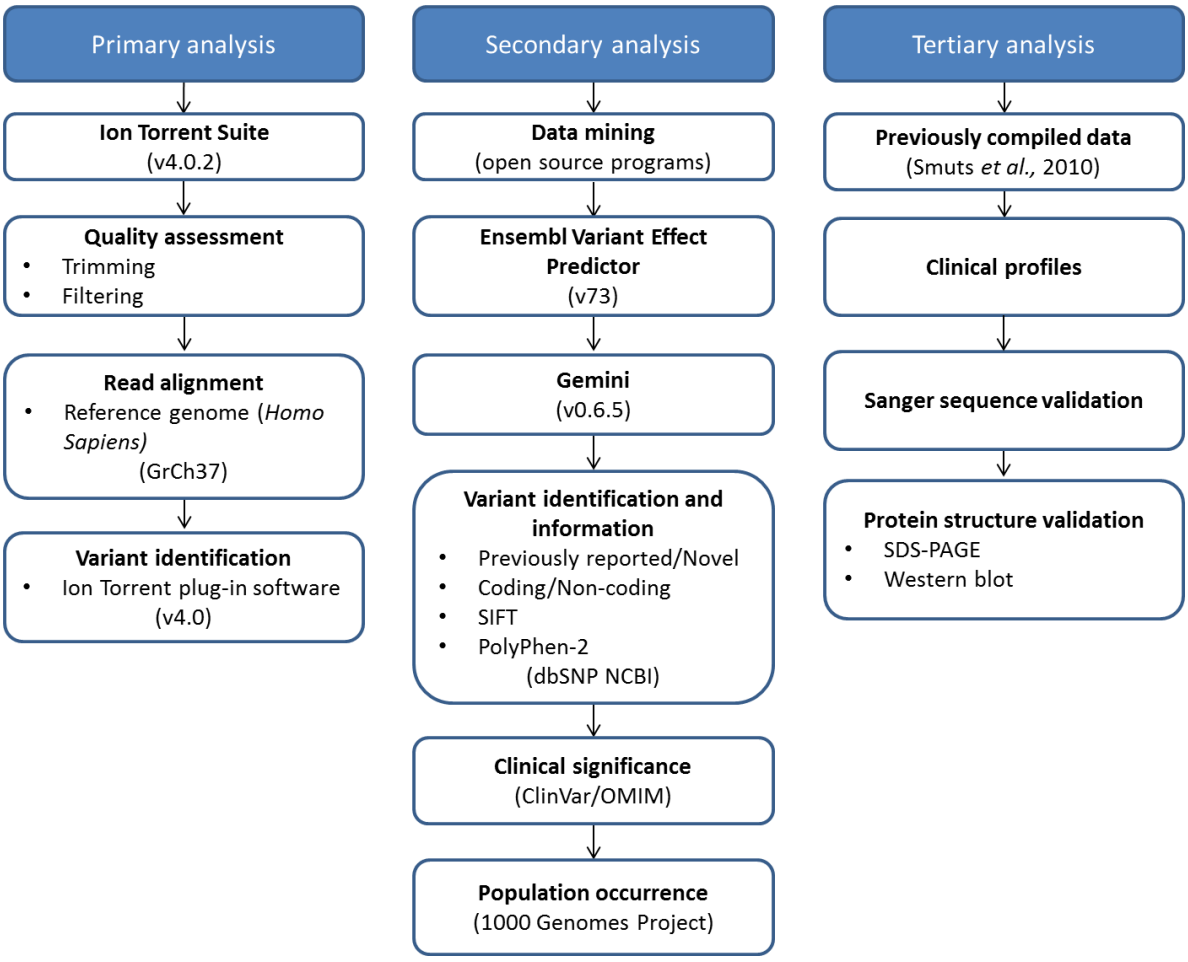

Figure 3-2: Schematic presentation of the workflow followed to perform data analysis. Consecutive steps follow below the main analysis. In short, the pathway starts with primary analysis performed using the Ion Torrent suite, followed by data mining as secondary analysis by making use of various available open source databases, leading to tertiary data analysis for the ultimate interpretation and validation of the results conducted from the study by making use of previous compiled clinical data and immunoblotting. SIFT: sorting intolerant from tolerant; PolyPhen-2: Polymorphism Phenotyping, v2; dbSNP: SNP database; SDS-PAGE: sodium dodecyl sulphate polyacrylamide gel-electrophoresis.

3.8 Data validation

The validation of the three compound heterozygous variants identified in the three patients (P2, P43 and P78) with probable pathogenic predictions were confirmed by Sanger (dideoxy-) sequencing on the ABI 3500xL Genetic Analyser (Inqaba Biotech, Pretoria, South Africa).

3.9 SDS-PAGE/western blot validation

In order to validate the effect that three possible compound heterozygous mutations identified during the NGS bioinformatics workflow had on three patients, protein structure validation measuring relative protein quantities were performed using SDS-PAGE and western blot analyses. SDS–PAGE was performed based on the original Laemmli method (Laemmli, 1970). A Bio-Rad Trans-Blot Turbo Transfer System (Bio-Rad, 170-4155) and V3 Western Workflow, together with enhanced chemiluminescence (ELC) were used to detect the individual targeted proteins during western blots immunodetection. The Bio-Rad V3 Western Workflow allows rapid, efficient and reproducible transfer of proteins from gels to membranes with minimal preparation time and normalization made easy using stain-free technology. The stain-free technology involves UV-induced haloalkane modifications in the pre-cast gels which react with trytophan residues of the proteins for fluorescent protein visualization (Bio-Rad, http://www.bio-rad.com/).

3.9.1 Materials

For these analyses 600 *g* supernatants were used as described in Section 3.3.3. The samples were prepared based on the total protein content present in each patient sample that was determined by BCA method (Smith *et al.,* 1985). Some of the contents used for these analyses (such as buffers) are not revealed and can therefore not be presented in this section.

The following materials and reagents were used: SDS (Sigma-Aldrich, 862010); Tris.HCl-[Tris(hydroxymethyl)aminomethane] (Melford Laboratories, Ipswich, United Kingdom, B2005); dithiothreitol (DTT) (Sigma-Aldrich, D0632); glycerol (Sigma-Aldrich, G-6279), bromophenol blue (Sigma-Aldrich, 114391); Milli-Q prepared water; Precision Plus Protein WesternC Standards (Bio-Rad, 161-0376); Trans-Blot Turbo Mini PDVF Transfer Pack (Bio-Rad, 170-4156); monoclonal and polyclonal antibodies (Abcam, Cambridge, United Kingdom, ab97051, ab107069, ab131376, ab6789); Mini-PROTEAN TGX Stain-Free precast Gels (Bio-Rad, 456-8123); 10x TGS running buffer (Bio-Rad, 161-0732); Clarity Western ELC Substrate (Bio-Rad, 170-5060S); methanol (Honeywell, Morristown, United States, 230-4); PBS (1x) (Sigma-Aldrich, P4417); Tween 20 (Merck Chemicals, 822184); milk powder; Milli-Q prepared water. HCl (Sigma-Aldrich, 318949) was used to adjust pH values.

Instrumentation used throughout immunoblotting: Mini Protean 3 System (Bio-Rad, 165-8000), Trans-Blot Turbo Transfer System (Bio-Rad, 170-4155), PowerPac Basic Power Supply (Bio-Rad, 164-5050), ChemiDoc MP System (Bio-Rad, 170-8280).

3.9.2 Methods

Experimental runs were performed in advance to determine the appropriate amount of protein concentrations to be loaded on the gel. From the 600 *g* supernatant samples of each patient 25 µg protein were re-suspended in a sample buffer (0.2 M Tris–HCl (pH 6.8), 10% (w/v) SDS, 10 mM DDT, 0.05% (w/v) bromophenol blue and 20% (v/v) glycerol) in a ratio of 1:4, and filled with water to a total of 20 µl. Samples were boiled for 5 min at 100ºC and were loaded along with a molecular weight standard. Using a 1X running buffer the samples were separated on Mini-PROTEAN TGX Stain-Free precast gels at 300 V for 15 min. After separation of the protein complexes during SDS-PAGE the gel was removed from the cassette and the protein separation visualised on the ChemiDoc MP imager. After the imaging of the gel the proteins on the gel were transferred to a polyvinylidenedifluoride (PVDF) membrane at 25 V for 3 min. The membrane was imaged to verify protein transfer and saved for normalisation purposes. The membrane was blocked overnight at 4ºC with 5% (w/v) milk powder and PBS. A 10 min wash in 0.05% (v/v) Tween 20 and PBS followed. The individual proteins were bound by monoclonal mouse anti-ETFDH antibody (ab131376), rabbit anti-COQ6 antibody (ab128652) and rabbit anti-COQ7 antibody (ab107069) for 1.5 hours at room temperature in 1% (w/v) milk powder and PBS on a shaking platform. This was followed by 3 x 5 min washes in 0.05% (v/v) Tween 20 in PBS. The individual proteins bound by monoclonal mouse- and rabbit antibodies were detected using goat-anti-mouse and goat-anti-rabbit IgG horse radish conjugated polyclonal antibodies for 1.5 hours at room temperature in 1% (w/v) milk powder and PBS on a shaking platform. Afterwards, 3 x 5 min washes in 0.05% (v/v) Tween 20 in PBS followed prior to detection of the individual proteins. This was accomplished by chemilumenescenes by mixing the substrates in the Clarity Western ELC kit in a ratio 1:1 in which the membrane was incubated for and imaged using a high sensitivity automated setting to detect the individual proteins using the Bio-Rad ChemiDoc system.

The western blot image was further visualised and normalised using the provided Image Lab Software (Bio-Rad, v5.1). Firstly the lanes and bands representing the individual proteins of each sample were detected and then normalised to the total protein content using the stain-free blot image.

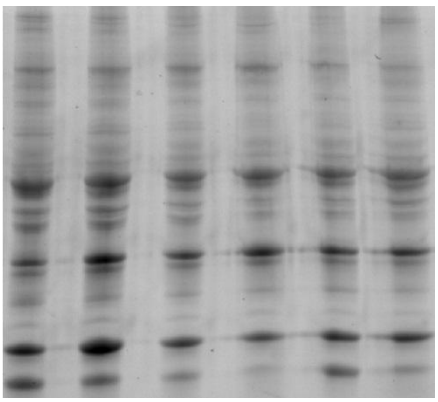

Figure 3-3: Example of a stain-free image during analysis using the ChemiDoc MP system. This image shows six samples of various sources of 25 µg 600 *g* supernatant run on a Mini-PROTEAN TGX Stain-Free precast gel. It represents the total protein content in the samples of UV-induced trihalo compounds reacting with tryptophan residues of the protein.

Chapter 4

Results and Discussion

4.1 Introduction

A thorough description of the experimental design and methods, including the data analysis workflow used to generate the results, was provided in Chapter 3. The results obtained during this study is presented and discussed in this chapter. If relevant, supporting and other data are provided in the appendixes indicated in each section. Further discussion of the results and the conclusions drawn from the results is provided in Chapter 5.

4.2 RC enzyme analysis

4.2.1 Enzyme diagnostic criteria

As described in Section 3.2, 25 patients and 20 CRC (five patients of each race and gender) were selected for the study based on previous enzymatic diagnosis of RC enzyme deficiencies. The relevant RC complexes were measured spectrophotometrically by making use of 600 *g* supernatants available from muscle homogenates. Enzyme analyses were performed in triplicate and compared to a reference sample as indicated in Section 3.3. As described in Section 3.4 the average of each sample was calculated before being normalised to CS and CII activities respectively.

The normalised values of the CRC were used to determine a distribution range which was then used to compare the normalised enzymatic activities of the 25 patients to the 5^{th} and 10^{th} percentiles of this distribution range. It should be noted that, although the distribution of the CRC enzyme data were normally distributed (p value >0.05 using Shapiro Wilk's test), the non-parametrical method (Transformation Kernel Density Estimation) was used to determine the distribution range. The reason for this approach was that the number of CRC cases (20) was relatively low and this procedure has been in use for this cohort study (Smuts *et al.,* 2010), thus allowing comparison of the distribution data with the existing diagnostic data (Table 4-1). An enzyme deficiency was identified according to *similar* criteria used by the Mitochondria Research Laboratory as published in Smuts *et al.* (2010). The difference is that, unlike Smuts *et al.* (2010), CIV was not also included as a third possible marker in this study as the CRC group in this study is a referred patient group with possible involvement and instability of the RC super complex, including CIV. The original reference group used in Smuts *et al.* (2010) was "healthy" control muscle obtained from children who have undergone orthopaedic surgery. Thus, based on similar principles according to these criteria, an enzyme (CII + III in this case) deficiency was

identified when the enzyme activity was lower than the 5th or 10th percentile of the CRC values when expressed on both CS and CII. At least one of these values should be equal or lower than the 5th percentile of the CRC range and the second at least equal or lower than the 10th percentile of the CRC range provided that the marker enzyme activities were not deficient. If a marker was also deficient (almost exclusively CII as a CS deficiency is not viable) only CS was used as marker.

4.2.1.1 *Reference Values*

Figure 4-1 illustrates the distributed data of the normalised reference values (CRC) using the Transformation Kernel Density Estimation for CII + III/CS and CII + III/CII respectively. The procedure performed on the limited number CRC were used to determine the estimated distribution ranges (including 5th and 10th percentiles) of the reference values which were used for the specific diagnostic criteria as discussed respectively in Sections 3.4 and 4.2.1. These estimated distribution ranges, and other calculated parameters from the current study and of previous years, are presented in Table 4-1.

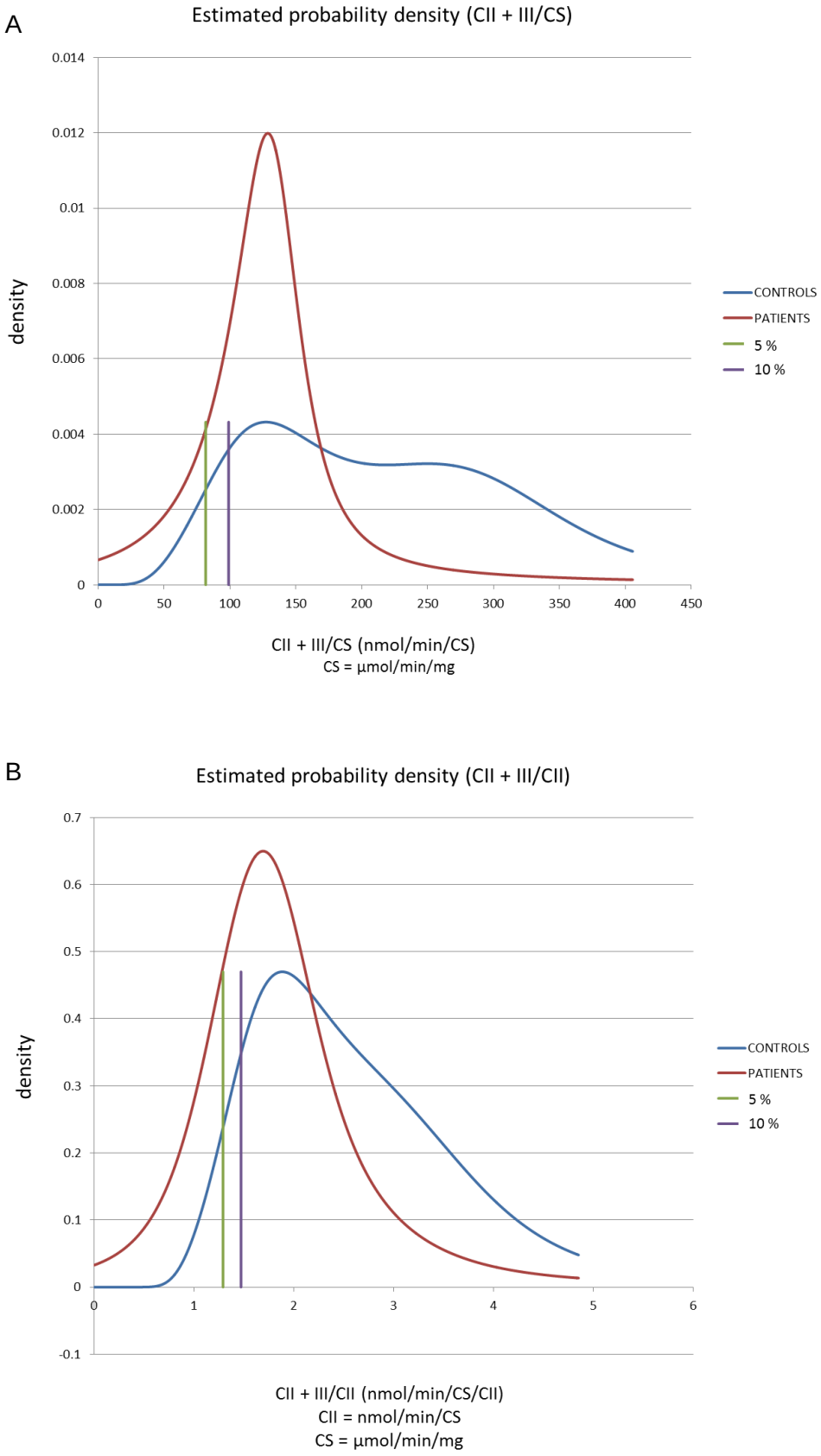

Figure 4-1: Transformation Kernel Density Estimation distribution for normalised reference values. Figure 4-1 illustrates the Kernel Density curve of the transformed A) CII + III/CS data, and

transformed B) CII + III/CII data, representing the probability density of the reference values (control group indicated in blue) and patient group (indicated in red) used in this study to calculate the distribution range for diagnostic purposes. The green line indicates the 5th percentile (5%), and purple line the 10th percentile (10%) calculated from the reference values (CRC) density estimation to which a combined CII + III deficiency was identified according to the specific diagnostic criteria used in this study.

Table 4-1: Parameters calculated using the re-analysed normalised reference values (CRC) obtained in the current study (top section of table), and CRC samples data before the onset of this study (bottom section of table), using the Transformation Kernel Density Estimation distribution.

CRC samples data as compiled in this study				
Variable	**Mean (n = 20)**	**5th percentile**	**10th percentile**	**SD**
CII + III/CS **(nmol/min/CS)**	223	81.5	98.9	94.8
CII + III/CII **(nmol/min/CS/CII)**	2.49	1.29	1.47	0.78
CRC samples data before onset of this study				
Variable	**Mean (n = 20)**	**5th percentile**	**10th percentile**	**SD**
CII + III/CS **(nmol/min/CS)**	296	115	121	304
CII + III/CII **(nmol/min/CS/CII)**	3.40	0.96	1.22	3.46

CII: (nmol/min/CS); CS: µmol/min/mg; n: number of CRC; SD: standard deviation.

4.2.2 Results

The data points measured in triplicate, including the average SD and CoV of each sample and reference, are illustrated in Appendix B. As mentioned in Section 3.3 the initial velocity/linear rate increase was determined and assessed visually and the accuracy confirmed with R^2-values >0.99. The CoV mean for the reference samples included to adjust the batch-to-batch variations were 12% and 6% for CII + III/CS and CII + III/CII respectively, and the CoV mean for all parameters measured for the samples during the enzymatic analysis were 16%, which indicated acceptable repeatability (desired CoV <10%).

The final processed enzyme specific activities for CII + III and outcomes (applying the diagnostic criteria as described in Section 4.2.1) are summarized in Table 4-2, consisting of the selected patient cases that were subsequently sequenced (patients with low and normal CoQ_{10}

levels), of which the patients with confirmed CII + III deficiency according to the specified criteria is highlighted in yellow. It also summarises selected clinical data obtained from previous studies (Smuts *et al.,* 2010), as well as the previous RC enzyme data. The patients in the table are sorted by increasing levels of CoQ_{10}, of which the first nine patients highlighted in grey represent the patient cases with reduced CoQ_{10} levels.

Table 4-2: Previously obtained clinical profiles and enzyme deficiency diagnosis, as well as CII + III enzymatic status from newly obtained results.

Original patient ID	Sex, age in years	Population group	Clinical profile	CoQ_{10} level (nmol/CS)*	RC CII + III/CS (nmol/min/CS)	RC CII + III/CII (nmol/min/CS/ CII)	Position relative to CII + III/CS reference values	Position relative to CII + III/CII reference values	CII + III deficiency confirmed	Initial identified RC enzyme deficiency
P78	F, 8	C	DD, CNS, migraine, M, H	0.86	138	1.25	>10%	<5%	No	I,III,IV, II+III
P24	F, 11	A	M	1.20	104	2.39	>10%	>10%	No	PDH, I,II,III, IV, II+III
P98	F, 10	A	C	1.23	73	1.20	<5%	<5%	Yes	II+III
P99	M, 10	C	FTT, M, GS, LA	1.39	120	1.25	>10%	<5%	No	III, II+III
P54	M, 0	C	FTT, LA	1.54	133	1.80	>10%	>10%	No	-
P2	F, 4	A	DD, M	1.56	14	0.68	<5%	<5%	Yes	PDH, CII + III
P29	M, /	A	M	1.61	90	2.15	<10%	>10%	No	II+III
P14	F, 1	A	DD, DR, CNS, E, M	1.64	119	2.79	>10%	>10%	No	I, CII + III
P95	F, 0	A	FTT, DD, CNS, E, M, H	1.65	144	1.47	>10%	>10%	No	III, IV, II+III
P64	M, 11	C	DR, CNS, E, migraine, M, H	1.86	101	1.52	>10%	>10%	No	I, II, III, IV, II+III
P85	F, 8	C	DR, CNS, E, M	2.12	241	2.85	>10%	>10%	No	II+III
P94	M, 2	A	FTT, DD, CNS, E, M, LA	2.34	157	1.70	>10%	>10%	No	II+III
P88	M, 1	C	DD, DR, CNS, E, M, GS	2.38	135	1.89	>10%	>10%	No	III, II+III
P21	F, 1	A	DD, M, LA	2.47	205	3.92	>10%	>10%	No	I, II, III, IV, II+III
P101	F, 2	A	FTT, DD, M	2.47	86	1.23	<10%	<5%	Yes	III, II+III
P34	F, 3	A	DD, M	2.52	150	2.67	>10%	>10%	No	I, III, IV, II+III
P43	F, 3	A	FTT, DD, M, LA	2.59	93	1.24	<10%	<5%	Yes	IV, II+III

Original patient ID	Sex, age in years	Population group	Clinical profile	CoQ_{10} level (nmol/CS)*	RC CII + III/CS (nmol/min/CS)	RC CII + III/CII (nmol/min/CS/ CII)	Position relative to CII + III/CS reference values	Position relative to CII + III/CII reference values	CII + III deficiency confirmed	Initial identified RC enzyme deficiency
P72	M, 8	A	M	2.95	0	0	<5%	<5%	Yes	II+III
P37	M, 1	A	DD, CNS, E, M, H, LA	3.14	87	1.48	<10%	>10%	No	III, IV, II+III
P32	M, 3	A	FTT, DD, M, GS, C, H, LA	3.30	138	1.99	>10%	>10%	No	I, III, IV, II+III
P36	F, 1	A	DD, CNS, M	3.51	166	1.87	>10%	>10%	No	III, II+III
P87	F, 1	C	FTT, DD, DR, CNS, E, M, RD	3.78	130	1.96	>10%	>10%	No	III, II+III
P86	M, 5	A	FTT, DD, DR, CNS, M, GS, LA	3.90	261	2.36	>10%	>10%	No	I, III, IV, II+III
P39	M, 3	A	DD, M, RD	3.97	132	1.99	>10%	>10%	No	II, III, II+III
			Controls, mean ± SD	3.4 ± 0.3	228 ± 94.8	2 ± 0.8				

Clinical and demographic data provided by Prof I Smuts (Department of Paediatrics, University of Pretoria); *Data obtained from Wilsenach (2014). ID: identification; F: female; M: male; /: not known; A: African; C: Caucasian; FTT: failure to thrive; DD: developmental delay; DR: developmental regression; CNS: central nervous system involvement; E: epilepsy; M: myopathy; C: cardiac involvement; C: cardiomyopathy; GS: gastrointestinal dysmotility; H: hepatopathy; RD: renal disease; LA: lactic acidosis; <5%: enzyme activity below 5th percentile of CRC; <10%: enzyme activity below 10th percentile of CRC; >10%: enzyme activity above 10th percentile of CRC; SD: standard deviation. The patient cases with reduced CoQ_{10} levels are indicated on a grey background; the dark grey background representing the one patient case (P54) with reduced CoQ_{10} levels, but normal enzyme activities.

4.2.3 Discussion

The enzyme analyses were repeated in order to validate the diagnosis of a combined CII + III deficiency in the selected patient cohort that was subsequently sequenced. In Smuts *et al.* (2010) it is reported that a relatively high prevalence of CII + III deficiency (37%) is found in this cohort and the possibility that this high prevalence could subsequently prove to be inaccurate has been considered (personal communication: Prof F.H. van der Westhuizen, Centre for Human Metabonomics, NWU). General variables that influence enzyme activity, such as pH and salt concentrations, are automatically improved when eliminating batch-to-batch variation as much as possible. It was generally accepted that the results obtained from this repeated analyses were more reliable due to the eliminating attempt regarding batch-to-batch, and even day-to-day variations. The Transformation Kernel Density Estimation distribution used on the data of the normalised reference values allowed accurate density estimation by requiring only a small number of samples (Sheather & Marron, 1990; Smuts *et al.,* 2010).

A total of 25 patients of whom four (16%) were Caucasian males, eight (32%) African males, three (12%) Caucasian females and 10 (40%) African females were included in the analyses. The diagnostic outcomes of the repeated enzyme activity measurements differed from the initial CII + III enzyme diagnosed deficiencies which were compiled over the past years (Table 4-2) which were more prone to batch-to-batch variation. As can be seen in Table 4-2, 5/25 (20%) of the patients in the selected cohort confirmed with a CII + III enzyme deficiency based on the control group selected for this study. This included 2/8 (25%) of the patient cases with reduced levels CoQ_{10} ($<5^{th}$ percentile of the CRC) and 3/16 (19%) of the patient cases with relative normal levels CoQ_{10} ($<5^{th}$ percentile of the CRC) measured from the previous study performed by Wilsenach (2014). It was interesting to note that 7/7 (100%) of the Caucasian patients in the patient cohort did not present CII + III deficiency in the repeated analysis, especially since Smuts *et al.* (2010) observed that it was patients of African descendants that significantly presented with the combined CII + III deficiencies. Based on the newly obtained enzymatic results 20/25 (80%) patients presented with no CII + III enzyme activity deficiency based on the specific criteria used in this study.

It should be taken into account that CII + III was not normalised to additional RC complexes (notably CIV) as included in the previous diagnosis and the control group used in this current study were referred patients with “normal” CII + III activities compared to previously healthy controls, as mentioned in Section 3.2. It was not an objective of this study to do a detailed investigation of the reference ranges of RC enzymes. However, based on the data obtained in

this study it illustrates the importance of the selection of reference controls and the impact it has on the diagnosis of RCD. With referred patient samples (albeit having previously normal CII + III activities) included in this study to determine the reference range, it is possible that this reference range distribution would be biased towards the lower ranges of CII + III relative to healthy controls as initially used in this cohort. This would tend to omit false positives but could increase the possibility for false negatives when identifying CII + III deficiencies. On the other hand the opposite should also be considered, i.e. the effect of the reference range distribution being biased towards the higher values of CII + III when using healthy controls. Such an effect would then support the mentioned concern raised about the unusual high number and cases identified with CII + III deficiencies previously reported in this cohort. Whatever the case may be, this again shows the importance of RC enzyme diagnostic reference ranges (which is beyond the scope of this study) and highlights the ethical dilemma when key previous diagnostic data are re-evaluated and a different outcome results. Importantly, it also illustrates the great importance of molecular genetic investigation, such is this study, to identify beyond doubt the primary RCD deficiencies in this cohort.

4.3 Sequencing data

The coding sequences of all 18 nuclear-encoded CoQ_{10} associated genes mentioned in Section 3.6 (Table 3-4) of the 24 initially identified combined CII + III RC deficiency patients (eight patients with low levels of CoQ_{10}, one patient (P54) with low CoQ_{10} but no enzyme deficiency, as well as 15 patients with normal levels of CoQ_{10}), were sequenced. This was done using a NGS (Ion Torrent) approach as described in Section 3.6. From the sequence report it was summarised that the ISP loading (chip loading) obtained was 68% (or 32% empty wells) of which a final library of 56% was obtained after the filtering of polyclonal and low quality reads that allowed a total of 1,900,543 (46%) usable reads. A mean amplicon read length ranging from 137 bp - 146 bp was obtained at an average of 167x base pair coverage, and an average accuracy of 98%. From the sequencing run, and after comparison against reference sequences, a total of 1368 variants were identified in 24 patients. Most of the variants were common (frequently found in populations) SNPs. Section 3.7 gives a detailed description of the complex bioinformatics workflow used to analyse the large amount of data obtained from the sequence runs. Through the primary and secondary data mining processes common single nucleotide variations (SNVs) were filtered from the data to ultimately present possible disease-causing variants important to this study, of which the focus was on the CoQ_{10} biosynthesis coding genes.

4.3.1 Results

4.3.1.1 *Overview of all variants*

From the bioinformatics investigation 16 possible disease-causing variants (if not counting the same variant more than once) at any allele frequency were identified in the patient cohort. Of the 16 variants 12 were previously reported variants and 4 possible novel variants that could cause LoF, or be probably damaging. All 16 possible disease-causing variants identified were substitutions and are presented in detail in Appendix C. Table 4-3 summarizes these findings as well as other relevant information for the variants identified in the patients with reduced CoQ_{10} (patients ID indicated in red), and patients with normal CoQ_{10} levels (patient ID indicated in blue). The HGVS names in Table 4-3 indicated the exact position in which the variants occurred in the specified chromosome of the gene of which a single nucleotide change altered the amino acid change, or formed a stop (termination) codon (*) at the variant site in the given gene. Important also for interpretation is to note the percentage provided by the 1000G data, which presents the frequency of the variant occurring in the African (AFR) and European (EUR) (which includes Caucasian) populations, if recorded at all.

Table 4-3: Possible disease causing variants identified in output sequenced data during bioinformatics analysis.

Gene	Variant HGVS ID / Transcript ID	PPS	SPS	1000G	Allele call	RefSNP	Patient ID
COQ6	c.41G>A p.Trp14* ENST00000394026	Not known	Not known	AFR: 5% EUR: 0% ALL: 1%	Heterozygous	rs17094161	**P2**, **P24**, P39, P68, **P98**, P101
ETFB	c.462G>A p.Thr154Met ENST00000309244	B: 0.4 B: 0.82	D: 0.03	AFR: 49% EUR: 56% ALL: 46%	Homozygous[#] Heterozygous	rs1130426	P29[#], **P2**, **P14**, P21, P32, P36, P37, **P54**, P68, **P78**, P85, P87, P88, P94, **P95**, **P98**, P101
ETFA	c.513G>A p.Thr171Ile ENST0000055794	PRD: 1	D: 0	AFR: 0% EUR: 9% ALL: 5%	Heterozygous	rs1801591	**P99**
COQ10A	c.376C>T p.Arg126Cys ENST00000308197	PRD: 0.99	D: 0.02	AFR: 0 % EUR: 0% ALL: 0%	Heterozygous	rs187670547	P32
APTX	c.597C>T p.Arg200His ENST00000436040	B: 0.32	D: 0.05	AFR: 2% EUR: 0% ALL: 0%	Heterozygous	rs150886026	**P24**, P101
APOE	c.388T>C p.Cys130Arg ENST00000252486	B: 0.002	T: 1	AFR: 26% EUR: 14% ALL: 15%	Heterozygous	rs429358	**P2**, **P24**, P39, P87, **P99**

Gene	Variant HGVS ID / Transcript ID	PPS	SPS	1000G	Allele call	RefSNP	Patient ID
COQ6	c.859G>T p.Ala287Ser ENST00000394026	PRD: 0.98	T: 0.15	AFR: 3% EUR: 0% ALL: 1%	Heterozygous	rs61743884	**P2**
COQ6	c.283G>A p.Gly92Asp ENST00000334571	PRD: 0.96	D: 0	AFR: 5% EUR: 0% ALL: 0%	Heterozygous	rs61743864	**P29**
COQ9	c.304C>T p.Arg102Cys ENST00000262507	PRD: 0.97	D: 0	AFR: - EUR: - ALL: -	Heterozygous	rs143043228	**P29**
COQ6	c.475G>T p.Val163Phe ENST00000334571	PD: 0.67	D: 0	AFR: 6% EUR: 0% ALL: 1%	Heterozygous	rs111833521	**P43**
COQ7	c.308C>T p.Thr102Met ENST00000569312	PD: 0.68	D: 0.03 CH: 0	AFR: 30% EUR: 64% ALL: 61%	Homozygous[#] Heterozygous	rs11074359	**P86**[#], **P99**[#], **P101**[#], **P14**, **P24**, **P39**, **P43**, **P54**, **P88**, **P94**, **P95**, **P98**
COQ7	c.371C>T p.Thr123Met ENST00000569312	B: 0.029	D: 0	AFR: 3% EUR: 0% ALL: 1%	Heterozygous	rs114401101	**P43**
APOE	c.326C>T p.Ala109Val ENST00000252486	PRD: 0.98	D: 0.03	AFR: - EUR: - ALL: -	Heterozygous	Novel	**P34**

Gene	Variant HGVS ID / Transcript ID	PPS	SPS	1000G	Allele call	RefSNP	Patient ID
ETFDH	c.1067G>A p.Gly356Glu ENST00000511912	PRD: 0.99	D: 0	AFR: - EUR: - ALL: -	Heterozygous	Novel	**P78**
ETFDH	c.1448C>T p.Pro483Leu ENST00000511912	PRD: 0.99	D: 0.01	AFR: - EUR: - ALL: -	Heterozygous	Novel	**P78**
COQ10A	c.523C>T p.Arg175* ENST00000308197	Not known	Not known	AFR: - EUR: - ALL: -	Heterozygous	Novel	**P99**

HGVS: Human Genome Variation Society; ID: identification; A: adenine; T: thymine; C: cytosine; G: guanine; Trp: tryptophan; *: created stop codon; Met: methionine; Arg: arginine; Cys: cysteine; His: histidine; Thr: threonine; Ile: isoleucine; Ala: alanine; Ser: serine; Gly: glycine; Asp: aspartic acid; Val: valine; Phe: phenylalanine; Glu: glutamic acid; Pro: proline; Leu: leucine; PPS: polyphen prediction score; SPS: SIFT prediction score; B: benign; PD: possibly damaging; PRD: probably damaging; D: deleterious; T: tolerated; -: not available; 1000G: allele frequency of the variant of samples in African (AFR), European (including Caucasian) (EUR) and all populations (ALL) based on the 1000 Genomes project; P: patient; $^{\#}$: patients with homozygous allele calls; RefSNP: reference SNP. The patient case IDs with reduced CoQ_{10} levels are indicated in red, and the patient control IDs with normal CoQ_{10} levels are indicated in blue.

As represented in Figure 4-2, where the data from Table 4-1 is illustrated in another way, it can be seen that the number of possible disease-causing variants identified in the patient cohort with low and normal CoQ_{10} levels did not vary in terms of the amount of possible disease-causing variants identified. For the patients with reduced CoQ_{10} levels a total of 537 variants were identified during a sequencing run, the number variants ranging from 47-70 (mean ± SD; 60 ± 9) per patient. Of these variants, 12/16 (75%) of the possible disease-causing variants were at least identified once in each patient case (100%). For the patients with normal CoQ_{10} levels, a total of 831 variants were identified during the sequencing run, the number variants ranging from 48-71 (mean ± SD; 55 ± 7) for each patient. Of these variants, 9/16 (56%) were from the possible disease-causing variants identified in 14/15 (93%) of the control patient cohort. From this it was concluded that there are no obvious high prevalence of specific variants in the low CoQ_{10} group, which could then have been investigated statistically, which has prompted further interpretation of the data to mainly focus on a case-by-case manner.

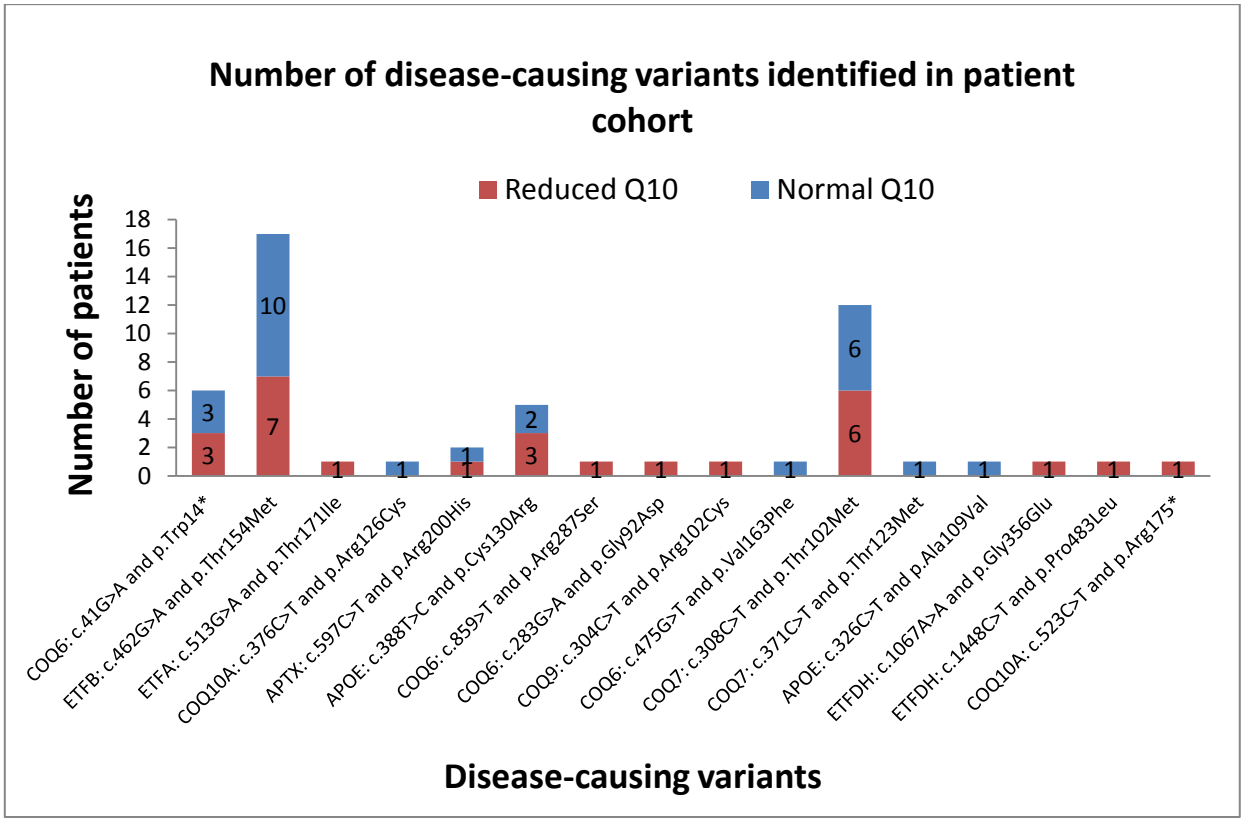

Figure 4-2: Bar plot representing the number of possible disease-causing variant identified in the patient cohort. The number of variants in patients with reduced CoQ_{10} levels is compared to those in patients with normal CoQ_{10} levels. Q10: coenzymeQ_{10}; A: adenine; T: thymine; C: cytosine; G: guanine; Trp: tryptophan; *: created stop codon; Thr: threonine; Met: methionine; Arg: arginine; Cys: cysteine; His: histidine; Ile: isoleucine; Ala: alanine; Ser: serine; Gly: glycine; Asp: aspartic acid; Val: valine; Phe: phenylalanine; Glu: glutamic acid; Pro: proline; Leu: leucine.

4.3.2 Discussion

The 16 variants on genes encoding the enzymes responsible for CoQ_{10} biosynthesis responsible for primary CoQ_{10} deficiency, and candidate genes likely to cause secondary CoQ_{10} deficiency in the patient cohort, were presented in Table 4-3. These variants are further discussed by considering the *previously reported* possible disease-causing variants ("known mutations") and *novel* possible disease-causing variants (candidate "novel mutations") in a case-by-case manner.

4.3.2.1 *Overview of previously reported possible disease-causing variants*

Section 2.6 gave a detailed description concerning the different mutations leading to CoQ_{10} deficiency. From the 16 variants identified in the patient cohort there was no obvious case with a homozygous pathogenic variant that could easily be identified as clearly disease-causing. Of the variants listed, and considering all other data available from the patients, a number of variants were identified as possibly resulting in a CoQ_{10} deficiency which merited further investigation. These variants are further discussed in the following sections with the aim of comparing previous findings (e.g. clinical and biochemical profile) of these variants with the available data from cases in this study.

4.3.2.1.1 Heterozygous truncate variants in the COQ6 gene

Although the precise enzyme functions for most *COQ* genes are still not known *COQ6* encodes a mono-oxygenase necessary for the biosynthesis of CoQ_{10}. Pathogenic mutations in *COQ6* have been identified as homozygous and compound heterozygous, associated with the multi-systemic nephrotic syndrome, which includes sensorineural deafness and renal failure. *COQ6* gene analysis performed in families also identified two patients with a nephrotic syndrome with heterozygous truncating mutations (Heeringa *et al.,* 2011).

A heterozygous truncate variant (rs17094161) on chromosome 14 was identified in the *COQ6* gene in six patients (P2, P24, P39, P68, P98 and P101). It is a 41G>A transition in exon 1, resulting in a Trp14Ter substitution in the monooxygenase area. The identified variation in *COQ6* was identified to cause LoF by introducing a stop codon at the variant site, predicting a high severity impact. The substitutions observed in the six patients all belong to the African population (four female, two male) at *heterozygous* allele frequencies. There have not yet been a clinical phenotype associated with this specific variant (c.41G>A / p.Trp14Ter), and it has been reported to be present in 5% of the African population and 1% of all populations according to the reference 1000G database. Patients P2, P24, P39, P68 and P101 presented with development delay and myopathy (P98 had cardiomyopathy). P101 also had failure to thrive,

dysmorphic features and ptosis resembling King-Denborough syndrome. P68 had migraine, nystagmus and visual loss, and P39 presented with nystagmus and renal disease. P24, P39 and P101 had multiple different RC enzyme deficiencies. Although all six patients with this variant was originally diagnosed with a CII + III deficiency, only P2 and P98 had CII + III deficiencies when repeated in this study (Table 4-2).

LoF variants vary in the degree that influences normal gene function. Homozygous genotypes are lethal in the sense of producing no functional gene product resulting in different alteration events including the block of transcription and lack of enzyme activity. LoF variants are usually recessive that only influence some protein function and was therefore excluded from further investigation (except for the possible compound heterozygous mutation in P2) based on a single good copy of the gene which is most probably sufficient (Griffiths *et al.,* 2000, since this specific variant has been reported to be recorded in 5% of the African population.

4.3.2.1.2 Compound heterozygous variant in COQ6 gene

A molecular genetic investigation of mtDNA involvement was performed on the same cohort of patients used in this study in the doctoral study of van der Walt (2011). In this thesis African female patient (P2) was identified to have muscle mtDNA depletion suspected to have a possible nuclear gene variant present to result in the given depletion. In this present study two probably damaging variants (rs17094161, rs61743884) identified in P2 in gene *COQ6* supports the possible identification of a compound heterozygous gene mutation (c.[41G>A] + [859G>T] and p.[Trp14Ter] + [Ala287Ser]). The first truncate variant (rs17094161) just mentioned before (Section 4.3.2.1.1) was compounded by the additional 859G>T variant (rs61743884) in exon 1, which results in an Ala287Ser substitution in the mono-oxygenase domain (c.859G>T / p.Ala287Ser). The amino acid change is a non-synonymous coding with a predicted medium severity impact of the affected transcript. The amino acid change is furthermore indicated to be probably damaging owing to its predicted PPS and SPS scores. The variant is not recognized to cause any known clinical phenotype and have been recorded to be present in 3% of the African population and 1% of all populations according to the 1000G data. P2 presented a clinical profile consisting of developmental delay as well as myopathy which is not associated with the typical *COQ6* deficient multi-systemic phenotype. Lipid bodies serving as reservoirs for cholesterol were identified in the tissue of P2, which can be ascribed to CoQ_{10} biosynthesis impairment by incorporating cholesterol from the excess substrates available through the mevalonate pathway. Although the variants identified in gene *COQ6* in this study have formerly been reported as a dbSNP, it still lacks any detailed information such as clinical phenotypes as well as SIFT and PolyPhen-2 scores used to predict pathogenicity. Nevertheless, this case was further investigated as described in Section 4.4.

4.3.2.1.3 Possible compound heterozygous variant in COQ7 gene

The required enzyme function of *COQ7* in the biosynthesis of CoQ_{10} is not yet known. An African female patient (P43) presented with a possible compound heterozygous variant (c.[308C>T] + [371C>T] and p.[Thr102Met] + [Thr123Met]) identified on chromosome 16 of the *COQ7* gene. The first variant identified in the *COQ7* gene was a heterozygous variant (rs11074359), with a 308C>T transition in exon 3, resulting in a Thr102Met substitution with a severity impact predicted to be possibly damaging and deleterious. The second variant identified in the *COQ7* gene, with 371C>T transition in exon 3, resulted in a Thr123Met substitution (rs114401101) with a medium severity impact predicted to be possibly damaging and deleterious according to the PPS and SPS scores. These identified variants have respectively been identified in 30% and 3% of the African population according to the 1000G database. There are no recorded CoQ_{10} deficiencies resulting from pathogenic variants in the *COQ7* gene, therefore also no clinical presentations that could be compared to these variants in P43. P43 presented a clinical phenotype which included failure to thrive, developmental delay, myopathy, ptosis, as well as lactic acidosis and CIV deficiency. It therefore also warranted further investigation as described in Section 4.4.

4.3.2.1.4 Other known heterozygous variants of note identified

P29 (African male) contained a reported heterozygous variant rs61743864 on chromosome 14 in gene *COQ6* (c.283G>A / p.Gly92Asp) in addition to the reported variant rs13043228 on chromosome 7 in gene *COQ9* (c.304C>T / p.Arg102Cys) with no known clinical phenotype associated with this variant. According to the PPS and SPS severity scores it was predicted that the variant in gene *COQ6,* which occurs in 5% of the African population, and the variant in the *COQ9* gene, which has no available frequency in any population, are probably damaging in the most affected transcript. P29 revealed a clinical profile consisting of only muscle myopathy, a clinical profile that is not associated with the multi-systemic disease reported associated with *COQ6* and *COQ9* deficient genes. This specific case was not included in further investigations for having heterozygous variants on two distinct loci.

An African male (P32) presented with a previous reported heterozygous variant (rs187670547) in the *COQ10A* gene (c.376C>T / p.Arg126Cys). No CoQ_{10} deficiency has been associated with the *COQ10* gene and although the variant has been reported, it has only been identified in 0.2% of the African population. It is predicted to be probably damaging and deleterious according to its predicted PPS and SPS scores. P32 presented a clinical phenotype which included failure to thrive, developmental delay, myopathy, gastrointestinal dysmotility, cardiac involvement, hypothyroidism, short stature, hepatopathy and lactic acidosis. P32 further included multiple complex deficiencies including CI, CII + III, CIII and CIV, of which a CI mutations in the *NDUFV2*

gene previous reported by Bénit *et al.,* (2003) has been identified recently in the patient with a homozygous allele call in an ongoing doctoral study at the NWU, Potchefstroom. The African patient described by Bénit and colleagues presented a phenotype including cardiomyopathy, hyptonia and encephalomyopathy, a phenotype which is comparable to that of P32 (Bénit *et al.*, 2003). This specific case was not included in further investigations because of its heterozygous allele frequency.

4.3.2.2 *Overview of novel possible disease-causing variants*

A number of novel variants which could possibly be disease-causing (pathogenic) were identified in several cases of the study cohort. These novel variants are listed in the following sub-sections.

4.3.2.2.1 Novel heterozygous variant in APOE gene

A novel variant on chromosome 19 in the *APOE* gene (c.326C>T / p.Ala109Val) was identified in P34 although at a heterozygous allele frequency. The PPS and SPS of the amino acid change predict the variant to be both deleterious and probably damaging when homozygous. The phenotype of *APOE* diplotypes have been related to be involved with cholesterol metabolism, atherosclerosis as well as lipid peroxidation. Although *APOE* is not a gene involved in CoQ_{10} deficiency it is said to be involved in CoQ_{10} metabolism, connected to CoQ_{10} biosynthesis which initiates with the mevalonate pathway. In a study by Fischer *et al.* (2011), two common genetic SNPs associated with the *APOE* gene were investigated on CoQ_{10} status which established that *APOE* polymorphisms may indeed affect CoQ_{10} status. P34 presented a clinical phenotype which included eye involvement, developmental delay and myopathy, with further multiple complex (CI, CII + III, CIII and CIV) deficiencies initially identified via spectrophotometric enzymatic analysis. Before attention was shifted to nDNA mutations causing RCDs, it was primarily directed and focused on maternally inherited mtDNA mutation in the late 1980's (Musumeci *et al.,* 2011). Decreased levels of CoQ_{10} have been reported to be associated with mtDNA mutations. It was interesting to note that a novel mtDNA homoplasmic (m.13790A>G) variant was identified in this patient in the doctoral thesis of van der Walt (2011). Further investigation was recommended in order to determine the pathogenicity of this found mtDNA variation which indicated *in silico* predictions of moderate pathogenicity.

4.3.2.2.2 Heterozygous truncate variant in gene COQ10A

There are currently no described pathogenic variants associated with *COQ10A*. P99 presented a novel heterozygous truncate variant identified on chromosome 12 in the *COQ10A* gene (c.523C>T / p.Arg175Ter). The identified variant in *COQ10A* was identified to cause loss of

gene function by introducing a stop codon at the variant site, predicting a high severity impact. P99 presented with failure to thrive, myopathy, gastrointestinal dysmotility, lactic acidosis and CIII deficiency. As LoF variants differ in the degree that it influences normal gene function, it is not certain to what degree the lack of gene activity occur. The heterozygous allele frequency and low quality score affected the decision to exclude this identified variant from further investigation. The intracellular transport of CoQ is a field in which there is a relative lack of information. By using the model of yeast CoQ biosynthesis, the *COQ10* gene in which deletions occur, is known to still produce the necessary amounts of CoQ needed although the CoQ cannot be used for respiration. The gene encodes a small lipophilic protein which binds and transports CoQ from the site of synthesis to the location where needed for electron transport. The two human homologs (*COQ10A, COQ10B*) are yet to be characterised (Trevisson *et al.,* 2011).

4.3.2.2.3 Novel compound heterozygous variant in ETFDH gene

In a female Caucasian patient (P78) with the lowest detected CoQ_{10} levels of the cohort, a possible novel compound heterozygous variant c.[1067G>A] + c.[1448C>T] and p.[Gly356Glu] + p.[Pro483Leu] on the *ETFDH* gene (located on chromosome 14) was identified. Both variants were predicted to be severely deleterious and probably damaging according to the PPS and SPS sores. P78 presented with developmental delay, myopathy, hepatopathy, central nervous system involvement and migraine that is more extensive than the previously reported secondary CoQ_{10} deficiency identified in gene *ETFDH* associated with an isolated myopathy (Gempel *et al.,* 2007). P78 also had CIII and CIV deficiencies with a more than 200% increased citrate synthase activity of the normal control patients' activity. *ETFDH* deficiency can result in CoQ_{10} deficiency by the lack of the produced enzyme, or the faulty binding to $CoQ_{10,}$ potentially resulting in degradation of CoQ_{10} which is the direct acceptor molecule of electrons from the electron transferring flavoproteins (Olsen *et al.,* 2007). This case was also included in further investigation as described in the next two sections.

4.4 Sanger Sequence validation

Sanger sequencing (dideoxy- or chain termination sequencing) was used for the validation of the variants found during the NGS data analysis. As mentioned in Section 3.8 validation of the three compound heterozygous variants identified in the three patients with probable pathogenic predictions in Sections 4.3.2.1.2 (P2), 4.3.2.1.3 (P43), and 4.3.2.2.3 (P78), respectively, were confirmed by Sanger sequencing by Inqaba Biotech (Pretoria, South Africa). This was done in order to confirm if the three variants investigated further were present and at heterozygous allele frequencies. Figures 4-3, 4-4 and 4-5 represents the validated heterozygous nucleotide

base changes identified in each patient and gene respectively, by illustrating partial alignment of the reference and patient sequences.

Table 4-4: Possible compound heterozygous variants identified by sequence analysis.

Patient ID	Gene	refSNP	HGVS ID	CII+ III deficiency confirmed
P2	*COQ6*	rs17094161	c.41G>A p.Trp14Ter	Yes
		rs61743884	c.859G>T p.Ala47Ser	
P43	*COQ7*	rs11074359	c.308C>T p.Thr103Met	Yes
		rs114401101	c.371C>T p.Thr124Met	
P78	*ETFDH*	NA	c.1067G>A p.Gly356Glu	No
		NA	c.1448C>T p.Pro483Leu	

ID: identification; RefSNP: reference SNP; NA: not available; HGVS: Human Genome Variation Society; A: adenine; T: thymine; C: cytosine; G: guanine; Trp: tryptophan; Ter: termination; Ala: alanine; Ser: serine; Thr: threonine; Met: methionine; Gly: glycine; Glu: glutamic acid; Pro: proline; Leu: leucine.

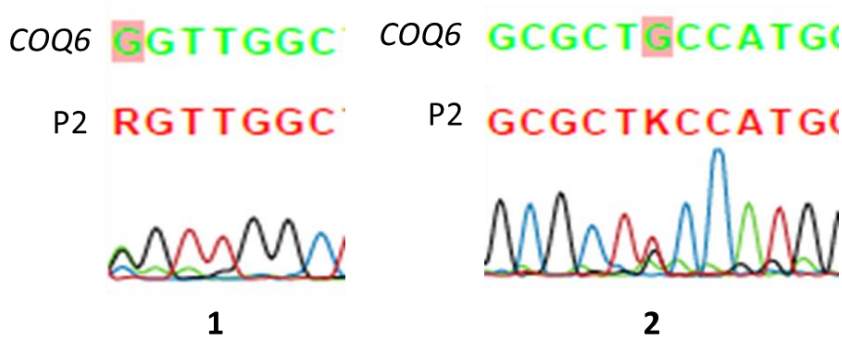

Figure 4-3: Partial sequence alignment of reference sequence *COQ6* (wild type) and P2 (variant type). This figure shows 1) the heterozygous G to A nucleotide base change (**R**, c.[41G>A]), and 2) the heterozygous G to T nucleotide base change (**K**, c.[859G>T]). The reference sequence was the *COQ6* gene sequence as published on Ensembl (ENS00000119723). Green line: adenine (A); black line: guanine (G); red line: tyrosine (T): blue line: cytosine (C).

The compound heterozygous variants c.[41G>A] + [859G>T] and p.[Trp14Ter] + [Ala287Ser] identified in gene *COQ6* of P2 during the sequence analysis was validated by forward and reverse DNA sequencing. The sequence of P2 was aligned to reference sequence *COQ6*, and was identical with only the single base changes from G to A (1) and G to T (2), respectively.

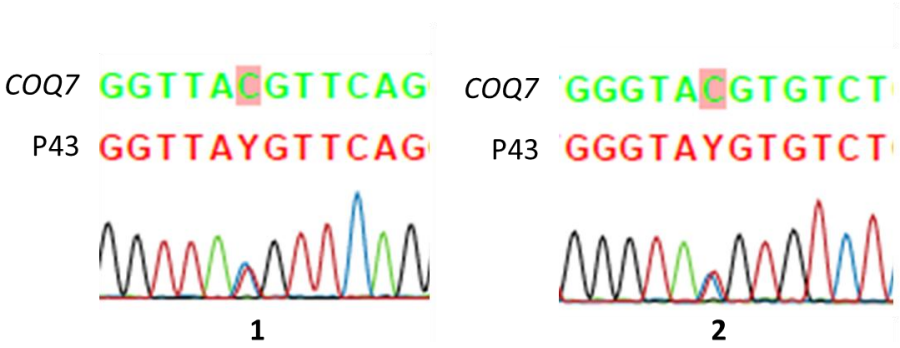

Figure 4-4: Partial sequence alignment of reference sequence *COQ7* (wild type) and P43 (variant type). This figure shows the heterozygous C to T nucleotide base changes (**Y**, c.[308C>T] and c.[371C>T]) in the sequence represented by 1 and 2 respectively. The reference sequence was the *COQ7* gene sequence as published on Ensembl (ENSG00000167186). Green line: adenine (A); black line: guanine (G); red line: tyrosine (T): blue line: cytosine (C).

The compound heterozygous variants c.[308C>T] + [371C>T] and p.[Thr102Met] + [Thr123Met] identified in gene *COQ7* of P43 during the sequence analysis was validated by forward and reverse DNA sequencing. The sequence of P43 was aligned to reference sequence *COQ7*, and was identical with only the single base changes at both variant sites from C to T.

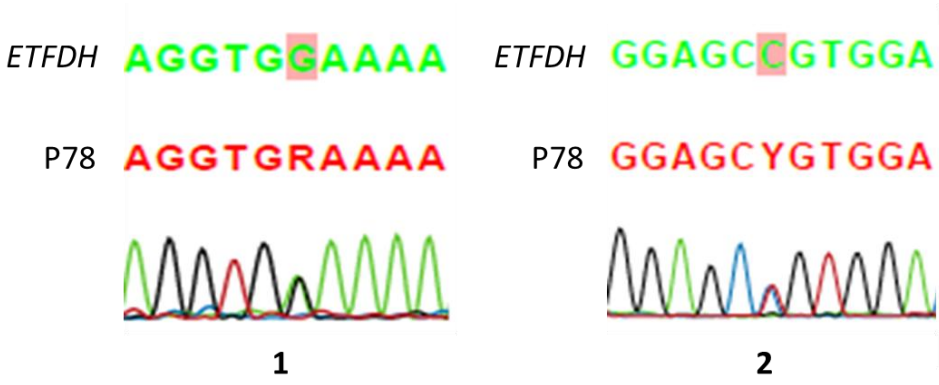

Figure 4-5: Partial sequence alignment of reference sequence *ETFDH* (wild type) and P78 (variant type). This figure shows 1) the heterozygous G to A nucleotide base change (**R**, c.[1067G>A]), and 2) the heterozygous C to T nucleotide base change (**Y**, c.[1448C>T]). The reference sequence was the *ETFDH* gene sequence as published on Ensembl (ENS00000171503). Green line: adenine (A); black line: guanine (G); red line: tyrosine (T): blue line: cytosine (C).

The *novel* compound heterozygous variants c.[1067G>A] + c.[1448C>T] and p.[Gly356Glu] + p.[Pro483Leu] identified in gene *ETFDH* of P78 during the sequence analysis was validated by forward and reverse DNA sequencing. The sequence of P78 was aligned to reference sequence *ETFDH*, and was identical with only the single base changes from G to T (1) and C to T (2), respectively.

From this section, describing an external validation of selected sequence data, it was concluded that the NGS data of these three heterozygous variants were correct and that the further protein analyses could be performed.

4.5 SDS-PAGE and western blot analysis

As discussed in Sections 3.9 and 4.4 the three possible compound heterozygous variants identified in genes *COQ6*, *COQ7 and ETFDH* in patients P2, P43 and P78 respectively (Table 4-4) were identified for further investigation. With the material available and without the possibility of obtaining additional material (e.g. fibroblasts) from these cases at the time of the study, the further investigation was limited to measuring the steady state levels (inferring expression and stability) of these proteins using SDS-PAGE and western blot analysis. The remaining 600 *g* supernatants available from these patients were used as discussed in Section 3.3.3. The three patient samples, together with five control samples (CRC) where no CoQ_{10} or RCD enzyme deficiencies were detected, were analysed on a SDS-PAGE gel and then transferred to a polyvinylidenedifluoride (PVDF) membrane during western blotting and individually analysed. The relative quantities of the individual proteins were detected by either a monoclonal mouse anti-ETFDH antibody or monoclonal rabbit anti-COQ6/COQ7 antibodies, and then detected using goat-anti-mouse or goat-anti-rabbit horse radish conjugated polyclonal antibodies respectively. This allowed for detection of the individual proteins by the secondary antibodies labelled with horseradish peroxide (HRP) enzyme. The HRP emits light in the presence of substrates for the enzyme which are then captured and quantified. The band intensities which are equivalent to the amount of proteins were directly measured by chemiluminescence produced and quantified relative to total amount of loaded protein normalisation. The normalised individual patient samples were compared to the patient control samples in each specific protein analysed.

4.5.1 Results

The western blots, stain free blots (used to normalize each sample individually as described in Section 3.9.2) and normalised relative band intensities can be seen in Figures 4-6, 4-7 and 4-8 showing the final analysed results of the individual genes of 600 *g* muscle tissue supernatant of each patient respectively.

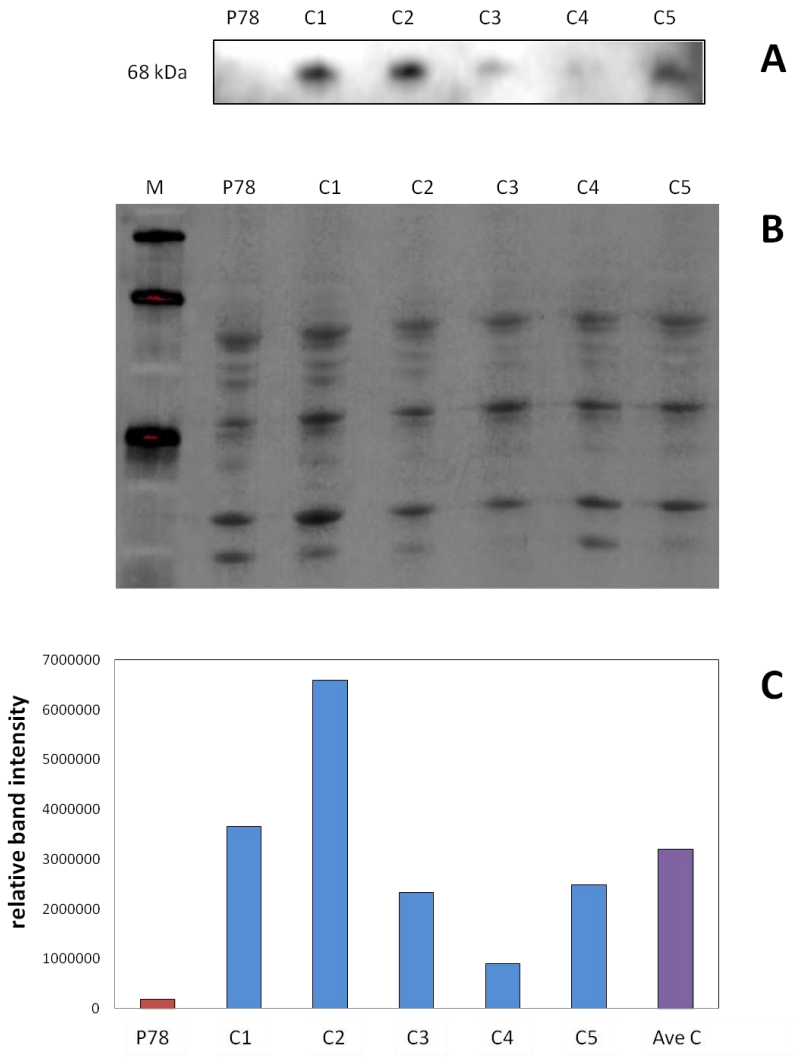

Figure 4-6: Immunoblotting analysis of ETFDH expression in P78 and patient controls. 600 *g* supernatant protein (25 µg) from muscle tissue, separated using SDS-PAGE and transferred using western blotting, was subjected to immunoblot analysis using monoclonal mouse anti-ETFDH antibodies and goat-anti-mouse horse radish conjugated polyclonal antibodies. A) The ETFDH protein was specifically detected as a band at ~68 kDa of the molecular weight marker. B) Stain free blot used for total protein normalization of P78 and patient controls**.** C) The relative band intensities of which the expression level of P78 was notably lower from that of the average- and relative band intensities of the individual controls (CRC). P: patient; C: control; kDa: kilodalton; M: protein molecular weight marker; Ave: average.

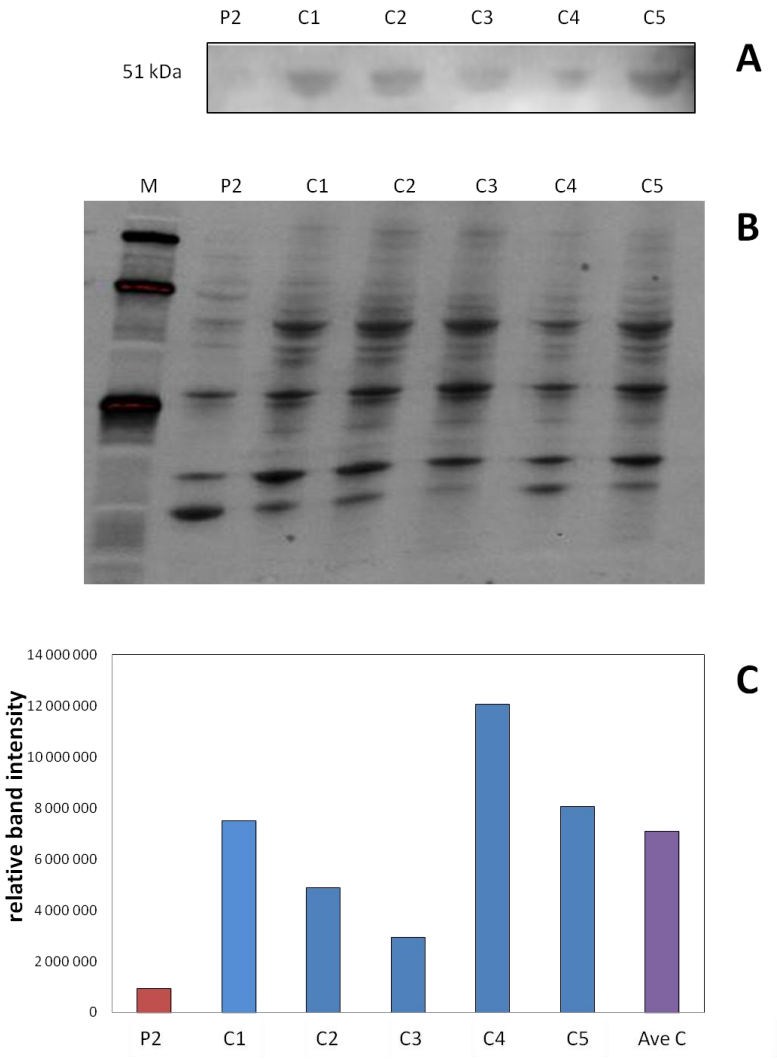

Figure 4-7: Immunoblotting analysis of COQ6 expression in P2 and patient controls. 600 *g* supernatant protein (25 µg) from muscle tissue, separated using SDS-PAGE and transferred using western blotting, muscle tissue was subjected to immunoblot analysis using monoclonal rabbit anti-COQ6 antibodies and goat-anti-rabbit horse radish conjugated polyclonal antibodies. A) The COQ6 protein was specifically detected as a band at ~51 kDa of the molecular weight marker. B) Stain free blot used for total protein normalization of P2 and patient controls. C) The relative band intensities of which the expression level of P2 was notably lower from that of the average- and relative band intensities of the individual controls (CRC). P: patient; C: control; kDa: kilodalton; M: protein molecular weight marker; Ave: average.

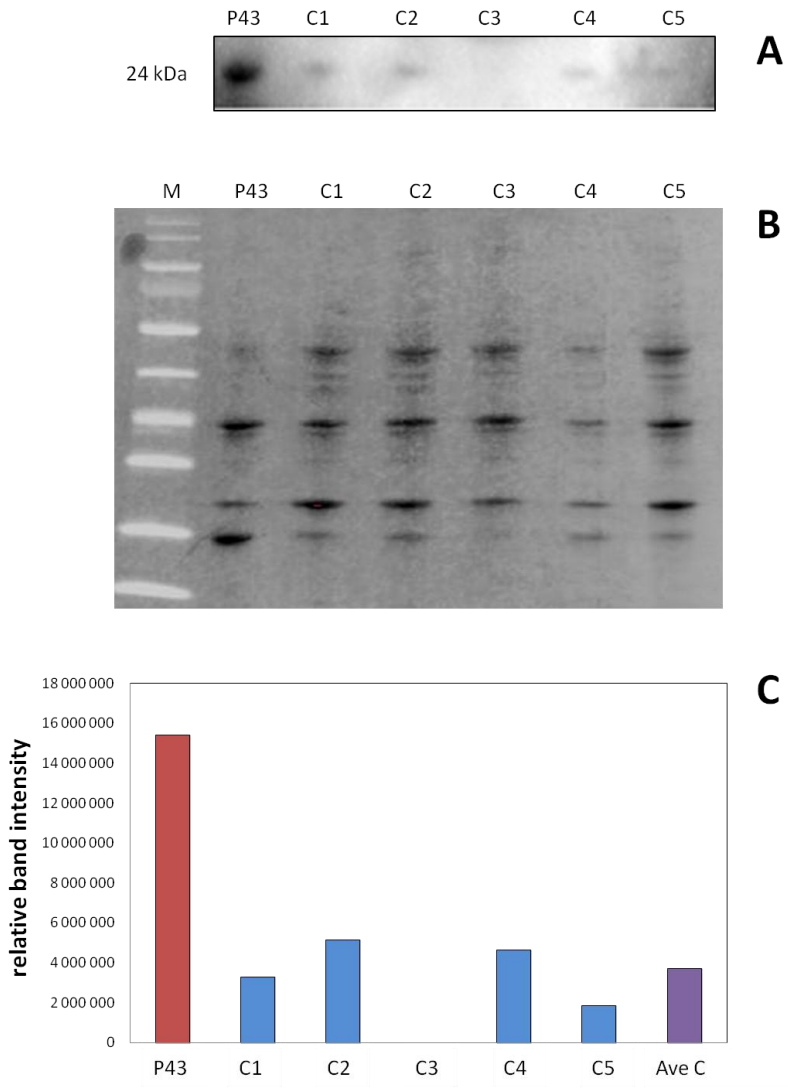

Figure 4-8: Immunoblotting analysis of COQ7 expression in P43 and patient controls. 600 *g* supernatant protein (25 µg) from muscle tissue, separated using SDS-PAGE and transferred using western blotting, muscle tissue was subjected to immunoblot analysis using monoclonal rabbit anti-COQ7 antibodies and goat-anti-rabbit horse radish conjugated polyclonal antibodies. A) The COQ7 protein was specifically detected as a band at ~24 kDa of the molecular weight marker. B) Stain free blot used for total protein normalization of P43 and patient controls. 25 µg 600 *g* muscle supernatant was loaded per lane, followed by each individual sample in each lane being normalized to its total protein content. C) The relative band intensities of which the expression level of P43 was notably higher from that of the average- and relative band intensities of the individual controls (CRC). P: patient; C: control; kDa: kilodalton; M: protein molecular weight marker; Ave: average.

4.5.2 Discussion

Immunoblotting is a very popular and powerful tool used to detect the effect that variants have on gene expression and protein stability in a relative affordable way. Recent studies have focused on the concerns regarding submissions of quantitative western blotting when using housekeeping genes (such as *GAPDH* and actin) as loading controls regarding their variability

(Eaton *et al.,* 2013). Many studies show that housekeeping genes are not only differently expressed quantitatively in different tissues, but also within the same tissue due to cell treatments and environmental changes (Barber *et al.,* 2005; Eaton *et al.,* 2013). Therefore, by eliminating the pitfalls that possibly occur when using housekeeping genes, the Bio-Rad stain-free technology, together with the Bio-Rad Clarity western ECL substrates used for the immunodetection, allowed for accurate gene expression analysis through sensitive protein imaging and total protein normalisation (Figure 4-6).

In this study a very limited amount of available material was present for these analyses although the results presented here could be repeated at least once. As shown in Figures 4-4 and 4-5 the western blot analysis revealed a quantitatively reduced amount of detected proteins ETFDH and COQ6 in P78 and P2 respectively when normalised using total protein analysis. In P78 the residual ETFDH protein amount was 17.6% compared to of the average of controls, and in P2, the residual COQ6 protein amount was 7.7% compared to the average of controls. The fluctuations observed in the protein expressions between the different controls were most probably due to the different steady state proteins present in the muscle tissue of each individual control since these patients were referred patient controls (CRC) and not healthy cases. The decrease in protein ETFDH in P78 was paralleled by the extremely reduced CoQ_{10} levels, lower than that of any patient in this study. As mentioned in Section 4.3.2.2.3, the clinical profile of P78 reported here resembled the isolated myopathy phenotype of previous secondary CoQ_{10} deficiency reported in gene *ETFDH*. The findings observed during the immunoblotting analysis therefore served as validation to the suggestion that the *novel* compound heterozygous variants that were identified via the sequencing analysis in gene *ETFDH* of P78, are the cause of the CoQ_{10} deficiency. One unsolved yet interesting observation was that P78 presented with a CII + III enzyme deficiency when expressed on CII but not when expressed on CS. This same observation was observed in another Caucasian patient (P99), of which a CII + III enzyme deficiency was presented when expressed only on CII. P99 was identified with a heterozygous variant in gene *ETFA* during the bioinformatics workflow (Table 4-3), a gene which also encodes as gene *ETFDH* for flavoprotein dehydrogenase.

The decrease in protein COQ6 in P2 was paralleled by deficient CoQ_{10} levels and the lack of CII + III enzyme activity diagnosed in previous years and confirmed in the current study (Section 4.2). P2 presented with a clinical profile consisting of developmental delay and myopathy, which is not the typical phenotype associated with the multi-systemic phenotype of *COQ6* deficiency. The results observed after the immunoblotting analysis was consistent with the prediction that the compound heterozygous variants identified via the sequencing analysis in gene *COQ6* of P2 is probably the cause of a primary CoQ_{10} deficiency. It can be possible that the distinct

phenotype reported here potentially expands the spectrum of deficiencies associated with *COQ6* dysregulation.

The western blot analysis in Figure 4-6 revealed a drastically increased amount of detected protein COQ7 in P43 when normalised to total protein content. Notably P43 was one of five patients that confirmed with a CII + III enzyme deficiency in this current study. In P43 the over-expressed COQ7 protein amount was more than 400% of that of the average of the controls. The reason for this striking result is clearly not evident from the current data. Molecular mechanisms ensure that certain genes are expressed at the right levels and under the right conditions, and when reduction occurs it is obvious that a certain mutant phenotype is observed concerning the loss of gene function. It is, however, not instinctive to accept that overexpression of a gene can also result in gene and organism disruption. Shastry (1995) reported numerous examples of how not only the absence, but also the overexpression of normal genes, causes human deficiency which can only be fully understood when understanding the variety of mechanisms by which overexpression can possibly cause mutant phenotypes (Prelich, 2012). This observation can also consequently be the result of disrupted stability of the believed multi-enzymatic complex consisting of COQ proteins, by creating non-functional sub-assemblies that are not capable of performing normal CoQ_{10} biosynthesis (Trevisson *et al.,* 2011; Prelich, 2012). It is therefore reasonable to accept the possibility of the compound heterozygous variants identified in gene *COQ7* in P43 to be responsible for the clinical phenotype (which included failure to thrive, developmental delay, myopathy and eye involvement) associated with the patient. It should, however, also be taken into account that the function and biogenesis of mitochondria is dependent on the mysterious interaction of its dual genome (mtDNA and nDNA) (Elstner & Turnbull, 2012) of which protein overexpression can be a potential mechanism that allows adaption to genetic stressors.

Chapter 5

Summary and Conclusions

5.1 Introduction

Chapter 2 presented a literature background concerning aspects of mitochondrial function and dysfunction. CoQ_{10} and its role in the RC and MDs were highlighted, serving as the motivation for this study. The problem statement, aims and objectives were also presented in Chapter 2, while Chapter 3 provided a description of the materials and detailed methods used to perform this study. The final results obtained were given and discussed in Chapter 4. In this chapter the approach, methodology used and the results obtained are discussed and interpreted to come to substantiated conclusions, outcomes and suggestions for future studies.

5.2 Problem statement, aim and objectives

The main role of CoQ_{10} as electron transporter in the RC is only one of the many functions attributed to CoQ_{10}. In addition to the impairment of CoQ_{10} as electron transporter is the very complicated biosynthesis of CoQ_{10} that contributes to the diverse clinical phenotypes associated with CoQ_{10} deficiency. Reduced levels of CoQ_{10} measured in muscle tissue and combined CI + III, and especially CII + III deficiencies are linked to possible CoQ_{10} deficiency (Montero *et al.,* 2008; Lenaz *et al.,* 2007; Quinzii *et al.,* 2007). A recent study by Wilsenach (2014) at the NWU established the positive correlation of this particular hallmark identified within a South African cohort with RCDs. Serving as problem statement was the suspicion that primary CoQ_{10} deficiency was most likely affecting this cohort South African patients identified with combined CII + III deficiencies and reduced levels of CoQ_{10}. The *aim* of the study was *to identify nuclear-encoded mutations in genes associated with CoQ_{10} deficiencies in a cohort of South African patients diagnosed with RCDs* which was achieved by pursuing the following objectives:

5.2.1 Enzyme activity assays in RC deficiency diagnosis

Spectrophotometric enzyme activity assays were performed as the *first objective* in order to confirm if CII + III deficiency was present in the patient cohort selected for this study (since the previous diagnosis stretched over a period of 10 years). The detailed methodology used was described in Section 3.3 and the results obtained described in Section 4.2. In an attempt to minimise potential influencing factors the enzyme analyses were performed over two days, a single complex analysis completed on the same day, eliminating variation in a specific complex and possible day-to-day variation. As discussed in Section 3.3 the samples were analysed in triplicate and contained a control sample used to normalize the inter-batch variations. All the

precautions taken contributed to reliable and high quality RC enzyme activity assay results obtained, although it should be taken into account that the margin for human error will always be existent in manual systems. Even though the results presented good repeatability, in some cases it differed from the previous diagnosis compiled over foregoing years. As presented in Section 4.2.2 in Table 4-2, 20% of the previous diagnosed CII + III deficient patients confirmed the CII + III deficiency when using the diagnostic criteria described. It was concluded that the controls used in this study (CRC), which were different form the healthy controls used in the past, could have been the main reasons for this outcome. Although it was not an objective of this study to do a detailed investigation of the reference ranges of RC enzymes (and even though the protocols used for routine work at the institute are very well defined) the importance and impact that the selection of reference controls have on the diagnosis of RCDs was highlighted in Section 4.2.3.

5.2.2 Sequence analysis

The *second objective* was to generate DNA sequence data of targeted nuclear-encoded CoQ_{10}-associated genes for molecular genetic testing of the selected patient cohort by using Ion Ampliseq custom panel design and the Ion Torrent PGM NGS platform. This was performed to identify if known and/or possible novel variants were the cause of the reduced CoQ_{10} levels measured in the patient cohort. The methods used were described in Section 3.6 in detail and the results represented in Section 4.3. The Ion PGM System software used identified 1368 variants (~57 variants per patient) via the primary data analyses.

5.2.3 Data analysis of sequence analysis

As *third objective* the in-house developed bioinformatics pipeline explained in detail in Section 3.7 was used to identify potential known or novel disease-causing (possible pathogenic) variants obtained in the large amount data generated from the sequencing run. The bioinformatics pipeline identified 16 possible disease-causing CoQ_{10} associated variants (Section 4.3) out of the total of 1368 variants which were largely common SNVs. Of these 16 possible disease-causing variants identified, 12 variants were previously reported and four variants were novel. Two variants were detected to cause loss of gene function and three patients presented possible compound heterozygous mutations. Unfortunately practically all of the rare possible disease-causing variants presented heterozygous allele calls and was therefore excluded from further analysis. It is important to know that there is no correct NGS work base and various databases should be used to compare results to establish which is most maintained to deliver the best reliable results. Fortunately, the combination of tools and information available continues to evolve in order to improve the automated predictions of variants identified during NGS (Pabinger *et al.,* 2014; Ohtake *et al.,* 2014). In this regard it was

also fortunate and very useful that the 1000G dataset was available to inform on allele frequencies in various populations (1000 Genomes Project Consortium, 2012). This, in addition to the *in silico* prediction tools, helped to identify the three compound heterozygous variants as indicated in Table 5-1, which were considered likely to be pathogenic.

5.2.4 Clinical assessment

The *fourth objective* was the comparison of clinical profiles of patients found with CoQ_{10}-associated disease-causing variants for genotype-phenotype correlation contributing to genetic diagnosis. Section 4.4.2 gave a detailed description of each patient's clinical phenotype associated with the different possible disease-causing mutations identified during the bioinformatics workflow. The novel compound heterozygous variant identified in gene *ETFDH* in female Caucasian P78 presented with a similar myopathy phenotype when compared to previous clinical presentations associated with secondary *ETFDH* CoQ_{10} deficient patients, but with added (and more severe) central nervous system involvement (Gempel *et al.,* 2007). The distinct phenotype reported in P2 in gene *COQ6,* which were not similar to previous reported *COQ6* deficiencies, can potentially expand the clinical spectrum of deficiencies associated with *COQ6* deregulation. The same applies to the possible compound heterozygous variants identified in P43 in *COQ7*. Both female African P2 and P43 confirmed with a CII + III deficiency and represented with a strong myopathic phenotype as observed in the before mentioned study conducted by Smuts *et al*. (2010). Distinct difference regarding the ethnicities of a cohort South African patients was recognized, the African patients mostly presented with combined complex deficiency and varying degrees of myopathy in contrast to non-African patients that predominantly presented forms of central nervous system involvement. If indeed the *COQ7* compound heterozygous variant is validated, it will be the first reported *COQ7*-associated primary CoQ_{10} deficiency. Not only will this contribute to the spectrum of clinical phenotypes, it will also add another gene to the list of potential genes to cause primary CoQ_{10} deficiency.

5.2.5 Sanger sequence and SDS-PAGE/western blot validation

Sanger sequence confirmation and protein structure analysis for validation purposes on selected candidate pathogenic variants using SDS-PAGE and western blot analysis served as the *fifth and final objective*. Compound heterozygous mutations are most commonly the pattern of inheritance in rare genetic disorders where parents are non-consanguineous and cannot be determined with confidence through a single DNA sequence analysis of one patient (Kamphan & Krawits, 2012; Kamphans *et al.,* 2013). Since there were no sequences or any DNA samples available for any family members of the three patients identified with possible compound heterozygous mutations (and will not be able to obtain since the families disappeared from the network), it was, unfortunately, not possible to establish with complete certainty whether the

patients inherited one variant from each parent. Evidence was however provided through the validated Sanger sequencing and immunoblot analysis given the reduced relative quantity protein expressions observed in P78 and P2 during ETFDH and COQ6 protein structure validation respectively. P43, on the other hand, presented the interesting observation of significantly overexpressing protein COQ7. Further investigation such as functional enzyme analysis of *COQ7* and/or information surrounding the exact position of the identified variant to the active site on the enzyme, could help validate pathogenicity regarding P43. Although the quality of the immunostaining of SDS-PAGE western blots was not ideal, the data for all three investigations were still considered sufficiently accurate and could be repeated at least once more. These analyses were mostly based on protocols and conditions optimised during an honours study (Jonck, 2013). The background featuring on the blots in Section 4.5.1 is most probably due to insufficient blocking and/or inadequate washing periods and should be experimented with in order to obtain more desirable results in the future.

In conclusion, three compound heterozygous variants identified in genes *COQ6* (P2), *COQ7* (P43) and *ETFDH* (P78) in this study are suggested to be disease-causing given the summary of data presented in Table 5.1.

Table 5-1: Evidence in support of and not in support of concluding three compound heterozygous variants as disease-causing.

Patient ID / Gene / Variant HGVS ID	Evidence supporting cause of disease	Evidence not supporting cause of disease
P2 / *COQ6* / c.[41G>A] + [859G>T] and p.[Trp14Ter] + [Ala287Ser]	• Enzyme data supportive (validated in current study) • CoQ_{10} levels supportive • Immunoblotting supportive	• Some *in silico* predictions indicate tolerated effect • No parental confirmation
P43 / *COQ7* / c.[308C>T] + [371C>T] and p.[Thr102Met] + [Thr123Met]	• Enzyme data supportive (validated in current study) • All *in silico* predictions indicate deleterious effect • Immunoblotting possibly supportive	• CoQ_{10} levels not supportive • No parental confirmation
P78 / *ETFDH* / c.[1067A>A] + [1448C>T] and p.[Gly356Glu] + [Pro483Leu]	• Enzyme data supportive (previous diagnosis) • All *in silico* predictions indicate deleterious effect • CoQ_{10} levels supportive • Immunoblotting supportive	• No parental confirmation

ID: identification; HGVS: Human Genome Variation Society; A: adenine; T: thymine; C: cytosine; G: guanine; Trp: tryptophan; Ter: termination; Ala: alanine; Ser: serine; Thr: threonine; Met: methionine; Gly: glycine; Glu: glutamic acid; Pro: proline; Leu: leucine; CoQ_{10}: coenzymeQ_{10}.

5.3 Final conclusions and recommendations

The following conclusions were made and recommendations proposed from the findings in this study:

1) *Lack of universal standards regarding enzyme analysis.* There is no doubt that the variation in assay protocols and tissue specificity in enzyme analysis still lack universal standards. This includes different concentrations and composition of substrates used, as well as the variation in control references and human controls for each complex with confirmed enzyme deficiency. This consequently hinders the ability to effectively diagnose children RCDs (Thornburn, 2004; Smuts *et al.,* 2010; Smuts & van der Westhuizen, 2010). Attention should therefore be focused on reassessing the RC enzyme diagnostic references used at the institute since the outcome of the diagnosis was clearly influenced when using the exact same protocol, but a different control group as used in previous years.

2) *Irregularity of CoQ_{10} levels and RC complex enzyme activities.* CoQ_{10} deficiency is often associated with varying levels of CoQ_{10} and combined RC complex enzyme activities. CoQ_{10} deficiency has been reported to often have very little or even no effect on CoQ_{10} levels and RC complex enzyme activities as reported in patients with proven *APTX* and *ETFDH* secondary CoQ_{10} deficiencies (Montero *et al.,* 2008; Trevisson *et al.,* 2011). Presented in this study was P78 identified with a novel compound heterozygous variant in gene *ETFDH*, whom had the lowest measured CoQ_{10} levels, but did not confirm with a combined CII + III deficiency. Further evidence provided by this study regarding the irregularity of CoQ_{10} levels and RC enzyme activities was the number of identified possibly disease-causing variants which did not vary between the patient cases with reduced CoQ_{10} levels, and the cases with normal CoQ_{10} levels (Section 4.2.1 and 4.3.1.1). Even though CoQ_{10} levels do play a role in CoQ_{10} deficiency suspicion, it is definitely in conjunction with other factors that regulate secondary CoQ_{10} levels so that reduced CoQ_{10} levels alone are not deleterious to the human genome and not enough evidence to suspect CoQ_{10} deficiency.

3) *Unknown effect of SNPs and LoF variants on CoQ_{10} status.* Low priority attention is given to genes that are particularly associated with heterozygous allele frequencies for being associated with healthy control patients (Ohtake *et al.,* 2014). The majority of the 16

possibly disease-causing variants identified in this study were not prioritised or submit for further investigation due to the heterozygous allele frequencies. These possible disease-causing variants should however not be disregarded, especially by pointing out the unknown effect of common genetic SNPs on CoQ_{10} status. Evidence that common SNPs influence CoQ_{10} levels have been reported (Fischer *et al.,* 2011) and larger cohorts are needed to determine the associations of common SNPs which can possibly influence CoQ_{10} levels. This further includes heterozygous variants causing loss of gene function since it cannot be predicted by a single DNA sequence of how much gene function is influenced by heterozygous allele frequencies. The effect of a heterozygous LoF mutation has been reported to result in possibly CoQ_{10} deficiency (Zhang *et al.*, 2014). This emphasizes the unidentified status of heterozygous LoF variants. The possible CoQ_{10}-associated heterozygous LoF variants identified in this study (both previously reported and novel) should therefore be subjected to further investigation. Studies like these would also benefit from better population genetics information from Southern African populations. Databases such as the 1000G used in this study include African data, but not populations from the Southern part of Africa.

4) *Lack of biochemical markers.* There is no doubt that biochemical profiles for CoQ_{10} deficient patients are lacking and routine laboratory tests have been reported as being inconsistent and not very supportive (Lamperti *et al.,* 2003; Horvath *et al.,* 2011). An intensive metabolic investigation concerning CoQ_{10} deficient patients vs. controls can in the future shed light on these shortcomings as studies by Reinecke *et al.* (2012), Smuts *et al.* (2013) and Venter *et al.* (2014) identified certain possible metabolic markers for RCD patients through a metabolomics investigation. This will not only minimize the selection for candidate genes for genetic analysis, but also contribute to eliminating the invasiveness of muscle biopsy procedures by possibly developing a desired non-invasive screening test for the determination of possible CoQ_{10} deficiency.

5.4 Concluding remarks

Any genes involved in CoQ_{10} biosynthesis can theoretically cause CoQ_{10} deficiency. It is important to recall that this study contributed to a larger overall study on molecular genetic data of causative DNA variants responsible for MDs in South African patients of whom data are lacking. Although studies, including biochemical and biochemical data have been well defined (Smuts *et al.,* 2010), molecular genetic investigations and services characterizing MDs in SA still remain limited. The aim of this study was reached to a large extent by pursuing the five formulated objectives, and the final results obtained from this molecular genetic investigation

could be published as possibly the first reported cases of CoQ_{10} deficiency in South Africa with molecular genetic evidence.

NGS technologies are becoming more popular since the era of genetic medicine which relies on the precise cause of genetic disorders (Rabanni *et al.,* 2013) which highlights the early recognition of potential treatable CoQ_{10} deficiencies. NGS will revolutionize the diagnostics process in years to come but, in order for this to happen, NGS must be commercialised and be made far more affordable by becoming more popular and commonly used in smaller laboratories (Trevisson *et al.,* 2011; Merriman *et al.,* 2012). Although exome sequencing has become such a popular way of sequencing (when taken into account that approximately 85% of all disease-causing variant are located in the exomes), the importance of commercializing NGS is highlighted in order to decrease the gap (of the remaining ~15%) there exists on structural and non-coding regions that can only be accomplished by whole genome sequencing (Rabanni *et al.,* 2013). With the revolutionising technology of NGS that is becoming more promising and affordable, it might soon be possible to better understand the aetiology concerning MDs in South African populations, which will ultimately contribute to accurate molecular diagnosis for genetic counselling supportive aspects.

In conclusion, the aim of this study was *to identify nuclear-encoded mutations in genes associated with CoQ_{10} deficiencies in a cohort of South African patients diagnosed with RCDs*, especially reduced CII + III activity. It was possible to consider pathogenicity of the identified variants identified in the selected cohort by using the different key information including the enzyme-, clinical- and molecular data, as well as the open source databases used during the bioinformatics workflow. Two compound heterozygous variants in *COQ6* (P2) and *ETFDH* (P78), identified in one African female patient and one Caucasian female patient, were identified as pathogenic. Even though the variants lack supportive evidence of pathogenicity regarding genetic data concerning family variation segregation, it was validated by Sanger sequencing as well as supported through the protein structural analysis performed. The motivation for the overexpression of protein COQ7 identified in African female P43 identified with possible disease-causing compound heterozygous variant in *COQ7* remains unsubstantiated and should be further investigated to establish pathogenicity.

Since most of the patients in the cohort used in this study presented not only isolated combined complex deficiency but a combination of enzyme deficiency, it is important to investigate nuclear involvement in all the cases where no pathogenic mtDNA variants were identified during the study conducted by van der Walt (2011). This mtDNA characterisation and a current PhD study underway investigating the other RC nuclear structural gene involvement in this cohort at the centre of human metabonomics will soon provide insight into the matter. The data should be

considered together upon completion of the studies to get a better view of the aetiology of these patients.

Bibliography

1000 GENOMES PROJECT CONSORTIUM. 2012. An integrated map of genetic variation from 1,092 human genomes. *Nature,* 491(7422):56-65.

ADZHUBEI, I.A., SCHMIDT, S., PESHKIN, L., RAMENSKY, V.E., GERASIMOVA, A., BORK, P., KONDRASHOV, A.S., SUNYAEV, S.R. 2010. A method and server for predicting damaging missense mutations. *Nature methods,* 7(4):248-249.

ANDERSON, S., BANKIER, A.T., BARRELL, B.G., DE BRUIJN, M., COULSON, A.R., DROUIN, J., EPERON, I., NIERLICH, D., ROE, B.A., SANGER, F. 1981. Sequence and organization of the human mitochondrial genome. *Nature,* 290(5806):457-465.

AURE, K., BENOIST, J.F., OGIER DE BAULNY, H., ROMERO, N.B., RIGAL, O., LOMBES, A. 2004. Progression despite replacement of a myopathic form of coenzyme Q10 defect. *Neurology,* 63(4):727-729.

BARBER, R.D., HARMER, D.W., COLEMAN, R.A., CLARK, B.J. 2005. GAPDH as a housekeeping gene: Analysis of GAPDH mRNA expression in a panel of 72 human tissues. *Physiological genomics,* 21(3):389-395.

BÉNIT, P., BEUGNOT, R., CHRETIEN, D., GIURGEA, I., LONLAY-DEBENEY, D., ISSARTEL, J., CORRAL-DEBRINSKI, M., KERSCHER, S., RUSTIN, P., RÖTIG, A. 2003. Mutant NDUFV2 subunit of mitochondrial complex I causes early onset hypertrophic cardiomyopathy and encephalopathy. *Human mutation,* 21(6):582-586.

BENTINGER, M., TEKLE, M., DALLNER, G. 2010. Coenzyme Q–biosynthesis and functions. *Biochemical and biophysical research communications,* 396(1):74-79.

BERG, J.M., TYMOCZKO, J.L., STRYER, L. 2002. The respiratory chain consists of four complexes: Three proton pumps and a physical link to the citric acid cycle. Biochemistry. Fifth edition. New York: W.H. Freeman and Company. http://www.ncbi.nlm.nih.gov/books/NBK21154/ Date of access: 16 February 2014.

BHAGAVAN, H.N. & CHOPRA, R.K. 2006. Coenzyme Q10: Absorption, tissue uptake, metabolism and pharmacokinetics. *Free radical research,* 40(5):445-453.

CHI, C. 2014. Diagnostic approach in infants and children with mitochondrial diseases. *Pediatrics & neonatology,* http://dx.doi.org/10.1016/j.pedneo.2014.03.009.

CHIPUK, J., BOUCHIER-HAYES, L., GREEN, D. 2006. Mitochondrial outer membrane permeabilization during apoptosis: The innocent bystander scenario. *Cell death & differentiation,* 13(8):1396-1402.

CHU, H.T., HSIAO, W.W., TSAO, T.T., CHANG, C.M., LIU, Y.W., FAN, C.C., LIN, H., CHANG, H.H., YEH, T.J., CHEN, J.C., HUANG, D.M., CHEN, C.C., KAO, C.Y. 2012. Quantitative assessment of mitochondrial DNA copies from whole genome sequencing. *BMC genomics,* 13(Suppl 7):S5.

CRANE, F., SUN, I., CLARK, M., GREBING, C., LÖW, H. 1985. Transplasma-membrane redox systems in growth and development. *Biochimica et biophysica acta (BBA)-reviews on bioenergetics,* 811(3):233-264.

DEHAHN, T., BARR, R., MORRÉ, D.J. 1997. NADH oxidase activity present on both the external and internal surfaces of soybean plasma membranes. *Biochimica et biophysica acta (BBA)-biomembranes,* 1328(2):99-108.

DEN DUNNEN, J.T. & ANTONARAKIS, S.E. 2000. Mutation nomenclature extensions and suggestions to describe complex mutations: A discussion. *Human mutation,* 15(1):7-12.

DESBATS, M.A., LUNARDI, G., DOIMO, M., TREVISSON, E., SALVIATI, L. 2014. Genetic bases and clinical manifestations of coenzyme Q10 (CoQ10) deficiency. *Journal of inherited metabolic disease,* 1-12.

DIMAURO, S. & SCHON, E.A. 2003. Mitochondrial respiratory-chain diseases. *New england journal of medicine,* 348(26):2656-2668.

DIMAURO, S., QUINZII, C.M., HIRANO, M. 2007. Mutations in coenzyme Q10 biosynthetic genes. *The journal of clinical investigation,* 117(3):587-589.

DÖRING, F., SCHMELZER, C., LINDNER, I., VOCK, C., FUJII, K. 2007. Functional connections and pathways of coenzyme Q10-inducible genes: An in-silico study. *International Union of Biochemistry and Molecular Biology (IUBMB) life,* 59(10):628-633.

DU TOIT, H. 2007. Biochemical Analyses of Deficiencies in the Oxidative Phosphorylation System in Human Muscle. Potchefstroom: NWU. (Dissertation - MSc).

EATON, S.L., ROCHE, S.L., HURTADO, M.L., OLDKNOW, K.J., FARQUHARSON, C., GILLINGWATER, T.H., WISHART, T.M. 2013. Total protein analysis as a reliable loading control for quantitative fluorescent western blotting. *PloS one,* 8(8):e72457.

ECHTAY, K.S., WINKLER, E., FRISCHMUTH, K., KLINGENBERG, M. 2001. Uncoupling proteins 2 and 3 are highly active H(+) transporters and highly nucleotide sensitive when activated by coenzyme Q (ubiquinone). *Proceedings of the national academy of sciences of the united states of america,* 98(4):1416-1421.

EL-NAJJAR, N., GALI-MUHTASIB, H., KETOLA, R.A., VUORELA, P., URTTI, A., VUORELA, H. 2011. The chemical and biological activities of quinones: Overview and implications in analytical detection. *Phytochemistry reviews,* 10(3):353-370.

ELSTNER, M. & TURNBULL, D.M. 2012. Transcriptome analysis in mitochondrial disorders. *Brain research bulletin,* 88(4):285-293.

EMMANUELE, V., LÓPEZ, L.C., BERARDO, A., NAINI, A., TADESSE, S., WEN, B., D'AGOSTINO, E., SOLOMON, M., DIMAURO, S., QUINZII, C. 2012. Heterogeneity of coenzyme Q10 deficiency: Patient study and literature review. *Archives of neurology,* 69(8):978-983.

ENDO, T., YAMANO, K., KAWANO, S. 2011. Structural insight into the mitochondrial protein import system. *Biochimica et biophysica acta (BBA)-biomembranes,* 1808(3):955-970.

ERNSTER, L. & DALLNER, G. 1995. Biochemical, physiological and medical aspects of ubiquinone function. *Biochimica et biophysica acta (BBA)-molecular basis of disease,* 1271(1):195-204.

ERNSTER, L., LEE, I., NORLING, B., PERSSON, B. 1969. Studies with Ubiquinone-Depleted submitochondrial particles. *European journal of biochemistry,* 9(3):299-310.

FILLER, K., LYON, D., BENNETT, J., MCCAIN, N., ELSWICK, R., LUKKAHATAI, N., SALIGAN, L.N. 2014. Association of mitochondrial dysfunction and fatigue: A review of the literature. *BBA clinical,* 1:12-23.

FINKEL, T. & HOLBROOK, N.J. 2000. Oxidants, oxidative stress and the biology of ageing. *Nature,* 408(6809):239-247.

FISCHER, A., SCHMELZER, C., RIMBACH, G., NIKLOWITZ, P., MENKE, T., DORING, F. 2011. Association between genetic variants in the coenzyme Q10 metabolism and coenzyme Q10 status in humans. *BMC research notes,* 4(245):1-7.

FREI, B., ENGLAND, L., AMES, B.N. 1989. Ascorbate is an outstanding antioxidant in human blood plasma. *Proceedings of the national academy of sciences of the united states of america,* 86(16):6377-6381.

GARRET, R.H., GRISHAM, C.M. 2010. Biochemistry. Fourth edition. Boston: Brooks/Cole. p. 592-607.

GEMPEL, K., TOPALOGLU, H., TALIM, B., SCHNEIDERAT, P., SCHOSER, B.G., HANS, V.H., PALMAFY, B., KALE, G., TOKATLI, A., QUINZII, C., HIRANO, M., NAINI, A., DIMAURO, S., PROKISCH, H., LOCHMULLER, H., HORVATH, R. 2007. The myopathic form of coenzyme Q10 deficiency is caused by mutations in the electron-transferring-flavoprotein dehydrogenase (ETFDH) gene. *Brain : A journal of neurology,* 130(Pt 8):2037-2044.

GIRONI, M., LAMPERTI, C., NEMNI, R., MOGGIO, M., COMI, G., GUERINI, F.R., FERRANTE, P., CANAL, N., NAINI, A., BRESOLIN, N., DIMAURO, S. 2004. Late-onset cerebellar ataxia with hypogonadism and muscle coenzyme Q10 deficiency. *Neurology,* 62(5):818-820.

GOMEZ-DIAZ, C., RODRIGUEZ-AGUILERA, J.C., BARROSO, M.P., VILLALBA, J.M., NAVARRO, F., CRANE, F.L., NAVAS, P. 1997. Antioxidant ascorbate is stabilized by NADH-coenzyme Q10 reductase in the plasma membrane. *Journal of bioenergetics and biomembranes,* 29(3):251-257.

GOODEVE, A., REITSMA, P., MCVEY, J. 2011. Nomenclature of genetic variants in hemostasis. *Journal of thrombosis and haemostasis,* 9(4):852-855.

GRIFFITHS, A.J., MILLER, J.H., SUZUKI, D.T., LEWONTIN, R.C., GELBART, W.M. 2000. Gene mutation. An introduction to Genetic Analysis. Seventh edition. New York: W.H. Freeman. http://www.ncbi.nlm.nih.gov/books/NBK21766/ Date of access: 25 February 2014.

HAAS, R.H., PARIKH, S., FALK, M.J., SANETO, R.P., WOLF, N.I., DARIN, N., WONG, L., COHEN, B.H., NAVIAUX, R.K. 2008. The in-depth evaluation of suspected mitochondrial disease. *Molecular genetics and metabolism,* 94(1):16-37.

HAAS, R.H., PARIKH, S., FALK, M.J., SANETO, R.P., WOLF, N.I., DARIN, N., COHEN, B.H. 2007. Mitochondrial disease: A practical approach for primary care physicians. *Pediatrics,* 120(6):1326-1333.

HAMILTON, S.J., CHEW, G.T., WATTS, G.F. 2007. Therapeutic regulation of endothelial dysfunction in type 2 diabetes mellitus. *Diabetes & vascular disease research : Official journal of the international society of diabetes and vascular disease,* 4(2):89-102.

HAUSS, T., DANTE, S., HAINES, T.H., DENCHER, N.A. 2005. Localization of coenzyme Q10 in the center of a deuterated lipid membrane by neutron diffraction. *Biochimica et biophysica acta,* 1710(1):57-62.

HEERINGA, S.F., CHERNIN, G., CHAKI, M., ZHOU, W., SLOAN, A.J., JI, Z., XIE, L.X., SALVIATI, L., HURD, T.W., VEGA-WARNER, V., KILLEN, P.D., RAPHAEL, Y., ASHRAF, S., OVUNC, B., SCHOEB, D.S., MCLAUGHLIN, H.M., AIRIK, R., VLANGOS, C.N., GBADEGESIN, R., HINKES, B., SAISAWAT, P., TREVISSON, E., DOIMO, M., CASARIN, A., PERTEGATO, V., GIORGI, G., PROKISCH, H., ROTIG, A., NURNBERG, G., BECKER, C., WANG, S., OZALTIN, F., TOPALOGLU, R., BAKKALOGLU, A., BAKKALOGLU, S.A., MULLER, D., BEISSERT, A., MIR, S., BERDELI, A., VARPIZEN, S., ZENKER, M., MATEJAS, V., SANTOS-OCANA, C., NAVAS, P., KUSAKABE, T., KISPERT, A., AKMAN, S., SOLIMAN, N.A., KRICK, S., MUNDEL, P., REISER, J., NURNBERG, P., CLARKE, C.F., WIGGINS, R.C., FAUL, C., HILDEBRANDT, F. 2011. COQ6 mutations in human patients produce nephrotic syndrome with sensorineural deafness. *The journal of clinical investigation,* 121(5):2013-2024.

HENZE, K. & MARTIN, W. 2003. Evolutionary biology: Essence of mitochondria. *Nature,* 426(6963):127-128.

HERRMANN, J.M. & NEUPERT, W. 2000. Protein transport into mitochondria. *Current opinion in microbiology,* 3(2):210-214.

HORVATH, R. 2012. Update on clinical aspects and treatment of selected vitamin-responsive disorders II (riboflavin and CoQ10). *Journal of inherited metabolic disease,* 35(4):679-687.

HORVATH, R., CZERMIN, B., GULATI, S., DEMUTH, S., HOUGE, G., PYLE, A., DINEIGER, C., BLAKELY, E.L., HASSANI, A., FOLEY, C., BRODHUN, M., STORM, K., KIRSCHNER, J., GORMAN, G.S., LOCHMULLER, H., HOLINSKI-FEDER, E., TAYLOR, R.W., CHINNERY, P.F. 2012. Adult-onset cerebellar ataxia due to mutations in CABC1/ADCK3. *Journal of neurology, neurosurgery, and psychiatry,* 83(2):174-178.

ITKONEN, O., SUOMALAINEN, A., TURPEINEN, U. 2013. Mitochondrial coenzyme Q10 determination by isotope-dilution liquid chromatography-tandem mass spectrometry. *Clinical chemistry,* 59(8):1260-1267.

JANSSEN, A.J., TRIJBELS, F.J., SENGERS, R.C., SMEITINK, J.A., VAN DEN HEUVEL, L.P., WINTJES, L.T., STOLTENBORG-HOGENKAMP, B.J., RODENBURG, R.J. 2007. Spectrophotometric assay for complex I of the respiratory chain in tissue samples and cultured fibroblasts. *Clinical chemistry,* 53(4):729-734.

JONCK, L. 2013. Analysis of the native structure of mitochondrial complex IV in metallothionein knockout mice. Potchefstroom: NWU. (Mini-dissertation - BSc Hons).

KAMPHANS, T., SABRI, P., ZHU, N., HEINRICH, V., MUNDLOS, S., ROBINSON, P.N., PARKHOMCHUK, D., KRAWITZ, P.M. 2013. Filtering for compound heterozygous sequence variants in non-consanguineous pedigrees. *PloS one,* 8(8):e70151.

KAMPHANS, T. & KRAWITZ, P.M. 2012. GeneTalk: An expert exchange platform for assessing rare sequence variants in personal genomes. *Bioinformatics (oxford, england),* 28(19):2515-2516.

KAPOOR, P. & KAPOOR, A. 2013. Coenzyme Q10–A novel molecule. *Journal, indian academy of clinical medicine,* 14(1):37-45.

KRAUSS, S. 2001. Mitochondria: *Structure and role in respiration. Beth Israel Deaconess Medical Center and Harvard Medical School, USA.* Encyclopedia of Life Sciences. Nature Publishing Group, p.1-6.

KUMAR, P., HENIKOFF, S., NG, P.C. 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature protocols,* 4(7):1073-1081.

LAEMMLI, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* 227(5259):680-685.

LALANI, S.R., VLADUTIU, G.D., PLUNKETT, K., LOTZE, T.E., ADESINA, A.M., SCAGLIA, F. 2005. Isolated mitochondrial myopathy associated with muscle coenzyme Q10 deficiency. *Archives of neurology,* 62(2):317-320.

LAMPERTI, C., NAINI, A., HIRANO, M., DE VIVO, D.C., BERTINI, E., SERVIDEI, S., VALERIANI, M., LYNCH, D., BANWELL, B., BERG, M., DUBROVSKY, T., CHIRIBOGA, C., ANGELINI, C., PEGORARO, E., DIMAURO, S. 2003. Cerebellar ataxia and coenzyme Q10 deficiency. *Neurology,* 60(7):1206-1208.

LANDRUM, M.J., LEE, J.M., RILEY, G.R., JANG, W., RUBINSTEIN, W.S., CHURCH, D.M., MAGLOTT, D.R. 2014. ClinVar: Public archive of relationships among sequence variation and human phenotype. *Nucleic acids research,* 42(Database issue):D980-5.

LAPUENTE-BRUN, E., MORENO-LOSHUERTOS, R., ACÍN-PÉREZ, R., LATORRE-PELLICER, A., COLÁS, C., BALSA, E., PERALES-CLEMENTE, E., QUIRÓS, P.M., CALVO, E., RODRÍGUEZ-HERNÁNDEZ, M. 2013. Supercomplex assembly determines electron flux in the mitochondrial electron transport chain. *Science,* 340(6140):1567-1570.

LENAZ, G., FATO, R., FORMIGGINI, G., GENOVA, M.L. 2007. The role of coenzyme Q in mitochondrial electron transport. *Mitochondrion,* 7:S8-S33.

LESHINSKY-SILVER, E., LEVINE, A., NISSENKORN, A., BARASH, V., PERACH, M., BUZHAKER, E., SHAHMUROV, M., POLAK-CHARCON, S., LEV, D., LERMAN-SAGIE, T. 2003. Neonatal liver failure and leigh syndrome possibly due to CoQ-responsive OXPHOS deficiency. *Molecular genetics and metabolism,* 79(4):288-293.

LÓPEZ, L.C., QUINZII, C.M., AREA, E., NAINI, A., RAHMAN, S., SCHUELKE, M., SALVIATI, L., DIMAURO, S., HIRANO, M. 2010. Treatment of CoQ10 deficient fibroblasts with ubiquinone, CoQ analogs, and vitamin C: Time-and compound-dependent effects. *PLoS one,* 5(7):e11897.

LÓPEZ, L.C., SCHUELKE, M., QUINZII, C.M., KANKI, T., RODENBURG, R.J., NAINI, A., DIMAURO, S. & HIRANO, M. 2006. Leigh syndrome with nephropathy and CoQ10 deficiency due to decaprenyl diphosphate synthase subunit 2 (PDSS2) mutations. *The american journal of human genetics,* 79(6):1125-1129.

MAHMOOD, T. & YANG, P. 2012. Western blot: Technique, theory, and trouble shooting. *North american journal of medical sciences,* 4(9):429.

MANNELLA, C.A. 2006. Structure and dynamics of the mitochondrial inner membrane cristae. *Biochimica et biophysica acta (BBA)-molecular cell research,* 1763(5):542-548.

MARGULIES, M., EGHOLM, M., ALTMAN, W.E., ATTIYA, S., BADER, J.S., BEMBEN, L.A., BERKA, J., BRAVERMAN, M.S., CHEN, Y., CHEN, Z. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature,* 437(7057):376-380.

Marín-García, J. 2013. Methods to study mitochondrial structure and function. Mitochondria and their role in cardiovascular disease. New-York: Springer. http://www.springer.com/medicine/cardiology/book/978-1-4614-4598-2 Date of access: 10 January 2014.

MATHEWS, C.K., HOLDE, K.E., AHERN, K.G. 2002. Electron Transport, Oxidative Phosphorylation, and Oxygen Metabolism. Biochemistry. Third edition. New Jersey: Pearson. p. 523-544.

MCBRIDE, H.M., NEUSPIEL, M., WASIAK, S. 2006. Mitochondria: More than just a powerhouse. *Current biology,* 16(14):R551-R560.

MCLAREN, W., PRITCHARD, B., RIOS, D., CHEN, Y., FLICEK, P., CUNNINGHAM, F. 2010. Deriving the consequences of genomic variants with the ensembl API and SNP effect predictor. *Bioinformatics (oxford, england),* 26(16):2069-2070.

MELDRUM, C., DOYLE, M.A., TOTHILL, R.W. 2011. Next-generation sequencing for cancer diagnostics: A practical perspective. *The clinical biochemist.reviews / australian association of clinical biochemists,* 32(4):177-195.

MERRIMAN, B., TORRENT, I., ROTHBERG, J.M., R&D TEAM. 2012. Progress in ion torrent semiconductor chip based sequencing. *Electrophoresis,* 33(23):3397-3417.

METZKER, M.L. 2005. Emerging technologies in DNA sequencing. *Genome research,* 15(12):1767-1776.

MILES, M.V., MILES, L., TANG, P.H., HORN, P.S., STEELE, P.E., DEGRAUW, A.J., WONG, B.L., BOVE, K.E. 2008. Systematic evaluation of muscle coenzyme Q10 content in children with mitochondrial respiratory chain enzyme deficiencies. *Mitochondrion,* 8(2):170-180.

MILLAT, G., CHANAVAT, V., ROUSSON, R. 2014. Evaluation of a new NGS method based on a custom AmpliSeq library and ion torrent PGM sequencing for the fast detection of genetic variations in cardiomyopathies. *Clinica chimica acta,* 433:266-271.

MITCHELL, P. 1975. The protonmotive Q cycle: A general formulation. *FEBS letters,* 59(2):137-139.

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

MOLLET, J., GIURGEA, I., SCHLEMMER, D., DALLNER, G., CHRETIEN, D., DELAHODDE, A., BACQ, D., DE LONLAY, P., MUNNICH, A., ROTIG, A. 2007. Prenyldiphosphate synthase, subunit 1 (PDSS1) and OH-benzoate polyprenyltransferase (COQ2) mutations in ubiquinone deficiency and oxidative phosphorylation disorders. *The journal of clinical investigation,* 117(3):765-772.

MONTERO, R., SÁNCHEZ-ALCÁZAR, J.A., BRIONES, P., HERNÁNDEZ, Á.R., CORDERO, M.D., TREVISSON, E., SALVIATI, L., PINEDA, M., GARCÍA-CAZORLA, A., NAVAS, P. 2008. Analysis of coenzyme Q10 in muscle and fibroblasts for the diagnosis of CoQ10 deficiency syndromes. *Clinical biochemistry,* 41(9):697-700.

MUSUMECI, O., NAINI, A., SLONIM, A.E., SKAVIN, N., HADJIGEORGIOU, G.L., KRAWIECKI, N., WEISSMAN, B.M., TSAO, C.Y., MENDELL, J.R., SHANSKE, S., DE VIVO, D.C., HIRANO, M., DIMAURO, S. 2001. Familial cerebellar ataxia with muscle coenzyme Q10 deficiency. *Neurology,* 56(7):849-855.

NEUSTADT, J. & PIECZENIK, S.R. 2008. Medication-induced mitochondrial damage and disease. *Molecular nutrition & food research,* 52(7):780-788.

NICHOLLS, D.G. & FERGUSON, S.J. 2002. The chemiosmotic proton circuit. Bioenergenetics 3. London: Academic Press. p. 57-174.

NIEDRINGHAUS, T.P., MILANOVA, D., KERBY, M.B., SNYDER, M.P., BARRON, A.E. 2011. Landscape of next-generation sequencing technologies. *Analytical chemistry,* 83(12):4327-4341.

OGASAHARA, S., ENGEL, A.G., FRENS, D., MACK, D. 1989. Muscle coenzyme Q deficiency in familial mitochondrial encephalomyopathy. *Proceedings of the national academy of sciences of the united states of america,* 86(7):2379-2382.

OGINO, S., GULLEY, M.L., DEN DUNNEN, J.T., WILSON, R.B. 2007. Standard mutation nomenclature in molecular diagnostics: Practical and educational challenges. *The journal of molecular diagnostics,* 9(1):1-6.

OHTAKE, A., MURAYAMA, K., MORI, M., HARASHIMA, H., YAMAZAKI, T., TAMARU, S., YAMASHITA, Y., KISHITA, Y., NAKACHI, Y., KOHDA, M. 2014. Diagnosis and molecular basis of mitochondrial respiratory chain disorders: Exome sequencing for disease gene identification. *Biochimica et biophysica acta (BBA)-general subjects,* 1840(4):1355-1359.

OLSEN, R.K., OLPIN, S.E., ANDRESEN, B.S., MIEDZYBRODZKA, Z.H., POURFARZAM, M., MERINERO, B., FRERMAN, F.E., BERESFORD, M.W., DEAN, J.C., CORNELIUS, N., ANDERSEN, O., OLDFORS, A., HOLME, E., GREGERSEN, N., TURNBULL, D.M., MORRIS, A.A. 2007. ETFDH mutations as a major cause of riboflavin-responsive multiple acyl-CoA dehydrogenation deficiency. *Brain : A journal of neurology,* 130(Pt 8):2045-2054.

PABINGER, S., DANDER, A., FISCHER, M., SNAJDER, R., SPERK, M., EFREMOVA, M., KRABICHLER, B., SPEICHER, M.R., ZSCHOCKE, J., TRAJANOSKI, Z. 2014. A survey of tools

for variant analysis of next-generation genome sequencing data. *Briefings in bioinformatics,* 15(2):256-278.

PAILA, U., CHAPMAN, B.A., KIRCHNER, R., QUINLAN, A.R. 2013. GEMINI: Integrative exploration of genetic variation and genome annotations. *PLoS computational biology,* 9(7):e1003153.

PAPA, S. & SKULACHEV, V. 1997. Reactive oxygen species, mitochondria, apoptosis and aging. *Molecular and Cellular Biochemistry,* 174(1-2):305-319.

PAPUCCI, L., SCHIAVONE, N., WITORT, E., DONNINI, M., LAPUCCI, A., TEMPESTINI, A., FORMIGLI, L., ZECCHI-ORLANDINI, S., ORLANDINI, G., CARELLA, G., BRANCATO, R., CAPACCIOLI, S. 2003. Coenzyme q10 prevents apoptosis by inhibiting mitochondrial depolarization independently of its free radical scavenging property. *The journal of biological chemistry,* 278(30):28220-28228.

POTGIETER, M., PRETORIUS, E., PEPPER, M.S. 2013. Primary and secondary coenzyme Q10 deficiency: The role of therapeutic supplementation. *Nutrition reviews,* 71(3):180-188.

PRELICH, G. 2012. Gene overexpression: Uses, mechanisms, and interpretation. *Genetics,* 190(3):841-854.

QUINZII, C.M., DIMAURO, S., HIRANO, M. 2007. Human coenzyme Q10 deficiency. *Neurochemical research,* 32(4-5):723-727.

QUINZII, C.M. & HIRANO, M. 2011. Primary and secondary CoQ10 deficiencies in humans. *Biofactors,* 37(5):361-365.

QUINZII, C.M. & HIRANO, M. 2010. Coenzyme Q and mitochondrial disease. *Developmental disabilities research reviews,* 16(2):183-188.

QUINZII, C., NAINI, A., SALVIATI, L., TREVISSON, E., NAVAS, P., DIMAURO, S., HIRANO, M. 2006. A mutation in Para-hydroxybenzoate-polyprenyl transferase (COQ2) causes primary coenzyme Q10 deficiency. *The american journal of human genetics,* 78(2):345-349.

QUINZII, C., LÓPEZ, L., NAINI, A., DIMAURO, S., HIRANO, M. 2008. Human CoQ10 deficiencies. *Biofactors,* 32(1):113-118.

RABBANI, B., TEKIN, M., MAHDIEH, N. 2013. The promise of whole-exome sequencing in medical genetics. *Journal of human genetics,* 59(1):5-15.

RAHMAN, S., BLOK, R., DAHL, H., DANKS, D., KIRBY, D., CHOW, C., CHRISTODOULOU, J., THORBURN, D. 1996. Leigh syndrome: Clinical features and biochemical and DNA abnormalities. *Annals of neurology,* 39(3):343-351.

RAZALI, N.M. & WAH, Y.B. 2011. Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of statistical modeling and analytics,* 2(1):21-33.

REINECKE, C.J., KOEKEMOER, G., VAN DER WESTHUIZEN, FRANCOIS H, LOUW, R., LINDEQUE, J.Z., MIENIE, L.J., SMUTS, I. 2012. Metabolomics of urinary organic acids in respiratory chain deficiencies in children. *Metabolomics,* 8(2):264-283.

RODENBURG, R.J. 2011. Biochemical diagnosis of mitochondrial disorders. *Journal of inherited metabolic disease,* 34(2):283-292.

ROSS, J.S. & CRONIN, M. 2011. Whole cancer genome sequencing by next-generation methods. *American journal of clinical pathology,* 136(4):527-539.

ROTHBERG, J.M., HINZ, W., REARICK, T.M., SCHULTZ, J., MILESKI, W., DAVEY, M., LEAMON, J.H., JOHNSON, K., MILGREW, M.J., EDWARDS, M. 2011. An integrated semiconductor device enabling non-optical genome sequencing. *Nature,* 475(7356):348-352.

RÖTIG, A., APPELKVIST, E., GEROMEL, V., CHRETIEN, D., KADHOM, N., EDERY, P., LEBIDEAU, M., DALLNER, G., MUNNICH, A., ERNSTER, L. 2000. Quinone-responsive multiple respiratory-chain dysfunction due to widespread coenzyme Q10 deficiency. *The lancet,* 356(9227):391-395.

SAIKI, R., OGIYAMA, Y., KAINOU, T., NISHI, T., MATSUDA, H., KAWAMUKAI, M. 2003. Pleiotropic phenotypes of fission yeast defective in ubiquinone-10 production. A study from the abc1Sp (coq8Sp) mutant. *Biofactors,* 18(1):229-235.

SALVIATI, L., TREVISSON, E., RODRIGUEZ HERNANDEZ, M.A., CASARIN, A., PERTEGATO, V., DOIMO, M., CASSINA, M., AGOSTO, C., DESBATS, M.A., SARTORI, G., SACCONI, S., MEMO, L., ZUFFARDI, O., ARTUCH, R., QUINZII, C., DIMAURO, S., HIRANO, M., SANTOS-OCANA, C., NAVAS, P. 2012. Haploinsufficiency of COQ4 causes coenzyme Q10 deficiency. *Journal of medical genetics,* 49(3):187-191.

SANGER, F., NICKLEN, S. & COULSON, A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the national academy of sciences of the united states of america,* 74(12):5463-5467.

SCARPULLA, R.C. 1997. Nuclear control of respiratory chain expression in mammalian cells. *Journal of bioenergetics and biomembranes,* 29(2):109-119.

SCHÄGGER, H. 2002. Respiratory chain supercomplexes of mitochondria and bacteria. *Biochimica et biophysica acta (BBA)-bioenergetics,* 1555(1):154-159.

SCHÄGGER, H. & PFEIFFER, K. 2000. Supercomplexes in the respiratory chains of yeast and mammalian mitochondria. *The European Molecular Biology Organization (EMBO) journal,* 19(8):1777-1783.

SHASTRY, B. 1995. Overexpression of genes in health and sickness. A bird's eye view. *Comparative biochemistry and physiology part B: Biochemistry and molecular biology,* 112(1):1-13.

SHEATHER, S.J. & MARRON, J.S. 1990. Kernel quantile estimators. *Journal of the american statistical association,* 85(410):410-416.

SHENDURE, J., PORRECA, G.J., REPPAS, N.B., LIN, X., MCCUTCHEON, J.P., ROSENBAUM, A.M., WANG, M.D., ZHANG, K., MITRA, R.D., CHURCH, G.M. 2005. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science,* 309(5741):1728-1732.

SHEPHERD, D. & GARLAND, P. 1969. The kinetic properties of citrate synthase from rat liver mitochondria. *Biochemical Journal,* 114(3):597-610.

SHERRY, S.T., WARD, M.H., KHOLODOV, M., BAKER, J., PHAN, L., SMIGIELSKI, E.M., SIROTKIN, K. 2001. dbSNP: The NCBI database of genetic variation. *Nucleic acids research,* 29(1):308-311.

SMITH, P., KROHN, R.I., HERMANSON, G., MALLIA, A., GARTNER, F., PROVENZANO, M., FUJIMOTO, E., GOEKE, N., OLSON, B., KLENK, D. 1985. Measurement of protein using bicinchoninic acid. *Analytical biochemistry,* 150(1):76-85.

SMUTS, I., LOUW, R., DU TOIT, H., KLOPPER, B., MIENIE, L.J., VAN DER WESTHUIZEN, FRANCOIS H. 2010. An overview of a cohort of south african patients with mitochondrial disorders. *Journal of inherited metabolic disease,* 33(3):95-104.

SMUTS, I., VAN DER WESTHUIZEN, F.H, LOUW, R., MIENIE, L.J., ENGELKE, U.F., WEVERS, R.A., MASON, S., KOEKEMOER, G., REINECKE, C.J. 2013. Disclosure of a putative biosignature for respiratory chain disorders through a metabolomics approach. *Metabolomics,* 9(2):379-391.

SMUTS. I. & VAN DER WESTHUIZEN. F.H. 2010 Mitochondrial disorders diagnostic approaches and their application in the South African context. *South African Paediatric Reviews* (*SAPR)*, 7:6-15.

SPINAZZOLA, A. & ZEVIANI, M. 2009. Disorders from perturbations of nuclear-mitochondrial intergenomic cross-talk. *Journal of internal medicine,* 265(2):174-192.

SUE, C.M. & SCHON, E.A. 2000. Mitochondrial respiratory chain diseases and mutations in nuclear DNA: A promising start? *Brain pathology,* 10(3):442-450.

TAANMAN, J. 1999. The mitochondrial genome: Structure, transcription, translation and replication. *Biochimica et biophysica acta (BBA)-bioenergetics,* 1410(2):103-123.

THORBURN, D. 2004. Mitochondrial disorders: Prevalence, myths and advances. *Journal of inherited metabolic disease,* 27(3):349-362.

TRAN, U.C. & CLARKE, C.F. 2007. Endogenous synthesis of coenzyme Q in eukaryotes. *Mitochondrion,* 7(Suppl):S62-S71.

TREVISSON, E., DIMAURO, S., NAVAS, P., SALVIATI, L. 2011. Coenzyme Q deficiency in muscle. *Current opinion in neurology,* 24(5):449-456.

TURUNEN, M., OLSSON, J., DALLNER, G. 2004. Metabolism and function of coenzyme Q. *Biochimica et biophysica acta (BBA)-biomembranes,* 1660(1):171-199.

TURUNEN, M., WEHLIN, L., SJöBERG, M., LUNDAHL, J., DALLNER, G., BRISMAR, K., SINDELAR, P. 2002. b2-Integrin and lipid modifications indicate a non-antioxidant mechanism for the anti-atherogenic effect of dietary coenzyme Q10. *Biochemical and Biophysical Research Communications*, 296:255–260

VAN DER WALT, E.M. 2011. Molecular genetic characterization of mitochondrial DNA in a cohort of South African patients with mitochondrial disorders. Potchefstroom: NWU. *(*Thesis - PhD).

VAN MALDERGEM, L., TRIJBELS, F., DIMAURO, S., SINDELAR, P.J., MUSUMECI, O., JANSSEN, A., DELBERGHE, X., MARTIN, J., GILLEROT, Y. 2002. Coenzyme Q–responsive leigh's encephalopathy in two sisters. *Annals of neurology,* 52(6):750-754.

VENTER, L., LINDEQUE, Z., VAN RENSBURG, P.J., VAN DER WESTHUIZEN, F., SMUTS, I., LOUW, R. Untargeted urine metabolomics reveals a biosignature for muscle respiratory chain deficiencies. *Metabolomics,* 1-11.

VOELKERDING, K.V., DAMES, S.A., DURTSCHI, J.D. 2009. Next-generation sequencing: From basic research to diagnostics. *Clinical chemistry,* 55(4):641-658.

WALTER, L., NOGUEIRA, V., LEVERVE, X., HEITZ, M.P., BERNARDI, P., FONTAINE, E. 2000. Three classes of ubiquinone analogs regulate the mitochondrial permeability transition pore through a common site. *The journal of biological chemistry,* 275(38):29521-29527.

WANG, L.L., LI, Y., ZHOU, S.F. 2009. A bioinformatics approach for the phenotype prediction of nonsynonymous single nucleotide polymorphisms in human cytochromes P450. *Drug metabolism and disposition: The biological fate of chemicals,* 37(5):977-991.

WEI, P., LIU, X., FU, Y. 2011. Incorporating predicted functions of nonsynonymous variants into gene-based analysis of exome sequencing data: A comparative study. *BMC proceedings,* 5(Suppl 9):S20.

WILSENACH, K. 2014. Quantification of coenzymeQ_{10} in South African paediatric patients with electron transport chain deficiencies. Potchefstroom: NWU. (Dissertation - MSc). (Submitted for examination)

WOLF, D.E., HOFFMAN, C.H., TRENNER, N.R., ARISON, B.H., SHUNK, C.H., LINN, B.O., MCPHERSON, J.F., FOLKERS, K. 1958. Coenzyme QI structure studies on the coenzyme Q group. *Journal of the american chemical society,* 80(17):4752-4752.

WONG, L.C. 2013a. Mitochondrial disorders caused by nuclear genes. Biochemical and molecular methods for the study of mitochondrial disorders. New York: Springer. http://www.springer.com/978-1-4614-3721-5 Date of access: 22 February 2014.

WONG, L.C. 2013b. Next generation molecular diagnosis of mitochondrial disorders. *Mitochondrion,* 13(4):379-387.

YOSHIMURA, H., IWASAKI, S., NISHIO, S., KUMAKAWA, K., TONO, T., KOBAYASHI, Y., SATO, H., NAGAI, K., ISHIKAWA, K., IKEZONO, T. 2014. Massively parallel DNA sequencing facilitates diagnosis of patients with usher syndrome type 1. *PloS one,* 9(3):e90688.

ZENG, S., YANG, J., CHUNG, B.H., LAU, Y.L., YANG, W. 2014. EFIN: Predicting the functional impact of nonsynonymous single nucleotide polymorphisms in human genome. *BMC genomics,* 15(1):455.

ZHANG, K., LIN, J.W., WU, X., GAO, H., HSIEH, Y.C., HWU, P., SU, L., CHIOU, H.Y., WANG, D., YUAN, Y.C., WHANG-PENG, J., CHIU, W.T., YEN, Y. 2014. A germline missense mutation in COQ6 is associated with susceptibility to familian schwannomatosis. *Genetics in Medicine,* 16(10):787-792.

ZHANG, W., CUI, H., WONG, L.C. 2014. Application of next generation sequencing to molecular diagnosis of inherited diseases. Chemical diagnostics. New York: Springer. http://www.springer.com/978-3-642-39941-1 Date of access: 4 June 2014.

ZSURKA, G. & KUNZ, W.S. 2013. Mitochondrial involvement in neurodegenerative diseases. *International Union of Biochemistry and Molecular Biology (IUBMB) life,* 65(3):263-272.

Appendix A

Table A-1: Ion AmpliSeq Custom Panel design IAD53496_133 details using the AmpliSeq online web designer (v.3.0.1).

Gene Name	Chr	Chr Start	Chr End	Nr Amplicons	Total Bases	Covered Bases	Missed Bases	Overall Coverage	Nr Exons	100 %	>95 %	>90 %	>80 %	0 %
ADCK3	chr1	227127927	227175256	29	3224	3162	62	98.08	15	14	0	0	0	0
APOE	chr19	45409028	45412660	10	1260	1260	0	100	4	4	0	0	0	0
APTX	chr9	32972593	33001649	17	4081	4081	0	100	14	14	0	0	0	0
COQ10A	chr12	56660631	56664760	13	1925	1909	16	99.17	6	5	1	0	0	0
COQ10B	chr2	198318220	198339861	13	2074	1997	77	96.29	5	4	0	1	0	0
COQ2	chr4	84184966	84206077	16	1781	1781	0	100	7	7	0	0	0	0
COQ3	chr6	99817337	99842092	11	1405	1405	0	100	7	7	0	0	0	0
COQ4	chr9	131084780	131096361	13	1712	1712	0	100	7	7	0	0	0	0
COQ5	chr12	120941071	120966974	11	1649	1622	27	98.36	7	5	2	0	0	0
COQ6	chr14	74416626	74429823	15	2106	1968	138	93.45	13	11	1	0	0	1
COQ7	chr16	19078906	19091427	21	2893	2866	27	99.07	7	6	1	0	0	0
COQ9	chr16	57481326	57495197	14	1869	1869	0	100	9	9	0	0	0	0
ETFA	chr15	76508618	76603820	16	1594	1593	1	99.94	12	12	0	0	0	0
ETFB	chr19	51848398	51869682	14	1749	1749	0	100	7	7	0	0	0	0
ETFDH	chr4	159593266	159629851	23	2608	2552	56	97.85	13	11	0	2	0	0
BRAF	chr7	140433802	140624574	29	3306	3258	48	98.55	18	16	1	1	0	0
PDSS1	chr10	26986584	27035736	15	1858	1858	0	100	12	12	0	0	0	0
PDSS2	chr6	107473750	107780789	24	3712	3416	296	92.03	8	6	1	0	1	0

Table A-1 includes the custom panel design details (Section 3.6) including the gene on the specified chromosome, start and en position of the chromosome, number of amplicons designed to cover the exon and untranslated regions of the genes, the total number of basis- covered and -missed, the overall predicted coverage of the targeted gene regions, number of exons included in the covered regions, and the percentage of exons covered (0 – 100%). Chr: chromosome; Nr: number.

Appendix B

Table B-1: Measured data points of citrate synthase (CS) enzyme activity analysis.

P. ID	[prot] (µg/µl)	A1	A2	A3	AVE	SD	CoV (%)	Citrate synthase activity µmol/min/mg (CS)
P2	2.5	10.0	8.5	7.0	8.5	1.5	17.6	0.0
P14	2.9	36.0	31.0	31.5	32.8	2.8	8.4	0.1
P15 (C)	4.3	49.5	67.0	61.5	59.3	8.9	15.1	0.1
P16 (C)	3.7	85.5	82.5	77.0	81.7	4.3	5.3	0.2
P19 (C)	2.8	47.5	38.0	37.5	41.0	5.6	13.7	0.1
P21	3.9	70.5	58.0	58.0	62.2	7.2	11.6	0.1
P22 (C)	4.0	73.5	66.5	63.0	67.7	5.3	7.9	0.2
Ref	3.0	70.5	60.5	57.5	62.8	6.8	10.8	0.2
P24	3.6	20.5	23.5	25.0	23.0	2.3	10.0	0.1
P29	2.8	41.5	39.5	44.5	41.8	2.5	6.0	0.1
P32	1.7	38.5	39.5	40.0	39.3	0.8	1.9	0.2
P34	3.1	55.0	54.5	58.5	56.0	2.2	3.9	0.2
P36	4.5	77.5	73.5	62.0	71.0	8.0	11.3	0.1
P37	3.6	30.5	30.0	41.0	33.8	6.2	18.4	0.1
P39	3.0	55.0	56.0	56.5	55.8	0.8	1.4	0.2
Ref	3.0	59.0	59.0	56.0	58.0	1.7	3.0	0.2
P40 (C)	2.9	51.0	48.5	47.5	49.0	1.8	3.7	0.2
P42 (C)	3.6	63.0	76.5	67.0	68.8	6.9	10.1	0.2
P43	3.5	43.0	45.5	37.5	42.0	4.1	9.7	0.1
P44 (C)	3.2	40.5	47.0	37.5	41.7	4.9	11.7	0.1
P52 (C)	4.7	87.5	87.5	88.0	87.7	0.3	0.3	0.2
P54	2.5	62.5	57.0	46.5	55.3	8.1	14.7	0.2
P61 (C)	3.1	57.5	52.0	54.5	54.7	2.8	5.0	0.2

P. ID	[prot] (µg/µl)	A1	A2	A3	AVE	SD	CoV (%)	Citrate synthase activity µmol/min/mg (CS)
Ref	3.0	58.5	60.0	56.0	58.2	2.0	3.5	0.2
P64	3.3	67.5	70.5	73.0	70.3	2.8	3.9	0.2
P68	3.6	51.0	53.5	52.5	52.3	1.3	2.4	0.1
P72	2.6	13.0	4.5	3.5	7.0	5.2	74.6	0.0
P78	2.9	89.0	98.5	94.0	93.8	4.8	5.1	0.3
P85	3.5	39.0	36.5	40.0	38.5	1.8	4.7	0.1
P86	3.3	63.5	67.0	65.5	65.3	1.8	2.7	0.2
P87	3.4	46.5	46.0	47.5	46.7	0.8	1.6	0.1
Ref	3.0	57.0	57.5	59.5	58.0	1.3	2.3	0.2
P88	3.4	58.5	61.5	52.0	57.3	4.9	8.5	0.2
P89 (C)	4.2	36.5	39.0	35.5	37.0	1.8	4.9	0.1
P94	2.6	39.0	45.0	42.0	42.0	3.0	7.1	0.1
P95	3.3	43.0	42.0	45.5	43.5	1.8	4.1	0.1
P98	2.1	52.0	50.0	47.0	49.7	2.5	5.1	0.2
P99	2.6	64.0	58.5	53.5	58.7	5.3	9.0	0.2
P101	2.8	52.0	53.0	52.0	52.3	0.6	1.1	0.2
Ref	3.0	62.0	55.0	60.5	59.2	3.7	6.2	0.2
P103 (C)	4.6	69.5	60.5	64.0	64.7	4.5	7.0	0.1
P107 (C)	4.2	31.5	26.0	23.5	27.0	4.1	15.2	0.1
P110 (C)	3.8	58.0	61.0	67.0	62.0	4.6	7.4	0.1
P111 (C)	3.4	70.0	69.0	70.5	69.8	0.8	1.1	0.2
P117 (C)	3.4	46.0	54.5	56.0	52.2	5.4	10.3	0.1
P136 (C)	3.7	59.5	64.5	64.0	62.7	2.8	4.4	0.2
P152 (C)	2.7	53.5	52.5	52.0	52.7	0.8	1.5	0.2
Ref	3.0	55.0	62.0	54.5	57.2	4.2	7.3	0.2
P162 (C)	4.5	49.5	44.5	44.0	46.0	3.0	6.6	0.1
P166 (C)	4.1	46.5	47.5	52.0	48.7	2.9	6.0	0.1

P. ID	[prot] (µg/µl)	A1	A2	A3	AVE	SD	CoV (%)	Citrate synthase activity µmol/min/mg (CS)
P168 (C)	3.3	39.5	37.5	47.0	41.3	5.0	12.1	0.1
Ref	3.0	59.5	64.0	65.5	63.0	3.1	5.0	0.2

Table B-2: Measured data points of CII enzyme activity analysis.

P. ID	A1	A2	A3	AVE	SD	CoV (%)	CII activity µmol/min/mg	CII activity nmol/min/CS
P2	1.0	1.2	0.8	1.0	0.2	20.0	0.00	20.7
P14	8.3	7.7	7.8	7.9	0.3	4.1	0.00	42.6
P15 (C)	34.2	31.6	30.8	32.2	1.8	5.5	0.01	95.6
P16 (C)	24.6	22.5	24.0	23.7	1.1	4.6	0.01	51.1
P19 (C)	10.5	9.6	10.1	10.1	0.5	4.5	0.01	43.3
P21	18.2	18.3	18.9	18.5	0.4	2.1	0.01	52.3
P22 (C)	23.6	23.1	25.5	24.1	1.3	5.3	0.01	62.7
Ref	30.6	31.0	31.3	31.0	0.4	1.1	0.02	86.8
P24	6.0	5.6	5.4	5.7	0.3	5.4	0.00	43.4
P29	9.9	10.4	9.4	9.9	0.5	5.1	0.01	41.7
P32	15.9	15.9	14.5	15.4	0.8	5.2	0.01	69.1
P34	17.8	19.3	16.7	17.9	1.3	7.3	0.01	56.4
P36	36.5	38.3	32.4	35.7	3.0	8.5	0.01	88.7
P37	11.4	12.2	10.4	11.3	0.9	8.0	0.00	59.0
P39	22.3	22.0	19.0	21.1	1.8	8.6	0.01	66.6
Ref	29.7	32.3	29.6	30.5	1.5	5.0	0.02	92.7
P40 (C)	23.5	18.3	18.0	19.9	3.1	15.5	0.01	71.7
P42 (C)	52.8	38.3	38.3	43.1	8.4	19.4	0.02	110.4
P43	21.3	16.5	15.8	17.9	3.0	16.8	0.01	74.9
P44 (C)	20.8	16.8	17.5	18.4	2.1	11.6	0.01	77.7
P52 (C)	85.5	58.2	59.8	67.8	15.3	22.6	0.02	136.3

P. ID	A1	A2	A3	AVE	SD	CoV (%)	CII activity µmol/min/mg	CII activity nmol/min/CS
P54	26.0	21.7	21.9	23.2	2.4	10.5	0.01	73.9
P61 (C)	23.4	22.8	20.8	22.3	1.4	6.1	0.01	72.0
Ref	35.5	34.5	33.5	34.5	1.0	2.9	0.02	104.5
P64	18.6	43.7	17.0	26.4	15.0	56.7	0.01	66.2
P68	16.3	38.6	15.1	23.3	13.2	56.7	0.01	78.5
P72	1.5	1.8	1.7	1.7	0.2	9.2	0.00	41.9
P78	57.8	62.9	55.9	58.9	3.6	6.1	0.03	110.5
P85	16.5	20.2	18.8	18.5	1.9	10.1	0.01	84.7
P86	38.8	42.8	41.6	41.1	2.1	5.0	0.02	110.7
P87	16.8	17.7	18.3	17.6	0.8	4.3	0.01	66.4
Ref	31.8	33.9	33.1	32.9	1.1	3.2	0.02	100.0
P88	24.1	23.5	22.3	23.3	0.9	3.9	0.01	71.6
P89 (C)	25.2	24.4	23.9	24.5	0.7	2.7	0.01	116.7
P94	20.5	23.4	21.8	21.9	1.5	6.6	0.01	91.9
P95	22.3	26.0	24.4	24.2	1.9	7.7	0.01	98.1
P98	16.4	18.6	16.4	17.1	1.3	7.4	0.01	60.8
P99	27.7	32.1	35.7	31.8	4.0	12.6	0.02	95.6
P101	18.0	21.9	22.6	20.8	2.5	11.9	0.01	70.1
Ref	30.2	34.2	32.7	32.4	2.0	6.2	0.02	96.4
P103 (C)	16.2	16.4	17.4	16.7	0.6	3.9	0.01	45.4
P107 (C)	10.9	10.3	11.2	10.8	0.5	4.2	0.00	70.5
P110 (C)	31.6	29.0	31.1	30.6	1.4	4.5	0.01	86.9
P111 (C)	28.5	20.6	34.8	28.0	7.1	25.4	0.01	70.6
P117 (C)	21.0	20.7	23.6	21.8	1.6	7.3	0.01	73.5
P136 (C)	55.2	56.7	59.2	57.0	2.0	3.5	0.02	160.3
P152 (C)	17.3	20.1	17.2	18.2	1.6	9.0	0.01	60.9
Ref	31.0	31.4	32.4	31.6	0.7	2.3	0.02	97.4

P. ID	A1	A2	A3	AVE	SD	CoV (%)	CII activity µmol/min/mg	CII activity nmol/min/CS
P162 (C)	30.1	27.3	28.2	28.5	1.4	5.0	0.01	109.3
P166 (C)	28.8	24.8	25.5	26.4	2.1	8.1	0.01	95.4
P168 (C)	38.6	34.9	33.6	35.7	2.6	7.3	0.02	152.2
Ref	39.2	35.0	33.8	36.0	2.8	7.9	0.02	100.7

Table B-3: Measured data points of CII + III enzyme activity analysis.

P. ID	A1	A2	A3	AVE	SD	CoV (%)	CII + III activity µmol/min/mg	CII + III activity nmol/min/CS
P2	1.3	-0.8	-0.1	0.1	1.1	802.0	0.00	14.0
P14	3.9	4.7	4.5	4.4	0.4	9.5	0.01	118.8
P15 (C)	20.6	22.7	18.6	20.6	2.1	9.9	0.04	310.5
P16 (C)	24.7	22.9	20.4	22.7	2.2	9.5	0.05	247.8
P19 (C)	6.6	6.0	6.5	6.4	0.3	5.0	0.02	138.7
P21	15.1	13.5		14.3	1.1	7.9	0.03	205.4
P22 (C)	11.8	13.2		12.5	1.0	7.9	0.02	165.0
Ref	18.3	15.4		16.9	2.1	12.2	0.04	239.5
P24	1.8	2.9	3.3	2.7	0.8	29.1	0.01	103.5
P29	3.4	5.5	3.7	4.2	1.1	27.0	0.01	89.7
P32	5.3	5.6	7.3	6.1	1.1	17.8	0.03	137.7
P34	8.5	11.1	8.7	9.4	1.4	15.3	0.02	150.4
P36	15.5	11.6	12.5	13.2	2.0	15.5	0.02	166.0
P37	4.6	1.4	3.9	3.3	1.7	51.0	0.01	87.1
P39	7.0	9.7	8.1	8.3	1.4	16.4	0.02	132.2
Ref	16.6	14.4		15.5	1.6	10.0	0.04	238.6
P40 (C)	7.8	8.8	8.4	8.3	0.5	6.0	0.02	151.9
P42 (C)	17.9	18.8	15.2	17.3	1.9	10.8	0.04	224.4
P43	5.3	4.0	3.8	4.4	0.8	18.7	0.01	92.8

P. ID	A1	A2	A3	AVE	SD	CoV (%)	CII + III activity µmol/min/mg	CII + III activity nmol/min/CS
P44 (C)	8.0	7.5	5.8	7.1	1.2	16.2	0.02	152.2
P52 (C)	24.3	21.0	23.0	22.8	1.7	7.3	0.04	231.9
P54	9.0	8.9	6.8	8.2	1.2	15.1	0.03	132.9
P61 (C)	19.5	16.3	17.8	17.9	1.6	9.0	0.05	291.8
Ref	15.0	13.3	16.1	14.8	1.4	9.5	0.04	227.2
P64	7.1	9.0	7.7	7.9	1.0	12.2	0.02	100.7
P68	7.1	6.3	7.6	7.0	0.7	9.4	0.02	119.4
P72	-0.7	0.4	0.2	0.0	0.6	-1757.8	0.00	-4.3
P78	14.3	15.4	13.9	14.5	0.8	5.3	0.04	138.3
P85	10.5	10.6	10.1	10.4	0.3	2.5	0.02	241.2
P86	18.3	17.7	21.3	19.1	1.9	10.1	0.05	261.0
P87	6.4	6.9	7.1	6.8	0.4	5.3	0.02	130.1
Ref	13.3	17.2	13.4	14.6	2.2	15.2	0.04	225.3
P88	9.2	8.5	8.4	8.7	0.4	5.0	0.02	135.5
P89 (C)	10.7	11.6	11.6	11.3	0.5	4.6	0.02	272.7
P94	8.2	7.9	6.0	7.4	1.2	16.2	0.02	156.6
P95	7.0	7.9	6.2	7.0	0.9	12.1	0.02	144.4
P98	5.4	4.5	2.3	4.1	1.6	39.2	0.02	73.1
P99	9.3	6.8	7.5	7.9	1.3	16.4	0.02	119.7
P101	5.3	5.2	4.7	5.1	0.3	6.3	0.01	86.5
Ref	14.6	13.3	13.7	13.9	0.7	4.8	0.04	209.3
P103 (C)	11.2	10.8	11.0	11.0	0.2	1.8	0.02	151.9
P107 (C)	3.1	3.2	2.8	3.0	0.2	6.9	0.01	100.3
P110 (C)	18.0	18.4	22.5	19.6	2.5	12.7	0.04	282.8
P111 (C)	11.0	9.5	10.3	10.3	0.8	7.3	0.02	131.3
P117 (C)	6.9	5.5	6.5	6.3	0.7	11.4	0.01	107.8
P136 (C)	14.8	26.0	28.5	23.1	7.3	31.6	0.05	329.2

P. ID	A1	A2	A3	AVE	SD	CoV (%)	CII + III activity µmol/min/mg	CII + III activity nmol/min/CS
P152 (C)	6.1	5.1	6.4	5.9	0.7	11.6	0.02	99.5
Ref	15.2	12.7	17.1	15.0	2.2	14.7	0.04	234.3
P162 (C)	19.2	20.3	20.6	20.0	0.7	3.7	0.04	388.9
P166 (C)	14.7	13.8	14.9	14.5	0.6	4.1	0.03	265.4
P168 (C)	19.1	18.1	19.1	18.8	0.6	3.1	0.05	405.4
Ref	23.3	20.2	20.4	21.3	1.7	8.1	0.06	301.9

In Tables B-1, B-2 and B-3 the triplicate measured data points (Section 3.3) are given for the spectrophotometric enzyme activity analyses. CII is normalised to CS, and CII + III is normalised to citrate synthase (CS) and CII. The blank spaces (spaces without values) indicate were not enough sample was available to complete the analysis in triplicate. P. ID: patient identification; P: patient; C: control; Ref: reference sample; A1: first measurement; A2: second measurement; A3: third measurement; AVE: average of triplicate measurements; SD: standard deviation of triplicate measurements; CoV: co-efficient of variance.

Appendix C

Table C-1: Descriptive details of the possible disease-causing variants identified in patient cohort.

P. ID	CHR	START	END	GENE	QUAL	COD	CODON CHG	A.A CHG	IMPACT	IMPACT SEV	LOF	ALLELE CALL	FREQ	ORG COV
P32	chr19	45411940	45411941	*APOE*	26.7	1	Tgc/Cgc	Cys/Arg	non_syn_coding	MED	0	Heterozygous	33.3	21
	chr12	56662936	56662937	*COQ10A*	361.71	1	Cgt/Tgt	Arg/Cys	non_syn_coding	MED	0	Heterozygous	65.6	95
	chr19	51850289	51850290	*ETFB*	260.53	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	46.2	143
P86	chr16	19085297	19085298	*COQ7*	1627.6	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Homozygous	100	167
	chr19	51850289	51850290	*ETFB*	482.4	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	50.7	218
P98	chr14	74416835	74416836	*COQ6*	51.83	1	tGg/tAg	Trp/*	stop_gain	HIGH	1	Heterozygous	50	20
	chr16	19085297	19085298	*COQ7*	67.35	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	41.5	41
	chr19	51850289	51850290	*ETFB*	170.44	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	50.7	75
P99	chr12	56663291	56663292	*COQ10A*	18.2	1	Cga/Tga	Arg/*	stop_gain	HIGH	1	Heterozygous	19.7	61
	chr19	45411940	45411941	*APOE*	22.75	1	Tgc/Cgc	Cys/Arg	non_syn_coding	MED	0	Heterozygous	50	8
	chr15	76578761	76578762	*ETFA*	130.46	1	aCa/aTa	Thr/Ile	non_syn_coding	MED	0	Heterozygous	39	100
	chr16	19085297	19085298	*COQ7*	723.84	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Homozygous	100	76
P54	chr16	19085297	19085298	*COQ7*	406.08	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	49.7	191
	chr19	51850289	51850290	*ETFB*	705.74	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	52.1	304
P85	chr19	51850289	51850290	*ETFB*	735.63	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	51.5	325
P2	chr14	74416835	74416836	*COQ6*	652.98	1	tGg/tAg	Trp/*	stop_gain	HIGH	1	Heterozygous	50	306
	chr19	45411940	45411941	*APOE*	27.69	1	Tgc/Cgc	Cys/Arg	non_syn_coding	MED	0	Heterozygous	40	15
	chr14	74417179	74417180	*COQ6*	821.74	1	Gcc/Tcc	Ala/Ser	non_syn_coding	MED	0	Heterozygous	52.3	351
	chr19	51850289	51850290	*ETFB*	3825.3	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Homozygous	100	524
P29	chr14	74420251	74420252	*COQ6*	57.44	1	gGc/gAc	Gly/Asp	non_syn_coding	MED	0	Heterozygous	39.5	38
	chr16	57486773	57486774	*COQ9*	322.65	1	Cgc/Tgc	Arg/Cys	non_syn_coding	MED	0	Heterozygous	48.7	158

P. ID	CHR	START	END	GENE	QUAL	COD	CODON CHG	A.A CHG	IMPACT	IMPACT SEV	LOF	ALLELE CALL	FREQ	ORG COV
	chr19	51850289	51850290	*ETFB*	3675.9	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Homozygous	100	388
P94	chr16	19085297	19085298	*COQ7*	89.69	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	41.4	57
	chr19	51850289	51850290	*ETFB*	421.02	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	54.6	163
P88	chr16	19085297	19085298	*COQ7*	380.86	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	60.7	116
	chr19	51850289	51850290	*ETFB*	942.49	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	52.6	445
P21	chr19	51850289	51850290	*ETFB*	706.21	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	47.1	379
P24	chr14	74416835	74416836	*COQ6*	580.98	1	tGg/tAg	Trp/*	stop_gain	HIGH	1	Heterozygous	52.9	241
	chr19	45411940	45411941	*APOE*	24.78	1	Tgc/Cgc	Cys/Arg	non_syn_coding	MED	0	Heterozygous	35.3	17
	chr9	32984802	32984803	*APTX*	612.24	1	cGt/cAt	Arg/His	non_syn_coding	MED	0	Heterozygous	46.9	324
	chr16	19085297	19085298	*COQ7*	449.9	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	48.7	222
P14	chr14	74416831	74416832	*COQ6*	541.63	1	Gtc/Ttc	Val/Phe	non_syn_coding	MED	0	Heterozygous	46.9	290
	chr16	19085297	19085298	*COQ7*	545.1	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	49.4	262
	chr19	51850289	51850290	*ETFB*	836.67	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	49.8	648
P95	chr16	19085297	19085298	*COQ7*	518.27	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	51.1	233
	chr19	51850289	51850290	*ETFB*	1006.3	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	54.9	390
P101	chr14	74416835	74416836	*COQ6*	216.96	1	tGg/tAg	Trp/*	stop_gain	HIGH	1	Heterozygous	40.1	162
	chr9	32984802	32984803	*APTX*	660.75	1	cGt/cAt	Arg/His	non_syn_coding	MED	0	Heterozygous	47.7	341
	chr16	19085297	19085298	*COQ7*	1757.4	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Homozygous	97.9	194
	chr19	51850289	51850290	*ETFB*	749.94	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	52.2	324
P34	chr19	45411878	45411879	*APOE*	19.11	1	gCa/gTa	Ala/Val	non_syn_coding	MED	0	Heterozygous	25	28
P43	chr14	74416831	74416832	*COQ6*	536.81	1	Gtc/Ttc	Val/Phe	non_syn_coding	MED	0	Heterozygous	59.4	170
	chr16	19085297	19085298	*COQ7*	420.78	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	50.5	189
	chr16	19085360	19085361	*COQ7*	460.95	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	52	197
P68	chr14	74416835	74416836	*COQ6*	363.55	1	tGg/tAg	Trp/*	stop_gain	HIGH	1	Heterozygous	47.8	184
	chr19	51850289	51850290	*ETFB*	633.02	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	47.7	328

P. ID	CHR	START	END	GENE	QUAL	COD	CODON CHG	A.A CHG	IMPACT	IMPACT SEV	LOF	ALLELE CALL	FREQ	ORG COV
P37	chr14	74416831	74416832	*COQ6*	550.08	1	Gtc/Ttc	Val/Phe	non_syn_coding	MED	0	Heterozygous	51.7	235
	chr16	19085297	19085298	*COQ7*	390.51	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	45.2	226
	chr19	51850289	51850290	*ETFB*	1065.2	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	56.2	391
P36	chr19	51850289	51850290	*ETFB*	596.73	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	48.7	297
P87	chr19	45411940	45411941	*APOE*	39.46	1	Tgc/Cgc	Cys/Arg	non_syn_coding	MED	0	Heterozygous	60	10
	chr19	51850289	51850290	*ETFB*	744.21	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	47.7	389
P39	chr14	74416835	74416836	*COQ6*	357.86	1	tGg/tAg	Trp/*	stop_gain	HIGH	1	Heterozygous	45.6	202
	chr19	45411940	45411941	*APOE*	51.06	1	Tgc/Cgc	Cys/Arg	non_syn_coding	MED	0	Heterozygous	70	10
	chr16	19085297	19085298	*COQ7*	449.84	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	45.2	260
P78	chr4	159620232	159620233	*ETFDH*	107.4	1	gGa/gAa	Gly/Glu	non_syn_coding	MED	0	Heterozygous	41.4	70
	chr4	159627502	159627503	*ETFDH*	286.91	1	cCg/cTg	Pro/Leu	non_syn_coding	MED	0	Heterozygous	49.6	133
	chr16	19085297	19085298	*COQ7*	571.41	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	52.7	238
	chr19	51850289	51850290	*ETFB*	721.54	1	aCg/aTg	Thr/Met	non_syn_coding	MED	0	Heterozygous	46.5	405

P. ID: patient identification; CHR: the chromosome on which the variant exists; START: reference start position on chromosome; END: reference end position on chromosome; GENE: gene affected by variant; QUAL: quality score predicted by phred; COD: indicates if variant falls in coding region (1: yes, 0: no); CODON CHG: codon change; A/a: adenine; T/t: thymine; G/g: guanine; C/c: cytosine; A.A. CHG: amino acid change; Trp: tryptophan; *: created stop codon; Met: methionine; Arg: arginine; Cys: cysteine; His: histidine; Thr: threonine; Ile: isoleucine; Ala: alanine; Ser: serine; Gly: glycine; Asp: aspartic acid; Val: valine; Phe: phenylalanine; Glu: glutamic acid; Pro: proline; Leu: leucine; IMPACT: consequence of most severe affected transcript; non_syn_coding: non synonomous substitution altering amino acid sequence; stop_gain: variant changing amino acid to a stop codon; IMPACT SEV: impact severity indicating the severity of the highest order observed for variant; MED: medium; LOF: indicates if variant causes loss of gene function (1: yes, 0: no); FREQ: Frequency of allele call; QUAL: quality score of variant; ORG COV: original coverage of variant.