Clinical and Genetic Studies in Paediatric Mitochondrial Disease

Yehani N.G.G. Wedatilake
Genetics and Genomic Medicine Programme
UCL Great Ormond Street Institute of Child Health
University College London
January 2017

A thesis submitted for the degree of Doctor of Philosophy awarded by University College London

Declaration

I, Yehani N.G.G Wedatilake, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed

_____

Abstract

Paediatric mitochondrial disease is a clinically and genetically heterogeneous group of disorders. Prior to the advent of next generation sequencing, many patients did not receive a genetic diagnosis. Genetic diagnosis is important for prognostication and prenatal diagnosis.

The aim of this thesis was to study the genetic aetiology of paediatric mitochondrial disease. Paediatric patients with suspected mitochondrial disease were grouped to facilitate studying the genetic aetiology. This included grouping patients by respiratory chain enzyme deficiency, by affected end-organ type (cardiomyopathy) or by collectively investigating patients presenting with a similar, genetically undefined clinical syndrome. Whole exome sequencing (WES) was used to investigate genetic causes in patients with suspected mitochondrial disease associated with cytochrome *c* oxidase (COX) deficiency. In another paediatric cohort with mitochondrial cardiomyopathy, mitochondrial DNA sequencing, candidate gene sequencing or WES were used to identify genetic causes. Two families with a unique syndrome (febrile episodes, sideroblastic anaemia and immunodeficiency) were investigated using WES or homozygosity mapping. Candidate genes were functionally evaluated including the use of animal models. Furthermore a natural history study was performed in a monogenic mitochondrial disease (SURF1 deficiency) using clinical data from 12 centres.

WES in COX-deficient patients revealed candidate genes for 14/30, including genes not targeted to the mitochondrion. In a mitochondrial cardiomyopathy (n=30) cohort, 18 genetic causes were identified. *ATAD3A* and *MDH2* were identified as novel causes of human disease. Novel *AARS2* and *VARS2* mutations were found in patients with new clinical presentations. Tubular aggregates were identified in SERAC1 deficiency, as a novel histological finding in this disease. *TRNT1* mutations were identified as causing a unique clinical syndrome. Two candidate genes (*IGHMBP2* and *SCYL1)* were investigated as the cause for a mitochondrial DNA depletion syndrome.

WES is a powerful tool for gene identification in mitochondrial disease leading to discovery of previously unrecognised, novel genetic causes.

Acknowledgements

I am grateful to my supervisors Professor Shamima Rahman and Dr Jan-Willem Taanman for their support, enthusiasm and guidance. In addition I would like to thank Dr Elisa Fassone and Dr Rojeen Shahni postdoctoral researchers at the UCL Institute of Child Health for their time.

I would like to express my thanks to Dr Iain Hargreaves who performs respiratory chain enzyme assays at the diagnostic lab at Neurometabolic Unit, Queen Square, London. In addition I am grateful to Mary Sweeney and Cathy Woodward for mitochondrial DNA sequencing, Suzie Drury for her supervision during the targeted sequencing project, Vincent Plagnol and Warren Emmett for their help with bioinformatics and Latifa Chentouf, clinical research nurse for her assistance with recruitment. I wish to thank all collaborators including Dr Peter Krawtiz, Dr Stephane Pelletier and Dr Michal Minczuk.

I am especially grateful to the Wellcome Trust for awarding me a research training fellowship to pursue this exciting project. I am also grateful to CLIMB for providing funding for the targeted next generation sequencing panel.

I am grateful to Kari and Magnus for their help and support during this time, especially with childcare.

I also wish take this opportunity to express my gratitude to my parents who have always put us before themselves, and still continuously support me in everything I do.

I thank my dearest E, who always encourages and supports me with my work; I thank him for his unwavering belief in me and for sharing his scientific point of view. Last but not least - thank you to our daughter who was born during this period of study, for all the joy she brings into our lives.

Publications arising from this work

Wedatilake Y, Niazi R, Fassone E, Powell CA, Pearce S, Plagnol V, Saldanha JW, Kleta R, Chong WK, Footitt E, Mills PB, Taanman JW, Minczuk M, Clayton PT, Rahman S. TRNT1 deficiency: clinical, biochemical and molecular genetic features. Orphanet J Rare Dis. 2016 Jul 2;11(1):90.

Wedatilake Y, Plagnol V, Anderson G, Paine SM, Clayton PT, Jacques TS, Rahman S. Tubular aggregates caused by serine active site containing 1 (SERAC1) mutations in a patient with a mitochondrial encephalopathy. Neuropathol Appl Neurobiol. 2015 Apr;41(3):399-402

Shahni R, **Wedatilake Y**, Cleary MA, Lindley KJ, Sibson KR, Rahman S. A distinct mitochondrial myopathy, lactic acidosis and sideroblastic anemia (MLASA) phenotype associates with YARS2 mutations. Am J Med Genet A. 2013 Sep;161A(9):2334-8.

Wedatilake Y, Brown RM, McFarland R, Yaplito-Lee J, Morris AA, Champion M, Jardine PE, Clarke A, Thorburn DR, Taylor RW, Land JM, Forrest K, Dobbie A, Simmons L, Aasheim ET, Ketteridge D, Hanrahan D, Chakrapani A, Brown GK, Rahman S. SURF1 deficiency: a multi-centre natural history study. Orphanet J Rare Dis. 2013 Jul 5;8:96.

Table of contents

List of figures

List of tables

List of abbreviations

3-MGA	3-methyl glutaconic aciduria
A site	acceptor site
ADP	adenosine diphosphate
AICAR	5-Aminoimidazole-4-carboxamide ribonucleotide
ALP	alkaline phosphatase
AMP	adenosine monophosphate
ANS	ataxia neuropathy syndromes
APS	ammonium persulphate
ARS	aminoacyl-transfer RNA synthetase
ATP	adenosine triphosphate
B2M	β2 microglobulin
BNGE	blue native gel electrophoresis
BSA	bicinchoninic acid assay
CCCP	carbonyl cyanide 3-chlorophenylhydrazone
cDNA	complementary deoxynucleic acid
CoQ10	co enzyme Q10 (also known as ubiquinone)
COX	cytochrome *c* oxidase
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
cRT–PCR	circularisation reverse transcription PCR
CS	citrate synthase
CSF	cerebrospinal fluid
CT	computerized tomography
D-loop	displacement loop
DAPI	4',6-diamidino-2-phenylindole
DCM	dilated cardiomyopathy
DCMA	dilated cardiomyopathy and ataxia
DD	developmental delay
DMEM	Dulbecco's Modified Eagle's Medium
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphates
DR	developmental regression
DTNB	5,5'-dithio bis-2-nitrobenzoid acid
E site	exit site
EDTA	ethylene diamine triacetic acid
EEG	electroencephalogram
EF-Tu_{mt}	elongation factor thermo unstable (mitochondrial)
EMG	electromyogram
ExAC	The Exome Aggregation Consortium
F	female
f^5C	5-formylcytidine
$FADH_2$	reduced flavin adenine dinucleotide
FBS	fetal bovine serum
Fe-S	Iron sulphur
FLAER	Fluorescent labelled aerolysin
fMet-tRNA	formyl methionine tRNA
GATK	Genome Analysis Tool Kit
gDNA	genomic DNA
GDP	guanosine diphosphate
GGT	gamma glutamyl transferase
GPI	Glycosylphosphatidylinositol

GRACILE	Growth retardation, aminoaciduria, cholestasis, iron overload, lactic acidosis
GTP	guanosine triphosphate
H strand	heavy strand
Hb	haemoglobin
HCM	hypertrophic cardiomyopathy
HMN6	distal hereditary motor neuronopathy type VI
Homo	homozygous
HSP	heavy chain promoter region
$IF2_{mt}$	initation factor 2 (mitochondrial)
$IF3_{mt}$	initation factor 3 (mitochondrial)
IgA	immunoglobulin A
IgG	immunoglobulin G
IgM	immunoglobulin M
IMM	inner mitochondrial membrane
IMS	inter membrane space
KP_i	potassium phosphate
KSS	Kearns Sayre syndrome
L strand	light strand
LHON	Leber hereditary optic neuropathy
LOD	logarithm of the odds
LSP	light chain promoter region
m	months
M	male
MAM	mitochondria-associated membranes
MCV	mean corpuscular volume
MEGDEL	3-methylglutaconic aciduria, deafness, encephalopathy and Leigh-like
MELAS	mitochondrial encephalomyopathy, lactic acidosis with stroke-like episodes
MEMSA	Myoclonic epilepsy, myopathy, sensory ataxia, without ophthalmoplegia
MERRF	Myoclonic epilepsy with ragged red fibres
MIDD	Maternally inherited diabetes and deafness
MLASA	Myopathy, lactic acidosis, sideroblastic anaemia
MNGIE	mitochondrial neuro-gastrointestinal involvement and encephalopathy
MRC	Medical Research Council
MRI	Magnetic resonance imaging
mt-tRNAs	mitochondrial transfer ribonucleic acid
mtDNA	mitochondrial DNA
mtEFG1	Elongation Factor G 1, Mitochondria
mtSSB	mitochondrial single-stranded DNA-binding protein
NA	not available
NADH	reduced nicotinamide dinucleotide
NAG	N-acetyl-beta-D-glucosaminidase
NARP	neuropathy, ataxia and retinitis pigmentosa
NGS	next generation sequencing
NHBLI	National Heart Lung and Blood Institute
OA	optic atrophy
OMM	outer mitochondrial membrane
OXPHOS	oxidative phosphorylation
P site	peptidyl transferase site
PBS	Phosphate Buffered Saline
PCR	polymerase chain reaction
PEO	progressive external ophthalmoplegia
PF	poor feeding
POLRMT	mitochondrial RNA polymerase

PPi	pyrophosphate
PTH	parathyroid hormone
PVDF	polyvinylidene fluoride
PW	poor weight gain
Q	ubiquinone
QH2	ubiquinol
qPCR	quantitative PCR
RBP	retinol binding protein
RCE	respiratory chain enzymes
rCT	received cardiac transplant
RF	respiratory failure
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
SD	standard deviation
SDH	succinate dehydrogenase
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SIFD	sideroblastic anaemia, immunodeficiency, fevers, and developmental delay
SIFT	Sorting Intolerant from Tolerant
SMARD1	Spinal Muscular Atrophy with Respiratory Distress
SNP	Single nucleotide polymorphisms
SQ	semiquinone
Sz	seizures
T1FP	type 1 muscle fibre predominance
T2	type 2 muscle fibres
TALENs	transcription activator-like effector (TALE) nuclease
TE	Tris EDTA
TEMED	N,N,N',N'-tetramethylethylenediamine
TMRM	tetramethylrhodamine, methyl ester
tRNA	transfer ribonucleic acid
TRNT1	CCA-adding transfer RNA nucleotidyl transferase
UCL	University College London
UD	undetectable
V	vomiting
VLCFA	very long chained fatty acids
WES	whole exome sequencing
WGS	whole genome sequencing
τm5s2U	5-taurinomethyl-2-thiouridine
τm5U	taurinomethyluridine

Chapter 1

1. Introduction

1.1 The discovery of mitochondria

Mitochondria are primordial intracellular organelles. Their evolutionary origins trace back to when aerobic proteobacteria were engulfed by eukaryotic cells two billion years ago. Thereafter a lasting endosymbiotic relationship (Margulis 1970; Grivell 1983) developed where the bacteria evolved into mitochondria thereby providing a more efficient aerobic system for cells to produce energy. More recent revisions of this theory postulate that the host was an archaeon (an archaebacterium) and not a eukaryotic cell (Martin 2016). One of the earliest descriptions of mitochondria was in 1890 when Richard Altmann, a German pathologist termed them bioblasts (Altmann 1890). The term mitochondrion was first used by Carl Benda, when he viewed these organelles using light microscopy; it is a combination of the Greek words '*mitos'* and '*chondros*' meaning thread-like and granules respectively (Benda 1898). In 1913, Warburg reported that respiration is linked to particles termed "grana," in extracts of guinea-pig liver (Warburg 1913). In 1949 Kennedy and Lehninger noted that the oxidation of fatty acids and citric acid cycle metabolites, with accompanying ATP formation from inorganic phosphate and ADP, occurred in rat liver mitochondria (Kennedy & Lehninger 1949). Today, we know mitochondria to primarily be the energy producing compartments of eukaryotic cells, with additional roles in programmed cell death (apoptosis), cell signalling, production of reactive oxygen species, haem biosynthesis and steroid synthesis (Smeitink et al. 2012).

1.2 The basic structure of mitochondria

Mitochondria have two membranes termed the inner and outer membranes (IMM and OMM respectively). The IMM forms invaginations called cristae extending into the central lumen or matrix (Palade 1953). The cristae significantly increase the surface area for organellar functions, especially oxidative phosphorylation. In between the membranes lies the intermembrane space (Figure 1.1).

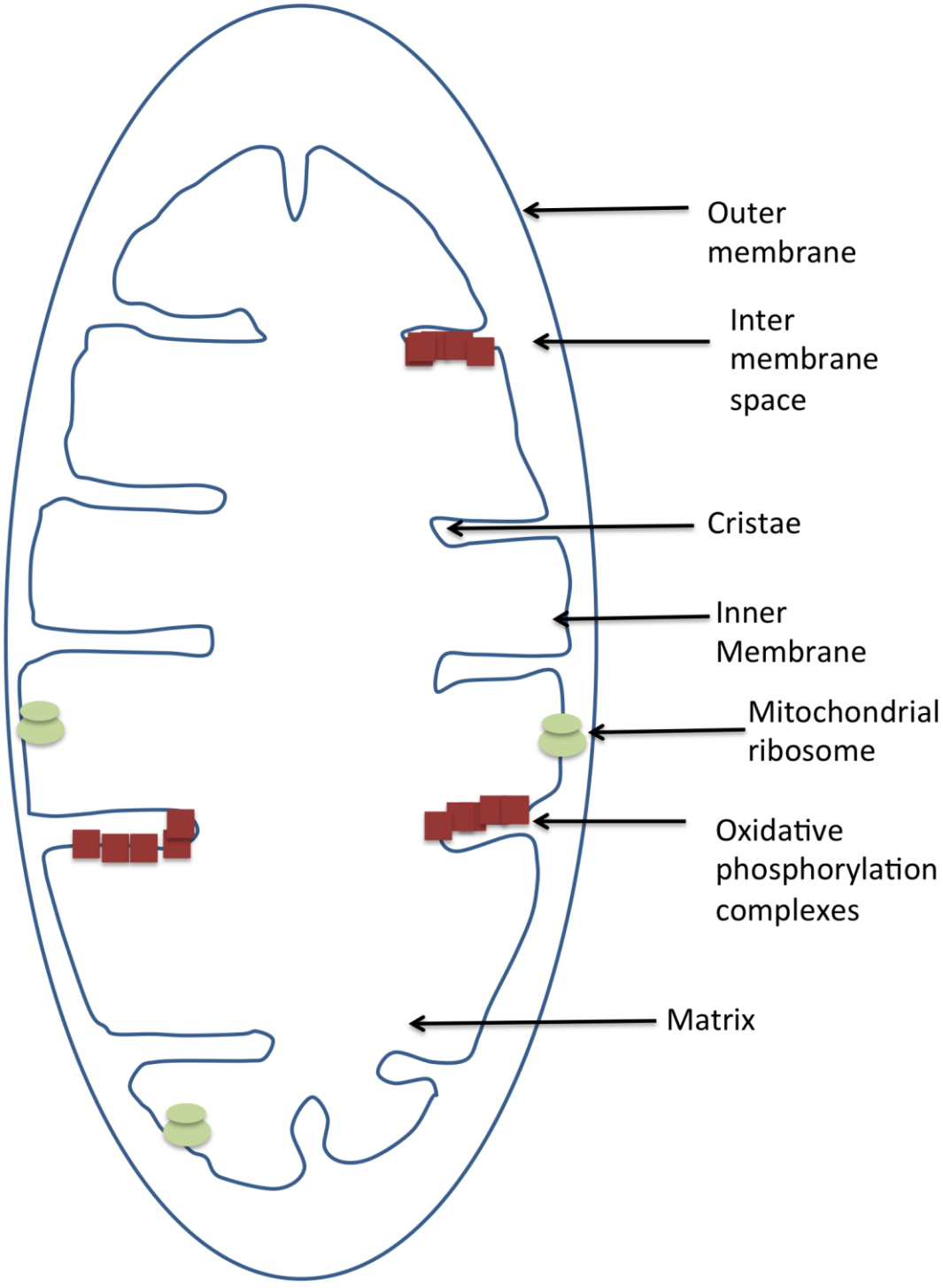


Figure 1:1: The mitochondrial structure.

In between the outer and inner membranes, lies the inter membrane space. The matrix is encapsulated by the inner membrane.

1.3 Mitochondria are dynamic organelles

In 1915 it was noted that mitochondria constantly change shape, with rods or threads transforming into granules, while 'threads' either fuse or branch into networks (Lewis & Lewis 1915). Live cell imaging techniques have demonstrated that within the cell, mitochondria form a dynamic network constantly dividing and fusing with one another (Bereiter-Hahn 1990). In recent years this dynamic process, referred to as 'fission and fusion', has been the focus of much interest. These dynamic processes allow the mitochondrion to be finely tuned with the physiological needs of its cellular environment and enable exchange of contents between mitochondria. The presence of fused, intercommunicating mitochondria are useful in energy consuming cells, whereas cells in inertia often have mitochondria that are relatively fragmented (Skulachev 2001; Collins et al. 2002). The more discrete fragmented mitochondria found in the fission stage generate more reactive oxygen species and are more likely to undergo mitophagy - the selective removal of damaged mitochondria by autophagy. In addition mitochondria are connected to the endoplasmic reticulum (Lewis & Tata 1973) via specialised membrane regions of close apposition between the two organelles called mitochondria-associated membranes (MAM) which facilitate calcium transport into the mitochondria (Szabadkai et al. 2006).

Mitochondrial fusion is complex, as it requires that the coalescing mitochondria merge their respective outer and inner membranes in a coordinated fashion. Optic atrophy protein (OPA1) is involved in IMM fusion whereas Mitofusin-1 (MFN1) and Mitofusin-2 (MFN2) perform OMM fusion (Santel & Fuller 2001; Cipolat et al. 2004).

The key mediator of fission is Dynamin related protein 1 (DRP1), a cytoplasmic protein which upon activation moves to the OMM (Cribbs & Strack 2009). It then multimerises, and forms a ring-like constriction dividing the mitochondrion. In addition there are three proteins, which aid the targeting of DRP1 to the OMM; they are mitochondrial fission protein 1 (Fis1), mitochondrial fission factor (MFF), and mitochondrial elongation factor 1 (Mozdy et al. 2000; Otera et al. 2010; J. Zhao et al. 2011). Furthermore the activity of DRP1 activity is subtly controlled by phosphorylation and dephosphorylation at two key serines residues (Taguchi et al. 2007; Cribbs & Strack 2009). DRP1 activity is increased by phosphorylation of serine 616 while its activity is decreased by phosphorylation of serine 637. More recently a protein termed signal transducer and activator of transcription factor-2 (STAT2) has been found to regulate DRP1 phosphorylation at serine 616 (Shahni et al. 2015).

1.4 The mitochondrial genome

Besides the nucleus, the mitochondrion is the only cellular organelle which contains its own DNA (mtDNA) encoding 13 proteins. As a result, mitochondria also have their own accompanying protein translation machinery to synthesise these 13 proteins. Another unique feature of mtDNA is uni-parental inheritance; mtDNA is inherited exclusively along the maternal line (Giles et al. 1980).

The existence of a mitochondrial genome was detected as early as the 1960s (Nass & Nass 1963) and mtDNA was first fully sequenced in 1981 (Anderson et al. 1981). The mitochondrial genome consists of a 16,569 base pair (bp) circular, double-stranded DNA. The two strands separate according to different buoyant densities into heavy (H) and light (L) strands. This occurs because the heavy strand consists mainly of higher molecular weight purines and the light chain consists of lower molecular weight pyrimidines (Kasamatsu & Vinograd 1974). Unlike nuclear DNA, mtDNA is very compact and intron-less which means that virtually the entire sequence is coding. The only non-coding regions are in the Displacement loop (D-loop) which contains the H strand origin of replication and the L strand origin of replication. MtDNA has a 10-20 fold mutation rate compared to nuclear DNA (Brown et al. 1979). Although there are mtDNA repair mechanisms, these are thought to be inadequate to circumvent damage due to the lack of histones. In addition mtDNA is more vulnerable to damage due to its proximity to reactive oxygen species generated by oxidative phosphorylation complexes (Dianov et al. 2001; de Souza-Pinto et al. 2009).

Within the mitochondrial matrix, mtDNA arranges into small complexes with mitochondrial proteins to form nucleoids (Miyakawa et al. 1987; Holt et al. 2007). Within the nucleoid the most abundant protein is mitochondrial transcription factor A (TFAM) which compacts and packages mtDNA, by forming loops (Alam et al. 2003). TFAM is also a key transcriptional activator and regulates mtDNA copy number in cells (Fisher & Clayton 1985; Larsson et al. 1998).

MtDNA contains 37 genes, of which 13 encode mRNAs for OXPHOS subunits, 22 encode transfer RNAs (tRNAs) and 2 encode ribosomal RNAs (rRNAs) (Figure 1.2). There are thousands of copies of mtDNA in each cell and if a mutation in the mtDNA occurs there will be both copies of the mutant and the wild type mtDNA in a given cell; this phenomenon is known as heteroplasmy (Holt et al. 1990). There is a critical threshold for the proportion of mutant mtDNAs which must be present in a cell, before respiratory chain dysfunction occurs (Schon et al. 1997). Another important concept in mtDNA genetics is the bottleneck phenomenon; during oogenesis only a small subgroup of mtDNA molecules in a primordial oogonial cell are selected to undergo amplification and subsequent transmission to offspring (Hauswirth & Laipis 1982; Marchington et al. 1998). The bottleneck is thought to protect future generations from the high mutation rate of mtDNA, since this selection process prevents the accumulation of deleterious mutations (Carling et al. 2011).

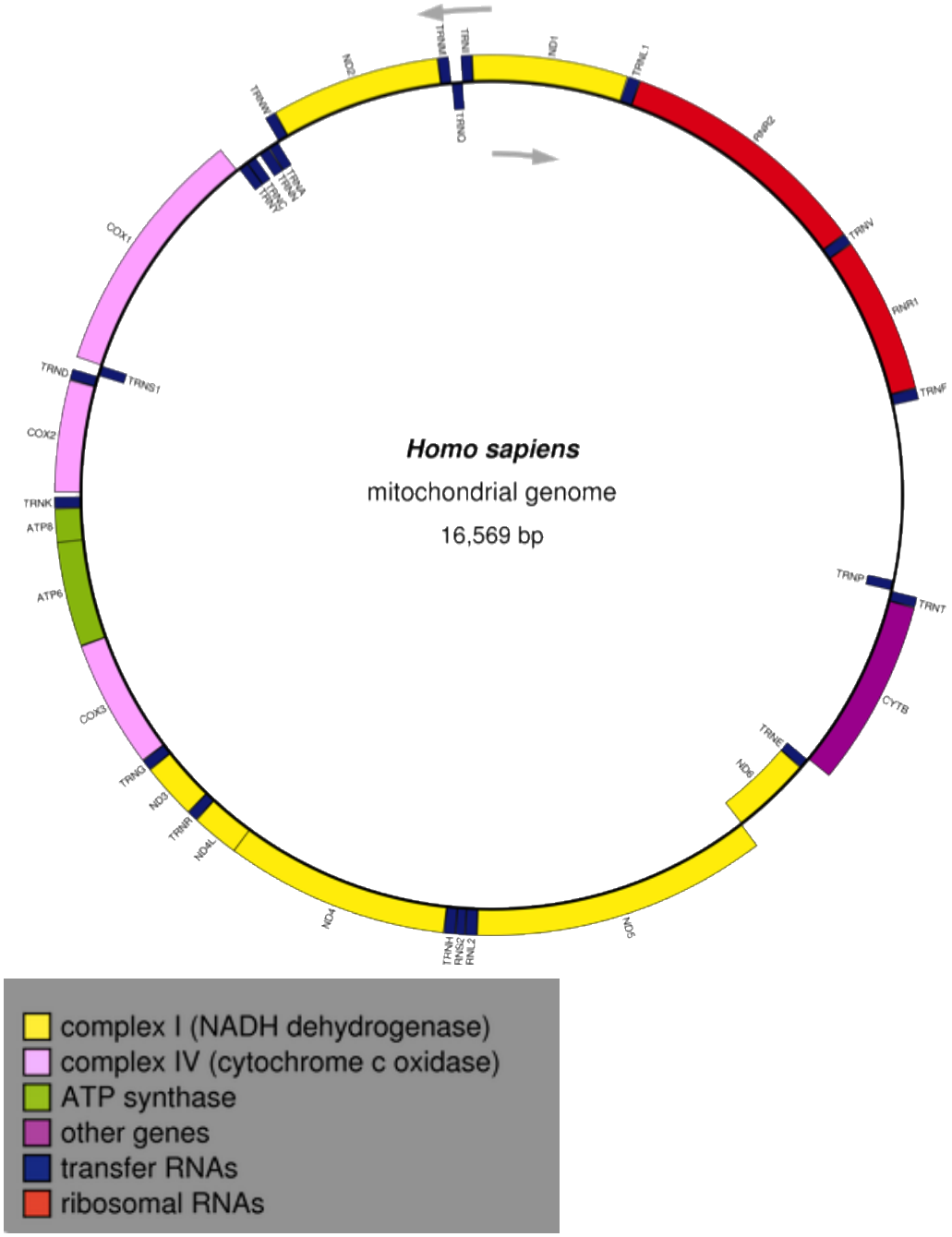


Figure 1:2: Human mitochondrial DNA.

Human mitochondrial DNA encodes for 13 mRNAs, 22 tRNAs and 2 rRNAs. The arrows indicate the direction of transcription. (Figure produced using http://ogdraw.mpimp-golm.mpg.de)

1.5 Mitochondrial DNA replication

Having its own genome would imply that mtDNA replication and transcription are processes independent of the cell nucleus. However in reality several nuclear encoded factors are necessary for mtDNA replication; DNA polymerase gamma, Twinkle (a DNA helicase), mitochondrial single-stranded DNA-binding protein (mtSSB) and mitochondrial RNA polymerase (POLRMT). POLRMT generates initiation primers for mtDNA replication and is also involved in transcription.

The only DNA polymerase in mammalian mitochondria is polymerase γ. Polymerase γ is a heterotrimer consisting of a catalytic subunit which is encoded by the *POLG* gene, and a homodimer of two identical subunits encoded by *POLG2*. The homodimer binds DNA and enhances processivity of replication. Twinkle is a DNA helicase which unwinds double-stranded DNA whereas mtSSB is postulated to maintain single-stranded DNA at replication forks while enhancing the activity of Twinkle and polymerase γ. In addition two factors that are also involved in nuclear DNA replication, Ribonuclease H1 and DNA ligase III, also participate in mtDNA replication. Other factors which are thought to be essential for replication include TFAM (mtDNA replication and packaging) and factors involved in mtDNA base excision repair pathways; the RNA and DNA 5' flap endonuclease (FEN1), RecB-type mitochondrial genome maintenance 5'-3' exonuclease 1 (MGME1), and the helicase/nuclease, DNA2 (Kalifa et al. 2009; Kornblum et al. 2013; Ngo et al. 2014).

Each cell contains thousands of mtDNA molecules which replicate continuously throughout and independently of the cell cycle (Bogenhagen & Clayton 1977), unlike in the biosynthesis of nuclear DNA which replicates only during cell division. There have been two schools of thought explaining the process of mtDNA replication. One mode is where there is a lag between the initiation of synthesis of the two strands and the other is similar to nuclear DNA where both strands are synthesised in parallel.

The classical asynchronous strand displacement model proposed in the 1970s states that the H and L strand synthesis is uncoupled i.e. H and L strands are not replicated simultaneously (Kasamatsu & Vinograd 1973). On the H strand synthesis commences from the origin of H strand replication (O_H), displacing the other strand. Synthesis continues in one direction until it exposes the origin of L strand replication (O_L), approximately after two thirds of the H strand is synthesised. When, the O_L site is exposed, the L strand can initiate synthesis which proceeds in the opposite direction.

At the beginning of the last decade, Holt *et al* proposed a new mechanism, where mtDNA is synthesised in a symmetric manner in both directions (Holt et al. 2000). Here synthesis is postulated to commence at one of two sites O_H or Ori-b. While the leading strand DNA replication progresses from O_H or Ori-b, RNA is simultaneously incorporated on to the lagging

strand. When O_L is reached on the lagging strand, DNA synthesis is initiated, where the lagging strand RNA is converted to DNA.

Mitochondrial deoxynucleotides (dNTPs) are replenished by active transport of cytoplasmic dNTPs, or via the purine and pyrimidine salvage pathways by two mitochondrial deoxyribonucleoside kinases, deoxyguanosine kinase (DGUOK) and thymidine kinase 2 (TK2) (Copeland 2012). In addition thymidine phosphorylase also forms part of the pyrimidine salvage pathway, catalysing the formation of thymidine and phosphate to thymine and deoxyribose-1-phosphate (Nishino et al. 1999).

1.6 Mitochondrial DNA transcription

MtDNA transcription commences on the light and heavy chains at three promoter sites: they are termed light and heavy chain promoter regions (LSP and HSP) (Montoya et al. 1982). In addition the HSP contains two areas of initiation (HSP1 and HSP2) whereas the LSP contains only one area (LSP1) (Taanman 1999). HSP2 and LSP1 are found in the D-loop whereas HSP1 is located closer to the 12S rRNA gene. The HSP and LSP regions function independently during transcription. While HSP1 is responsible for the synthesis of the two rRNAs, $tRNA^{Phe}$ and $tRNA^{Val}$, HSP2 and LSP1 transcribe full-length polycistronic RNAs in the form of 12 mRNAs, 14 tRNA and 2 rRNAs (for HSP2) or 1 mRNA and 8 tRNAs (for LSP). In addition LSP serves as the site for the generation of RNA primers to synthesise the H strand during replication (Chang et al. 1985).

Mitochondrial RNA polymerase (POLRMT), is a nuclear encoded polymerase involved in mtDNA transcription (Tiranti et al. 1997). There are three other nuclear encoded proteins, TFAM and transcription factor B homologous protein 2 (TFB2M) and the termination factor MTERF1 which are thought to be essential for transcription. TFB1M was previously thought to be essential for transcription but is probably not required in mammalian mtDNA transcription (Metodiev et al. 2009). POLRMT contains a promoter recognition loop. Together with POLRMT, TFAM and TFB2M bind the HSP and LSP to initiate transcription. Once transcription has been initiated, the synthesis proceeds around the mtDNA molecule, where polycistronic RNA transcripts are produced from each strand. For transcripts initiated at HSP1, binding of mitochondrial transcription termination factor (mTERF1) aids termination at the 22 bp sequence within the $tRNA^{Leu(UUR)}$ gene (Christianson & Clayton 1986; Daga et al. 1993).

The polycistronic strands are subsequently cleaved at specific points into separate tRNA, mRNA and rRNA species. This cleavage system was previously termed the ‘tRNA punctuation model’ because it was apparent that almost every tRNA is flanked by mRNA and hence, cleaving of tRNAs led to production of mRNA (Ojala et al. 1981; Anderson et al. 1981). The cleavage process is facilitated by cloverleaf structure of tRNAs which functions as a mode of recognition for the cleavage enzymes. Cleavage is thought to require mitochondrial RNaseP

(for 5' cleavage which is a protein complex composed of MRPP 1, 2, and 3) and tRNase Z (for 3' cleavage, ELAC2) (Doersen et al. 1985; Rossmanith 2011). Soon after the tRNAs are cleaved, the rRNAs and mRNAs are polyadenylated by mitochondrial poly(A) polymerase (Tomecki et al. 2004). CCA-adding tRNA nucleotidyl transferase1 (TRNT1) adds a CCA to the 3' tail of the tRNA which enables binding with its cognate amino acid (Nagaike et al. 2001).

Only a few proteins have been identified in the regulation of mitochondrial mRNA stability. LRPPRC (Leucine-rich PPR motif-containing protein) interacts with SLIRP (stem-loop interacting RNA binding protein) to polyadenylate and stabilise mRNAs, especially mRNAs coding for cytochrome *c* oxidase (Ruzzenente et al. 2012). Suv3 helicase interacts with polynucleotide phosphorylase (PNPase, encoded by *PNPT1*) and aids mRNA stability (Minczuk et al. 2002; Piwowarski et al. 2003; Powell et al. 2015).

As described above the role of the mitochondrial tRNA is to recruit its cognate amino acid to join the growing polypeptide chain on the mitoribosome. In order for mitochondrial tRNAs to be recognised by their cognate amino acids, tRNAs undergo extensive editing and modification after being transcribed.

1.7 Post- transcriptional modification of mitochondrial tRNAs

Mitochondrial tRNAs require fewer post transcriptional modification steps than cytosolic tRNAs. Mitochondrial tRNAs also have different secondary structures which deviate from the typical cloverleaf tRNA structure found in their cytoplasmic counterparts. For example mammalian mitochondrial $tRNA^{Ser(AGY)}$ does not possess a D-loop. Base modifications play an important role in maintaining tertiary structure by facilitating folding and stability. In addition the decoding properties of tRNA are regulated by modification of bases at the wobble position (the first position of the anticodon).

Many post-transcriptional base modifications have been identified of which three are mitochondrial specific base modifications: 5-formylcytidine (f^5C), 5-taurinomethyluridine (τm5U), and 5-taurinomethyl-2-thiouridine (τm5s2U) (Suzuki et al. 2002). Four types of modified bases have been identified in wobble positions: 5-formylcytidine f^5C, queuosine, and 5-taurinomethyluridine (τm5U), and 5-taurinomethyl-2-thiouridine (τm5s2U) (Suzuki & Suzuki 2007). Approximately 34 nuclear genes are thought to be responsible for these base modifications (Suzuki et al. 2011). PUS1 encodes a pseudouridine synthase which catalyses the conversion of uridine to pseudouridine, stabilising the secondary and tertiary tRNA structures (Patton et al. 2005). TRMU (MTU1) causes 2-thiolation of the wobble base uridine in $tRNA^{Lys}$, $tRNA^{Glu}$ and $tRNA^{Gln}$ (Sasarman et al. 2011). MTO1 is thought to be responsible in yeast for 5-carboxymethylaminomethylation of the wobble uridine base in some specific mitochondrial tRNAs [$tRNA^{Gln}$, $tRNA^{Glu}$, $tRNA^{Lys}$, $tRNA^{Leu(UUR)}$ and possibly $tRNA^{Trp}$] (Wang et al.

2010). In addition GTPBP3 and MTO1 are predicted to be the enzymes responsible for the τm5U modification (Li & Guan 2002; Li et al. 2002).

1.8 Mitochondrial protein translation

Human mtDNA encodes 22 tRNAs, 13 mRNAs and 2 rRNAs required to synthesise polypeptide subunits of complexes I, III, IV, and V (Anderson et al. 1981).

One of the unique characteristics of mitochondrial protein translation is that mitochondria use a genetic code which differs on several aspects from the universal code. In addition to the universal stop codons UAA and UAG, human mitochondria use AGG and AGA for termination (Osawa et al. 1992). The universal stop codon UGA codes for tryptophan in mitochondria, and the AUA sequence codes for methionine in mitochondria (Barrell et al. 1979). Mitochondrial mRNAs do not have a cap and the poly-A tail sometimes forms part of the stop codon. Crick's wobble hypothesis (Crick 1966) predicts the use of 31 tRNAs but mitochondria use only 22 tRNAs by using a more simple decoding method. There are two species of tRNA methionine for initiation and elongation in other eukaryotic cytoplasmic translation systems whereas mammalian mitochondria use just one tRNA methionine (Mikelsaar 1983).

The current understanding of the mitochondrial translation mechanism stems from our knowledge of the bacterial protein translation system. Currently an *in vitro* mitochondrial ribosome does not exist. Broadly there are three steps in mitochondrial translation and they are described below and shown in Figure 1.3.

1.8.1 Initiation

The mitochondrial ribosome (mitoribosome) has two subunits: the 39S and the 28S subunits. There are several key players in the initiation of translation: the 28S ribosomal subunit, formylmethionine tRNA and initiation factors $IF2_{mt}$, $IF3_{mt.}$ The two mitochondrial ribosomal subunits (39S and 28S) are separated by $IF3_{mt}$ and a complex is formed between $IF3_{mt}$ 28S (Liao & Spremulli 1989). The dissociation of these ribosomal subunits is important for their interface to be accessible by mRNA and formylmethionine tRNA (fMet-tRNA). The mRNA enters into the complex by binding the 28S subunit while the fMet-tRNA attaches to the first mRNA codon. $IF2_{mt}$ then binds the complex which is thought to enable the joining of the larger 39S ribosomal subunit (Liao & Spremulli 1990; Ma & Spremulli 1996).

1.8.2 Elongation

The elongation phase consists of the sequential addition of amino acids onto the growing polypeptide chain. There are two essential parts of this process: codon-anticodon pairing between mRNA and aminoacyl-tRNAs, and subsequent peptide bond formation between amino acids. Initially a ternary complex is formed between elongation factor $EF\text{-}Tu_{mt}$, GTP and aminoacyl-tRNA (Nagao et al. 2007). This complex enters the acceptor (A) site of the

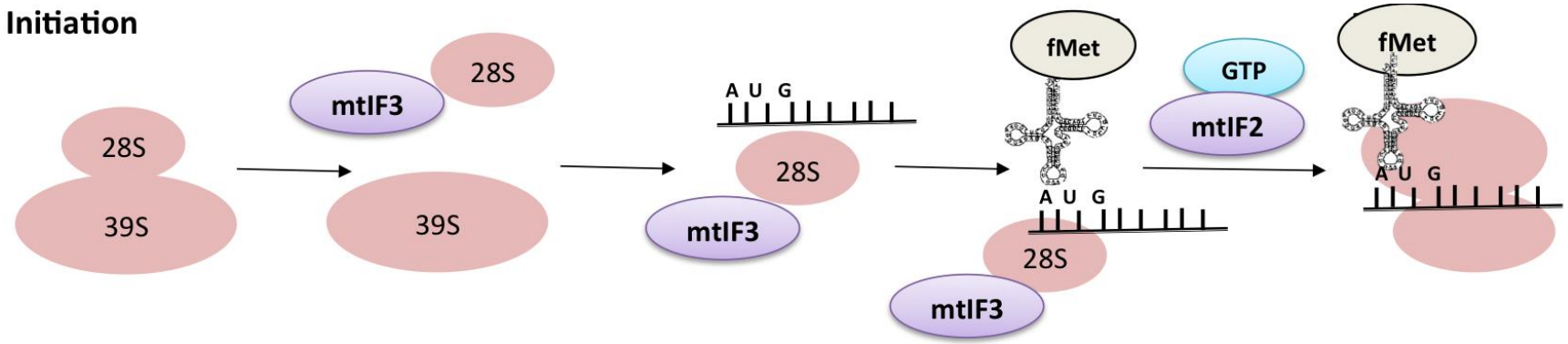
mitochondrial ribosome. When tRNA-mRNA pairing takes place at the A site, GTP is hydrolysed releasing EF-Tu_{mt}-GDP (Woriax et al. 1996; Christian & Spremulli 2012). mtEFTs then reconverts mtEFTu·GDP back to its active form mtEFTu·GTP.

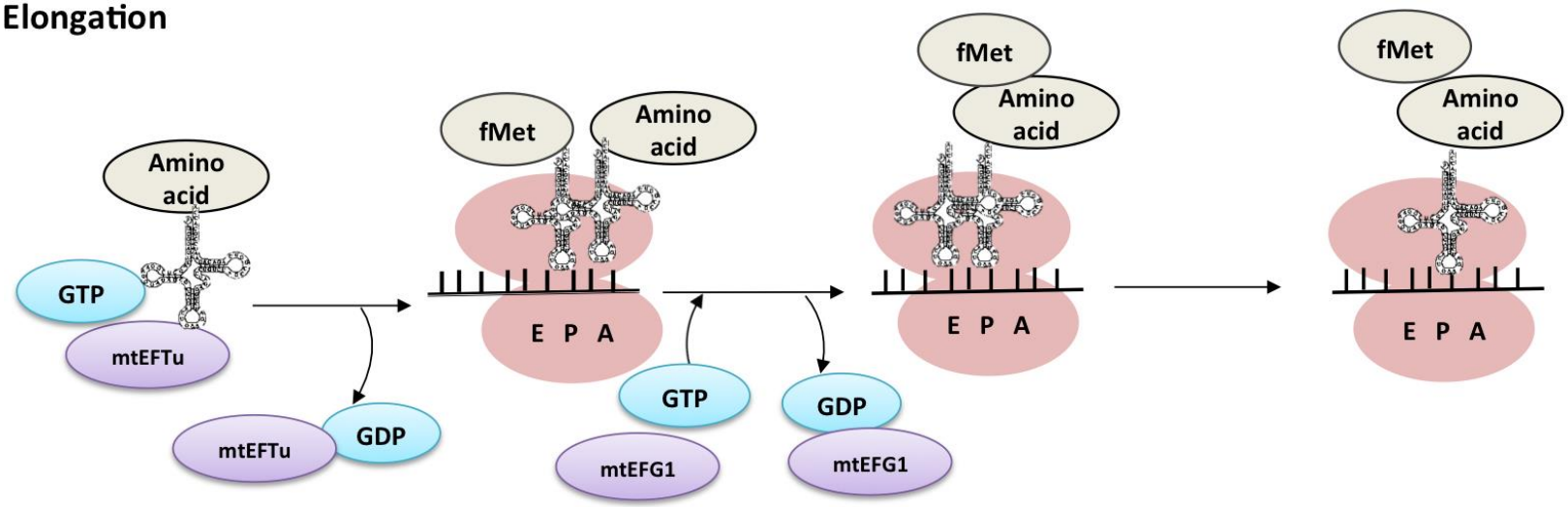
After exiting the ternary complex, the 3' end of the aminoacyl-tRNA moves towards the peptidyl transferase (P) site of the 39S subunit. Here a peptide bond forms as the tRNA adds a new amino acid to the growing polypeptide chain. Meanwhile, the movement of the mRNA chain along the ribosome occurs with the aid of mtEFG1-GTP hydrolysis. This step-by-step movement results in tRNAs which are at the A and P sites moving over to the P and E (exit) sites of the ribosome. The tRNA can exit via the E site, although there is uncertainty about the actual existence of an E site on the mammalian mitochondrial ribosome (Smits et al. 2010). The mRNA thus progresses with the reading of one codon with the above process repeating cyclically until a stop codon (UAA, UAG, AGA or AGG) is found.

1.8.3 Termination

When a stop codon is encountered at the A site, mtRRF1a, a mitochondrial release factor (Soleimanpour-Lichaei et al. 2007) recognises it and this leads to the release of the polypeptide from the last aminoacyl-tRNA in the P site. The enzymatic activity of the 39S subunit is thought to aid the dissociation of the ester bond between the tRNA and the peptide in a GTP hydrolysis dependent manner. The final steps consist of mtRRF1 and mtEFG2 recycling factors to release the mRNA, aminoacyl-tRNA and the ribosomal subunits for the next cycle of protein translation (Smits et al. 2010; Richter et al. 2010).

Mitochondrial Protein Translation





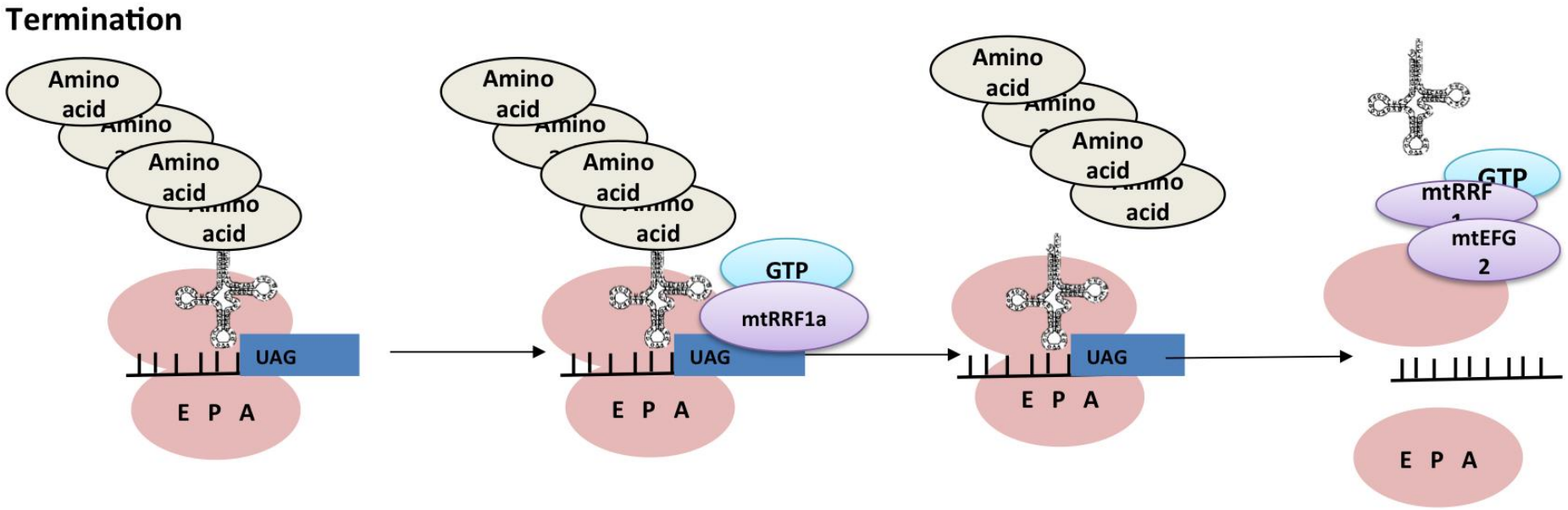


Figure 1:3: Mitochondrial protein translation.

There are three phases of mitochondrial translation—initiation, elongation and termination. The initiation, elongation and termination factors are shown in lilac, amino acids in beige, and GTP or GDP is shown in pale blue. (Redrawn using illustration in (Smits et al. 2010))

1.8.4 Aminoacyl tRNA synthetases

Aminoacyl tRNA synthetases are enzymes which catalyse the binding of respective amino acids to their cognate tRNA which bear the correct triplet anticodon. Human mitochondria utilise 20 different amino acid-specific aminoacyl-transfer RNA synthetase (ARS) enzymes; 17 function exclusively in the mitochondria, and three are bi-functional in mitochondria and cytoplasm (Konovalova & Tyynismaa 2013). The mitochondrial amino acyl synthetases are AARS2, CARS2, DARS2, EARS2, FARS2, HARS2, IARS2, LARS2, MARS2, NARS2, PARS2, RARS2, SARS2, TARS2, VARS2, WARS2 and YARS2 with the first letter denoting their cognate amino acids. GARS, KARS and QARS function both as cytoplasmic and mitochondrial aminoacyl synthetases.

Aminocylation is a two-step reaction where the ARS enzyme binds its respective amino acid and ATP to form an aminoacyl adenylate while releasing inorganic pyrophosphate (PPi). In the second step, the tRNA binds the ARS enzyme while the amino acid is the transferred to the 3' end of the tRNA, after which both AMP and the ARS enzyme are released. The reactions can be summarised as follows (Ibba & Soll 2000):

ARS + AminoAcid + ATP -> ARS(AminoAcid-AMP) + PPi
ARS(AminoAcid-AMP)tRNA -> AminoAcid-tRNA + AMP + ARS

1.9 Mitochondrial oxidative phosphorylation

ATP is generated by five multimeric OXPHOS complexes (complexes I, II, III, IV and V) embedded in the IMM as shown in Figure 1.4. The respiratory chain consists of four complexes (named I, II, III and IV) with two electron carriers, coenzyme Q_{10} (ubiquinone, CoQ_{10}, Q) and cytochrome *c*.

The oxidation of pyruvate, fatty acids and the Krebs cycle produce electrons which are transferred to the respiratory chain via NADH (reduced nicotinamide dinucleotide) and $FADH_2$ (reduced flavin adenine dinucleotide). The electrons produced during oxidation of pyruvate and fatty acids are transferred to complex I via NADH, while the electrons from succinate oxidation and **β oxidation of fatty acids** are transferred to complex II via $FADH_2$. Pyruvate is produced by glycolysis in the cytoplasm and it is actively transported in to the mitochondrion. Within the mitochondrial matrix, pyruvate is oxidised into acetyl CoA by the pyruvate dehydrogenase enzyme, generating NADH and CO_2. Acetyl CoA then enters the Krebs cycle (Krebs & Johnson 1937), and in an eight-step process generates CO_2, three molecules of NADH, one molecule of $FADH_2$ and one molecule of ATP or GTP.

The mitochondrial OXPHOS system is demonstrated in Figure 1.4. Within the respiratory chain enzyme complexes (I, II, III and IV), electron carriers exist in the form of flavins, iron-sulphur (Fe-S) clusters, quinones and the haem groups in cytochromes. In addition, in between the OXPHOS complexes there are mobile carriers of electrons: ubiquinone and cytochrome *c*. During OXPHOS, electron transfer through the complexes is coupled to proton pumping by complexes I, III and IV across the IMM. The electrons are transported from complexes I and II to complex III via ubiquinone. Finally the electrons are transported from complex III to complex IV via cytochrome *c*, where they reduce molecular oxygen to water. Coupled to this electron transfer, protons are pumped into the intermembrane space by complexes I, III and IV and the proton gradient generated (proton motive force, ~150mV) during this process is utilised by ATP synthase (complex V) to produce ATP from ADP and inorganic phosphate using its unique rotatory mechanism (Mitchell 1961; Walker 1998). In addition the OXPHOS complexes are amassed into functionally interconnected units called supercomplexes which allow more efficient transmission of substrates, electron transport, reducing the diffusion time required by quinol or cytochrome *c* and boosting proton pumping (Dudkina et al. 2010).

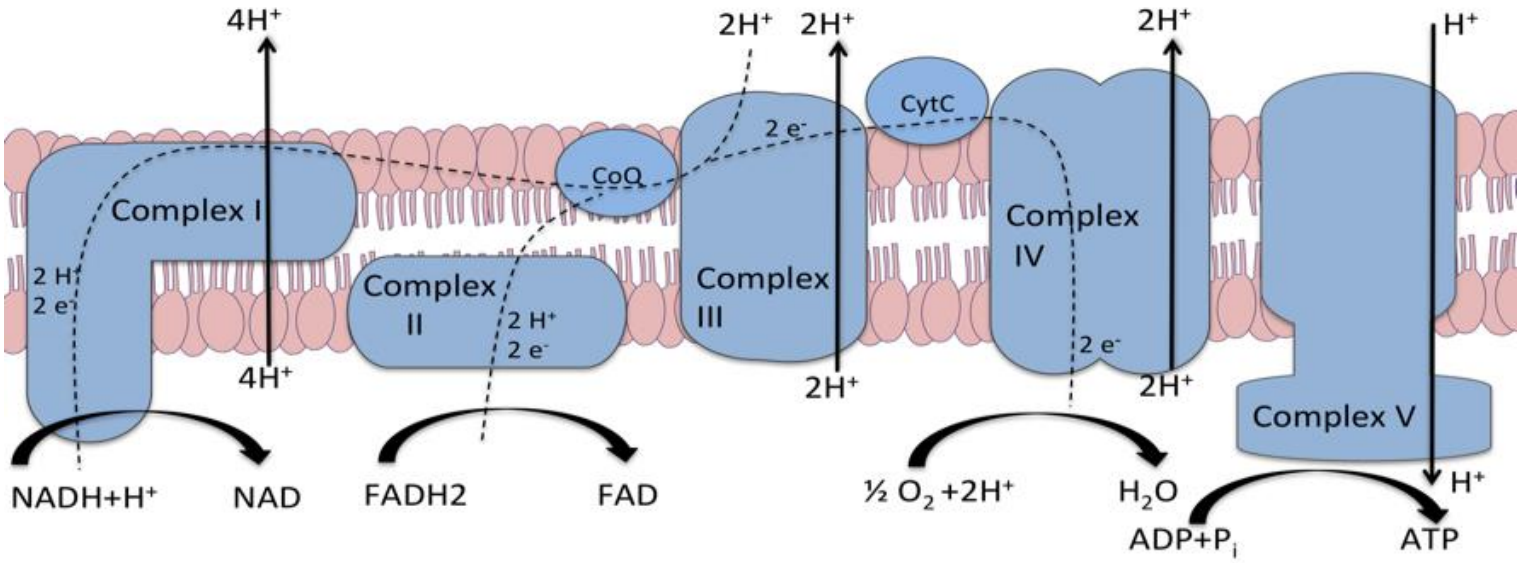


Figure 1:4: Mitochondrial oxidative phosphorylation system:

Complexes I, II, III, IV and V are positioned in the inner mitochondrial membrane. Electrons from NADH are accepted by complex I and then transferred to ubiquinone/coenzyme Q_{10} (CoQ). This electron transfer results in the translocation of protons from the matrix to the intermembrane space. CoQ then delivers the electrons to complex III, which are subsequently transported to cytochrome *c* (CytC). Complex III also transfers H^+ ions from the matrix to the intermembrane space. Cytochrome *c* is oxidised by complex IV and molecular O_2 is reduced to water. Electrons from $FADH_2$ are transferred by complex II to CoQ, but it does not translocate any protons across the inner mitochondrial membrane. The channelling of protons from the intermembrane space through complex V results in the synthesis of ATP. (Redrawn using an illustration in Vademecum Metabolicum, third edition)

1.9.1 Complex I

Complex I (NADH:ubiquinone oxidoreductase) is the largest of the OXPHOS complexes and consists of 44 subunits (~980 kDa) of which 37 are nuclear DNA encoded and the rest are encoded by mtDNA (Baradaran et al. 2013; Vinothkumar et al. 2014). It also has the highest proton pumping capacity with production of ~40% of the proton motive force which is utilised by complex V (ATP synthase).

Complex I transfers electrons between NADH and ubiquinone and translocates four protons across the membrane. It has an L-shaped structure with a membrane arm embedded in the IMM and a peripheral arm that projects into the mitochondrial matrix. The membrane arm contains the hydrophobic mtDNA encoded subunits whereas the nuclear encoded subunits are housed in the peripheral arm. NADH oxidation takes place on the peripheral arm where electrons are transferred through flavin mononucleotide and a group of Fe-S clusters ultimately to the ubiquinone reduction site at the base of the peripheral arm. The membrane domain utilises the energy which is released by electron transfer to pump hydrogen ions into the inter membrane space (Vinothkumar et al. 2014).

1.9.2 Complex II

Complex II (succinate-ubiquinone oxidoreductase) differs from the other OXPHOS enzymes in several aspects: it is the only OXPHOS enzyme which is encoded solely by nuclear genome, it does not pump H^+ ions across the IMM and it also has a role in the Krebs cycle. It is the smallest enzyme consisting of only four subunits (124 kDa). The membrane domain contains the hydrophobic succinate dehydrogenase C and D (SDHC and SDHD) subunits. The hydrophilic domain extends into the matrix and contains the two catalytic subunits succinate dehydrogenase A and B (SDHA and SDHB). According to the crystal structure, Complex II has five prosthetic groups FAD, three iron-sulphur cluster and a haem moiety (Sun et al. 2005). SDHA has a covalently attached FAD cofactor with a succinate binding site, whereas SDHB contains three Fe-S clusters.

SDHA has an enzymatic role in the Krebs cycle by oxidising succinate to fumarate with concurrent FAD^{2+} reduction to $FADH_2$. The electrons released from succinate oxidation first reduce FAD^{2+} to $FADH_2$. $FADH_2$ is then re-oxidised to FAD and the released electrons pass through three Fe-S clusters in SDHB to the membrane integral subunits, and finally to ubiquinone.

1.9.3 Complex III

Complex III (ubiquinol-cytochrome c oxidoreductase) (~240 kDa) is composed of 11 subunits of which only one subunit, cytochrome *b*, is encoded by mtDNA. It contains three haem moieties b_L, b_H and c_1 which are housed within the two cytochrome subunits cytochrome *b* and cytochrome *c1* (Rieske 1976). It also has one Fe-S cluster termed the Rieske centre and two ubiquinol binding sites. The transfer electrons from ubiquinol to cytochrome *c* coupled with

proton transport is known as the Q cycle and first described by Peter Mitchell (Mitchell 1976). The modified version of the Q cycle (Trumpower 1990) demonstrates how the transfer of two electrons to cytochrome *c* leads to the net pumping of four protons across the IMM by complex III. There are two binding sites for ubquinone (Q) and ubiquinol (QH_2) termed Q_i and Q_o respectively. Therefore both Q and QH_2 serve as substrates and products for complex III.

Ubiquinone (Q) binds Q_i site while ubiquinol (QH_2) binds to the Q_o site releasing the first electron to the Rieske centre and the second electron to the b_L haem. The first electron travels from the Rieske centre, next to cytochrome c_1, and finally to oxidised cytochrome *c*. The reduced cytochrome *c* molecule then dissociates away from complex III. The second electron is transferred initially to b_L, next to b_H and ultimately to the Q_i site where it reduces the bound Q to a semiquinone (SQ). The semiquinone is an unstable radical and in order to complete the reaction of converting the semiquinone to ubiquinol (QH_2), a second cycle occurs. Through this cycle complex III is also responsible for the production of superoxide ions.

1.9.4 Complex IV

Complex IV (cytochrome *c* oxidase, COX) is the terminal enzyme of the respiratory chain which is also the only oxygen accepting site. It couples the translocation of four H^+ ions across the IMM to the concomitant oxidation of cytochrome *c* and reduction of O_2 to H_2O. A sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) system identified 13 structural subunits in this enzyme (Kadenbach et al. 1983), but more recent data (Balsa et al. 2012; Pitceathly et al. 2013) indicate that complex IV consists of 14 subunits. This was a surprising finding given that the bovine crystal structure of COX, published in 1996, revealed only 13 subunits (Tsukihara et al. 1996). It is thought that the newly identified nuclear-encoded subunit, NDUFA4 disassociates from the COX holoenzyme during routine purification procedures involving n-dodecyl-β-D-maltoside (Pitceathly et al. 2013).

The mammalian complex IV exists as a dimer and according to the bovine heart crystal structure it contains haem a and a_3, Cu_A, Cu_B, zinc and magnesium (Tsukihara et al. 1996) The haem a and copper moieties comprise its redox centre which is responsible for electron transfer. Oxygen reduction takes place on the matrix side of the IMM and the energy thus released drives the pumping of four protons from the matrix into the IMS. For the purpose of this thesis the COX enzyme is described in more detail.

1.9.4.1 The structure of cytochrome *c* oxidase

The bacterial COX enzyme has only four subunits whereas as described above there are 14 subunits in the mammalian COX enzyme. One subunit was originally thought to be a complex I subunit and therefore has atypical nomenclature, NDUFA4. There are two numbering systems where the other subunits are numbered according to decreasing molecular weight on SDS-PAGE: I, II, III, IV, Va, Vb, VIa, VIb, VIc, VIIa, VIIb, VIIc and VIII (Kadenbach et al. 1983). More recent nomenclature following the Hugo genome nomenclature committee uses MT-CO1, MT-

CO2, MT-CO3, COX4, COX5A, COX5B, COX6A, COX6B, COX6C, COX7A, COX7B, COX7C and COX8 respectively. In addition the following subunits have isoforms (1) COX4 has COX4I1 and COX4I2, (2) COX6A has COX6A1 and COX6A2, (3) COX6B has COX6B1 and COX6B2, (4) COX7A has COX7A1 and COX7A2 (5) COX7B has COX7B and COX7B2 and COX8 has COX8A and COX8C.

The first three subunits (I-III or MT-CO1-3) are mtDNA encoded, are the largest subunits and form the catalytic core of the COX enzyme. The haem a and the haem a_3-Cu_B centre is found in MT-CO1, while MT-CO2 contains the Cu_A centre. Electrons from cytochrome *c* are transferred from Cu_A via haem a to the haem a_3-Cu_B binuclear centre. There is also evidence of an allosteric regulation system within COX where an effector molecule binds to a site other than the catalytic site of an enzyme. An ATP binding pocket has been postulated to be formed by the matrix domain of COX4, MT-CO1 and MT-CO2 (Huttemann et al. 2001). When there is a high ATP/ADP ratio in the mitochondrial matrix, ATP is able to bind and downregulate COX (Follmann et al. 1998).

1.9.4.2 COX assembly

COX assembly has been studied using blue native gel electrophoresis (BNGE); a technique which has allowed the delineation of respiratory chain enzyme complexes in their native form and the demonstration of subassembly intermediates (Schägger & Von Jagow 1999). The COX assembly process appears to be a sequential process where there is a co-ordinated and orderly addition of subunits and co-factors to form the complete holoenzyme. The assembly commences with MT-CO1 as the nucleus with addition of the haem a and Cu_B centres to form the first subassembly intermediate. The next subassembly intermediate is formed by the addition of COX4 and COX5A. Subsequently MT-CO2, MT-CO3, COX5B, COX6B, COX6C, COX7A, COX7C and COX8 subunits are sequentially added. COX6A and COX7B are added into form the final subassembly intermediate (Nijtmans et al. 1998; Williams et al. 2004; Mick et al. 2011). The monomer then undergoes dimerisation to form the enzymatically active holoenzyme. In addition there is evidence in mammalian cells that the addition of the these subunits can occur directly in to supercomplexes (Moreno-Lastres et al. 2012). The electron transfer is performed centrally by MT-CO1 and MT-CO2 (Wilson & Prochaska 1990). The function of the nuclear-encoded surrounding subunits in the mammalian enzyme is still unclear but they may play a role in stabilisation of the mtDNA encoded catalytic subunits (Li et al. 2006; Galati et al. 2009) and modulation of the enzyme activity (Taanman & Capaldi 1993).

During the assembly process a number of accessory factors are recruited for Cu and haem insertion. Haem *a* is a compound present only in COX and its synthesis has been extensively studied in yeast. Haem *b* is converted to Haem *o* by the farnesyl-transferase COX10 and haem *o* is then converted into Haem *a* by COX15 (Brown et al. 2002). SURF1 appears to assist in the insertion of haem a_3 into MT-CO1 and in the maturation of the haem a_3-Cu_B centre in *Parracoccus dentrificans* and *Rhodobacter sphaeroides* (Hannappel et al. 2012).

As described above the COX enzyme also contains two Cu centres (Cu_A in MT-CO2 and Cu_B in MT-CO1). Several accessory factors (COX17, COX11, SCO1, SCO2, COX19, COA6, COX20) are thought to play a role in insertion of these copper moieties (Leary et al. 2004; Leary 2010; Leary et al. 2013; Ghosh et al. 2014; Bourens et al. 2014). COX17 delivers Cu to COX11, SCO1 and SCO2. S- and SCO2 form the Cu_A centre and COX11 is thought to be involved in forming the Cu_B centre (Glerum et al. 1996; Leary et al. 2004). In addition FAM36A appears to assist in the early biogenesis of COX during the incorporation of the MTCO2 subunit (Szklarczyk et al. 2013). The orthologue of the *S. cerevisiae* gene PET191, c2orf64, and c12orf62 have also been implicated in early COX assembly (Huigsloot et al. 2011).

In addition, LRPPRC, a transcriptional factor of COX mRNAs and TACO1 a specific translational activator of the *MT-CO1*, have also been found to cause isolated COX deficiency (Sasarman et al. 2010; Weraarpachai et al. 2009). FASTKD2 appears to be involved in mitochondrial apoptosis rather than in COX assembly (Ghezzi et al. 2008).

1.9.5 Complex V

Complex V (F_o F_1 ATP synthase) is the final component in mitochondrial oxidative phosphorylation. It is a large complex (~500 kDa) consisting of 16 subunits of which two (ATP6 and ATP8) are mtDNA encoded. It has two domains, the catalytic domain (F_1) and a membrane bound domain (F_o). F_1 consists of five types of subunits termed α, β, γ, δ and ε (Abrahams et al. 1994). There is only one each of the γ, δ and ε subunits whereas there are three each of the α and β subunits. ATP synthesis occurs on the β units. F_o consists of one ‘a’ subunit, two ‘b’ subunits and nine to twelve ‘c’ subunits. The ‘a’ and ‘b’ subunits form a peripheral stalk connecting F_1 to the IMM. A central stalk is form by the γ subunit.

ATP synthase has the ability to harness the transmembrane proton motive force developed by the electron transport chain. It has been termed a molecular machine (Boyer 1997) with a central rotor which turns ~150 times per second, while producing three ATP molecules per rotation and using eight H^+ ions per rotation (Walker 1998). The protons flow from the IMS into the matrix via subunit ‘a’ in F_o while the F_1 domain synthesises ATP. When the hydrogen ions flow through F_o, it is forced to twist around along with the central stalk (γ subunit). The three α and the three β beta subunits in the F_1 part cannot rotate, and are locked in a fixed position by the b subunit in the peripheral stalk, which anchors F_1 to the membrane. The central stalk rotates within F_1 causing conformational changes within the β subunits leading to synthesis of ATP from ADP and inorganic phosphate using ‘the binding changing mechanism’.

1.10 Mitophagy

Mitochondria possess their own quality control mechanism where damaged and defective mitochondria are removed by a process called mitophagy. Mitochondria were first visualized within lysosomes in 1962 (Ashford & Porter 1962) but this pathway has become of much interest since the identification of mitophagy proteins in the pathogenesis of Parkinson's disease (Kitada et al. 1998; Valente et al. 2004). During mitophagy defective mitochondria are targeted to autophagosomes, which are subsequently engulfed and degraded by lysosomes. Healthy mitochondria will fuse with the aid of MFN1, MFN2 and OPA1, while DRP1 and Fis1 mediate fission in damaged mitochondria before they undergo mitophagy.

In healthy mitochondria, phosphatase and tensin homolog–induced putative kinase 1 (PINK1) is maintained at very low levels on the OMM by import into the IMM and degradation by the mitochondrial protease presenilins-associated rhomboid-like (PARL). When the membrane potential decreases, PINK1 import and degradation is prevented causing it to accumulate on the OMM (Jin et al. 2010). The accumulation of PINK1 on the OMM leads to recruitment of the E3 ubiquitin ligase Parkin to mitochondria (Matsuda et al. 2010). The presence of Parkin promotes the removal of mitochondria by autophagy (Narendra et al. 2008) by aiding ubiquitination of multiple targets including MFN1 and MFN2 in the mitochondrial membrane. In addition FBXO7 has been shown to activate mitophagy acting together with Parkin in response to mitochondrial membrane depolarisation (Burchell et al. 2013).

The ubiquitination of mitochondrial proteins attracts ubiquitin binding receptors such as p62 (Sequestosome-1) and optineurin, which go on to recruit autophagosomes bearing the LC3 protein (Geisler et al. 2010; Wong & Holzbaur 2014). The cytoplasmic form of LC3 (LC3-I) is converted to the LC3-phosphatidylethanolamine conjugate (LC3-II), which is then recruited to autophagosomal membrane. When the autophagosome fuses with the lysosome, LC3-II is degraded by lysosomal hydrolases.

1.11 Mitochondria and human disease

In 1959 Rolf Luft, a Swedish endocrinologist, first described a case of a young female with "euthyroid hypermetabolism", which was characterised by increased perspiration, polydipsia without polyuria, hyperphagia, decreased body weight, with a 'defect in mitochondrial respiratory control'; these were the beginnings of the study of mitochondrial disease (Luft et al. 1962). Since then over 300 different mtDNA mutations (www.mitomap.org) (Schon et al. 2012) and over 200 (Ng & Turnbull 2015) nuclear genes have been linked to human disease.

Conventionally the term primary mitochondrial disease is used for clinical disorders due to defects in the mitochondrial respiratory chain enzymes (Chinnery 2010) which are estimated to affect 1 in 5000 live births (Skladal et al. 2003). Mitochondrial diseases can be extremely variable in their clinical presentation, giving rise to a variety of clinical syndromes which include

the more commonly recognised diseases such as Leigh syndrome, Pearson's marrow pancreas syndrome, Kearns Sayre syndrome, Myoclonic Epilepsy with Ragged Red Fibres (MERRF) or Mitochondrial Encephalomyopathy, Lactic Acidosis and Stroke-like episodes (MELAS) (Rahman & Hanna 2009). Many more syndromes are increasingly being recognised but most patients present with dysfunction involving multiple organs which do not fall into a specific syndromic pattern, often with predominant involvement of the nervous system. In general mitochondrial disease is suspected if the symptoms and signs fall into a specific syndromic pattern or where there are two or more organs affected or where there is persistent lactic acidosis. However there are some patients with single organ involvement such as mitochondrial cardiomyopathy, Leber hereditary optic neuropathy and epileptic encephalopathy. This clinical variability underpins the difficulties encountered in recognising and diagnosing mitochondrial diseases.

Due to its dual genetic control, OXPHOS defects can be caused by mutations in the mtDNA or nuclear DNA. In addition genetic disorders involving the mitophagy, mitochondrial fission and fusion pathways have also been identified (Baloh et al. 2007; Ashrafian et al. 2010; Shahni et al. 2015). Mitochondrial disorders are therefore a genetically heterogeneous group of diseases and the study of human mitochondrial disease has provided insights into mitochondrial biology and mitochondrial pathways (Vafai & Mootha 2012).

The diagnosis of mitochondrial disease often requires identification of clinical features combined with a complete laboratory diagnostic work-up and a comprehensive three-generation family history. Clinical diagnosis of mitochondrial disease is challenging since there can be a range of different presentations for a given genetic defect.

Laboratory diagnostic tests which are generally found to be useful are: raised lactate levels in blood and cerebrospinal fluid (CSF), raised plasma alanine and proline, raised levels of certain metabolites in the urine organic acids (e.g. methyl glutaconic acid, methymalonate) and abnormal species of acylcarnitines (e.g: C3 carnitine, C4 dicarboxylic carnitine). In a small minority of cases a genetic diagnosis may be evident based on clinical phenotyping and non-invasive laboratory testing. However the majority of patients undergo a muscle biopsy for histology (ragged red fibres, cytochrome *c* oxidase staining) and respiratory chain enzyme analysis for biochemical assay of complexes I, II, III and IV. If the above investigations lead to a presumed diagnosis of mitochondrial disease, the next step is to exclude mtDNA mutations. In the UK, mtDNA sequencing is performed in the diagnostic laboratory setting whereas currently there is only a limited diagnostic service for nuclear gene disorders. The type of respiratory chain enzyme deficiency is often used to guide candidate gene selection in mitochondrial disease. The type of enzyme deficiency is also often used to categorise the type of disease, in the absence of a genetic diagnosis for e.g.: complex I deficient mitochondrial disease.

1.11.1 Disorders due to Mitochondrial DNA mutations

The first breakthrough in mitochondrial DNA genetic disorders was in 1988, when a pathogenic point mutation m.11778G>A was associated with Leber hereditary optic neuropathy (LHON) and single large-scale deletions were linked to Kearns-Sayre syndrome (KSS) (Wallace et al. 1988; Holt et al. 1988). MtDNA mutations are divided into two groups: point mutations and large-scale rearrangements. Large-scale rearrangements are usually sporadic while point mutations are usually maternally inherited. There have been mutations have been reported in each of the 37 mtDNA genes (http://www.mitomap.org/MITOMAP). While most adults have mitochondrial disease due to mtDNA mutations, childhood onset disease is caused mostly by nuclear genetic defects.

1.11.2 Disorders due to nuclear gene mutations

In 1995, the first nuclear gene mutation was found in mitochondrial disease when *SDHA* mutations were found in a patient with Leigh syndrome and complex II deficiency (Bourgeron et al. 1995). The vast majority of nuclear gene mitochondrial disorders are inherited in an autosomal recessive manner, with a few disorders being dominant, or X-linked. Disorders due to nuclear gene mutations can be categorised according to the pathway or process affected as below:

- Mutations in genes encoding OXPHOS subunits and assembly factors
- Mutations in nuclear genes involved in mtDNA replication, maintenance and nucleoside metabolism
- Mutations in genes required for mitochondrial protein synthesis
- Mutations in genes encoding proteins involved in mitochondrial protein and solute import
- Mutations of enzymes required for cofactor biosynthesis (e.g. CoQ_{10}, lipoic acid and iron-sulphur clusters)
- Mutations in genes encoding enzymes required for lipid biosynthesis
- Mutations of proteins involved in mitochondrial fission and fusion
- Mutations of proteins involved in mitophagy
- Mutations in enzymes that cause accumulation of toxic metabolites that disrupt the OXPHOS system

1.11.3 The clinical syndromes of mitochondrial disease

Classical mitochondrial syndromes have been described with some syndromes occurring in childhood and others occurring in adulthood. Syndromes diagnosed in adults include MERRF, MELAS, mitochondrial neuro-gastrointestinal involvement and encephalopathy (MNGIE), neuropathy, ataxia and retinitis pigmentosa (NARP) and progressive external ophthalmoplegia (PEO). Childhood onset syndromes include Leigh syndrome, 3-methylglutaconic aciduria, deafness, encephalopathy and Leigh-like disease (MEGDEL syndrome), Alpers-Huttenlocher Syndrome, Pearson syndrome, MELAS, Sengers syndrome, Barth syndrome and KSS. The

clinical phenotypes and the associated genetic causes of classical mitochondrial syndromes are summarised in Table 1.1.

Table 1.1: Clinical syndromes associated with mitochondrial disease

Syndrome	Clinical features	Onset	Genetic Defect
Leigh syndrome	Developmental delay, neurological regression, optic atrophy, dystonia, tremor, ataxia and central respiratory failure	Infancy, early childhood	>70 different causes; commonly mtDNA mutations involving complex I subunits (especially *MT-ND5*) and complex V subunits (especially *MT-ATP6*)
MEGDEL syndrome	3-Methylglutaconic aciduria, deafness, encephalopathy and Leigh-like disease	Early infancy	*SERAC1*
MtDNA depletion syndromes			
Hepatocerebral	Vomiting, hypoglycaemia, lactic acidosis, liver failure, often with developmental delay and/or regression	Infancy	*POLG, c10orf2, DGUOK MPV17, SUCLG1*.
Myopathic	Hypotonia, muscle weakness, bulbar weakness	Childhood adulthood	*TK2*
Encephalomyopathic	Leigh-like features, sensorineural deafness, dystonia, methylmalonic aciduria	Infancy	*SUCLA2, FBXL4*
MNGIE	Mitochondrial Neurogastrointestinal Encephalomyopathy. Gastrointestinal dysmotility, weight loss, peripheral neuropathy and leukoencephalopathy	Early- late adolescence	*TYMP*
Alpers	A form of hepatocerebral depletion syndrome. Severe	Early childhood	*POLG* (a few cases with

Syndrome	Clinical features	Onset	Genetic Defect
	encephalopathy with intractable epilepsy and hepatic failure**.**		*C10orf2, FARS2, NARS2* and *PARS2*)
MEMSA	Myoclonic epilepsy, myopathy, sensory ataxia, without ophthalmoplegia	Early adulthood	Autosomal recessive *POLG*
ANS	Cerebellar or sensory ataxia and neuropathy, seizures or ophthalmoplegia in some cases	Early adulthood	Autosomal recessive *POLG*
PEO	Bilateral ptosis, ophthalmoparesis, proximal limb weakness, peripheral neuropathy, Sensorineural hearing loss, depression, premature ovarian failure	Early adulthood	Autosomal recessive PEO: *POLG, POLG2, c10orf2, SLC25A4*, RRM2B. Autosomal dominant PEO: RRM2B
MELAS	Mitochondrial encephalomyopathy, lactic acidosis, stroke-like episodes	Childhood-adolescence	Majority: m.3243A>G *(MT-TL1*), others :*MT-ND5, POLG*
Leber hereditary optic neuropathy	Rapidly progressive painless visual loss, centrocecal scotoma	Teenage or 20's	Homoplasmic m.3460G>A, m.11778G>A, m.14484T>C affecting complex I subunits
Optic Atrophy type 1	Gradual decrease in visual acuity), bilateral and symmetric optic nerve pallor	Childhood	*OPA1*
Kearns-Sayre syndrome	Diagnostic triad : onset < 20 years, progressive ophthalmoplegia and one of :heart block, cerebellar ataxia, high CSF protein	Childhood (<20 years)	Large scale mtDNA deletion
Pearson syndrome	Sideroblastic anaemia/ pancytopaenia, exocrine pancreatic deficiency, liver dysfunction	Infancy	Large scale mtDNA deletion
MERRF	Myoclonic epilepsy with ragged	Childhood-	Heteroplasmic

Syndrome	Clinical features	Onset	Genetic Defect
	red fibres. Also dementia and encephalopathy	adolescence	m.8344A>G
NARP	Neuropathy, ataxia, retinitis pigmentosa	Childhood to adulthood	m.8993T>G/C affecting MT-ATP6
MIDD	Maternally inherited diabetes and deafness	20-30 years	m.3243A>G (*MT-TL1)*
Infantile reversible myopathy	Severe muscle weakness, hypotonia, lactic acidosis, respiratory muscle weakness.	Infancy	Homoplasmic m.14674T>C/G (*MT-TE*)
Reversible liver disease	Acute liver failure, lactic acidosis	Infancy	*TRMU*
Barth syndrome	Cardiomyopathy, cyclical neutropaenia, short stature	Infancy	*TAZ* (X linked)
DCMA	Dilated cardiomyopathy and ataxia	Childhood	*DNAJC19*
Sengers syndrome	Congenital cataract, cardiomyopathy	Neonatal	*AGK*
MLASA	Myopathy, lactic acidosis, sideroblastic anaemia	Infancy/ childhood	*PUS1, YARS2*
GRACILE	Growth retardation, aminoaciduria, cholestasis, iron overload, lactic acidosis	Neonatal/ infancy	*BCS1L*

1.11.4 Cytochrome *c* oxidase deficient mitochondrial disease

Patients with COX-deficient mitochondrial disease have heterogeneous clinical phenotypes which include Leigh syndrome, hypertrophic cardiomyopathy, liver disease, sideroblastic anaemia and encephalomyopathy. While mutations have been found in all three mtDNA-encoded catalytic subunits, of the nuclear encoded subunits only COX6A1, COX6B1, COX412, COX7B and COX8A have been associated with human disease. Of the nuclear-encoded assembly factors SURF1, SCO1, SCO2, COX10, COX15, FAM36A, C2orf64, COA3 and C12orf62 have been associated with disease. However in many patients, the cause of isolated COX deficiency remains unknown.

The clinical phenotypes and genetic causes associated with isolated COX deficiency are extremely variable; the disorders described thus far have been caused by mutations in COX assembly factors, subunits and other translational and apoptotic factors summarised in Table 1.2.

Table 1.2: Genotypes and phenotypes associated with isolated COX deficiency in the literature

Gene	Phenotype	References
MtDNA encoded subunits		
MT-CO1	Sideroblastic anaemia	(Gattermann et al. 1997)
	Myoclonic epilepsy	(Bruno et al. 1999)
	Rhabdomyolysis	(Karadimas et al. 2000)
	Epilepsia partialis continua	(Varlamov et al. 2002)
	Myopathy	(Valente et al. 2009)
	MELAS	(Lamperti et al. 2012)
	Motor neurone disease	(Comi et al. 1998)
MT-CO2	Myopathy	(Rahman et al. 1999)
	Encephalopathy	(Horváth et al. 2005; Clark et al. 1999)
	Alpers-Huttenlocher like	(Uusimaa et al. 2003)
	Rhabdomyolysis	(McFarland et al. 2004) (Vissing et al. 2013)
	Strokes, seizures, and lactic acidosis	(Tam et al. 2008)
	Hypertrophic cardiomyopathy	(Wei et al. 2009)
	asthenozoospermy	(Siwar et al. 2014)
MT-CO3	MELAS	(Manfredi et al. 1995)

Gene	Phenotype	References
	Encephalopathy, proximal myopathy	(Hanna et al. 1998)
	Leigh-like	(Tiranti et al. 2000; Mkaouar-Rebai et al. 2011)
	Myopathy and lactic acidosis	(Horváth et al. 2002)
	Bilteral sporadic optic neuropathy	(Bosley et al. 2008)
	Rhabdomyolysis	(Keightley et al. 1996; Marotta et al. 2011)
	MIDD	(Tabebi et al. 2015)
Nuclear-encoded subunits		
COX6A1	Charcot-Marie Tooth disease	(Tamiy et al. 2014)
COX6B1	Encephalomyopathy	(Massa et al. 2008)
	Neonatal lactic acidosis, cystic leukodystrophy	(Calvo et al. 2012)
COX4I2	Exocrine pancreatic insufficiency, dyserythropoeitic anaemia, calvarial hyperostosis	(Shteyer et al. 2009)
COX7B	Microphthalmia, linear skin lesions	(Indrieri et al. 2012)
COX8A	Leigh-like	(Hallmann et al. 2016)
Nuclear-encoded assembly factors		
SURF1	Leigh syndrome or Leigh- like	(Zhu et al. 1998; Tiranti et al. 1998)
	Demyelinating Charcot Marie Tooth Disease	(Echaniz-Laguna et al. 2013)
SCO2	Hypertrophic cardiomyopathy	(Papadopoulou et al. 1999; Vesela et al. 2004; Mobley et al. 2009)
	Spinal muscular atrophy	(Pronicki et al. 2010) Pronicka et al. 2013) (Tarnopolsky et al. 2004) (Gurgel-Giannetti et al. 2013)
SCO1	Hepatic failure and encephalopathy	(Valnot et al. 2000)
	Fatal Infantile Encephalopathy and Lactic Acidosis	(Leary et al. 2004)

Gene	Phenotype	References
COX10	Leigh syndrome	(Antonicka et al. 2003; Coenen et al. 2004)
	Encephalopathy	(Valnot et al. 2000)
	Myopathy, neuropathy, premature ovarian failure	(Pitceathly et al. 2013)
COX15	Hypertrophic cardiomyopathy	(Antonicka et al. 2003) (Alfadhel et al. 2011)
	Leigh syndrome	(Bugiani et al. 2005)
C12orf62	Neonatal lactic acidosis	(Weraarpachai et al. 2012)
C2orf64	Hypertrophic cardiomyopathy	(Huigsloot et al. 2011)
FAM36A	Muscle hypotonia, cerebellar ataxia	(Szklarczyk et al. 2013)
COA6	Hypertrophic cardiomyopathy	(Baertling et al. 2015)
COA3	Neuropathy, exercise intolerance	(Ostergaard et al. 2015)
PET100	Leigh syndrome	(Lim et al. 2014)
Other causes		
TACO1	Leigh syndrome	(Weraarpachai et al. 2009)
LRPPRC	French–Canadian Leigh syndrome	(Debray et al. 2011)
FARS2	Infantile onset epilepsy	(Elo et al. 2012)
RARS2	Pontocerebellar hypoplasia	(Glamuzina et al. 2012)
FASTKD2	Encephalomyopathy	(Ghezzi et al. 2008)
ETHE1	Encephalopathy, microangiopathy, chronic diarrhoea	(Tiranti et al. 2009)
APOPT1	Cavitating leukoencephalopathy	(Melchionda et al. 2014)
AIFM1	Infantile onset motor neurone disease	(Diodato et al. 2015)

1.12 Next generation sequencing technology

The genetic investigation of human disease has been revolutionised by next generation sequencing (NGS). NGS became available after the completion of the human genome project at the turn of the 21st century (Lander et al. 2001; Venter et al. 2001). The double helix structure of DNA was first described by Watson and Crick in 1953 (Watson & Crick 1953). Robert Holley, was the first to sequence a nucleic acid when he sequenced tRNA in 1964 (Holley et al. 1964). In 1977 Frederick Sanger introduced the dideoxyribonucleotide chain-termination method for sequencing DNA which has since been known as Sanger sequencing (Sanger et al. 1977). Sanger and colleagues employed their technique to sequence the first genome, of bacteriophage ϕX174 ('PhiX'), often used as a control genome in next generation sequencing platforms today (Sanger et al. 1977). Mullis first developed the polymerase chain reaction (PCR) in 1983. Sanger sequencing is often termed 'first-generation' technology, whereas newer methods are referred to as NGS. NGS encompasses the techniques of targeted exon sequencing ('gene panels'), whole exome sequencing and whole genome sequencing.

1.12.1 Whole exome sequencing

Whole exome sequencing (WES) involves sequencing of the coding region of the genome (1-2% of the genome). In comparison to traditional Sanger sequencing methods where only one gene at a time is analysed, whole exome sequencing offers the opportunity to analyse all the coding regions of the human genome simultaneously. While this method only provides genetic data of the coding regions, the majority of the mutant alleles found in Mendelian disorders disrupt these protein-coding sequences, which make it a high throughput, cost effective model for causative gene identification (Stenson et al. 2009).

WES uses approximately 1.5-3 µg of DNA, which is broken or 'sheared' randomly into small fragments of 300-400 bp. Adaptors (which are short DNA sequences with universal priming sites within them) are ligated to the ends of each fragment. These adaptors are used to hybridise DNA fragments onto a flow cell or bead and also for sequencing using universal primers. The region of interest (i.e. the exome) is then captured using DNA or RNA oligonucleotides called probes or baits. The sheared DNA (and ligated adaptors) are then hybridised with biotinylated probes. Streptavidin beads or a solid plate are used to physically separate the exome captured by the baits from the remaining DNA.

These massively parallel sequencing methods have been in development since the 1990's. In 2005 the first commercially available NGS method was released called the pyrosequencing method by 454 Life Sciences (currently Roche) (Margulies et al. 2005). It uses a 'sequencing by synthesis' method where the DNA fragments are attached onto resin beads for amplification. The fragments together with the beads are transferred on to wells of a plate for sequencing. One pyrophosphate labelled nucleotide is incorporated at a time, which is followed by pyrophosphate release and light emission.

The Illumina (Solexa) Genome Analyzer IIx (GAIIx) sequencing platform and the Illumina HiSeq sequencers have been the most used platforms and have the advantage of providing cheaper sequencing per base than the 454 method (Carroll et al. 2014). The Illumina/Solexa method (Bennett 2004) uses the sequencing by synthesis method where adaptors at both ends of the DNA fragment bind on to oligonucleotides on the flow cell forming a bridge. The DNA is amplified in clusters termed isothermal bridge amplification before sequencing takes place. It uses a reversible dye terminator method which enables detection of single bases as they are incorporated into growing DNA strands. The Illumina MiSeq sequencer is a benchtop version often used in diagnostic laboratories.

In the SOLiD method (Shendure et al. 2005) the sequencing primers hybridise to the adaptor sequences in the library and four fluorescently labelled di-base probes compete to ligate to the sequencing primer. Multiple cycles of ligation, fluorescence detection and cleavage take place while the number of cycles determines the read length.

The IonTorrent method from Life Technologies detects hydrogen ions which are released when a nucleotide is incorporated.

1.12.2 Gene identification using whole exome sequencing

Next generation sequencing is rapidly superseding other genetic techniques to become the first line investigation in monogenic nuclear gene disorders. A large volume of data is generated, and to overcome this, the development of a clear bioinformatics approach is necessary from the outset. The reads that are obtained after sequencing are aligned against, and compared with, the reference genome (current version hg_19). During this process 20,000-30,000 variants are typically identified per exome and the key challenges lie in filtering the disease-related variants from non-pathogenic polymorphisms and sequencing artefacts.

Since its first diagnostic use in Miller syndrome (Ng et al. 2010) WES has been used successfully to identify the causative gene defect in many Mendelian disorders. Before embarking on WES it is highly desirable to have recognised the likely inheritance pattern within a pedigree i.e. autosomal dominant, recessive, X linked or maternal inheritance. Common strategies for bioinformatics filtering include filtering out the commonly observed polymorphisms and focusing on variants which are predicted to be pathogenic according to mutation prediction software. Generally, variants are filtered against single nucleotide polymorphism (SNP) databases such as dbSNP, the 1000 genome project and the Exome variant server to filter benign variants. In addition several algorithm-based software exist to predict if a missense variant is deleterious or damaging. Examples include SIFT (Kumar et al. 2009), PolyPhen v2 (Adzhubei et al., 2010) and MutationTaster (Schwarz et al. 2010) which primarily use protein conservation data across species or the predict the consequences on the protein conformation.

Additionally in the case of mitochondrial disorders, there is an inventory of genes where the protein products are known or predicted to be primarily localised to the mitochondria (often termed the mitochondrial proteome or mitoproteome) called Mitocarta 2 (Pagliarini et al. 2008; Calvo & Mootha 2010; Calvo et al. 2015). This inventory contains 1158 human genes predicted to produce proteins localised to the mitochondrion which were identified by proteomic screening of mitochondrial fractions from mouse tissues and eight different prediction methods. The knowledge of the mitoproteome has been used in targeted NGS. Baits were designed to capture the MitoCarta genes and the 37 mtDNA genes in the diagnosis of childhood mitochondrial disease (Calvo et al. 2012). In addition the MitoCarta Compedium has been used during bioinformatics filtering of exome sequencing in mitochondrial disease to identify the causative variants (Haack et al. 2012; Carroll et al. 2014).

Some of the approaches to filtering variants and finding the causative gene are 1) filtering rare or novel variants shared by two or more affected siblings with identical clinical phenotypes; 2) filtering rare or novel variants shared by several unrelated, probands with identical clinical phenotypes; 3) analysis of variants from exomes of both parents and an affected child ('trio' analysis), particularly useful to identify *de novo* heterozygous mutations and 4) homozygosity mapping/linkage analysis combined with analysis of rare variants found in whole exome sequencing, used in consanguineous families.

1.12.3 Whole genome sequencing

Although proof of principle data for whole genome sequencing (WGS) in rare diseases has been available for several years (Lupski et al. 2010), the cost of whole genome sequencing (WGS) has been prohibitive. The human genome project cost ~US$ 2.7 billion, although the cost fell sharply afterwards, in 2011 the cost per genome was still ~ US$10,000. In 2014 Illumina announced that their HiSeq X Ten technology would offer WGS for < US$1000 per genome. In addition to enhanced clustering, the DNA templates are packed more densely in nanowells (rather than in the flow cells) on the HiSeq X, enabling the read out of more data per run. The HiSeq X Ten also uses newly developed polymerase enzymes with faster reaction times and a faster camera (Check Hayden 2014).

WGS offers complete coverage of coding and non-coding regions of the genome. However additional challenges with WGS include bioinformatics handling of the millions of variants which are generated and methods for secure storage of the large volumes of raw data where ~200 GB of data is generated per genome.

In the UK genome sequencing has already commenced on a large scale. In 2012 Genomics England launched the 100,000 Genomes Project with the aim of sequencing 100,000 genomes from patients with rare diseases and cancer by 2017.

1.13 Aims of this thesis

The aims of this thesis are:

- To identify a cohort of patients with suspected mitochondrial disease due to isolated COX deficiency
- To investigate the genetic basis of their disease by using next generation sequencing
- To establish pathogenicity of any novel mutations in patient fibroblast cell cultures
- To identify a cohort of patients with mitochondrial cardiomyopathy and study the genetic aetiology of their disease
- To investigate the genetic basis of a disease in two families presenting with a suspected mitochondrial disease consisting of fever, diarrhoea and vomiting, pancreatic insufficiency, sideroblastic anaemia and developmental delay
- To study the natural history of a mitochondrial disease due to a single gene disorder

The following work was performed between June 2012 – June 2016 (including one year of maternity leave).

Chapter 2

2. Patients and methods

2.1 Recruitment of a patient cohort with isolated COX deficiency

Patients were recruited through the paediatric mitochondrial clinical service at Great Ormond Street Hospital, London UK. This study was approved by the National Research Ethics Committee - London Bloomsbury (09/H0713/76).

Patients were suspected as having a mitochondrial disease if they displayed a phenotype consistent with a canonical mitochondrial syndrome (as described in table 1.1 chapter 1) or where mitochondrial disease clinical criteria (Morava et al. 2006) pointed towards a diagnosis of possible or probable mitochondrial disease. In addition patients were required to have isolated COX deficiency in muscle, where COX deficiency was defined as COX activity/citrate synthase ratio of <2 standard deviations (SD's) (<0.014) from the control mean (0.024). The control mean has been established in the Neurometabolic unit, a national diagnostic laboratory, Queen Square, London. In the case of similarly affected siblings, there were some cases where only one sibling underwent muscle biopsy. In such cases if isolated COX deficiency was found in one sibling, both siblings were recruited to the study. Fibroblast COX activity was not routinely available in all patients and therefore was not established as a recruitment criterion. Using these criteria, 74 patients were recruited to the study between 2011 and 2016, excluding cases with a common mtDNA mutation.

Upon recruitment, where a candidate gene was presumed to be the cause, Sanger sequencing of the gene was carried out, often in the diagnostic laboratory. In some recruited cases, it was apparent that investigations were already underway in another research laboratory and these cases were not investigated further. The phenotypic findings in the cohort are described in Tables 2.1- 2.3. Table 2.1 shows patients from families with more than one affected sibling, Table 2.2 shows patients from consanguineous families with only one affected child and Table 2.3 shows patients from non-consanguineous families with only one affected child. Within each table, the patients are listed according to date of birth of the patient or the older sibling.

Table 2.1: Patients with suspected COX deficient mitochondrial disease in families with more than one affected child

(includes consanguineous families with more than one affected child).

Patient	Sex	Clinical features	MRI brain	Muscle histology	Highest plasma lactate mmol/L (0.7-2.1)	CSF lactate mmol/L (<2)	COX citrate synthase ratio (0.014-0.034)
Family 1							
S1	F	Sideroblastic anaemia, failure to thrive and hepato-renal failure, Pearson-like	NA	NA	NA	NA	NA
S2	F	Developmental delay, deafness, partial villous atrophy, inflammatory enteropathy, bone marrow dysplasia, Pearson-like	NA	Slightly increased lipid	2.2	1.1	0.008
Family 2							
S3	F	Developmental delay, renal tubular wasting, hypocalcaemia	NA	NA	NA	NA	NA
S4	F	Congenital lactic acidosis, microcephaly, dysmorphic features, gastroesophageal reflux, failure to thrive	Microcephaly	Normal	3.3	1.2	0.006
S5	F	Renal tubulopathy, hypocalcaemia, developmental delay	NA	Normal	1.3	1.7	0.003
Family 3							
S6	F	Developmental delay, dilated cardiomyopathy	Normal	NA	1.7	NA	NA
S7	M	Developmental delay, dilated cardiomyopathy	Normal	COX negative fibres	2.1	NA	0.008
Family 4							

Patient	Sex	Clinical features	MRI brain	Muscle histology	Highest plasma lactate mmol/L (0.7-2.1)	CSF lactate mmol/L (<2)	COX citrate synthase ratio (0.014-0.034)
S8	F	Hereditary sensory motor neuropathy, cholestatic liver disease, mtDNA depletion in muscle	Normal	Neurogenic changes, pale staining of COX	0.7	NA	0.010
S9	F	Hereditary sensory motor neuropathy, cholestatic liver disease, mtDNA depletion in muscle	Normal	Pale staining with SDH and COX stains	1.0	NA	0.011
Family 5							
S10	F	Fever, sideroblastic anaemia, sensorineural deafness, retinopathy	NA	NA	2	NA	0.007
S11	M	Fever, sideroblastic anaemia, sensorineural deafness, retinopathy	NA	NA	2.4	NA	NA
Family 6							
S12	F	Lactic acidosis, cardiomyopathy	NA	Normal	26	NA	NA
S13	M	Lactic acidosis, cardiomyopathy, adrenal insufficiency	NA	Normal	5.2	NA	0.009
Family 7							

Patient	Sex	Clinical features	MRI brain	Muscle histology	Highest plasma lactate mmol/L (0.7-2.1)	CSF lactate mmol/L (<2)	COX citrate synthase ratio (0.014-0.034)
S14	F	Leigh syndrome, hypothyroidism, demyelinating polyneuropathy	Symmetrical changes in lenticular nuclei	Mild myopathic features	1.1	NA	0.007
S15	M	Leigh syndrome, acute encephalopathy, motor and speech delay, seizures	Symmetrical changes in lenticular nuclei	NA	2.4	NA	NA
Family 8							
S16	F	Ponto cerebellar atrophy, microcephaly, cortical visual impairment, epilepsy	Pontocerebellar atrophy	Excess lipid, mitochondrial accumulation	1.6	NA	0.007
S17	M	Profound visual impairment, nystagmus, microcephaly, infantile spasms	NA	NA	3.8	NA	NA
Family 9							
S18	F	Ataxia, dystonia	Slightly enlarged ventricles and sulci	Excess lipid, mild myopathic changes	1.2	3.1	0.006

Patient	Sex	Clinical features	MRI brain	Muscle histology	Highest plasma lactate mmol/L (0.7-2.1)	CSF lactate mmol/L (<2)	COX citrate synthase ratio (0.014-0.034)
S19	M	Ataxia, dystonia	NA	NA	2.4	NA	NA
Family 10							
S20	M	Developmental delay, seizures, hypotonia cerebellar atrophy	Cerebellar atrophy	No diagnostic features	1	1.3	0.009
S21	M	Developmental delay, hypotonia, cerebellar atrophy	Cerebellar atrophy	No diagnostic features	2.3	1.4	0.013
Family 11							
S22	M	Epileptic encephalopathy, developmental delay	Lack of white matter	NA	1.1	1	0.014
S23	M	Epileptic encephalopathy, liver failure, (Alpers-like)	Bilateral perisylvian insular polymicrogyria	Normal	1.8	1.2	0.009

Table 2.2: Patients with suspected COX deficient mitochondrial disease in consanguineous families with one affected child

Patient	Sex	Clinical features	MRI brain	Muscle histology	Highest plasma lactate mmol/L (0.7-2.1)	CSF lactate mmol/L (<2)	COX citrate synthase ratio (0.014-0.034)
S24	M	Generalised epilepsy, Lennox-Gastaut spectrum, severe learning difficulties, low CSF folate	Cerebral atrophy	NA	1.7	1.3	0.007
S25	F	Leigh syndrome, history of regression, microcephaly, seizures, ataxic gait, motor neuropathy	Brain stem involvement	Excess lipid	2.3	2.4	0.004
S26	M	Global developmental delay, lactic acidosis, microcephaly	Leukodystrophy	Slightly increased lipid	5.7	6.1	0.008
S27	M	Diplegic motor disorder, expressive language delay, motor neuronopathy	Pontocerebellar hypoplasia	Neurogenic changes	1.6	1.1	0.006
S28	F	Sensorineural hearing loss, central hypotonia, retinal dystrophy	Bilateral symmetrical increased signal in the globus pallidus	COX negative fibres.	2.2	2.7	0.012
S29	F	Acute encephalopathy, focal seizure with secondary generalisation, dystonia, pyramidal	Abnormal signal in the right frontal corona radiata	Excess lipid	6.9	3.2	0.005

Patient	Sex	Clinical features	MRI brain	Muscle histology	Highest plasma lactate mmol/L (0.7-2.1)	CSF lactate mmol/L (<2)	COX citrate synthase ratio (0.014-0.034)
		signs, pyrexia, hepatomegaly, microcytic anaemia, liver biopsy: microvesicular fatty change					
S30	M	Muscle weakness, recurrent chest infections	NA	NA	3.6	NA	0.006
S31	F	Lactic acidosis, sensorineural deafness, developmental delay, cerebral cyst, dystonia	Putaminal lesions	Excess lipid	4.3	NA	0.007

CSF: cerebrospinal fluid, F: female, M: male, MRI magnetic resonance imaging, NA not available

Table 2.3: Patients with suspected COX deficient mitochondrial disease in non-consanguineous families with one affected child

Patient	Sex	Clinical features	MRI brain	Muscle histology	Highest plasma lactate mmol/L (0.7-2.1)	CSF lactate mmol/L (<2)	COX citrate synthase ratio (0.014-0.034)
S32	M	Cardiomyopathy, Sensorineural deafness, retinal dystrophy, renal failure	NA	Non specific features	1.6	2.6	0.010
S33	F	Cerebellar atrophy, sensorimotor neuropathy, septo-optic dysplasia	Cerebellar atrophy	Neurogenic changes	NA	NA	0.010
S34	M	Sideroblastic anaemia, sensory motor axonal neuropathy, congenital deformity of hands and fingers, refractory bicuspid aortic valve, Retinal dystrophy	Cerebellar atrophy	Myopathic changes	1	NA	0.009
S35	M	Chronic inflammatory demyelinating polyneuropathy, epilepsy, chronic anaemia, colon inflammation	Normal	Targeted and core-like lesions	2.6	1.7	0.009
S36	M	Basal ganglia lesions, ptosis, apnoea	Basal ganglia lesions	NA	4.2	3.9	
S37	M	Giant axonal neuropathy, unsteady gait and poor motor skills, progressive motor and sensory neuropathy, central precocious puberty, progressive scoliosis	Leukoencephalopathy, abnormal signal in the dentate nuclei and globus pallidi	Abnormality of fibre size	2.6	1.9	0.008
S38	M	Developmental delay, lactic acidosis	Prominent CSF spaces and lateral ventricles	Small type 1 fibres	5.2	3.1	43% of lowest control

Patient	Sex	Clinical features	MRI brain	Muscle histology	Highest plasma lactate mmol/L (0.7-2.1)	CSF lactate mmol/L (<2)	COX citrate synthase ratio (0.014-0.034)
S39	F	Hypopituitarism, retinal dystrophy, fatigue	Neurodegenerative disorder with progressive leukodystrophy	Normal	1.9	1	0.010
S40	F	Chronic renal disease treated by renal transplant, dystonic movement disorder, mild left ventricular hypertrophy	Bilateral striatal necrosis and mild cerebellar hypertrophy	Increased lipid. clusters of mitochondria.	2.3	NA	0.010
S41	F	Recurrent episodes of encephalopathy, hypoglycaemia, rhabdomyolysis, episodic ataxia	Normal	NA	1.2	NA	0.008
S42	M	Nephrotic syndrome, mitral valve regurgitation, transient ischaemic attacks with dysphasia	Left caudate lesion	Normal	2.7	1.1	0.008
S43	M	Lactic acidosis, hyperammonaemia, sensorineural deafness, 3-methyl glutaconic aciduria, Leigh-like	Abnormal striatal signal	Tubular aggregates and excess lipid	7.4	2	0.013
S44	M	Hypotonia, ptosis, bulbar palsy, seizures	Multifocal areas of swelling in cortex and white matter	Very pale COX prominent lipid	2.6	1.2	0.004
S45	F	Congenital ptosis, partial agenesis of corpus callosum, headaches	Partial agenesis of corpus callosum	Irregular large mitochondria	1	NA	0.007

Patient	Sex	Clinical features	MRI brain	Muscle histology	Highest plasma lactate mmol/L (0.7-2.1)	CSF lactate mmol/L (<2)	COX citrate synthase ratio (0.014-0.034)
S46	M	Leigh-like, developmental delay, growth failure	Abnormal signal in medulla, middle cerebellar peduncles, dentate and caudate	Excess lipid	3	5.4	0.004
S47	M	Sideroblastic anaemia, raised lactate, biventricular hypertrophy	NA	Excess lipid	3.5	2.3	0.003
S48	M	Developmental delay, hypotonia, low serum copper	Cerebellar atrophy, decreased bulk of white matter	Normal	2.7	1.6	0.009
S49	M	Global developmental delay with motor regression, progressive feeding difficulties (partially gastrostomy fed), dislocated hips	Cerebellar atrophy, mild white matter signal abnormalities	Neurogenic changes	1.3	NA	0.011
S50	F	Non-obstructive hypertrophic cardiomyopathy, Cardiovascular collapse, mild right hemiparesis speech delay	Middle cerebral artery area infarction	NA	1.5	2	0.009
S51	M	Leukoencephalopathy, epilepsy, dystonia, learning difficulties, exocrine pancreatic insufficiency	Leukoencephalopathy	Excess lipid	2.4	NA	0.012
S52	F	Epilepsy, right sided weakness, learning difficulties	Normal	NA	1.5	1.4	0.005
S53	F	Motor delay, joint laxity, poor weight gain, fatiguability	Bilateral periventricular	Normal	1.3	NA	0.006

Patient	Sex	Clinical features	MRI brain	Muscle histology	Highest plasma lactate mmol/L (0.7-2.1)	CSF lactate mmol/L (<2)	COX citrate synthase ratio (0.014-0.034)
			signal abnormality				
S54	F	Poor feeding, faltering growth, hypotonia, motor delay, developmental regression, slurred speech	Slightly reduced white matter bulk	Mild myopathic changes	3.1	1.5	0.011
S55	M	Deafness, choreoathetosis	Symmetrical changes in globus pallidi	Normal	1.2	1.6	0.009
S56	F	Ataxia, ptosis, myopathy, movement disorder	NA	Non specific features	1.1	1.3	0.007
S57	F	Myopathy, poor coordination, ataxia, speech delay	Cerebellar hypoplasia	Type 1 fibre predominance	1.8	NA	0.010
S58	F	Epileptic encephalopathy, global developmental delay, gastro-oesophageal reflux, motor disorder, visual impairment	Normal	Normal	1.3	1.2	0.006
S59	M	Sensorineural deafness, hypermetropia, developmental delay, sensorimotor neuropathy, truncal ataxia	Leukodystrophy	Excess lipid	1.7	1.4	0.007
S60	M	Limb dystonia, myopathy on EMG	Dorsal putaminal lesions	Mild myopathic changes	1.4	1.3	0.005
S61	M	Global developmental delay, dystonia, eosinophilic colitis, delayed visual maturation, seizures, faltering growth	Abnormal signal lesion within the right frontal corona radiata	Myopathic changes, necrotic fibres,	1.4	1.1	0.007

Patient	Sex	Clinical features	MRI brain	Muscle histology	Highest plasma lactate mmol/L (0.7-2.1)	CSF lactate mmol/L (<2)	COX citrate synthase ratio (0.014-0.034)
				hypertrophic fibres.			
S62	F	Four limb dystonia, motor skill regression, developmental delay	Bilateral striatal lesions also involving left caudate	Myopathic changes. Excess lipid	2.3	2.1	0.006
S63	M	Myoclonic epileptic encephalopathy, developmental delay,	Normal	Excess lipid	1.1	NA	0.011
S64	F	Peripheral neuropathy, auditory neuropathy, bilateral ptosis, gastrostomy fed, respiratory failure	NA	Excess lipid, prominent mitochondria	1.0	1.1	0.010
S65	M	Generalised hypotonia and weakness, mild to moderate sensorineural deafness, bulbar dysfunction, gastrooesophageal reflux disease	Normal	Variation in fibre size, increased lipid	2.3	1.3	0.006
S66	F	Speech delay, four limb motor disorder, ataxia, behavioural problems, ventrolateral thalamic changes, Leigh-like	Bilateral ventro thalamic changes	Excess lipid	1.2	NA	0.004
S67	F	Seizures, dystonia, congenital lactic acidosis	Pontocerebellar hypoplasia	Normal	9.5	2.4	0.011
S68	F	Global developmental delay, ataxia, fatigue, myopathic	Delayed myelination	Inflammatory	1.2	1.2	0.005

Patient	Sex	Clinical features	MRI brain	Muscle histology	Highest plasma lactate mmol/L (0.7-2.1)	CSF lactate mmol/L (<2)	COX citrate synthase ratio (0.014-0.034)
		EMG, hypermobility		changes			
S69	M	Infantile epileptic encephalopathy, infantile spasms, sensorineural deafness	Mild reduction in white matter bulk	Occasional prominent mitochondria	1.7	1.0	0.004
S70	M	Profound hypotonia with developmental delay, visual impairment, myopathic EMG	Lack of white matter bulk, small optic chiasm	Non specific features	1.0	1.0	0.012
S71	F	Cataracts, developmental delay, hypertrophic cardiomyopathy, 3-methylglutaconic aciduria	NA	Excess lipid	1.6	NA	0.012
S72	M	Vomiting, poor growth, developmental delay, infantile spasms, sensorineural deafness, pigmentary retinopathy	Bilateral extensive polymicrogyria	Patchy COX staining	5.9	NA	0.010
S73	M	Pancytopenia, seizures, hypotonia, immunodeficiency	Non-specific signal abnormality of deep white matter	NA	1.9	NA	0.006
S74	F	Infantile spasms, developmental delay	Lesions in thalami brainstem and dentate nuclei	Excess lipid myopathic features	3.8	2.7	0.005

CSF: cerebrospinal fluid, EMG: electromyogram, F Female, M Male, MRI magnetic resonance imaging, NA not available

2.2 Genetic methods

2.2.1 Polymerase chain reaction

Patient genomic DNA (gDNA) extracted from blood was obtained from the Regional Molecular Genetics diagnostic laboratory at Great Ormond Street Hospital. The DNA concentration and absorbance maximum was measured using a NanoDrop ND-1000 spectrophotometer (NanoDrop). A 2 µL aliquot of gDNA sample was loaded on to the pedestal of the machine. Nucleotides have an absorbance maximum at 260 nm, and if there is protein contamination the sample will have an absorbance maximum at 280 nm. Therefore the absorbance ratio 260:280 nm is used to assess the purity of the gDNA sample. An absorbance ratio (260:280) of approximately 1.8-2.0 is considered an acceptable level of purity.

Primers were designed using Primer3 software (http://frodo.wi.mit.edu/) to cover all exons and their flanking splice sites. The specificity of each primer was cross-checked using Primer blast software (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). Target regions were amplified by PCR using 12.5 µL of the Promega colourless master mix (50 units/mL of Taq DNA polymerase supplied in a proprietary reaction buffer (pH 8.5), 400µM dATP, 400µM dGTP, 400µM dCTP, 400µM dTTP, 3mM $MgCl_2$), 1 µL each of the forward and reverse primer (15 pmol/µL), 1 µL of gDNA (50-100 ng/µL) and 10.5 µL of nuclease free water. PCR plates were run on a Thermocycler and amplification was performed by denaturation at 95°C for 2 minutes, 35 cycles of 94°C for 30 seconds, annealing at 57°C for 30 seconds, 72°C for 40 seconds, and a final extension step at 72°C for 5 minutes. The presence of the correct length of PCR products was verified by electrophoresis of PCR products on a 1% agarose gel containing ethidium bromide and comparison with a 1000 bp ladder (Hyper ladder, Bioline). The DNA bands were visualised using a UV transilluminator. PCR products were then purified with MicroClean PCR clean-up kit (Microzone Ltd) by adding 5 µL of MicroClean to 5 µL of the PCR product.

In some instances PCR reactions were cleaned using ExoSAP-IT (Affymetrix) buffer by incubating at 37°C for 30 minutes, and then 80°C for 15 minutes to inactivate the enzymes and sequencing was outsourced to Source Biosciences, Cambridge UK.

2.2.2 Automated Sanger sequencing

Bi-directional sequencing was performed using a DNA sequencing kit (BigDye Terminator v1.1 Cycle Sequencing Kit (BDT -Applied Biosystems). After Microclean was added, the plate was vortexed and centrifuged at 4000 g for 30-40 minutes. The plate was then spun slowly for 2 minutes upside down briefly to collect the residual reagents. Afterwards 9 µL of sequencing mix (BDT 0.7 µL, nuclease free water 5.3 µL, 5X buffer 2 µL, betaine 1 µL) and the corresponding forward or reverse primer (1 µL of 1.5 pmol/µL) was added to each well. The plates were sealed, spun briefly and sequenced on the Thermocycler overnight. Cycling conditions were:

1) 96°C for 90 seconds

2) 96°C for 30 seconds (denaturation)
3) 57°C for 30 seconds (annealing)
4) 60°C for 4 minutes (extension)
5) Repeat step 2 to 4, 35 times
6) 15°C for 5 hours

The following day, 50 µL of 2 mol/l sodium acetate in 100% ethanol was added to each well, and the plate was vortexed and centrifuged at 4000 g for 40 minutes. The contents of the plate were then emptied by turning the plate upside down and centrifuging at 600 g briefly. Next 150 µL of 70% ethanol was added and the plate was centrifuged at 4000 g for 15 minutes. The contents were then emptied by centrifuging the plate upside down briefly. The plate was allowed to air dry at room temperature for 30-45 minutes after which 10 µL of 10 mM Tris, 1 mM EDTA (pH 7.4) (TE) was added to each well and the plate was loaded onto ABI3730 capillary machines (Applied Biosystems).

Sequencing data were analysed using Sequencher version 4.8 (Gene Codes) and the sequences were compared with reference sequences obtained from the Ensembl database (http://www.ensembl.org/Homo_sapiens/Info/Index).

2.2.3 Next generation sequencing: targeted exon sequencing

Identification of COX genes

All genes encoding human COX subunits and assembly factors were identified by performing literature searches for previous publications. Putative human homologues of yeast COX genes were identified by using BLAST (http://blast.ncbi.nlm.nih.gov/) and by previous publications using PubMed searches.

Bait design for capture of target regions

Baits are RNA oligonucleotides designed to match and bind regions of interest of gDNA. They are biotinylated to capture on to streptavidin-coated magnetic beads. The bait design protocol used is outlined below.

The list of genes of interest with their RefSeq transcript identifiers (e.g: NM_00001) was compiled in a Microsoft excel file. This list was uploaded to the UCSC genome browser http://genome.ucsc.edu/ (under tables tab) with the parameters set to:
clade: Mammal
genome: Human
assembly: Feb. 2009 (GCRCh37/hg19)
group: Genes and Gene Prediction Tracks
track: UCSC Genes
table: knownGene
region: genome

output format: BED-browser extensible data
file type returned: plain text
The output format for this was set to exons plus 50 bases at each end. The next step was to merge the co-ordinates by using the Galaxy software, http://main.g2.bx.psu.edu/. The file generated from UCSC was uploaded here and the overlapping regions of interest were merged. For example Chr3, 1 to 500 and Chr3, 400 to 600 were merged as Chr3, 1 to 600. Next the size of the entire target of interest was calculated by summation of the individual targets; this was useful to calculate coverage and to check how many samples can be run on the NGS platform. The format of the file was changed so that it was suitable for uploading onto eArray (from chr6 32793137 32796861 to chr6: 32793137-32796861) and this was copied onto a Notepad file (txt file). (https://earray.chem.agilent.com/earray/). eArray is a web-based application to create oligonucleotide libraries. On eArray the application type is set to SureSelect Target Enrichment with the following parameters:
Sequencing protocol: Single-End
Use optimised parameters (tick box)
Species: H.sapiens
Avoid standard repeat regions (tick box)
The text file was then uploaded for the earray software to design the baits. Once this process was complete, the files (with the designed baits for the target region) were downloaded and the coverage of the baits was checked. Coverage details were included in the files from earray which gives the length of target, number of baits per target and percentage of target region covered by baits. In addition the coverage was also assessed manually by using the IGV tool (http://www.broadinstitute.org/software/igv/). The bait tiling bed file from earray was uploaded to IGV. The baits with the RefSeq genes were visualised. Some GC rich regions were not covered adequately and the 'boosting' function on the earray software was used to increase the coverage.

SureSelectXT Target Enrichment System for Illumina Paired-End Sequencing Library Illumina HiSeq and MiSeq Multiplexed Sequencing Platforms Protocol Version 1.3.1, February 2012

DNA content and quality

DNA quality was assessed using a NanoDrop Spectrophotometer as described before, to ensure the optical density at 260/280 nm was 1.8-2.0. DNA quantity was measured using the Qubit system (broad range dsDNA assay) according to the manufacturer's protocol (ThermoFisher Scientific). A total of 3 µg of DNA per sample was diluted in low TE buffer to a total volume of 130 µL in a 1.5 mL LoBind tube. The workflow diagram of the SureSelect protocol is shown in Figure 2.1.

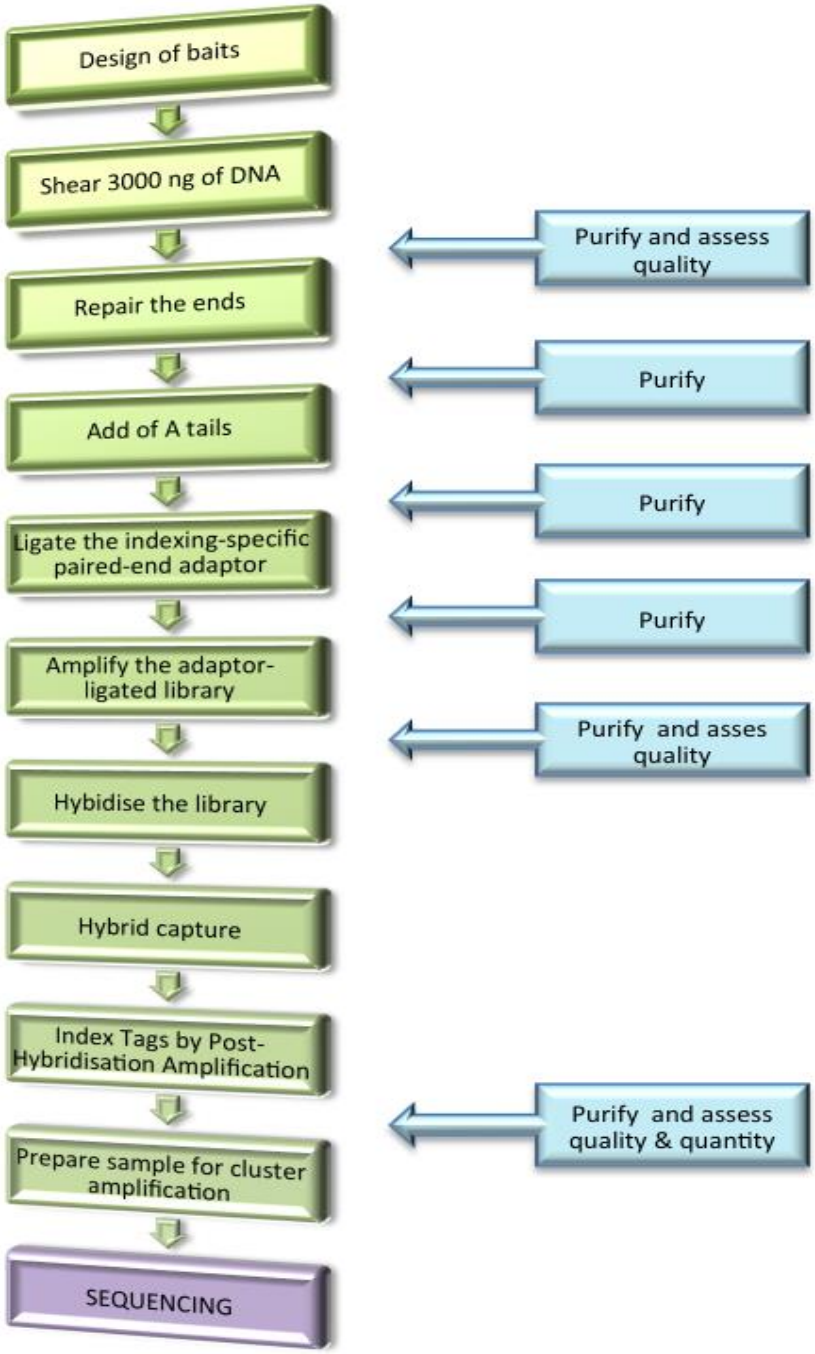


Figure 2:1: Workflow diagram of the SureSelectXT Target Enrichment System for Illumina Paired-End Sequencing Library Illumina HiSeq and MiSeq Multiplexed Sequencing Platforms

(Protocol Version 1.3.1, February 2012)

Shearing of DNA

The shearing was performed using Covaris focused-ultrasonicator. The water was filled up to level 12 on the tank and the temperature set to between 2°C to 5°C. The water in the instrument was degassed for at least 30 minutes before use since insufficient degas levels within the water bath may result in poor acoustic coupling. The 130 µL of DNA sample was slowly transferred to a Covaris microTube through the pre-split septa on the cap ensuring no bubbles were present. The Covaris microtube was loaded onto the tube holder and the Covaris was set to the following protocol:

Duty Cycle: 10%

Intensity: 5

Cycles per Burst: 200

Time: 6 cycles of 60 seconds each

Set Mode: Frequency sweeping

Temperature: 4° to 7° C

Afterwards the sheared DNA was carefully transferred into a new 1.5 mL Lo binding tube.

Purification of sheared DNA

AMPure XP beads were allowed to warm to room temperature for at least 30 minutes, mixed thoroughly using a vortex until the mixture looked homogeneous and the beads (180 µL) were added into each Lo-binding tube containing the sheared DNA. The sample was mixed and incubated at room temperature for 5 minutes after which the tubes were placed on a magnetic stand for 3-5 minutes. The magnetic AMPure XP beads are attracted to the magnet and the solution clears. The cleared solution was carefully discarded and 500 µL of freshly prepared 70% ethanol was added, allowed to sit for one minute and the ethanol was then discarded. The addition and removal of 70% ethanol was repeated twice. The samples were then air dried for 5-10 minutes until all the ethanol completely evaporated. Nuclease free water (50 µL per sample) was added; the beads were mixed on a vortex and allowed to incubate for 2 minutes at room temperature. The tube was put back on the magnetic stand and once the solution cleared, the cleared solution (50 µL of purified, sheared DNA) was collected into a new Lo-binding tube.

Assessing quality using the 2100 Bioanalyzer

The samples were loaded on to a Bioanalyzer DNA 1000 chip (Agilent) according to the manufacturer's protocol and the electropherogram was viewed to ensure that the distribution had peak heights between 150 and 200 bp to assess adequate shearing of DNA.

Repairing the ends

The samples were transferred to PCR tubes. The following end repair master mix (SureSelect Library Prep Kit, ILM) was prepared on ice and added to 48 µL of each sample.

Nuclease-free water	35.2 µL
10X End Repair Buffer	10 µL
dNTP Mix	1.6 µL

T4 DNA Polymerase	1 µL
Klenow DNA Polymerase	2 µL
T4 Polynucleotide Kinase	2.2 µL

The samples were then incubated on a thermal cycle at 20°C for 30 minutes.

Purification of the sample

Purification of the sample was carried out as described previously except that the volume of Ampure beads added was 90 µL rather than 180 µL. The beads were washed with 70% ethanol as described before. Also in the final step 32 µL of nuclease free water was added, incubated for 2 minutes and put on a magnetic stand. The cleared solution (~30 µL) was collected into a new Lo-binding tube.

Addition of 'A-tails' to the 3' end

The samples were transferred to PCR tubes. The following master mix (SureSelect Library Prep Kit, ILM) was prepared on ice and added to 30 µL of each sample.

Nuclease-free water	11 µL
10X Klenow Polymerase Buffer	5 µL
dATP	1 µL
Exo(-) Klenow	3 µL

The samples were the placed on a thermal cycler for 30 minutes at 37°C.

Purification of the sample

Purification of the sample was carried out as described previously except that the volume of Ampure beads added was 90 µL rather than 180 µL. Also in the final step 15 µL of nuclease free water was added, incubated for 2 minutes and put on a magnetic stand. The cleared solution was collected into a new Lo-binding tube.

Ligation of the indexing-specific paired-end adaptor

The following mix was (SureSelect Library Prep Kit, ILM) added to 13 µL of sample and left to incubate for 15 minutes on a thermal cycler at 20°C.

Nuclease-free Water	15.5 µL
5X T4 DNA Ligase Buffer	10 µL
SureSelect Adaptor Oligo Mix	10 µL
T4 DNA Ligase	1.5 µL

Purification of the sample

Purification of the sample was carried out as described previously except that the volume of Ampure beads added was 90 µL rather than 180 µL. Also in the final step 32 µL of nuclease free water was added, incubated for 2 minutes and put on a magnetic stand. The cleared solution was collected into a new Lo-binding tube.

Amplification of the adaptor-ligated library

The following PCR mix (SureSelect Library Prep Kit, ILM, SureSelect Target Enrichment Kit ILM Indexing Hyb Module Box #2, Herculase II Fusion DNA Polymerase - Agilent) was added to 15 µL of sample.

Nuclease-free water	21 µL
SureSelect Primer	1.25 µL
SureSelect ILM Indexing Pre Capture PCR	1.25 µL
5X Herculase II Rxn Buffer	10 µL
100 mM dNTP Mix	0.5 µL
Herculase II Fusion DNA Polymerase	1 µL

The samples were loaded on to a thermal cycler with the following conditions:

1) 98°C 2 minutes
2) 98°C 30 seconds
3) 65°C 30 seconds
4) 72°C 1 minute
5) Repeat Step 2 through Step 4 for a total of 4 to 6 times.
6) 72°C 10 minutes
7) 4°C Hold

Purification of the sample

Purification of the sample was carried out as described previously except that the volume of Ampure beads added was 90 µL rather than 180 µL. Also in the final step 30 µL of nuclease free water was added, incubated for 2 minutes and put on a magnetic stand. The cleared solution was collected into a new Lo-binding tube.

Assessing quality using the 2100 Bioanalyzer

The samples were loaded on to a Bioanalyzer DNA 1000 chip according to the manufacturers protocol and the electropherogram was viewed to ensure that the distribution had peak heights between 250 and 275 bp to ensure that the adaptors were ligated. The concentration for each sample was measured by estimating the area under the peak.

Hybridisation reaction

The hybridisation reaction requires 500 ng of DNA in a maximum volume 3.4 µL (147 ng/µL). A vacuum concentrator (speedvac) was used to dehydrate the samples for about 30 minutes. Caps of the LoBind tubes were pierced with a narrow gauge needle to allow evaporation. The samples were then inserted in the vacuum concentrator set at 45°C. Since the initial concentrations of the samples were measured previously, the dehydrated samples were then reconstituted with the appropriate volume of nuclease-free water to bring the final concentration to 147 ng/µL.

The hybridisation buffer was prepared using SureSelect Hyb buffer #1, #2, #3, #4 in volume ratios of 25:1:10:13 respectively and warmed at 37°C.

The SureSelect Capture Library mix was prepared by mixing 2 µL of SureSelect bait Library (bait library) and 0.5 µL of RNAse block and 4.5 of nuclease free water.

The SureSelect Block Mix was prepared by adding the following reagents together on ice: SureSelect Indexing Block #1 2.5 µL, SureSelect Block #2 2.5 µL and SureSelect Indexing Block #3 0.6 µL per reaction.

In a PCR tube 3.4 µL of each 147 ng/µL sample was added followed by 5.6 µL of the SureSelect block mix. The thermal cycler was set at 95°C for 5 minutes and maintained at 65°C. This step was used to denature the libraries and anneal the oligonucleotides.

At room temperature 13 µL of the hybridisation buffer was added to the capture bait library mix (7 µL). This mix (20 µL) was added to the denatured libraries on the PCR machine after they had been at 65°C for 5 minutes. The PCR tubes were sealed and the hybridisation mixture was incubated for 24 hours at 65°C with a heated lid of 105 °C.

Preparation of the magnetic beads (Dynabeads MyOne Streptavidin T1)

The following day Sureselect wash 2 was heated at 65°C in a water bath. Dynabeads MyOne Streptavidin T1 were vigorously resuspended by using a vortex. Next 50 µL of the Dynabeads MyOne Streptavidin T1 was added to separate 1.5 mL microfuge tubes ensuring there were enough microfuge tubes for each hybridisation reaction. Next 200 µL of SureSelect Binding Buffer was added to each microfuge tube. The mixture was mixed on a vortex for 5 seconds and then put on the Dynal magnetic separator for 3-5 minutes until the solution cleared. The beads are attracted towards the magnet and the solution clears. The resulting clear supernatant was then carefully removed and discarded. The above addition of 200 µL of SureSelect Binding Buffer and removal of the supernatant was repeated twice.

Hybrid capture

After the 24-hour hybridisation the volume of hybridisation mixture remaining for each reaction was estimated using a pipette. While the reactions remained at 65°C on the thermal cycler, the hybridisation mixture was added to the beads prepared above. The hybrid-capture/bead solution was mixed for 30 minutes on a BD Clay Adams Nutator Mixer at room temperature. The tubes were briefly centrifuged and put on the Dynal magnetic separator for 5 minutes. The cleared solution was removed and discarded. The beads were resuspended in 500 µL of SureSelect Wash 1 by mixing on a vortex. The samples were left to incubate at room temperature for 15 minutes, mixing occasionally. The tubes were briefly centrifuged and put on magnetic separator and left for 5 minutes. The resulting clear solution was removed.

The beads were resuspended in 500 µL of prewarmed (65°C) SureSelect Wash 2 and mixing on a vortex. The samples were left to incubate for 10 minutes at 65°C and then the beads and buffer were separated using the magnetic separator. The supernatant was removed. The addition, incubation and removal of 500 µL of 65°C pre-warmed SureSelect Wash 2 was repeated twice. Next 30 µL of nuclease free water was added to each tube and mixed on a vortex mixer. The samples were incubated for 2 minutes at room temperature, occasionally mixing using a vortex mixer. The tubes were briefly spun on a centrifuge and the bead and buffer were separated using a magnetic separator for 5 minutes. The supernatant contained the captured DNA and this was transferred to a new microfuge tube.

Addition of Index Tags by Post-Hybridisation Amplification

The addition of a unique index tag to each sample allows the identification of the individual samples when they are pooled together. The following PCR mix (volumes are per sample) was prepared on ice and 14 µL captured DNA was added to the mix in PCR tubes.

Nuclease-free water	22.5 µL
5X Herculase II Rxn Buffer	10 µL
100 mM dNTP Mix	0.5 µL
Herculase II Fusion DNA Polymerase	1 µL
SureSelect ILM Indexing Post Capture Forward PCR Primer	1 µL
PCR Primer Index 1 - Index 16 (sample specific)	1 µL

(A separate index is added to each sample; e.g: index 1 to sample 1, index 2 to sample 2 etc.)

The following programme was run on a thermal cycler.

1) 98°C 2 minutes
2) 98°C 30 seconds
3) 57°C 30 seconds
4) 72°C 1 minute
5) Repeat step 2 through step 4 for a total of 15 times.
6) 72°C 10 minutes
7) 4°C Hold

Purification of the sample

AMPure XP beads were allowed to come to room temperature for at least 30 minutes and mixed on a vortex until homogeneous in appearance. To a 1.5 mL LoBind tube, 90 µL of homogenous AMPure XP beads was added together with the post-hybridised amplified library (50 µL). The samples were mixed on a vortex mixer and left to incubate for 5 minutes. The tubes were then put on a magnetic separator for 5 minutes and the supernatant was carefully removed. While the tubes were on the stand 500 µL of 70% ethanol was added, allowed to sit for 1 minute and then removed. The addition and removal of ethanol was repeated twice. The samples were air dried for 5-10 minutes ensuring all residual ethanol was removed while not allowing the bead pellet to dry excessively and crack. Next 30 µL of nuclease-free water was

added to each sample, mixed on a vortex mixer and left to incubate for 2 minutes at room temperature. The tubes were then put on a magnetic separator for 5 minutes and the resulting supernatant (30 µL) was transferred to a new Lo-binding tube. This contains the purified library captured by baits.

Assessing quality with the 2100 Bioanalyzer

The samples were loaded on to a Bioanalyzer DNA high sensitivity chip according to the manufacturer's protocol. The electropherogram was viewed to assess quality and quantity after PCR, and to assess the distribution with peak heights between 300 and 400 bp.

Assess the quantity of each index-tagged library

The quantity of the index tagged library concentration was determined using the Qubit high sensitivity assay (ng/µL) according to the manufacturers protocol. The mean length of the library was calculated in bp for each sample using the electropherogram generated by the 2100 Bioanalyzer. The library concentration (ng/µL) and mean length derived above were used to calculate the library concentration in nmol/l using a pre-programmed excel spreadsheet. The samples were diluted in TE to give a final concentration of 2 nmol/L. The pooled library was then re-quantified using the Qubit HS assay ensuring the pool was 2 nmol/L. The concentration of the pools was adjusted to 2 nmol/L if necessary.

Preparation of samples for cluster amplification

The MiSeq cartridge was taken out of the freezer and left to thaw for 1 hour in a deionised water bath. The following reagents were prepared: Hybridisation buffer (HT1), HP3 (2 N NaOH) and PhiX control. HT1 buffer was thawed on ice.

The 10 M NaOH was diluted 1:5 to 2 M NaOH. Next 2 µL of the 2 nmol/l multiplexed sample pool was added into 10 µL of the 2 M NaOH and left to incubate at room temperature for 5 minutes to denature the DNA. After 5 minutes 980 µL of pre-chilled HT1 buffer was added and the mix was kept on ice. For an 8 pmol final concentration 400 µL of the 20 pmol denatured library was mixed with 600 µL of the HT1 buffer.

The PhiX control was prepared by adding 2 µL of the 10 nmol PhiX library to 8 µL of the 10mmolTris-Cl, pH 8.5 and 0.1% Tween 2 and 10 µL of 0.2 M NaOH. This was left to incubate at room temperature for 5 minutes. Then HT1 buffer (980 µL) was added to give 20 pmol PhiX library. To dilute to an 8 pmol solution 400 µL was further diluted using 600 µL of HT1.

The prepared PhiX control (10 µL) was added to 990 µL of the denatured sample libraries.

The thawed MiSeq cartridge was retrieved, the load samples reservoir was pierced and 600 µL of the final solution prepared above was pipetted in. The cartridge was loaded on the MiSeq analyser and run according to manufacturer's instructions.

2.2.4 Next generation sequencing: whole exome sequencing

Whole exome sequencing was outsourced to different providers, which included Perkin Elmer (Agilent Sureselect), Oxford Gene Technology (Agilent Sureselect), UCL Institute of Neurology (Nextera capture kit) and Beijing Genomics Institute (Agilent Sureselect).

The Nextera method was used by Dr Deborah Hughes at UCL Institute of Neurology: library construction using Nextera chemistry (Illumina Inc., San Diego, CA) was used to prepare dual indexed, paired-end libraries from 50 ng of genomic DNA that were subsequently enriched for the exome (62 megabases) according to standard protocols supplied by Illumina (Illumina Inc., San Diego, CA). The enriched DNA libraries were sequenced, 2x101 cycles on a HiSeq2500 sequencing system (Illumina, Inc., San Diego, CA).

The other samples were prepared according to Agilent's SureSelect Protocol. The concentration of each library was determined using Agilent's QPCR NGS Library Quantification Kit (G4880A). Samples were pooled prior to sequencing with each sample at a final concentration of 10nM. Read files (Fastq) were generated from the sequencing platform via runs to an average 20 x coverage on a HiSeq 2000 system (Illumina).

Alignment, variant calling and annotation: This analysis was carried out by Dr Vincent Plagnol, UCL Genetics Institute. The sequencing data was mapped to the human reference assembly, hg19 (GRCh37; downloaded from the UCSC genome browser) by Novoalign Software. After removal of PCR duplicates (Picard) and reads without a unique mapping location, variants were extracted using the Maq model in SAMtools and filtered by the following criteria: consensus quality >30, SNP quality >30 and root mean square mapping quality >30. Calling was performed using the haplotype caller module of GATK (https://www.broadinstitute.org/gatk), creating gVCF formatted files for each sample. The individual gVCF files were combined into merged gVCF files for each chromosome containing on average 100 samples each. The final variant calling was performed using the GATK Genotype gVCFs module jointly for all samples (cases and controls). Variants quality scores were then re-calibrated according to GATK best practices separately for indels and SNPs. These variants were further filtered against dbSNP build 135 (UCSC genome browser), 1000genomes by use of Annovar.

2.2.5 Homozygosity mapping

Homozygosity mapping was performed entirely by Kerra Pearce of UCL Genomics. Patients born to consanguineous families are likely to carry two identical recessive copies of a given allele from a common ancestor. These alleles are referred to as identical-by-descent alleles. Small regions of chromosomes tend to be transmitted as whole sections and individuals from consanguineous families. Single nucleotide polymorphism markers near these homozygous regions will display homozygosity at these markers. In first cousin parents 1/16 of their

offspring's genome will be homozygous. The principle of this method, first suggested in 1987 (Lander & Botstein 1987) is to search for regions of homozygosity that are shared by the affected individuals. The method is described below.

Illumina Human DNA Analysis Beadchip Procedure

All processing was carried out in accordance with the Infinium HD Ultra Assay protocol (Rev B, 2010, Illumina Inc, San Diego, USA). High quality genomic DNA (300 ng) was whole genome amplified at 37°C for 20-24 hours in a deep well plate, and then fragmented (37ºC for 1 hour and 15 minutes in hybridisation oven), precipitated and resuspended in hybridisation buffer. Samples were denatured and then taken from the plate and loaded onto the chips using a liquid handling robot (Freedom Evo, Tecan Ltd, Switzerland). Hybridisation took place overnight (16-20 hours) at 48°C. Single base extension and staining was carried out by the liquid handling robot. The chips contain probes which were extended by a single hapten-labelled dideoxynucleotide (ddNTP) base complementary to the hybridised DNA. The ddATP and ddTTP bases were labelled with 2,4-Dinitrophenol (DNP) whereas the ddCTP and ddGTP were labelled with Biotin. The DNA sample was then removed from the chip using formamide. The staining procedure involved signal amplification by multi-layer immunohistochemical staining. The haptens were detected simultaneously by Streptavidin and an anti-DNP primary antibody conjugated to green and red fluorophores respectively (STM reagent, Illumina). In order to amplify the fluorescent signals the chips were then counterstained with biotinylated anti-steptavidin and a DNP-labelled secondary antibody to the anti-DNP primary antibody (ATM reagent, Illumina). The final layer of stain is the STM reagent which contained fluorophores thus allowing signal detection. Finally the stained chips were coated in a glue-like substance in order to protect the dyes, and the chips scanned using the iScan scanner with autoloader (Illumina Inc, San Diego, USA).

The log R ratio and B allele frequency were studied and these two parameters help visualisation of homozygous areas. The log R ratio estimates copy number of each SNP and the B allele frequency estimates the proportion of times an allele is called B at each SNP locus: Therefore the expected values would be 1.0 (B/B), 0.5 (A/B) and 0.0 (A/A). The AA, AB and BB genotyping were exported as excel files so that shared regions could be analysed.

2.2.6 Real time quantitative PCR to quantitate mtDNA

DNA from mouse tissue was extracted using the Gentra Puregene Tissue Kit (4 g) using RNase A Solution, Puregene Proteinase K, and Reagents (Qiagen).

Mouse tissues were retrieved from the -80˚C freezer and were cut using a scalpel into small (~5mg) pieces. Each piece of tissue was weighed. The lysis mix was prepared by adding 600 µL of cell lysis solution per sample to 15 µL per sample of Puregene Proteinase K. The lysis mix was added to each tissue sample and the contents of the tubes were mixed by inverting several times. The samples were then placed on a rotator and incubated at 55˚C for 3 hours so that the

tissues appeared completely dissolved. Subsequently 3 µL of RNAse A solution (to remove RNA) was added to each sample and the samples were incubated for 60 minutes at room temperature. To remove the protein, samples were then incubated for 3 minutes on ice and 200 µL of protein precipitation solution was added to each sample. The samples were mixed using a vortex for 20 seconds and then kept on ice. The samples were then placed on a centrifuge at 4˚C for 5 minutes and centrifuged at high speed (13,000-16,000 g) where after the supernatant was transferred to new tubes. The supernatant was centrifuged again at high speed for 5 minutes and the supernatant was transferred to new tubes.

The volume of supernatant was calculated in each tube and an equal volume of isopropanol was added to each tube, which leads to precipitation of DNA. The samples were centrifuged at room temperature (13,000-16,000 g) for 3 minutes. The supernatant was removed carefully without disturbing the DNA pellet. The DNA pellet was washed with 600 µL of 70% ethanol, centrifuged at (13,000-16,000 g) for 3 minutes and the supernatant was again carefully removed taking care not to disturb the DNA pellet. The pellet was air-dried, 100 µL of DNA hydration solution was added and the samples were incubated overnight on a shaker. The following day the NanoDrop was used to quantify the amount of DNA in each sample. Each sample was diluted down to ~4 ng/µL.

Separate forward and reverse primers (concentration 20 µm) for mtDNAs and β2 microglobulin (B2M) were used.
mouse_mtDNA _Dloop_forward GACCAACATAACTGTGGTGTCA
mouse_mtDNA_Dloop_reverse ATTCTTCTCCGTAGGTGCGTCTAG
mouse_B2M_forward TGTCAGATATGTCCTTCAGCAAGG
mouse_B2M_reverse TGCTTAACTCTGCAGGCGTATG

Real time quantitative PCR (RT-qPCR)
Quantitative real-time polymerase chain reaction (RT-qPCR) is a technique to assess copy numbers of a targeted DNA sequence.

In real-time quantitation, the PCR product are considered by the point in time when amplification of a target achieves a fixed level of fluorescence rather than the amount of product accumulated after a fixed number of cycles. The PCR amplification is monitored in real time by measuring the fluoresence emitted when the SYBR Green reagent binds to the double stranded DNA. When a fixed level of fluorescence is achieved, the rate correlates with the amount of original DNA and the amount of DNA is further normalised against reference genes to allow comparison.

The qPCR reaction was run using the StepOne™ Real-Time PCR System (Applied Biosystems) and the *Power* SYBR® Green PCR Master Mix (Applied Biosystems). The SYBR Green chemistry uses a highly specific, double-stranded DNA binding dye to detect the PCR product as it accumulates during the PCR cycle. The qPCR reaction was set up as detailed in below

and each sample was analysed in quadruplicate in MicroAmp Fast Optical 96-well plates (Life Technologies). The RT-PCR cycle was 95˚C for 3 min, 40 cycles of 95˚C for 30 seconds, 60˚C for seconds, 72˚C for 30 seconds. Melting curve analysis was carried out at a temperature range from 60˚C to 95˚C, in 1˚C increments.

Reagent	µL per well
Power SYBR® Green PCR Master mix	10
Deionised H_2O	7.4
Forward primer (20 µm)	0.3
Reverse Primer (20 µm)	0.3
Sample	2
Total	20

The comparative CT method was used. In this method, the amplification of the target (mtDNA) or the endogenous control (β2 microglobulin) was measured in the samples. One control sample was defined as the reference sample. The measurements of mitochondrial DNA amplification were then normalised against the endogenous control, β2 microglobulin amplification. The relative quantity in each sample is obtained by comparison of the normalised target quantity in each sample to normalised target quantity in the reference sample. Wells that contain water instead of sample template were used to ensure no amplification of the target occurs in these negative control wells.

Data analysis

Using StepOne Software v2.2 (Applied Biosciences) flagged samples, which were outliers among the replicates, were excluded from analysis and the ΔCT values were obtained for each sample. The mean ΔCT values were calculated for the control mice. The ΔΔ CT value was obtained by subtracting the mean ΔCT values of the wild-type from the ΔCT of each (mutant mouse) sample. The $2^{\text{delta-delta CT}}$ values and the $1- 2^{\text{delta-delta CT}}$ values were calculated for the individual samples. The mean $1- 2^{\text{delta-delta CT}}$ value was calculated separately for the mutant and wild type mice.

2.2.7 Droplet digital PCR to quantitate mtDNA

The QX200 Droplet digital PCR (ddPCR) method uses microfluidics and surfactant chemistries to partition the PCR samples into ~20,000 water-in-oil droplets where PCR amplification occurs within the droplets. After PCR, each droplet is analysed in a flow cytometer and the proportion of PCR-positive and PCR-negative droplets in the original sample are calculated. Software using Poisson statistics are used to determine the target DNA template concentration in the original sample.

The concentrations of DNA of the samples for ddPCR were estimated using Qubit Fluorometer which measures a sample concentration from 100 pg/µL–1,000 ng/µL. It is highly selective for double-stranded DNA over RNA. The Qubit-iT™ High-Sensitivity dsDNA Assay Kit (ThermoFisher Scientific) was used. The working reagent was prepared making a 1:200 dilution of the Qubit dsDNA HS Reagent using the Qubit dsDNA HS Buffer. A 1:10 dilution of the gDNA was prepared. The assay tubes were prepared by mixing 190 µL of the working solution with 10 µL of the standards (Qubit dsDNA BR Standard #1 and #2) or 10 µL of the diluted DNA sample. The assay tubes were incubated at room temperature for 2 minutes. Standards were used to calibrate the Qubit® Fluorometer after which samples were measured individually.

The PCR reaction mixture (final total volume of 20 µL) consisted of 2× ddPCR Supermix (Bio-Rad), 20X primer, and FAM and VIC probes (final concentrations of 900 and 250 nM, respectively) and template (1ng/µL, variable volume). The sequences are indicated below,

Primer sequence (5′→3′)	**Probe sequence (5′→3′)**
Forward: CATCTGGTTCCTACTTCAGGG	6FAMCTTAAATAAGACATCACGATGGATC
Reverse: TGAGTGGTTAATAGGGTGATAGA	AC-TAMRA (np16549–6)
Forward: TGCTGTCTCCATGTTTGATGTATCT	VIC-TTGCTCCACAGGTAGCTC TAG GAG
Reverse:TCTCTGCTCCCCACCTCTAAGT	G-TAMRA (621–645)

Each 20 µL reaction was pipetted into a separate well of an eight-well droplet generator cartridge (Bio-Rad) followed by addition of 70 µL of droplet generation oil into each of the corresponding oil wells. The cartridge was then loaded into the droplet generator (QX200, Bio-Rad). A vacuum was applied to each of the outlet wells to generate droplets separating the mixture into an oil emulsion mix. Subsequently 40 µL of the oil emulsion mix of each sample was transferred using multichannel pipette into separate wells in a 96- well polypropylene plate (Eppendorf), sealed with aluminium sealer and amplified in a thermal cycler (C1000 Touch Thermocycler-Bio rad) following a 10 minute enzyme activation period at 95°C, 40 cycles of 30 second denaturation at 94°C and 40 cycles of 1 minute annealing/extension at 60°C and a final 10 minute inactivation step at 98°C.

Afterwards, the plates were transferred to QX200 droplet reader (Bio-Rad) containing droplet reader oil. The droplet reader aspirates the droplets and sends them in a single file past a fluorescence detector that is able to detect the fluorescent positive droplets and the fluorescent negative droplets. Quanta Soft (Bio-Rad), software package was used for data analysis. Each droplet from a sample was plotted on a graph depicting fluorescence intensity vs. droplet number.

2.3 Cell culture

Fibroblast cells were derived from patient skin biopsies and cultured in Dulbecco's Modified Eagle's Medium (DMEM) containing 4.5 g/L glucose, Glutamax (Life Technologies) and 110

mg/l sodium pyruvate (Gibco), supplemented with 10% heat inactivated Fetal Bovine Serum (FBS, Sigma), 50 µg/mL streptomycin and 50 u/mL penicillin (Gibco). Additionally, uridine was added (Sigma) to a final concentration of 50 mg/L because patient cells with respiratory chain enzyme deficiencies may be deficient in uridine (because Coenzyme Q_{10} is required for pyrimidine biosynthesis) and therefore cellular nucleic acid synthesis. Cells were cultured in 182 cm^2 flasks (Santa Cruz) containing 18 mL of supplemented DMEM or 10 cm plates containing 10 mL of supplemented DMEM. Flasks or plates were incubated at 37°C with 5% CO_2. Culture medium was changed every 3-4 days and when 70-80% confluency was reached, cells were detached after washing with 8 mL of Dulbecco's Phosphate Buffered Saline, (PBS) (Sigma) by incubating for 3 minutes with 1 mL of 0.25% trypsin (Sigma). For subculture, the trypsinised cells from one flask or plate were resuspended in 8 mL of culture medium to inactivate the trypsin digestion process and an equal volume of the cell suspension was transferred into two new flasks or plates.

When cell pellets were harvested for further experiments, the cells re-suspended in 8 mL of medium were transferred to 50 mL falcon tubes and centrifuged at 200 g for 10 minutes at room temperature. The supernatant was discarded and the pellet was resuspended in 10 mL of PBS and centrifuged again at 200 g for 10 minutes. The resulting supernatant was discarded and the cell pellet was again resuspended in 1 mL of PBS and transferred to a 1.5 mL reaction tube on ice. This tube was used for further experiments including Western blot and citrate synthase activity.

Cryostocks of cells were prepared for long-term storage in liquid nitrogen. DMEM was aspirated from a fully confluent 182 cm^2 flask or 10 cm plate and the cells were washed using 8 mL of PBS. The cells were detached by trypsinisation as described before. When the cells were completely detached 8 mL of FBS supplemented DMEM was added to the flask and the total volume was transferred into a 50 mL falcon tube. The falcon tube was then centrifuged for 10 minutes at 200 g. The supernatant was carefully removed and the cells were resuspended in 4 mL of DMEM and 100 µL of Dimethyl sulfoxide (DMSO Sigma). The 4 mL suspension of cells was split equally between 4 cryotubes, placed in a cryocontainer and kept at -80˚C for 24 hours. Subsequently the cells were transferred to liquid nitrogen for long-term storage.

2.4 Western blot

2.4.1 Preparation of whole cell lysates

Whole cell lysates were prepared from ~1 million cultured fibroblasts. All procedures were carried on ice or in a centrifuge cooled to 4˚C. The tube with the cells pellet resuspended in 1 mL of PBS was used for further experiments. The cells were pelleted by micro-centrifugation for 5 min, at 6,500 g, 4°C. The supernatant was removed carefully, without disturbing the pellet. Approximately 100-200 µL of 1% Triton X-100 in PBS with added protease inhibitors (2 mM Phenylmethanesulfonyl fluoride –PMSF, 2 µg/mL pepstatin A and 2 µg/mL leupeptin) was

added to each pellet. The 1% Triton X-100, PBS with three protease inhibitors mix was prepared using the following ratios.

- 800 µL water
- 100 µL 10x PBS
- 100 µL 10% Triton X-100 in water
- 2 µL leupeptin stock (1 mg/mL water)
- 2 µL pepstatin A stock (1 mg/mL methanol)
- 2 µL 1 M PMSF in acetone

Each sample was mixed by pipetting up and down and by using a vortex. The samples were left on ice for 20 minutes, mixing them using a vortex twice in between. Afterwards the samples were centrifuged at 13000 g for 10 minutes at 4°C and the supernatant was transferred to a new tube while the pellet (the cellular nuclear fraction) was discarded.

2.4.2 Measurement of protein concentration

To ensure equal protein loading, the protein concentration of each sample was quantified using Pierce bicinchoninic acid (BCA) assay (Thermofisher Scientific), which is based on the biuret reaction. In the alkaline medium, sample protein reduces Cu^{2+} to Cu^{1+} and subsequently Cu^{1+} is chelated with BCA resulting in a purple product. Diluted samples and standards (25 µL) were incubated at 37°C with the BCA working reagent (200 µL) for 30 minutes. The protein concentration is directly proportional to the amount of purple complex formed. The absorbance was detected on a plate reader by spectrophotometry at 562 nm. Unknown sample concentrations were extrapolated by comparison with the standard curve.

2.4.3 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

The volume of sample containing an equal protein quantity (e.g. 20 µg of protein per sample) was calculated for patient and control samples. To each sample 4x Laemmli Sample Buffer (Bio-Rad), 10X 500 mM Dithiothreitol (Invitrogen) and additional 1% Triton X-100 in PBS were added at room temperature to constitute a total volume of 50-100 µL per sample. SDS-PAGE was performed by loading samples on to 4–20% Mini-Protean TGX Precast Gels (Bio-Rad). A chromophore labelled Precision Plus Protein Dual Color Standards (Bio-Rad) was also loaded and run in parallel to the samples, to ensure identification of protein molecular weights in samples.

The gel was secured in the inner chamber of the Bio-Rad electrophoresis system, and filled with Tris/Glycine/SDS Electrophoresis running buffer (Bio-Rad). The proteins were separated using electrophoresis by applying 100 V for 10 minutes and then 300 V for 20-30 minutes. The proteins were then transferred using Trans-Blot Turbo Midi PVDF Transfer Packs and the Trans-Blot Turbo Transfer System (Bio-Rad), which uses a semi-dry blotting method for electroblotting of proteins. The membrane was dried overnight.

The following day the membrane was briefly soaked in methanol and equilibrated in PBS. The membrane was then transferred into a 50 mL Falcon tube containing 5% w/v skimmed milk powder/ PBS, and incubated for 2 hours as a blocking step. PBS, 0.3% Tween (PBS-T) was prepared by adding 1.5 mL of Tween 20 to 500 mL of PBS. The blots were then washed briefly using with PBS-T. The blots were then placed into the blocked falcon tubes with the protein side facing inwards. The membrane was then probed by incubating with primary antibodies contained in 3 mL of PBS-T usually overnight at 4°C or for 2 hours at rooms temperature followed by three 15 minute washes using PBS-T. The membrane was subsequently probed with the appropriate (anti-rabbit/ anti-mouse) horseradish peroxidase-conjugated secondary antibody. The primary and secondary antibodies used for each experiment are included under the relevant chapter. The blots were washed three times with PBS-T and developed after incubating with 2 mL of an enhanced chemiluminescence reagent (Clarity Western ECL Substrate, Bio-Rad). Western blots were exposed for 3 – 60 minutes and protein bands visualised using the Bio-Rad Gel Doc system and densitometry was analysed with Image Lab Software.

2.5 Cytochrome *c* oxidase activity in fibroblasts

Reduced ferrocytochrome-*c* was prepared by reducing 1 g of horse ferricytochrome-*c* (Sigma) using 13 mg of l-ascorbic acid in 100 mL of 100 mM potassium phosphate buffer (KP_i) (pH 7.0). This was followed by extensive dialysis in 0.01 M KP_i (pH 7.0) at 4°C, and was stored at -80°C.

Cytochrome *c* oxidase activity was measured by spectrophotometry at 30°C by monitoring the rate of ferrocytochrome-*c* oxidation at 550 nM in 10 mM KP_i (pH 7.0) and 50 mM horse ferrocytochrome-*c*. The activity of complex IV was calculated as the first-order velocity constant (Wharton & Tzagoloff 1967). The KP_i buffer was prepared by adding 0.1M KH_2PO_4, which is acidic, to 0.1M K_2HPO_4, which is basic, titrating to a pH of 7.0.

The cell pellets were prepared as described previously by harvesting cells from two 10 cm plates. Instead of resuspending the pellet in 1 mL of PBS, 1 mL of homogenisation medium (320 mM of sucrose, 10 mM of Tris-HCl and 1 mM of EDTA stored at 4°C) was used. The cells were then pelleted by microcentrifugation for 5 minutes at 13,000 g at 4°C. The supernatant was removed and the cell pellets were stored at -80°C. The cell pellets were lysed to expose the mitochondrial enzymes by freezing in liquid nitrogen briefly and then thawing by immersing in a 30°C water bath. The spectrophotometer was set to measure at 550 nm. Initially two identical cuvettes (reference and test cuvettes) were loaded with KP_i buffer (pH 7.0) and water. The cuvettes were warmed to 30°C in water bath. Reduced cytochrome *c* was added to both cuvettes, mixed thoroughly and the cuvettes were loaded on to the spectrophotometer. The spectrophotometer was set to zero and potassium ferricyanide was added to reference cuvette to oxidise the reduced cytochrome *c*. The absorbance difference between the between the test and reference cuvettes was at this point at its maximum, the absorbance reading 0.96 ± 0.02. The assay was commenced and 30 seconds later the patient sample was added to the test

cuvette, mixed thoroughly. The entire measurement cycle lasted for 5 minutes and the absorbance values were recorded at every one-minute interval. Sample was added to the test cuvette and the decrease in absorbance, at 550 nm, was measured by the HITACHI U-3310 spectrophotometer for five minutes. Measurements of each sample were made in quadruplicate. To calculate COX activity the rate constant, k, was calculated from a Microsoft excel spreadsheet by using the natural logarithm of change in absorbance over the five minute period. After correcting for the dilution the activity was expressed as k/min/mL and then expressed by the protein concentration in each sample and as a ratio to citrate synthase.

2.6 Citrate synthase activity in fibroblasts

Citrate synthase (CS) activity was measured according to the method of Coore *et al* (Coore et al. 1971). CS catalyses the reaction between oxaloacetate and acetyl CoA, which produces citrate and free thiol groups. DTNB (5,5'-dithio bis-2-nitrobenzoid acid) is used in this assay as it binds to CS producing DTNB-CoA, a yellow chromophore absorbing at 412 nm. CS is a mitochondrial marker and its activity provides an indication of the number of mitochondria in a given sample. It is often used to normalise mitochondrial respiratory chain enzyme activities, by expressing them as a ratio to citrate synthase. Since CS found in the mitochondrial matrix, it can be exposed in the sample by permeating the mitochondrial membranes with the detergent Triton-X100.

A master mix was prepared and warmed to 30°C ensuring adequate volume for the number of samples being assayed. The Master mix for one sample (1 mL) contained 500 µL of 200 mM Tris (pH 8.0), 20 µL of 10 mM Acetyl-CoA, 20 µL of 10 mM DTNB, 10 µL of 10% Triton X-100, 10 µL of 10 mM oxaloacetate and 435 µL of combined double distilled water. For each patient sample, 995 µL of the master mix was added to 5 µL of whole cell lysate and mixed. The reactions were carried out in a 96-well plate where 240 µL of the above solution (master mix and patient sample) was pipetted into a 96-well plate. Each sample was analysed in triplicate. Four blank samples were also analysed substituting double distilled water for the patient sample.

The reaction was then read for 30 minutes at 30°C, with absorbance at 412 nm, using a Bio-Tek Synergy HT plate reader. The mean of the triplicates was calculated and the enzymatic activity within each sample was extrapolated using the extinction coefficient of DTNB (13600 $M^{-1}cm^{-1}$), the path length (0.85 cm) and the dilution of the sample. This value was then divided by the protein content of the samples, and expressed as µmol/min/mg.

2.7 Immunocytochemistry of fibroblasts

Cultured skin fibroblasts were seeded on polylysine coated coverslips covered with 500-µL of medium at 37˚C, 5% CO_2 in wells of a 6-well plate 3 days before staining. On the day of staining, 5 µM MitoTracker Red CM-H2XRos (Invitrogen/Molecular Probes) in pre-warmed DMEM was added for 45 minutes. MitoTracker Red is a red-fluorescent stain used for staining

mitochondria in live cells. Its accumulation in mitochondria is dependent upon membrane potential and the dye is well-retained after paraformaldehyde fixation. The DMEM containing MitoTracker Red was then aspirated off and the cells were covered with DMEM (without dye) and cultured for a further 30 minutes.

The coverslips were then washed three times in beakers containing ~25 mL of PBS (without $MgCl_2$ and $CaCl_2$). The cells were then fixed by incubating with fresh pre-warmed 4% w/v paraformaldehyde in PBS at room temperature for 20 minutes. The coverslips were washed in PBS as described above and then incubated for 20 minutes in a container with 10 mM sodium citrate and citric acid buffer (pH 6.0) at 90°C. This step was utilised to increase specificity of the primary antibodies. The coverslips were subsequently incubated with methanol at -20˚C for 15 min to solubilise the membranes. Afterwards a blocking step was performed to reduce non-specific antibody binding where the coverslips were then incubated with 300 µL of 10% normal goat serum in PBS at 37˚C for 30 minutes. The coverslips were then incubated with 80 µL of diluted primary antibody in PBS, 2% normal goat serum for 45 minutes at 37°C in a humidified atmosphere. The coverslips were washed once again in PBS and probed with 80 □l of 100X diluted secondary antibody tagged with an Alexa fluorophore (Molecular Probes) in 2% normal goat serum in PBS for 45 min at 37ºC in a humidified atmosphere. The coverslips were then washed again in PBS and mounted with citifluor/PBS/glycerol (Agar Scientific) supplemented with 2 µg/mL of 4',6-diamidino-2-phenylindole (DAPI) (Sigma) and the coverslips were sealed with nail varnish. Fluorescent images were obtained with a Zeiss Axioplan microscope.

2.8 Live cell staining of fibroblasts with MitoTracker green

MitoTracker green FM (Life Technologies) is a green fluorescent mitochondrial dye which is utilised to visualise mitochondria in live cells. It contains a monovalent cation which permeates the cell membrane and accumulates into negatively charged mitochondria. Once accumulated, the chloromethyl group of MitoTracker Green FM binds covalently with membrane protein sulphhydryl groups– and is retained if the mitochondria subsequently depolarise.

Patient fibroblasts were seeded on sterile glass coverslips, for 3-4 days in DMEM (containing 4.5 g/l of glucose, pyruvate and fetal bovine serum). The coverslips were placed each within its individual well, in 6-well plates. The cells were seeded on coverslips at three different densities. On the day of the staining, DMEM and phenol red-free medium (Life Technologies) were pre-warmed to 37**°**C**.** Two tubes were prepared with pre-warmed culture medium containing 5 µg/mL of DAPI (Sigma 1000x stock solution: 5 mg/mL water).

The medium was aspirated off the coverslips and 2 mL of pre-warmed, DAPI-containing culture medium was added to each well, incubating the cells for 1.5–3 hours. After 1.5–3 hours of culturing in the DAPI containing medium, MitoTracker Green FM was added to the second tube with DAPI-containing medium, to a final concentration of 200 nM. Medium was aspirated off one coverslip and 2 mL of MitoTracker Green/DAPI-containing medium was added. The remaining

MitoTracker Green/DAPI-containing medium was kept protected from light at 37°C for the next coverslip. The coverslip was incubated for exactly 15 minutes in MitoTracker Green/DAPI containing medium. The coverslip was then removed using forceps and rinsed three times in beakers with phenol red-free medium, draining the coverslips on tissue paper in between washes. After the third rinse the coverslip was drained briefly and mounted on the slide, with the cell side facing downwards. The slide was examined immediately under a fluorescence microscope (Zeiss Axioplan) and images were taken within 15 minutes before the cells appeared to deteriorate. MitoTracker Green shows significant photo-bleaching, therefore, cells were not exposed to fluorescence longer than necessary.

2.9 PicoGreen staining of mtDNA in fibroblasts

Cells were cultured in DMEM in µ-Dish 35 mm, high Glass Bottom (Ibidi) for 4-5 days. The medium covering the cells was aspirated completely and the cells were rinsed with 1 mL of Hanks' Balanced Salt Solution (156 mM NaCl, 3 mM KCl, 2 mM $MgSO_4$, 1.25 mM KH_2PO_4, 2 mM $CaCl_2$, 10 mM glucose and 10 mM HEPES·NaOH pH 7.35). The Hanks' Balanced Salt Solution (HBSS) was aspirated completely and 1 mL of 2 µL/mL of PicoGreen (6 mL of HBSS + 12 µL PicoGreen stock- Life Technologies) was added and the cells were incubated for 15 minutes. The lid was used to prevent evaporation. After 15 minutes all the PicoGreen solution was aspirated and 1 mL of 25 nM Tetramethylrhodamine, methyl ester (TMRM) (6 mL of HBSS and 3 µL of 50 µM TMRM 2000× stock in ethanol–Life Technologies) was added. The lid was put on to prevent evaporation and visualised immediately under a confocal microscope (Nikon Eclipse Ti-E inverted microscope) using the following parameters: 30 z-stack slices, 30 second interval, duration of 60 minutes, 60× lens; Fast mode=½; size=1025; 'Nyquist' to get optimal resolution. The FITC (HV=90; offset=-15; 488=2.00) and TRITC channels (HV=100; offset=5; 561=3.00) were used to monitor PicoGreen and TMRM, respectively. A recording was made for 2-3 minutes and analysed using Image J Software.

2.10 Mitochondrial translation assay in fibroblasts

This work was performed entirely by Dr Michal Minczuk and Dr Sarah Pearce, Medical Research Council (MRC) Mitochondrial Biology Unit, University of Cambridge. Patient and control fibroblasts were plated on 6-well plates. At 80% confluence, they were washed twice with methionine and cysteine-free DMEM (1mL/well of a 6 well plate). Labelling medium (600 µL) was added to each well and incubated for 20 minutes at 37°C, before adding emetine (Sigma, 25mg/mL in PBS) to a final concentration of 100 µg/mL and with a further incubation at 37°C for 30 minutes. 100 µCi ^{35}S L-Methionine (Perkin Elmer - NEG009L005MC) was then added to each well to a final concentration of 166.6 µCi/mL per well and incubated for 30 minutes at 37°C. The cells were washed with PBS and trypsinised, protein was isolated and concentration determined using BCA assay. The gel was stained with SimplyBlue SafeStain (Thermo Fisher Scientific), dried at 80°C under vacuum for 1 hour and exposed to a phosphorimage screen for 72 hours.

2.11 Statistical analysis

Where statistical analysis was required it was performed using Stata (StataCorp LP, Texas, USA).

Chapter 3

3. Next generation sequencing in a cohort of patients with cytochrome *c* oxidase deficiency

3.1 Introduction

Prior to the advent of NGS technology, patients with mitochondrial disease were investigated using a combination of mtDNA sequencing and a candidate gene approach for nuclear genes. However since mutations in ~1,158 nuclear genes encoding a mitochondrial protein (Calvo et al. 2015) can potentially cause a mitochondrial disease the candidate gene sequencing approach has limited yield. Thus, in the past many patients were classified as a probable mitochondrial disease according to their respiratory chain enzyme deficiency and did not receive a genetic diagnosis.

Initial NGS studies used gene panels targeting OXPHOS subunits and assembly factors of a given mitochondrial OXPHOS complex (e.g. complex I) investigating patients with the corresponding isolated OXPHOS deficiency (Calvo et al. 2010) . Subsequently gene panels were used, targeting the genes in the Mitocarta compendium to investigate various types of respiratory chain enzyme deficiencies (Calvo et al. 2012; Lieber et al. 2013; Legati et al. 2016). Since WES was adopted it has revolutionised gene discovery in mitochondrial disease. This study was undertaken 2011-2016 during a time where NGS, especially WES, was being applied in the study of these patients thus rapidly expanding the field of mitochondrial medicine.

This chapter reports how a cohort patients displaying cytochrome *c* oxidase deficiency (COX) were investigated using NGS.

3.1.1 Declaration of work

I recruited and collated all clinical details of patients for the study. I performed the laboratory benchwork for the targeted NGS panel analysis. Dr Vincent Plagnol and Dr Warren Emmett of UCL Genetics Institute performed the bioinformatics alignment of the raw exome data. Dr Vincent Plagnol designed and executed the software algorithms for the bioinformatics pipeline using 2,600 in-house exomes. Filtering and prioritising of candidate genes using WES data was performed by me. Sanger sequencing, mutation segregation studies and Western blots were performed by me. Homozygosity mapping was performed by Dr Kerra Pearce of UCL Genomics. Muscle histochemistry was performed by Dr Thomas Jacques. Since I collected the control samples, Dr Suzie Drury performed blinded bioinformatics analysis for control samples in the targeted NGS panel. Dr Jose Saldanha performed *in silico* protein modelling.

3.2 Methods

3.2.1 Targeted next generation sequencing: patients and study design

A cohort of patients with isolated COX deficiency was recruited as detailed in chapter 2. In 2011, a pilot study was designed to use a gene panel encompassing 39 nuclear genes encoding COX subunits and known or putative COX assembly factors. In order to evaluate the

panel, and establishing its ability to detect disease-causing mutations, control samples (patient samples with known pathogenic mutations) were analysed, using a blinded approach where the investigator was unaware of the genes or the mutations. This gene panel was used in a pilot study of nine consecutively recruited patients with isolated COX deficiency to determine its efficacy in identifying the underlying genetic cause. At the time of the design of the study, it was decided that depending on the outcome from this study it would be decided whether to proceed with screening patients using the targeted NGS method or proceed to WES.

Ethical approval for this study was obtained from the National Research Ethics Committee, London Bloomsbury, UK and informed consent was obtained from parents/ legal guardians.

3.2.2 Targeted next generation sequencing: panel of genes

A total of 39 nuclear genes were identified in 2011, which were previously or putatively linked to isolated COX deficiency. Baits were designed for these genes for targeted sequencing (table 3.1). Of these genes, 17 encoded for COX subunits (including tissue specific isoforms), 18 for known human assembly factors and three were putative human assembly factors identified by using NCBI BLAST against known yeast assembly factors (table 3.1). In addition DNA polymerase gamma (*POLG*), which causes mtDNA depletion, was included since it has been found in association with isolated COX deficiency in some cases in our centre. The method is described in detail in chapter 2 (2.2.3).

Table 3.1: Genes identified for targeted next generation exon sequencing method

Human gene	Yeast gene	Protein
Nuclear-encoded subunits		
COX4I1	COX5A	cytochrome *c* oxidase subunit IV isoform 1
COX4I2	COX5B	cytochrome *c* oxidase subunit IV isoform 2 (lung)
COX5A	COX6	cytochrome *c* oxidase subunit Va
COX5B	COX4	cytochrome *c* oxidase subunit Vb
COX6A1 (L)	COX13	cytochrome *c* oxidase subunit VIa polypeptide 1 (non-muscle tissues)
COX6A2 (H)		cytochrome *c* oxidase subunit VIa polypeptide 2 (striated muscles)
COX6B1	COX12	cytochrome *c* oxidase subunit VIb polypeptide 1 (ubiquitous)
COX6B2	-	cytochrome *c* oxidase subunit VIb polypeptide 2 (testis)
COX6C	COX9	cytochrome *c* oxidase subunit VIc
COX7A1 (H)	COX7	cytochrome *c* oxidase subunit VIIa polypeptide 1 (muscle)
COX7A2 (L)		cytochrome *c* oxidase subunit VIIa polypeptide 2 (liver)
COX7B	-	cytochrome *c* oxidase subunit VIIb
COX7B2	-	cytochrome *c* oxidase subunit VIIb2 (poorly expressed)
COX7C	COX8	cytochrome *c* oxidase subunit VIIc
COX8A	-	cytochrome *c* oxidase subunit VIIIA (ubiquitous)
COX8C	-	cytochrome *c* oxidase subunit VIIIC
COX7A2L		cytochrome *c* oxidase subunit VIIa polypeptide 2 like
Nuclear encoded assembly factors		
SURF1	SHY1	surfeit 1
SCO1	SCO1	SCO cytochrome *c* oxidase deficient homolog 1

Human gene	**Yeast gene**	**Protein**
SCO2	-	SCO cytochrome *c* oxidase deficient homolog 2
COX11	COX11	COX11 cytochrome *c* oxidase assembly homolog
COX19	COX19	COX19 cytochrome *c* oxidase assembly homolog
COX17	COX17	COX17 cytochrome *c* oxidase assembly homolog
C2orf64	PET191	chromosome 2 open reading frame 64
CMC1	CMC1	COX assembly mitochondrial protein homolog
COX10	COX10	cytochrome *c* oxidase assembly protein, haem A: farnesyltransferase
COX15	COX15	COX15 homolog, cytochrome *c* oxidase assembly protein
COX16	COX16	COX16 cytochrome *c* oxidase assembly homolog
COX18	COX18	COX18 cytochrome *c* oxidase assembly homolog
OXA1L	OXA1	oxidase (cytochrome *c*) assembly 1-like
IMMP1L	IMP1	IMP1 inner mitochondrial membrane peptidase-like
IMMP2L	IMP2	IMP2 inner mitochondrial membrane peptidase-like
Putative assembly factors		
AL050321.1	PET117	cytochrome *c* oxidase assembly factor-like
CHCHD7	COX23	coiled-coil-helix-coiled-coil-helix domain containing 7
CHCHD8	COA4	coiled-coil-helix-coiled-coil-helix domain containing 8
Others genes		
TACO1	YGR021W	translational activator of mitochondrially encoded cytochrome *c* oxidase
LRPPRC		leucine-rich PPR-motif containing
FASTKD2		Apoptotic protein
POLG	MIP1	DNA polymerase gamma, catalytic subunit

3.2.3 Selection of patients for whole exome sequencing

Within the cohort of patients with isolated COX deficiency patients were prioritised where there was adequate DNA available (at the time the requirements were 3 µg of DNA per patient). In addition subjects were prioritised where there were characteristics that suggested that the COX deficiency was a primary genetic deficiency rather than a secondary deficiency. These characteristics included the presence of parental consanguinity, or the presence of two or more affected siblings, or the presence of features of a canonical mitochondrial disease phenotype (such as Leigh syndrome, Pearson syndrome, Alpers syndrome, Sengers syndrome, pontocerebellar hypoplasia as found in *RARS2* mutations) or the presence of COX deficiency in more than one tissue, or the presence of profoundly reduced COX activity (≤ 3 SDs from the control mean i.e. COX activity expressed as a ratio to citrate synthase ≤0.008). MtDNA sequencing was performed excluding maternally inherited mtDNA mutations.

Using the above criteria, WES was performed in 30 patients and WES was performed as and when genomic DNA and parental consent was available. The sequencing was outsourced to Perkin Elmer, Oxford Gene Technologies, UCL Institute of Neurology and Beijing Genetic Institute. The methods used are discussed in chapter 2 (2.2.4). Whole exome sequences were visualised using Integrative Genomic Viewer. https://www.broadinstitute.org/igv/

3.2.4 Development of methods for bioinformatics filtering

WES was performed in 30 patients with isolated COX deficiency. During the course of the study new bioinformatics filtering methods were developed and adopted, therefore in the initial four patients a different method was used to identify candidate genes. This method consisted of comparing variants in the patient exomes against publicly available SNP databases (dbSNP and the 1000 genome project). Subsequently due to confounding factors such as pathogenic variants existing in the dbSNP database (albeit as a rare variant), lack of population frequencies in the 1000 genome database for some variants, new methods were developed for filtering. This method was developed by research groups using exome sequencing at our institution, in collaboration with UCL Genetics Institute (Dr Vincent Plagnol) to form a cohort of ~2,600 exomes. This enabled us to compare allele frequencies within the cohort and prioritise rare variants within the cohort. A minor allele frequency of <0.5% in our internal control group, as well as the National Heart Lung and Blood Institute (NHLBI) exome sequencing data set was taken as the cut-off to define a rare variant. All exome data previously analysed using public databases was re-analysed using this second method.

3.2.5 Prioritisation of candidate genes

Since all patients had unaffected parents, and in some cases patients had other unaffected siblings, an autosomal recessive mode of inheritance was assumed. Since variants which cause Mendelian diseases are expected to be rare variants, polymorphisms which are commonly found within the population can be disregarded. The filtering pipeline is depicted in Figure 3.1.

After filtering for rare variants, genes encompassing the mitochondrial proteome were prioritised in the first instance. The list of genes encoding the mitochondrial proteome was compiled by using Mitocarta 1.0, the most up-to-date version of this at the time (Pagliarini et al. 2008). MitoCarta 1.0 is a compendium of 1098 mouse genes encoding proteins where there is strong evidence for mitochondrial localisation based on mass spectrometry of mitochondria isolated from fourteen mouse tissues where localisation of the proteins was assessed through large-scale GFP tagging/microscopy. These data (curated by Broad Institute, Cambridge, USA) were integrated with six other genome-scale datasets of mitochondrial localisation, utilising a Bayesian approach. https://www.broadinstitute.org/publications/broad807s. In addition lists of genes were curated in-house using previous publications of genes linked to mitochondrial disease.

Where there were no genes localising to the mitochondrial proteome, the search was expanded to genes with a previous link to a human disease and/or genes where an animal model displayed a similar phenotype to the patient. If no plausible variants were found an attempt was made to look at the heterozygous variants and X-linked variants looking for associations with disease in OMIM (http://www.omim.org) and Pubmed searches.

All missense variants found were categorised using software tools to predict likelihood of pathogenicity, although these tools was not utilised to filter out variants. SIFT (Sorting Intolerant From Tolerant) PolyPhen-2 and Mutation Taster which calculate probabilities of the impact of amino acid changes established on protein conservation data within different species. In addition Polyphen-2 also utilises known biochemical properties of proteins to produce the predictions in a Bayesian fashion.

A novel variant was considered likely to be pathogenic, if: a) it was a truncating variant b) in the case of missense variants the affected amino acid is evolutionarily conserved in different species; c) the variant was confirmed on Sanger sequencing and it segregated with disease in affected family members, or the variant was shared among unrelated patients displaying a similar disease phenotype d) *In silico* molecular modelling of the variant predicted functional consequences e) functional complementation studies provided supporting evidence: where there was a defect of the cells (such as COX deficiency), expression of the wild type protein in the patient cells would be expected to correct the defect in the patient cells. Functional complementation studies were considered only when reporting novel disease genes and where the fibroblast cells demonstrated a functional defect.

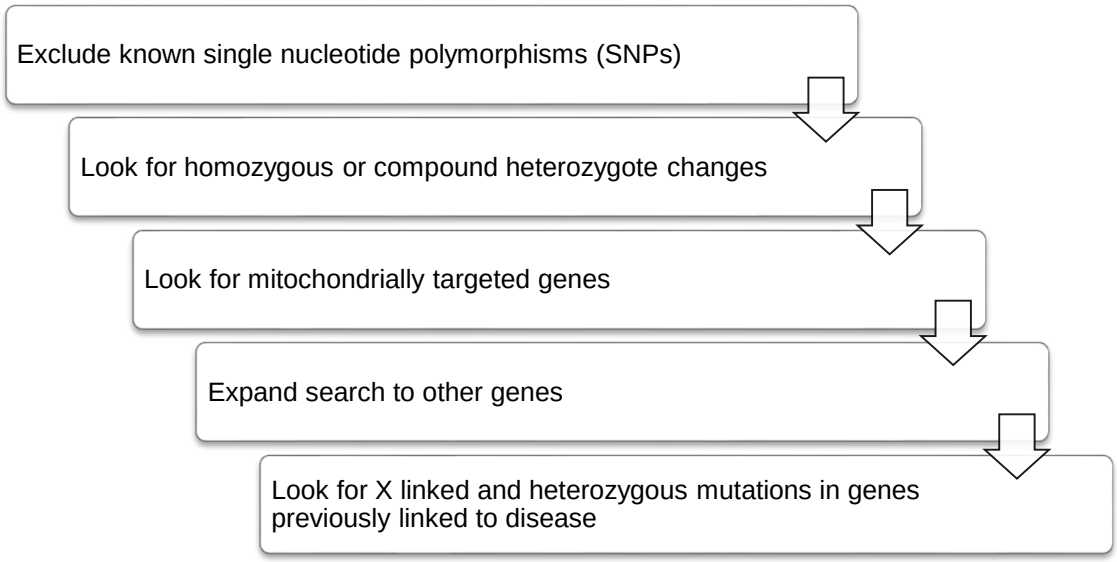


Figure 3:1: Outline of strategy used for filtering whole exome sequencing data

3.2.6 Homozygosity mapping

This method is described in chapter 2.

3.2.7 Sanger sequencing

Sanger sequencing was performed as described in chapter 2 (2.2.2). The following primer sequences were used.

AARS2 exon 4 Forward CCAGGATATTGTGGGGTCAG

AARS2 exon 4 Reverse TCAGGGATTCTCATCCTGCT

AARS2 exon 13 Forward CTTCAGACCGTGGCTACCTG

AARS2 exon 13 Reverse CTGCAGGGTATTGGGAACAG

YARS exon 4 Forward AACCGAGATCCCTGGTTACA

YARS exon 4 Reverse TACACAGGGTTGCCACACAC

YARS exon 10 Forward TCCCTCCCCGCTAGTCTAAT

YARS exon 10 Reverse ACAAACCTGACCCCAATCTG

SERAC1 exon 14 Forward TTATAGTGCAAGAGGCCTTTAGTT

SERAC1 exon 14 Reverse CACTTCCCATCCAAGAGGAC

MRPL44 exon 2 Forward ACAAGCTTTTGGACATCGGT

MRPL44 exon 2 Reverse ACAGGGCTCCAATAACTGCA

AKR1B15 exon 2 Forward GGCAACATGGCAAGGTCT

AKR1B15 exon 2 Reverse CCCGATCCAGATACTTCTGC

AKR1B15 exon 10 Forward GGAAGATCACAGGGTTGTCTCT

AKR1B15 exon 10 Reverse CCCCCAGGATGTTCTTAGC

c12orf65 exon 2 Forward AAGAGCAGTTTGTGAAAGGACA

c12orf65 exon 2 Reverse CTCAGGGATGGGACGAGAT

3.2.8 Western blot

Western blot was performed as described chapter 2 (2.4) using the following antibodies AARS2 (rabbit, sigma HPA035636, 1:500) OXPHOS cocktail (mouse, Abcam, ab110411, 1:200) and YARS (goat, Santa Cruz, C-18, 1:500).

3.2.9 Protein modeling

The effects of the nonsynonymous mutation (p.Val616Met) and the frameshift deletion mutant that caused a premature stop codon (p.Asp242fs) on the structure and function of AARS2 were examined by three-dimensional molecular modelling. Two models were created, one for the AARS2 editing domain, the template being the crystal structure of alanyl-tRNA synthetase from *Pyrococcus horikoshii* at a resolution of 2.16Å, pdb code 2ZZE (Sokabe et al. 2009) and the other for the frameshift deletion mutant, the template being the crystal structure of alanyl-tRNA from *Aquifex aeolicus* at 2.08Å, pdb code 1YFS (Swairjo & Schimmel 2005) using the program Modeller (Sali & Blundell 1993). The Modeller program builds protein models using homology-derived restraints of dihedral angles and distances for main-chain and side-chain atoms. Regions lacking template coordinates were built by an energy function optimization approach based on energy minimization using conjugate gradients and molecular dynamics with simulated annealing. The stereochemical quality of the models was assessed by the ProCheck program (Laskowski et al. 1993).

3.3 Results

3.3.1 Results of targeted next generation sequencing

Eighteen samples were analysed using this method of which 11 were disease controls (*SURF1* n=3, *POLG* n=4, *COX10* n=2, *COX15* n=1, *SCO1* n=1) and seven were unknown patient samples.

The size of the target region was 95 kilobases and more than 30 times read depth was achieved in 99.7% of the target regions. The bioinformatic analysis and prioritisation of candidate genes was performed blinded i.e. the identity of control samples and the genetic mutations were unknown at the time of analysis by the investigator. Using this method it was possible to identify all the mutations in the disease control samples which included one insertion deletion, one deletion and three splice variants (Table 3.2). No potential pathogenic mutations were detected in any of the patient samples. The filtering of variants is depicted in Table 3.2.

The results of the initial targeted gene sequencing project found that the panel was effective at detecting pathogenic mutations in the control samples, but ineffective at finding the genetic cause in the seven “unknown” patients analysed. This suggested that the patients had causative mutations in genes not included in the panel. Therefore WES was used to investigate the molecular basis of COX deficiency.

Table 3.2: Filtering of variants after targeted next generation sequencing of 39 genes in 20 patients

Patient	Total variants	Variants <0.01 in 1000 genome	Excluding variants in UTR	Significant variants found
Control1	953	207	67	*SURF1* homozygous c.516-2A>G
Control 2	972	131	35	*SURF1* c.311insATdelTCTGCCAGCC; c.751+5 G>A
Control 3	1176	302	62	*SURF1* homozygous c.792_793delG p.Arg264Serfs*27
Control 4	1152	208	74	*COX10* homozygous c.612C>A p.Asn204Lys
Control 5	1084	186	62	*COX10* c.1007A>T p.Asp336Val; c.1015C>T p.Arg339Trp
Control 6	943	204	56	*COX15* homozygous c.649C>T p.Arg217Trp
Control 7	1098	184	61	*POLG* c.1399G>A p.Ala467Thr; c.2740A>C p.Thr914Pro
Control 8	1039	236	52	*POLG* homozygous c.3286C>T p.Arg1096Cys
Control 9	990	216	49	*POLG* homozygous c.1399G>A p.Ala467Thr
Control 10	984	209	64	*POLG* c.2542G>A het (p.Gly848Ser), c.1399G>A het (p.Ala467Thr)
Control 11	1071	252	71	*SCO1* c.364+6T>C; c.880A>G p.Met294Val
S25	527	65	12	None found (only intronic variants and 1 heterozygous missense mutation in *POLG*- not significant)
S53	1028	134	43	None found (only intronic variants – not significant)
S29	864	128	29	None found (only intronic variants – not significant)
S41	926	104	39	None found (intronic variants, 1 heterozygous missense mutation in *COX18* - not significant)
S35	756	98	30	None found (intronic variants, 1 heterozygous missense mutation in *COX8C* - not significant)
S51	956	140	53	None found (intronic variants, 1 heterozygous missense mutation in *COX4I1* - not significant)
S13	871	169	58	None found (only intronic variants – not significant)

3.3.2 Whole exome sequencing: patient demographics and clinical characteristics

Within the Tables 3.3-3.5 the patients are listed in the order that WES was performed in the study. Sixteen (53%) of these patients were of white European origin, seven (23%) of Pakistani origin, three of Indian origin (10%) one (3%) each of African, Bengali, Turkish and mixed ethnicity. Male to female ratio was 17:13. As demonstrated in Table 3.4, the cohort represented broad clinical heterogeneity with overlapping clinical phenotypes ranging from peripheral neuropathy (n=5), cardiomyopathy (n=6), hypocalcaemia (n=2), nephrotic syndrome (n=1), Leigh-like syndrome (n=3), phenotype consistent with Pearson syndrome (n=1), Sengers syndrome (n=1), Alpers-like (n=1), pontocerebellar hypoplasia (n=2), cerebellar atrophy (n=3) and seizure disorders (n=6). All patients demonstrated reduced COX activity in muscle (median 0.008, range 0.003-0.013) and one patient had reduced COX in fibroblast cell cultures as well reduced muscle COX.

3.3.3 Candidate gene identification

WES and bioinformatics analysis was performed on 30 patients with isolated COX deficiency. Using the in-house bioinformatics pipeline and a cut-off of 0.5% allele frequency within in-house controls and the NHLBI dataset, rare variants within each exome ranged from 253 to 993 variants. Potential candidate genes were identified for 15/30 (50%) patients. Six subjects were from consanguineous families and in these subjects a candidate gene was identified for 4/6 subjects. Of the 30 patients analysed, 14 patients were analysed as sibling pairs and rare variants shared between siblings were prioritised. Of the seven sibling pairs candidate genes were identified for four sibling pairs (8 subjects).

Of the gene defects identified, a known pathogenic mutation was identified in four subjects. Interestingly none of the patients carried COX subunit or assembly factor mutations. A candidate gene with evidence for mitochondrial function/localisation was identified in 9/15 (60%) subjects.

.

Table 3.3: Clinical and biochemical characteristics of 30 patients with COX deficiency selected for whole exome sequencing

Patient	Age of onset	Current age	Clinical features	Plasma lactate mmol/L	CSF lactate mmol/L	Muscle COX activity
Family 4						
S8	2 y	20 years	Hereditary sensory motor neuropathy, cholestatic liver disease, Mitochondrial DNA depletion in muscle	0.7	NA	0.010
S9	2 y	18 years	Hereditary sensory motor neuropathy, cholestatic liver disease, Mitochondrial DNA depletion in muscle	1.0	NA	0.011
Family 3						
S6	neonatal	21 years	Developmental delay, dilated cardiomyopathy	1.7	NA	ND
S7	Infancy	Died at 17 years	Developmental delay, dilated cardiomyopathy	2.1	NA	0.008
Family 8						
S16	infancy	13 years	Ponto cerebellar atrophy, microcephaly, cortical visual impairment, epilepsy	1.6	NA	0.007
S17	6 weeks	1.5 years	Profound visual impairment, nystagmus, microcephaly, infantile spasms	3.8	NA	ND
Family 1						
S2	neonatal	16 years	Lactic acidosis, failure to thrive, developmental delay, microcephaly, hypocalcaemia	3.3	1.2	0.006
S3	3.5 m	13 years	Mild developmental delay, hypocalcaemia,	1.3	1.7	0.003
Family 10						

Patient	Age of onset	Current age	Clinical features	Plasma lactate mmol/L	CSF lactate mmol/L	Muscle COX activity
S20	3 m	10 years	Developmental delay, seizures, hypotonia cerebellar atrophy	1.0	1.3	0.009
S21	4 m	6 years	Developmental delay, hypotonia cerebellar atrophy	1.6	NA	0.011
Family 6						
S12	10 m	Died 18 m	Lactic acidosis, cardiomyopathy	26	NA	NA
S13	Day 1	10 years	Lactic acidosis, cardiomyopathy, adrenal insufficiency	5.2	NA	0.009
Family 9						
S18	21 m	11 years	Ataxia, dystonia increased lipid on muscle biopsy	1.2	3.1	0.006
S19	1 y	9 years	Ataxia, dystonia	2.4	NA	NA
Single index cases						
S2	Infancy	Died at 1 year	Failure to thrive, motor developmental delay, bone marrow dysplasia, sideroblastic anaemia, partial villous atrophy	2.2	1.1	0.008
S32	3 m	21 years	Cardiomyopathy, sensorineural deafness, retinal dystrophy, renal failure	1.6	2.6	0.010
S34	18 m	20 years	Sideroblastic anaemia, sensory motor axonal neuropathy,congenital deformity of hands and fingers, bicuspid aortic valve, retinal dystrophy	1.0	NA	0.009
S66	2 m	6 years	Speech delay, four limb motor disorder, ataxia, behavioural problems, ventrolateral thalamic changes, Leigh-like	1.2	NA	0.004
S27	4 m	11 years	Diplegic motor disorder, expressive language delay, motor	1.6	1.1	0.006

Patient	Age of onset	Current age	Clinical features	Plasma lactate mmol/L	CSF lactate mmol/L	Muscle COX activity
			neuronopathy, pontocerebellar hypoplasia			
S42	18 m	15 years	Nephrotic syndrome, severe mitral valve regurgitation, transient ischaemic attacks with dysphasia, caudate lesion	2.7	1.1	0.008
S25	3.5 y	13 years	Leigh syndrome, history of regression, microcephaly, seizures, ataxic gait, motor neuropathy	2.3	2.4	0.004
S49	2 y	12 years	Global developmental delay with motor regression, gastrostomy fed, pontocerebellar atrophy	1.3	NA	0.011
S43	Day 1	14 years	Lactic acidosis, sensorineural deafness, 3-methyl glutaconic aciduria, Leigh-like, striatal lesions	7.5	2.0	0.013
S74	4 m	5 years	Infantile spasms, developmental delay, basal ganglia lesions	3.8	2.7	0.005
S71	6 weeks	Died at 17 m	Cataracts, developmental delay, hypertrophic cardiomyopathy, 3-methylglutaconic aciduria	1.6	NA	0.012
S59	10 months	9 years	Sensorineural hearing loss, hypermetropia, developmental delay, sensorimotor neuropathy, limb and truncal ataxia	1.7	1.4	0.007
S36	Infancy	Died at 8 years	Kearns Sayre spectrum symptoms and basal ganglia lesions, ptosis, apnoea	4.2	3.9	<0.014
S33	1 year	20 years	Cerebellar atrophy, sensorimotor neuropathy, septo-optic dysplasia	NA	NA	0.010
S23	Day 6	Died	Epileptic encephalopathy, liver failure (Alpers-like)	1.8	1.2	0.009
S73	3 months	Died at 1.5 years	Pancytopenia, seizures, hypotonia, immunodeficiency	1.9	NA	0.006

CSF: cerebrospinal fluid, m:months

Table 3.4: Results of bioinformatics filtering of whole exome sequencing data in 7 sibling pairs (using a minor allele frequency threshold of ≤ 0.5% in the 2,600 exomes in the UCL exome consortium database)

Patient	Total variants	Rare variants	Rare compound heterozygous variants	Rare homozygous variants	Number of genes with rare variants shared between siblings	Mitochon-drial genes	Candidate genes
Family 4							
S8	24,108	506	52	35	13	0	**IGHMBP2, SCYL1*
S9	24,305	531	58	25	13	0	**IGHMBP2, SCYL1*
Family 3							
S6	24,012	360	14	3	5	1	*AARS2*
S7	23,748	358	20	3	5	1	*AARS2*
Family 8							
S16	25,440	541	87	4	11	0	Variants of uncertain significance
S17	20,861	652	81	4	11	0	Variants of uncertain significance
Family 1							
S2	21,087	467	41	17	6	0	Variants of uncertain significance
S3	21,937	488	43	31	6	0	Variants of uncertain significance
Family 10							
S20	22,883	276	11	2	6	0	*PIGW*

Patient	Total variants	Rare variants	Rare compound heterozygous variants	Rare homozygous variants	Number of genes with rare variants shared between siblings	Mitochon-drial genes	Candidate genes
S21	21,375	272	11	2	6	0	*PIGW*
Family 6							
S12	24,764	421	14	2	12	1	*MRPL44*
S13	24,492	389	10	3	12	1	*MRPL44*
Family 7							
S18	25,314	616	58	0	17	0	Variants of uncertain significance
S19	25,360	591	65	1	17	0	Variants of uncertain significance

*In view of consanguinity only genes shared between siblings with homozygous variants were prioritised as candidate gene

Table 3.5: Results of bioinfomatic filtering of whole exome sequencing data in 16 individuals (using a minor allele frequency threshold of ≤ 0.5 % data in the 2,600 exomes in the UCL exome consortium database)

Patient	Total variants	Rare variants	Rare compound heterozygous variants	Rare homozygous variants	X-linked variants	Mitochondrial genes	Candidate genes
S2	24,488	313	9	0	8	0	*YARS*
S32	27,465	993	67	14	13	0	Variants of uncertain significance
S34	23,037	420	49	20	1	0	Variants of uncertain significance
S66	21,798	263	14	0	7	0	Variants of uncertain significance
S27	22,963	540	73	20	7	2	OGDH, FBXO7, GTPBP1*
S42	22,510	315	42	3	3	0	Variants of uncertain significance
S25	22,486	419	33	18	13	2	*C12orf65, POLRMT**
S49	22,623	304	38	4	3	1	*AKR1B15*
S43	23,719	595	55	4	4	1	*SERAC1*
S74	22,329	488	54	4	10	0	*ALG13*
S71	24,889	314	14	0	9	1	*ATAD3A*
S59	24,662	329	33	2	6	0	Variants of uncertain significance
S36	24,849	320	17	2	8	0	Variants of uncertain significance
S33	24,661	336	21	0	9	0	Variants of uncertain significance
S23	24,793	314	24	0	3	0	Variants of uncertain significance
S73	25,418	253	51	2	4	0	Variants of uncertain significance

*In view of consanguinity only genes with homozygous variants were prioritised as candidate gene

3.3.4 Strategies used to prioritise of variants

Tables 3.4 and 3.5 demonstrated the filtering of variants in all cases where WES was performed (n=30). Candidate genes were identified in 15 subjects. In the following section the patients where variants of potential significance were found, are discussed in further detail.

Family 4

In this consanguineous family homozygous mapping was performed in the parents, two unaffected siblings and two affected siblings. Both affected siblings (S8 and S9) had a hereditary sensory motor neuropathy and cholestatic liver disease. Using homozygosity mapping, chromosomal regions which were shared between the affected siblings but not between the unaffected siblings or the parents were identified. These regions were found in chromosome 3,5,6 and 11 and are depicted in Table 3.6.

Table 3.6: Homozygous regions shared between affected siblings in family 1 but not between unaffected siblings or parents

Chromosome no	Homozygous regions	Size of region (bp)	No of genes in the region
3	32,994,420-33,989,932	995512	12
5	137,810,014-157,165,220	19355206	282
6	168,286,152-170,740,474	2454322	28
11	61,038,926-69,262,785	8223859	314

As demonstrated in the table there were 636 genes in the region of interest. The Maestro score was used to rank variants and genes with positive maestro scores were shortlisted. Maestro uses 8 data sets to compute the likelihood of mitochondrial localisation of 33,860 proteins listed in the Ensembl human genome database. A Maestro score >1 (i.e. positive maestro scores) is regarded as potentially localising into the mitochondrion. Of these 43 genes had maestro score of >1.

Due to the large number of potential candidate genes on homozygosity mapping, WES was performed analysing rare homozygous variants shared between the two probands within the homozygous regions. Two approaches were used for filtering. In our initial bioinfomatic pipeline of the 24,108 and 24,305 variants present in S8 and S9 exomes of which there were 6401 homozygous variants shared between affected siblings in the homozygous regions described in Table 3.6. When synonymous variants were excluded there were 3023 variants left. Of these there were only 19 variants not present in dbSNP and none of the variants had evidence of potential mitochondrial localisation. Two candidate genes *IGHMBP2* and *SCYL1* were shortlisted due to previous disease association in humans or in animal models.

The second approach to filtering was after the development of the UCL exome consortium. There were 24,108 and 24,305 variants in each of the siblings of which 506 and 531 were rare variants. Of these 35 and 25 variants in S8 and S9's exomes respectively were homozygous variants. There were a total of 13 rare homozygous variants shared between the affected siblings in the following genes: (*ST6GALNAC5, ELTD, COL24A1, KCNA10, WDR27, TAF1L, POLR1E, TJP2, HRASLS5, SCYL1, IGHMBP2, CAPN5*). Of these only *IGHMBP2, SCYL1* and *WDR27* were within the shared areas of homozygosity considered above. Of these three genes at the time of analysis *IGHMBP2* was associated with spinal muscular atrophy with respiratory distress (SMARD1). The *SCYL1* gene was not previously associated with human disease but a knock-out mouse model demonstrated an early onset progressive motor neuron disorder (Pelletier et al. 2012).

There were two homozygous missense variations in the *IGHMBP2* gene (c.1591C>A:p.Pro531Thr and c.1363A>C:p.Thr455Pro) and one homozygous missense mutation (c.462G>C: p.Glu154Asp) in *SCYL1*. Both homozygous mutations in the *IGHMBP2* gene were predicted to be benign or a polymorphism using Mutation Taster, PolyPhen and SIFT. The *SCYL1* variant was predicted to be damaging using Mutation Taster but benign or tolerated using PolyPhen or SIFT. Further work on these two genes is described in chapter 4.

Family 3

These two siblings had an autistic spectrum disorder and dilated cardiomyopathy. S6 is an 18-year-old female, born at term to non-consanguineous, white European parents. She developed respiratory grunting within the first few hours of life, requiring supplementary oxygen therapy and nasogastric feeding. She continued to have poor feeding and choking episodes during the following months. Motor and speech milestones were delayed. At 3 years, ligamentous laxity and hypotonia were noted and she was reported to be clumsy and bumping into objects. At 6 years she was found to have learning difficulties, autistic spectrum behavioural problems and at 12 years, short-term memory impairment was noted. Three years later, she developed generalised tonic-clonic seizures followed by neurological regression. An echocardiogram at the time showed a dilated left ventricle with no evidence of hypertrophy. Electromyogram and brain MRI at 13 years were normal. Currently she is able to walk short distances.

Her younger brother, S7 demonstrated a similar phenotype with early onset developmental delay, learning difficulties, ligamentous laxity, long tract signs and cardiomyopathy. He was born at term after an uncomplicated pregnancy. The neonatal period was uneventful but subsequently motor and speech development were delayed. He was able to walk independently at two years and was noted to have vacant episodes lasting approximately fifteen minutes but electroencephalogram (EEG) was normal. At 12 years, he could walk for short distances with a spastic gait. MRI brain was normal. During this time, since his sister was found to have cardiomyopathy, echocardiography was performed which revealed evidence of mild left

ventricular dilation, abnormal echodensity and a speckled pattern in myocardial muscle possibly indicative of a metabolic or mitochondrial disorder. The cardiac function was at the lower end of normal, but he was largely asymptomatic. He died at 17 years of age.

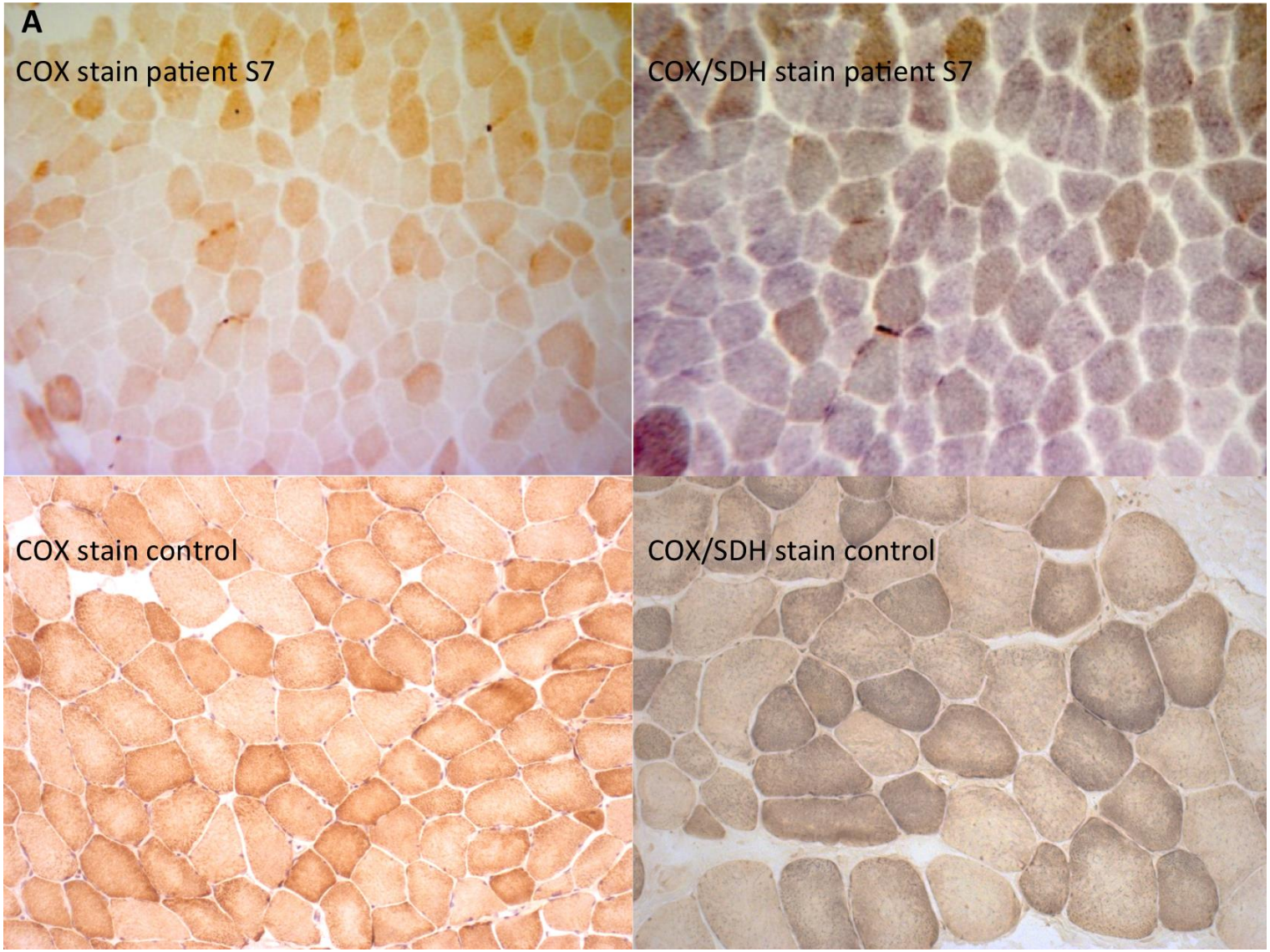
Blood investigations including lactate, copper, caeruloplasmin, manganese, acylcarnitines, purine metabolites, leucocyte lysosomal enzyme activities and urine organic acids were normal in both siblings. Mild plasma alanine elevation 578 mmol/L (reference range 150-450) was noted in S6. Following the echocardiographic findings, a mitochondrial disorder was suspected and a skeletal muscle biopsy was performed in S7. Haematoxylin and eosin stain demonstrated variation in fibre size. COX histochemical staining showed pale staining COX negative fibres which appeared blue on COX-succinate dehydrogenase (SDH) stain (Figure 3.2 A).

For these two siblings, two approaches were used to analyse exome data as they were analysed both prior to, and after the development of the bioinformatics pipeline with 2,600 in-house exomes. Analysis of exome sequencing results revealed that S6 and S7 had a total of 24,012 and 23,748 variants respectively. Since both siblings had a rare, serious disorder, with unaffected parents the assumption was that the disease was caused by rare variants not present in SNP databases (dbSNP135 or 1000 Genome project). Rare, homozygous and compound heterozygous variants shared between both siblings were filtered. There were 13,575 shared variants of which 830 were novel variants. Synonymous variants were excluded leaving 601 variants. These included variants in only three mitochondrially-targeted genes of which only one gene (*AARS2*) had two changes. Two novel variants in *AARS2* (c.725delA, p.Asn242fs and c.1846G>A, p.Val616Met) were identified in both siblings and were confirmed by Sanger sequencing. At the time these variants were not present in 6500 exomes in European or African Americans in the Exome Variant Server database. Both parents were carriers of a single heterozygous mutation (Figure 3.2).

In addition the exome data from these patients was analysed using the UCL exome consortium pipeline. It was found that S6 and S7 had 360 and 358 rare variants (<0.5% mean allele frequency) of which there were five genes shared between the siblings with compound heterozygous or homozygous variants (*AARS2, LAMA5, MYO5C, ZNF623*). Only one gene *AARS2*, encoded a product targeted to the mitochondrion and was mutations in this gene have previously associated with infantile onset hypertrophic cardiomyopathy (Gotz et al. 2011).

Of the two novel *AARS2* mutations identified, one was a frameshift mutation c.725delA, p.Asn242fs which produces a premature stop codon truncating the 985 amino acid AARS2 protein to 308 amino acids. The N-terminal domain is preserved, with absence of the editing, aminoacylation and C terminal domains. The missense mutation was located within a highly conserved region of the protein and is phylogenetically preserved from *H.sapiens* to *D.rerio* (Figure 3.2C). In addition *in silico* protein modelling demonstrated that the p.Val616Met

mutation affects the editing domain which is responsible for translational fidelity resulting in a clash with the opposing p.L610 residue and leading to a conformational change in the protein (Figure 3.3). Western blotting demonstrated reduced AARS2 protein in S7 (cells were not available in S6) (Figure 3.2D). In addition the following OXPHOS proteins were affected: NDUFB8 and UQCRC2 with preserved SDHB and ATP5A.


A
COX stain patient S7
COX/SDH stain patient S7
COX stain control
COX/SDH stain control

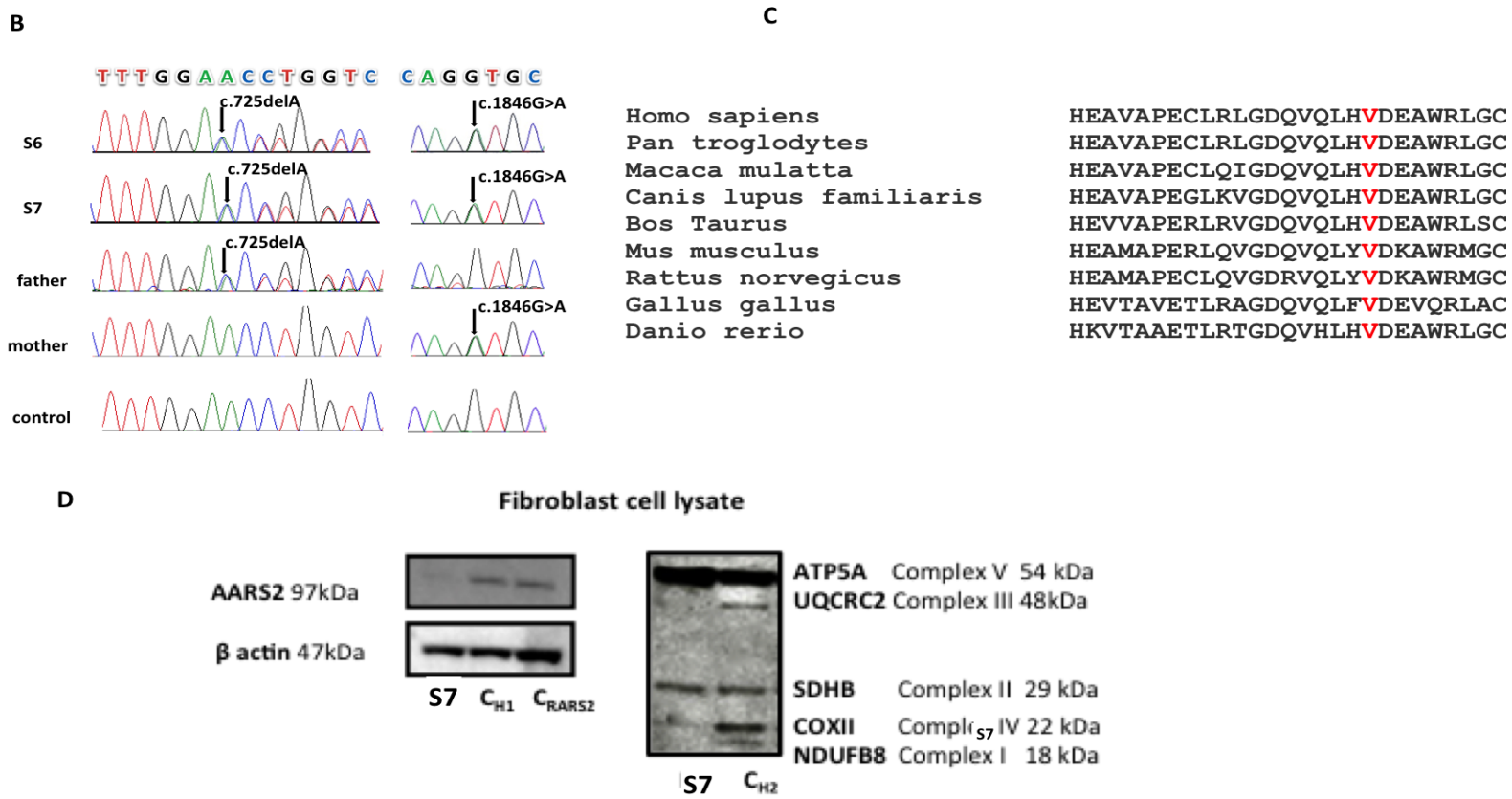


Figure 3:2: Findings in patients with *AARS2* mutations

(A) Skeletal muscle histology sections in S7 demonstrating cytochrome oxidase (COX) and COX/succinate dehydrogenase (SDH) histochemical staining compared to a control (B) Sanger sequencing results of patients, parents and a control subject (C) AARS2 amino acid sequence alignment between different species demonstrating the valine residue affected by the missense p.Val616Met substitution. (D) Western blot analysis of fibroblast whole cell lysates in patient (S7), healthy controls (CH1 and CH2) and a disease control (CRARS2), a patient with RARS2 mutations.

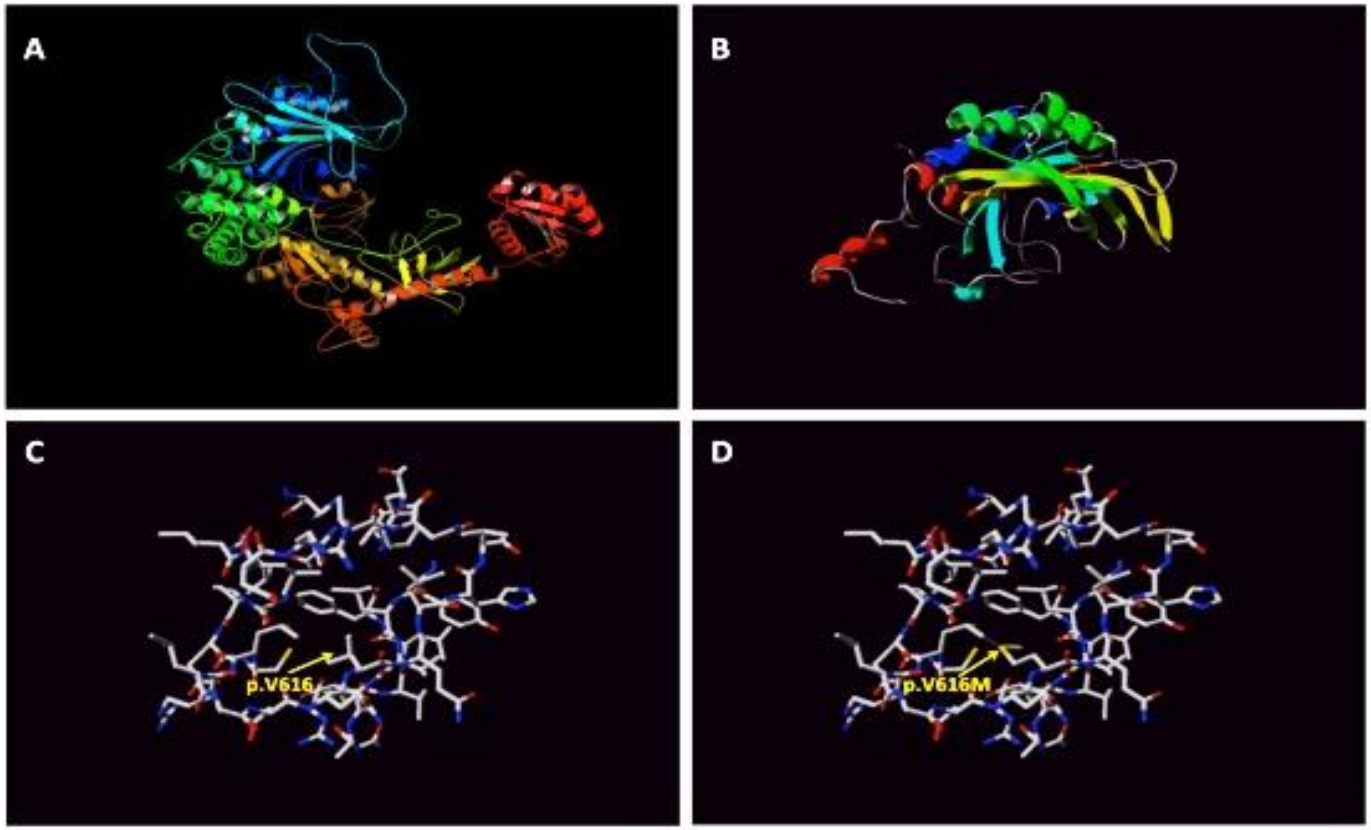


Figure 3:3: Modelling of AARS2 protein

(A) A ribbon diagram of the complete AARS2 protein (B) A ribbon diagram of the AARS2 protein with the deletion. The N-terminal domain of AARS2 is complete and is likely to fold and perform its function, but the remaining domains are missing. (C and D) Modelling of the wild type human AARS2 protein (C) and then with the p.Val616Met (D) mutation which affects the editing domain resulting in a clash with the opposing p.L610 residue leading to a conformational change in the protein.

Family 10

These two siblings (S20 and S21) born to unrelated white European parents had developmental delay, hypotonia, and cerebellar atrophy with the older sibling developing a seizure disorder. There were 22,883 and 21,375 variants in each exome of which 276 and 272 variants respectively, were rare variants. There were no genes with rare homozygous variants. There were six genes with compound heterozygous variants (*FRAS1, CMYA5, PIGW, ZNF90, FAM81B* and *TLDC2*) shared between siblings. Of these none were targeted to the mitochondrion. The *FRAS1* gene is associated with Fraser syndrome (characterised by cryophthalmos and syndactyly), which did not correlate with the clinical phenotype. Of the genes remaining, *PIGW,* which at the time of exome analysis was not associated with disease seemed a likely candidate given its role in GPI anchor synthesis and hence protein glycosylation. This genetic finding is further explored in chapter 4.

Family 6

WES was performed on 2 siblings with lactic acidosis, cardiomyopathy and COX deficiency in the younger sibling who had a muscle biopsy. S12 was born to unrelated white European parents. S12 presented with lactic acidosis (lactate 11-26 mmol/L), was found to have hypertrophic cardiomyopathy at 10 months of age and subsequent multi-organ failure at 18 months. S13 is her 10-year-old male sibling who was investigated for cardiomyopathy from an early age due to his sister's history. He was found to have lactic acidosis in cord blood and 2D echocardiography at the time showed structurally normal heart with no evidence of significant chamber wall hypertrophy. He had a muscle biopsy in the first week of life, which did not show any histological abnormalities, but COX was reduced at 0.009 (0.014-0.034). At 8 months he was developing normally and echocardiography demonstrated a pericardial effusion which required drainage. The fractional shortening was 36%, and the LV wall was 9mm at its thickest point with moderate concentric left ventricular hypertrophy. At the age of 15 months following a 10-day history of symptoms of heart failure, his echocardiogram showed a fractional shortening of 8-9%. He was put on extracorporeal membrane oxygenation and subsequently on a Berlin heart biventricular assist device. At 16 months he demonstrated neurological deterioration, was not fixing or following and neuroimaging showed bilateral infarctions in the temporo-parietal regions. He recovered from this episode and went on to have an orthotopic cardiac transplant at 21 months of age. In the immediate postoperative period he developed lactic acidosis which resolved with bicarbonate infusions. Currently he has short stature (height <0.4th centile) is on hydrocortisone therapy for adrenal insufficiency, reports some limitation on physical exercise and is progressing reasonably well in school.

There were 24,764 and 24,292 variants in S12 and S13 respectively. Of these there were 12 genes with rare homozygous or compound heterozygous variants which were shared between siblings. Only one of these genes was predicted to be targeted to the mitochondrial proteome (MRPL44); a homozygous mutation (c.467T>G, p.Leu156Arg) was found in this gene. Mutation Taster, PolyPhen and SIFT all predicted this variant as 'damaging'. This mutation has since been published in 2013 (Carroll et al. 2013) and annotated as

pathogenic. Therefore this was considered to be the cause of disease. Parents were found to be heterozygous (Figure 3.4).

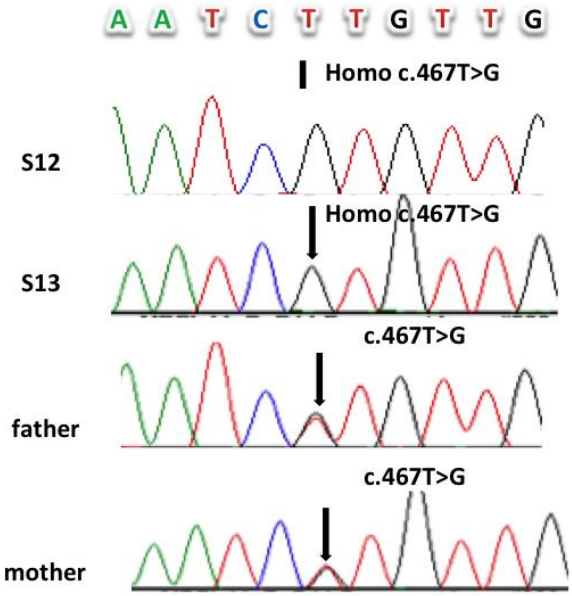


Figure 3:4: Sanger sequencing of *MRPL44* mutation found in family 6
(patients S12 and S13).

S2

WES was performed on a patient (S2) born to unrelated white British parents who presented with early-childhood onset severe failure to thrive, motor developmental delay, bone marrow dysplasia, partial villous atrophy and inflammatory enteropathy with low COX activity in muscle (0.008). Bone marrow biopsy showed trilineage dysplasia with 1 to 2% ringed sideroblasts. Stool elastase (a measure of pancreatic exocrine function) was abnormal. Renal tubular function was abnormal with a urinary N-acetyl-beta-D-glucosaminidase (NAG):Creatinine ratio of 546, but normal Retinal binding protein (RBP):creatinine ratio. Blood lactate was raised on three occasions (2.9 to 5.4 mmol/L but was normal at 2.2 mmol/L on one occasion) and CSF lactate was normal at 1.0 mmol/L. She also had low serum copper 7.6 -11.5 µmol/L (reference range 12.5-26.8), found on 4 occasions. Gamma glutamyl transferase (GGT) was slightly elevated in 111 u/L. Immunoglobulin IgG subclasses (IgG1 and IgG2) were low. Upper gastrointestinal biopsy revealed partial villous atrophy. Open muscle biopsy revealed only a slight increase in intracytoplasmic fat on histology. MtDNA rearrangements and m.3243A>G, m.8344A>G or m.8993T>G/C point mutations were excluded in muscle.

Her sister S1 (biological samples unavailable) died at the age of 10 months after a similar illness, including sideroblastic anaemia, failure to thrive and hepato-renal failure. Her liver biopsy showed evidence of fatty change, probably related to severe malnutrition, and the jejunal biopsy demonstrated partial villous atrophy.

In view of the clinical phenotype reminiscent of Pearson syndrome together with COX deficiency it was presumed she had a mitochondrial disease. Autosomal recessive inheritance was assumed and WES was performed. Bioinformatic analysis revealed 24,488 variants, of which 1652 variants which were not present in dbSNP. The 1652 variants were filtered to exclude variants present in 42 in-house exomes available at the time. Of the 90 variants left there was only one gene with two changes (YARS). In addition the exome data were analysed using the UCL exomes consortium pipeline. Variants with a frequency of <0.5% within the exome consortium database were prioritised as potentially disease causing. There were 24,488 variants of which 313 were rare variants. Of these there were no rare homozygous variants but there were compound heterozygous variant in the following nine genes: *MAPT, NXN, SETD1A, ZNF384, TTN, VWA3B, PCLO, ARHGAP30* and *YARS.* The remaining variants were prioritised according to pathogenicity and function of the genes, and YARS (Figure 3.5) was considered the most likely candidate gene (mutations: c.389C>T, p.Thr130Ileu and c.1094delC, p.365fs figure 5B) given its previous association to human disease. Western blot analysis was performed and showed reduced YARS protein in the patient compared to controls. The missense mutation was evolutionarily conserved from *Homo sapiens* to *Danio rerio* (Figure 3.5). Due to senescence of fibroblast cells, further studies were not carried out.

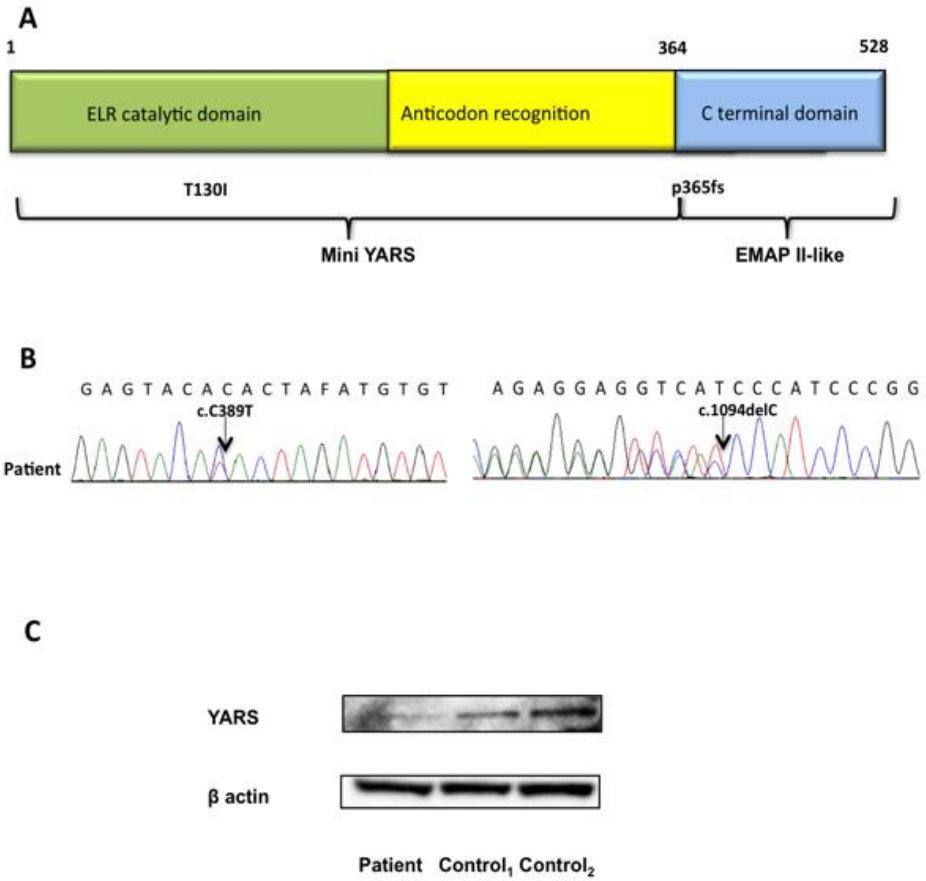


Figure 3:5: Findings in patient with *YARS* mutations

(A) Diagrammatic representation of the YARS protein denoting the position of the variants found. (B) Sanger sequencing of the YARS variants found in S2. (C) Western blot using fibroblast cell lysates from patient S2 compared to controls. (D) Cross species conservation of YARS amino acids surrounding the p.Thr130Ileu change.

S27

P19 was born at full-term by normal vaginal delivery following an uneventful pregnancy to first cousin parents. Peripheral neurophysiology showed neurogenic change in bulbar and peripheral motor neurones

with no evidence of a neuropathy; this was thought to be consistent with a neuronopathy. Muscle biopsy showed some neurogenic features and reduced COX activity (0.006, reference range 0.014-0.034). MRI brain at 5 years of age showed marked cerebellar hypoplasia with a small pons and very small medulla with a thin cervical cord. Candidate gene sequencing of the *EXOSC3* gene, (encoding a component of the exosome, which plays a role in RNA processing and degradation) was performed and no mutations were found. This gene was sequenced due to its association with pontocerebellar hypoplasia and spinal motor neuron degeneration.

Of the 22,953 variants present in S27's exome, 540 were rare variants. Of these due to parental consanguinity, homozygous changes were prioritised. There were 20 homozygous variants of which two had evidence of mitochondrial function/localisation *FBXO7* and *OGDH*. In addition there was another candidate gene guanosine-triphosphate-binding protein 1 (*GTPBP1*), which associates with both the exosome and target mRNAs, possibly similar to *EXOSC3*. Further work related to these prioritised variants is described in chapter 4.

S25

S25 is 11 years of age, from a consanguineous Turkish family and presented with global developmental delay, history of regression, microcephaly, seizures, ataxic gait, a progressive motor neuropathy, abnormal MRI brain with evidence of brainstem involvement and reduced COX activity in both muscle biopsy and cultured skin fibroblasts. She has three unaffected siblings. Thymidine and deoxyuridine were not detected in plasma or urine, excluding the possibility of MNGIE syndrome. Sequence analysis of *POLG, SURF1, COX10* and *COX15* were performed in the laboratory in Oxford (Dr Garry Brown) and found to be normal.

WES was performed in S25 assuming autosomal recessive inheritance due to a homozygous variant. There were 22,486 variants of which 419 were rare variants. Due to consanguinity, homozygous variants were prioritised; there were 18 homozygous variants of which two variants were in genes targeted to the mitochondrial proteome: *POLRMT* and *c12orf65*. Of these two genes, *c12orf65* carried a known disease causing variant c.248delT. Mutation segregation in the family (parents and two unaffected siblings) was confirmed by Sanger sequencing (Figure 3.6).

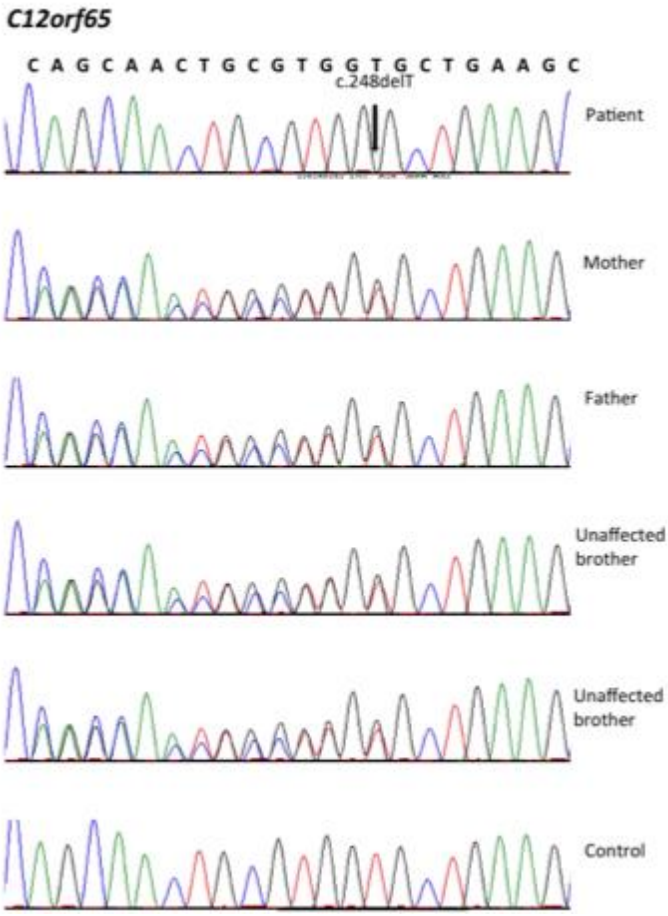


Figure 3:6: Sanger sequencing of *c12orf65* c.248delT in S25 parents and unaffected brothers

S49

S49 is a 12-year-old boy with global developmental delay, motor regression, progressive feeding difficulties (partially gastrostomy fed), neurogenic changes on muscle biopsy and pontocerebellar hypoplasia with mild white matter signal abnormalities. S49 was born after a normal pregnancy at 39 weeks gestation to unrelated white European parents. He has an older brother who has Asperger's syndrome. He also has two maternal half-brothers that are 27 and 23 years old and well.

On day 2 feeding difficulties were noted. At 10 months it was noticed that he crawled in an unusual way and by 18 months he was unable to walk with sudden falls while cruising. Neurophysiology and muscle biopsy suggested the possibility of either spinal muscular atrophy or Charcot Marie Tooth disease. By 3 years he became progressively weaker and initially needing a walking frame, and eventually stopping walking altogether. By 10 years of age he had progressive kyphoscoliosis for which he underwent surgery, after which the post-operative period was complicated by recurrent respiratory tract infections due to respiratory muscle weakness. His EMG at the time showed evidence of a myopathy. Muscle COX activity was 0.011

(0.014-0.034). The *EXOSC3* gene (due to possible small pons and cerebellum) was sequenced and no mutations were found.

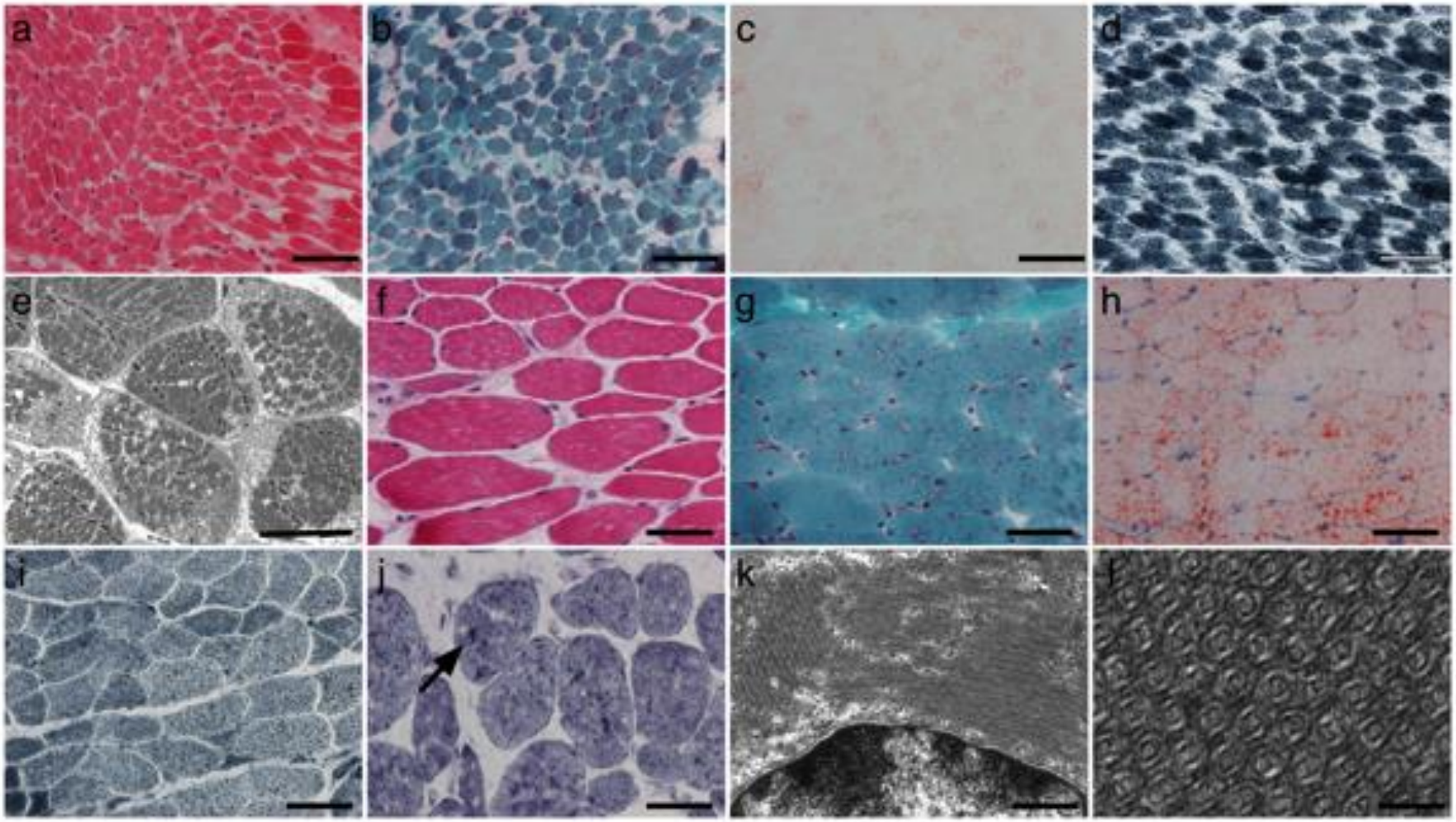
There were 22,623 variants of which, 304 were rare variants (<0.5% UCL exome consortium). Of these there were 38 compound heterozygous variants and four homozygous variants. Only one gene, *AKR1B15*, had evidence of being targeted to the mitochondrial proteome and has previously been found in a still born fetus in the Mitoexome study (Calvo et al. 2012). However segregation studies showed that the mother did not have either of the missense variants (c.856G>A:p.Val286Met and c.206G>A:p.Arg69His). DNA from the father was not available. Therefore this not was considered the cause of disease in this patient.

S43

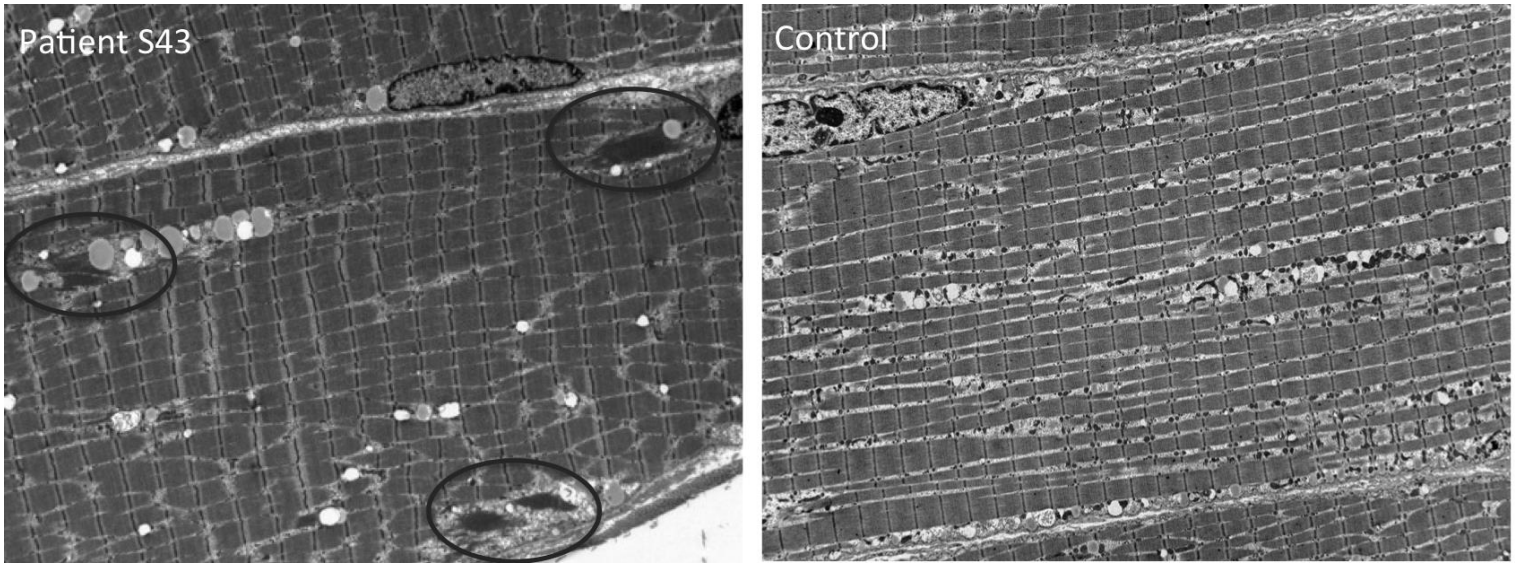
S43 is a patient presenting with developmental delay, poor feeding, dystonia, optic atrophy and microcephaly born to non-consanguineous Bengali parents. He initially presented to the intensive care unit at 3 days of age with hyperammonaemia (420 µmol/L) and lactic acidosis, requiring haemofiltration and ventilation. He appeared to make a very good recovery and was discharged home at the age of 1 month. Since then he was found to be slow to feed and had poor weight gain. The MRI scan of the brain showed abnormal signal in the striatum bilaterally, appearances consistent with mitochondrial disorder. There was also associated atrophy, most marked in the caudate heads. He also had evidence of bilateral sensorineural hearing loss. A muscle biopsy was performed during the neonatal period demonstrated complex II and complex III deficiency. However a repeat muscle biopsy was performed at 12 years of age and he was found to have a slightly reduced COX activity of 0.013 (0.014-0.034) with normal complex I and normal complex II + III activities. His muscle histology showed evidence of tubular aggregates on electron microscopy but no other specific histological changes (Figure 3.7). He was also found to have 3-methylglutaconic aciduria. He has one unaffected sibling.

WES was performed and he was found to have 21,388 variants. Variants with a frequency of <0.5% within the exome consortium database were prioritised as potentially disease causing. With this approach, initially mitochondrially targeted genes were assessed. There were 55 rare compound heterozygous calls and four rare homozygous calls. One of the mutations in the *SERAC1* gene was a known pathogenic mutation and was homozygous (c.1403+1G>C). This was therefore was thought to be the most likely candidate gene. In addition, the clinical presentation was in keeping with the phenotype described thus far (MEGDEL syndrome) for *SERAC1* mutations, a recessive disorder of dystonia and deafness with Leigh-like syndrome, impaired oxidative phosphorylation and 3-methylglutaconic aciduria. Sanger sequencing confirmed segregation of the mutation with disease in the family (Figure 3.7).

A



B



C

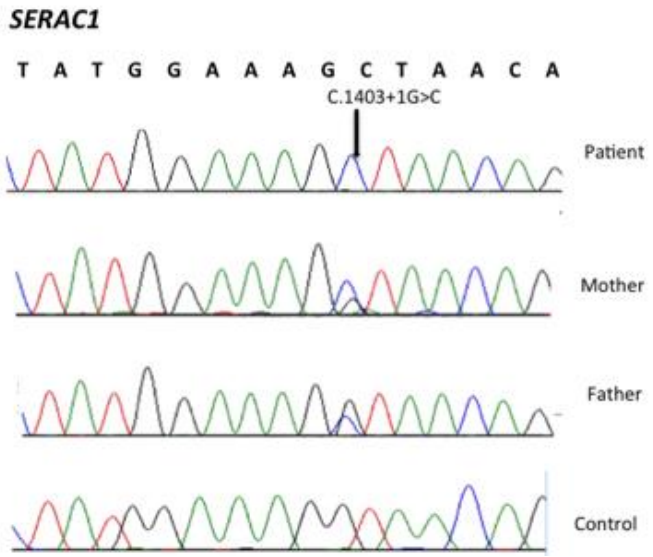


Figure 3:7: Muscle biopsy and Sanger sequencing in a patient with *SERAC1* mutations

A: Muscle biopsies taken at 15 days (**a**–**e**) and 12 years of age (**f**–**i**). At 15 days, the biopsy did not show significant diagnostic abnormalities (**a**, H&E; **b**, Gomori trichrome; **c**, Oil red O; **d**, NADH-TR). Ultrastructural examination revealed prominent nonmembrane-bound glycogen in the sarcoplasm (**e**). Specifically, there were no tubular aggregates. At 12 years of age (**f**–**i**), there were small, predominantly angular, fibres and scattered fine vacuolation and occasional deposits highlighted by the Gömöri stain (**f,** H&E; **g**, Gomori trichrome). There was excess lipid for the age (**h**, Oil red O). There was a little granularity on NADH-TR staining (**i**) but without well-defined aggregates. Histochemical staining for adenylate deaminase (**j**) showed scattered dense staining structures (arrow). Electron microscopy revealed very frequent collections of tubular aggregates, many of which are relatively small (indicated by the circles) (**k**). These showed a typical pattern of concentric double-lumen tubules (**l**). Scale bars, **a**, **b**, **c, d**, **f**, **g**, **h**, **i**, **j**: 50 µm; **e**: 10 µm; **k**: 4 µm; **l**: 100 nm. (as published in Wedatilake *et al* 2015)

B: Electron microscopy of muscle revealing very frequent collections of tubular aggregates, (indicated by the circles) in patient S43 compared to control on the left.

C: Sanger sequencing of SERAC1 mutation in patient S43, parents and a control.

S72

S72 is a female born to unrelated parents of Indian origin. She was born after a pregnancy where intrauterine growth retardation was noted at 20 weeks. Poor feeding was noted in the first few weeks but her early development was normal during the time. At four months she stopped fixing her gaze and smiling and developed episodes of 'eye rolling'. MRI brain at the time demonstrated some reduced white matter. Her legs started to jerk and she developed infantile spasms. Her seizures responded to high dose prednisolone but recurred when the dose was reduced. EEG was consistent with infantile spasms and hypsarrythmia. At this time she was also treated with vigabatrin to which she responded poorly. At the age of 8 months her feeding difficulties worsened and faltering growth was noted. There was a poor response to several anti-epileptics and she experienced global developmental delay. Blood lactate ranged between 1.5 and 3.8 mmol/L (reference range <2) and metabolic testing revealed elevated blood alanine (684 µmol/L reference range 150-450) and CSF alanine 65 µmol/L (reference range 14-40) suggestive of a respiratory chain defect. Muscle respiratory chain enzyme activities revealed COX 0.005 (0.014-0.034). S72 is currently 5 years old, is able to walk and feed herself and is seizure free since commencing nitrazepam.

There were 22,329 variants of which 488 were rare variants. There were no plausible candidate genes with recessive mutations affecting the mitoproteome or causing inborn errors of metabolism. The possibilities of X-linked and *de novo* dominant mutations was considered and a previously reported pathogenic X-linked de novo *ALG13* mutation (c.320A>G, p.Asn107Ser) was found (Allen et al. 2013). Sanger sequencing of parents demonstrated that neither parent carried the mutation i.e. it was a *de novo* mutation.

S71

S71 presented with motor delay, faltering growth, cataracts and cardiomyopathy and decreased COX 0.012 (0.014-0.034). This patient is investigated further in chapter 4.

Metabolic investigations were normal except for 3-methylglutaconic aciduria and lactic aciduria. Because of the Sengers syndrome (cataracts, cardiomyopathy and 3-methylglutaconic aciduria) phenotype the *AGK* gene was sequenced. No mutation was found, so WES was performed. The exome data demonstrated 24,889 variants of which 329 were rare variants. There were 14 rare compound heterozygous changes encompassing seven genes (*ATAD3A, SPRR1A, ATRAID, PRKRA, ADAMTS9, LRSAM1* and *ATP84*) and no rare homozygous changes. Of these there was only one gene with evidence for mitochondrial localisation, *ATAD3A,* harbouring a nonsense mutation c.511C>T: p.Arg171X and a three bp deletion p.615_616del. Further details are discussed in chapter 4.

3.4 Discussion

This chapter describes the use of NGS in uncovering the underlying genetic cause in patients with isolated COX deficiency lacking a molecular diagnosis. Initially a targeted NGS approach was tested in seven subjects, but no potential disease causing mutations were found. Therefore WES was utilised to uncover the genetic basis of COX deficiency. The patients analysed using WES consisted of either two affected siblings (n=14) or single affected cases (n=16) within a given family.

3.4.1 Targeted NGS versus whole exome sequencing

The targeted NGS panel was designed in 2011-12 at which time nearly all the patients with COX deficiency published in the literature carried either a COX subunit or assembly factor mutation (table 1.2 in chapter 1). In addition targeted NGS was successfully employed in complex I deficiency using a panel encompassing complex I subunits and assembly factors (Calvo et al. 2010). Therefore a similar strategy was employed in this study to uncover the genetic basis in COX deficient patients. However when employed in a pilot study of seven patients, no mutations were identified. One of the reasons that the panel used here did not identify mutations was that in our centre, due to clinical expertise in mitochondrial disease, patients are often referred for candidate gene sequencing based on their clinical and biochemical phenotype. This may have resulted in the population studied here being skewed.

3.4.2 Challenges in data analysis

The main challenge with the WES data was the large number of variants generated during this process and filtering strategies were developed to arrive at the pathogenic gene. During initial attempts of analysing WES data the filtering strategies used were to identify variants which were not present in public databases such as dbSNP and the 1000 genome project. It soon became apparent that one of the pitfalls of this approach was that pathogenic mutations can also be annotated with an ‘rs’ number in dbSNP, leaving room for pathogenic mutations to be filtered out. It was apparent that population frequencies may be more useful to filter out common polymorphisms. However using these approaches there were often >1000 variants left in a given patient. It became evident that in-house control data would be useful to filter out non-disease causing variants. Consequently when the UCL exome consortium with data from 2,600 exomes was developed each patient exome was filtered against this data, prioritising variants which were only present with a mean allele frequency of <0.5%. This process appeared more efficacious in filtering out possible polymorphisms since the number of variants after filtering was considerably reduced to 200-900 variants and most patients had 200-400 rare variants. Interestingly, the exomes that contained >450 rare variants were from patients of ethnicities other than of a white European background. Therefore more exome data from different ethinic groups (which would contain ethnicity specific polymorphisms) is needed to have a more robust filtering strategy when analysing data from patients with different ethnic origins. In the case of family 1, the combination of homozygosity mapping and exome sequencing i.e. focusing on the variants in the genes in the homozygous areas shared by the affected siblings (but not in the

unaffected family members) in a consanguineous family, helped to reduce the number of candidate genes.

3.4.3 Identification of candidate genes

A plausible candidate gene was identified for 14/30 (46%) cases analysed here. This compares with previous studies in patients with multiple respiratory chain enzyme deficiencies (diagnostic rate 53%) and other mitochondrial disease cohorts (diagnostic rate of 34.5%) (Taylor et al. 2014; Kohda et al. 2016). Fifteen cases were largely unsolved and the possible causes for this could be: (a) the inheritance pattern is not autosomal recessive as assumed in the bioinformatics analysis; this may be the case in families with a single affected child. Indeed one case had a *de novo* X-linked *ALG13* mutation although initial assumptions were that it was an autosomal recessive disease; (b) the disease causing mutation lies in one of the deep intronic areas which are not targeted by WES; (c) in the case of patients from ethnic minorities (other than white European) the lack of exome data from controls of a similar ethnic background led to a large number of variants being left after filtering. It was difficult to delineate ethnic group specific polymorphisms from disease causing mutations in these patients; (d) the disease was due to a copy number variant which may be missed on WES

3.4.4 Mitochondrial versus non-mitochondrial genes

Three of the mitochondrial genes identified here are involved in the mitochondrial translation pathway (*AARS2, MRPL44* and *c12orf65*). Mitochondrial translation defects should in theory cause multiple OXPHOS deficiencies with preserved complex II activity rather than isolated COX activity as found in this study. Biochemically, this finding is similar to other mitochondrial aminoacyl synthetase deficiencies (RARS2 and YARS2), where biochemical findings ranged from normal to multiple respiratory chain enzyme (RCE) defects or isolated COX deficiency (Riley et al. 2010; Glamuzina et al. 2012). In contrast to the biochemical findings, reduced complex I, III (ubiquinol:cytochrome *c* reductase) and IV subunits with preserved complex II subunit was found on immunoblotting in the *AARS2* patient, consistent with a mtDNA translation disorder. The observation of mtDNA translation defects with biochemical evidence of isolated COX deficiency adds to the complexity of unravelling the genetic basis of COX deficiency which may arise from defects of mtDNA or nuclear-encoded COX subunits, mtDNA maintenance factors or COX assembly factors.

In 6/14 cases, the candidate genes identified did not have evidence of mitochondrial localisation. In these patients there were no genes targeted to the mitochondrion after filtering for rare variants. Two patients had candidate genes, which were in the protein glycosylation pathway (*PIGW* and *ALG13*).

3.4.5 Novel findings

A novel mitochondrial gene (*ATAD3A*) previously not linked to human disease was found as a candidate gene for Sengers syndrome and two possible candidate genes were found for mtDNA depletion syndrome (*SCYL1* and *IGHMBP2*). These findings are further explored in chapter 4.

A multisystem mitochondrial disease caused by novel mutations in the *AARS2* gene

Two novel mutations in the *AARS2* gene were uncovered in this study. Pathogenicity was supported by demonstrating that p.Val616Met affects an evolutionarily conserved amino acid and by showing reduced AARS2 protein on immunoblotting. Previously AARS2 deficiency was associated with infantile hypertrophic cardiomyopathy and an adult presentation with leukodystophy and ovarian failure (Gotz et al. 2011; Dallabona et al. 2014). These findings show an atypical presentation of the *AARS2* mutations, where a mitochondrial disease was not initially obvious, with survival to adulthood, and highlight the importance of actively seeking systemic clues in unexplained neurological diseases.

SERAC1 deficiency and tubular aggregates in muscle

The patient who carried a *SERAC1* mutation had a COX level which was just below the reference range (+/- 2SD's from the control mean) had basal ganglia lesions, lactic acidosis and 3-methyl glutaconic aciduria consistent with MEGDEL syndrome, the clinical phenotype associated with SERAC1 deficiency (Wortmann et al. 2012). In this patient a new phenotypic manifestation, tubular aggregates, was found in muscle which was previously not associated with the disease. Tubular aggregates are cytosolic aggregates of membranous tubules derived from the endoplasmic reticulum.

The discovery of SERAC1 deficiency has heralded the study of a new type of phospholipid biosynthesis and remodelling disorders involving endoplasmic reticulum mitochondria-associated endoplasmic reticulum membrane (ER-MAM). SERAC1 is a phospholipase which is thought to be a phosphatidylglycerol remodelling protein found at the ER-MAM, where it facilitates phospholipid exchange and intracellular cholesterol trafficking (Wortmann et al. 2012). Tubular aggregates were present in the patient's muscle at 12 years but not in his neonatal biopsy, which may be due to ineffective lipid remodelling due to SERAC1 deficiency, leading to the accumulation of ER components. Previous reports of electron microscopy in patients with *SERAC1* mutations found abnormal mitochondrial architecture in liver and muscle but tubular aggregates were not noted (Wortmann et al. 2012; Sarig et al. 2013). This may be due the fact that previously reported muscle biopsies were performed in infancy or very early childhood, which likely explains the absence of tubular aggregates in previous cases. This case describes the changes in muscle pathology over time in SERAC1 deficiency, and adds to the differential diagnosis of tubular aggregates seen on histology.

***YARS* as a candidate gene for a Pearson syndrome-like phenotype**

YARS is the cytoplasmic tyrosyl tRNA synthetase which catalyses the aminoacylation of tyrosine tRNA with tyrosine and therefore plays an essential role in cytoplasmic protein synthesis. The N-terminal fragment of YARS has two domains: the ELR domain and the anticodon recognition domain and together are known as mini-YARS. The ELR domain contains the catalytic site, and the C terminal fragment (which is only found in the mammalian enzyme) is a human endothelial monocyte-activating polypeptide II (EMAP II)-like cytokine. In addition under apoptotic conditions YARS splits into two; the N terminal domain has Interleukin 8 (IL-8) like activities and is thought to have immune mediated actions such as acting as a chemoattractant for polymorphonuclear lymphocytes. The EMAPII like C terminal domain may have angiogenic properties (http://www.ncbi.nlm.nih.gov/gene/8565). Heterozygous YARS mutations have been implicated in dominant intermediate Charcot Marie Tooth disease. Here the mutant proteins were thought to have a dominant-negative effect on the wild type protein, possibly by reducing enzymatic activity due to dimerisation between the wild-type and mutant subunits (Jordanova et al. 2006). Jordanova *et al* found two heterozygous missense mutations (p.Gly41Arg and p.Glu196Lys) and one *de novo* deletion (153–156del) in the *YARS* gene in three unrelated families with dominant intermediate Charcot Marie Tooth disease. Further functional studies in yeast (yeast orthologue TYS1) show partial reduction in aminoacylation activity and reduced yeast growth.

The patients described here had a different phenotype to Charcot Marie Tooth disease, where patients have an axonal or demyelinating peripheral neuropathy. The patient described here has a severe phenotype with evidence of haematopoietic dysfunction, liver disease and intestinal partial villous atrophy. Two novel mutations in the YARS protein were found which were presumably recessively inherited (parental DNA not available to verify compound heterozygosity) causing loss-of-function of the protein. It is possible that the phenotypic effects in the patients are caused by a combination of loss of aminoacylation activity and loss of cytokine activities of the YARS protein. The precise cytokine activities of YARS remain unclear at present.

The mechanism of COX deficiency in this patient is likely to be a secondary mechanism; COX is a copper dependent enzyme and this patient was found to have low serum copper on several occasions (serum caeruoplasmin not measured). Low COX is a well-known secondary effect in disorders of copper metabolism such as Menkes disease (Maehara et al. 1983) and Wilson's disease (Shokeir & Shreffler 1969).

3.4.6 Summary

In summary this study identified a probable or definitive candidate gene in 14/30 patients using whole exome sequencing in a previously undiagnosed cohort of COX deficient patients. Only 8/14 cases had a probable or definitive disease gene targeted to the mitochondrial proteome. Targeted gene sequencing of COX genes did not reveal any mutations in seven patients and in

the 30 patients analysed using WES, no COX subunit or assembly factor mutations were found. WES appears to be a powerful tool in the genetic diagnosis of rare disorders such as mitochondrial diseases, and may yield surprising results.

Chapter 4

4. Functional evaluation of selected candidate genes

4.1 Introduction

In the previous chapter the application of NGS to uncover the genetic causes of COX deficient mitochondrial disease, was described. WES was performed in 30 patients. In this chapter genetic changes were functionally evaluated where a gene previously not linked to disease was found, or where there was more than one candidate gene. Additionally it was possible only to investigate patients who had stored biological material i.e. cultured patient fibroblasts. Accordingly in this chapter the following families/patients are investigated further: family 4, family 10, S27 and S71. The studies here focused on attempting to find evidence of pathogenicity for the candidate genes and on trying to uncover a phenotype in the patient fibroblast cells, which is necessary to perform functional complementation studies.

4.1.1 Declaration of work

I performed all Sanger sequencing with the exception of *ATAD3A* mutation sequencing. The Clustered regularly interspaced short palindromic repeats -associated protein-9 nuclease (CRISPR-Cas9) mouse model for Scyl1 was developed by Dr Stephane Pelletier, St Jude's Hospital, Memphis, USA. The flow cytometry for GPI anchored proteins was performed Dr Peter Krawtiz, Institut für Medizinische Genetik und Humangenetik, Berlin, Germany. Sanger sequencing of *ATAD3A* mutations was performed by Dr Jan-Willem Taanman. I performed all western blots with the exception of the western blot for the ATAD3A protein. I performed all mitophagy treatment studies on fibroblast cells. The whole cell lysates for ATAD3A western blot was prepared by me and the SDS-PAGE electrophoresis, transfer and probing with antibodies was performed by Dr Jan-Willem Taanman. Dr Michal Minczuk and Dr Sarah Pearce of the Mitochondrial Biology unit performed the mitochondrial translation assay. I performed all immunocytochemistry and live cell imaging experiments. Mitochondrial DNA quantitation using Real time qPCR , droplet digital PCR and COX activity assays were performed by me.

4.2 Methods

The methods are described in detail in chapter 2. Where there are specific differences or additional information, it is described below.

4.2.1 Sanger sequencing

The method is described chapter 2 (2.2.2). The following primer sequences were used for PCR and Sanger sequencing

IGHMBP2 exon 4 Forward GAGTTGGGTCCTGAGGTTTG

IGHMBP2 exon 4 Reverse TTGCTGCGTAGAAACAGGTG

IGHMBP2 exon 11 Forward GCCTTCATTGCTGATGTGAA

IGHMBP2 exon 11 Reverse CAGCCAACAAATCACCCTTT

SCYL1 exon 4 Forward GGTGATGAGAGCAGGGATG

SCYL1 exon 4 Reverse GAACAGGGGAGTGCCAACTA

FBXO7 exon 2 Forward TGATTCTTCAAGATGACATTCCA
FBXO7 exon 2 Reverse CTGGAACAGAATTCAACTGAGG
GTPBP1 exon 3 Forward GCACCTGCTTCATCAGCTC
GTPBP1 exon 3 Reverse TGCCAGCCATCTGGGAGT
ATAD3A exon 5 Forward ACACATGGGCACAATCACC
ATAD3A exon 5 Reverse CCCATGAAGTGAAGCAGCTC
ATAD3A exon 16 Forward GGAAGCCTGTGTTTCACGCTC
ATAD3A exon 16 Reverse GAGCCAGACGGAAAATGCATCTC
PIGW exon 2 Forward AAGCAGATGAAGGAAGCTTTTG
PIGW exon 2 Reverse AGCCCTGCCCCAAAGATAAT

4.2.2 Western blot

To investigate the effects on the mitophagy pathway, cells were treated with the mitochondrial uncoupler carbonyl cyanide 3-chlorophenylhydrazone (CCCP). When treated with CCCP, depolarisation of mitochondria occurs leading to Parkin recruitment and thereby mitophagy. PINK1 is degraded in mitochondria under basal conditions but it accumulates on the surface of uncoupled mitochondria. Parkin recruitment leads to ubiquitination of several targets, including mitofusins, MFN1 and MFN2.

The glucose containing cell culture medium was replaced with 4.5g/L of D-(+)-galactose for 5 days prior to harvesting. Galactose was used in place of glucose to induce cells to use OXPHOS rather than glycolysis. Six 10 cm plates of the following fibroblast cells were cultured: S27, three healthy controls and two disease controls (one PINK1-deficient patient and one Parkin-deficient patient). When ~80% cell confluency was achieved, two of the 10 cm plates were treated for 1 hour (t=0) and another two plates were treated for 2 hours (t=2) with 20 µmol of CCCP. Two plates were left untreated (t=0). After the treatment times, cell lysates were harvested as indicated previously for Western blot.

Antibodies used were: anti-IGHMBP2 1:300 dilution (ProteinTech, 23945-1-AP, rabbit) 1:500 dilution, anti-SCYL1 (Sigma, HPA015015, rabbit) 1:500 dilution, anti-beta Actin (Abcam, ab6276, mouse) 1:60,000 dilution, anti-FBXO7 (GeneTex GTX101986, rabbit) 1:500 dilution, anti-GTPBP1 (ProteinTech, 16374-1-AP, rabbit) 1:300 dilution, anti-MFN1 (Abcam, ab57602, mouse) 1:500 dilution, OXPHOS cocktail (Abcam, ab110411, mouse) 1:200 dilution, anti-TOM20 (Santa Cruz, sc-11415, rabbit) 1:500 dilution, anti-ATAD3A (Abcam, ab188386, rabbit) 1:500 dilution, anti-TFAM (Thermoscientific, MA5-16148, mouse)

4.2.3 Mitochondrial depletion studies in patient tissue

Dr Carl Fratter of the Oxford Medical Genetics Laboratories, Oxford University Hospitals NHS Trust, Oxford UK performed depletion studies in patient muscle and liver.

4.2.4 Droplet digital PCR for quantitating mtDNA in fibroblasts

The mtDNA copy number was quantitated using the ddPCR method described in chapter 2 (2.2.7). DNA was isolated from fibroblast cell pellets using the Gentra Puregene Tissue Kit (Qiagen) according to manufacturers instructions. MtDNA copy number was measured in controls and two POLG-deficient patients (1: p.Trp748Ser, p.His1110Tyr and 2: p.Arg232His, p.Gly848Ser). The POLG deficient patient cells were used as disease controls.

4.2.5 Real time qPCR for quantitating mtDNA in mouse tissue

The method is described in chapter 2 (2.2.6). The *Scyl1* knock out mouse mtDNA copy number was compared against the wild type mouse mtDNA copy number using the Mann-Whitney U test in muscle, nerve, liver, cardiac, cerebral and cerebellar tissues.

4.2.6 COX activity assay

COX activity was measured in patient and control fibroblasts grown on galactose medium as described in chapter 2 (2.5).

4.2.7 Live staining of nucleoids using PicoGreen/TMRM and immunocytochemistry of fixed cells

Detailed methods are in chapter 2 (2.7-2.9). Mitochondrial nucleoids were visualised using PicoGreen/TMRM staining in live cultured skin fibroblasts. Antibodies used for immunocytochemistry included anti-MTCO1 antibody mouse (ab14705) 1:500 and anti-TOM20 (sc-11415) rabbit 1:500.

4.2.8 Mitochondrial translation assay

This assay was performed entirely by Dr Michal Minzcuk and Dr Sarah Pearce of the Mitochondrial Biology Unit Cambridge and the method is described in chapter 2 (2.10).

4.2.9 Electron microscopy of fibroblast cells

Electron microscopy was performed by Dr Rosalind King at the UCL Institute of Neurology.

4.2.10 Fibroblast cell flow cytometry to detect GPI anchor proteins

This work was performed by our collaborators Dr Peter Institut für Medizinische Genetik und Humangenetik, Berlin, Germany. Fibroblasts were obtained from deep skin biopsies, treated with collagenase and grown in a Dulbecco's Modified Eagles medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 1% glutamine, 1% penicillin/streptomycin and grown in a 5% CO_2 incubator.

For investigation of GPI-anchored proteins, cultured skin fibroblasts derived from patients S20, S21 and controls were detached from the dish with 0.01% Trypsin in EDTA, pipetted several times to get a single cell suspension, counted, washed and stained with FLAER (Fluorescent labelled aerolysin (AF-488)) and antibodies against GPI-linked proteins (CD87-APC, CD73-PE-

Cy7) in PBS without calcium or magnesium and supplemented with 2% FBS for 20 minutes at room temperature. Cells were then centrifuged at 300 g for 3 minutes and washed twice in the above solution. Labelled cells were analysed using a CantoII flow cytometer (BD Biosciences) and FlowJo™ v. 9.8.2 software.

4.3 Results and discussion

4.3.1 Family 4: Case histories

S8 and S9 are female siblings born to Pakistani first cousin parents. S8, who is currently aged 19 years, and S9 aged 17 years have similar problems with hereditary sensory motor neuropathy, cholestatic liver disease, reduced COX activity in muscle and mtDNA depletion in muscle. They have two other male and female siblings, who are both well.

S8 was well until the age of 22 months when she started walking with a limp and falling. She subsequently developed dropped foot arches, puffy red palms and clawing of her hands. She remained alert and cognitively normal throughout this episode at 22 months. She required splinting of her feet to allow her to walk but she continued to develop muscle wasting and joint pains. Intercurrent viral and bacterial infections led to exacerbation of her lower limb symptoms. She was thought to have Guillain-Barre syndrome initially and she recovered her ability to walk, albeit with bilateral foot drop. Over the next 15 years she developed 3-4 episodes of stepwise deterioration with acute weakness of her upper and lower limbs. These episodes of deterioration coincided with intercurrent illness.

At age 13, S8 developed jaundice, abdominal pain and progressive pruritus. She was walking independently until at 13 years, she deteriorated to the point where she was struggling to stand and her hands became very contracted. A diagnosis of progressive familial intrahepatic cholestasis type 1 was made. She was prescribed ursodeoxycholic acid, a zinc supplement and a reduced fat diet.

She currently uses a wheelchair full-time for mobility. Her respiratory status was unaffected as indicated by forced vital capacity of 3.99 litres (109% predicted). On examination of the upper limbs, tone and proximal power was normal. Distal power was reduced with fixed flexion of the fingers bilaterally. There was thenar, hypothenar and interosseous muscle wasting. Biceps and triceps jerks were present. In the lower limbs, tone was normal. However, there was marked wasting of the lower legs. There was weak dorsiflexion and extension of the ankle and minimal movement of the toes and absent ankle jerks.

S9 was well until age of 20 months when she developed difficulty walking, puffiness and redness of her hands and clawing of the fingers. At the age of four years, she developed severe occipital headaches. During the ages of five and seven, she had two episodes during which her motor function deteriorated sharply. On the first occasion, she woke up to find she could not lift her head off the pillow and she had no control of her legs or arms. This episode

left her with reduced muscle tone and bulk in the legs, arms and shoulders. Muscle tone never fully returned to normal. She subsequently had a viral infection which led to further deterioration and she sustained fractures of her little toe and radius.

She developed jaundice with abdominal pain and pruritus at the age of 10 years. The tests for hepatitis were negative and following liver biopsy a diagnosis of familial intrahepatic cholestasis type 1 was made. She was started on treatment with ursodeoxycholic acid, and zinc in view of low plasma zinc level. She was also put on reduced fat diet. At age 15 years she developed a further episode of jaundice and weight loss.

She continued to have occipital headaches, hair loss and developed the appearance of a squint. She also developed exacerbations of jaundice when she was fatigued.

On examination, she was using a wheelchair. There was weakness in the hands. The grip strength was MRC grade 3/5. She was able to elevate the arms with full range antigravity at the shoulders and elbows. Knee extensors strength was MRC grade 3+/5, but there was marked weakness and no movement at the ankles or the toes.

The results are summarised in Table 4.1. These investigations were performed in the diagnostic laboratories.

Table 4.1: Investigations in patients S8 and S9

Investigation	**S8**	**S9**
Nerve conduction studies	Progressive and severe axonal sensory motor neuropathy	Progressive and severe axonal sensory motor neuropathy.
Muscle biopsy	Number of atrophic fibres and some hypertrophic fibres with evidence of group atrophy and some focal fibre type grouping, supporting a neurogenic process. There was irregularity of staining with SDH and pale staining of COX, with no true COX negative fibres.	Neurogenic changes and mildly increased lipid content - no COX-negative fibres.
Nerve biopsy	Chronic inflammatory neuropathy	Not available
Liver biopsy	Marked canalicular cholestasis with scanty evidence of fibrosis.	Marked canalicular and hepatocellular cholestasis, with mild lobular inflammation
MRI brain	Normal	Normal
Blood and CSF	Normal	Normal

Investigation	S8	S9
lactate		
Platelet thymidine phosphorylase activity.	Normal	Normal
Muscle COX/CS ratio (reference range 0.014-0.034)	0.010	0.011
Liver RCEs	Normal	Liver sample was insufficient for mitochondrial respiratory chain enzyme assay
Sequencing of candidate genes	POLG, PEO1, DGUOK, SUCLA2, MPV17 and MFN2 normal	POLG, PEO1, DGUOK, SUCLA2, MPV17 and MFN2 normal
Muscle mtDNA	Quantitation: 29% compared to normal control levels	Quantitation: 32% compared to normal control levels
Liver mtDNA	Quantitation: normal	Quantitation: normal

4.3.2 Family 4: Candidate genes from whole exome sequencing

Two candidate genes were thought to be possible causes of the disease and are analysed individually below. The analyses of these candidate genes were performed during 2012 -2014.

4.3.3 Family 4: Immunoglobulin mu binding protein 2 protein (*IGHMBP2*) as a candidate gene

The *IGHMBP2* gene is located in chromosome 11q13.2-q13.4 and encodes a 993 amino acid protein. It contains three domains: a helicase domain, a R3H domain and a zinc finger domain. IGHMBP2 is thought to localise in the cytoplasm and *in vitro* studies indicate that it is an ATP dependent 5'-3' helicase which uncoils DNA and RNA duplexes (Guenther et al. 2009). In addition Guenther *et al* noted that IGHMBP2 interacts with the large and the small ribosomal subunits. According to the crystal structure of the helicase domain (residues 1-652) it is structurally similar to Upf1 where both are noted to have nucleotide and nucleic acid-binding sites (Lim et al. 2012).

In 2001 Grohmann *et al* found that spinal muscular atrophy with respiratory distress (SMARD1) and distal hereditary motor neuronopathy type VI (HMN6) was caused by mutations in the *IGHMBP2* gene (Grohmann et al. 2001). This is a severe childhood-onset disorder which affects the motor neurons resulting diaphragmatic weakness and respiratory failure usually in the first year of life. The majority of the *IGHMBP2* mutations causing SMARD1 affect the helicase

domain. The remaining mutations found outside the helicase domain were truncating mutations such as deletions or nonsense mutations. Only one missense mutation outside the helicase domain has been found in association with SMARD1 but it was a compound heterozygous mutation in association with a deletion (Grohmann et al. 2003).

The 'nmd' mouse is a naturally occurring mouse with recessive mutations (single amino acid deletion and a splice donor mutation) in the *IGHMBP2* gene (Cox et al. 1998). The mouse demonstrates axonal degeneration with respiratory failure during the late stages of the disease. A recent study demonstrated that adeno-associated virus mediated IGHMBP2 administration resulted in amelioration of the motor phenotype and increased the life span (Nizzardo et al. 2015).

4.3.4 Family 4: Sanger sequencing, protein conservation and Western blot of *IGHMBP2*

In family 4, two homozygous *IGHMBP2* mutations (c.1363A>C and c.1591C>A) were found on WES. For both mutations SIFT, PolyPhen and Mutation Taster predicted these changes to be benign or a polymorphism. They were not found in the dbSNP, 1000 genome or exome variant server databases at the time of analysis. More recently the allele frequency in the ExAC database is as follows: c.1363A>C, p.Thr455Pro variant was 0.00006 and for the c.1591C>A, p.Pro531Thr variant was 0.00009 indicating they are rare variants. Sanger sequencing of the mutations was performed to verify the mutations and the results are depicted in Figure 4.1. Both mutations lie within the helicase domain. Protein conservation data showed that the amino acids affected by the p.Thr455Pro and the p.Pro531Thr changes were only conserved in *P tronglodytes* and *M mullata* and not in other species. Western blotting showed that the IGHMBP2 protein levels in the patients were lower than in three controls, but the levels in the controls showed a large variation. Therefore, the levels in the patients were not significantly lower (Figure 4.1).

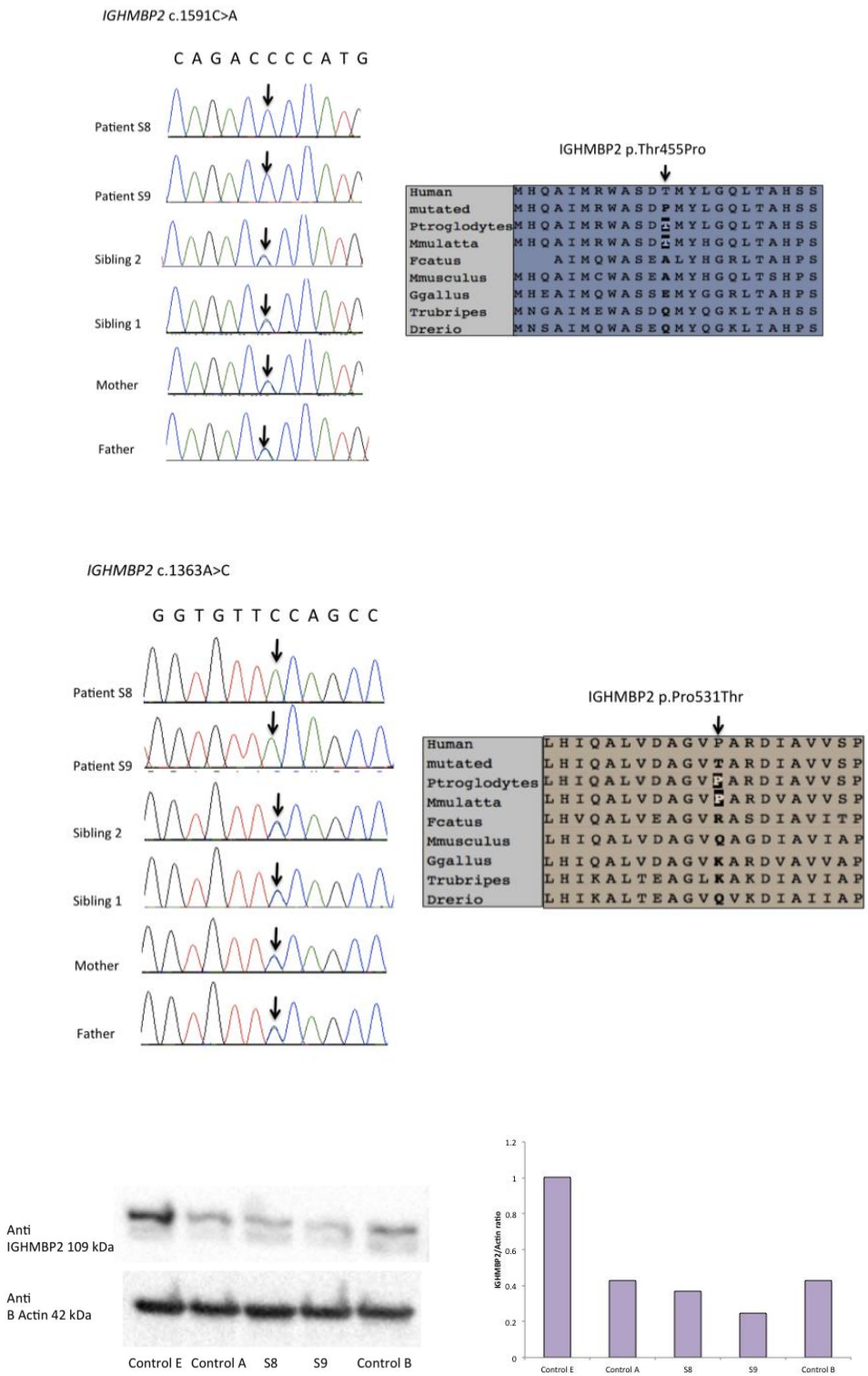


Figure 4:1: Sanger sequencing of *IGHMBP2* c.1592C>A and c.1363A>C mutations in family 6.

Sibling 1 and 2 are unaffected. Cross species amino acid conservation for the respective mutation is demonstrated alongside the sequences. Below: western blot depicting IGHMBP2 protein levels in S8 and S9 compared to healthy controls.

4.3.5 Family 4: SCY1 like pseudokinase 1 (*SCYL1*) as a candidate gene

The *SCYL1* gene is located on chromosome 11q13. The SCYL protein family has three members: SCYL1, SCYL2, and SCYL3 (Pelletier 2016). SCYL1 is catalytically inactive protein kinase which contains a serine-threonine kinase-like domain at the N-terminal, a HEAT repeat domain in the centre and a C-terminal containing a coiled-coil region. Published work thus far indicates that SCYL1 is involved in COPI-mediated retrograde protein trafficking and has a role in between the Golgi body and the membrane trafficking machinery mediated by COPI-coated vesicles (Burman et al. 2008). It also has a role in Golgi morphology (Burman et al. 2010). SCYL1 is also thought to be involved in shuttling of tRNAs between the nucleus and the cytoplasm by interacting with the nuclear pore complex proteins Nup98 and Nup107, XPOT and XPO5 (nuclear tRNA export receptors) and GTP-bound form of Ran (a GTPase involved in the nuclear cytoplasmic shuttling of RNA) (Pelletier 2016).

In addition it is thought to be a transcriptional regulator which has the ability to bind specific DNA sequences through its C-terminal domain. It is considered to be a transcriptional activator of the telomerase reverse transcriptase and DNA polymerase beta gene (http://www.ncbi.nlm.nih.gov/gene). It is a highly conserved gene among eukaryotes.

The *mdf* (muscle deficient) mouse phenotype was found to be caused by a homozygous 1-bp insertion (c.1169_1170insT) in exon 8 of the *Scyl1* gene (Schmidt et al. 2007). The *mdf* phenotype is associated with a spinocerebellar ataxia, characterised by progressive gait ataxia, cerebellar vermis atrophy, purkinje cell loss and optic-nerve atrophy. In *Scyl1* knock out mice (Pelletier et al. 2012) growth abnormalities and an early onset progressive motor neurone disease was found together with mislocalisation of TDP-43 and ubiquilin 2, similar to amytrophic lateral sclerosis (Neumann et al. 2006). In addition Scyl1 was found to be mainly expressed in cortical neurons, brainstem neurons, and anterior horn spinal cord neurons (Schmidt et al. 2007).

4.3.6 Family 4: Sanger sequencing, protein conservation and Western blot of *SCYL1*

In family 4, a homozygous *SCYL1* mutation (*SCYL1* c.462G>C, p.Glu154Asp) was found on WES. SIFT and PolyPhen predicted that the mutation was tolerated or benign whereas Mutation Taster predicted it to be disease causing. At the time this variant was not found in dbSNP, 1000 genome or exome variant server databases. Since then it has appeared in the ExAC database with an allele frequency of 0.0007 indicating it is a rare variant. Sanger sequencing of the mutation was performed to verify the mutations and these are depicted in Figure 4.2. Amino acid conservation across species showed that the affected amino acid was conserved up to *Danio rerio*. Western blotting did not demonstrate an appreciable decrease in SCYL1 protein in the patients compared to healthy controls and a disease control (SURF1 deficient patient) (Figure 4.2).

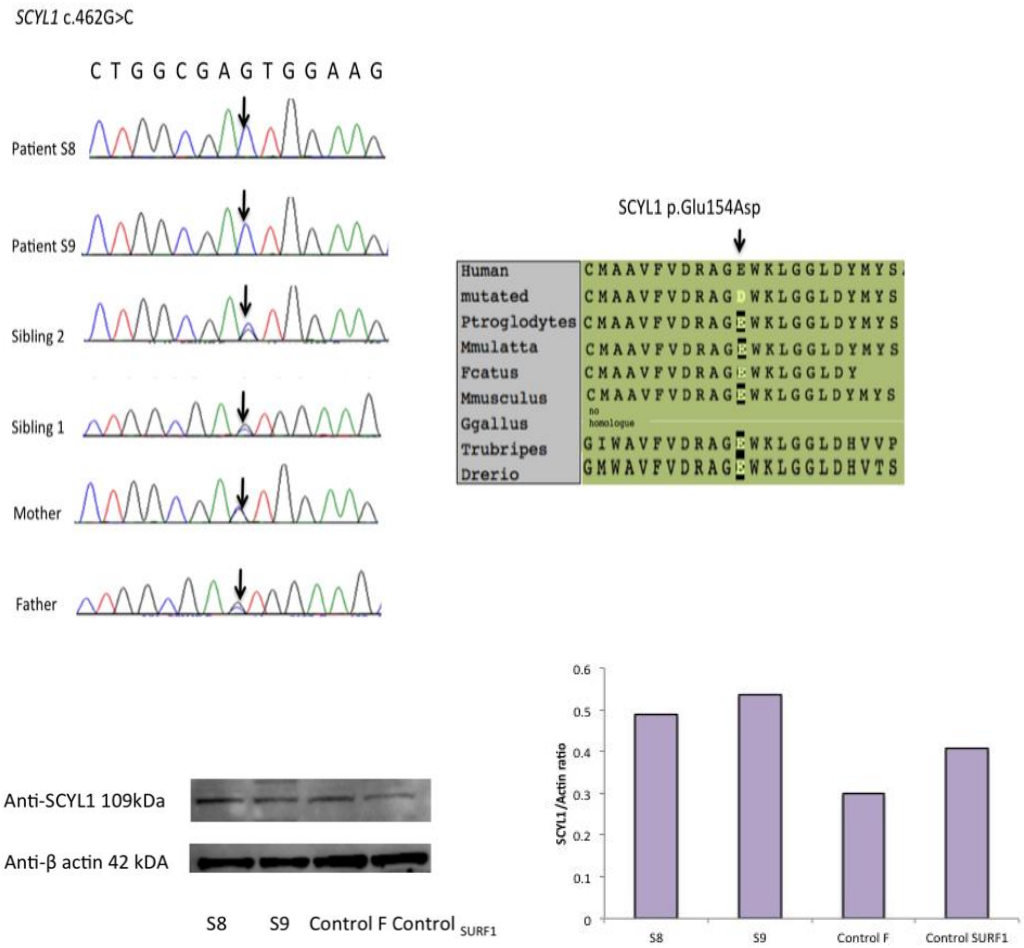


Figure 4:2: Sanger sequencing of *SCYL1* c.462G>C in family 6.

Sibling 1 and 2 are unaffected. Cross species amino acid conservation the respective mutation is demonstrated in alongside the sequences. Western blot depicting SCYL1 protein levels in S8 and S9 compared to a healthy control and a SURF1 deficient patient.

4.3.7 Family 4: Mitochondrial DNA levels in *Scyl1* knock out mouse tissue

The patients (S8 and S9) demonstrated mtDNA depletion in muscle tissue. Therefore it was hypothesised that if *SCYL1* was the genetic cause of the disease, then the *Scyl1* knock out mouse should also demonstrate mtDNA depletion. To this end collaboration with Stephane Pelletier, St. Jude Children's Research Hospital, Memphis, TN, United States was established. Samples of the following tissues were obtained: muscle, nerve, liver, heart, cerebrum and cerebellum from 8 week old *Scyl1* knock out mice (n=5) and control wild type littermates (n=7). As demonstrated in Figure 4.3 no significant difference in mitochondrial DNA copy number was found between the *Scyl1* knock out mice and control wild-type littermates in muscle (p=0.2), nerve (p=0.6), liver (p=0.1), heart (p=0.06), cerebellum (p=0.46), and cerebrum (p=0.46).

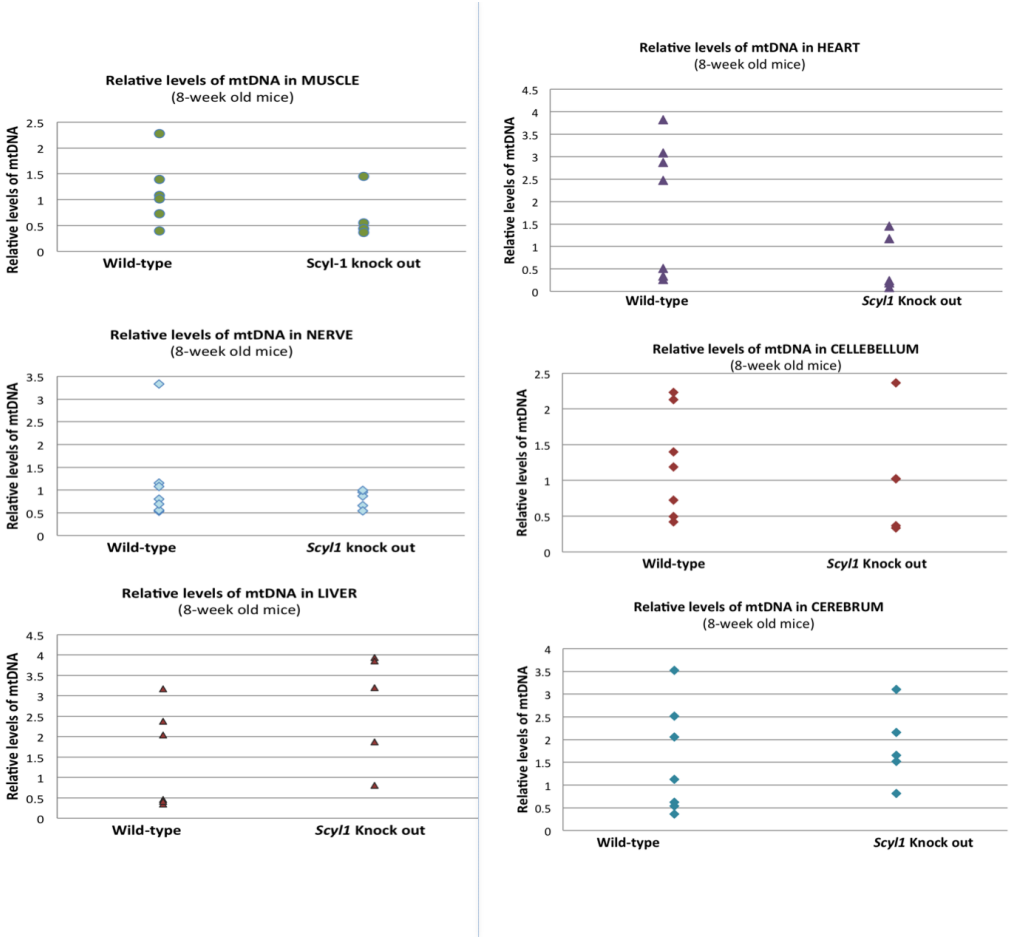


Figure 4:3: Relative mtDNA levels in *Scyl1* knock out mice compared to wild type littermate controls.

(In muscle nerve, liver tissue, cardiac, cerebellar and cerebral tissue)

4.3.8 Family 4: *Scyl1*-Glu154Asp mouse model

This work was performed by entirely by Stephane Pelletier, St. Jude Children's Research Hospital, Memphis, TN, United States using CRISPR-Cas9 technology, genotyping (PCR digest and direct sequencing), protein expression and assessing motor functions. Dr Pelletier and colleagues developed a *Scyl1* homozygous p.Glu154Asp mouse (the same mutation as the patients S8 and S9). The mouse however did not show an apparent motor phenotype demonstrating normal body weight and grip test at 15 weeks (Figure 4.4). The liver also appeared normal.

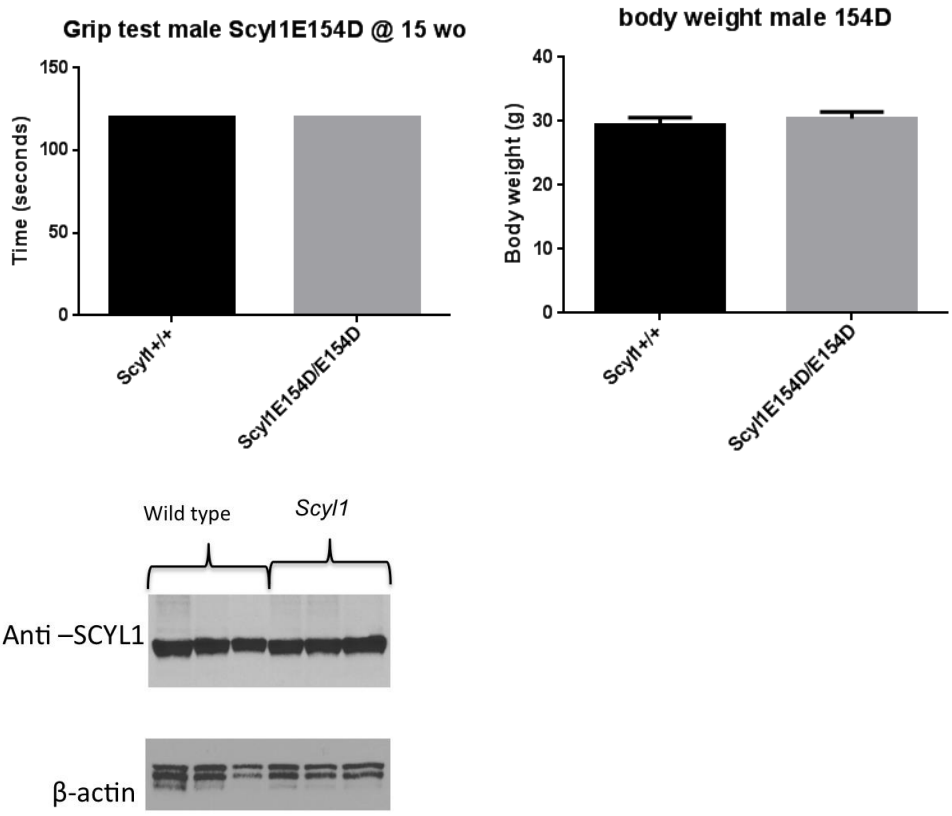


Figure 4:4: Findings in *Scyl1* homozygous p.Glu154Asp mice.

Grip test and body weight in *Scyl1* homozygous p.Glu154Asp mice compared to wild type littermates (n=5 males in each genotype). Western blot demonstrating SCYL1 protein levels in wild type and mutant mice.

4.3.9 Family 4: Fibroblast COX activity, mtDNA quantitation and PicoGreen/TMRM staining of nucleoids

The following investigations were performed, in order to detect a phenotype in fibroblast cells.

There was no evidence of COX deficiency in patient fibroblasts. COX activity (in cells grown in galactose medium, normalised to protein concentration) for controls (n=4) was mean±SD 0.640±0.279 k/min/mg. For S8 and S9 COX values were 0.656 and 0.626 k/min/mg. COX activity (on cells grown in galactose medium, normalised to citrate synthase activity) for controls (n=4) was mean±SD 0.014±0.006. For S8 and S9 COX/citrate synthase values were 0.017 and 0.012 respectively. MtDNA quantitation in S8 and S9 fibroblasts was normal compared to controls. The results are shown graphically in Figure 4.5 where patients with *POLG* mutations were used as disease controls. There was no significant difference in the appearance of nucleoids between patients and controls (Figure 4.6).

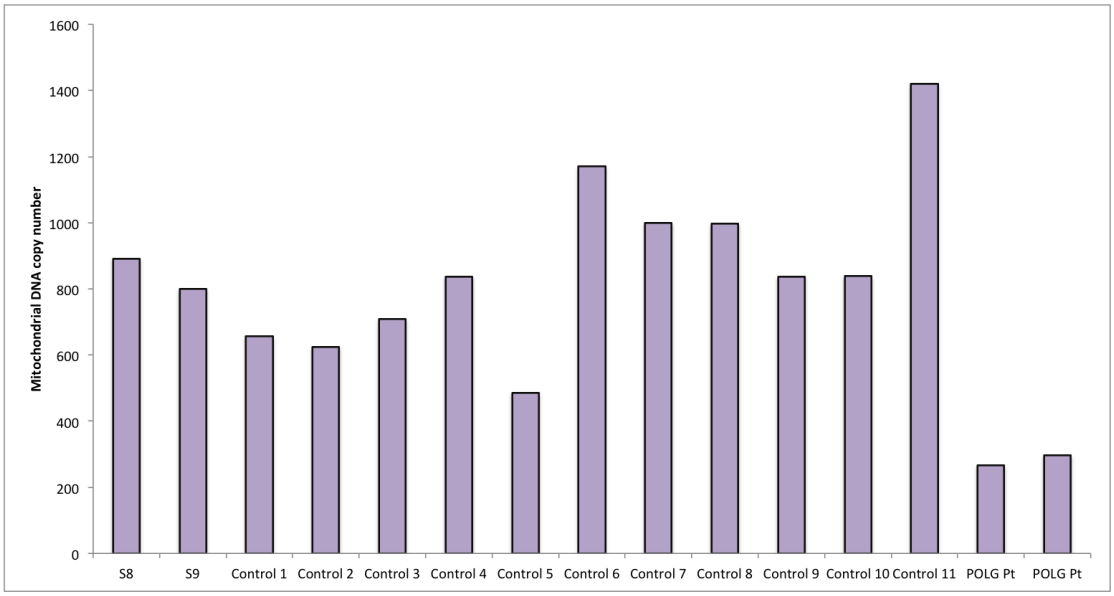


Figure 4:5: mtDNA copy number in patients S8 and S9, compared to controls and POLG deficient patients

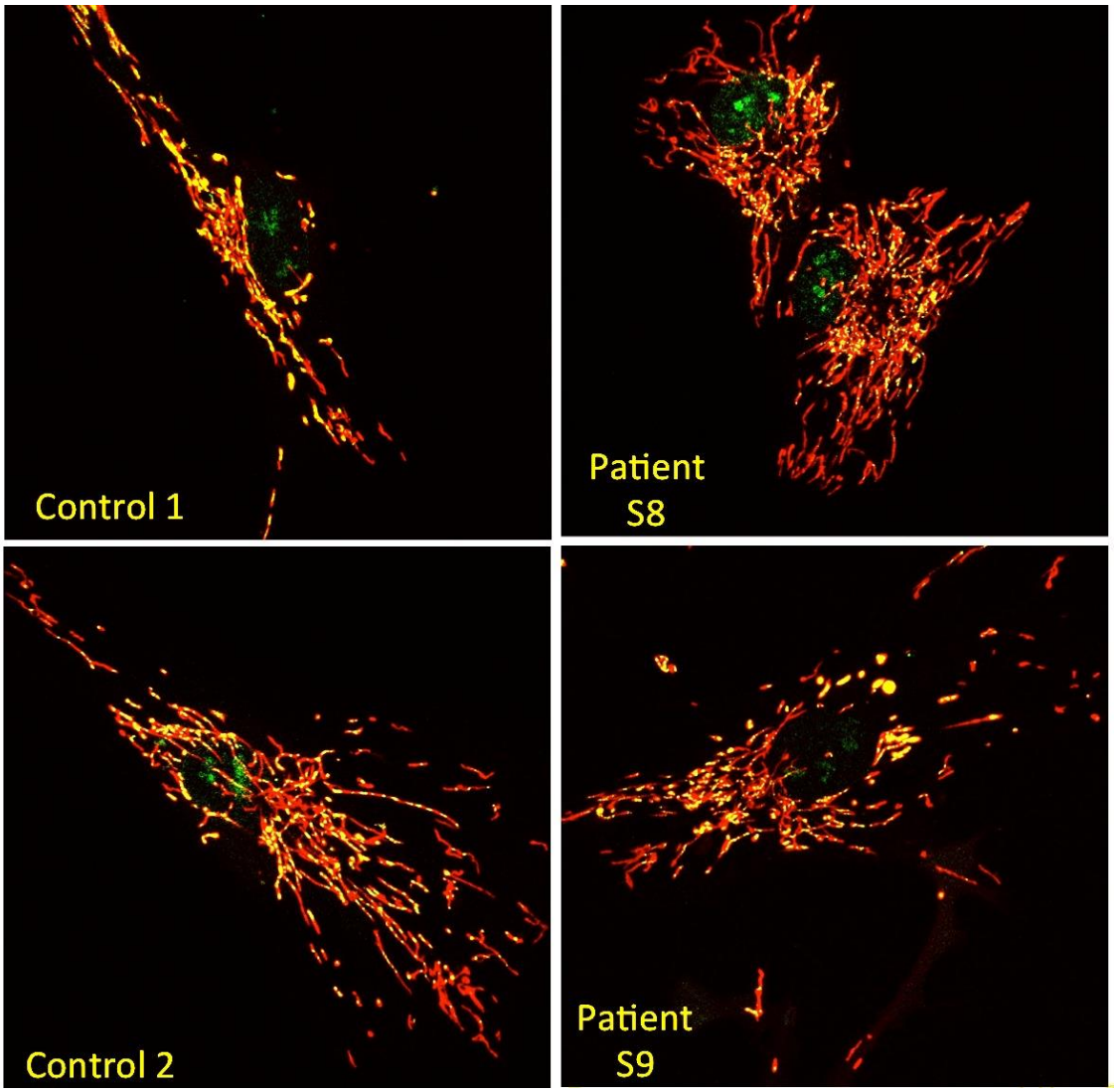


Figure 4:6: Confocal microscopy of TMRM/PicoGreen stained live cells in patients S8 and S9, compared to controls.

Red: TMRM (mitochondria), green: picogreen (DNA), yellow (green + red): picogreen + TMRM (DNA colocalising with mitochondria i.e. mtDNA)

4.3.10 Family 4: Discussion

In the case of family 4, two affected siblings born to Pakistani consanguineous parents with a mixed hereditary motor sensory neuropathy and recurrent cholestatic liver dysfunction were investigated. There was no evidence of respiratory compromise. In addition both siblings had evidence of significant mtDNA depletion in muscle while there were normal mtDNA copy numbers in liver. Two candidate genes identified by WES and homozygosity mapping were investigated as the possible cause of the disease. *IGHMBP2* was previously associated with SMARD1, a disease affecting the motor neurones in infants causing respiratory failure. The encoded protein has a helicase domain and is a DNA/RNA helicase and therefore could possibly be linked to mtDNA maintenance. No appreciable reduction in IGHMBP2 protein levels was found on Western blotting compared to controls.

We also investigated a gene which had not been linked to human disease before but where a mouse model demonstrated a similar phenotype. We were able to establish a collaboration with the group who had developed a knock out *Scyl1* mouse model. In this mouse model, mtDNA depletion was not found in a number of tissues investigated, which included muscle and liver. In addition a collaborator developed a mouse with the homozygous mutation found in these patients (*Scyl1*- a mouse model) but the mouse did not display a motor phenotype. At this point *SCYL1* seemed an unlikely cause of the patients' disease.

At the end of 2014, Cottenie *et al* published a study stating that 'truncating and missense mutations in *IGHMBP2* cause Charcot-Marie Tooth disease type 2' (Cottenie et al. 2014). One of the *IGHMBP2* mutations found here (c.1591C>A, p.Pro531Thr) was found in a Pakistani family (as a compound heterozygous mutation) with c.1738G>A, p.Val580Ile mutation. Their 20-year-old patient also had trisomy 21 mosaicism and demonstrated distal muscle wasting and weakness in her lower limbs with less involvement of her hands. Nerve conduction studies confirmed an axonal neuropathy with both motor and sensory involvement. The motor phenotype of S8 and S9 correlated with the findings in this study. However none of the patients had liver dysfunction. In addition only a few patients who had SMARD1 due to *IGHMBP2* mutations had mild hepatomegaly but no firm evidence of liver disease (Jędrzejowska et al. 2014).

Schmidt *et al* (who described the mdf mouse) published a study in 2015 where they found that *SCYL1* mutations were associated with recurrent episodes of liver failure, mixed axonal peripheral neuropathy, cerebellar atrophy and ataxia (Schmidt et al. 2015). These patients had recurrent, severe episodes of liver failure, precipitated by febrile infection. The episodes occurred between infancy and 11 years of age, liver biopsy demonstrated focal bridging fibrosis and the patients had residual hepatosplenomegaly.

In the case of family 6, the affected siblings display both features of *IGHMBP2* and *SCYL1* deficient patients (i.e. axonal mixed neuropathy and liver disease) published thus far. Both genes lie in the same homozygous region in chromosome 11. At present none of the experimental evidence point towards a pathomechanism associated with either gene to explain the mtDNA depletion observed in S8 and S9. However one of the *IGHMBP2* mutations in these patients has been associated with disease in the Cottenie *et al* series. Since the patients are of consanguineous background, it is also plausible that they could have two diseases due to homozygous changes in two genes.

4.3.11 Patient S27: Case history

S27 was born at full-term by normal vaginal delivery to first cousin parents following an uneventful pregnancy. He did not require any resuscitation and the perinatal period was uneventful. However he had feeding difficulties since 6 weeks of age and his development was delayed at 4 months. His older sibling is well. At 1.5 years of age he developed four focal febrile seizures, but has been seizure free since 2 years of age. Examination at 4 years of age did not demonstrate dysmorphic features. His face appeared slightly asymmetric with prominence of the nasolabial furrow and smaller palpebral fissure on the left side. Upper limbs appeared hypotonic while deep tendon reflexes were normal on the right and brisk on the left side. His lower limbs were hypertonic and there was marked contraction of the tendo- Achilles on both sides with limitation of joint movements; deep tendon reflexes were exaggerated in both lower limbs with sustained ankle clonus and extensor plantar responses bilaterally.

Peripheral neurophysiology showed neurogenic change in bulbar and peripheral motor neurones with no evidence of a neuropathy; this was thought to be consistent with a neuronopathy. He was noted to have spastic lower limbs with a diplegic motor disorder, with a relative expressive language delay where relatively better language understanding was reported. Muscle biopsy showed some neurogenic features and reduced COX activity (0.006 reference range 0.014-0.034). MRI brain at 5 years of age showed marked cerebellar hypoplasia with a small pons and very small medulla with a thin cervical cord. Urine organic acids, plasma amino acids and blood and CSF lactate were normal.

Plasma creatine kinase, urate, glucose, lactate, ammonia, liver function, thyroid function, serum folate, vitamin B12, plasma amino acids and white cell enzyme studies were normal.

In this patient, as noted in chapter 3, three candidate genes were identified. *OGDH,* oxoglutarate (alpha ketogutarate) dehydrogenase, is a Krebs cycle enzyme which catalyses the conversion of 2-oxoglutarate (alpha-ketoglutarate) to succinyl-CoA and CO_2. A congenital deficiency in 2-oxoglutarate dehydrogenase activity has been reported to lead to hypotonia, metabolic acidosis and high lactate (Guffon et al. 1993). In Krebs cycle defects, it would be expected that there would be significant lactic acidosis with Krebs cycle metabolites in urine organic acids. S27 did not demonstrate these abnormalities and it was therefore considered to be unlikely to be the disease-causing gene.

4.3.12 Patient S27: Sanger sequencing and western blotting of FBXO7 and GTPBP1

Two genes *GTPBP1* and *FBXO7,* were explored. Homozygous variants in both these genes were confirmed on Sanger sequencing (Figure 4.7). Parental DNA was not available in this case. Western blotting was performed to assess the effect of the mutation on the protein level for both FBXO7 and GTPBP1 proteins. GTPBP1 was not significantly reduced, while FBXO7 protein was significantly reduced in the patient (S27) (Figure 4.7). (The western blot revealed two bands: a minor upper band present at similar levels in all samples, including the patient,

and a major lower band only seen in the controls. It is likely that the lower band is FBXO7, whereas the upper band is a nonspecific band). Given this finding and its previous association with a neurological disease, *FBXO7* was considered to be the likely candidate gene and its function was further explored.

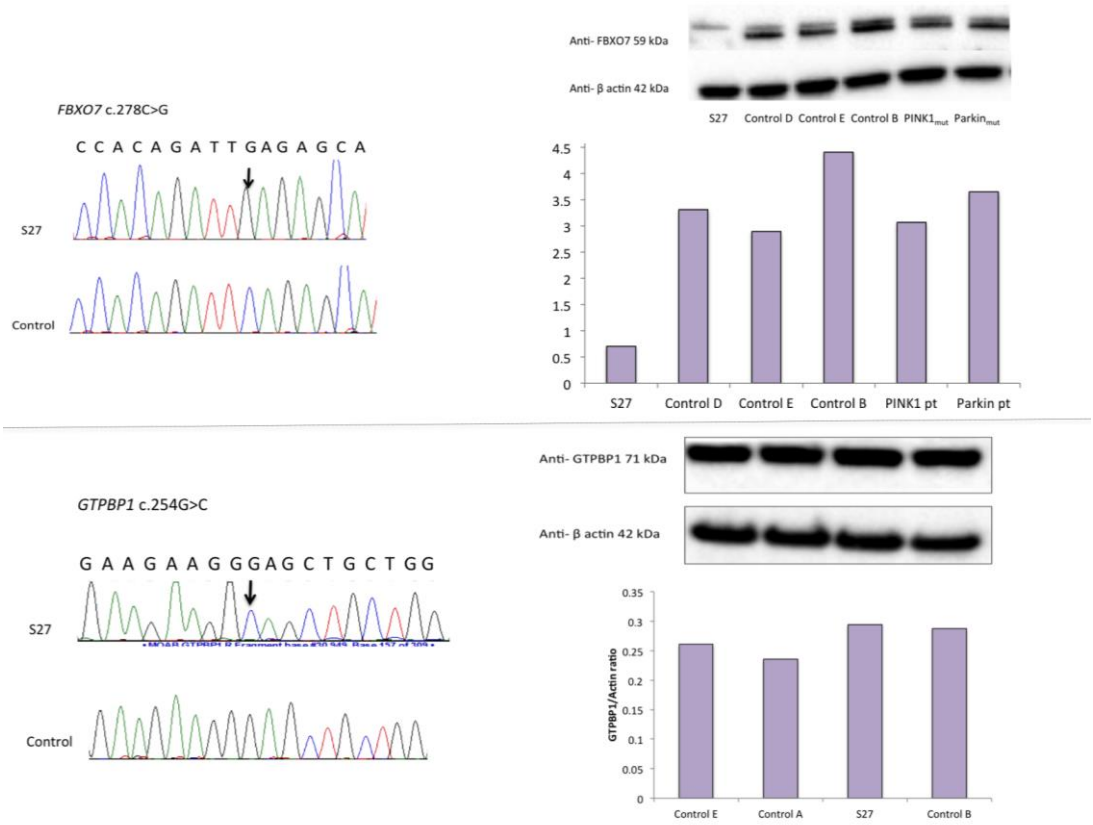


Figure 4:7: Sanger sequencing and Western blot findings in S27.
Left: Sanger sequencing of *FBXO7* c.278C>G and *GTPBP1* c.254G>C in patient S27 and a healthy control. Right: Western blot of S27 fibroblast extracts compared to controls, using anti-FBXO7 and anti-GTPBP1.

4.3.13 Patient S27: *FBXO7* as a candidate gene

The homozygous *FBXO7* nonsense mutation (c. 278C>G, p.Ser93X) found in patient S27 was not present in dbSNP, exome variant server or 1000 genome databases and it does not appear in the ExAC database. The mutation terminates this protein at amino acid 90. FBXO7 is a F-box protein which forms components of modular E3 ubiquitin protein ligases called SCFs (Skp1–Cul1–F-box) and functions in phosphorylation-dependent ubiquitination. The *FBXO7* gene lies in the 22q12.3 chromosomal location and its main isoform encodes a 522 amino acid protein while the second isoform encodes a 443 amino acid protein.

It contains a ubiquitin-like domain at the N-terminus which interacts with Parkin, followed by a Cdk6 domain, a dimerization domain and an F-box domain (Skowyra et al. 1997; Winston et al. 1999). At the C-terminal end of the protein there is a proline rich region (Figure 4.8). N-terminal

ubiquitin-like domain is present in the main FBXO7 isoform but not in the second isoform. In skin fibroblasts, isoform 1 is the more abundant form (T. Zhao et al. 2011).

When *FBXO7* isoforms 1 and 2 are over expressed, cells are protected against stress. However if p.Thr22Met, pArg378Gly or p.Arg498X mutated FBXO7 was over expressed, stress-induced toxicity was aggravated. In addition if any form of FBXO7 isoform was overexpressed (i.e. mutated or wild type), it resulted in formation of FBXO7 aggregates in mitochondria with the mutant FBXO7 cells forming more aggregates than the wild type FBXO7 (Zhou et al. 2015). Also FBXO7 has been found to act together with Parkin and PINK1 to induce mitophagy in response to mitochondrial depolarisation (Burchell et al. 2013). Furthermore Fbxo7 expression has been described to rescue the disease phenotype of *Drosophila* Parkin mutants (Burchell et al. 2013).

FBXO7 mutations have been identified as the cause of autosomal recessive early-onset Parkinson's disease, a disease similar to that caused by mutations in *PINK1* or *PARK2* genes. *FBXO7* mutations were first associated with disease in an Iranian family where affected family members presented with parkinsonian-pyramidal syndrome symptoms, in the third decade of life (Shojaee et al. 2008). All patients had a preceding phase where they developed spasticity mostly in the lower limbs, exaggerated deep tendon reflexes and a positive Babinksi sign.

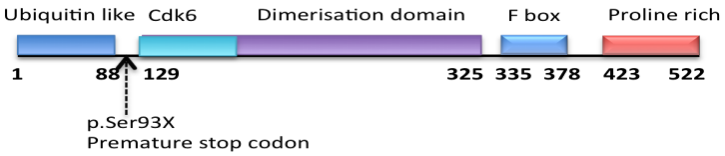


Figure 4:8: Schematic representation of the FBXO7 protein (isoform 1).
The position of the p.Ser93X change found in S27 in demonstrated.

4.3.14 Patient S27: MFN1 ubiquitination and OXPHOS proteins

For this experiment, cells from patients with known *PARK2* (homozygous deletion affecting exon 2 and 3) and *PINK1* (homozygous PINK1 p.Try90Leufsx12) mutations were used as disease controls, since FBXO7 has been linked to the mitophagy pathway (Burchell et al. 2013). MFN1 and MFN2 ubiquitination appeared reduced compared to healthy controls but not to the extent which the disease controls, PINK1 and Parkin deficient cells, displayed reduced ubiquitination. In the latter two cell cultures the ubiquitinated band appeared virtually absent (Figure 4.9A).

There were no differences in OXPHOS protein levels using antibodies to OXPHOS subunits {NDUFB8 (complex I), SDHB (complex II), UQCRC2 (complex III), MT-CO2 (complex IV) and ATP5A (complex V)} compared to controls on Western blotting (Figure 4.9B).

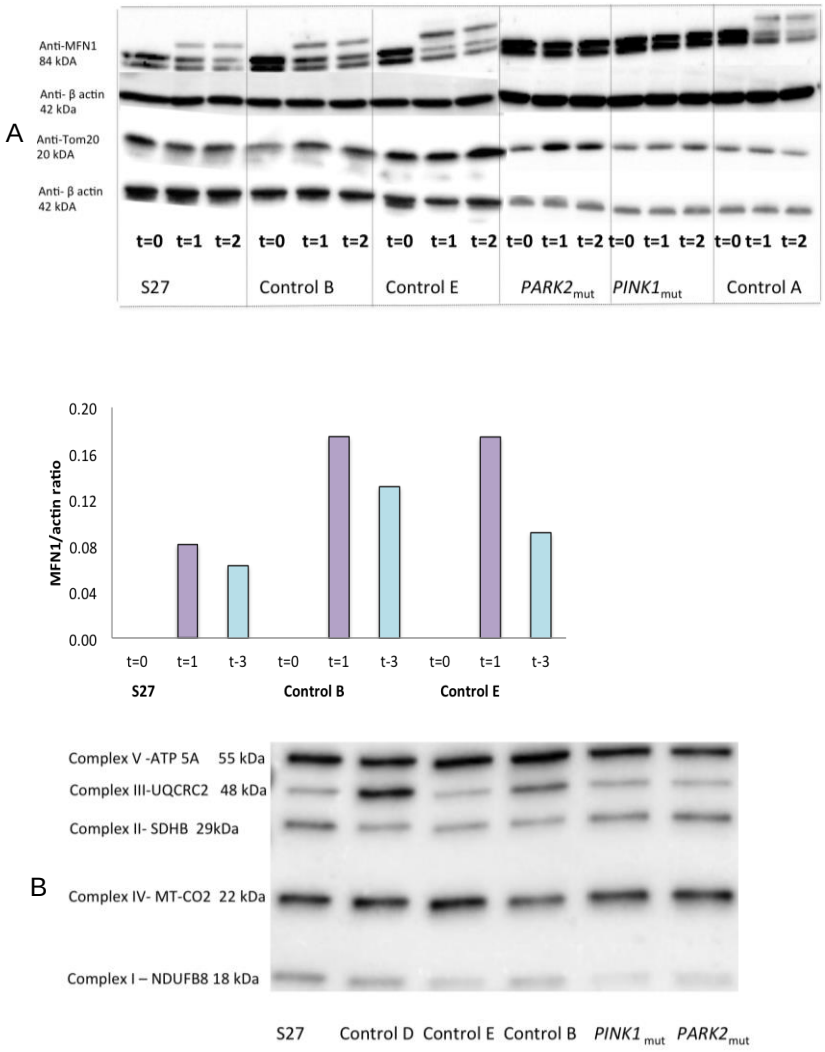


Figure 4:9: Findings in CCCP treated fibroblast cells and basal OXPHOS protein levels in fibroblast cells.

(A) Fibroblast cells were treated for 1 hour (t=1) and 2 hours (t=2) with 20 µmol CCCP. Untreated cells are denoted as t=0. Western blot was performed with Anti-MFN1 and Anti-TOM20 antibodies. The ubiquitinated band is the top band among the three bands in the anti-MFN1 blot. Below: graphical representation of MFN1 (ubiquitinated band)/actin ratios) in controls and S27. (B) Immunoblotting for OXPHOS subunits in patient S27, controls, PINK1 and Parkin deficient patients.

4.3.15 Patient S27: Fibroblast cytochrome *c* oxidase activity and mtDNA quantitation

Cytochrome *c* oxidase activity did not appear to be decreased compared to controls (cells grown in galactose, mean±SD 0.640±0.279 k/min/mg). For S27 COX values were 0.891 k/min/mg. COX activity (cells grown in galactose medium, normalised to citrate synthase activity) for controls (n=4) was mean±SD 0.014±0.006. In S27 COX/citrate synthase activity was 0.019. Mitochondrial DNA copy number in fibroblast cells appeared elevated compared to controls (Figure 4.10).

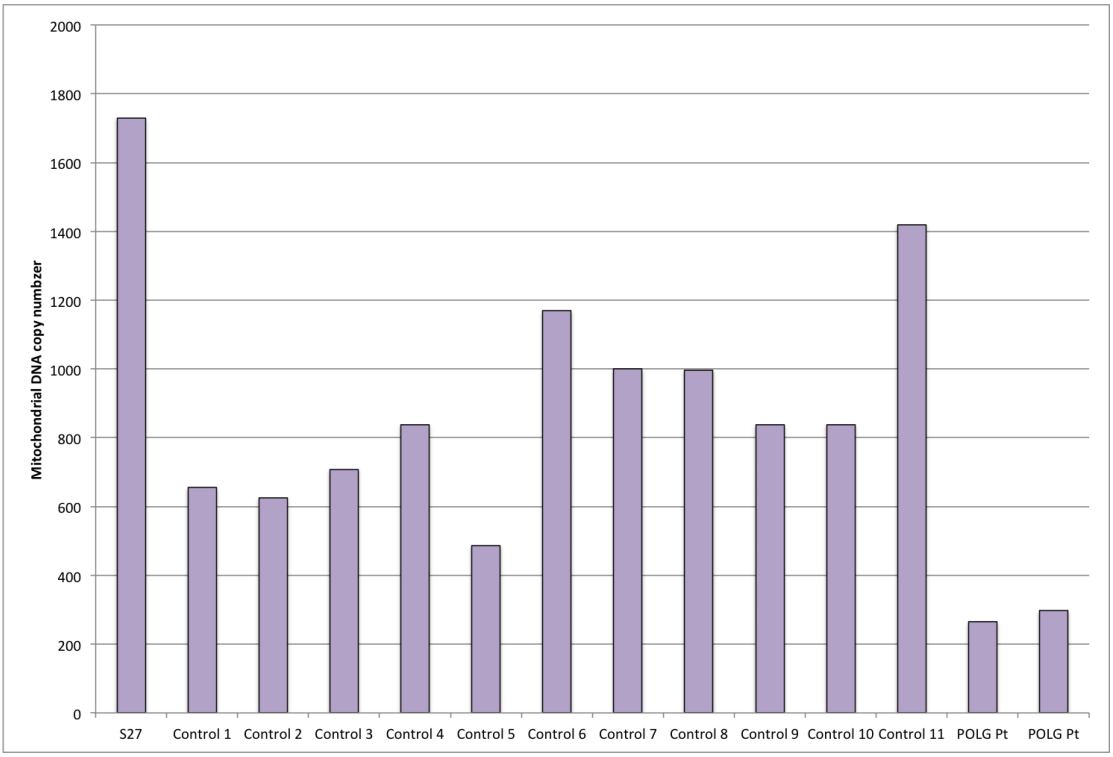


Figure 4:10: mtDNA copy number in patient S27, compared to controls and POLG deficient patients

4.3.16 Patient S27: Immunocytochemical staining of fibroblasts for MT-CO1 and TOM20

There was no apparent reduction staining in the COX subunit MT-CO1 and outer membrane protein TOM20 compared to controls (Figure 4.11).

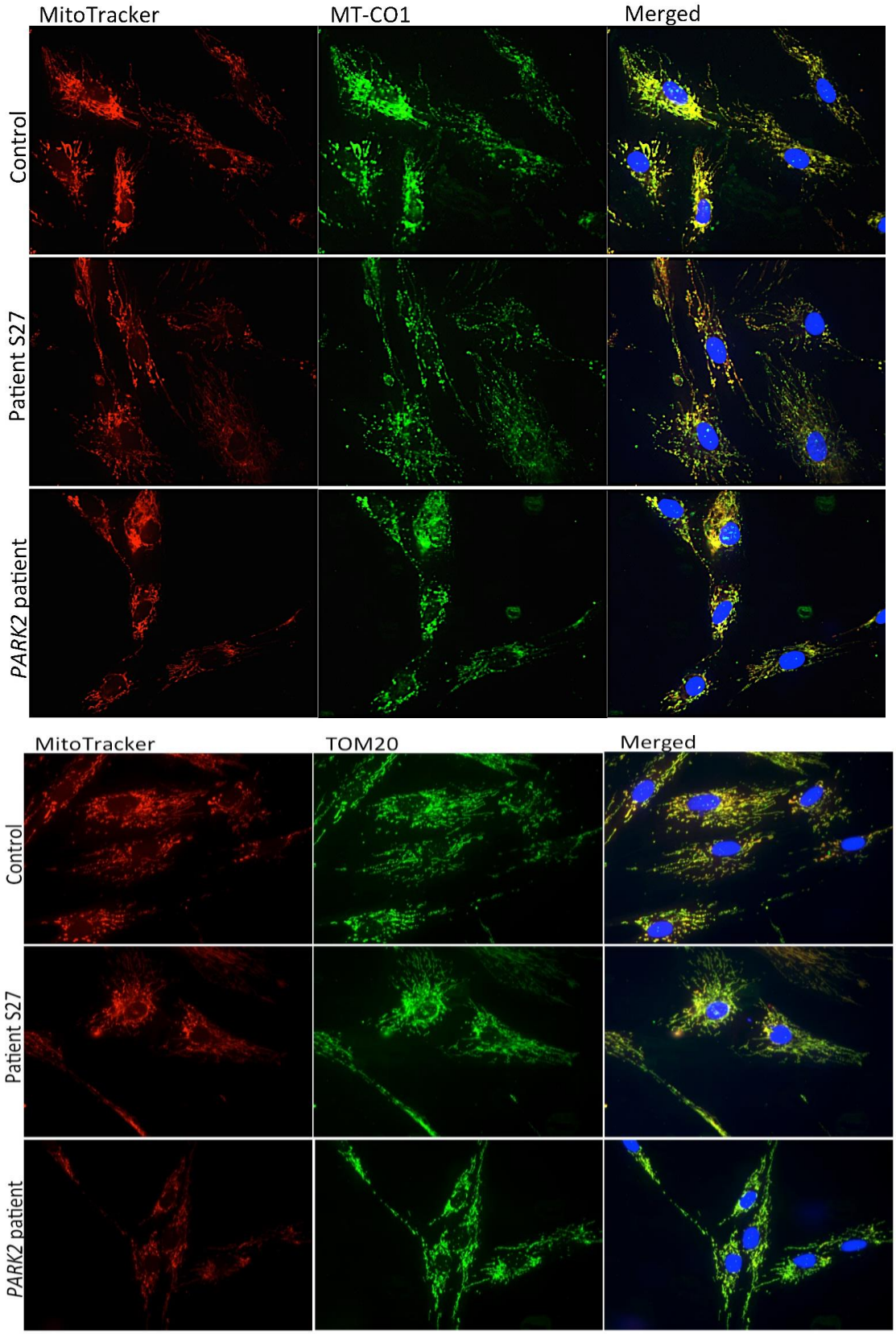


Figure 4:11: Immunocytochemistry of S27, *PARK2* patient and control fibroblast cells

Above: immunocytochemistry of fibroblast cells using MT-CO1 antibody and MitoTracker Red. Below: immunocytochemistry of fibroblast cells using TOM20 antibody and MitoTracker Red.

4.3.17 Patient S27: Mitochondrial morphology in fibroblast cells

Mitochondrial morphology did not appear different compared to control cells on Mitotracker green live cell staining of fibroblast cells (Figure 4.12).

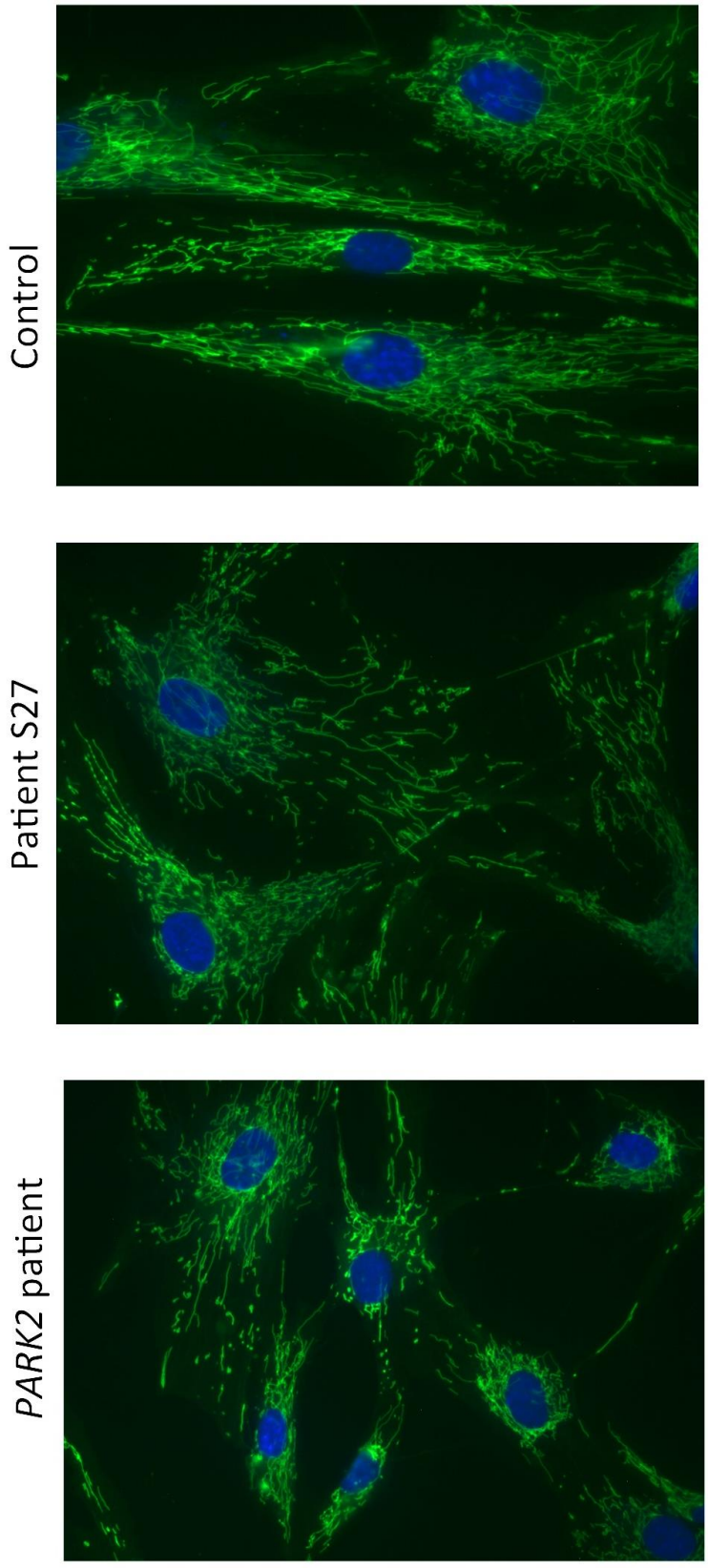


Figure 4:12: MitoTracker Green staining of mitochondria in live cells in S27, a *PARK2* patient and a healthy control.

4.3.18 Patient S27: Discussion

Here, a patient presenting with diplegic spastic motor disorder from infancy, pontocerebellar hypoplasia and evidence of a motor neuronopathy was investigated. He was found to have a homozygous nonsense *FBXO7* mutation which terminates the protein at 93 amino acids, likely leading to nonsense mediated decay as evidenced by Western blot data evidenced in Figure 4.7. Previous *FBXO7* mutations causing early onset Parkinson's disease were either missense mutations (p.Thr22Met, p.Arg378Gly) or a nonsense mutation at the C terminal end (p.Arg498X).

The focus of the studies described above was to uncover a phenotype in the patient cells and to assess if several different mitochondrial pathways (OXPHOS, morphology, mitophagy, mtDNA maintenance) were affected by the *FBXO7* mutations. There was evidence that MFN1 ubiquitination was reduced but not to the extent seen in the PINK1 and Parkin deficient cells. Reduced ubiquitination of MFN1 is in keeping with a previous study using FBXO7 mutant cells from a Parkinson's disease patient (Burchell et al. 2013). However given that the mutation found here (homozygous nonsense mutation) is more severe, it would be expected to find a more severe loss of ubiquitination, akin to PINK1 and Parkin deficient cells. Recent evidence suggest that *FBXO7* mutations can lead to FBXO7 protein aggregation and impairment of FBXO7-linked ubiquitin proteasome system functions leading to neuronal degeneration (Zhou et al. 2016). The effect of this homozygous mutation on these two functions has not been investigated here.

The elevation of mtDNA copy number in patient fibroblasts compared to control fibroblasts was an interesting observation. One possibility is that if there was impaired mitophagy, defective mitochondria would accumulate and therefore would result in elevated mtDNA copy number. Given the previous link to a neurodegenerative disease, *FBXO7* is the likely causative gene in this patient. S27, presented here, has a novel phenotype which has not previously been linked with *FBXO7* mutations but the pathogenicity of the mutation in this case has not been fully established here.

4.3.19 Patient S71: Case history

A patient presenting with motor delay, faltering growth, cataracts and cardiomyopathy was investigated. S71 was born at term weighing 2.9 kg to unrelated white European parents and she has a maternal half brother who is well. Initially she was well but at six weeks cataracts were detected. Until 4 months of age her weight was stable along the 0.4th centile but thereafter faltering growth was noted in her first year. At 13 months motor developmental delay was observed and her parents noticed coarse crackling noises when she was breathing. She was referred for an echocardiogram, which showed biventricular hypertrophy with good systolic function and no evidence of left ventricular outflow tract obstruction. Her MRI brain was normal. She fed poorly and she was supplemented with nasogastric feeding. At 16 months she was admitted with respiratory failure following pneumonia. She died a few days later. Muscle histology showed excess lipid and slow fibre predominance with an isolated COX deficiency 0.012 (0.014-0.034). Fibroblast COX activity measured in the diagnostic laboratory in Oxford, was normal (84 nmol/mg protein/min, reference range 30-90). Due to the Sengers syndrome phenotype the *AGK* gene was sequenced and no mutations were found (Dr Marta Kanabus, former PhD student at UCL Great Ormond Street Institute of Child Health).

4.3.20 Patient S71: ATPase family, AAA domain containing 3A (*ATAD3A)* as a candidate gene

ATAD3A and ATAD3B are mitochondrial inner membrane proteins and both are located in tandem on chromosome 1p36.3. All multicellular organisms possess *ATAD3A*, whereas *ATAD3B* is only present in the human genome (Hubstenberger et al. 2008).

In addition members of the AAA proteases family contain a proteolytic domain which is well conserved, however ATAD3A does not contain this domain and therefore is unlikely to function as a mitochondrial protease. The ATAD3A protein has been isolated from the inner mitochondrial membrane in mouse liver mitochondria (Da Cruz et al. 2003) and is postulated to stabilise mitochondrial nucleoids (He et al. 2007). He *et al* also found that *ATAD3* knockdown resulted in impaired mitochondrial translation.

There is a high degree of sequence similarity between *Drosophila atad3a* and human *ATAD3A*; in homozygous *atad3a Drosophila* mutants larval growth was arrested (Gilquin et al. 2010). Studies in a human steroidogenic cell line NCI-H295R (where only ATAD3A but not ATAD3B is expressed) suggested that the ATAD3A protein is required for steroidogenesis, probably by cholesterol channeling at contact sites (Rone et al. 2012). In addition transfection of U373 cells with plasmids encoding mutant *ATAD3A* resulted in mitochondrial fragmentation. When these cells were treated with *DRP1* siRNA (i.e. knock down of mitochondrial fission regulator, DRP1), the fragmented phenotype was rescued suggesting that *ATAD3A* loss of function is detected by the fission apparatus (Gilquin et al. 2010). ATAD3A has also been found to interact with ccdc56 (MITRAC12) a subunit of the COX assembly complex (Ban-Ishihara et al. 2015).

ATAD3A is tethered to the inner membrane and has been found to have contact sites with the outer membrane. The N terminal domain (amino acids 1-245) is located in the intermembrane space whereas the transmembrane part of the protein (amino acids 246-264) is located in the inner membrane and the C terminal AAA+ ATPase domain is found in the mitochondrial matrix. The C terminal domain contains the mitochondrial targeting sequence. ATAD3A is therefore thought to have an interactive role between membranes and maybe involved in mitochondrial dynamics and cholesterol trafficking. This gene has not previously been associated with human disease. In this section, the effect of the mutation on protein level and mitochondrial function is investigated.

4.3.21 Patient S71: Sanger sequencing and Western blot

Sanger sequencing of the c.655C>T, p.Arg219X and c.1844_1846del, mutations demonstrated that the parents each were heterozygous (Figure 4.13A). The mutations were not present in dbSNP, exome variant server or the 1000 genome project databases, at the time of analysis. The c.1844_1846delAGA mutation was present in the ExAc database as a very rare variant with a frequency of 0.000008. The c.655C>T mutation leads to a stop codon at amino acid 219 in this 634 amino acid protein. In addition *in silico* analysis of the AGA deletion at c.1844_1846 revealed a loss of a lysine residue at p.616. Western blot of ATAD3A protein revealed reduced levels in the patient (Figure 4.13B). TFAM, the most abundant protein in nucleoids, was not reduced in the patient.

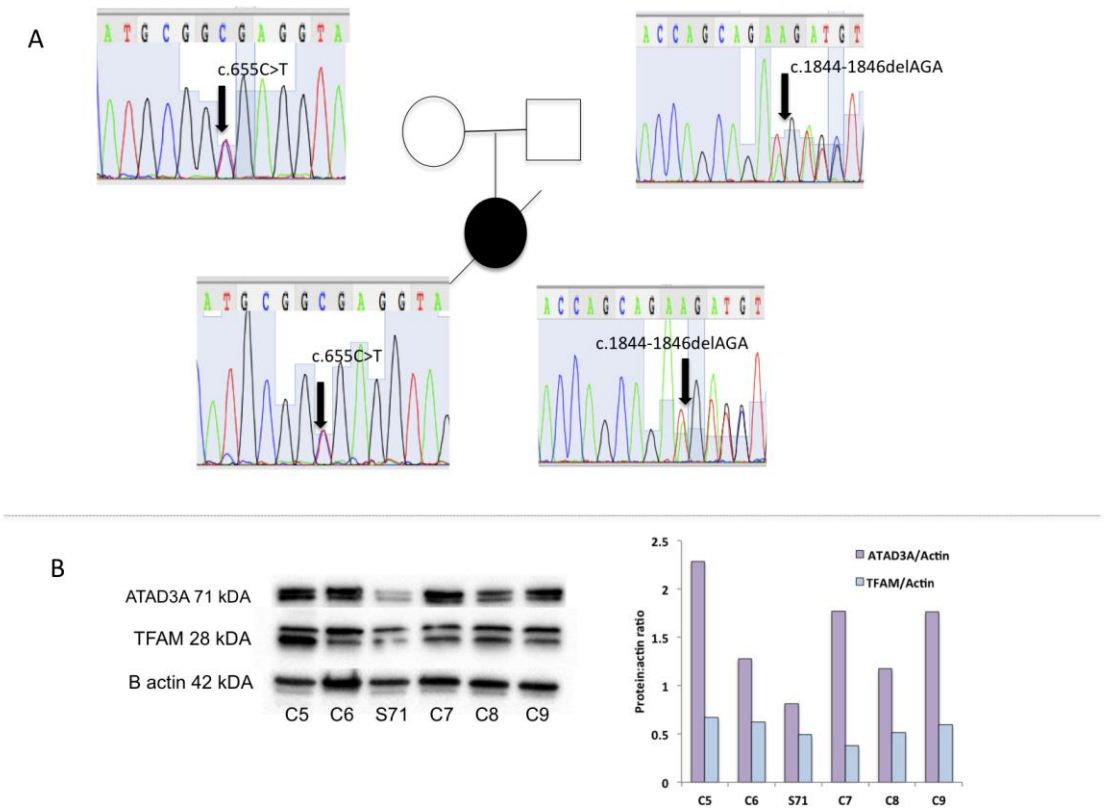


Figure 4:13: Findings in the patient with *ATAD3A* mutations

(A) Sanger sequencing of *ATAD3A* mutations (B) Western blot of ATAD3A and TFAM protein; right quantitation of Western blot data where proteins ATAD3A and TFAM are expressed as a ratio to actin. S71: patient and C5- C9: controls.

4.3.22 Patient S71: PicoGreen/TMRM staining of nucleoids

PicoGreen/TMRM staining was used to visualise mitochondrial nucleoids. No significant differences were noted in the patient (S71) cells (Figure 4.14).

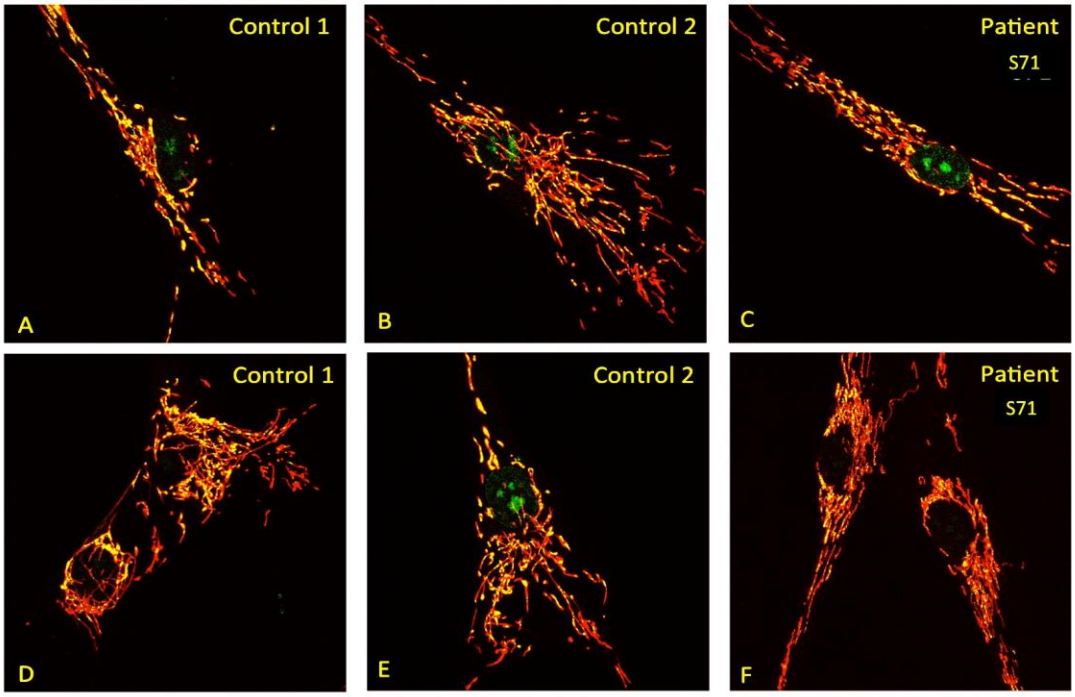


Figure 4:14: Confocal microscopy of PicoGreen/TMRM stained of live fibroblast cells.
Red: TMRM (mitochondria), green: picogreen (DNA), yellow (green + red): picogreen + TMRM (DNA co-localising with mitochondria i.e. mtDNA).

4.3.23 Patient S71: Mitochondrial depletion studies in patient muscle

Residual mtDNA in patient muscle was 61% of the mean normal level; i.e. there was no evidence of mtDNA depletion.

4.3.24 Patient S71: Mitochondrial translation assay

No reduction in mitochondrial translation products was observed (Figure 4.15).

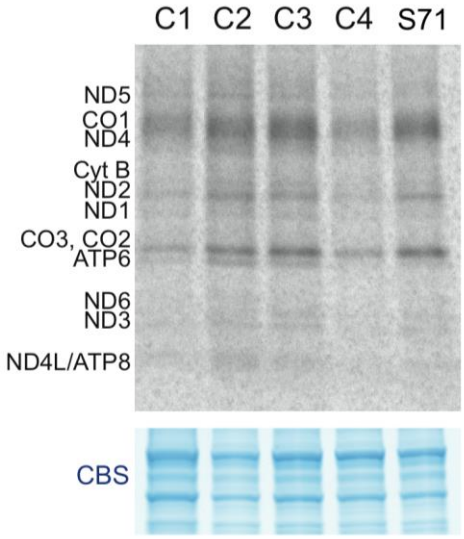


Figure 4:15: Metabolic ^{35}S Methionine labelling of mitochondrial translation products in patient S71 compared to controls C1-C4.

4.3.25 Patient S71: Discussion

In this section a patient with cataracts, cardiomyopathy, developmental delay and 3-methylglutaconic aciduria was investigated. Her clinical phenotype was consistent with Sengers syndrome, which has been associated with AGK deficiency, an acyl glycerol kinase involved in biosynthesis of cardiolipin in the inner mitochondrial membrane. As discussed in the beginning of this section the postulated genetic cause, *ATAD3A*, is also an inner membrane protein.

There was evidence of mutation segregation within the family and clear reduction in the ATAD3A protein levels on immunoblot. ATAD3A, a mitochondrial membrane protein, has been postulated to stabilise mtDNA-protein complexes (nucleoids) and have a role in protein translation (He et al. 2007). These functions were investigated here and there was no evidence of disruption of the nucleoids, mtDNA translation or mtDNA depletion. These findings indicate that ATAD3A has another mitochondrial function.

The patient had 3-methylglutaconic aciduria which is seen in phospholipid remodelling defects and other disorders where the mitochondrial membrane is disrupted. For example Barth syndrome due to Tafazzin deficiency, Sengers syndrome due to AGK deficiency, and MEGDEL syndrome due SERAC1 deficiency are all characterised by 3-methylglutaconic aciduria, and are mitochondrial membrane disorders (Wortmann et al. 2013). As discussed previously, ATAD3A is also required for steroidogenesis, probably by cholesterol trafficking at mitochondrial membrane contact sites (Gilquin et al. 2010). The presence of 3-methylglutaconic aciduria indicates that mitochondrial membrane disruption is a likely patho-mechanism in ATAD3A

deficiency akin to AGK, CLPB, SERAC1 and TAZ deficiencies (Kanabus et al. 2015; Wortmann et al. 2012). The finding that ATAD3A, was deficient on Western blot in patient S71 who has a Sengers syndrome phenotype including 3-methylglutaconic aciduria (present in other mitochondrial membrane disorders), is supportive of *ATAD3A* mutations being the cause of disease in this patient.

4.3.26 Family 10: Case histories

S20 and S21 are siblings born to unrelated white European parents. S20 is now 10 years old and was born at term after a normal pregnancy. The neonatal period was uneventful, but by four months developmental delay, hypotonia, hypermetropia, intermittent squint and nystagmus had been noted. Since 8 years of age he has had frequent vomiting and difficulty in maintaining his weight necessitating gastrostomy feeding. He has a myopathic face and developed generalised seizures.

EMG and nerve conduction studies were normal. Blood and CSF lactate were both normal but COX activity was low in skeletal muscle at 0.009 (reference range 0.014-0.034). Transferrin electrophoresis was normal. CSF pyridoxal phosphate was normal but at the lower end at 12 nmol/L (reference range 10-37) with normal plasma pyridoxal phosphate. CSF 5-methyltetrahydrofolate was decreased at 52 nmol/L (reference range 72-172). Oral Vitamin B6 and folinic acid supplementation did not lead to seizure control. Serum alkaline phosphatase (ALP) was slightly elevated at 4.5 years at 441 u/L (reference range 150-380) and was normal at 318 u/L (reference range 175-420) at 8 years.

His younger brother S21 is currently six years old and was born at term by normal vaginal delivery. The neonatal period was uneventful except for transient jaundice. At 3 months 'eye rolling' was reported. At 9 months mild hypotonia and global developmental delay were noted. At 2 years he was found to have myopathic facies, intermittent convergent squint and hypermetropia. Plasma lactate was mildly elevated (2.3 mmol/L, reference range <2) but CSF lactate was normal at 1.4 mmol/L (reference range<2. CSF pyridoxal phosphate was low at 7 nmol/L (reference range 10-37) with normal plasma pyridoxal phosphate. Oral B6 and folinic acid supplementation was commenced and so far he has not experienced any seizures. Serum ALP was elevated at 480 u/L (reference range 145 - 320) at 1 year of age and remained elevated at 590 u/L (reference range 150-380) at 5 years. Muscle COX was decreased at 0.013 (reference range (0.014-0.034) with normal fibroblast COX activity (57, reference range 30-40). MRI brain scans demonstrated progressive cerebellar atrophy in both cases and are shown in Figure 4.16.

The third child in the family, a girl, is well, but the fourth child is currently one year old, has hypotonia, poor head control, nystagmus and an intermittent squint. Her ALP was elevated at 405 u/L (reference range 60-330) at 6 months.

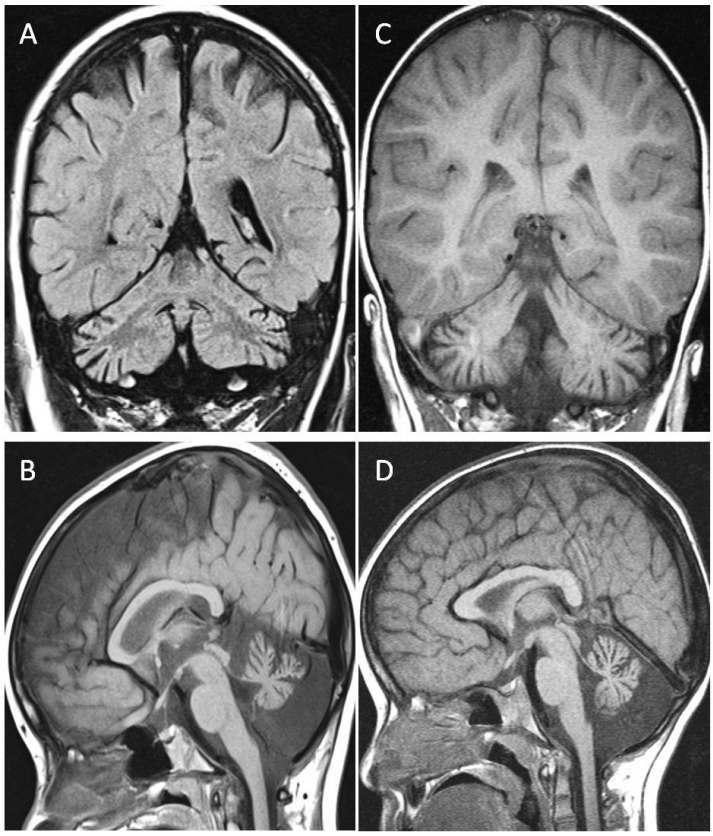


Figure 4:16: MRI brain scan findings in S20 and S21.

A&B Sagittal T1-weighted and coronal FLAIR MRI scans showing atrophy of the cerebellar hemispheres and vermis with relative sparing of the pons in S20. C&D Sagittal T1-weighted and coronal MRI scans showing atrophy/hypoplasia of the cerebellar hemispheres and vermis with relative sparing of the pons in S21.

4.3.27 Family 10: Phosphatidylinositol Glycan Anchor Biosynthesis Class W Protein (*PIGW*) as a candidate gene

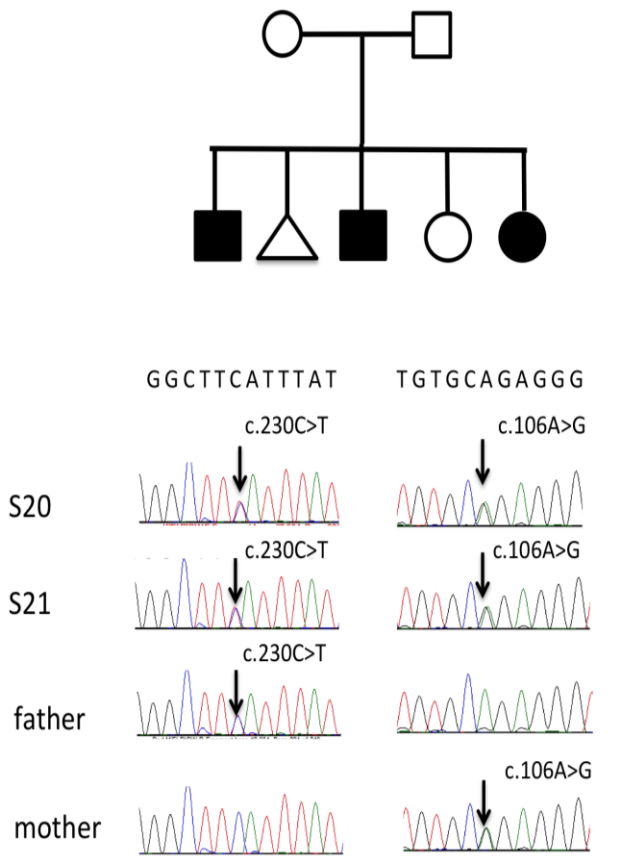
The *PIGW* gene is located on the chromosome 17q12 and encodes a 504-amino acid protein. It is involved in the glycosylation pathway in the synthesis of glycosylphosphatidylinositol (GPI) anchors. GPI anchors are glycolipids which tether proteins to the cell surface. PIGW acylates the inositol ring of phosphatidylinositol which forms the third step in the GPI anchor biosynthetic pathway (Murakami et al. 2003). The acylation of inositol is a critical step for bridging ethanolamine phosphate to the third mannose of the GPI anchor. The acylation step is, however, not critical for the addition of mannoses or the ethanolamine phosphate side-branch. Therefore, in PIGW deficiency it would be expected that the GPI anchor intermediates which accumulate do not contain the inositol-linked acyl chain, but possess mannose residues and the ethanolaminephosphate side-branch. It has been postulated that these intermediary metabolites activate GPI transamidase causing a high ALP i.e. hyperphosphatasia (Murakami et al. 2012; Chiyonobu et al. 2014).

In 2014 Chiyonobu *et al* identified compound heterozygous mutation in a Japanese patient with hyperphosphatasia (high serum alkaline phosphatase) and West syndrome (infantile spasms, electroencephalogram pattern termed hypsarrhythmia and mental retardation) (Chiyonobu et al. 2014).

4.3.28 Family 10: Sanger sequencing and amino acid conservation data

Sanger sequencing verified the mutations in S20 and S21 and that the parents were each heterozygous for one of the *PIGW* variants (c.106A>G:p.Arg36Gly and c.230C>T:p.Ser77Leu) (Figure 4.18). At the time of analysis the c.106A>G, p.Arg36Gly mutation was present with an allele frequency of 0.0005 (1000 genome database) and 0.0003 (Exome variant server), making it a rare allele. The c.230C>T, p.Ser77Leu mutation was not present in these databases. The arginine residue at 36 is conserved in mammals and Xenopus and the serine residue at 77 is conserved in mammals (Figure 4.17).

Subsequently their younger sister who appeared affected and was born after the WES was performed on S20 and S21 was found to have these *PIGW* mutations. The younger sister was born during the time when functional analysis was still in progress and parents did not seek prenatal diagnosis at the time.



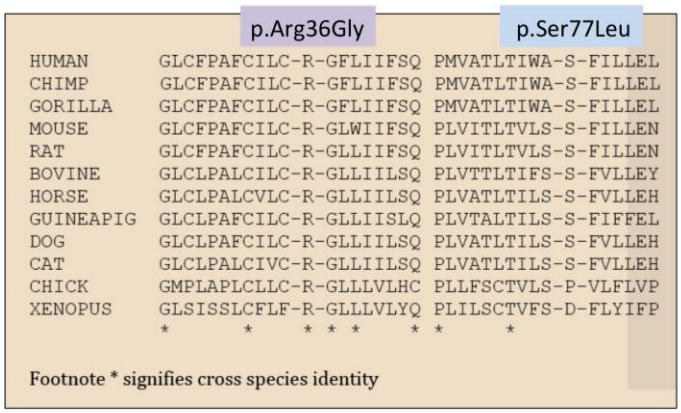


Figure 4:17: Sanger sequencing and amino acid conservation across species in S20 and S21.

Upper panel: pedigree of family 10 and Sanger sequencing of the *PIGW* mutations. Lower panel amino acid conservation across species of residues affected by the c.106A>G:p.Arg36Gly and c.230C>T:p.Ser77Leu variants.

4.3.29 Family 10: Flow cytometry of fibroblast cells for GPI-linked proteins

This work was performed entirely by a collaborator Dr Peter Krawitz, Institut für Medizinische Genetik und Humangenetik, Berlin, Germany. GPI linked proteins CD87-APC and CD73-PE-

Cy7 were reduced in S20 and S21 compared to healthy controls indicating a GPI anchor protein biosynthetic defect (Figure 4.18).

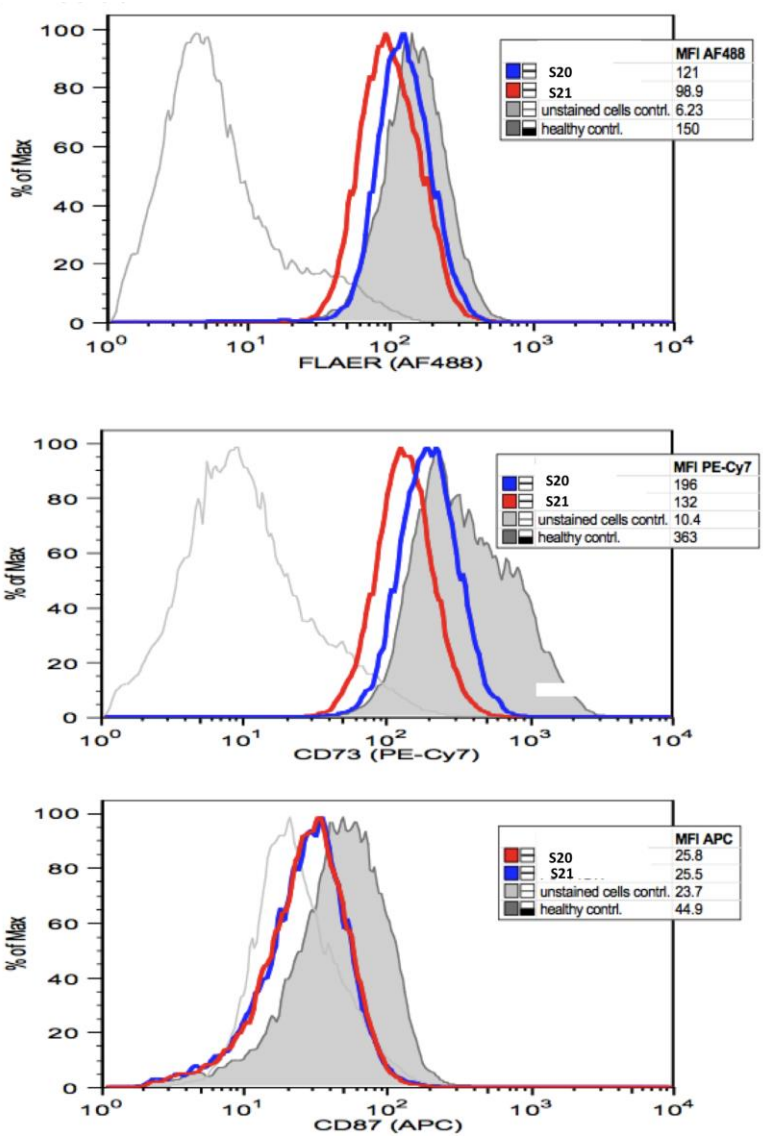


Figure 4:18: Flow cytometry of fibroblast cells. Top panel: FLAER is an Alexa Fluor® 488 labelled aerolysin, which is a protein which binds specifically to mammalian GPI anchors. In GPI anchor deficiencies, FLAER cannot bind.

Middle and lower panel: CD87 is a urokinase plasminogen activator-receptor and is a GPI anchored glycoprotein. CD73 (ecto-5'-nucleotidase) is a GPI anchored purine salvage enzyme. CD8 (APC) and CD73 (PE-Cy7) antibodies were used to detect surface expression of these proteins on fibroblast cells in patients S20 and S21 compared to unstained and healthy controls. The white plot with the light grey outline depicts unstained cells whereas the grey-filled plot depicts healthy controls.

4.3.30 Family 10: Discussion

This genetic finding was unexpected as it was hypothesised at the outset that the patients had a COX-deficient mitochondrial disease. Here evidence for pathogenicity of the *PIGW* mutations is supported by the clinical phenotype which is similar to the patient published in 2014 (Chiyonobu

et al. 2014), the presence of a raised ALP, and flow cytometry of fibroblast cells which demostrated deficient GPI anchor proteins.

Currently mutations in nine genes (*PIGA, PIGL, PIGM, PIGQ, PIGN, PIGO, PIGT, PIGV, PGAP2, PGAP3* and *PIGW)* which encode GPI biosynthetic proteins have been identified as disease causing (Ng & Freeze 2014). Affected patients have multisystem phenotypes which include developmental delay, intellectual disability, cerebellar atrophy, seizures and dysmorphism. This is the second report of *PIGW* mutations causing human disease.

Prior to the advent of NGS, glycosylation defects were thought to be under-diagnosed owing to its presentations mimicking mitochondrial disease (Briones et al. 2001). Given that both mitochondrial diseases and disorders of glycosylation have multisystem phenotypes this is not surprising. Indeed for the GPI anchor biosynthesis disorders and for disorders of O-glycosylation, transferrin electrophoresis, a common test available for screening for N-glycosylation disorders is usually normal, as was for the patients described here. COX deficiency is presumably a secondary defect although its precise mechanism has not been investigated here.

4.4 Conclusions and summary

In this chapter four different families and five possible genetic causes (*IGHMBP2, SCYL1, FBXO7, ATAD3A* and *PIGW*) were investigated further. Collaborations were formed to investigate these findings more thoroughly.

Of these there was definitive functional evidence for pathogenicity in the *PIGW* mutations. For the novel disease gene, *ATAD3A*, the majority of the functions previously postulated for the encoded protein were normal in the ATAD3A deficient patient cells. It seems probable that a role in cholesterol trafficking is the most likely function of ATAD3A. In the consanguineous family 4 it was possible that there were two disease genes (*IGHMBP2* and *SCYL1)* given the publications that arose during the time of the study. For the patient, with the homozygous nonsense *FBXO7* mutations, while there was a clear deficiency of FBXO7 on western blot making it the likely cause, there was no clear cellular phenotype to inform functional complementation studies. This was an interesting finding since it is a gene causing Parkinson's disease and here it was found in a childhood-onset neurological disorder.

Chapter 5

5. The study of mitochondrial cardiomyopathy in childhood

5.1 Introduction

Cardiomyopathy in childhood is aetiologically diverse, with identified causes including viral cardiomyopathy, inborn errors of metabolism such as mitochondrial diseases, fatty acid oxidation defects and lysosomal storage disorders, mutations involving sarcomeric proteins and Noonan's syndrome (Weintraub & Semsarian 2013). The worst outcomes have been reported in cardiomyopathy due to an inborn error of metabolism, where a mortality or heart transplantation rate of 57% has been found (Lipshultz et al. 2013). While there have been two large registries with epidemiological data on paediatric cardiomyopathy in the last decade, its genetic aetiology still requires elucidation (Nugent et al. 2003; Lipshultz et al. 2003). Having a causal diagnosis is crucial for prognostication, detection of extra-cardiac symptoms and prenatal diagnosis (Ware 2011). Previously, patients with inherited cardiomyopathy have only had the opportunity for limited genetic screening of a few genes associated with a specific disease phenotype; with the advent of next generation sequencing technologies such as whole exome and genome analysis, the genetic diagnostic rate for inherited cardiomyopathies is improving vastly, at least in a research laboratory setting.

Mitochondrial diseases affect 1 in 5000 individuals and at present there are limited therapeutic options for patients (Kanabus et al. 2013). As described in chapter 1, mitochondria are dynamic organelles which produce ATP by the concerted actions of proton pumping and electron transfer by the OXPHOS enzyme complexes I-V. The individual subunits which constitute the OXPHOS complexes are encoded by both the nuclear genome and the maternally-inherited mitochondrial genome with the exception of complex II which is solely nuclear-encoded. Where there is failure of OXPHOS due to a nuclear or mtDNA mutation, organs such as the heart are particularly vulnerable to the resulting energy failure because of high bioenergetic requirements.

While cardiac dysfunction is better recognised in some mitochondrial syndromes such as KSS, Barth syndrome and MELAS, new disease mechanisms have been revealed by next generation sequencing. Nuclear genes previously associated with mitochondrial cardiomyopathy are shown in table 5.1. More recently reported causes of childhood mitochondrial cardiomyopathy include *AARS2, MRPL44, MTO1 and GTPBP3* mutations (Gotz et al. 2011; Carroll et al. 2013; Ghezzi et al. 2012; Kopajtich et al. 2014).

Table 5.1: Nuclear genes previously associated with mitochondrial cardiomyopathy

Genetic cause	Type of cardiomyopathy predominantly found
TAZ	DCM
AGK	HCM
CLPB	HCM
DNAJC19	DCM

Genetic cause	Type of cardiomyopathy predominantly found
SLC25A3	HCM
FXN	HCM
DGUOK	DCM
ANT1	HCM
NDUFV2	HCM
NDUFA11	HCM
NDUFS2	HCM
NDUFS8	HCM
NDUFAF1	HCM
SCO2	HCM
COX10	HCM
COX15	HCM
SDHA	DCM
TMEM70	HCM
TSFM	HCM
MRPS22	HCM
MRPL44	HCM
AARS2	HCM
GTPBP3	HCM
MTO1	HCM
SLC25A3	HCM
ACAD9	HCM
ELAC2	HCM

DCM: dilated cardiomyopathy, HCM: hypertrophic cardiomyopathy

This aim of this study was to:

(a) Identify a paediatric cohort with suspected mitochondrial cardiomyopathy

(b) Use stratified genetic testing to uncover the underlying genetic causes in this cohort

(c) To study the clinical, biochemical, echocardiographic and histological characteristics in patients with suspected paediatric cardiomyopathy

5.1.1 Declaration of work

I performed filtering of WES data for all COX deficient patients, collated all clinical data and analysed all demographic and clinical data. MtDNA sequencing and analysis was performed by Cathy Woodward and Mary Sweeney at the diagnostic laboratory, Neurogenetics Unit, National Hospital for Neurology and Neurosurgery, London. Sanger sequencing of mutations found by WES was carried out by Dr Elisa Fassone (patients with isolated complex I deficiency) and Dr Rojeen Shahni (multiple respiratory chain enzyme deficiencies). Barth syndrome testing was carried out in the genetics laboratory at Bristol University NHS Trust. Candidate gene

sequencing was performed by Sanger sequencing, in the diagnostic laboratory or by Dr Rojeen Shahni in two cases.

5.2 Methods

Ethical approval for the study was obtained from the National Research Ethics Committee London Bloomsbury, UK. All studies were performed after informed parental consent.

5.2.1 Study population

A cohort of patients with mitochondrial cardiomyopathy was identified at a tertiary referral centre Great Ormond Street Hospital UK. The study protocol involved recruiting all patients with cardiomyopathy due to suspected mitochondrial disease referred to the mitochondrial or cardiac transplant service. The type of cardiomyopathy was classified using echocardiography by a single paediatric cardiologist in a tertiary cardiac transplant centre for children. The echocardiographic data was initially collated by me but a comprehensive echocardiographic evaluation was carried out by a single paediatric cardiologist Dr Jacob Simmonds, at Great Ormond Street Hospital.

A proportion of this cohort was recruited retrospectively (n=13) by recruiting cases from the cardiac transplant service where a genetic diagnosis of mitochondrial disease was made and cardiomyopathy was also noted echocardiographically. The remaining patients were recruited prospectively (n=17). For the cases recruited prospectively a diagnosis of possible or probable mitochondrial disease was classified where there was a canonical mitochondrial syndrome or by using mitochondrial disease criteria (Morava et al. 2006) and the presence of cardiomyopathy was established by a echocardiographic evaluation by a paediatric cardiologist.. Clinical data including age of onset, clinical features, age of death, echocardiographic findings, lactate levels and RCE activities were collated using a structured database. Some of the COX deficient patients with cardiomyopathy in chapter 2 fulfilled the recruitment criteria for this study and are therefore also included here. Systematic genetic testing was performed as described below.

5.2.2 Genetic diagnostic approach

Mitochondrial DNA analysis

MtDNA was sequenced in muscle in all patients, unless there was an obvious nuclear candidate gene. MtDNA analysis was carried out by Sanger sequencing, GeneChip® resequence analysis or polymerase chain reaction-fragment length polymorphism for mtDNA deletions (Rahman et al. 2012).

Candidate gene sequencing

In some cases where the clinical presentation suggested a specific genetic cause, single genetic testing was performed. This included screening for: *AGK* in the presence of cataracts

and 3-methylglutaconic aciduria, *TMEM70* founder mutation in Romanian patients, *YARS2* founder mutation in Lebanese patients with myopathy lactic acidosis and sideroblastic anaemia (MLASA) and *TAZ* mutations in the presence of cardiomyopathy and neutropaenia in a male patient.

Whole exome sequencing

Where mtDNA mutations were not found or a candidate nuclear gene was not apparent, WES was performed, assuming autosomal recessive inheritance. The method is described in chapter 2. Only cases that had followed this stratified genetic testing pathway were included in this study.

5.2.3 Western blotting

The western blot for the VARS2 protein was performed by Dr Rojeen Shahni. The method was as described in chapter 2 (2.4.1) with the exception of using Radioimmunoprecipitation assay buffer (RIPA) buffer for cell lysis and the use of 10% Bis-Tris pre-cast gels (Invitrogen) and the iBlot transfer system (Invitrogen) for transfer of proteins on to PVDF membranes.

5.2.4 Mitochondrial translation assay

This was performed as described in chapter 2 (2.10) and was performed by Dr Sarah Pearce and Dr Michal Minczuk at the MRC Mitochondrial Biology Unit, Cambridge.

5.3 Results

5.3.1 General characteristics

A total of 30 patients underwent genetic investigations (Table 5.2). Female:male ratio was 11:19. Twenty-four patients (83%) were of white European origin, of which three patients were of Roma traveller background. Three patients were of South Asian, two of Middle Eastern and one of African origin. In this cohort only two patients were from consanguineous families.

5.3.2 Presenting features

Most patients (25/30, 83%) presented <1 year of age of which 14 (43%) of these cases presented in the neonatal period (presentation age range birth-6 years). Among the 30 patients, thirteen patients (45%) demonstrated symptoms/signs related to cardiomyopathy (cardiomegaly, cardiac murmur, cyanotic episodes, shortness of breath, cardiogenic shock) at initial presentation. The other patients presented with symptoms related to systemic energy failure such as poor feeding/faltering growth (n=6, 20%), developmental delay (n=2, 7%) or congenital lactic acidosis without evidence of initial cardiac involvement (n=2, 7%). Two cases demonstrated hypotonia as an initial presentation and five had mitochondrial syndromic presentations (ptosis/ophthalmoplegia in KSS, sideroblastic anaemia in MLASA, migraine in MELAS, cataracts and 3-methylglutaconic aciduria in Sengers syndrome and neutropaenia in Barth syndrome).

5.3.3 Disease course

The clinical features are described in Table 5.2. P1, P2, P5, P6 and P22 are patients S6, S7, S12, S13 and S71 in chapter 2. P1 and P2 were initially diagnosed with a non-specific autistic spectrum behavioural disorder with global developmental delay, ligamentous laxity, hypotonia and learning difficulties, until asymptomatic cardiomyopathy was detected in adolescence, triggering suspicion of a mitochondrial disease.

P5 developed hypertrophic cardiomyopathy (HCM) at 10 months and died of multi-organ failure at 18 months while awaiting cardiac transplantation. As a result, her younger brother P6 was screened at birth, when lactic acidosis with muscle COX deficiency was found. He developed HCM by 8 months, underwent cardiac transplantation at 21 months and remains well at 10 years. Of the six patients with cardiomyopathy at birth, two died within the first 3 months, two died within the first year and two are currently alive at four and six years. One patient with HCM presented in cardiogenic shock after respiratory syncytial virus bronchiolitis and died at 6.5 months. Interestingly in both cases with Friedreich's ataxia, cardiomyopathy preceded the onset of neurological features.

Five patients developed specific symptomatology associated with mitochondrial syndromes (KSS, Sengers syndrome, Barth syndrome MLASA and MELAS) prior to developing cardiomyopathy. P12, a patient with known MELAS developed dilated cardiomyopathy (DCM) required inotropic support to maintain cerebral perfusion following a hemiplegic stroke after which he underwent cardiac transplantation at 13 years. P13 with KSS developed complete heart block requiring permanent pacing, prior to developing cardiomyopathy. Two maternal cousins (found to have the same mtDNA mutation) developed cardiomyopathy at different ages (7 months and 2.75 years) were diagnosed with different types of cardiomyopathy (hypertrophic vs. dilated) and symptoms such as developmental delay and hypotonia preceded the onset of cardiac dysfunction.

Three Romanian traveller patients presented at birth with cardiomyopathy, lactic acidosis and hyperammonaemia. Four of five patients with Barth syndrome presented with cardiac symptomatology (one had an incidental neutropaenia) and 4/5 patients had DCM. All five of these patients with Barth syndrome were referred for transplant; two died, two received cardiac transplants and in one patient, cardiomyopathy diagnosed at 2 months of age improved spontaneously and he is well at 9 years.

Table 5.2: Clinical phenotypes and genetic causes of mitochondrial cardiomyopathy in patients referred to Great Ormond Street Hospital between 1997-2015

Case	Sex	Age of onset	Clinical features at onset	Age at HCM/ DCM diagnosis	Other clinical features	Cardio-myopathy type	Age of death	Lactate mmol/L (0.7-2.1)	RCEs	Genetic cause	Mutation	Inherit-ance
P1	F	Perinatal	Grunting, poor feeding	14 years	Learning difficulties, lax ligaments developmental delay	DCM	Alive at 21 years	1.7	NA	*AARS2*	c.725delA, p.Asn242fs c.1846G>A p.Val616Met	AR
P2	M	Early infancy	Developme ntal delay	12 years	Learning difficulties, lax ligaments, spasticity, dystonia	DCM	Died at 17 years	2.1	Low CIV	*AARS2*	c.725delA, p.Asn242fs c.1846G>A p.Val616Met	AR
P3	M	Birth	Hypotonia, lactic acidosis, cardiomega ly	Birth	Seizures	HCM	Died on day 19	14-16	Low cardia c CI+IV, muscle normal	*VARS2*	c.1358G>C p.Arg453Pro c.2922_2924 del, p.974_975de l	AR

Case	Sex	Age of onset	Clinical features at onset	Age at HCM/ DCM diagnosis	Other clinical features	Cardio-myopathy type	Age of death	Lactate mmol/L (0.7-2.1)	RCEs	Genetic cause	Mutation	Inherit-ance
P4	M	1 year	Pallor, jaundice	12 years	Myopathy, ptosis, anaemia, pericardial effusion (MLASA)	HCM	Alive at 17 years	5.6-10.1	NA	*YARS2*	Homozygous c.156C>G, p.Phe52Leu	AR
P5	F	10 months	Lactic acidosis, cardiomyopathy	10 months	Poor feeding, faltering growth	HCM	Died 18 months	11-26	NA	*MRPL44*	Homozygous c.467T>G, p.Leu156Arg	AR
P6	M	Day 1	Lactic acidosis	8 months	cerebral infarct, hypoadrenalism	HCM+D+ ↓contractility	Alive at 10 years (rCT)	4.7-5.2	Low CI	*MRPL44*	Homozygous c.467T>G, p.Leu156Arg	AR
P7	M	Early infancy	Faltering growth	18 months	Lactic acidosis, Motor delay	DCM	Alive at 5 years	5.3	Low CI+IV	*MTO1*	Homozygous c.1307C>T, p.Thr436Ile	AR
P8	F	Birth	Lactic acidosis, cardiomyopathy	Birth		HCM	Died at 3 months	9.8	Low CI	*ELAC2*	Homozygous c.460T>C, p.Phe154Leu	AR

Case	Sex	Age of onset	Clinical features at onset	Age at HCM/ DCM diagnosis	Other clinical features	Cardio-myopathy type	Age of death	Lactate mmol/L (0.7-2.1)	RCEs	Genetic cause	Mutation	Inherit-ance
P9	M	6 weeks	Heart murmur	18 months	Leigh syndrome, developmental delay, pericardial effusion	HCM	Died at 2.5 years	2.1-3.0	NA	*c12orf65*	Homozygous c.210delA, p.Pro70fs	AR
P10	F	4 months	Poor feeding, faltering growth, Hypotonia	7 months	Developmental delay, lactic acidosis	HCM+↓contr-actility	Died at 3.5 years	5.4	Low CI+IV	*MT-TL1*	m.3303C>T	Maternal
P11	M	10 months	Hypotonia	2 years 9 months	Motor delay, lactic acidosis, fatigue	DCM	Died at 4 years	3-4.5	Low CI+IV	*MT-TL1*	m.3303C>T	Maternal
P12	M	6 years	Migraine	13 years	Right hemiplegia, migraines (MELAS)	HCM+D+ ↓contractility	Alive at 15 years (rCT)	High, value NA	NA	*MT-TL1*	m.3243A>G	Maternal

Case	Sex	Age of onset	Clinical features at onset	Age at HCM/ DCM diagnosis	Other clinical features	Cardio-myopathy type	Age of death	Lactate mmol/L (0.7-2.1)	RCEs	Genetic cause	Mutation	Inherit-ance
P13	F	2 years	Ptosis, ophthalmoplegia	10 years	Pigmentary retinopathy, deafness, complete heart block (KSS)	DCM	Died at 12 years	0.99-4.3	NA	*Single mtDNA deletion*	Large scale deletion	Sporadic
P14	F	18 months	Mild developmental delay	3 years	Dyspnoea	DCM	Alive at 14 years (rCT)	1.0-1.8	Low CI+IV in heart and muscle	*MT-CO1*	m.7317A>G	Maternal
P15	F	6 months	Cardiogenic shock after viral illness	6.5 months	Pallor, poor feeding, sweating, faltering growth	HCM+↓co-ntractility	Died at 7 months	2-18	Low CI	*NDUFAF1*	c.631C>T, p.Arg211Cys c.733G>A, p.Gly245Arg	AR
P16	M	12 months	Poor feeding and lethargy	12 months	Lactic acidosis, renal failure, encephalopathy	HCM+D+↓contracti-lity	Died at 13 months	22	Low CI	*ACAD9*	c.1795_1798del, p.599_600del c.1250G>A, p.Arg417His	AR

Case	Sex	Age of onset	Clinical features at onset	Age at HCM/ DCM diagnosis	Other clinical features	Cardio-myopathy type	Age of death	Lactate mmol/L (0.7-2.1)	RCEs	Genetic cause	Mutation	Inherit-ance
P17	F	Birth	Lactic acidosis, HCM	Birth	Hyperammonaemia, absent corpus callosum	HCM	Alive at 4 years	9.7-23	Normal	*TMEM70*	317-2A>G	AR
P18	M	Birth	Lactic acidosis, hypertrophic cardiomyopathy	Birth	Hypoglycaemia faltering growth, hyperammonaemia	HCM	Alive at 6 years	21	NA	*TMEM70*	317-2A>G	AR
P19	M	Birth	Hypothermic, pale, mixed severe acidosis	Birth	Respiratory distress	HCM	Died on day 2	NA	Normal	*TMEM70*	317-2A>G	AR
P20	F	Birth	Cyanotic episodes, fatigue	5 years 9 months	Oedema, ataxia, tremors (Friedreich)	HCM+↓co-ntractility	Died at 16 years	2.1	Low CIV	*FXN*	c.2T>C p.Met1Thr GAA repeat expansion in intron 1	AR

Case	Sex	Age of onset	Clinical features at onset	Age at HCM/ DCM diagnosis	Other clinical features	Cardio-myopathy type	Age of death	Lactate mmol/L (0.7-2.1)	RCEs	Genetic cause	Mutation	Inherit-ance
P21	F	4.5 years	Fatigue, shortness of breath	4.5 years	Cerebellar signs, 3-MGA (Friedreich)	HCM+D+ ↓contracti-lity	Alive at 13 years	1.8-2.6	Normal	*FXN*	Homozygous GAA repeat expansion of intron 1	AR
P22	F	6 weeks	Congenital cataracts	13 months	Motor delay, faltering growth (Sengers)	HCM+↓co-ntractility	Died at 17 months	1.6	Low CIV	*ATAD3A*	c.655C>T, p.Arg219X c.1844_1846del p.615_616del	AR
P23	M	Birth	Lactic acidosis	5 months	Sweating, cataracts (Sengers)	HCM	Alive at 19 months	13	NA	*AGK*	Homozygous c.1047-A>G	AR
P24	M	2 years	DCM	2 years	Neutropenia, myopathy, faltering growth (Barth)	DCM	Alive at 22 years (rCT)	NA	NA	*TAZ*	c.207C>G, p.His69Gln	X-linked
P25	M	17 days	Poor oxygen saturation	17 days	Neutropaenia, motor delay, poor feeding (Barth)	DCM	Died 34 months	1.9	NA	*TAZ*	c.337-9delTTC, p.Phe113del	X-linked

Case	Sex	Age of onset	Clinical features at onset	Age at HCM/ DCM diagnosis	Other clinical features	Cardio-myopathy type	Age of death	Lactate mmol/L (0.7-2.1)	RCEs	Genetic cause	Mutation	Inherit-ance
P26	M	3 weeks	Poor feeding, irritability	3 weeks	Neutropaenia, hypoglycaemia, VF (Barth)	DCM	Alive at 7 years (rCT)	NA	NA	*TAZ*	c.170G>T, p.Arg57Leu	X-linked
P27	M	Birth	Heart murmur	Birth	Neutropaenia (Barth)	HCM+↓contractility	Died at 9 months	NA	NA	*TAZ*	c.284+1G>C	X-linked
P28	M	2 weeks	Neutropaenia	2 months	Faltering growth (Barth)	DCM	Alive at 9 years	1.4	NA	*TAZ*	c.281G>A p.Arg94His	X-linked
P29	M	Birth	Hypotonia, lactic acidosis	Day 9	Abnormal neuroimaging, deafness	DCM	Died at 1 year	4-6.9	Normal	*MDH2*	c.368delC p.Thr123fs, c.759+5G>A	AR
P30	M	3 months	Faltering growth	3 years	Dyspnoea, renal failure, deafness, retinal dystrophy	HCM+D+↓contractility	Alive at 21 years (rCT)	1.6	Low CIV	Not identified	NA	NA

AR Autosomal recessive, CI Complex I, CIV Complex IV, D dilatation, DCM Dilated cardiomyopathy, F female, HCM Hypertrophic cardiomyopathy, M Male, NA not available, rCT received cardiac transplant, VF ventricular fibrillation, 3-MGA 3-methyl glutaconic aciduria

5.3.4 Mortality and transplantation

The median age of cardiac symptoms/signs was 0.9 years (range birth–14 years). In this cohort sixteen patients (53%) died (median age 1.5 years; range 2 days-17 years) with five patients dying <1 year of cardiac failure. Of the patients who died (n=16) the median interval between cardiac presentation and death was 9 months (range 1 day-5 years). Six patients were recipients of orthotopic cardiac transplants and are currently alive at 7, 10, 14, 15, 21 and 22 years. Median age at transplantation or death was 2 years (range 2 days-17 years). Median 'event free survival time' (where both death and transplant were events) was 3.3 years (CI 1.75-14). Kaplan Meier survival analysis is demonstrated in Figure 5.1.

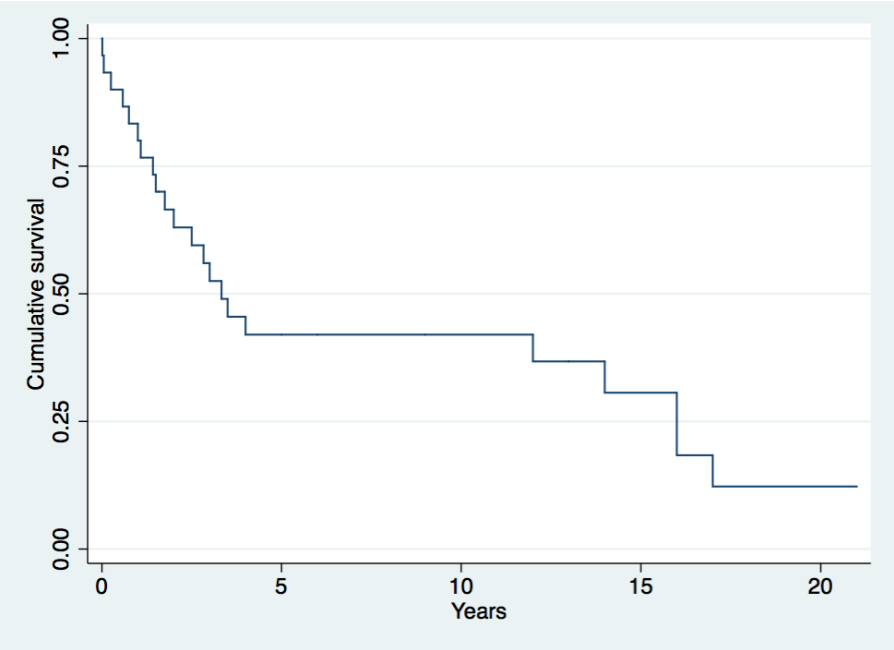


Figure 5:1: Kaplan Meier survival analysis of a paediatric cohort with mitochondrial cardiomyopathy (n=30)

5.3.5 Echocardiography

In the cohort, nine patients demonstrated features of HCM while 11 developed DCM. Ten had a 'mixed' picture developing HCM with impaired contractility of which five also had dilatation (Table 5.3). None of the patients met the criteria for LV non-compaction (LV non-compaction ratio>2) but seven met the criteria for hypertrabeculation (ratio>1). Of the seven who met the criteria for hypertrabeculation, six had DCM.

Table 5.3: Echocardiographic features in a cohort with mitochondrial paediatric cardiomyopathy

	HCM	DCM	HCM with impaired contractility
Number	9	11	10
Age median(range) years	0.05(0-12.7)	2.8(0.03-14.3)	1.4(0.6-15.1)
LV non-compaction ratio median(range)	0.62(0.43-0.89)	1.08(0.83-1.45)	0.71(0.34-1.02)
Compacted diameter mm median(range)	0.7(0.40-0.95)	0.62(0.33-0.81)	0.81(0.51-1.11)
Non-compacted diameter mm median(range)	0.3(0.25-0.85)	0.68(0.34-0.89)	0.51(0.38-1.06)
Right atrium Z score, median(range)	-1.03(-2.77-3.53)	-1.42(-3.42-(-0.52))	-0.76(-2.38-0.06)
Left atrium Z score, median(range)	-0.96(-1.87-3.1)	-0.34(-1.22-(-0.31))	0.52(-0.24-2.89)
Interventricular septum Z score, median(range)	4.02(0.70-12.25)	0.29(-1.78-2.35)	2.97(0.54-11.11)
LV posterior wall dimensions Z score, median(range)	5.49(2.97-17.20)	0.23(0.53-2.2)	7.11(2.39-12.88)
LV diameter (diastolic) Z score, median(range)	-2.04(-4.85-(-0.94))	3.18(-0.10-7.79)	3.77(-3.44-8.42)
LV diameter(systolic) Z score, median(range)	-2.64(-4.63-0.23)	4.78(1.78-16.23)	8.09(-1.07-13.88)
LV shortening	37.7(29-60)	15(7.4-33.3)	14(9-35.7)

	HCM	DCM	HCM with impaired contractility
fraction %, median(range)			
LV ejection fraction %, median(range)	60(59-61)	37(23-50)	38(19-57)

DCM: dilated cardiomyopathy, HCM: hypertrophic cardiomyopathy, LV: Left ventricular

5.3.6 Metabolic investigations

Where plasma lactate was measured (n=27), elevated plasma lactate was noted in most cases (20/27, 74%). Plasma amino acid analysis demonstrated raised alanine in 10/16 cases (range 480–3503 µmol/L, reference range 150-450). Urine organic acid analysis revealed increased lactate and pyruvate metabolites (10/20), Krebs cycle intermediates (e.g. fumarate, malate and 2-oxoglutarate in 3/20) and 3-methyl glutaconic aciduria (3-MGA) (7/20). Seven patients demonstrated normal organic acid profiles. All Barth syndrome cases demonstrated diagnostically elevated monolysocardiolipin/cardiolipin ratios.

5.3.7 Muscle biopsy

Muscle biopsies were performed in 18 patients and revealed increased lipid in 11 cases, COX-negative fibres in three patients, and ragged-red and ragged-blue fibres in P16. Muscle RCE activities were measured in 17 patients and showed isolated COX deficiency (5/17), isolated complex I deficiency (3/17), or combined complex I and IV deficiency (4/17). Five patients had normal muscle RCE activities (Table 5.4).

Table 5.4: Mitochondrial respiratory chain enzyme activities in 17 patients with mitochondrial cardiomyopathy

Patient	Complex I (0.104-0.268)	Complex II+III (0.048-0.204)	Complex IV (0.014-0.034)	Genetic cause
P2	0.153	0.107	0.008	AARS2
P3	0.262	0.221	0.017	VARS2
P3 cardiac	0.033	0.100	0.002	VARS2
P6	0.114	0.109	0.009	MRPL44
P7	0.074	0.095	0.007	MTO1
P8	0.057	0.152	0.017	ELAC2
P10	0.049	0.093	0.008	MT-TL1
P11	0.022	0.055	0.006	MT-TL1
P14	0.060	0.130	0.007	MT-CO1
P14 cardiac	0.008	0.126	0.002	MT-CO1
P15	0.026	0.075	0.019	NDUFAF1
P16	0.017	0.107	0.019	ACAD9
P17	0.136	0.159	0.018	TMEM70
P19	0.132	0.202	0.027	TMEM70
P20	0.220	0.136	0.007	FXN
P21	0.263	0.133	0.030	FXN
P22	0.245	0.106	0.012	ATAD3A
P29	0.187	0.145	0.025	MDH2
P30	0.134	0.041	0.010	Unknown

All measurements made on skeletal muscle unless otherwise stated. All complex activities expressed as a ratio to citrate synthase. Reference ranges stated in brackets.

5.3.8 Cardiac biopsy

Cardiac biopsies were analysed by Dr Michael Ashworth, Histopathologist, Great Ormond Street Hospital London, UK. Cardiac histology of eight patients (P3, P5, P6, P12, P14, P15, P16 and P30) revealed hypertrophic cardiomyocytes, with cytoplasmic vacuolation and giant mitochondria in 7/8 biopsies (Figure 5.2). In P16 there were non-specific changes with endocardium showing mild fibroelastic thickening. There was a small amount of interstitial fibrosis of the myocardium but no other abnormality was noted.

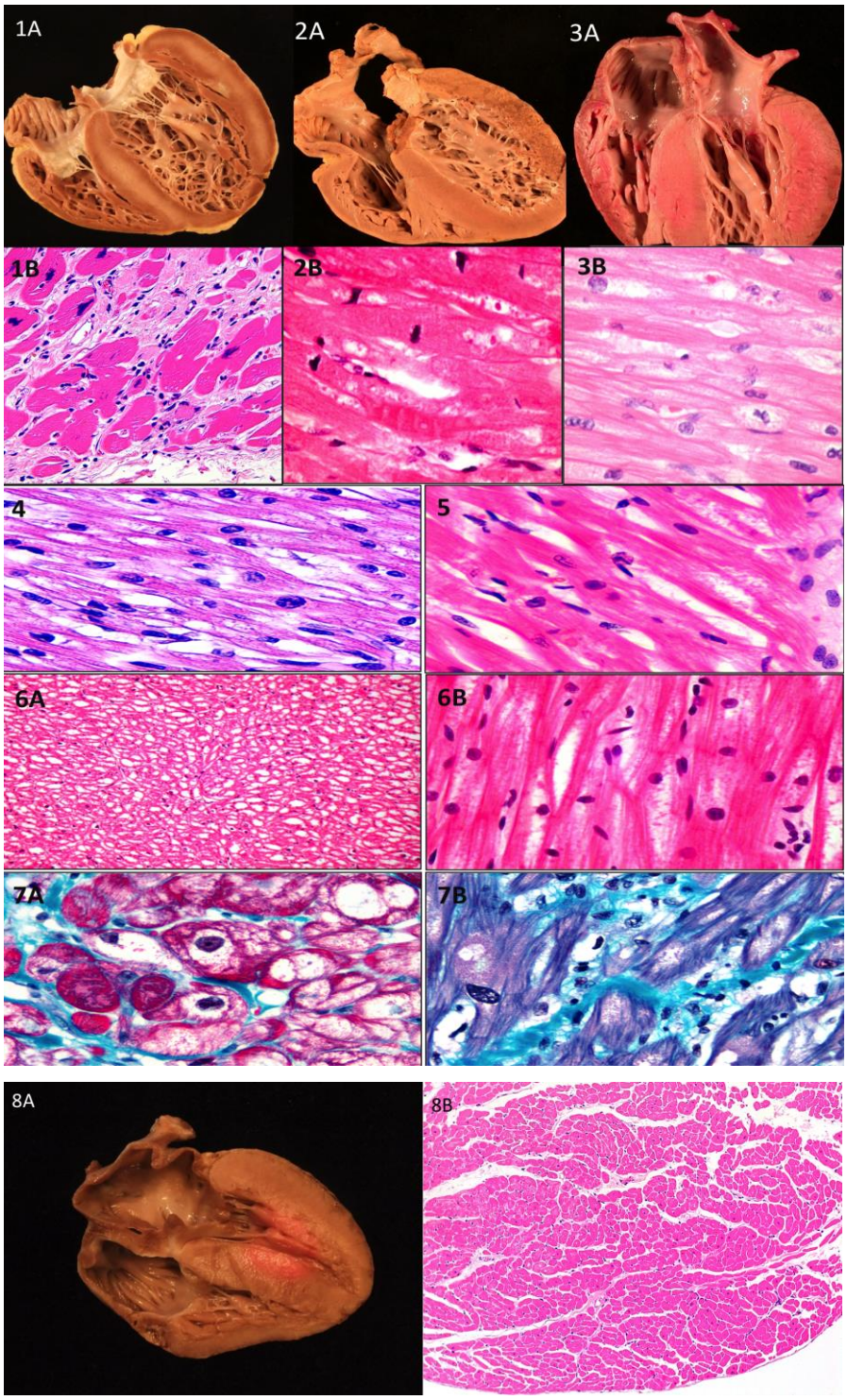


Figure 5:2: Macroscopic and microscopic findings in cardiac tissue from mitochondrial cardiomyopathy patients.

Explanted hearts from:

1A, B P12 (MELAS syndrome). Thickened LV wall with dilated LV. . Myocardium shows hypertrophy of the myocytes, many with enlarged, irregular and hyperchromatic nuclei

2A, B P30 (genetic diagnosis unknown). Dilated right atrium, right ventricle and dilated left ventricle. Microscopy shows hypertrophic myocardium enlarged myocytes are enlarged with,hyperchromatic nuclei.
3A, B P15 NDUFAF1 deficiency. Thickened RV, LV and septum. Microscopy shows vacuolation and irregular nuclei.
4 P3 VARS2- deficiency. Hypertrophic myocytes with cytoplasmic vacuolation
5 P5 MRPL44 deficiency 6A, B P6 MRPL44-deficiency. Microscopy shows hypertrophied myocytes with enlarged, irregular and hypertrophic nuclei.
7A, B P14 MT-CO1 deficiency. Microscopy shows enlarged myocytes showing central vacuolation and lacy cytoplasm
8A heart from a 6 month old without cardiomyopathy (control). 8B a haematoxylin and eosin stain from a healthy heart

5.3.9 Genetic findings

A genetic diagnosis was made in 29/30, 97% of cases. MtDNA sequencing was used in the diagnosis of 5/30 (17%) cases; a candidate gene approach was used in 12/30 (40%) identifying *YARS2, TMEM70, FXN, AGK* and *TAZ* mutations; and WES was used in 10/30 (33%) identifying the novel genetic causes *ATAD3A* and *MDH2,* along with *AARS2, VARS2, MRPL44, ELAC2, c12orf65* and *ACAD9* mutations (Table 5.2). Novel mutations were found in the *ACAD9, AGK, VARS2* and *AARS2* genes.

In P15, a patient who demonstrated isolated complex I deficiency, Sanger sequencing of complex I subunits and assembly factors was performed and revealed *NDUFAF1* mutations (Fassone et al. 2011). The *MTO1* mutation was screened because an affected cousin was diagnosed with an *MTO1* mutation at another centre. Interestingly an unrelated patient not reported here, with the same *MTO1* mutation revealed by WES had congenital lactic acidosis and liver failure with Complex I and IV deficiency, but not cardiomyopathy. Similarly the patient with Leigh syndrome and peripheral neuropathy with a homozygous *c.12orf65* (c.248delT, p.V83fs) mutation in the COX deficient cohort in chapter 3 did not have cardiomyopathy. The genetic cause was not found in P30. There was no biological material available in P29 and pathogenicity for the *MDH2* mutations was verified by familial segregation of mutations. Pathogenicity for novel *ATAD3A*, *AARS2* and *VARS2* mutations was verified by familial segregation and Western blot analysis. Segregation studies and Western blot for AARS2 and ATAD3A mutations have already been illustrated in chapter 3.

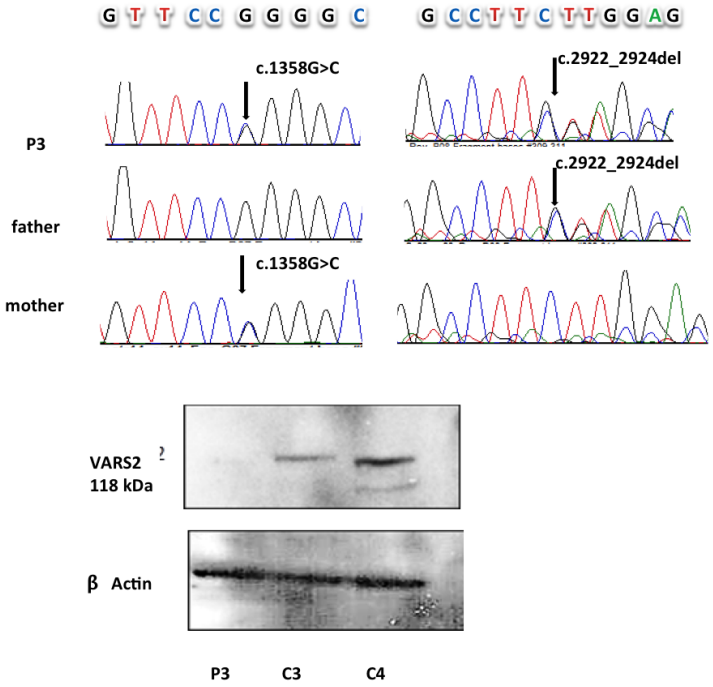


Figure 5:3: Sanger sequencing and western blotting in P3 with novel *VARS2* mutations

(C3 and C4: controls)

Novel genetic findings: ATAD3A is a novel cause of Sengers syndrome

The evidence for pathogenicity and investigation of postulated functions in the ATAD3A deficient patient has been described in chapter 4.

Novel genetic findings: MDH2 is a novel genetic cause of mitochondrial cardiomyopathy

P29 was a male patient born by spontaneous vaginal delivery, at term, weighing 3.18 kg. It was noticed that he was very floppy immediately at birth. He was not latching properly during the next couple of hours and also remained floppy. He had a poor cry and stopped breathing twice after starting to cry. He did not require ventilation, but was noted to have hypothermia, lactic acidosis and an abnormal cranial ultrasound. Subsequently MRI brain scan showed absent corpus callosum, polymicrogyria, cystic changes and a poorly developed cerebellum.

On day nine of life numerous investigations were performed including an echocardiogram, which showed left ventricular hypertrophy. His condition remained fairly stable although cardiac function continued to deteriorate. He never made developmental progress and at 8 months he was not responsive at all and did not smile. He was found to have cortical blindness and bilateral sensorineural hearing loss. He died at 1 year of age of congenital lactic acidosis. Plasma amino acids were normal on the first occasion they were measured but a repeat measurement showed that the plasma serine was just below the normal range on the first occasion and on the second was significantly reduced at 53 µmol/L (reference 90 to 290). CSF serine was low at 28 µmol/L (reference range 50-90).

Urine organic acids showed strongly raised lactate and moderately raised pyruvate with moderately raised fumarate, 2-oxoketoglutarate and malate and mildly raised 3 hydroxybutyrate.

Transferrin electrophoresis was normal. Muscle histology showed excess lipid with no features suggestive of a mitochondrial disorder. Muscle respiratory chain enzymes were normal (Complex I 0.187 (0.104-0.268), Complex II+III 0.145 (0.040-0.204), COX 0.025 (0.014-0.034)).

WES revealed compound heterozygous mutations (Table 5.5) in *MDH2* encoding malate dehydrogenase, a Krebs cycle enzyme. Familial segregation of disease, where parents were heterozygous for each mutation, was verified but biological material was not available for further functional studies. However his urine organic profile described above is consistent with a Krebs cycle defect such as MDH2 deficiency with the presence of malate, fumarate and 2-oxoketoglutarate. At the time of writing MDH2 deficiency had not previously been associated with human disease.

Table 5.5: Bioinformatic analysis of whole exome sequencing data in patients with mitochondrial cardiomyopathy

	P1	P2	P3	P5	P6	P9	P8	P16	P22	P29	P30
Total variants	24,012	23,748	24,247	24,764	24.492	24,248	19,566	18,289	24,889	25,361	27,465
Rare variants after filtering	360	358	365	421	389	329	474	517	314	578	993
Genes with rare compound heterozygous variants	14	20	16	14	10	5	24	28	7	19	67
Genes with rare homozygous variants	3	3	3	2	3	3	4	0	0	3	14
Genes with shared variants between siblings	5	5	NA	12	12	NA	NA	NA	NA	NA	NA
Genes with mitochondrially targeted products	1	1	1	1	1	1	1	1	1	1	-
Candidate gene	**AARS2**	**AARS2**	**VARS2**	**MRPL44**	**MRPL44**	**C12orf65**	**ELAC2**	**ACAD9**	**ATAD3A**	**MDH2**	**Not found**

5.3.10 Mitochondrial pathways

In this cohort, 13/30 (43%) patients had defects involving the mitochondrial protein translation system, two patients had *FXN* mutations (a mitochondrial iron chaperone involved in iron-sulphur cluster biogenesis), three patients had homozygous founder Romanian *TMEM70* mutations (biogenesis and stabilisation of ATP synthase), two had complex I assembly factor mutations (*ACAD9* and *NDUFAF1*) and one had a complex IV subunit mutation (*MT-CO1*) (Figure 5.4). In addition ACAD9 has been found to have a possible role in fatty acid oxidation. Seven patients had defects of mitochondrial membrane metabolism (TAZ, AGK and ATAD3A). Cardiolipin (two phosphatidic acid moieties connected by a central glycerol backbone) is an essential component of the inner mitochondrial membrane where the OXPHOS complexes lie and has a key role in maintaining its integrity. It is almost exclusively found in the mitochondrial inner membrane and is synthesised using phosphatidylglycerol and cytidinediphosphate-diacylglycerol. *TAZ* encodes tafazzin and *AGK* encodes acyl glycerol kinase, both which are involved in cardiolipin biosynthesis. One patient had a Krebs cycle defect (MDH2).

AARS2, YARS2 and VARS2 are mitochondrial aminoacyl synthetases catalysing tRNA binding to their cognate amino acids alanine, tyrosine and valine respectively. *MT-TL1* encodes a mitochondrial tRNA which transfers leucine to the polypeptide chain at the mitoribosome during translation. ELAC2 is an endoribonuclease which catalyses the removal of the 3' end from precursor tRNAs. MTO1 is involved in mitochondrial tRNA modification, MRPL44 is a mitoribosomal protein, whilst C12orf65 contributes to peptide chain termination.

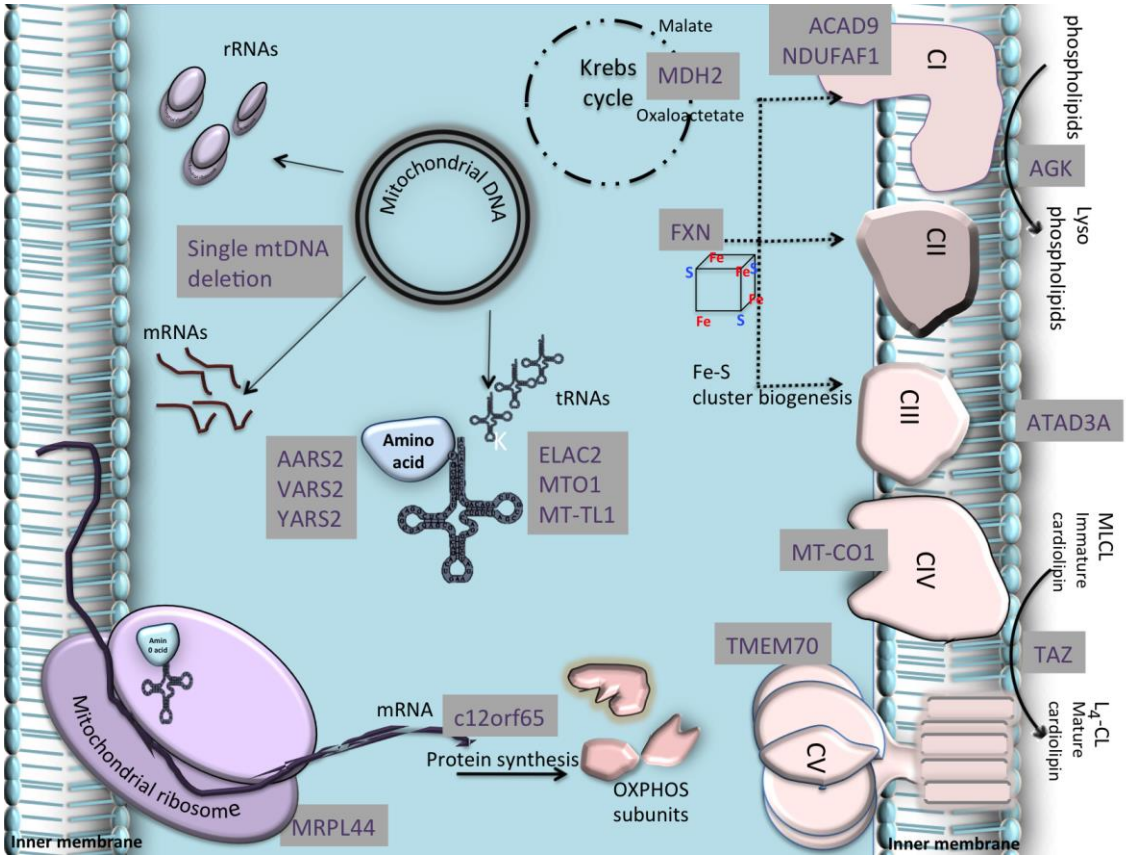


Figure 5:4: Molecular mechanisms of mitochondrial cardiomyopathy revealed by the investigation of 30 patients with mitochondrial cardiomyopathy.

Mitochondrial DNA encodes 37 genes: 13 mRNAs for the protein subunits of the OXPHOS enzymes, 2 ribosomal RNAs and 22 transfer RNAs (tRNA). Protein translation takes place on the mitoribosome. The aminoacyl synthetases catalyse the binding of the tRNA to their cognate amino acids (e.g. AARS2, VARS2, YARS2). MT-TL1 is a mitochondrial tRNA which transfers leucine. MTO1 and ELAC2 post transcriptionally modify tRNA. C12orf65 is involved in peptide chain termination. MRPL44 is a mitoribosomal protein. ACAD9 and NDUFAF1 are complex I assembly factors and MT-CO1 is a complex IV subunit. TMEM70 is a membrane protein which is thought to be involved in biogenesis of complex V. TAZ, AGK and ATAD3A are mitochondrial membrane proteins. FXN is involved in mitochondrial Fe-sulphur cluster biogenesis and MDH2 encodes a Krebs cycle enzyme, malate dehydrogenase.

5.4 Discussion

Mitochondrial medicine is a rapidly expanding clinical field and to date, > 250 mtDNA mutations and >200 causative nuclear genes have been recognised involving various biosynthetic pathways (http://www.mitomap.org/MITOMAP) (Vafai & Mootha 2012; Kohda et al. 2016). In this paediatric cohort with mitochondrial cardiomyopathy the genetic diagnosis was found in 29/30 (97%) where 18 different genetic mechanisms were established due to both mtDNA and nuclear genes. Novel genetic causes of cardiomyopathy, *ATAD3A*, a new cause of Sengers syndrome and *MDH2*, encoding malate dehydrogenase in the Krebs cycle, were found. The results here demonstrate that dysfunction of the mitochondrial translation system may play a key mechanistic role in mitochondrial cardiomyopathy.

According to the findings, mitochondrial cardiomyopathy resulted from defects in diverse mitochondrial biosynthetic pathways; the most commonly involved pathway (9/18 genes) was mitochondrial protein translation, which should typically cause reduced synthesis of OXPHOS complexes I, III, IV and V. OXPHOS deficiency is more frequently associated with the HCM phenotype (caused by reduced myocardial relaxation and aberrant tissue formation) than DCM (Watkins et al. 2011). However, wall thickening probably occurs alongside or is followed by dilatation (Cahill et al. 2013). To this end, it was observed that the same mtDNA defect (*MT-TL1*) caused dilated and hypertrophic presentations in two cousins. Furthermore ten patients had ‘mixed’ echocardiographic findings such as HCM with impaired contractility. Other interesting findings included observing the same *MTO1* mutation causing cardiomyopathy in one patient but not in a second unrelated patient. This kind of clinical heterogeneity is typical of mitochondrial disease and highlights the difficulties of drawing genotype phenotype correlations. For mtDNA mutations organ-specific disease manifestations are explained by tissue-specific mutant DNA load, but the cause for tissue specificity in nuclear gene mutations is less clear. Considering for example, the protein translation pathway, mitochondrial tRNA requires extensive modification to progress into its functional form. This process requires >30 mitochondrial enzymes, cofactors and substrates where cellular availability may vary between tissues (Powell et al. 2015).

Emerging metabolomics data have suggested that there may be a ‘metabolic signature’ in mitochondrial syndromes (Thompson Legault et al. 2015). When establishing the diagnosis of a mitochondrial disease, metabolic markers may provide diagnostic clues. In the novel Krebs cycle defect MDH2, the presence of malate and 2-oxoglutarate organic aciduria aided confirmation of the pathogenicity of the mutations. Mitochondrial metabolic markers such as hyperlacticacidaemia, hyperalaninaemia and Krebs cycle metabolites provided corroborative evidence. Also, 3-MGA, is a biomarker in Barth, Sengers and MEGDEL syndromes, CLPB deficiency and, as identified here, ATAD3A deficiency (Wortmann et al. 2012; Kanabus et al. 2015). Given that 3-MGA is frequently associated with genetic cardiomyopathy, its presence should prompt cardiovascular examination. On the other hand, if a patient presents with cardiomyopathy and 3-MGA an underlying mitochondrial disease should be suspected.

If hypertrophy with decreased systolic function is observed in childhood, a mitochondrial disease or storage myocardial disease should be suspected (Rapezzi et al. 2013). Also hypertrabeculation is particularly suggestive of mitochondrial disease (Meyers et al. 2013). Mitochondrial cardiomyopathy may be misdiagnosed as viral cardiomyopathy, since viral illnesses can precipitate acute decompensations in patients with mitochondrial disease, producing a confusing clinical picture.

Most patients had decreased muscle RCEs, which was regarded as the gold-standard test to delineate mitochondrial disease prior to the advances in genetic testing. However the cases here illustrate the interpretative difficulties encountered when using muscle RCEs as a diagnostic test. In P3 (*VARS2*), where echocardiography was suggestive of mitochondrial aetiology, muscle RCEs were normal but with clear reduction in cardiac tissue. One patient with Friedreich ataxia unexpectedly had low COX, since impaired iron-sulphur cluster biosynthesis in this condition should specifically affect complexes I, II and III. A COX subunit mutation atypically produced both complex I and IV deficiency; complex I deficiency here is presumably secondary to perturbation of supercomplexes. On the other hand, 10 cases received a genetic diagnosis without a muscle biopsy for RCEs, facilitated by knowledge of mitochondrial syndromes and founder mutations in certain populations.

The majority of the cardiac biopsies demonstrated similar features (hypertrophic cardiomyocytes, vacuolated cytoplasm and giant mitochondria). While cytoplasmic vacuolation is suspicious of either a metabolic or mitochondrial aetiology, the presence of giant mitochondria in association with cardiomyopathy is very suggestive of a mitochondrial aetiology.

This study encompasses the largest cohort of genetically defined paediatric mitochondrial cardiomyopathy reported to date. While cardiomyopathy is well-documented in Friedreich's ataxia and Barth syndrome, in many cases described here the causative gene was associated with human disease only in the last 5 years. WES enabled the diagnosis of 10/30 (33%) patients. The cause of disease was not found in one case (P30), which may be due to limitations of WES, which does not target deep intronic regions. In addition it may be due to alternative inheritance patterns than hypothesised here, such as *de novo* dominant.

Between 30-40% of mitochondrial patients have cardiac involvement (Limongelli et al. 2010; Wahbi et al. 2015) and a 10% incidence of major adverse cardiac events over 7 years was reported in adult mitochondrial disease (Wahbi et al. 2015). Adults largely had mtDNA mutations (Wahbi et al. 2015) whereas nuclear gene mutations were more prevalent (82%) in the paediatric cases described here. Survival rate without transplantation is as poor as previously reported, where the survival rate at 16 years was 18% (also 18% in this study) (Scaglia et al. 2004). However because patients from a transplant centre were studied this may have skewed the population to severe cases referred for cardiac transplantation. The six

patients who underwent cardiac transplantation had four different genetic causes (*MRPL44, MT-TL1, MT-CO1, TAZ*) of cardiomyopathy and positive outcomes. Riboflavin responsive ACAD9 deficiency was identified by WES, in a deceased patient. This is a relatively new treatable disorder and emphasises the importance of rapid genetic diagnosis.

In summary this study elucidates the causes of paediatric mitochondrial cardiomyopathy in a substantial cohort, illustrating the accompanying clinical and genetic diversity and complex disease course. Furthermore by successful cardiac transplantation, steps towards successful therapeutic intervention in a disease group widely regarded as untreatable, is demonstrated here.

Chapter 6

6. TRNT1 deficiency: clinical, biochemical and molecular genetic features

6.1 Introduction

WES and more recently WGS, has enabled the diagnosis of many mitochondrial disease patients worldwide. New syndromes have been discovered (Steenweg et al. 2012; Kanabus et al. 2015) and previously known syndromes such as Sengers syndrome have been genetically redefined (Mayr et al. 2012).

At the start of this study in 2011/2012 two families (four patients) with an unusual multi system phenotype (infantile-onset cyclical, aseptic febrile episodes with vomiting and diarrhoea, global electrolyte imbalance during these episodes, sideroblastic anaemia, B lymphocyte immunodeficiency and retinitis pigmentosa) were identified. At the time, the genetic diagnosis was unknown. WES was performed only in one family because it was apparent that another research group had begun investigating the second family.

In 2013 Wiseman *et al* defined this syndrome as 'SIFD' congenital sideroblastic anaemia associated with immunodeficiency, fevers, and developmental delay (Wiseman et al. 2013). In 2014 the genetic cause of this syndrome was identified as *TRNT1* by Chakraborty *et al.* TRNT1 is a CCA-adding tRNA nucleotidyl transferase enzyme, which performs an essential post-transcriptional modification step by adding on the cytosine-cytosine-adenine (CCA) trinucleotide sequence to the 3′ end of both cytoplasmic and mitochondrial tRNAs. This TRNT1-dependent tRNA modification is essential for both cytosolic and mitochondrial tRNAs (mt-tRNAs) to participate in protein synthesis. The CCA trinucleotide sequence is required to precisely attach amino acids, to position the tRNA on the ribosome correctly and also to terminate protein translation (Hou 2010).

Since its first description in 2014 two different TRNT1 related disease entities have been reported; a haematological description of a congenital sideroblastic anaemia associated with immunodeficiency, fevers, and developmental delay (SIFD) in childhood (Wiseman et al. 2013; Chakraborty et al. 2014; Sasarman et al. 2015) and an adult presentation of retinitis pigmentosa with a focus on ophthalmological features (DeLuca et al. 2016). This chapter reports a new family with TRNT1 deficiency and studies the functional effects of the mutations on CCA addition to mt-tRNAs. In addition this chapter also reviews all the features reported in the cases published thus far and summarises the mutations published to date.

6.1.1 Declaration of work

Sanger sequencing was performed by Dr Rojeen Shahni and the translation assay, Northern blot and circularisation experiments were performed by Dr Michal Minczuk and his team at the MRC Mitochondrial Biology Unit, University of Cambridge. Kerra Pearce performed homozygosity mapping and protein modelling was performed by Dr Jose Saldanha. I collated clinical details and performed filtering of WES data to uncover the genetic cause of disease.

6.2 Methods

6.2.1 Whole exome sequencing

WES was performed as described in chapter 2 using the Agilent SureSelect Capture kit. Rare variants (defined as having an allele frequency <0.5% in our in-house control samples and NHBLI data set) which were splicing, nonsynonymous, frameshift or presumed loss-of-function and present in both affected siblings (A1 and A2) were prioritised. Since patients were unrelated, compound heterozygous inheritance was assumed.

6.2.2 Homozygosity mapping

Homozygosity mapping was performed as described in the chapter 2 (2.2.5). Whole genome-wide SNP genotyping was performed in S10 and S11, their parents and one unaffected sibling using the Illumina HumanCytoSNP-12 microarray.

6.2.3 Sanger sequencing

The *TRNT1* mutations identified by WES in A1 and A2 were confirmed by Sanger sequencing as described in chapter 2, using the following primer sequences for *TRNT1* exons 3 and 6

Exon 3:

Forward ATAGCAGGAGGAGCAGTGAG

Reverse GACTGCAGGGTTTATGACGG

Exon 6:

Forward CAGATTTTGCTTGTGATATGCCA

Reverse ACCCCATAACCCAAACTTTGTC

6.2.4 Mitochondrial translation assay

This analysis was performed by Dr Michal Minzcuk, Christopher Powell and Sarah Pearce at the MRC Mitochondrial Biology Unit, University of Cambridge, UK. The method is described in chapter 2 (2.10).

6.2.5 Circularisation reverse transcription–PCR assay

This analysis was performed by Dr Michal Minzcuk, Christopher Powell and Sarah Pearce at the MRC Mitochondrial Biology Unit, University of Cambridge, UK. Circularisation reverse transcription PCR (cRT–PCR) and sequencing were used to determine both 5’ and 3’ ends of mt-tRNA as described previously (Rorbach et al. 2011). In this method, 2.5 µg of total RNA was circularised with T4 RNA ligase (New England Biolabs) in the presence of DNase I (Roche) followed by phenol–chloroform extraction and ethanol precipitation. Reverse transcription was carried out using Omniscript Reverse Transcriptase (Qiagen) with a gene-specific reverse primer (tRNA_His_R1).

Newly synthesised cDNA was used as a template for a PCR reaction, primed with R1 primer and a forward primer, F1. The region amplified contained the junction of the 5’ and 3’

extremities. Products were cloned using TOPO TA Cloning Kit (Life Technologies) and sequenced.

The following primers were used in the analysis:

Mt-tRNA_His_Reverse: CAATCTGATGTTTTGGTTAAACTATA,

Mt-tRNA_His_Forward: AATCTGACAACAGAGGCTTACGACCC,

Mt-tRNA_SerAGY_Reverse: CATGAGTTAGCAGTTCTTGTGAGCTT,

Mt-tRNA_SerAGY_Forward: CCCCATGTCTAACAACATGGCTTTC,

Mt-tRNA_Cys_Reverse: TGCAATTCAATATGAAAATCACCTCG,

Mt-tRNA_Cys_Forward: TTCGAAGAAGCAGCTTCAAACCTGCC.

The following TOPO clones were sequenced:

mt-tRNA^{His} C1 29 clones, A1 32 clones, A2 29 clones

mt-tRNA^{Cys} 16 clones for C1, A1 and A2

mt-tRNA^{SerAGY} C1 13 clones, A1 16 clones, A2 15 clones.

6.2.6 Northern blotting

This analysis was performed by Dr Michal Minzcuk and Christopher Powell at the Mitochondrial Biology Unit, Cambridge, UK. Northern blot analysis was performed as described previously (Minczuk et al. 2011). Briefly, total RNA was isolated from primary skin fibroblasts using TRIzol® LS Reagent (Life Technologies) according to the manufacturer's instructions. RNA was resolved by Urea-PAGE on a 10% gel, transferred to a nylon membrane, UV cross-linked to the membrane and hybridised with radioactively labelled T7-transcribed RNA probes.

6.2.7 Protein modelling

This analysis was performed by Dr Jose Saldanha, National Institute of Medical Research, UK. The molecular graphics program, Molecular Operating Environment, 2012 running on an Ubuntu Linux workstation was used to visualize, mutate and analyse the crystal structure of full-length human mitochondrial CCA-adding enzyme (Kuhn et al. 2015) solved at a resolution of 1.9Å.

6.2.8 Identification of all reported cases

A PubMed search was performed using the terms 'TRNT1', 'SIFD' and 'sideroblastic anaemia'. Since the gene *TRNT1* was only associated with human disease in 2014 publications, from 2014 to date were searched.

6.3 Results

6.3.1 Case histories

This study identifies one new family and also studies the phenotype of one previously reported family (Wiseman et al. 2013; Chakraborty et al. 2014). Both families attended the metabolic clinic at Great Ormond Street Hospital, London UK for diagnosis and treatment.

A1 was born to unrelated white European parents after a normal pregnancy. At two weeks of age she presented with a febrile illness associated with poor feeding and diarrhoea. At the time an infective cause was not found and she recovered from this episode. Over the next three years she had multiple recurring acute episodes of fever, diarrhoea and vomiting, metabolic acidosis, electrolyte imbalance (hyponatraemia, hypokalaemia, hypocalcaemia, hypomagnesaemia and hypophosphataemia), elevated hepatic transaminases and raised inflammatory markers. While being investigated for an immune deficiency she was found to have hypogammaglobulinaemia with B cell maturation arrest and received sandoglobulin supplementation every three weeks. A transfusion dependent sideroblastic anaemia with variable neutropaenia and thrombocytopaenia was also detected. Immunological investigations revealed low IgG, IgA and IgM (Table 6.1) while blood films showed anisocytosis, neutropaenia, pencil cells and elliptocytes. Bone marrow examination demonstrated that most erythroblasts had abnormal siderosis with around two-thirds being ring or crescent forms. While there was dysplastic nuclear morphology megakaryopoiesis was abundant.

During the course of the illness, she developed episodes of encephalopathy associated with nystagmus and photophobia. Development was globally delayed and retinal dystrophy was noted. Neuroimaging was initially normal but subsequently progressive cerebellar atrophy was observed. In addition, in the last stages of her disease, brain MRI showed widespread multiple lesions resembling infection. There were recurrent bacterial infections, gastrointestinal involvement with chronic gastritis, partial villous atrophy, hepatosplenomegaly and pancreatic insufficiency as demonstrated by an undetectable stool elastase.

Biochemical investigations revealed mildly elevated blood lactate at 2.5 mmol/L (reference range <2.1) and muscle mitochondrial respiratory chain COX activity at the lower end of the reference range (ratio to citrate synthase 0.015, reference range 0.014-0.034). Muscle histology was normal. In particular, neither ragged-red nor COX-negative fibres were found. She died at 3 years and 3 months after a progressive encephalopathic episode.

A2 is the older sibling of non-identical twins (Figure 6.1A) and is the younger sibling of A1. At 3 weeks of age he developed poor feeding, loose stools and pallor after which he appeared to recover but at 4 weeks he was admitted to hospital with poor feeding, fever and anaemia (haemoglobin 66 g/L). His blood film demonstrated occasional nucleated red cells with possible sideroblastic granules and his bone marrow examination at one month showed dyserythropoeisis but no ringed sideroblasts. He had central hypotonia, retinal pigmentation,

hepatosplenomegaly and exocrine pancreatic insufficiency. Immunological investigations revealed initially low B cells which normalised and low IgG, IgA and IgM (Table 6.1). He had recurrent seizures and his MRI brain at 10 months demonstrated multiple focal lesions in the cerebral hemispheres and cerebellum. He died at 10 months of age because of intractable status epilepticus.

S10 and S11 (recruited to the COX deficient cohort), reported as cases 1 and 2 in Wiseman *et al* (Wiseman et al. 2013) were born to first cousin Pakistani parents (Figure 6.1D). They had five children (one was a stillborn baby at 32 weeks' gestation). Two other daughters are well except for radioulnar fusion due to Cenani-Lenz syndrome caused by a homozygous *LRP4* mutation (Li et al. 2010).

S10 was found to have sideroblastic anaemia at 8 months. From 1 year of age she developed recurrent episodes of fever during which global developmental delay, retinitis pigmentosa and bilateral sensorineural deafness were noted. Low B cells with hypogammaglobulinaemia with features of sideroblastic anaemia on bone marrow examination were found. In order to maintain her haemoglobin (Hb) above 70 g/L she required multiple blood transfusions. During the febrile episodes she developed episodic hypokalaemia, hyponatraemia, hypocalcaemia, hypophosphataemia hypomagnesaemia and generalised aminoaciduria. Plasma amino acid analysis revealed elevated proline 347 µmol/L (85-290), alanine 749 µmol/L (150-450) and glutamine 1014 µmol/L (480-800). Plasma urate was mildly raised and at four years of age, she was found to have right-sided hydronephrosis due to a urate renal calculus. Porphyria screen was normal and no mutation was found when sequencing of exons 9 and 11 of the mevalonate kinase gene (performed because of the recurrent fevers). Muscle biopsy demonstrated markedly reduced COX activity expressed as a ratio to citrate synthase at 0.007 (0.014-0.034) (Table 6.1). She died after multi-organ failure and presumed cephalosporin-induced toxic epidermal necrolysis at 14 years.

Her younger brother S11 presented in infancy with global developmental delay and hypotonia was found to have a microcytic anaemia. Initially he presented at two months with fever, vomiting and diarrhoea. Anaemia (Hb 71 g/L) and pan-hypogammaglobulinaemia were noted (Table 6.1). His blood film demonstrated red cell abnormalities such as poikilocytosis, microcytosis, elliptocytosis, hypochromia and polychromasia, similar to the findings in his sister S10. His fever recurred weekly, and was associated with vomiting and diarrhoea. Pearson's syndrome was excluded in both S10 and S11 by screening for mitochondrial DNA deletions. S11 was treated using bone marrow transplantation; after this the febrile episodes resolved, he made some developmental progress and his growth has improved. He remains well 3 years after bone marrow transplantation. Currently at 10.5 years of age he has begun to walk independently and is making progress with speaking. He continues to have moderate deafness and retinopathy.

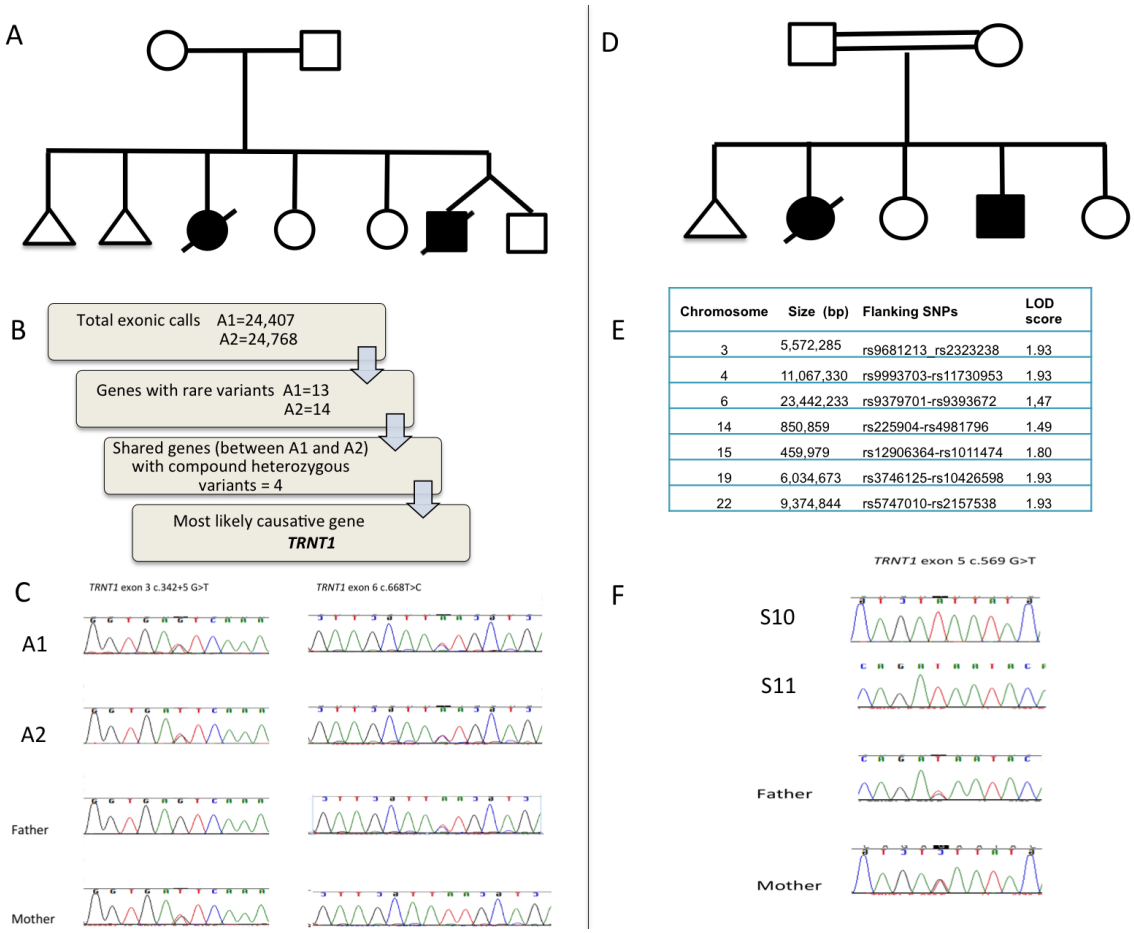


Figure 6:1: Genetic findings in patients A1, A2. S10 and S11.

(A) Pedigree of family with A1 and A2. (B) Bioinformatic analysis of whole exome data (C) Sanger sequencing results of A1 and A2 and parents (D) Pedigree of family of S10 and S11. (E) Homozygosity mapping data showing sizes of the chromosomal regions with the highest LOD scores. (F) Electropherogram demonstrating the *TRNT1* mutation in S10 and S11 (c.569G>T p.Arg190Ile) found by Chakraborty *et al.*

(As published in (Wedatilake et al. 2016))

Table 6.1: TRNT1 deficient patients: biochemical and other laboratory data

Analyte	A1	A2	S10	S11	S11: 2 years post bone marrow transplant	Reference range
Plasma lactate mmol/L	**2.5**	**3.1**	1.2-**2.8**	1.4-**2.7**	1.1	<2
CSF lactate mmol/L	**2.5**	NA	NA	NA	NA	<2
CSF protein g/L	**0.41-1.79**	NA	NA	NA	NA	<0.3
Plasma glycine µmol/L	**504**	248	280-332	310-**344**	235	100 - 330
Plasma threonine µmol/L	**256**	**305**	168-**266**	329-**556**	260	70 - 220
Plasma proline µmol/L	**246**	255	291-**452**	304-**414**	249	85 - 290
Plasma leucine µmol/L	91	133	140-**245**	96-198	129	65 - 220
Plasma isoleucine µmol/L	46	73	63-118	69-**114**	68	26 - 100
Plasma valine µmol/L	**142**	216	244-**404**	197-**323**	255	90 - 300
Plasma alanine µmol/L	296	**458**	562-**875**	572-**665**	346	150 - 450
Plasma ornithine µmol/L	101	107	83-**210**	112-**241**	145	25 - 120
White cell ubiquinone pmol/mg protein	NA	57.0	**10.0**	NA	NA	37 - 133
Creatine kinase u/L	NA	87-383	88	41-137	NA	75 - 230

Analyte	A1	A2	S10	S11	S11: 2 years post bone marrow transplant	Reference range
Electrolytes mmol/L	Episodic hyponatraemia (131), hypokalaemia (2.7), hypocalcaemia (1.7), hypomagnesaemia (0.51), hypophosphataemia (0.89)	Normal	Episodic hyponatraemia, hypokalaemia, hypocalcaemia, hypomagnesaemia, hypophosphatemia	Not done during acute episode of illness	Normal	
Intact PTH pmol/L	3.3	NA	8.0	NA	NA	1.1 - 5.4
Gamma GT u/L	300	normal	NA	NA	NA	10-20
Urine organic acids	Mildly raised 3- hydroxy butyrate with moderately raised adipate, suberate and 3OH-sebacate and mildly raised sebacate, C8:1 and C10:1 dicarboxylates, 3OH (C14:0, C8:1, C10:1) dicarboxylates	No abnormality	Strongly raised 3-hydroxybutyrate and acetoacetate	Mildly raised pyruvate	NA	NA
Urine amino acids	NA	NA	Generalised aminoaciduria	NA	NA	NA
Urine NAG/creatinine	612 – 2606	217	666	NA	NA	3.5 - 27.3

Analyte	A1	A2	S10	S11	S11: 2 years post bone marrow transplant	Reference range
Urine RBP/creatinine	4531 – 25714	961	419	NA	NA	1.9 - 42.6
Hb g/L	Lowest 79 (frequent transfusions)	66, 67	65-74	71, 68	NA (Sample clotted)	114-145
Blood film	Anisocytosis, neutropaenia dimorphic red cells, pencil cells. Elliptocytes on film.	Marked anisopoikilocytosis with hypochromic red cells, elliptocytes and fragments	Marked red cell anisopoikilocytosis. Many hypochromic, microcytic red cells and elliptocytes. Numerous red cell fragments. Occasional target cells	Neutrophils: toxic granulation. Poikilocytosis, microcytosis, eliptocytosis, hypochromic red cells.	NA	
IgG g/L	**2.24**-5.42 (3.7 – 15.8)	**0.86** (3 - 10.9)	**3.95**, 10.6 (5.4 - 16.1)	**4.11-4.32** (4.9 - 16.1)	10.90	
IgA g/L	**<0.03** (0.3 – 1.3)	**<0.06** (0.2 - 0.7)	**0.15**-0.62 (0.5 - 1.8)	**0.17**-0.41 (0.4 - 2.0)	0.76	

Analyte	A1	A2	S10	S11	S11: 2 years post bone marrow transplant	Reference range
IgM g/L	**0.11-0.23** (0.5 - 2.2)	**0.17** (0.6 - 2.1)	**0.09, 0.20** (0.5 - 2.2)	**0.29**-1.19 (0.5 - 2.0)	0.58	
Lymphocyte Subsets	Absent B cells	Low B cells	Low B cells	NA	NA	
Other	Moderate exocrine pancreatic insufficiency (stool elastase 182 µg/g, normal >200); later undetectable Hair shafts normal histology. Duodenal biopsy: Partial villous atrophy: Liver Iron 143 µg/100mL dry weight Normal VLCFA	Stool elastase undetectable	Normal ammonia, blood spot carnitine profile, VLCFA, copper, caeruloplasmin, plasma methylmalonic acid, autoimmune profile; Large bowel histology – mild patchy increase in eosinophils in lamina propria	Normal transferrin glycoforms, blood spot carnitine profile, VLCFA, urate, copper, caeruloplasmin, plasma methyl malonic acid	NA	
Muscle histology	No ragged-red fibres or COX negative fibres	NA	NA	NA	NA	

Analyte	A1	A2	S10	S11	S11: 2 years post bone marrow transplant	Reference range
Muscle respiratory chain enzymes (ratio to citrate synthase)	Complex I 0.139 Complex II+III 0.064 COX **0.015**	NA	Complex I 0.228 Complex II+III 0.068 COX **0.007**	NA	NA	Complex I 0.104-0.268 Complex II+III 0.040-0.204 COX 0.014-0.034

NA Not available, NAG N-acetyl-beta-D-glucosaminidase, PTH parathyroid hormone, RBP retinol binding protein, VLCFA very long chain fatty acids

(table as published in (Wedatilake et al. 2016))

6.3.2 Genetic findings: whole exome sequencing

After filtering, only four genes (*SRRM, CEL, SLC12A1, TRNT1*) with compound heterozygous variants shared between the affected siblings A1 and A2, were found (Figure 6.1B). Of these given the clinical findings, *TRNT1* was the most likely causative gene. A novel splice mutation c.342+5G>T and a known pathogenic missense mutation, c.668T>C p.Ile223Thr were found. Parents were heterozygous for each mutation, confirming segregation with disease in this family (Figure 6.1C).

6.3.3 Genetic findings: homozygosity mapping

Homozygosity mapping was performed prior to the adoption of WES techniques in S10, S11, parents and one sibling unaffected by this syndrome. The mapping data revealed seven homozygous regions shared between the two affected siblings encompassing 56.8 megabases, including four regions with a LOD score (logarithm of the odds to the base 10) of 1.93 (Figure 1E) containing 716 genes. There were 33 genes with a Maestro (predicted mitochondrial localisation) score >1 of which 11 genes had a Maestro score >10 (*TRNT1*, *GRPEL1*, *LONP1*, *NDUFA11*, *CLPP*, *TIMM44*, *NDUFA7*, *ATP6V1E1*, *SLC25A1*, *TXNRD2* and *CHCHD10*). The *TRNT1* mutation (c.569G>T, p. Arg190Ile) illustrated in Figure 6.1F in this family has previously been published by another group (Chakraborty et al. 2014) who were also investigating the genetic cause of disease in this family, unbeknownst to our research group.

6.3.4 Mitochondrial translation and analysis of mitochondrial tRNAs

Mitochondrial translation rate was normal in A1 and A2 compared to a healthy control and to a disease control (patient without mitochondrial disease) (Figure 6.2A). The purpose of this experiment was to identify if *TRNT1* mutations resulted in reduced CCA addition to mitochondrial tRNAs (mt-tRNAs). High resolution northern blotting (Figure 6.2C) was performed using total RNA isolated from A1 and A2 primary fibroblasts. There was partial lack of CCA in mt-tRNA^{Cys} (A1: 5%, A2: 5%, C1: <1%, C2: <1%) and mt-tRNA^{LeuUUR} (A1: 2%, A2: 2%, C1: <1%), when compared to controls. However, there were no obvious differences in CCA addition for other mt-tRNAs tested, including mt-tRNA^{Ile}, mt-tRNA^{Phe}, - mt-tRNA^{His}, mt-tRNA^{Arg}, mt-tRNA^{Pro}, - mt-tRNA$^{Ser(AGY)}$. To investigate subtle differences in maturation of the 3' end of mt-tRNA, which were possibly beyond Northern blotting detection limits, the ends of mt-tRNA after circularisation, reverse transcription, PCR amplification and cloning (cRT–PCR) were analysed. The cRT-PCR analysis of mt-tRNA^{SerAGY} and mt-tRNA^{Cys} was consistent with the high-resolution northern blotting analysis where there were a proportion of incorrectly processed mt-tRNA^{Cys} (Figure 6.2D) was found. Also in keeping with the findings in the northern blots, there were no defects in maturation of mt-tRNA^{SerAGY}. In addition mt-tRNA^{His} was tested using cRT-PCR and a proportion of incorrectly processed mt-tRNA^{His} was found.

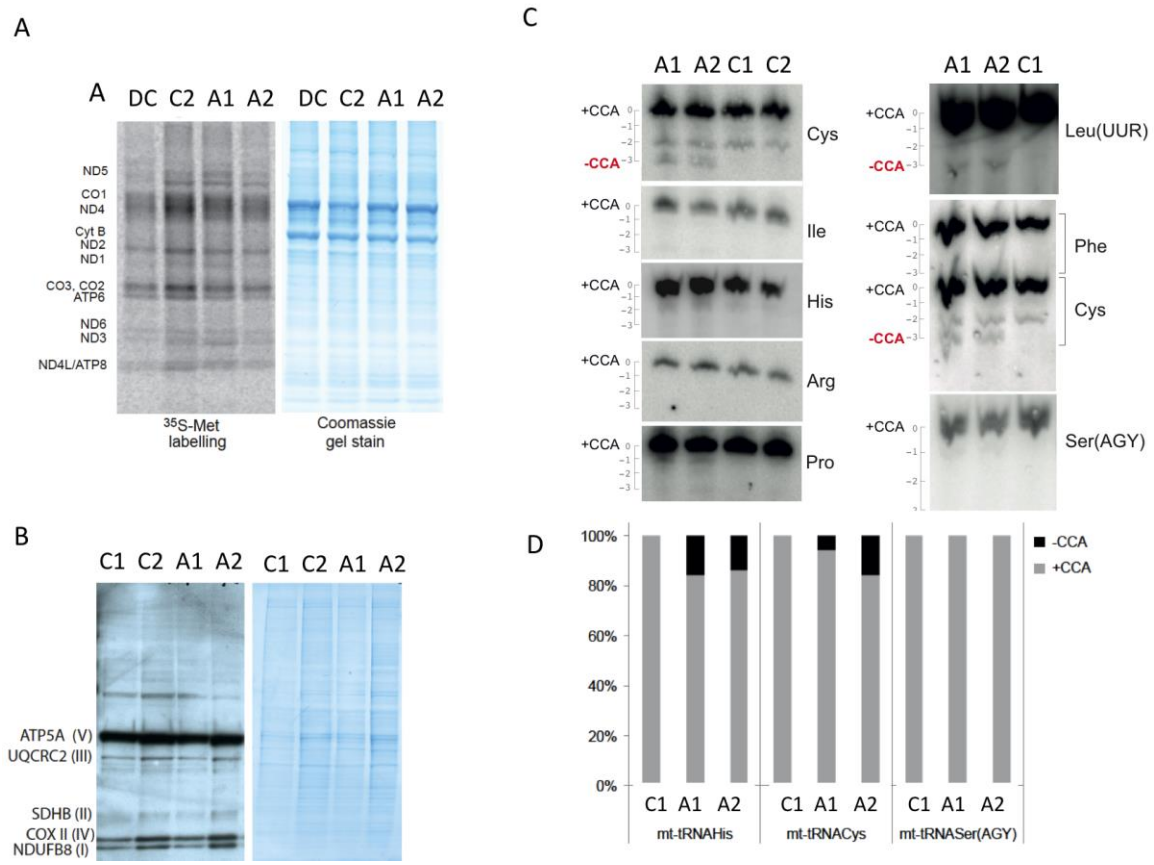


Figure 6:2: Mitochondrial translation assay, western blot, northern blot and cRT-PCR findings in A1 and A2.

(A) Mitochondrial translation assay in disease control (DC), healthy control (C2) and patient (A1, A2) fibroblasts. Coomassie blue used as a loading control. The mitochondrial translation products are shown on the left: four subunits of complex I (ND), one subunit of complex III (cyt b), three subunits of complex IV (CO) and two subunits of complex V (ATP). (B) Western blot for OXPHOS proteins (NDUFB8, MT-CO2 (COXII), SDHB, UQCRC2 and ATP5A) in A1 and A2 compared to controls C1 and C2. The OXPHOS subunits detected are shown to the left with their respective complexes in parentheses. (C) High-resolution northern blot analysis of total RNA isolated from the TRNT1 deficient patients (A1 and A2) or control (C1) primary fibroblasts. The blots were probed with the mt-tRNA-specific RNA probes as indicated. Full-length mt-tRNAs are indicated as +CCA. Mt-tRNAs lacking CCA are highlighted by "-CCA". (D) The results of cRT-PCR analysis for mt-tRNA^{His}, mt-tRNA^{Ser} and mt-tRNA^{Cys} in control and TRNT1 deficient patients A1 and A2. Data provided by Dr Michal Minczuk. (as published in (Wedatilake et al. 2016))

6.3.5 Analysis of all reported patients: initial presentation

The clinical features in 18 subjects reported to date were studied (including the cases in this study). There is clearly a spectrum of disease ranging from a neonatal-onset disease to an adult presentation. The earliest initial presentation was three hours after birth when pallor and splenomegaly was noted in a newborn. More than half (10/18, 55%) of the 18 cases reported to date presented in the first two months of life (range 1 day-19 years). The salient initial feature

(9/18, 50%) was a systemic 'inflammatory' illness where an infective cause was not identified. This 'inflammatory' episode was accompanied by gastrointestinal manifestations such as poor feeding, diarrhoea (8/18, 44%) and vomiting (7/18, 39%). In one adult patient (DeLuca et al. 2016) a febrile illness was recorded in childhood, and was attributed to juvenile rheumatoid arthritis. In one patient with childhood-onset disease (Sasarman et al. 2015) lactic acidosis and vomiting was noted but fever was not reported. The 'inflammatory' episodes appeared to occur cyclically. Other initial presentations included anaemia and developmental delay. The adult patients presented clinically between 18-21 years with visual impairment. Table 6.2 shows the clinical features and laboratory findings in all published patients.

Table 6.2: **Clinical features of TRNT1 deficiency**

Clinical feature	**Total**	**%**
Symptom/sign		
Recurrent 'inflammatory' episodes	14/18	78
Developmental delay	14/18	78
Sideroblastic anaemia	13/18	72
Diarrhoea	8/18	44
Vomiting	7/18	39
Sensorineural deafness	7/18	39
Seizures	7/18	39
Retinitis pigmentosa	6/18	33
Splenomegaly	6/18	33
Ataxia	5/18	28
Brittle hair	5/18	28
Hypotonia	5/18	28
Nephrocalcinosis	5/18	28
Renal tubulopathy	4/18	22
Hepatomegaly	4/18	22
Pancreatic insufficiency	3/18	17
Villous atrophy	2/18	11
Acute encephalopathy	2/18	11
Cardiomyopathy	1/18	6
Laboratory investigations		
Low or low-normal haemoglobin	16/18	89
Microcytosis	16/18	89
B lymphopaenia/ hypogammaglobulinaemia	12/18	67
Anisocytosis	9/18	50
High lactate	6/18	33
Metabolic acidosis	5/18	28
High alanine	3/18	17

6.3.6 Analysis of all reported patients: blood and immune system

Sideroblastic anaemia (13/18, 72%) is a prominent feature of the clinical syndrome. In most cases anaemia was found during the investigation of an unexplained febrile illness. Interestingly the three adult patients had a Hb in the lower end of the reference range with an asymptomatic microcytosis (~130 g/L) In contrast, the childhood-onset cases developed a severe anaemia (Hb range 48-82 g/L). One child was reported to present with a neurological phenotype with no febrile episodes or sideroblastic anaemia (Sasarman et al. 2015). Sixteen (89%) cases had evidence of anaemia (mild in the adult cases) with microcytosis (median Hb 72 g/L, median MCV 63.4 fl). The majority of cases (12/18, 67%) also developed B cell lymphopaenia with or without an accompanying hypogammaglobulinaemia. Only two cases were found to have a low total white cell count.

6.3.7 Analysis of all reported patients: retinitis pigmentosa

Of the four cases (A1-A4) followed up by the metabolic service at our institution, one patient developed retinal dystrophy and two other cases were found to have pigmentary retinopathy. In all 18 cases studied, six cases (33%) had retinitis pigmentosa; three of these were adults. In the adult patients this was the only apparent symptomatic feature. One adult case reported night blindness as a child.

6.3.8 Analysis of all reported patients: neurological features

Most patients (14/18, 78%) reported delayed development; this was global except in one case reported to have isolated motor delay. Seven patients developed seizures, six patients had ataxia and five patients had hypotonia. Sensorineural deafness was found in 7/18 (39%) patients. Two of the cases identified at our institution (A1 and A2) developed episodes of encephalopathy. One patient was reported to present with ptosis and ophthalmoplegia (Sasarman et al. 2015). Where neuroimaging was performed, four patients developed cerebral atrophy, two were noted to have changes in the brainstem, two had cerebellar changes (atrophy in one, cerebellar hypoperfusion in the other) and one had changes in the thalamus. Two of the cases reported here (A1 and A2) also demonstrated multiple cerebral and cerebellar focal lesions reminiscent of infective changes (Figure 6.3).

Figure 6:3: MRI brain scan findings in A2.

Axial (A) and Coronal (B) T2-weighted images and Sagittal (C) post-contrast T1 weighted images of the in A2, showing multiple focal lesions in the cerebral hemispheres and cerebellum involving both grey and white matter with contrast enhancement. Bilateral subdural collections were also present. (as published in (Wedatilake et al. 2016))

6.3.9 Analysis of all reported patients: kidney, liver and spleen involvement

One case followed up at our institution (S10) developed a urate calculus. Six subjects (33%) were found to have splenomegaly. Five subjects developed nephrocalcinosis (28%) and four (22%) had hepatomegaly.

6.3.10 Analysis of all reported patients: gastrointestinal features

As described above cyclical vomiting (7/18) and diarrhoea (8/18) accompanied the febrile episodes. Three cases had pancreatic insufficiency and two had partial villous atrophy on duodenal biopsy.

6.3.11 Analysis of all reported patients: cardiomyopathy features

Only one case developed (dilated) cardiomyopathy.

6.3.12 Analysis of all reported patients: additional features

Brittle hair was found in five cases.

6.3.13 Analysis of all reported patients: survival

The median age of death in all patients analysed was 37.5 months (range 10 months - 14 years). The three adult patients were alive and systemically well at 18, 19 and 21 years. Two paediatric patients (Wiseman et al. 2013) underwent myeloablative allogenic bone marrow transplant. Both these patients are alive with no recurrence of fever, at three years post bone marrow transplantation.

6.3.14 Analysis of all reported patients: metabolic testing

When metabolic and biochemical findings were analysed in all cases, an elevated blood lactate was found in six cases, being associated with metabolic acidosis in five cases, an elevated plasma alanine was noted in three cases and evidence of renal tubulopathy in four cases.

Detailed metabolic investigations were performed in the patients seen at our institution (Table 6.1). Blood lactate was mildly elevated (2.4-3.1 mmol/L, reference range <2.1) in all four subjects described here in the current study. Plasma alanine was raised in three patients (patient range 458-875; reference range 150-450 µmol/L) and proline was raised in two patients (patient range 304-452; reference range 85–290 µmol/L). Branched chain amino acids and ornithine were elevated in S10 and S11. Plasma threonine was elevated in all four patients (patient range 256-556; reference range 70–220 µmol/L) and plasma glycine was raised in three patients (patient range 332-504; reference range 100–330 µmol/L). S11 was observed to have a normal amino acid profile post bone marrow transplant.

Electrolyte abnormalities in A1 and S10 observed during acute febrile episodes included hypokalaemia, hyponatraemia, hypocalcaemia, hypomagnesaemia and hypophosphataemia. Renal tubular dysfunction was found in three patients (A1, A2 and S10; not done in S11), where

high urinary NAG:creatinine and RBP:creatinine ratios were found. In one patient (A1) elevated hepatic transaminases were consistently noted.

6.3.15 Analysis of all reported patients: respiratory chain enzymes

Muscle RCEs were measured in 2/4 patients (A1 and S10) described here in detail. S10 had definitively low COX of 0.007, (reference range 0.014-0.034) whereas A1 had a normal value which was at lower end of the reference range (0.015) (Table 6.1). In one patient (N) reported previously (Sasarman et al. 2015) COX activity was decreased in muscle (0.25, control range 0.58-2.50) and liver (0.25, control range 0.60-2.50). In addition in their patient, complex I+III (0.15, control range 0.58-2.5) and complex II+III (0.06, control range 0.50-1.90) were also decreased in muscle.

6.3.16 Analysis of all reported patients: genotypes and protein modelling

All *TRNT1* mutations published to date are demonstrated in Figure 6.4A. All missense mutations reported to date are modelled *in silico* on a ribbon diagram of the protein in Figure 6.4B. The homozygous p.Arg190Ile missense mutation found in S10 and S11 is predicted to cause a severe conformational change in the protein because the hydrogen bond network formed by arginine is disrupted by isoleucine. This leaves the protein either unfolded, or folded yet dysfunctional (Figure 6.4C).

The missense mutation found in A1 and A2 (c.688C>T, p.Ile223Thr) was also found in paediatric patients from eight unrelated pedigrees of white Caucasian, Afro-Caribbean and Hispanic origins reported previously (Chakraborty et al. 2014), suggestive of a mutational hotspot in this region. *In silico* modelling of p.Ile223Thr showed that in the wild type protein isoleucine is a buried amino acid (Figure 6.4D). The threonine which is substituted in the missense mutation can also be buried satisfying its H-bond and therefore would result in a mild phenotype.

All three adult patients demonstrated a single nucleotide deletion or insertion of an adenine at the same position (c.1246) on one allele. These changes affected the last exon in the gene resulting in a truncated protein. Interestingly the c.1246del found in the adult patients was also found in a paediatric patient with the more severe disease (Chakraborty et al. 2014). In the adult patients the opposite allele contained either a deletion or splicing mutation; the authors predicted that the deletion only caused a minor conformational change in the protein and that the splicing was affected only in a subgroup of the transcripts (DeLuca et al. 2016).

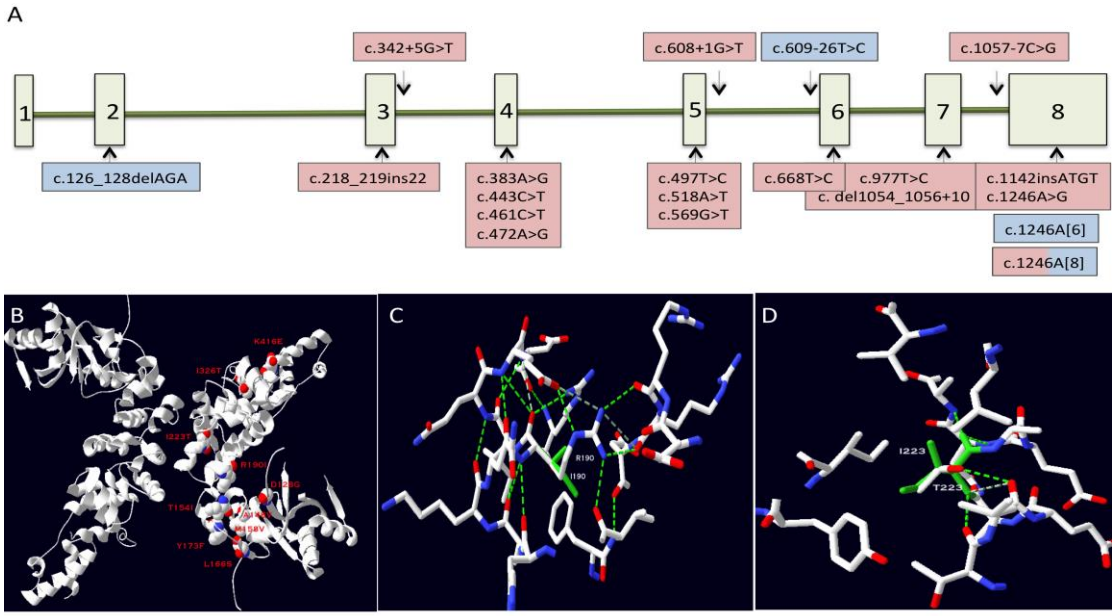


Figure 6:4: Schematic diagram of the *TRNT1* gene and modelling of mutations in the TRNT1 protein.

(A) Schematic diagram of the *TRNT1* gene demonstrating the mutations reported to date. Mutations found in the milder adult phenotype are shaded in blue and those found in childhood-onset severe disease are shaded in pink. (B) The human TRNT1 protein is shown in ribbon representation. All missense mutations published to date are shown as space fill atoms, white is for carbon, red is for oxygen and blue is for nitrogen. The mutant residue is shown at each position and labelled. (C) A representation of the environment surrounding the p.Arg190Ile mutation. Carbon atoms in white, oxygen in red and nitrogen in blue. Ile190 is shown in green throughout. Hydrogen bonds are shown as dotted lines. The extensive hydrogen bond network stabilising the wild type p.Arg190 is completely lost following mutation to the hydrophobic Ile190. (D) A representation of the p.Ile223Thr mutation. Carbon atoms are in white, oxygen is in red and nitrogen is in blue. Wild type p.Ile223 is shown in green. Hydrogen bonds are shown as dotted lines. The protein is able to accommodate the Thr223 mutation in place of the buried, hydrophobic p.Ile223 through the hydrogen bond from the p.Thr223 side chain. (as published in (Wedatilake et al. 2016))

6.4 Discussion

This study identified a new family with *TRNT1* mutations, including a novel splicing mutation, and showed that *TRNT1* mutations cause a complex multisystem disease leading to manifestations in most organs. In this chapter a review of all previously reported cases was used to typify the phenotypic presentation of TRNT1 deficiency, a new metabolic disorder. This study provides evidence of mitochondrial pathogenicity in this disease, demonstrating functional impaired addition of CCA to mt-tRNAs in patient fibroblasts, presumably leading to defective mt-tRNA post-transcriptional modification.

The published reports and the findings in this study indicate that there are two ends of the spectrum of TRNT1-related disease. On one hand there is a severe childhood-onset disorder which starts usually in infancy with cyclical, aseptic, febrile episodes accompanied by gastrointestinal manifestations and sideroblastic anaemia. On the other hand the adult-onset disease presents with retinitis pigmentosa as the only discernible feature and asymptomatic microcytosis on haemological indices (Wiseman et al. 2013; Chakraborty et al. 2014; DeLuca et al. 2016; Sasarman et al. 2015). Pigmentary retinopathy, the only sign found in adult-onset disease was also found in three paediatric cases suggesting that it may be a key clinical feature of TRNT1 deficiency. Indeed ophthalmological features in a multisystem disorder may be undetected and under reported unless specialist screening is undertaken by an ophthalmologist. Other clinical features found in this syndrome included splenomegaly, nephrocalcinosis and hepatomegaly (reported in six, five and four cases respectively).

There was evidence of OXPHOS impairment in TRNT1 deficiency, with evidence of COX being low in one patient and in the lower end of the reference range for one patient. Among the patients published to date muscle RCE assays were not routinely performed. Therefore it is not known whether OXPHOS abnormalities are a definitive feature of TRNT1 deficiency. Given the pivotal role of TRNT1 in CCA addition to mt-tRNAs, multiple RCE abnormalities would be expected. Similar to AARS2 deficiency in chapter 3, the observation of mitochondrial translation disorders with isolated COX deficiency is similar to other mitochondrial translation disorders (e.g. RARS2 and YARS2 deficiency) where biochemical findings ranged from normal RCE activities to multiple RCE defects or isolated COX deficiency (Glamuzina et al. 2012; Sasarman et al. 2012). TRNT1 enzyme is a ubiquitous enzyme and therefore multi-organ involvement in these patients is perhaps not unexpected.

There were variable changes in the plasma amino acid profiles from the four children described with TRNT1 deficiency in this chapter. One explanation could be that impaired muscle metabolism leads to decreased breakdown of branched chain amino acids. On the other hand the amino acid abnormalities could be related to TRNT1 deficiency where there could conceivably be a backlog from disordered cellular protein translation. Further studies will be needed to establish this as it is difficult to determine the reasons for these changes given the small number of patients studied here.

Patients with inherited metabolic disease often decompensate with intercurrent febrile illness. During these acute phases metabolic physicians are involved in the care of these patients. In the four cases discussed here, febrile illness was a component of the disease rather than a precipitant. In addition the clinical syndrome found in TRNT1 deficiency can be confused with severe sepsis given the febrile episodes with raised inflammatory markers. In the first instance, sepsis must obviously be excluded. However indications at this stage that this may be TRNT1 disease are: the presence of anisocytosis, microcytic anaemia and B cell lymphopaenia (typically with hypogammaglobulinaemia).

The childhood onset TRNT1-deficient disease enters into the differential diagnosis of congenital sideroblastic anaemia, genetic immune deficiency and autoimmune inflammatory disease. Other congenital sideroblastic anaemias which present similarly and may need exclusion include Pearson syndrome (sideroblastic anaemia with exocrine pancreatic failure and pancytopaenia) (Broomfield et al. 2015); and PUS1 and YARS2 deficiency which cause myopathy, lactic acidosis and sideroblastic anaemia (MLASA) syndrome (Riley et al. 2010; Shahni et al. 2013; Bykhovskaya et al. 2004). These diseases can be distinguished clinically from TRNT1 deficiency since cyclical febrile episodes and B cell lymphopaenia are not features of these diseases.

The acute management of the patients studied here consisted of fluid and electrolyte replacement and blood transfusion for anaemia. So far two patients have received bone marrow transplantation and preliminary results appear encouraging; however more long-term data are needed. Ideal evidence, in the form of a randomised double-blinded placebo-controlled clinical trial is unlikely to be possible in this very rare disease.

The TRNT1 enzyme is bifunctional in both the cytosol and the mitochondrion. The functional implications on cytosolic translation have not been investigated here. TRNT1 deficiency presents with a multisystem disease which in some aspects resembles Pearson syndrome, a mitochondrial disease, suggesting that post-transcriptional modification (CCA addition) of mitochondrial tRNAs, may be critical to disease pathogenesis. The impact of the TRNT1 mutations on tRNA modification in primary patient fibroblast cell cultures was investigated and the findings show that mt-tRNAs partially lacked 3' CCA in mt-tRNA^{Cys}, mt-tRNA^{LeuUUR} and mt-tRNA^{His}. This finding contrasts with a previous study which suggested that only mt-tRNA$^{Ser(AGY)}$ was affected (Sasarman et al. 2015). The reasons for these differences in CCA addition are currently unclear with one plausible explanation being that it is related to severity of the *TRNT1* mutation.

In summary *TRNT1* mutations were identified in patients with a complex, childhood-onset, multisystem disorder who fall into the SIFD spectrum of TRNT1 disease. In addition adult-onset retinitis pigmentosa has been found TRNT1 deficiency. TRNT1 deficiency has an impact on mt-

tRNA post-transcriptional modification in affected patient cells. However the precise pathogenic mechanism of TRNT1 deficiency in this multisystem disease requires further elucidation and the cytosolic aspect of its function requires further study. Review of all cases reported to date helped study the complete clinical phenotype to aid recognition of further cases. Early bone marrow transplantation should be considered in TRNT1 deficiency to treat the febrile episodes and metabolic decompensations, in order to improve quality of life.

Chapter 7

7. The natural history of SURF1 deficiency, a monogenic mitochondrial disease causing isolated cytochrome *c* oxidase deficiency

7.1 Introduction

SURF1 deficiency is a recessively inherited mitochondrial disease and is the most common cause of Leigh syndrome associated with COX deficiency. As discussed in chapter 1, a number of assembly factors are required for the intricate biogenesis of the COX holoenzyme and SURF1 has been identified as a key player in the early assembly of the COX enzyme.

The *SURF1* gene is located on chromosome 9p34 in a region which is well conserved from yeast to humans. In vertebrates, it is found among a collection of genes, which are termed the surfeit genes. The *SURF1* gene encodes a 300 amino acid protein, which has been localised to the inner mitochondrial membrane. Although *SURF1* has been associated with human disease since 1998, (Tiranti et al. 1998; Zhu et al. 1998) its precise function remains poorly understood. SURF1 is believed to be involved in the formation of a COX assembly intermediate (S2) (Williams et al. 2004). Studies in *Paracoccus denitrificans* and *Saccharomyces cerevisiae* have identified that *Shy1*, the yeast orthologue, plays a role in haem 'a' insertion in the COX enzyme (Hannappel et al. 2012). In addition MITRAC12 (CCDC56), interacted with SURF1 and MT-CO1 forming a mitochondrial translation regulation assembly intermediate of MT-CO1 (Mick et al. 2012).

While SURF1 deficiency in humans is associated with a severe childhood-onset neurodegenerative disease the *Surf1* knock-out mice have a mild phenotype with paradoxically increased lifespan and mild COX deficiency (Dell'Agnello et al. 2007). This has been attributed to differences in the way that the COX enzyme behaves between species. For example, in both *Surf1* knock-out and wild-type mouse tissue the COX enzyme is less incorporated into supercomplexes and the COX monomer is more stable in *Surf1* knock out mouse tissue (Kovářová et al. 2016) than in humans.

Leigh syndrome, a fatal subacute necrotising encephalomyelopathy, is a genetically heterogeneous mitochondrial disorder which can be associated with any OXPHOS deficiency (Rahman et al. 1996). Since *SURF1* mutations were found to be disease causing (Tiranti et al. 1998; Zhu et al. 1998) isolated case reports and reports of mutation have been reported (Péquignot *et al*., 2001, Lee *et al*., 2012). However a complete description of the natural history including biochemical and radiological features of this disease has not yet been described. Additionally, there is limited data in the literature to provide prognostic information to patients. Comprehensive natural history studies will also inform future clinical trials.

In the previous chapters the focus was to uncover the genetic cause of suspected mitochondrial disease. This chapter is a study of cases with an established monogenic mitochondrial disorder,

in order to characterise its natural history. For this purpose patients with SURF1 deficiency, the most common cause of isolated COX deficiency, was identified as a suitable cohort. A multi-centre study of SURF1 deficient cases found from 1998-2013 were identified to study phenotypes, genotypes, biochemical data, natural disease progression and survival data.

7.2 Methods

7.2.1 Study cohort

The study design consisted of a multi-centre retrospective case notes review of all patients diagnosed since 1998. All patients diagnosed with *SURF1* mutations were identified through the NHS Specialised Services-funded mitochondrial molecular diagnostic laboratory in Oxford, UK, which provides a national diagnostic service for *SURF1* mutations in the UK, and through the Australasian paediatric diagnostic centre (Murdoch Childrens Research Institute, Melbourne). Subjects were included if they had a confirmed genetic diagnosis of SURF1 deficiency, defined as biallelic pathogenic variants in the *SURF1* gene. Patients who were deceased before the *SURF1* gene was identified were included, if subsequent analysis of stored samples revealed two pathogenic *SURF1* mutations. In some deceased patients stored biological material was not available, and they were included if they had an affected sibling with confirmed *SURF1* mutations and a similar disease course.

The following clinical data were collected using a structured questionnaire: age-of-onset of symptoms, nerve conduction studies, neuroimaging, muscle histology and histochemistry, plasma lactate, CSF lactate, and RCEs in muscle biopsies and cultured skin fibroblasts. The data were collected using a structured questionnaire completed by the responsible clinician at each centre. The questionnaire was developed and tested in a pilot study conducted at Great Ormond Street Hospital, London on ten patients with genetically confirmed SURF1 deficiency.

Informed consent of parents/legal guardians of the patients was obtained for genetic testing. Ethical approval for the study was obtained from the National Research Ethics Committee London Bloomsbury, UK and the Royal Children's Hospital Human Research Ethics Committee, Melbourne, Australia (HREC 32188A).

7.2.2 Systematic literature review

Previously published cases were identified by a systematic literature search in PubMed. The search was first conducted using the search terms "SURF1", "SURF-1", "Leigh syndrome", "cytochrome *c* oxidase deficiency", "complex IV deficiency", and "cox deficiency". The search was limited to human studies published after 1.1.1998 (the year SURF1 deficiency was first described) with no language restrictions. There were 2135 records, of these 1969 were excluded after reviewing titles, 110 were excluded after reviewing abstracts, and 56 full-text records were retrieved. Of these, 43 records describing 129 cases with SURF1 deficiency were reviewed. Individual bibliographies were searched manually for additional records. The cases identified in the literature were used to establish all previously identified phenotypes and

genotypes, and for Kaplan-Meier survival analysis. These cases in the literature were not included in the analysis of the natural history, as data were incomplete with the chronological onset of symptoms often being unavailable. Authors were contacted and raw data was obtained for survival rates of LRPPRC-deficient Leigh syndrome patients (Debray *et al*., 2011) and nuclear-encoded complex I-deficient Leigh syndrome patients (Fassone and Rahman, 2012) in order to compare survival rates between the different diseases.

The literature search was updated in June 2016 when a further 688 records were identified. Of these only 12 records were deemed relevant after reviewing all titles. Eleven were excluded on the basis of abstracts.

7.2.3 Statistical analysis

Survival rates for the SURF1 deficiency, LRPPRC-deficient Leigh syndrome and nuclear-encoded complex I-deficient Leigh syndrome patients were compared using Kaplan-Meier survival analysis. Mortality rates were compared using the logrank test. Statistical analysis was performed using Stata 12.1 software (Stata, College Station, Tex).

7.3 Results

7.3.1 Demography

Fifty-seven patients with SURF1 disease born between 1976 and 2010 were identified at ten centres in the UK (n=52) and two centres in Australia (n=5). In 54 patients *SURF1* mutant genotypes were established. Three deceased patients were presumed to have the same genotype as their similarly affected sibling with genetically confirmed SURF1 deficiency. Clinical data to document natural history were available in 45 patients. In the remaining 12 cases it was not possible to trace clinical records and/or the responsible clinician. One patient from a consanguineous pedigree was found to have genetically confirmed Zellweger syndrome in addition to SURF1 disease (homozygous c.799_800delCT) and was excluded from further study. There were 44 patients (24 males, 55%) from 37 pedigrees, of which 10 (23%) patients were from consanguineous families. The largest ethnic group was white European (n=20, 45%) after which Bangladesh Bengali (n=5, 11%) Indian Gujarati (n=5, 11%) and Pakistani (n=4, 9%) were the next most frequent ethnicities. There were two Turkish Cypriot patients, one Chinese, one Sri Lankan and six cases where the ethnicity was unavailable.

7.3.2 Initial symptomatology

The onset of first symptoms was at a median age of 9.5 months (range 0-60 months). Most of the patients (32/44, 73%) presented in the first year. The most frequently noted initial symptoms were poor feeding/vomiting and poor weight gain (Table 7.1, Figure 7.1). Poor feeding and vomiting was often thought to be due to gastro-oesphageal reflux disease. In the majority of cases the neonatal period was relatively asymptomatic (41/44, 93%) except in two patients with feeding problems and one patient found to be hypotonic. Interestingly developmental regression (loss of cognitive or motor skills) was the initial symptom in 3/44 (7%) patients. The majority of

cases (26/44, 59%) developed a combination of gastro-intestinal symptoms, poor weight gain and hypotonia as the initial presentation. There were three subjects who had to undergo a general anaesthetic in the first two months of life; they experienced significant problems during the post-operative recovery period. Case 30 (Table 7.1) who underwent major cardiovascular surgery for transposition of the great arteries on day 9, developed persistent vomiting and required enteral feeding. Case 26 developed hypotonia after surgery for pyloric stenosis at 8 weeks. Case 4 developed a movement disorder after an upper gastro-intestinal endoscopy.

7.3.3 Clinical features

Of the 44 cases, 32 patients met the criteria for Leigh syndrome (Rahman *et al*., 1996). Twelve patients were classified as “Leigh-like” since they had atypical or normal neuroradiology, or neuroimaging was unavailable or had not been performed. Figure 7.1 graphically demonstrates the development of clinical features over time. Poor feeding and vomiting preceded the onset of the most commonly observed symptoms, poor growth and hypotonia (Figure 7.1). Thus in 21/44 (48%) long term enteral feeding was necessary. The median age of onset for developmental regression was 19 months and was noted in 27/38 (71%) patients. Developmental regression was precipitated by an intercurrent viral illness in 12 patients.

In 18/44 (41%) patients, hypertrichosis was observed. Nystagmus was noted in 25/42 (60%) and ophthalmoplegia was found in and 22/42 (52%) patients. These signs tended to occur later in the disease course (median age-of-onset 29 months). A movement disorder was found in 22/42 (52%) patients (median age-of-onset 2 years). Thirteen patients had an isolated intention tremor, three patients had choreoathetoid movements and two patients had dystonia whereas four patients demonstrated a combination of these abnormal movements. Seizures were observed in only 6/44 (14%) (generalised tonic clonic n=5, myoclonic n=1) and four of these cases were from Australia. Other less frequently occurring clinical signs included optic atrophy 10/44 (23%), encephalopathy 9/44 (20%) and hypertrophic cardiomyopathy 1/44 (2%). Interestingly none of the cases studied were reported to have pigmentary retinopathy or sensorineural deafness.

Table 7.1: Clinical features and genotypes of 44 patients with SURF1 deficiency

Case	Sex	Consan-guinity	Clinical features	Age at onset* (months)	Age at death* (months)	Muscle histology	Muscle COX	Fibrobl-ast COX	*SURF1* Mutation
1a	M	Yes	PF/V, PW, nystagmus, RF	8	2y 1m	NA	NA	29% of LLR	Homo c.792_793delAG
1b	F	Yes	PW, hypotonia, DR, DD, hypertrichosis, nystagmus, RF	9	3y	T1FP, increased lipid	Low	UD	Homo c.792_793delAG
2	F	No	Ataxia, hypotonia, nystagmus, ophthalmoplegia, PW, tremor	18	Alive 12y	Smaller T1 fibres	Low	19% of LLR	c.240+1G>T, c.575G>A
3	F	No	DD, hypotonia, nystagmus, PW, tremor, ophthalmoplegia, ataxia	birth	Alive 19y	Normal	Low	37% of LLR	c.312_320del10insATc.751+5G>A
4	F	No	PF, movement disorder, PW, DD, RF	7	16	NA	Low	57% of LLR	Homo c.312_320del10insAT
5a	M	No	PF, ataxia, nystagmus. ophthalmoplegia, tremor, PW,	10	5y 2m	Absent COX	Low	UD	Homo c.516-2A>G
5b	M	No	PF/V DD, DR, RF	3 days	20	Increased lipid	NA	NA	Homo c.516-2A>G

Case	Sex	Consan-guinity	Clinical features	Age at onset* (months)	Age at death* (months)	Muscle histology	Muscle COX	Fibrobl-ast COX	*SURF1* Mutation
6	M	No	V, PW, hypotonia, DR ophthalmoplegia, ataxia, encephalopathy	9	21	Increased lipid reduced	NA	12% of LLR	Homo c.324-11T>G
7	M	No	V, PW, DR, tremor, nystagmus, hypotonia, ophthalmoplegia, ataxia	10	4y	Increased lipid	NA	27% of LLR	Homo c.312_320del10insAT
8a	F	No	Hypertrichosis, DD, ataxia, hypotonia, nystagmus, ophthalmoplegia, encephalopathy, PF, OA, RF	9	11y 9m	Reduced COX T1FP, increased lipid	NA	UD	Homo c.516-2A>G
8b	F	No	V, DD, PW, hypertrichosis, hypotonia, nystagmus, ophthalmoplegia, ataxia, RF	10	6y 6m	NA	NA	NA	Homo c.516-2A>G
9a	M	No	V, PW, DD nystagmus, ophthalmoplegia, hypotonia	2	24m	Increased lipid reduced	NA	27% of LLR	Homo c.312_320del10insAT
9b	F	No	Hypotonia, PF/V, PW, DD, hypertrichosis, motor delay, ophthalmoplegia, RF, OA	1.5	4y	NA	NA	NA	Homo c.312_320del10insAT

Case	Sex	Consan-guinity	Clinical features	Age at onset* (months)	Age at death* (months)	Muscle histology	Muscle COX	Fibrobl-ast COX	*SURF1* Mutation
10	M	Yes	Hypotonia, PF/V, DD, nystagmus, PW, ataxia, RF	4	24m	T1FP	Low	50% of LLR	Homo c.516-2A>G
11	F	Yes	DD, DR, PF/V, PW, tremor, OA, PW, hypotonia, nystagmus, ataxia, RF	9	7y 1m	Absent COX increased lipid, T1FP	NA	7% of LLR	Homo c.751C>T
12	M	No	PF/V, DD, DR, hypotonia, encephalopathy, PW, RF	6	13m	NA	NA	NA	Homo c.312_320del10insATc.688 C>T ,
13a	F	Yes	DD, DR, PF/V, nystagmus, dystonia, PW, Sz, hypertrichosis, RF	12	5y 5m	T1FP, Reduced COX	Low	53% of LLR	Homo c.324-11T>G
13b	F	Yes	PF/ V, PW, hypotonia, DD, DR	15	NA	NA	Low	NA	Homo c.324-
14	F	Yes	DD, DR, hypotonia, PF/V, chorea, PW, nystagmus	18	NA	Reduced COX, increased lipid	Low	UD	Homo c.792_793delAG
15	M	No	Hypotonia, DD, DR, dystonia, ophthalmoplegia, PF/V, RF	2y 6m	4y 6m	Reduced COX, T1FP	Low	UD	Homo c.312_320del10in
16	M	Yes	PF/V, PW, hypotonia, DD, DR, encephalopathy, RF	9	18m	Reduced COX.	Low	17% of LLR	Homo c.324-11T>G

Case	Sex	Consan-guinity	Clinical features	Age at onset* (months)	Age at death* (months)	Muscle histology	Muscle COX	Fibrobl-ast COX	*SURF1* Mutation
17	M	No	Short stature, ataxia, PW, hypertrichosis, nystagmus, DD	5y	Alive 15y	Increased lipid, reduced COX	Low	NA	Homo c.312_320del10in sAT
18	F	Yes	PF/V, DR, PW, DD, hypotonia, tremor, dystonia, OA, ophthalmoplegia	10	8y 9m	Absent COX T1FP	Low	40% of LLR	Homo c.833+1G>A
19	M	No	Hypotonia, PF/V, PW, choreoathetoid movements, dystonia, DR, RF	10	3y 6m	Reduced COX T1FP	Low	NA	Homo c.312_320del10in sAT
20	M	No	Hypotonia, apnoeic episodes, motor delay, ataxia, DR, DD, tremor, hypertrichosis, dystonia, PF/V, PW, RF ophthalmoplegia, nystagmus, hypertrichosis	10	Alive 16y	Reduced COX	Low	NA	Homo c.704C>T
21a	M	No	V, PW, generalized hypotonia, nystagmus, ophthalmoplegia, DD, encephalopathy, RF	10	2y 3m	Absent COX, Increased lipid	Low	23% of LLR	Homo c.792_793delAG
21b	M	No	PF/V, PW, DD, DR, hypotonia, OA, ophthalmoplegia, nystagmus, hypertrichosis, RF	9	2y 5m	NA	NA	NA	Homo c.792_793delAG

Case	Sex	Consan-guinity	Clinical features	Age at onset* (months)	Age at death* (months)	Muscle histology	Muscle COX	Fibrobl-ast COX	*SURF1* Mutation
22	M	No	DD, DR, PW, ataxia, dysarthria, choking episodes, V, hypotonia, nystagmus, hypertrophic cardiomyopathy	4y 3m	NA	Absent COX	Low	NA	c.240+1G>T, c.574C>T
23	F	No	PF/V, DR, DD, PW, hypotonia, nystagmus, tremor, dystonia, hypertrichosis, ophthalmoplegia, encephalopathy, ataxia, RF	15	21m	Absent COX	Low	NA	Homo c.312_320del10insAT
24	F	No	PW, V, hypertrichosis, DD, hypotonia, ophthalmoplegia, tremor, Sz, ataxia, encephalopathy, RF	2	2y 10m	Absent COX	Low	NA	Homo c.324-11T>G
25a	M	No	PW, DD, ataxia, falls, OA, hypotonia, hypertrichosis, ophthalmoplegia, nystagmus, ataxia, RF	20	14y	NA	NA	NA	Homo c.871insT
25b	M	No	DD, PF/V, PW, choreoathetosis, ataxia, hypotonia, OA, encephalopathy, ataxia, RF	24	11y	Normal	NA	30% of LLR	Homo c.871insT

Case	Sex	Consan-guinity	Clinical features	Age at onset* (months)	Age at death* (months)	Muscle histology	Muscle COX	Fibrobl-ast COX	*SURF1* Mutation
26	M	No	Hypotonia, V, PW, ophthalmoplegia, DD, nystagmus, dyskinesia, choreoathetoid movements, RF	2	3y 10m	Increased lipid	NA	UD	c.312_320del10insATc.240+1G>T
27	M	No	DD, hypotonia, PF, PW, DR, RF	10	4y 9m	Reduced COX	Normal	UD	c.845_846delCT, c.240+1G>T
28	F	No	Hypotonia, PF/ V, PW, hypertrichosis, DD, DR, encephalopathy, nystagmus	3	Alive 24m	NA	NA	NA	Homo c.312_320del10insAT
29	F	Yes	PF, PW	birth	24m	NA	NA	17% of LLR	Homo c.799_800delCT
30	M	No	PW, DD, tremor, hypertrichosis. Post operative V, ophthalmoplegia, hypotonia	9 days	5y	NA	Low	NA	c.312_320del10insAT, c.871insT
31	F	No	PF/V, PW, DD, DR, tremor	1.5	3y 7m	Normal	Low	NA	Homo c.792_793delAG
32	F	No	PW, DD, DR, PF, hypotonia	12	2y 3m	Increased lipid	Low	NA	Homo c.754_755delAG
33	M	No	PF, PW, hypertrichosis, ataxia, DR, hypotonia, tremors, nystagmus, OA	14	7y 2m	Reduced COX	Low	32% of LLR	Homo c.312_320del10insAT

Case	Sex	Consan-guinity	Clinical features	Age at onset* (months)	Age at death* (months)	Muscle histology	Muscle COX	Fibrobl-ast COX	*SURF1* Mutation
34	M	No	PW, PF/V, DD, DR, hypertrichosis, tremor, hypotonia, ophthalmoplegia, ataxia, RF, Sz	13	24m	Absent COX increased lipid	Low	NA	c.312_320del10insA c.574insCTGC
35	F	No	PW, PF/V, DD, DR, hypertrichosis, hypotonia, ophthalmoplegia, OA, nystagmus, tremor, ataxia, RF, Sz	9	2y 6m	Absent COX, increased lipid	Post mortem	12% of LLR	c.312_320del10insAT c.751C>T
36	F	No	PW, PF/V, DD, DR, hypertrichosis, hypotonia, ophthalmoplegia, OA, nystagmus, tremor, ataxia, RF, Sz,	15	7y 10m	Reduced COX, increased lipid	Post mortem	44% of LLR	c.312_320del10insATc.688C>T
37	M	No	PF, ophthalmoplegia ataxia, hypotonia, Gross motor delay, PW, DR, DD, RF, Sz, hypertrichosis	birth	6y	Reduced COX	Low	NA	c.312_320del10insAT c.752 -2A>G

*Age in months unless otherwise stated; a, b denote siblings; y/m years and months; NA not available

Key to clinical features: COX cytochrome *c* oxidase, DD developmental delay, DR developmental regression, Homo homozygous, LLR lower limit of reference range, NA not available, PF poor feeding, PW poor weight gain, RF Respiratory failure, OA optic atrophy, Sz seizures, T1 type 1 muscle fibres, T1FP Type 1 muscle fibre predominance, T2 type 2 muscle fibres, UD undetectable, V vomiting.

(As published in (Wedatilake et al. 2013))

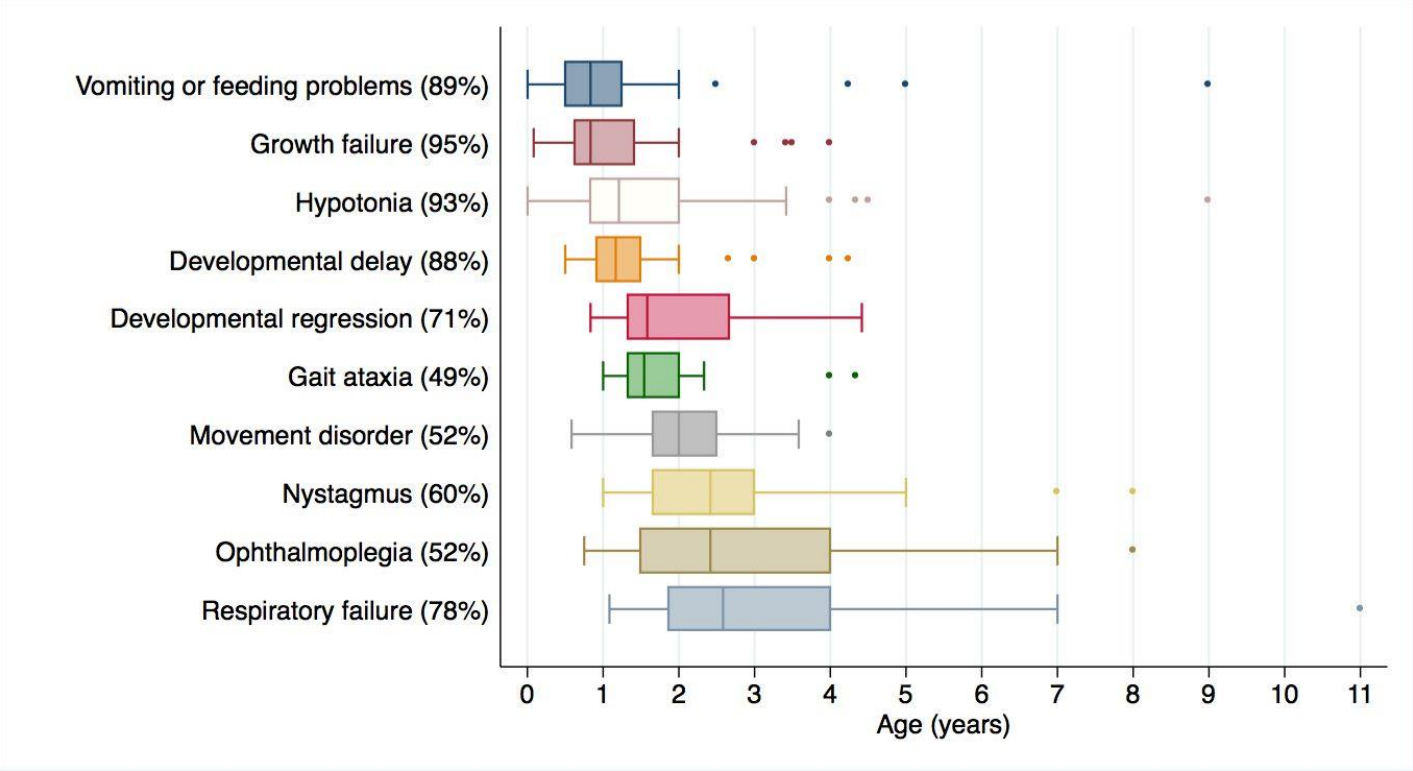


Figure 7:1: Clinical features in 44 patients with SURF1 deficiency.

The percentages represent the proportion of patients with each clinical feature, and box-and-whisker plots show the age of onset. The vertical line within the boxes indicates median age of onset. Boxes represent upper and lower quartiles, whiskers represent extreme values, and dots represent outliers which are ≥1.5 times the interquartile range from the median. Less frequently observed features not included in this figure included hypertrichosis (41%), optic atrophy (23%), encephalopathy (20%), seizures (14%) and cardiomyopathy (2%). (As published in (Wedatilake et al. 2013))

7.3.4 Survival

Of the 44 patients, five were alive at the time of analysis in 2013 (age range 2-19 years). In three of these cases the vital status at the time was not recorded. In the 36 deceased patients where the cause of death was available, 29/36 (80%) patients died of respiratory failure. Only seven patients survived beyond 10 years of age (patients 2, 3, 8a, 17, 20, 25a, 25b in Table 7.1). Six of these longer surviving patients presented with symptoms such as ataxia and motor developmental delay; gastro-intestinal symptoms were not the major presenting feature in these cases. In addition these six subjects also did not experience developmental regression during their clinical course.

Searching the literature as described above identified 98 SURF1-deficient cases with available survival data. This data was pooled together with the survival data from the 44 cases identified in this study. The survival of SURF1-deficient cases (n=142) was compared with two other groups with Leigh syndrome due to nuclear gene mutations; 56 patients with LRPPRC deficiency (Debray *et al*., 2011) and 63 patients with nuclear-encoded complex I-deficient Leigh syndrome/"Leigh-like" disease (Fassone and Rahman, 2012) (Figure 7.2). Median survival length was longer in the SURF1 deficient group (median 5.4 years, 25th centile 3.0, 75th centile 10 years) than in the LRPPRC deficient group (median 1.8 years, 25th centile 1.0, 75th centile 4 years), and the nuclear-encoded complex I-deficient Leigh syndrome group (median 1.6 years, 25th centile 1.0, 75th centile 10 years) ($p<0.001$ for difference across groups; logrank test).

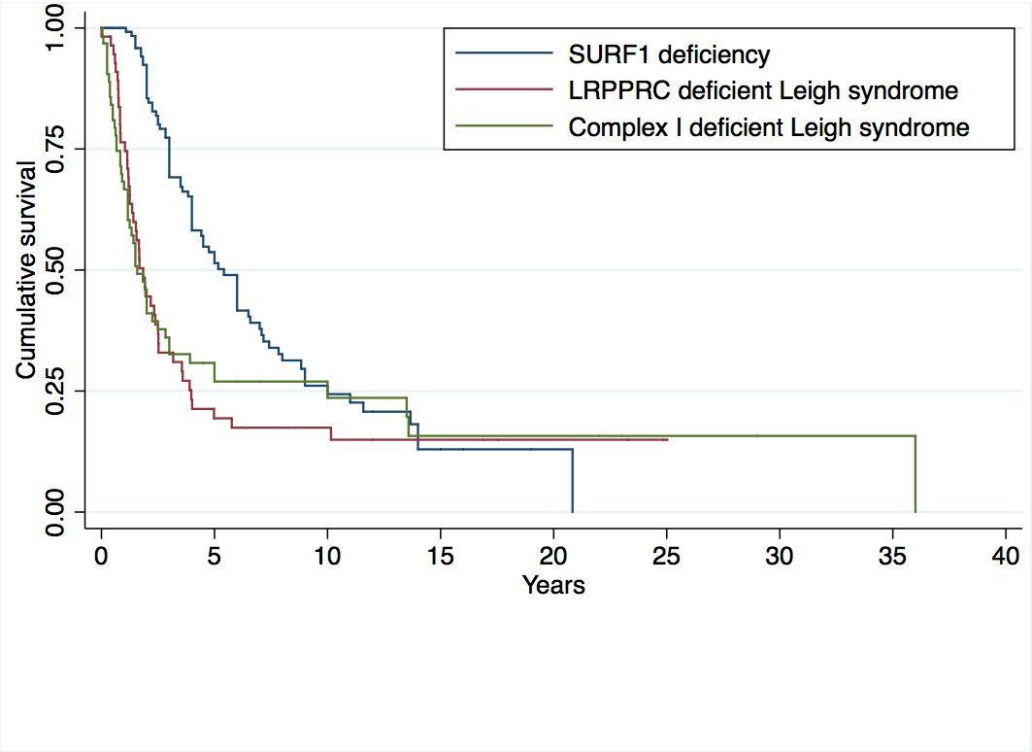


Figure 7:2: Kaplan-Meier survival curves of SURF1 deficient patients, LRPPRC-deficient patients and nuclear-encoded complex I deficient Leigh/Leigh-like syndrome patients.

Blue line: SURF1 deficient patients, n = 142 (patients from this chapter and literature), red line: LRPPRC deficient patients, n = 56, green line: complex I deficient Leigh/Leigh-like syndrome n = 63. Logrank test $p < 0.001$. (As published in (Wedatilake et al. 2013))

7.3.5 Peripheral nervous system

Sixteen patients were identified where nerve conduction studies were performed. Abnormal peripheral nerve conduction studies were found in 13/16 (81%). The majority (7/16) had a demyelinating neuropathy whereas an axonal neuropathy was found in 2/16 subjects. The type of neuropathy was not specified in 4/16 patients. The majority of the cases (9/13, 69%) demonstrated a mixed sensori-motor involvement whereas two subjects had a pure sensory neuropathy.

7.3.6 Neuroimaging

In the cohort MRI and computerised tomography scan (CT) data were available in 39 patients (MRI n=33 and CT scans n=6). The affected brain areas in SURF1 deficiency are shown in Table 7.3. The majority of the MRIs demonstrated radiological signs characteristic of Leigh syndrome (28/33, 85%), where there were symmetrical hyperintense lesions on T2-weighted images in the brainstem and/or basal ganglia (table 7.2). There were two patients with normal MRI scans (performed at 1 year of age) and subsequent imaging was not performed. Two patients demonstrated leukoencephalopathy (a disorder that predominantly or exclusively affects the white matter of the brain) (Schiffmann & Van Der Knaap 2009), while one patient demonstrated only cerebellar atrophy and involvement of the dentate nucleus. Of the patients with CT scans (n=6), three scans were normal whereas three patients had bilateral symmetrical hypodensities in the basal ganglia.

Table 7.2: Laboratory and magnetic resonance imaging (MRI) findings in SURF1 deficiency

Laboratory or neuroimaging finding	Number of patients	%
Blood		
Metabolic acidosis	21/33	64
Elevated lactate	31/38	81
Cerebrospinal fluid		
Elevated lactate	30/30	100
Muscle		
Reduced/absent COX histochemistry	23/33	70
Type I fibre predominance	8/33	24
Elevated muscle lipid content	16/33	48
Reduced muscle COX activity	25/26	96
Fibroblast		
Reduced fibroblast COX activity	25/25	100
Nerve conduction studies		
Peripheral neuropathy	13/16	81
MRI lesions		
Midbrain	12/33	36

Laboratory or neuroimaging finding	Number of patients	%
Pons	10/33	30
Medulla	15/33	45
Putamen	16/33	48
Globus pallidus	14/33	42
Caudate nucleus	12/33	36
Subthalamic nucleus	4/33	12
Periaqueductal grey matter	4/33	12
Olivary nuclei	3/33	9
Red nuclei	2/33	6
Cerebellar white/grey matter	6/33	18
Cerebellar atrophy	3/33	9
Dentate nucleus	13/33	39
Cerebellar peduncles	8/33	24
Leukoencephalopathy	2/33	6

7.3.7 *SURF1* gene mutations

SURF1 mutations were identified in 57 patients. The homozygous c.312_320del10insAT insertion/deletion was found in 16/57 (28%), whereas 15/57(26%) were compound heterozygotes, 12/57 (21%) had homozygous splice site mutations, 10/57 (18%) had homozygous deletions, 2/57 (4%) had homozygous insertions, 1/57 (2%) had homozygous missense mutations and 1/57(2%) had homozygous nonsense mutations. The single most common mutation was c.312_321del10insAT (p.Leu105X) (16 homozygous and 11 compound heterozygous). The next most common mutations were c.792_793delAG which occurred only in Bangladesh Bengali subjects (n=5, 3 pedigrees) and the splice mutation c.240+1G>T which was observed in 5 individuals, all of white European origin. The c.792_793delAG mutation has previously only been found in a patient of Japanese origin (Teraoka *et al*., 1999). Individuals of Pakistani origin had the novel c.799_800delCT mutation or the c.324-11T>G mutation. The c.516-2A>G mutation was found in two different Indian Gujarati pedigrees; previously this mutation has only been found in one subject of unspecified mixed white Asian parenthood (Lee *et al*., 2012). The c.324-11T>G mutation was observed in 4 individuals from 2 different Pakistani pedigrees and one individual of Indian Gujarati origin. All five patients from the Australian cohort had at least one c.312_321del10insAT mutation. There were no evident genotype phenotype correlations. The patients who survived longer than 10 years had differing *SURF1* genotypes: one was homozygous for the c.312_321del10insAT mutation, two were compound heterozygotes, two had homozygous insertions and one had a splice site mutation.

Novel pathogenic SURF1 mutations

Of the 19 *SURF1* mutations (Figure 7.3) identified in 57 patients five novel pathogenic mutations (c.468_469delTC, c.799_800delCT, c.575G>A (p.Arg192Gln), c.751+5G>A. and c.752-2A>G) were found.

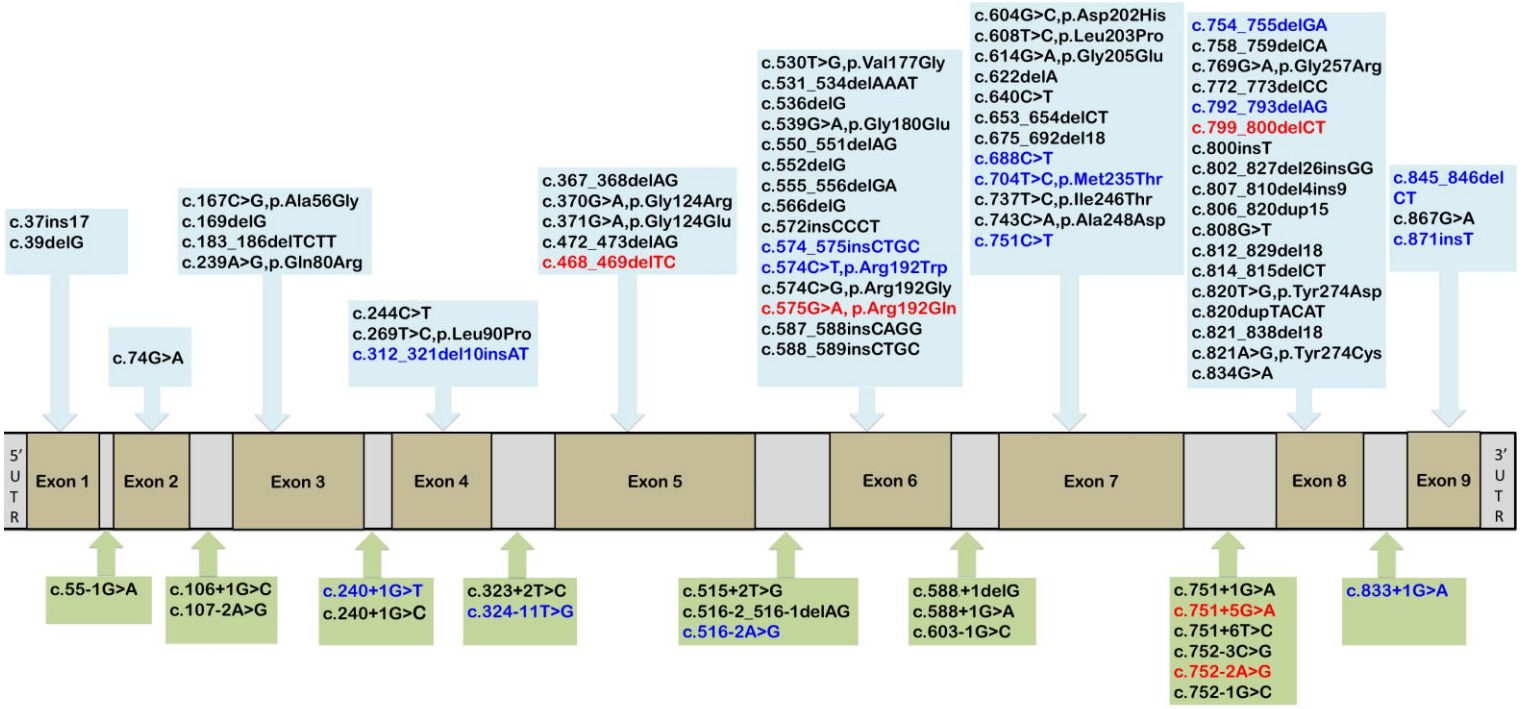


Figure 7:3: Mutations found in the human *SURF1* gene.

A schematic representation of the *SURF1* gene illustrating the 83 mutations associated with disease. Red: novel unreported mutations identified in this study, blue: other mutations which have been previously reported and found in this study, black: mutations previously reported in the literature. (As published in (Wedatilake et al. 2013))

7.3.8 Biochemical data

Metabolic acidosis was found in 21/33 (64%) patients (mean bicarbonate level 14.7, range 10-17 mmol/L). CSF lactate was elevated in all subjects who had this measured (mean 4.3, range 2.5-8.6 mmol/L, normal <2 mmol/L). Thirty one patients (31/38, 81%) demonstrated elevated plasma lactate (mean 4.4, range 2.3-7.3 mmol/L) (Table 7.2). Where fibroblast COX activity was measured, it was found to be low in all subjects (25/25, 100%) ranging from undetectable COX activity to 57% of the lower limit of the reference range.

7.3.9 Muscle biopsy

The histology and respiratory chain enzyme (RCE) findings of all patients where a muscle biopsy was performed (n=33) were analysed. Lipid content was increased in 15/33 (46%) and type 1 fibre predominance was found in 8 biopsies. Histochemical staining for COX demonstrated absent COX in 9 (27%) biopsies and reduced COX staining in 14 (42%) biopsies. Ten (30%) of the biopsies had normal COX histochemistry with uniformly staining fibres. RCE activities were determined in 28 muscle biopsies. There were two post-mortem biopsies which were excluded since post-mortem delay could lead to an artifactual decrease in RCE activities. There was an isolated decrease of COX activity in all except one biopsy where the COX activity was at the lower end of the reference range.

7.3.10 Additional cases in the literature

A literature review identified 129 SURF1-deficient patients where clinical information was available. Additional clinical features which were found in other published cases but not in the study cohort here included facial dysmorphism (5 patients), renal tubulopathy (4 cases), hepatic involvement (3 cases) (Poyau *et al*., 2000; Moslemi *et al*., 2003; Rossi *et al*., 2003; Darin *et al*., 2003; Tay *et al*.,2005; Yüksel *et al*., 2006, Lee *et al*., 2012) and demyelinating autosomal recessive Charcot Marie Tooth disease (2 cases) (Echaniz-Laguna et al. 2013).

7.4 Discussion

SURF1-deficient mitochondrial disease is the most common single cause of Leigh syndrome in the UK population. Frequently, it is difficult to characterise clinical phenotypes in mitochondrial disease because it is difficult to identify enough patients with a single gene mitochondrial disorder. In this study 57 SURF1-deficient patients were identified with wide-ranging phenotypic data on 44 patients. SURF1-deficient patients have a disease course which commences in late infancy typically with poor feeding, vomiting and faltering growth followed by episodic developmental regression, ophthalmoplegia, a movement disorder, and finally death due to central respiratory failure. This clinical phenotype of Leigh syndrome was however difficult to differentiate from other causes of Leigh syndrome, such as complex I deficiency or complex V deficiency.

7.4.1 Clinical features and survival

The natural history of Leigh syndrome frequently consists of stepwise neurological regression often precipitated by viral illnesses and other states of high-energy demand (Sofou et al. 2014). In this cohort six patients surviving longer than 10 years were identified. Interestingly they did not experience developmental regression, implying that if the precipitation factors leading to these regressive episodes are prevented or treated, it could be a way of improving outcome in SURF1 deficiency. However the prevention of metabolic decompensation is difficult, especially on occasions where an obvious precipitant cannot be identified.

Cardiomyopathy was a rare feature (one patient) in this cohort and this was similar to previously published cases where only 2/129 cases developed cardiomyopathy (Poyau et al. 2000; Tay et al. 2005). The low incidence of seizures is similar to previously published cases where only 7/129 (5%) cases were reported to have had seizures. Seizures occurred infrequently and were found primarily in the Australian patients, which may indicate that there were environmental triggers which precipitate seizures. However high environmental temperature is unlikely to be the precipitant since no seasonal variation in the occurrence of seizures was observed in the cases from Australia.

Sensorineural deafness was not found in this cohort and is comparable with the low occurrence in previously published cases where only two previous cases had sensorineural deafness (Poyau et al. 2000). In this cohort 10 (23%) cases were identified as developing optic atrophy whereas previously only three cases (2%) were noted to have optic atrophy (Péquignot *et al*., 2001; Tay *et al*., 2005). It is worth noting that visual and auditory signs could be underdiagnosed in many cases where formal testing was not done. Three patients in this study who underwent general anaesthesia experienced neurological decline. Therefore caution must be exercised when considering general anaesthesia for SURF1-deficient patients.

A demyelinating peripheral neuropathy has been found in previous cases in the literature (Santoro et al. 2000; Bruno et al. 2002). A more recent study found that SURF1 deficiency caused Charcot Marie Tooth Disease type 4 (Echaniz-Laguna et al. 2013). This study demonstrates that while in the majority of cases a demyelinating neuropathy was found, an axonal neuropathy can also be found in SURF1 deficiency.

This study found that long-term survival in SURF1 deficiency was better compared to two other groups with Leigh syndrome due to nuclear gene disorders (complex I deficient Leigh syndrome and LRPPRC deficient Leigh syndrome). While the reasons for a poorer outcome in complex I deficiency is unclear, it is conceivable that the acute metabolic crises experienced in LRPPRC deficient French Canadian Leigh syndrome, significantly contributes to mortality (Debray et al. 2011).

7.4.2 Genotypes

In addition to the 78 known pathogenic mutations found in the *SURF1* gene, this study found five novel mutations. Approximately 80% of the pathogenic variants found in the *SURF1* gene are truncating mutations due to aberrant splicing, frameshift deletions or nonsense mutations. This study demonstrates that certain mutations are found within particular ethnic groups. The c.311_312insATdel10 appears to be a founder mutation in white Europeans whereas the c.790_800delAG was found in three different Bangladesh Bengali pedigrees. The c.604G>C mutation has only been reported in individuals of Chinese origin. Previous reports (Piekutowska-Abramczuk et al. 2009) indicated that patients with missense *SURF1* mutations had a better prognosis. It has not been possible to confirm or refute this observation since only three subjects had missense mutations (1 homozygous, 2 compound heterozygous) in this cohort. However two of these patients harbouring missense mutations are currently alive at 12 years and 17 years. In this cohort no definitive genotype phenotype correlation was found. The seven patients with a less severe disease course had varied genotypes including homozygosity for the common c.312_320del10insAT mutation, suggesting that there are other factors which may influence survival.

7.4.3 Biochemistry and histochemistry

This study found a consistent biochemical phenotype in SURF1 deficiency including elevated CSF lactate, and decreased COX activity in fibroblasts and muscle. If a typical clinical presentation of Leigh syndrome is encountered with decreased fibroblast COX activity targeted sequencing of the *SURF1* gene should be undertaken. Since there are no distinctive clinical features to delineate SURF1 disease from other causes of Leigh syndrome it is prudent to perform a muscle biopsy especially if the patient is deteriorating rapidly.

Typically it would be expected that muscle histochemistry demonstrates uniformly and globally decreased COX staining. However this could be a subtle finding which could be missed since many centres use prolonged incubations to allow for maximal demarcation between the negative and positive fibres. Ragged red fibres were not noted in any patients in this cohort; they typically suggest a mitochondrial DNA mutation and have only been noted in two previous SURF1-deficient cases (Poyau et al. 2000; Tay et al. 2005).

7.4.4 Neuroimaging

Previous neuroimaging studies have found that the subthalamic nucleus was consistently involved on MRI in SURF1 deficiency (Savoiardo et al. 2002). Only four patients (12%) had lesions in the subthalamic nucleus in this study while 16/33 (48%) patients had putaminal lesions. It is not possible to say if these lesions are specific to SURF1 deficiency without comparing to other groups with Leigh syndrome. Two patients demonstrated leukoencephalopathic changes, an atypical finding (Rahman et al. 2001). Two patients had

normal MRI findings at one year of age; they did not have subsequent imaging and therefore it is unknown if abnormalities would have manifested later in the disease (Tay et al. 2005).

7.4.5 Conclusion

In conclusion, this natural history study demonstrates that SURF1 deficiency causes a homogenous phenotype of Leigh syndrome characterised by systemic COX deficiency. This study provides a complete clinical description of the disease to date, amalgamating the genotypes and phenotypes reported thus far. This study utilises a larger sample size than previous cohorts and therefore gives a better indication of prognosis and survival in this disease.

Chapter 8

8. General discussion and future work

8.1 Overview

This thesis focuses on exploring clinical phenotypes and genotypes in childhood-onset mitochondrial disease. This has been a study of mitochondrial disease in the era of next generation sequencing. At the time of commencement of this study the application of NGS technologies in the research setting was just beginning and made it an exciting and novel venture. It was an age of rapid development in the field of genetics and it was necessary to constantly keep abreast and explore new analyses. Towards the latter stage of the studies described here, many of the patients who remained undiagnosed were recruited for whole genome sequencing in the Genomics England 100,000 Genomes Project (100K) and therefore these cases were not investigated further using WES (The appendix describes which cases have been recruited to the 100K project).

Chapters 2-6 of this thesis focused on trying to uncover the genetic cause. In chapter 7 the focus shifted towards a natural history study. Having studied the causation, understanding the natural history of a mitochondrial disease seemed to be a natural direction for my studies to proceed. Although not explored in this thesis, the next clear step would be the study of therapies and clinical trials.

This work contributes to uncovering two novel disease genes *ATAD3A* and *MDH2*. This study also demonstrates novel clinical phenotypes in paediatric mitochondrial disease such as in AARS2 and VARS2 deficiency. A novel *TRNT1* mutation was found and a comprehensive phenotypic study is performed of this new disease. This study also identified a unique cohort of patients with paediatric mitochondrial cardiomyopathy due to either mtDNA or nuclear gene mutations. The first natural history study of SURF1 deficiency was performed as a part of this work.

Much of the discussion about the findings in each chapter has been included at the end of each chapter. The following section explores key findings, limitations and future work.

8.2 Identification of COX deficient cohort

In chapter 2, a substantial cohort with suspected mitochondrial disease and isolated muscle COX deficiency was recruited. Tables 2.1-2.3 exemplify the clinical heterogeneity encountered in these patients. Prior to the advent of advanced genetic testing, patients were often classified only using their RCE deficiency.

One of the limitations of this study was the use of a single measurement of muscle COX activity to classify the patient as COX deficient. Indeed given the invasive nature of a muscle biopsy it is not possible to have multiple samples to verify this finding. When classifying a patient as having anaemia, it is possible to have several blood haemoglobin estimations to verify the laboratory

result. One of the ways of circumventing this would have been to only recruit patients with COX deficiency both biochemically and on histochemical staining. However it must be noted that the latter method is operator dependent and different laboratories use different incubation times. Indeed using both recruitment criteria would have meant that the patient with the *MRPL44* mutations (S13) would not have been recruited thus decreasing the sensitivity of the recruitment criteria.

Another approach would have been to stipulate that there should be COX deficiency in more than one tissue/cells for example, COX deficiency in both fibroblasts and muscle. However fibroblast COX activity was not routinely performed on a diagnostic basis at the time, and therefore this was not considered a recruitment criterion. In addition OXPHOS defects found in muscle are often not expressed in fibroblast cells (since fibroblasts mainly use glycolysis rather than OXPHOS) (van den Heuvel et al. 2004) and some cases may be missed. Furthermore the best approach is considered to be to biopsy the affected tissue. In the cardiomyopathy case series this is exemplified where the patient with VARS2 deficiency did not have decreased muscle RCE's but had decreased complexes I and IV in cardiac tissue. However it is often not possible to perform invasive biopsies on affected tissue.

This work demonstrates that the clinical diagnosis of a mitochondrial disease is difficult even in the most skilled hands. Diagnostic criteria e.g. mitochondrial disease criteria utilised to classify probability of mitochondrial disease in chapter 2 (Morava et al. 2006), are not definitive; they often contain categories such as possible or probable mitochondrial disease. Elevated lactate in blood and CSF are often included in disease criteria; there can be false positives in patients with seizures where lactate can rise due to epileptiform activity and venous lactate levels can be spuriously elevated in difficult venepuncture on a struggling child.

8.3 Whole exome sequencing to identify the genetic basis of COX deficiency

Chapter 3 describes a study focusing on the use of next generation sequencing in COX deficient patients. In this study, one of the aims was to use WES in patients where primary genetic (rather than secondary) COX deficiency was likely. Therefore the inclusion criteria included several parameters. Some of the cases in the cohort (n=74) identified in chapter 2 met the inclusion criteria, but it was not possible to obtain DNA samples at the time. This was because some of the patients were recruited to the study by sending out study invitations and consent forms in the post and it was not possible to subsequently obtain the necessary biological samples.

At the time of recruitment, it was postulated that COX subunit and assembly factor mutations were likely causes, since these were the majority of the causes found in the published literature on human COX deficiency at the time. Additionally a targeted NGS study of complex I-deficient cohort demonstrated (Calvo et al. 2010) complex I assembly factors and subunit mutations in the majority of the cases. It was this evidence that led to the design of the targeted NGS panel in chapter 3. The main advantage of targeted NGS is that one can ensure adequate coverage of the genes of interest. While it was apparent that the targeted NGS method was able to identify the mutations in disease controls, it was not efficacious in identifying the cause of disease in a small sample of this cohort. Furthermore additional evidence through the WES study showed that the NGS panel would not have been successful in this cohort, given that none of the patients carried COX subunit or assembly factor mutations.

At the start of the study, WES was a novel technique and the main challenge was to use bioinformatics to filter variants appropriately. Thus in some cases two filtering strategies have been described, since the filtering methodology was continuously developed and revised during the course of the study. Indeed in the first batch of samples analysed by WES in 2011 (complex I deficient patients not described here) a disease control with a known mutation was included to interrogate the robustness of the bioinformatics pipeline. While many studies at the time often focused on investigating individual cases or pedigrees with suspected mitochondrial disease (Gotz et al. 2011) using NGS, the approach used here was to study a whole cohort with a single RCE deficiency.

In chapter 3 Sanger sequencing and Western blot analysis (in some cases) were performed for cases where only one candidate gene was thought to be plausible, or where a mutation linked to a previously published disease causing gene or a known pathogenic mutation, was found. In the case of the YARS-deficient patient the segregation studies and functional evaluation could not be completed owing to lack of parental DNA and cell senescence.

Perhaps the most interesting aspect of this thesis is the link of many ‘non-mitochondrial’ genes to patients with suspected mitochondrial disease. This has been noted previously by Pyle *et al*

and Coon *et al* where they observed mutations in non-mitochondrial genes (*NKX2-1, ORAI1, CAPN3, COLQ, EXOSC8*, and *ANO10)* in patients with suspected mitochondrial disease with RCE deficiencies (Pyle et al. 2015; Coon et al. 2016). There are at least two reasons for this. Firstly the clinical diagnosis of mitochondrial disease is not straightforward, with phenotypes overlapping with other disease groups such as glycosylation disorders (Haas et al. 2007). Secondly respiratory chain enzyme deficiencies do not definitively indicate primary mitochondrial disease, and can be secondary to another disease process as indicated below.

8.3.1 Secondary COX deficiency

While primary COX deficiency has been the focus of much research interest, secondary COX deficiency is a relatively under-researched area. Secondary COX deficiency has been found in valproate treatment of mice which induced a loss of activity of COX activity in liver (Ponchaut et al. 1992). Rats who developed cardiomyopathy after chronic doxorubicin administration were found to have low COX activity in cardiac tissue (Lebrecht et al. 2010). Wilson's disease, a disorder of copper metabolism also leads to a secondary decrease in COX activity in leucocytes because copper is an essential prosthetic group in the COX holoenzyme (Shokeir & Shreffler 1969). In addition secondary OXPHOS deficiency has been noted in sporadic epilepsy with the presence of COX negative neurones. In hippocampal sclerosis, a type of focal epilepsy, severely reduced complex I activity has been found in the CA3 hippocampal subfield in surviving neurones (Kunz et al. 2000).

The mechanisms of COX deficiency in the patients where non-mitochondrial genes were found is uncertain at present in this study, but it could be a secondary phenomenon due to muscle dysfunction.

8.4 Functional evaluation of candidate genes for COX deficient patients

In chapter 4 the most interesting or novel genetic findings were explored further, to verify pathogenicity. One of the limitations of this part of the study was the difficulty in finding definitive evidence of pathogenicity. There were several reasons for this, which are outlined below,

While COX deficiency was noted in patient muscle, none of the patient fibroblast cells displayed COX deficiency. This hindered further progress of verifying pathogenicity of the mutations as there was no functional readout in cells in culture. This also made it difficult to study the precise mechanism of COX deficiency in each of these cases (*IGHMBP2/SCYL*, *FBXO7, ATAD3A* and *PIGW* mutations). In addition other phenotypes in the cells were sought; for example impaired ubiquitination of MFN1 in the *FBXO7* mutations; impaired protein translation and nucleoid disruption in *ATAD3A* mutations and mtDNA depletion in *IGHMBP2/SCYL1* mutations. Again no compelling cellular phenotype was identified.

Another way of verifying pathogenicity is to accumulate a wealth of genetic evidence from different families with the same gene and similar clinical phenotypes. In this study no gene defect was shared between different families and therefore this avenue for verifying pathogenicity could not be pursued. However more recently websites such as Genematcher have become available https://genematcher.org to establish connections between researchers around the world sharing an interest in the same gene.

The identification of candidate genes, not linked directly to the mitochondrion directly meant that it was essential to form collaborations with other research groups internationally with expertise in these pathways, in order to make progress with these studies. This was an interesting albeit time consuming exercise and some of the studies are still on-going with these collaborative research groups. This collaborative approach opened up opportunities to study animal models of disease and to develop animal models with the same genetic variant as the patient. Animal models do not however always have the same clinical picture, as demonstrated by the Surf1 deficient mouse where a mild phenotype was found compared to severe necrotising encephalomyelopathy observed in humans (Dell'Agnello et al. 2007).

A patient with childhood onset pontocerebellar hypoplasia and neuronopathy harboured a truncating *FBXO7* mutation a gene which has previously been linked to a Parkinsonian pyramidal syndrome in adults. This is the reverse of the finding for *COQ2*, a gene initially linked with childhood onset mitochondrial disease but subsequently found to cause multisystem atrophy in adults (Diomedi-Camassei et al. 2007; M-SAR Collaboration 2013).

In addition to using the impact of the mutation at the protein level using Western blot, metabolic investigations were used to refute or support a candidate gene. For example, 3-methylglutaconic aciduria made it likely that *ATAD3A* was the genetic cause, Krebs cycle

intermediates made it likely that *MDH2* was the genetic cause, while *OGDH* was unlikely due to the absence of Krebs cycle metabolites. Use of such an approach is often encountered in clinical practice and indeed recent metabolomics evidence indicates that each mitochondrial syndrome may have its own metabolic thumbprint (Thompson Legault et al. 2015). Therefore it is likely that in the future metabolomics evidence may be used to complement genetic findings, and integration of large multi-omics data will be a bioinformatics challenge going forward.

8.5 The genetic heterogeneity of mitochondrial cardiomyopathy in children

In chapter 5 the causes of mitochondrial cardiomyopathy were explored. This cohort included patients with COX deficiency and other patients with complex I deficiency, combined complex I and IV deficiency and patients with normal RCEs.

In 2003 Holmgren published a cohort of paediatric patients with mitochondrial disease (n=101) of which 17 patients had cardiomyopathy and in 2007 Yaplito-Lee published 86 patients with mitochondrial disease with cardiomyopathy in 29 patients (Holmgren et al. 2003; Yaplito-Lee et al. 2007). Both these cohorts were published before the use of WES and therefore the genetic aetiology was incompletely defined. Since WES was introduced, Taylor *et al* performed WES in a cohort of patients with multiple respiratory chain enzyme deficiencies where cardiac involvement was noted in 19/53 cases. The causes in this cohort included *AARS2, MTO1, MGME, RMND1, MTFMT, ELAC2, GARS and PTCD1* (Taylor et al. 2014). A more recent study found that 11 patients had cardiomyopathy among 142 mitochondrial disease patients where WES was performed (Kohda et al. 2016). Legati *et al* used targeted NGS to investigate 125 patients with mitochondrial disease of which 10 cases had cardiomyopathy (Legati et al. 2016). These cohorts however did not include cases solved by mtDNA sequencing or by candidate gene sequencing. Thus the cohort in this study represents the largest genetically defined cohort of children with mitochondrial cardiomyopathy where the cardiomyopathy was due to a mtDNA or nuclear gene mutation.

One of the striking findings of this cohort was that all of the genetic causes found were localised to the mitochondrion, unlike in the COX deficient cohort. This is because some of the genetic approaches, such as a candidate gene approach, are inherently biased since by definition it means that the phenotype had previously been linked with a known mitochondrial disease. Also some cases had clinically apparent classical mitochondrial syndromes such as MELAS and Kearns Sayre syndrome which are known to be associated with mtDNA point mutations or deletions respectively. This leads to ascertainment bias. However it is interesting that in the 11 cardiomyopathy patients where WES was performed, the genetic cause was found in 10 cases (which included 2 sibling pairs) and they all had mutations in genes localised to the mitochondrion.

Whole exome sequencing allows an intrinsically unbiased approach where the main assumption is the mode of inheritance, i.e. autosomal recessive in the cases here. This could mean that the presence of cardiomyopathy in suspected mitochondrial disease is a ‘hard’ sign that it is a *bona fide* mitochondrial disease. As a part of the study design only cases which had followed the pathway (mtDNA sequencing and/or candidate gene sequencing and/or WES) have been included. Not all cases with suspected mitochondrial cardiomyopathy presenting to our institution have been included here. There were some cases who underwent mtDNA sequencing and/or candidate gene sequencing but did not have WES due to inability to contact

the patient to obtain consent for exome studies. It is not known if the inclusion of these cases would have led to a different genetic diagnostic rate. There were only two consanguineous families in this cohort, and the genetic diagnostic rate here is encouraging for patients from non-consanguineous backgrounds.

Chapter 5 focused on uncovering the genetic causes and mitochondrial pathways involved in the pathogenesis of cardiomyopathy. Eighteen different genetic causes were found and the majority of the pathogenic genes were involved in the mitochondrial translation pathway. This may be a reflection that there are an estimated 250-300 nuclear genes (~1/4 of the mitoproteome) involved in mitochondrial protein translation (Powell et al. 2015) therefore making it more likely to find one of these genes as the cause. The exact mechanism of the genetic causes underlying the cardiomyopathy was not explored in this chapter. The observation that patients carrying the same mutation (*MTO1*) or different mutations in the same gene (*c.12orf65*) did not always manifest cardiomyopathy implies that it is likely that the pathogenesis of mitochondrial cardiomyopathy is affected by factors other than the genetic change itself.

8.6 The functional implications of TRNT1 deficiency

Chapter 6 explores a relatively new disorder, TRNT1 deficiency which presents with an unusual multi-system phenotype. A novel mutation was found in the *TRNT1* gene and the phenotypic spectrum of this disorder is studied.

TRNT1 is a CCA adding enzyme which is bifunctional in both the cytoplasm and in the mitochondrion. One of the limitations of this study was that the cytoplasmic aspect of TRNT1 dysfunction was not explored. At the outset, before the *TRNT1* gene was identified as the cause of this unique syndrome, this disease was thought to be a mitochondrial disease due to its clinical resemblance to Pearson syndrome and YARS2 deficiency. The majority of the identified causes of genetic sideroblastic anaemia are due to dysfunction in mitochondrial pathways such as haem biosynthesis, iron-sulfur cluster biogenesis, and mitochondrial protein synthesis (Fleming 2011). Furthermore one of the cases described in chapter 6 had COX deficiency. Therefore after identifying *TRNT1* as the candidate gene, it was hypothesised that the disease mechanism involves impaired CCA addition to mitochondrial tRNAs. In addition in other diseases due to deficiencies of dually located enzymes such as PUS1, a primary mitochondrial disease presentation is encountered (Ghezzi et al. 2012; Bykhovskaya et al. 2004). This may indicate that the mitochondrial translation apparatus is more dependent on these modification enzymes than the more robust cytoplasmic system.

It is interesting to note that while CCA addition was impaired, the mitochondrial translation rate was normal in this study and in a previous study (Sasarman et al. 2015). Therefore the real functional significance of a subset of mitochondrial tRNAs lacking CCA-tail, still requires elucidation. In addition there is a need to investigate the impact on both cellular protein

translation systems given that cytoplasmic translation system disorders such as LARS deficiency (Casey et al. 2012) present with a multisystem phenotype (acute liver failure, anaemia, renal tubulopathy, developmental delay, seizures, failure to thrive and deterioration of liver function with illness) similar to TRNT1 deficiency.

8.7 The natural history of a monogenic mitochondrial disease

The natural history of a disease is defined as "The natural course of a disease from the time immediately prior to its inception, progressing through its presymptomatic phase and different clinical stages to the point where it has ended and the patient is either cured, chronically disabled or dead without external intervention" (de la Paz et al. 2010). Natural history studies are particularly important in rare diseases where it is often difficult to obtain an adequate sample size to conduct a double blinded placebo controlled randomised clinical trial. During drug development, the natural history data can be compared to the disease progression of patients where a new drug is administered, to observe if the natural progression is changed by the therapy (Augustine et al. 2013).

This is the first natural history study of SURF1 deficiency. Much of the work presented in this thesis focused on unravelling the genetic basis of COX deficiency. Therefore natural history of disease caused by mutations in *SURF1,* the first nuclear gene associated with COX deficiency, was an appropriate natural history study to undertake. By utilising a multi-centre approach to data collection it was possible to accumulate a sizeable cohort of patients in order to study the natural history. Furthermore at the time this study was conducted, natural history studies on nuclear genes causing mitochondrial disease were lacking. In order to maximise the number of potential outcome measures a wide array of clinical symptoms, metabolic and biochemical tests, neuroimaging and nerve conduction studies were studied.

This study has several limitations. While the study data were collected using a standardised questionnaire this is a retrospective study where clinical data were obtained by case note review and many of the patients are now deceased. It is however unlikely that a large prospective study will be conducted for a rare disease such as SURF1 deficiency. The MRI scans were performed at different hospitals and reviewed by different neuroradiologists making it difficult to assign specific disease patterns. Furthermore the neuroimaging findings are based on a single MRI or CT scan in most cases and therefore represent a snapshot of the neuroradiological changes but do not allow exploration of disease progression. Repeated neuroimaging in young children with mitochondrial disease usually involves general anaesthesia with the accompanying risks of decompensation, and therefore it is often not prudent to perform multiple scans.

8.8 Future work

This work represents the identification of patients with paediatric mitochondrial disease, clinical phenotyping, genetic investigation and efforts to prove pathogenicity of genetic findings. As a result of wide adoption of NGS technologies a vast number of new diseases are being recognised. For example it was estimated that in 2013, a new glycosylation disorder was being reported every 17 days (Ng & Freeze 2014).

As described at the beginning of this chapter the patients where a cause has not been found are currently being recruited for WGS in the 100,000 Genomes Study. In this national study patients are recruited in trios (parents and patient) which will presumably allow for more potent gene identification for unsolved cases in this study. In the future it is likely that the diagnostic pathway for inherited diseases will be completely reversed with genetic testing being at the forefront of the investigations, ahead of invasive tests such as muscle biopsies. The use of NGS has been particularly efficacious in mitochondrial diseases due to the well-defined mitoproteome, which makes it easier to prioritise potentially disease causing variants. The work in this thesis demonstrates that functional studies still retain their place especially where there is a lack of compelling genetic evidence; often it is not possible to find unrelated pedigrees with the same clinical phenotype and the same genetic defect due to the rarity of individual mitochondrial disorders.

Future work arising from this thesis includes proving pathogenicity of variants found in genes such as *IGHMBP2/SCYL1* and *FBXO7*. To this end a collaboration is underway with a group who have developed a *fbxo7* knock out mouse. The precise pathogenic gene i.e. whether it is *IGHMBP2* or *SCYL1* or both has been difficult to establish conclusively. As more and more cases with disease causing mutations in both these genes are published, the precise cause is likely to become clearer. There is also an on-going collaboration to study tissues from a Ighmbp2 deficient mouse.

One of the conundrums which remains unresolved is tissue specificity found in mitochondrial diseases. For mtDNA mutations, high loads of mutated mtDNA are often found in clinically affected patient tissue. This is especially puzzling in the case of nuclear gene mutations where consistent organ specific associations have been found in the amino acyl synthetase deficiencies (Konovalova & Tyynismaa 2013). While this thesis does not specifically address this question, an attempt is made to explore this conundrum by studying the genetic aetiology of a single organ manifestation in mitochondrial disease (mitochondrial cardiomyopathy).

In the genomic era, clinical phenotyping continues to retain its importance in the investigation of such patients. This work demonstrates that while clinical phenotyping is used as a tool to hypothesise the type of disease (e.g. a COX deficient mitochondrial disease) prior to NGS, when there are unexpected findings it is imperative to revisit the phenotype and evaluate the

exome findings to hand. As more and more phenotypic data becomes available for different genotypes, large structured databases are needed to curate phenotypic information and disease causing mutations. This would enable clinicians and scientists investigating these patients to ascribe genetic causes more easily and confidently. In addition new biomarkers other than RCEs for mitochondrial disease would be useful tools and would help confirm or refute genetic findings on NGS. Current biomarkers for mitochondrial disease include FGF21 and GDF-15 (Suomalainen et al. 2011; Montero et al. 2016). An exciting prospect is the emerging idea that each mitochondrial disease may have its own metabolomic thumbprint (Thompson Legault et al. 2015).

Uncovering the genetic cause is especially important for families who wish to pursue prenatal or pre-implantation genetic diagnosis. Nevertheless, having studied the aetiology of disease, a clear future direction would be to study therapies. While mitochondrial diseases (with the exception of a few vitamin and cofactor responsive disorders) are largely considered untreatable, more and more options for therapy are being explored. Given the heterogeneity of this condition, it is difficult to envisage 'a one size fits all' approach. Instead tailor made therapies, i.e. a personalised medicine approach, are likely to be most fruitful.

Where there is genuine OXPHOS dysfunction (rather than, for example, a disease due to a fission defect) hypoxia therapy has shown promise in preclinical studies (Jain et al. 2016). In addition the OXPHOS block can be bypassed by using substitute OXPHOS enzymes such as using the complex I yeast counterpart, NADH dehydrogenase/CoQ reductase (Ndi1) in complex I deficiency (El-Khoury et al. 2014). Current pharmacological therapies which have shown some promise, such as bezafibrate, resveratrol and 5-Aminoimidazole-4-carboxamide ribonucleotide (AICAR) act by increasing mitochondrial biogenesis (Wenz et al. 2008; Baur et al. 2006; Viscomi et al. 2011). It has also been demonstrated that the NAD^+ pool can be increased by dietary supplementation with its precursor nicotinamide riboside leading to improvement in the OXPHOS defect and exercise intolerance in *Sco2* knockout/knockin mouse models (Cerutti et al. 2014).

Other emerging therapies include gene therapy approaches such as adenovirus mediated gene transfer which have been used in mouse models of MNGIE and ethylmalonic encephalopathy (Torres-Torronteras et al. 2011). 'Genetic editing' with mitochondrially targeted transcription activator-like effector (TALE) nuclease (TALENs) and Zinc finger nucleases to selectively remove pathogenic mtDNA deletions or point mutations (Minczuk et al. 2008; Gammage et al. 2014; Bacman et al. 2013). Mitochondrial donation for patients with mtDNA mutations has garnered much attention from the media and the legislation required for this procedure has been the subject of a parliamentary debate in the UK (Gorman et al. 2015). This approach involves transferring the nucleus from fertilised embryos or oocytes (pro-nuclear transfer or metaphase II spindle transfer) to donor oocyte cytoplasm and has been successfully used in

rhesus macaque monkeys and human oocytes (Tachibana et al. 2009; Tachibana et al. 2012; Craven et al. 2010). In 2016 Yamada *et al* showed that although the low levels of mutant mtDNA present in human oocytes during nuclear transfer often disappear, sometimes there can be reversion to the original mutant genotype (Yamada et al. 2016). A recent study showed that transplantation of the pronucleus shortly after meiosis is completed (rather than just before the first mitotic division) results in mtDNA carryover <2% in the majority (79%) of blastocysts (Hyslop et al. 2016).

9. References

Abrahams, J.P. et al., 1994. Structure at 2.8 A resolution of F1-ATPase from bovine heart mitochondria. *Nature*, 370(6491), pp.621–8.

Alam, T.I. et al., 2003. Human mitochondrial DNA is packaged with TFAM. *Nucleic acids research*, 31(6), pp.1640–5.

Alfadhel, M. et al., 2011. Infantile cardioencephalopathy due to a COX15 gene defect: report and review. *American Journal of Medical Genetics. Part A*, 155A(4), pp.840–4.

Allen, A.S. et al., 2013. De novo mutations in epileptic encephalopathies. *Nature*, 501(7466), pp.217–21.

Altmann, R., 1890. Die Elementarorganismen Und Ihre Beziehungen Zu Den Zellen. *Leipzig*.

Anderson, S. et al., 1981. Sequence and organization of the human mitochondrial genome. *Nature*, 290(5806), pp.457–65.

Antonicka, H., Leary, S.C., et al., 2003. Mutations in COX10 result in a defect in mitochondrial heme A biosynthesis and account for multiple, early-onset clinical phenotypes associated with isolated COX deficiency. *Human Molecular Genetics*, 12(20), pp.2693–702.

Antonicka, H., Mattman, A., et al., 2003. Mutations in COX15 produce a defect in the mitochondrial heme biosynthetic pathway, causing early-onset fatal hypertrophic cardiomyopathy. *American Journal of Human Genetics*, 72(1), pp.101–14.

Ashford, T.P. & Porter, K.R., 1962. Cytoplasmic components in hepatic cell lysosomes. *The Journal of Cell Biology*, 12, pp.198–202.

Ashrafian, H. et al., 2010. A mutation in the mitochondrial fission gene Dnm1l leads to cardiomyopathy. *PLoS Genetics*, 6(6), p.e1001000.

Augustine, E.F., Adams, H.R. & Mink, J.W., 2013. Clinical trials in rare disease: challenges and opportunities. *Journal of Child Neurology*, 28(9), pp.1142–50.

Bacman, S.R. et al., 2013. Specific elimination of mutant mitochondrial genomes in patient-derived cells by mitoTALENs. *Nature Medicine*, 19(9), pp.1111–3.

Baertling, F. et al., 2015. Mutations in COA6 cause cytochrome c oxidase deficiency and neonatal hypertrophic cardiomyopathy. *Human Mutation*, 36(1), pp.34–8.

Baloh, R.H. et al., 2007. Altered axonal mitochondrial transport in the pathogenesis of Charcot-Marie-Tooth disease from mitofusin 2 mutations. *The Journal of Neuroscience*, 27, pp.422–430.

Balsa, E. et al., 2012. NDUFA4 is a subunit of complex IV of the mammalian electron transport chain. *Cell Metabolism*, 16(3), pp.378–386.

Ban-Ishihara, R. et al., 2015. COX assembly factor ccdc56 regulates mitochondrial morphology by affecting mitochondrial recruitment of Drp1. *FEBS Letters*, 589(20), pp.3126–3132.

Baradaran, R. et al., 2013. Crystal structure of the entire respiratory complex I. *Nature*, 494(7438), pp.443–448.

Barrell, B.G., Bankier, a T. & Drouin, J., 1979. A different genetic code in human mitochondria. *Nature*, 282(5735), pp.189–194.

Baur, J. a et al., 2006. Resveratrol improves health and survival of mice on a high-calorie diet.

Nature, 444(7117), pp.337–42.

Benda, C., 1898. *Ueber die Spermatogenese der Vertebraten und höherer Evertebraten, II*,

Bennett, S., 2004. Solexa Ltd. *Pharmacogenomics*, 5(4), pp.433–438.

Bereiter-Hahn, J., 1990. Behavior of Mitochondria in the Living Cell. *International Review of Cytology*, 122, pp.1–63.

Bogenhagen, D. & Clayton, D.A., 1977. Mouse L cell mitochondrial DNA molecules are selected randomly for replication throughout the cell cycle. *Cell*, 11(4), pp.719–727.

Bosley, T.M. et al., 2008. Sporadic bilateral optic neuropathy in children: The role of mitochondrial abnormalities. *Investigative Ophthalmology and Visual Science*, 49(12), pp.5250–5256.

Bourens, M. et al., 2014. Human COX20 cooperates with SCO1 and SCO2 to mature COX2 and promote the assembly of cytochrome c oxidase. *Human Molecular Genetics*, 23(11), pp.2901–2913.

Bourgeron, T. et al., 1995. Mutation of a nuclear succinate dehydrogenase gene results in mitochondrial respiratory chain deficiency. *Nature Genetics*, 11(2), pp.144–149.

Boyer, P.D., 1997. The ATP synthase - A splendid molecular machine. *Annual Review of Biochemistry*, 66, pp.717–749.

Briones, P. et al., 2001. Congenital disorders of glycosylation (CDG) may be underdiagnosed when mimicking mitochondrial disease. *European Journal of Paediatric Neurology*, 5(3), pp.127–31.

Broomfield, A. et al., 2015. Paediatric single mitochondrial DNA deletion disorders: an overlapping spectrum of disease. *Journal of Inherited Metabolic Disease*, 38(3), pp.445–57.

Brown, K.R. et al., 2002. Identification of novel hemes generated by heme A synthase: evidence for two successive monooxygenase reactions. *Biochemistry*, 41, pp.10906–10913.

Brown, W.M., George, M. & Wilson, A.C., 1979. Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 76(4), pp.1967–71.

Bruno, C. et al., 2002. A novel mutation in the SURF1 gene in a child with Leigh disease, peripheral neuropathy, and cytochrome-c oxidase deficiency. *J Child Neurol*, 17(3), pp.233–236.

Bruno, C. et al., 1999. A stop-codon mutation in the human mtDNA cytochrome c oxidase I gene disrupts the functional structure of complex IV. *American Journal of Human Genetics*, 65(3), pp.611–20.

Bugiani, M. et al., 2005. Novel mutations in COX15 in a long surviving Leigh syndrome patient with cytochrome c oxidase deficiency. *Journal of Medical Genetics*, 42(5), p.e28.

Burchell, V.S. et al., 2013. The Parkinson's disease-linked proteins Fbxo7 and Parkin interact to mediate mitophagy. *Nature Neuroscience*, 16(9), pp.1257–65.

Burman, J.L. et al., 2008. Scyl1, mutated in a recessive form of spinocerebellar neurodegeneration, regulates COPI-mediated retrograde traffic. *The Journal of Biological*

Chemistry, 283(33), pp.22774–86.

Burman, J.L., Hamlin, J.N.R. & McPherson, P.S., 2010. Scyl1 regulates Golgi morphology. *PloS one*, 5(3), p.e9537.

Bykhovskaya, Y. et al., 2004. Missense mutation in pseudouridine synthase 1 (PUS1) causes mitochondrial myopathy and sideroblastic anemia (MLASA). *American Journal of Human Genetics*, 74(6), pp.1303–8.

Cahill, T.J., Ashrafian, H. & Watkins, H., 2013. Genetic cardiomyopathies causing heart failure. *Circulation research*, 113(6), pp.660–75.

Calvo, S.E. et al., 2010. High-throughput, pooled sequencing identifies mutations in NUBPL and FOXRED1 in human complex I deficiency. *Nature Genetics*, 42(10), pp.851–8.

Calvo, S.E. et al., 2012. Molecular diagnosis of infantile mitochondrial disease with targeted next-generation sequencing. *Sci Transl Med*, 4(118), p.118ra10.

Calvo, S.E., Clauser, K.R. & Mootha, V.K., 2015. MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. *Nucleic Acids Research*, pp.1–7.

Calvo, S.E. & Mootha, V.K., 2010. The mitochondrial proteome and human disease. *Annual Review of Genomics and Human Genetics and Human Genetics*, 11, pp.25–44.

Carling, P.J., Cree, L.M. & Chinnery, P.F., 2011. The implications of mitochondrial DNA copy number regulation during embryogenesis. *Mitochondrion*, 11(5), pp.686–692.

Carroll, C.J. et al., 2013. Whole-exome sequencing identifies a mutation in the mitochondrial ribosome protein MRPL44 to underlie mitochondrial infantile cardiomyopathy. *Journal of Medical Genetics*, 50(3), pp.151–9.

Carroll, C.J., Brilhante, V. & Suomalainen, A., 2014. Next-generation sequencing for mitochondrial disorders. *British Journal of Pharmacology*, 171(8), pp.1837–1853.

Casey, J.P. et al., 2012. Identification of a mutation in LARS as a novel cause of infantile hepatopathy. *Molecular Genetics and Metabolism*, 106(3), pp.351–8.

Cerutti, R. et al., 2014. NAD(+)-dependent activation of Sirt1 corrects the phenotype in a mouse model of mitochondrial disease. *Cell Metabolism*, 19(6), pp.1042–9.

Chakraborty, P.K. et al., 2014. Mutations in TRNT1 cause congenital sideroblastic anemia with immunodeficiency, fevers, and developmental delay (SIFD). *Blood*, 124(18), pp.2867–71.

Chang, D.D., Hauswirth, W.W. & Clayton, D.A., 1985. Replication priming and transcription initiate from precisely the same site in mouse mitochondrial DNA. *The EMBO journal*, 4(6), pp.1559–67.

Check Hayden, E., 2014. Is the $1,000 genome for real? *Nature*.

Chinnery, P.F., 2010. Mitochondrial Disorders Overview [GeneReviews™. 1993] - PubMed - NCBI. , p.http://www.ncbi.nlm.nih.gov/books/NBK1224/. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20301403.

Chiyonobu, T. et al., 2014. Glycosylphosphatidylinositol (GPI) anchor deficiency caused by mutations in PIGW is associated with West syndrome and hyperphosphatasia with mental retardation syndrome. *Journal of Medical Genetics*, 51(3), pp.203–7.

Christian, B.E. & Spremulli, L.L., 2012. Mechanism of protein biosynthesis in mammalian

mitochondria. *Biochim Biophys Acta*, 1819(9-10), pp.1035–1054.

Christianson, T.W. & Clayton, D. a, 1986. In vitro transcription of human mitochondrial DNA: accurate termination requires a region of DNA sequence that can function bidirectionally. *Proceedings of the National Academy of Sciences of the United States of America*, 83(17), pp.6277–6281.

Cipolat, S. et al., 2004. OPA1 requires mitofusin 1 to promote mitochondrial fusion. *Proceedings of the National Academy of Sciences*, 101(45), pp.15927–15932.

Clark, K.M. et al., 1999. An mtDNA mutation in the initiation codon of the cytochrome C oxidase subunit II gene results in lower levels of the protein and a mitochondrial encephalomyopathy. *American Journal of Human Genetics*, 64, pp.1330–1339.

Coenen, M.J.H. et al., 2004. Cytochrome c oxidase biogenesis in a patient with a mutation in COX10 gene. *Annals of Neurology*, 56(4), pp.560–4.

Collins, T.J. et al., 2002. Mitochondria are morphologically and functionally heterogeneous within cells. *EMBO Journal*, 21(7), pp.1616–1627.

Comi, G.P. et al., 1998. Cytochrome c oxidase subunit I microdeletion in a patient with motor neuron disease. *Ann.Neurol.*, 43, pp.110–116.

Coon, E.A. et al., 2016. Expanding Phenotypic Spectrum of NKX2-1-Related Disorders-Mitochondrial and Immunologic Dysfunction. *JAMA neurology*, 73(2), pp.237–8.

Coore, H.G. et al., 1971. Regulation of adipose tissue pyruvate dehydrogenase by insulin and other hormones. *The Biochemical journal*, 125(1), pp.115–127.

Copeland, W.C., 2012. Defects in mitochondrial DNA replication and human disease. *Critical Reviews in Biochemistry and Molecular Biology*, 47(1), pp.64–74.

Cottenie, E. et al., 2014. Truncating and missense mutations in IGHMBP2 cause Charcot-Marie Tooth disease type 2. *American Journal of Human Genetics*, 95(5), pp.590–601.

Cox, G.A., Mahaffey, C.L. & Frankel, W.N., 1998. Identification of the mouse neuromuscular degeneration gene and mapping of a second site suppressor allele. *Neuron*, 21(6), pp.1327–1337.

Craven, L. et al., 2010. Pronuclear transfer in human embryos to prevent transmission of mitochondrial DNA disease. *Nature*, 465(7294), pp.82–85.

Cribbs, J.T. & Strack, S., 2009. Functional characterization of phosphorylation sites in dynamin-related protein 1. *Methods in enzymology*, 457, pp.231–53.

Crick, F.H., 1966. Codon--anticodon pairing: the wobble hypothesis. *Journal of Molecular Biology*, 19(2), pp.548–555.

Da Cruz, S. et al., 2003. Proteomic analysis of the mouse liver mitochondrial inner membrane. *The Journal of Biological Chemistry*, 278(42), pp.41566–41571.

Daga, a et al., 1993. Molecular characterization of the transcription termination factor from human mitochondria. *The Journal of biological chemistry*, 268(11), pp.8123–8130.

Dallabona, C. et al., 2014. Novel (ovario) leukodystrophy related to AARS2 mutations. *Neurology*, 82(23), pp.2063–71.

Debray, F.G. et al., 2011. LRPPRC mutations cause a phenotypically distinct form of Leigh

syndrome with cytochrome c oxidase deficiency. *Journal of Medical Genetics*, 48(3), pp.183–189.

Dell'Agnello, C. et al., 2007. Increased longevity and refractoriness to Ca2+-dependent neurodegeneration in Surf1 knockout mice. *Human Molecular Genetics*, 16(4), pp.431–444.

DeLuca, A.P. et al., 2016. Hypomorphic mutations in TRNT1 cause retinitis pigmentosa with erythrocytic microcytosis. *Human Molecular Genetics*, 25(1), pp.44–56.

Dianov, G.L. et al., 2001. Base excision repair in nuclear and mitochondrial DNA. *Prog Nucleic Acid Res Mol Biol*, 68, pp.285–297.

Diodato, D. et al., 2015. A novel AIFM1 mutation expands the phenotype to an infantile motor neuron disease. *European Journal of Human Genetics*, 24(3), pp.463–466.

Diomedi-Camassei, F. et al., 2007. COQ2 nephropathy: a newly described inherited mitochondriopathy with primary renal involvement. *Journal of the American Society of Nephrology*, 18(10), pp.2773–80.

Doersen, C.J. et al., 1985. Characterization of an RNase P activity from HeLa cell mitochondria. Comparison with the cytosol RNase P activity. *Journal of Biological Chemistry*, 260(10), pp.5942–5949.

Dudkina, N. V et al., 2010. Structure and function of mitochondrial supercomplexes. *Biochimica et biophysica acta*, 1797(6-7), pp.664–70.

Echaniz-Laguna, A. et al., 2013. SURF1 deficiency causes demyelinating Charcot-Marie-Tooth disease. *Neurology*, 22;81(17), pp.1523–1530.

El-Khoury, R. et al., 2014. Engineering the alternative oxidase gene to better understand and counteract mitochondrial defects: State of the art and perspectives. *British Journal of Pharmacology*, 171(8), pp.2243–2249.

Elo, J.M. et al., 2012. Mitochondrial phenylalanyl-tRNA synthetase mutations underlie fatal infantile Alpers encephalopathy. *Human Molecular Genetics*, 21(20), pp.4521–4529.

Fassone, E. et al., 2011. Mutations in the mitochondrial complex I assembly factor NDUFAF1 cause fatal infantile hypertrophic cardiomyopathy. *Journal of Medical Genetics*, 48(10), pp.691–7.

Fisher, R.P. & Clayton, D.A., 1985. A transcription factor required for promoter recognition by human mitochondrial RNA polymerase. Accurate initiation at the heavy- and light-strand promoters dissected and reconstituted in vitro. *Journal of Biological Chemistry*, 260(20), pp.11330–11338.

Fleming, M.D., 2011. Congenital sideroblastic anemias: iron and heme lost in mitochondrial translation. *Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. Education Program*, 2011(1), pp.525–31.

Follmann, K. et al., 1998. Cytochrome c oxidase from eucaryotes but not from procaryotes is allosterically inhibited by ATP. *Biochemistry and Molecular Biology International*, 45(5), pp.1047–55.

Galati, D. et al., 2009. Role of nuclear-encoded subunit Vb in the assembly and stability of

cytochrome c oxidase complex: implications in mitochondrial dysfunction and ROS production. *The Biochemical journal*, 420, pp.439–449.

Gammage, P.A. et al., 2014. Mitochondrially targeted ZFNs for selective degradation of pathogenic mitochondrial genomes bearing large-scale deletions or point mutations. *EMBO Molecular Medicine*, 6(4), pp.458–466.

Gattermann, N. et al., 1997. Heteroplasmic point mutations of mitochondrial DNA affecting subunit I of cytochrome c oxidase in two patients with acquired idiopathic sideroblastic anemia. *Blood*, 90, pp.4961–4972.

Geisler, S. et al., 2010. PINK1/Parkin-mediated mitophagy is dependent on VDAC1 and p62/SQSTM1. *Nature Cell Biology*, 12(2), pp.119–131.

Ghezzi, D. et al., 2008. FASTKD2 nonsense mutation in an infantile mitochondrial encephalomyopathy associated with cytochrome c oxidase deficiency. *American Journal of Human Genetics*, 83(3), pp.415–23.

Ghezzi, D. et al., 2012. Mutations of the mitochondrial-tRNA modifier MTO1 cause hypertrophic cardiomyopathy and lactic acidosis. *American Journal of Human Genetics*, 90(6), pp.1079–87.

Ghosh, A. et al., 2014. Copper supplementation restores cytochrome c oxidase assembly defect in a mitochondrial disease model of COA6 deficiency. *Human Molecular Genetics*, 23(13), pp.3596–3606.

Giles, R.E. et al., 1980. Maternal inheritance of human mitochondrial DNA. *National Academy of Sciences*, 77(11), pp.6715–6719.

Gilquin, B. et al., 2010. The AAA+ ATPase ATAD3A controls mitochondrial dynamics at the interface of the inner and outer membranes. *Molecular and Cellular Biology*, 30(8), pp.1984–1996.

Glamuzina, E. et al., 2012. Further delineation of pontocerebellar hypoplasia type 6 due to mutations in the gene encoding mitochondrial arginyl-tRNA synthetase, RARS2. *Journal of Inherited Metabolic Disease*, 35(3), pp.459–467.

Glerum, D.M., Shtanko, A. & Tzagoloff, A., 1996. Characterization of COX17, a yeast gene involved in copper metabolism and assembly of cytochrome oxidase. *The Journal of Biological Chemistry*, 271, pp.14504–14509.

Gorman, G.S. et al., 2015. Mitochondrial Donation — How Many Women Could Benefit? *New England Journal of Medicine*, 372(9), p.150128140051005.

Gotz, A. et al., 2011. Exome sequencing identifies mitochondrial alanyl-tRNA synthetase mutations in infantile mitochondrial cardiomyopathy. *American Journal of Human Genetics*, 88(5), pp.635–642.

Grivell, L.A., 1983. Mitochondrial DNA. *Scientific American*, 248(3), pp.78–89.

Grohmann, K. et al., 2003. Infantile Spinal Muscular Atrophy with Respiratory Distress Type 1 (SMARD1). *Annals of Neurology*, 54(6), pp.719–724.

Grohmann, K. et al., 2001. Mutations in the gene encoding immunoglobulin mu-binding protein 2 cause spinal muscular atrophy with respiratory distress type 1. *Nature Genetics*, 29(1),

pp.75–77.

Guenther, U.P. et al., 2009. IGHMBP2 is a ribosome-associated helicase inactive in the neuromuscular disorder distal SMA type 1 (DSMA1). *Human Molecular Genetics*, 18(7), pp.1288–1300.

Guffon, N. et al., 1993. 2-Ketoglutarate dehydrogenase deficiency, a rare cause of primary hyperlactataemia: report of a new case. *Journal of Inherited Metabolic Disease*, 16(5), pp.821–30.

Gurgel-Giannetti, J. et al., 2013. Mitochondrial cardioencephalomyopathy due to a novel SCO2 mutation in a Brazilian patient: case report and literature review. *JAMA Neurology*, 70(2), pp.258–61.

Haack, T.B. et al., 2012. Molecular diagnosis in mitochondrial complex I deficiency using exome sequencing. *Journal of Medical Genetics*, 49(4), pp.277–83.

Haas, R.H. et al., 2007. Mitochondrial disease: a practical approach for primary care physicians. *Pediatrics*, 120(6), pp.1326–1333.

Hallmann, K. et al., 2016. Loss of the smallest subunit of cytochrome c oxidase, COX8A, causes Leigh-like syndrome and epilepsy. *Brain*, 139(Pt 2), pp.338–45.

Hanna, M.G. et al., 1998. Cytochrome c oxidase deficiency associated with the first stop-codon point mutation in human mtDNA. *American Journal of Human Genetics*, 63, pp.29–36.

Hannappel, A., Bundschuh, F.A. & Ludwig, B., 2012. Role of Surf1 in heme recruitment for bacterial COX biogenesis. *Biochimica et biophysica acta*, 1817(6), pp.928–937.

Hauswirth, W.W. & Laipis, P.J., 1982. Mitochondrial DNA polymorphism in a maternal lineage of Holstein cows. *Proceedings of the National Academy of Sciences of the United States of America*, 79(15), pp.4686–90.

He, J. et al., 2007. The AAA+ protein ATAD3 has displacement loop binding properties and is involved in mitochondrial nucleoid organization. *The Journal of Cell Biology*, 176(2), pp.141–6.

van den Heuvel, L.P., Smeitink, J.A. & Rodenburg, R.J.T., 2004. Biochemical examination of fibroblasts in the diagnosis and research of oxidative phosphorylation (OXPHOS) defects. *Mitochondrion*, 4(5-6), pp.395–401.

Holley, R.W., Madison, J.T. & Zamir, A., 1964. A new method for sequence determination of large oligonucleotides. *Biochemical and Biophysical Research Communications*, 17(4), pp.389–394.

Holmgren, D. et al., 2003. Cardiomyopathy in children with mitochondrial disease; clinical course and cardiological findings. *European Heart Journal*, 24(3), pp.280–8.

Holt, I.J. et al., 1990. A new mitochondrial disease associated with mitochondrial DNA heteroplasmy. *American Journal of Human Genetics*, 46(3), pp.428–33.

Holt, I.J. et al., 2007. Mammalian mitochondrial nucleoids: Organizing an independently minded genome. *Mitochondrion*, 7(5), pp.311–321.

Holt, I.J., Harding, A.E. & Morgan-Hughes, J.A., 1988. Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. *Nature*, 331(6158), pp.717–9.

Holt, I.J., Lorimer, H.E. & Jacobs, H.T., 2000. Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA. *Cell*, 100(5), pp.515–24.

Horváth, R. et al., 2002. Childhood onset mitochondrial myopathy and lactic acidosis caused by a stop mutation in the mitochondrial cytochrome c oxidase III gene. *Journal of Medical Genetics*, 39(11), pp.812–6.

Horváth, R. et al., 2005. Mutations in mtDNA-encoded cytochrome c oxidase subunit genes causing isolated myopathy or severe encephalomyopathy. *Neuromuscular Disorders*, 15, pp.851–857.

Hou, Y.-M., 2010. CCA addition to tRNA: Implications for tRNA quality control. *IUBMB Life*, 62(4), p.NA–NA.

Hubstenberger, A. et al., 2008. ATAD 3A and ATAD 3B are distal 1p-located genes differentially expressed in human glioma cell lines and present in vitro anti-oncogenic and chemoresistant properties. *Experimental Cell Research*, 314(15), pp.2870–2883.

Huigsloot, M. et al., 2011. A mutation in C2orf64 causes impaired cytochrome c oxidase assembly and mitochondrial cardiomyopathy. *American Journal of Human Genetics*, 88(4), pp.488–93.

Huttemann, M., Kadenbach, B. & Grossman, L.I., 2001. Mammalian subunit IV isoforms of cytochrome c oxidase. *Gene*, 267(1), pp.111–123.

Hyslop, L.A. et al., 2016. Towards clinical application of pronuclear transfer to prevent mitochondrial DNA disease. *Nature*, 534(7607), pp.383–6.

Ibba, M. & Soll, D., 2000. Aminoacyl-tRNA synthesis. *Annual review of Biochemistry*, 69, pp.617–650.

Indrieri, A. et al., 2012. Mutations in COX7B cause microphthalmia with linear skin lesions, an unconventional mitochondrial disease. *American Journal of Human Genetics*, 91(5), pp.942–9.

Jain, I.H. et al., 2016. Hypoxia as a therapy for mitochondrial disease. *Science*, 352(6281), pp.54–61.

Jędrzejowska, M. et al., 2014. Severe phenotypes of SMARD1 associated with novel mutations of the IGHMBP2 gene and nuclear degeneration of muscle and Schwann cells. *European Journal of Paediatric Neurology*, 18(2), pp.183–192.

Jin, S.M. et al., 2010. Mitochondrial membrane potential regulates PINK1 import and proteolytic destabilization by PARL. *Journal of Cell Biology*, 191(5), pp.933–942.

Kadenbach, B. et al., 1983. Separation of mammalian cytochrome c oxidase into 13 polypeptides by a sodium dodecyl sulfate-gel electrophoretic procedure. *Analytical Biochemistry*, 129(2), pp.517–521.

Kalifa, L. et al., 2009. Evidence for a role of FEN1 in maintaining mitochondrial DNA integrity. *DNA Repair*, 8(10), pp.1242–1249.

Kanabus, M. et al., 2015. Bi-allelic CLPB mutations cause cataract, renal cysts, nephrocalcinosis and 3-methylglutaconic aciduria, a novel disorder of mitochondrial protein disaggregation. *Journal of Inherited Metabolic Disease*, 38(2), pp.211–9.

Kanabus, M., Heales, S.J. & Rahman, S., 2013. Development of Pharmacological Strategies for Mitochondrial Disorders. *British Journal of Pharmacology*, 171(8), pp.1798–1817.

Karadimas, C.L. et al., 2000. Recurrent myoglobinuria due to a nonsense mutation in the COX I gene of mitochondrial DNA. *Neurology*, 55, pp.644–649.

Kasamatsu, H. & Vinograd, J., 1974. Replication of circular DNA in eukaryotic cells. *Annual Review of Biochemistry*, 43, pp.695–719.

Kasamatsu, H. & Vinograd, J., 1973. Unidirectionality of replication in mouse mitochondrial DNA. *Nature: New biology*, 241(108), pp.103–5.

Keightley, J.A. et al., 1996. A microdeletion in cytochrome c oxidase (COX) subunit III associated with COX deficiency and recurrent myoglobinuria. *Nature Genetics*, 12(4), pp.410–6.

Kennedy, E.P. & Lehninger, A.L., 1949. Oxidation of fatty acids and tricarboxylic acid cycle intermediates by isolated rat liver mitochondria. *The Journal of Biological Chemistry*, 179(2), pp.957–72.

Kitada, T. et al., 1998. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature*, 392(6676), pp.605–608.

Kohda, M. et al., 2016. A Comprehensive Genomic Analysis Reveals the Genetic Landscape of Mitochondrial Respiratory Chain Complex Deficiencies. *PLoS Genetics*, 12(1), p.e1005679.

Konovalova, S. & Tyynismaa, H., 2013. Mitochondrial aminoacyl-tRNA synthetases in human disease. *Molecular Genetics and Metabolism*, 108(4), pp.206–211.

Kopajtich, R. et al., 2014. Mutations in GTPBP3 cause a mitochondrial translation defect associated with hypertrophic cardiomyopathy, lactic acidosis, and encephalopathy. *American Journal of Human Genetics*, 95(6), pp.708–20.

Kornblum, C. et al., 2013. Loss-of-function mutations in MGME1 impair mtDNA replication and cause multisystemic mitochondrial disease. *Nature Genetics*, 45(2), pp.214–9.

Kovářová, N. et al., 2016. Tissue- and species-specific differences in cytochrome c oxidase assembly induced by SURF1 defects. *Biochimica et Biophysica Acta*, 1862(4), pp.705–715.

Krebs, H.. & Johnson, W.., 1937. The Role of Citric Acid in Intermediate Metabolism in Animal Tissues. *Enzymologia*, 4, pp.148–156.

Kuhn, C.D. et al., 2015. On-enzyme refolding permits small RNA and tRNA surveillance by the CCA-adding enzyme. *Cell*, 160(4), pp.644–658.

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), pp.1073–1081.

Kunz, W.S. et al., 2000. Mitochondrial complex I deficiency in the epileptic focus of patients with temporal lobe epilepsy. *Annals of Neurology*, 48(5), pp.766–773.

de la Paz, M.P. et al., 2010. Rare Diseases Epidemiology Research. In Springer Netherlands, pp. 17–39.

Lamperti, C. et al., 2012. MELAS-like encephalomyopathy caused by a new pathogenic mutation in the mitochondrial DNA encoded cytochrome c oxidase subunit I. *Neuromuscular Disorders*, 22(11), pp.990–994.

Lander, E.S. et al., 2001. Initial sequencing and analysis of the human genome. *Nature*, 409(6822), pp.860–921.

Lander, E.S. & Botstein, D., 1987. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. *Science (New York, N.Y.)*, 236(4808), pp.1567–1570.

Larsson, N.G. et al., 1998. Mitochondrial transcription factor A is necessary for mtDNA maintenance and embryogenesis in mice. *Nature Genetics*, 18(3), pp.231–236.

Laskowski, R.A. et al., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26, pp.283–291.

Leary, S.C. et al., 2013. COX19 mediates the transduction of a mitochondrial redox signal from SCO1 that regulates ATP7A-mediated cellular copper efflux. *Molecular Biology of the Cell*, 24(6), pp.683–691.

Leary, S.C. et al., 2004. Human SCO1 and SCO2 have independent, cooperative functions in copper delivery to cytochrome c oxidase. *Human Molecular Genetics*, 13, pp.1839–1848.

Leary, S.C., 2010. Redox regulation of SCO protein function: controlling copper at a mitochondrial crossroad. *Antioxidants & Redox Signaling*, 13(9), pp.1403–1416.

Lebrecht, D. et al., 2010. Respiratory chain deficiency precedes the disrupted calcium homeostasis in chronic doxorubicin cardiomyopathy. *Cardiovascular Pathology*, 19(5), pp.e167–74.

Legati, A. et al., 2016. New genes and pathomechanisms in mitochondrial disorders unraveled by NGS technologies. *Biochimica et Biophysica Acta*, 1857(8), pp.1326–35.

Lewis, J.A. & Tata, J.R., 1973. A rapidly sedimenting fraction of rat liver endoplasmic reticulum. *Journal of Cell Science*, 13(2), pp.447–59.

Lewis, M.R. & Lewis, W.H., 1915. Mitochondria (and other cytoplasmic structures) in tissue cultures. *American Journal Of Anatomy*, 17(3), pp.339–401.

Li, X. et al., 2002. Isolation and characterization of the putative nuclear modifier gene MTO1 involved in the pathogenesis of deafness-associated mitochondrial 12 S rRNA A1555G mutation. *Journal of Biological Chemistry*, 277(30), pp.27256–27264.

Li, X. & Guan, M.-X., 2002. A human mitochondrial GTP binding protein related to tRNA modification may modulate phenotypic expression of the deafness-associated mitochondrial 12S rRNA mutation. *Molecular and Cellular Biology*, 22(21), pp.7701–11.

Li, Y. et al., 2006. Cytochrome c oxidase subunit IV is essential for assembly and respiratory function of the enzyme complex. *Journal of Bioenergetics and Biomembranes*, 38, pp.283–291.

Li, Y. et al., 2010. LRP4 mutations alter Wnt/beta-catenin signaling and cause limb and kidney malformations in Cenani-Lenz syndrome. *American Journal of Human Genetics*, 86(5), pp.696–706.

Liao, H.X. & Spremulli, L.L., 1990. Identification and initial characterization of translational

initiation factor 2 from bovine mitochondria. *Journal of Biological Chemistry*, 265(23), pp.13618–13622.

Liao, H.X. & Spremulli, L.L., 1989. Interaction of bovine mitochondrial ribosomes with messenger RNA. *Journal of Biological Chemistry*, 264(13), pp.7518–7522.

Lieber, D.S. et al., 2013. Targeted exome sequencing of suspected mitochondrial disorders. *Neurology*, 80(19), pp.1762–1770.

Lim, S.C. et al., 2014. A Founder Mutation in PET100 Causes Isolated Complex IV Deficiency in Lebanese Individuals with Leigh Syndrome. *American Journal of Human Genetics*.

Lim, S.C. et al., 2012. The Ighmbp2 helicase structure reveals the molecular basis for disease-causing mutations in DMSA1. *Nucleic acids research*, 40(21), pp.11009–22.

Limongelli, G. et al., 2010. Prevalence and natural history of heart disease in adults with primary mitochondrial respiratory chain disease. *European journal of heart failure*, 12(2), pp.114–21.

Lipshultz, S.E. et al., 2013. Risk stratification at diagnosis for children with hypertrophic cardiomyopathy: an analysis of data from the Pediatric Cardiomyopathy Registry. *Lancet*, 382(9908), pp.1889–97.

Lipshultz, S.E. et al., 2003. The incidence of pediatric cardiomyopathy in two regions of the United States. *The New England Journal of Medicine*, 348(17), pp.1647–55.

Luft, R. et al., 1962. A case of severe hypermetabolism of nonthyroid origin with a defect in the maintenance of mitochondrial respiratory control: a correlated clinical, biochemical, and morphological study. *The Journal of Clinical Investigation*, 41, pp.1776–804.

Lupski, J.R. et al., 2010. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *The New England Journal of Medicine*, 362(13), pp.1181–1191.

M-SAR Collaboration, T., 2013. Mutations in COQ2 in Familial and Sporadic Multiple-System Atrophy. *The New England journal of medicine*, pp.1–12.

Ma, J. & Spremulli, L.L., 1996. Expression, purification, and mechanistic studies of bovine mitochondrial translational initiation factor 2. *Journal of Biological Chemistry*, 271(10), pp.5805–5811.

Maehara, M. et al., 1983. Cytochrome c oxidase deficiency in menkes kinky hair disease. *Brain and Development*, 5(6), pp.533–540.

Manfredi, G. et al., 1995. A new mutation associated with MELAS is located in a mitochondrial DNA polypeptide-coding gene. *Neuromuscular Disorders*, 5, pp.391–398.

Marchington, D.R. et al., 1998. Evidence from human oocytes for a genetic bottleneck in an mtDNA disease. *The American Journal of Human Genetics*, 63(3), pp.769–775.

Margulies, M. et al., 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437(7057), pp.376–80.

Margulis, L., 1970. Origin of Eukaryotic Cells. *Yale University Press, New Haven, CT*. Available at: http://www.goodreads.com/book/show/6304471-origin-of-eukaryotic-cells.

Marotta, R. et al., 2011. Novel single base pair COX III subunit deletion of mitochondrial DNA associated with rhabdomyolysis. *Journal of Clinical Neuroscience*, 18(2), pp.290–292.

Massa, V. et al., 2008. *Severe infantile encephalomyopathy caused by a mutation in COX6B1, a nucleus-encoded subunit of cytochrome c oxidase.*,

Martin, W.F. et al 2016. Endosymbiotic theories for eukaryote origin. *Philos Trans R Soc Lond B Biol Sci.* 370(1678)

Matsuda, N. et al., 2010. PINK1 stabilized by mitochondrial depolarization recruits Parkin to damaged mitochondria and activates latent Parkin for mitophagy. *Journal of Cell Biology*, 189(2), pp.211–221.

Mayr, J.A. et al., 2012. Lack of the mitochondrial protein acylglycerol kinase causes Sengers syndrome. *American Journal of Human Genetics*, 90(2), pp.314–20.

McFarland, R. et al., 2004. A novel sporadic mutation in cytochrome c oxidase subunit II as a cause of rhabdomyolysis. *Neuromuscular Disorders*, 14(2), pp.162–166.

Melchionda, L. et al., 2014. Mutations in APOPT1, encoding a mitochondrial protein, cause cavitating leukoencephalopathy with cytochrome c oxidase deficiency. *American Journal of Human Genetics*, 95(3), pp.315–325.

Metodiev, M.D. et al., 2009. Methylation of 12S rRNA Is Necessary for In Vivo Stability of the Small Subunit of the Mammalian Mitochondrial Ribosome. *Cell Metabolism*, 9(4), pp.386–397.

Meyers, D.E., Basha, H.I. & Koenig, M.K., 2013. Mitochondrial cardiomyopathy: pathophysiology, diagnosis, and management. *Texas Heart Institute Journal*, 40(4), pp.385–94.

Mick, D.U. et al., 2012. MITRAC links mitochondrial protein translocation to respiratory-chain assembly and translational regulation. *Cell*, 151(7), pp.1528–1541.

Mick, D.U., Fox, T.D. & Rehling, P., 2011. Inventory control: cytochrome c oxidase assembly regulates mitochondrial translation. *Nature reviews. Molecular cell biology*, 12(1), pp.14–20.

Mikelsaar, R., 1983. Human mitochondrial genome and the evolution of methionine transfer ribonucleic acids. *Journal of Theoretical Biology*, 105(2), pp.221–232.

Minczuk, M. et al., 2008. Development of a single-chain, quasi-dimeric zinc-finger nuclease for the selective degradation of mutated human mitochondrial DNA. *Nucleic Acids Research*, 36(12), pp.3926–3938.

Minczuk, M. et al., 2002. Localisation of the human hSuv3p helicase in the mitochondrial matrix and its preferential unwinding of dsDNA. *Nucleic Acids Research*, 30(23), pp.5074–5086.

Minczuk, M. et al., 2011. TEFM (c17orf42) is necessary for transcription of human mtDNA. *Nucleic Acids Research*, 39(10), pp.4284–99.

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

Mitchell, P., 1976. Possible molecular mechanisms of the protonmotive function of cytochrome systems. *Journal of Theoretical Biology*, 62(2), pp.327–367.

Miyakawa, I. et al., 1987. Isolation of morphologically intact mitochondrial nucleoids from the yeast, Saccharomyces cerevisiae. *Journal of Cell Science*, 88 (Pt 4), pp.431–9.

Mkaouar-Rebai, E. et al., 2011. Molecular-clinical correlation in a family with a novel heteroplasmic Leigh syndrome missense mutation in the mitochondrial cytochrome c oxidase III gene. *Journal of Child Neurology*, 26(1), pp.12–20.

Mobley, B.C. et al., 2009. A novel homozygous SCO2 mutation, p.G193S, causing fatal infantile cardioencephalomyopathy. *Clinical Neuropathology*, 28(2), pp.143–9.

Montero, R. et al., 2016. GDF-15 is elevated in children with mitochondrial diseases and is induced by mitochondrial dysfunction. *PLoS ONE*, 11(2).

Montoya, J. et al., 1982. Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 79(23), pp.7195–7199.

Morava, E. et al., 2006. Mitochondrial disease criteria: Diagnostic applications in children. *Neurology*, 67(10), pp.1823–1826.

Moreno-Lastres, D. et al., 2012. Mitochondrial complex I plays an essential role in human respirasome assembly. *Cell Metabolism*, 15(3), pp.324–335.

Mozdy, A.D., McCaffery, J.M. & Shaw, J.M., 2000. Dnm1p GTPase-mediated mitochondrial fission is a multi-step process requiring the novel integral membrane component Fis1p. *Journal of Cell Biology*, 151(2), pp.367–379.

Murakami, Y. et al., 2012. Mechanism for release of alkaline phosphatase caused by glycosylphosphatidylinositol deficiency in patients with hyperphosphatasia mental retardation syndrome. *Journal of Biological Chemistry*, 287(9), pp.6318–6325.

Murakami, Y. et al., 2003. PIG-W is critical for inositol acylation but not for flipping of glycosylphosphatidylinositol-anchor. *Molecular biology of the cell*, 14(10), pp.4285–95.

Nagaike, T. et al., 2001. Identification and characterization of mammalian mitochondrial tRNA nucleotidyltransferases. *The Journal of Biological Chemistry*, 276(43), pp.40041–9.

Nagao, A., Suzuki, T. & Suzuki, T., 2007. Aminoacyl-tRNA surveillance by EF-Tu in mammalian mitochondria. *Nucleic Acids Symposium Series (2004)*, (51), pp.41–2.

Narendra, D. et al., 2008. Parkin is recruited selectively to impaired mitochondria and promotes their autophagy. *Journal of Cell Biology*, 183(5), pp.795–803.

Nass, S. & Nass, M.M., 1963. Intramitochondrial fibers with DNA characteristics II. Enzymatic and other hydrolytic treatments. *The Journal of Cell Biology*, 19, pp.613–29.

Neumann, M. et al., 2006. Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. *Science*, 314(5796), pp.130–133.

Ng, B.G. & Freeze, H.H., 2014. Human genetic disorders involving glycosylphosphatidylinositol (GPI) anchors and glycosphingolipids (GSL). *Journal of Inherited Metabolic Disease*, 38(1), pp.171–178.

Ng, S.B. et al., 2010. Exome sequencing identifies the cause of a mendelian disorder. *Nature Genetics*, 42(1), pp.30–35.

Ng, Y.S. & Turnbull, D.M., 2015. Mitochondrial disease: genetics and management. *Journal of Neurology*, 263(1), pp.179–91.

Ngo, H.B. et al., 2014. Distinct structural features of TFAM drive mitochondrial DNA packaging

versus transcriptional activation. *Nature Communications*, 5, p.3077.

Nijtmans, L.G. et al., 1998. Assembly of cytochrome-c oxidase in cultured human cells. *European Journal of Biochemistry / FEBS*, 254, pp.389–394.

Nishino, I., Spinazzola, A. & Hirano, M., 1999. Thymidine phosphorylase gene mutations in MNGIE, a human mitochondrial disorder. *Science*, 283(5402), pp.689–692.

Nizzardo, M. et al., 2015. Gene therapy rescues disease phenotype in a spinal muscular atrophy with respiratory distress type 1 (SMARD1) mouse model. *Science Advances*, 1(2), p.e1500078.

Nugent, A.W. et al., 2003. The epidemiology of childhood cardiomyopathy in Australia. *The New England Journal of Medicine*, 348(17), pp.1639–46.

Ojala, D., Montoya, J. & Attardi, G., 1981. tRNA punctuation model of RNA processing in human mitochondria. *Nature*, 290(5806), pp.470–474.

Osawa, S. et al., 1992. Recent Evidence for Evolution of the Genetic Code. *Microbiological Reviews*, 56(1), pp.229–264.

Ostergaard, E. et al., 2015. Mutations in COA3 cause isolated complex IV deficiency associated with neuropathy, exercise intolerance, obesity, and short stature. *Journal of Medical Genetics*, 52(3), pp.203–207.

Otera, H. et al., 2010. Mff is an essential factor for mitochondrial recruitment of Drp1 during mitochondrial fission in mammalian cells. *Journal of Cell Biology*, 191(6), pp.1141–1158.

Pagliarini, D.J. et al., 2008. A Mitochondrial Protein Compendium Elucidates Complex I Disease Biology. *Cell*, 134(1), pp.112–123.

Pagliarini, D.J. et al., 2008. A mitochondrial protein compendium elucidates complex I disease biology. *Cell*, 134, pp.112–123.

Palade, G.E., 1953. An electron microscope study of the mitochondrial structure. *Journal of Histochemistry & Cytochemistry*, 1(4), pp.188–211.

Papadopoulou, L.C. et al., 1999. Fatal infantile cardioencephalomyopathy with COX deficiency and mutations in SCO2, a COX assembly gene. *Nature Genetics*, 23(3), pp.333–7.

Patton, J.R. et al., 2005. Mitochondrial myopathy and sideroblastic anemia (MLASA): Missense mutation in the pseudouridine synthase 1 (PUS1) gene is associated with the loss of tRNA pseudouridylation. *Journal of Biological Chemistry*, 280(20), pp.19823–19828.

Pelletier, S. et al., 2012. An early onset progressive motor neuron disorder in Scyl1-deficient mice is associated with mislocalization of TDP-43. *The Journal of Neuroscience*, 32(47), pp.16560–73.

Pelletier, S., 2016. SCYL pseudokinases in neuronal function and survival. *Neural Regeneration Research*, 11(1), pp.42–4.

Piekutowska-Abramczuk, D. et al., 2009. SURF1 missense mutations promote a mild Leigh phenotype. *Clinical Genetics*, 76(2), pp.195–204.

Pitceathly, R.D.S., Taanman, J.-W., et al., 2013. COX10 Mutations Resulting in Complex Multisystem Mitochondrial Disease That Remains Stable Into Adulthood. *JAMA Neurology*, 70(12), pp.1556–61.

Pitceathly, R.D.S., Rahman, S., et al., 2013. NDUFA4 mutations underlie dysfunction of a cytochrome c oxidase subunit linked to human neurological disease. *Cell Reports*, 3(6), pp.1795–805.

Piwowarski, J. et al., 2003. Human polynucleotide phosphorylase, hPNPase, is localized in mitochondria. *Journal of Molecular Biology*, 329(5), pp.853–857.

Ponchaut, S., van Hoof, F. & Veitch, K., 1992. Cytochrome aa3 depletion is the cause of the deficient mitochondrial respiration induced by chronic valproate administration. *Biochemical Pharmacology*, 43(3), pp.644–7.

Powell, C.A., Nicholls, T.J. & Minczuk, M., 2015. Nuclear-encoded factors involved in post-transcriptional processing and modification of mitochondrial tRNAs in human disease. *Frontiers in Genetics*, 6, p.79.

Poyau, A. et al., 2000. Missense mutations in SURF1 associated with deficient cytochrome c oxidase assembly in Leigh syndrome patients. *Human Genetics*, 106(2), pp.194–205.

Pronicka, E. et al., 2013. The natural history of SCO2 deficiency in 36 Polish children confirmed the genotype–phenotype correlation. *Mitochondrion*, 13(6), pp.810–6.

Pronicki, M. et al., 2010. A homozygous mutation in the SCO2 gene causes a spinal muscular atrophy like presentation with stridor and respiratory insufficiency. *European Journal of Paediatric Neurology Society*, 14(3), pp.253–60.

Pyle, A. et al., 2015. Respiratory chain deficiency in nonmitochondrial disease. *Neurology. Genetics*, 1(1), p.e6.

Rahman, S. et al., 1999. A missense mutation of cytochrome oxidase subunit II causes defective assembly and myopathy. *American Journal of Human Genetics*, 65, pp.1030–1039.

Rahman, S. et al., 2001. A SURF1 gene mutation presenting as isolated leukodystrophy. *Annals of Neurology*, 49(6), pp.797–800.

Rahman, S. et al., 2012. Hearing in 44-45 year olds with m.1555A>G, a genetic mutation predisposing to aminoglycoside-induced deafness: a population based cohort study. *BMJ open*, 2, p.e000411.

Rahman, S. et al., 1996. Leigh syndrome: Clinical features and biochemical and DNA abnormalities. *Annals of Neurology*, 39(3), pp.343–351.

Rahman, S. & Hanna, M.G., 2009. Diagnosis and therapy in neuromuscular disorders: diagnosis and new treatments in mitochondrial diseases. *Journal of Neurology, Neurosurgery, and Psychiatry*, 80(9), pp.943–53.

Rapezzi, C. et al., 2013. Diagnostic work-up in cardiomyopathies: bridging the gap between clinical phenotypes and final diagnosis. A position statement from the ESC Working Group on Myocardial and Pericardial Diseases. *European Heart Journal*, 34(19), pp.1448–58.

Richter, R. et al., 2010. Translation termination in human mitochondrial ribosomes. *Biochemical Society transactions*, 38(6), pp.1523–6.

Rieske, J.S., 1976. Composition, structure, and function of complex III of the respiratory chain. *Biochimica et Biophysica Acta*, 456(2), pp.195–247.

Riley, L.G. et al., 2010. Mutation of the mitochondrial tyrosyl-tRNA synthetase gene, YARS2, causes myopathy, lactic acidosis, and sideroblastic anemia--MLASA syndrome. *American Journal of Human Genetics*, 87(1), pp.52–59.

Rone, M.B. et al., 2012. Identification of a dynamic mitochondrial protein complex driving cholesterol import, trafficking, and metabolism to steroid hormones. *Molecular Endocrinology*, 26(11), pp.1868–82.

Rorbach, J., Nicholls, T.J.J. & Minczuk, M., 2011. PDE12 removes mitochondrial RNA poly(A) tails and controls translation in human mitochondria. *Nucleic Acids Research*, 39(17), pp.7750–63.

Rossmanith, W., 2011. Localization of human RNase Z isoforms: dual nuclear/mitochondrial targeting of the ELAC2 gene product by alternative translation initiation. *PloS one*, 6(4), p.e19152.

Ruzzenente, B. et al., 2012. LRPPRC is necessary for polyadenylation and coordination of translation of mitochondrial mRNAs. *The EMBO journal*, 31(2), pp.443–456.

Sali, A. & Blundell, T.L., 1993. Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, 234(3), pp.779–815.

Sanger, F., Air, G.M., et al., 1977. Nucleotide sequence of bacteriophage phi X174 DNA. *Nature*, 265(5596), pp.687–95.

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, pp.5463–5467.

Santel, A. & Fuller, M.T., 2001. Control of mitochondrial morphology by a human mitofusin. *Journal of Cell Science*, 114(Pt 5), pp.867–874.

Santoro, L. et al., 2000. A novel SURF1 mutation results in Leigh syndrome with peripheral neuropathy caused by cytochrome c oxidase deficiency. *Neuromuscular Disorders*, 10(6), pp.450–3.

Sarig, O. et al., 2013. Infantile mitochondrial hepatopathy is a cardinal feature of MEGDEL syndrome (3-methylglutaconic aciduria type IV with sensorineural deafness, encephalopathy and Leigh-like syndrome) caused by novel mutations in SERAC1. *American Journal of Medical Genetics. Part A*, 161(9), pp.2204–15.

Sasarman, F. et al., 2012. A novel mutation in YARS2 causes myopathy with lactic acidosis and sideroblastic anemia. *Human Mutation*, 33(8), pp.1201–1206.

Sasarman, F. et al., 2011. The 2-thiouridylase function of the human MTU1 (TRMU) enzyme is dispensable for mitochondrial translation. *Human Molecular Genetics*, 20(23), pp.4634–4643.

Sasarman, F. et al., 2015. The 3’ addition of CCA to mitochondrial tRNASer(AGY) is specifically impaired in patients with mutations in the tRNA nucleotidyl transferase TRNT1. *Human Molecular Genetics*, 24(10), pp.2841–7.

Savoiardo, M. et al., 2002. MRI in Leigh syndrome with SURF1 gene mutation. *Annals of Neurology*, 51(1), pp.138–139.

Scaglia, F. et al., 2004. Clinical spectrum, morbidity, and mortality in 113 pediatric patients with mitochondrial disease. *Pediatrics*, 114(4), pp.925–31.

Schägger, H. & Von Jagow, G., 1999. Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem*, pp.223–231.

Schiffmann, R. & Van Der Knaap, M.S., 2009. Invited Article: An MRI-based approach to the diagnosis of white matter disorders. *Neurology*, 72(8), pp.750–759.

Schmidt, W.M. et al., 2015. Disruptive SCYL1 Mutations Underlie a Syndrome Characterized by Recurrent Episodes of Liver Failure, Peripheral Neuropathy, Cerebellar Atrophy, and Ataxia. *American Journal of Human Genetics*, 97(6), pp.855–61.

Schmidt, W.M. et al., 2007. Mutation in the Scyl1 gene encoding amino-terminal kinase-like protein causes a recessive form of spinocerebellar neurodegeneration. *EMBO reports*, 8(7), pp.691–7.

Schon, E. a, DiMauro, S. & Hirano, M., 2012. Human mitochondrial DNA: roles of inherited and somatic mutations. *Nature reviews. Genetics*, 13, pp.878–90.

Schon, E.A., Bonilla, E. & DiMauro, S., 1997. Mitochondrial DNA mutations and pathogenesis. *Journal of Bioenergetics and Biomembranes*, 29(2), pp.131–149.

Schwarz, J.M. et al., 2010. MutationTaster evaluates disease-causing potential of sequence alterations. *Nature Methods*, 7(8), pp.575–576.

Shahni, R. et al., 2013. A distinct mitochondrial myopathy, lactic acidosis and sideroblastic anemia (MLASA) phenotype associates with YARS2 mutations. *American Journal of Medical Genetics. Part A*, 161A(9), pp.2334–8.

Shahni, R. et al., 2015. Signal transducer and activator of transcription 2 deficiency is a novel disorder of mitochondrial fission. *Brain*, 138(Pt 10), pp.2834–46.

Shendure, J. et al., 2005. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science*, 309(5741), pp.1728–1732.

Shojaee, S. et al., 2008. Genome-wide Linkage Analysis of a Parkinsonian-Pyramidal Syndrome Pedigree by 500 K SNP Arrays. *American Journal of Human Genetics*, 82(6), pp.1375–1384.

Shokeir, M.H. & Shreffler, D.C., 1969. Cytochrome oxidase deficiency in Wilson's disease: a suggested ceruloplasmin function. *Proceedings of the National Academy of Sciences of the United States of America*, 62(3), pp.867–72.

Shteyer, E. et al., 2009. Exocrine pancreatic insufficiency, dyserythropoeitic anemia, and calvarial hyperostosis are caused by a mutation in the COX4I2 gene. *American Journal of Human Genetics*, 84(3), pp.412–7.

Siwar, B.G. et al., 2014. Two novel mutations in COII and tRNAHis mitochondrial genes in asthenozoospermic infertiles men. *Biochemical and Biophysical Research Communications*, 450(1), pp.610–615.

Skladal, D., Halliday, J. & Thorburn, D.R., 2003. Minimum birth prevalence of mitochondrial respiratory chain disorders in children. *Brain*, 126, pp.1905–1912.

Skowyra, D. et al., 1997. F-box proteins are receptors that recruit phosphorylated substrates to

the SCF ubiquitin-ligase complex. *Cell*, 91(2), pp.209–219.

Skulachev, V.P., 2001. Mitochondrial filaments and clusters as intracellular power-transmitting cables. *Trends in Biochemical Sciences*, 26(1), pp.23–29.

Smeitink, J.A.M., Koopman, W.J.H. & Willems, P.H.G.M., 2012. Monogenic Mitochondrial Disorders. *New England Journal of Medicine*, 366, pp.1132–1141.

Smits, P., Smeitink, J. & van den Heuvel, L., 2010. Mitochondrial translation and beyond: processes implicated in combined oxidative phosphorylation deficiencies. *Journal of Biomedicine & Biotechnology*, 2010, p.737385.

Sofou, K. et al., 2014. A multicenter study on Leigh syndrome: disease course and predictors of survival. Orphanet Journal of Rare Disease, 15(9), pp 52.

Sokabe, M. et al., 2009. The structure of alanyl-tRNA synthetase with editing domain. *Proceedings of the National Academy of Sciences of the United States of America*, 106(27), pp.11028–11033.

Soleimanpour-Lichaei, H.R. et al., 2007. mtRF1a Is a Human Mitochondrial Translation Release Factor Decoding the Major Termination Codons UAA and UAG. *Molecular Cell*, 27(5), pp.745–757.

de Souza-Pinto, N.C. et al., 2009. Novel DNA mismatch-repair activity involving YB-1 in human mitochondria. *DNA Repair*, 8(6), pp.704–719.

Steenweg, M.E. et al., 2012. Leukoencephalopathy with thalamus and brainstem involvement and high lactate 'LTBL' caused by EARS2 mutations. *Brain*, 135(Pt 5), pp.1387–1394.

Stenson, P.D. et al., 2009. The Human Gene Mutation Database: 2008 update. *Genome Medicine*, 1(1), p.13.

Sun, F. et al., 2005. Crystal structure of mitochondrial respiratory membrane protein Complex II. *Cell*, 121(7), pp.1043–1057.

Suomalainen, A. et al., 2011. FGF-21 as a biomarker for muscle-manifesting mitochondrial respiratory chain deficiencies: a diagnostic study. *The Lancet. Neurology*, 10(9), pp.806–18.

Suzuki, T. et al., 2002. Taurine as a constituent of mitochondrial tRNAs: New insights into the functions of taurine and human mitochondrial diseases. *EMBO Journal*, 21(23), pp.6581–6589.

Suzuki, T., Nagao, A. & Suzuki, T., 2011. Human Mitochondrial tRNAs: Biogenesis, Function, Structural Aspects, and Diseases. *Annual Review of Genetics*, 45, pp.299–329.

Suzuki, T. & Suzuki, T., 2007. Chaplet column chromatography: isolation of a large set of individual RNAs in a single step. *Methods in Enzymology*, 425, pp.231–9.

Swairjo, M.A. & Schimmel, P.R., 2005. Breaking sieve for steric exclusion of a noncognate amino acid from active site of a tRNA synthetase. *Proceedings of the National Academy of Sciences of the United States of America*, 102(4), pp.988–993.

Szabadkai, G. et al., 2006. Chaperone-mediated coupling of endoplasmic reticulum and mitochondrial Ca2+ channels. *Journal of Cell Biology*, 175(6), pp.901–911.

Szklarczyk, R. et al., 2013. A mutation in the FAM36A gene, the human ortholog of COX20,

impairs cytochrome c oxidase assembly and is associated with ataxia and muscle hypotonia. *Human Molecular Genetics*, 22(4), pp.656–67.

Taanman, J.W., 1999. The mitochondrial genome: structure, transcription, translation and replication. *Biochimica et Biophysica Acta*, 1410(2), pp.103–23.

Taanman, J.W. & Capaldi, R.A., 1993. Subunit VIa of yeast cytochrome c oxidase is not necessary for assembly of the enzyme complex but modulates the enzyme activity. Isolation and characterization of the nuclear-coded gene. *Journal of Biological Chemistry*, 268(25), pp.18754–18761.

Tabebi, M. et al., 2015. A novel mutation MT-COIII m.9267G>C and MT-COI m.5913G>A mutation in mitochondrial genes in a Tunisian family with maternally inherited diabetes and deafness (MIDD) associated with sever nephropathy. *Biochemical and Biophysical Research Communications*.

Tachibana, M. et al., 2009. Mitochondrial gene replacement in primate offspring and embryonic stem cells. *Nature*, 461(7262), pp.367–372.

Tachibana, M. et al., 2012. Towards germline gene therapy of inherited mitochondrial diseases. *Nature*, pp.1–5.

Taguchi, N. et al., 2007. Mitotic phosphorylation of dynamin-related GTPase Drp1 participates in mitochondrial fission. *Journal of Biological Chemistry*, 282(15), pp.11521–11529.

Tam, E.W.Y. et al., 2008. A novel mitochondrial DNA mutation in COX1 leads to strokes, seizures, and lactic acidosis. *Neuropediatrics*, 39, pp.328–334.

Tamiy, G. et al., 2014. A Mutation of COX6A1 causes a recessive axonal or mixed form of charcot-marie-tooth disease. *American Journal of Human Genetics*, 95(3), pp.294–300.

Tarnopolsky, M. a et al., 2004. Novel SCO2 mutation (G1521A) presenting as a spinal muscular atrophy type I phenotype. *American Journal of Medical Genetics. Part A*, 125A(3), pp.310–4.

Tay, S.K. et al., 2005. Unusual clinical presentations in four cases of Leigh disease, cytochrome C oxidase deficiency, and SURF1 gene mutations. *Journal of Child Neurology*, 20(8), pp.670–674.

Taylor, R.W. et al., 2014. Use of whole-exome sequencing to determine the genetic basis of multiple mitochondrial respiratory chain complex deficiencies. *JAMA*, 312(1), pp.68–77.

Thompson Legault, J. et al., 2015. A Metabolic Signature of Mitochondrial Dysfunction Revealed through a Monogenic Form of Leigh Syndrome. *Cell Reports*, 13(5), pp.981–9.

Tiranti, V. et al., 2000. *A novel frameshift mutation of the mtDNA COIII gene leads to impaired assembly of cytochrome c oxidase in a patient affected by Leigh-like syndrome.*,

Tiranti, V. et al., 1997. Identification of the gene encoding the human mitochondrial RNA polymerase (h-mtRPOL) by cyberscreening of the Expressed Sequence Tags database. *Human Molecular Genetics*, 6(4), pp.615–625.

Tiranti, V. et al., 2009. Loss of ETHE1, a mitochondrial dioxygenase, causes fatal sulfide toxicity in ethylmalonic encephalopathy. *Nature Medicine*, 15(2), pp.200–205.

Tiranti, V. et al., 1998. Mutations of SURF-1 in Leigh disease associated with cytochrome c

oxidase deficiency. *American Journal Human Genetics*, 63(6), pp.1609–1621.

Tomecki, R. et al., 2004. Identification of a novel human nuclear-encoded mitochondrial poly(A) polymerase. *Nucleic Acids Research*, 32(20), pp.6001–6014.

Torres-Torronteras, J. et al., 2011. Hematopoietic gene therapy restores thymidine phosphorylase activity in a cell culture and a murine model of MNGIE. *Gene therapy*, 18(8), pp.795–806.

Trumpower, B.L., 1990. The Protonmotive Q Cycle: Energy Transduction by Coupling of Proton Translocation to Electron Transfer by the Cytochrome bc1 Complex. *The Journal of Biological Chemistry*, 265(20), pp.11409–11412.

Tsukihara, T. et al., 1996. The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 A. *Science*, 272(5265), pp.1136–44.

Uusimaa, J. et al., 2003. A mutation in mitochondrial DNA-encoded cytochrome c oxidase II gene in a child with Alpers-Huttenlocher-like disease. *Pediatrics*, 111(3), pp.e262–8.

Vafai, S.B. & Mootha, V.K., 2012. Mitochondrial disorders as windows into an ancient organelle. *Nature*, 491(7424), pp.374–83.

Valente, E.M. et al., 2004. Hereditary early-onset Parkinson's disease caused by mutations in PINK1. *Science*, 304(5674), pp.1158–1160.

Valente, L. et al., 2009. Identification of novel mutations in five patients with mitochondrial encephalomyopathy. *Biochimica et Biophysica Acta*, 1787, pp.491–501.

Valnot, I., von Kleist-Retzow, J.C., et al., 2000. A mutation in the human heme A:farnesyltransferase gene (COX10) causes cytochrome c oxidase deficiency. *Human Molecular Genetics*, 9(8), pp.1245–9.

Valnot, I., Osmond, S., et al., 2000. Mutations of the SCO1 gene in mitochondrial cytochrome c oxidase deficiency with neonatal-onset hepatic failure and encephalopathy. *American Journal Human Genetics*, 67(5), pp.1104–9.

Varlamov, D.A. et al., 2002. Metabolic consequences of a novel missense mutation of the mtDNA CO I gene. *Human molecular genetics*, 11(16), pp.1797–805.

Venter, J.C. et al., 2001. The sequence of the human genome. *Science*, 291(5507), pp.1304–51.

Vesela, K. et al., 2004. Clinical, biochemical and molecular analyses of six patients with isolated cytochrome c oxidase deficiency due to mutations in the SCO2 gene. *Acta Paediatrica*, 93(10), pp.1312–17.

Vinothkumar, K.R., Zhu, J. & Hirst, J., 2014. Architecture of mammalian respiratory complex I. *Nature*, 515(7525), pp.80–84.

Viscomi, C. et al., 2011. In vivo correction of COX deficiency by activation of the AMPK/PGC-1?? axis. *Cell Metabolism*, 14(1), pp.80–90.

Vissing, C.R. et al., 2013. Recurrent myoglobinuria and deranged acylcarnitines due to a mutation in the mtDNA MT-CO2 gene. *Neurology*, 80(20), pp.1908–10.

Wahbi, K. et al., 2015. Long-term cardiac prognosis and risk stratification in 260 adults presenting with mitochondrial diseases. *European Heart Journal*, 36(42), pp.2886–93.

Walker, J.E., 1998. ATP synthesis by rotary catalysis (Nobel lecture). In *Angewandte Chemie - International Edition*. pp. 2308–2319.

Wallace, D.C. et al., 1988. Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. *Science*, 242(4884), pp.1427–1430.

Wang, X., Yan, Q. & Guan, M.-X., 2010. Combination of the loss of cmnm5U34 with the lack of s2U34 modifications of tRNALys, tRNAGlu, and tRNAGln altered mitochondrial biogenesis and respiration. *Journal of Molecular Biology*, 395(5), pp.1038–48.

Warburg, O., 1913. No Title. *Arch. Gesamte. Physiol.*, 154, pp.599–617.

Ware, S.M., 2011. Genetic diagnosis in pediatric cardiomyopathy: clinical application and research perspectives. *Progress in Pediatric Cardiology*, 31(2), pp.99–102.

Watkins, H., Ashrafian, H. & Redwood, C., 2011. Inherited cardiomyopathies. *The New England Journal of Medicine*, 364(17), pp.1643–56.

Watson, J.D. & Crick, F.H., 1953. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171(4356), pp.737–8.

Wedatilake, Y. et al., 2013. SURF1 deficiency: a multi-centre natural history study. *Orphanet Journal of Rare Diseases*, 8(1), p.96.

Wedatilake, Y. et al., 2015. Tubular aggregates caused by serine active site containing 1 (SERAC1) mutations in a patient with a mitochondrial encephalopathy. Neuropathol Appl Neurobiol. 41(3):399-402.

Wedatilake, Y. et al., 2016. TRNT1 deficiency: clinical, biochemical and molecular genetic features. *Orphanet Journal of Rare Diseases*, 11(1), p.90.

Wei, Y.-L. et al., 2009. Novel mitochondrial DNA mutations associated with Chinese familial hypertrophic cardiomyopathy. *Clinical and Experimental Pharmacology & Physiology*, pp.933–939.

Weintraub, R.G. & Semsarian, C., 2013. Paediatric cardiomyopathy: getting to the heart of the matter. *Lancet*, 382(9908), pp.1866–7.

Wenz, T. et al., 2008. Activation of the PPAR/PGC-1?? Pathway Prevents a Bioenergetic Deficit and Effectively Improves a Mitochondrial Myopathy Phenotype. *Cell Metabolism*, 8(3), pp.249–256.

Weraarpachai, W. et al., 2009. Mutation in TACO1, encoding a translational activator of COX I, results in cytochrome c oxidase deficiency and late-onset Leigh syndrome. *Nature Genetics*, 41(7), pp.833–7.

Weraarpachai, W. et al., 2012. Mutations in C12orf62, a factor that couples COX I synthesis with cytochrome c oxidase assembly, cause fatal neonatal lactic acidosis. *American Journal of Human Genetics*, 90(1), pp.142–51.

Wharton, D.C. & Tzagoloff, A., 1967. Cytochrome oxidase from beef heart mitochondria. *Methods in Enzymology*, 10(C), pp.245–250.

Williams, L. et al., 2004. Cytochrome c oxidase subassemblies in fibroblast cultures from patients carrying mutations in COX10, SCO1, or SURF1. *Journal of Biological Chemistry*, 279(9), pp.7462–7469.

Wilson, K.S. & Prochaska, L.J., 1990. Phospholipid vesicles containing bovine heart mitochondrial cytochrome c oxidase and subunit III-deficient enzyme: analysis of respiratory control and proton translocating activities. *Archives of Biochemistry and Biophysics*, 282(2), pp.413–20.

Winston, J.T. et al., 1999. A family of mammalian F-box proteins. *Current Biology*, 9(20), pp.1180–1182.

Wiseman, D.H. et al., 2013. A novel syndrome of congenital sideroblastic anemia, B-cell immunodeficiency, periodic fevers, and developmental delay (SIFD). *Blood*, 122(1), pp.112–23.

Wong, Y.C. & Holzbaur, E.L., 2014. Optineurin is an autophagy receptor for damaged mitochondria in parkin-mediated mitophagy that is disrupted by an ALS-linked mutation. *Proceedings of the National Academy of Sciences of the United States of America*, 111(42), pp.E4439–48.

Woriax, V.L., Spremulli, G.H. & Spremulli, L.L., 1996. Nucleotide and aminoacyl-tRNA specificity of the mammalian mitochondrial elongation factor EF-Tu.Ts complex. *Biochimica et Biophysica Acta*, 1307(1), pp.66–72.

Wortmann, S.B. et al., 2013. Inborn errors of metabolism with 3-methylglutaconic aciduria as discriminative feature: proper classification and nomenclature. *Journal of Inherited Metabolic Disease*, 36(6), pp.923–8.

Wortmann, S.B. et al., 2012. Mutations in the phospholipid remodeling gene SERAC1 impair mitochondrial function and intracellular cholesterol trafficking and cause dystonia and deafness. *Nature Genetics*, 44(7), pp.797–802.

Yamada, M. et al., 2016. Genetic Drift Can Compromise Mitochondrial Replacement by Nuclear Transfer in Human Oocytes. *Cell Stem Cell*, 18(6), pp.749–54.

Yaplito-Lee, J. et al., 2007. Cardiac manifestations in oxidative phosphorylation disorders of childhood. *The Journal of Pediatrics*, 150(4), pp.407–11.

Zhao, J. et al., 2011. Human MIEF1 recruits Drp1 to mitochondrial outer membranes and promotes mitochondrial fusion rather than fission. *The EMBO Journal*, 30(14), pp.2762–2778.

Zhao, T. et al., 2011. Loss of nuclear activity of the FBXO7 protein in patients with parkinsonian-pyramidal syndrome (PARK15). *PLoS ONE*, 6(2).

Zhou, Z.D. et al., 2015. F-box protein 7 mutations promote protein aggregation in mitochondria and inhibit mitophagy. *Human Molecular Genetics*, 24(22), pp.6314–6330.

Zhou, Z.D. et al., 2016. Linking F-box protein 7 and parkin to neuronal degeneration in Parkinson's disease (PD). *Molecular Brain*, 9(1), p.41.

Zhu, Z. et al., 1998. SURF1, encoding a factor involved in the biogenesis of cytochrome c oxidase, is mutated in Leigh syndrome. *Nature Genetics*, 20(4), pp.337–343.

10. Appendix

Patient	Clinical features	Candidate gene/s	Study identified/ recruited
Family 1			
S1	Sideroblastic anaemia, failure to thrive and hepato-renal failure, Pearson-like	*YARS*	This study
S2	Developmental delay, deafness, partial villous atrophy, inflammatory enteropathy, bone marrow dysplasia, Pearson-like	*YARS*	This study
Family 2			
S3	Developmental delay, renal tubular wasting, hypocalcaemia	*Not identified*	This study
S4	Congenital lactic acidosis, microcephaly, dysmorphic features, gastroesophageal reflux, failure to thrive	*Not identified*	This study
S5	Renal tubulopathy, hypocalcaemia, developmental delay	*Not identified*	This study
Family 3			
S6	Developmental delay, dilated cardiomyopathy	*AARS2*	This study
S7	Developmental delay, dilated cardiomyopathy	*AARS2*	This study
Family 4			

Patient	Clinical features	Candidate gene/s	Study identified/ recruited
S8	Hereditary sensory motor neuropathy, cholestatic liver disease, mtDNA depletion in muscle	*IGHMBP2/SCYL1*	This study
S9	Hereditary sensory motor neuropathy, cholestatic liver disease, mtDNA depletion in muscle	*IGHMBP2/SCYL1*	This study
Family 5			
S10	Fever, sideroblastic anaemia, sensorineural deafness, retinopathy	*TRNT1*	Another study
S11	Fever, sideroblastic anaemia, sensorineural deafness, retinopathy	*TRNT1*	Another study
Family 6			
S12	Lactic acidosis, cardiomyopathy	*MRPL44*	This study
S13	Lactic acidosis, cardiomyopathy, adrenal insufficiency	*MRPL44*	This study
Family 7			
S14	Leigh syndrome, hypothyroidism, demyelinating polyneuropathy	*ADA1*	Another study
S15	Leigh syndrome, acute encephalopathy, motor and speech delay, seizures	*ADA1*	Another study
Family 8			
S16	Ponto cerebellar atrophy, microcephaly, cortical visual impairment, epilepsy	Awaiting results	Recruited to 100,000GS
S17	Profound visual impairment, nystagmus, microcephaly, infantile spasms	Awaiting results	Recruited to 100,000GS

Patient	Clinical features	Candidate gene/s	Study identified/ recruited
Family 9			
S18	Ataxia, dystonia	Awaiting results	Recruited to 100,000GS
S19	Ataxia, dystonia	Awaiting results	Recruited to 100,000GS
Family 10			
S20	Developmental delay, seizures, hypotonia cerebellar atrophy	*PIGW*	This study
S21	Developmental delay, hypotonia, cerebellar atrophy	*PIGW*	This study
Family 11			
S22	Epileptic encephalopathy, developmental delay	Not identified	This study
S23	Epileptic encephalopathy, liver failure, (Alpers-like)	Not identified	This study
S24	Generalised epilepsy, Lennox-Gastaut spectrum, severe learning difficulties, low CSF folate	*FOLR1*	Candidate gene sequencing
S25	Leigh syndrome, history of regression, microcephaly, seizures, ataxic gait, motor neuropathy	*C12orf65*	This study
S26	Global developmental delay, lactic acidosis, microcephaly	*No DNA available*	

Patient	Clinical features	Candidate gene/s	Study identified/ recruited
S27	Diplegic motor disorder, expressive language delay, motor neuronopathy	*FBXO7*	This study
S28	Sensorineural hearing loss, central hypotonia, retinal dystrophy	*PNPT1*	100,000GS
S29	Acute encephalopathy, focal seizure with secondary generalisation, dystonia, pyramidal signs, pyrexia, hepatomegaly, microcytic anaemia, liver biopsy: microvesicular fatty change	Awaiting results	Recruited to 100,000GS
S30	Muscle weakness, recurrent chest infections	*TK2*	Candidate gene sequencing
S31	Lactic acidosis, sensorineural deafness, developmental delay, cerebral cyst, dystonia	Awaiting results	Recruited to 100,000GS
S32	Cardiomyopathy, Sensorineural deafness, retinal dystrophy, renal failure	Not identified	This study
S33	Cerebellar atrophy, sensorimotor neuropathy, septo-optic dysplasia	Not identified	This study
S34	Sideroblastic anaemia, sensory motor axonal neuropathy, congenital deformity of hands and fingers, refractory bicuspid aortic valve, Retinal dystrophy	Awaiting results	Recruited to 100,000GS
S35	Chronic inflammatory demyelinating polyneuropathy, epilepsy, chronic anaemia, colon inflammation	DNA available not adequate	Awating
S36	Basal ganglia lesions, ptosis, apnoea	Awaiting results	Recruited to 100,000GS
S37	Giant axonal neuropathy, unsteady gait and poor motor skills, progressive motor and sensory neuropathy, central precocious puberty, progressive scoliosis	No DNA/consent available	
S38	Developmental delay, lactic acidosis	*SCO1*	Another study

Patient	Clinical features	Candidate gene/s	Study identified/ recruited
S39	Hypopituitarism, retinal dystrophy, fatigue		To be recruited to 100,000 GS
S40	Chronic renal disease treated by renal transplant, dystonic movement disorder, mild left ventricular hypertrophy	Awaiting results	Recruited to 100,000GS
S41	Recurrent episodes of encephalopathy, hypoglycaemia, rhabdomyolysis, episodic ataxia	Awaiting results	Recruited to 100,000GS
S42	Nephrotic syndrome, mitral valve regurgitation, transient ischaemic attacks with dysphasia	Not identified	This study
S43	Lactic acidosis, hyperammonaemia, sensorineural deafness, 3-methyl glutaconic aciduria, Leigh-like	*SERAC1*	This study
S44			To be recruited to 100,000 GS
S45	Congenital ptosis, partial agenesis of corpus callosum, headaches	Awaiting results	Recruited to 100,000GS
S46	Leigh-like, developmental delay, growth failure	Awaiting results	Recruited to 100,000GS
S47	Sideroblastic anaemia, raised lactate, biventricular hypertrophy	No DNA/consent available	
S48	Developmental delay, hypotonia, low serum copper	No DNA/consent available	

Patient	Clinical features	Candidate gene/s	Study identified/ recruited
S49	Global developmental delay with motor regression, progressive feeding difficulties (partially gastrostomy fed), dislocated hips	Not identified	This study
S50	Non-obstructive hypertrophic cardiomyopathy, Cardiovascular collapse, mild right hemiparesis speech delay	No DNA/consent available	
S51	Leukoencephalopathy, epilepsy, dystonia, learning difficulties, exocrine pancreatic insufficiency		Part of another study
S52	Epilepsy, right sided weakness, learning difficulties		To be recruited to 100,000 GS
S53	Motor delay, joint laxity, poor weight gain, fatiguability		Part of another study
S54	Poor feeding, faltering growth, hypotonia, motor delay, developmental regression, slurred speech	Awaiting results	Recruited to 100,000GS
S55	Deafness, choreoathetosis	No DNA/consent available	
S56	Ataxia, ptosis, myopathy, movement disorder		To be recruited to 100,000 GS
S57	Myopathy, poor coordination, ataxia, speech delay		To be recruited to 100,000 GS
S58	Epileptic encephalopathy, global developmental delay, gastro-oesophageal reflux, motor disorder, visual impairment	*KCNQ2*	Candidate gene sequencing

Patient	Clinical features	Candidate gene/s	Study identified/ recruited
S59	Sensorineural deafness, hypermetropia, developmental delay, sensorimotor neuropathy, truncal ataxia	Awaiting results	Recruited to 100,000GS
S60	Limb dystonia, myopathy on EMG		
S61	Global developmental delay, dystonia, eosinophilic colitis, delayed visual maturation, seizures, faltering growth	Awaiting results	Recruited to 100,000GS
S62	Four limb dystonia, motor skill regression, developmental delay	Awaiting results	Recruited to 100,000GS
S63	Myoclonic epileptic encephalopathy, developmental delay,		To be recruited to 100,000 GS
S64	Peripheral neuropathy, auditory neuropathy, bilateral ptosis, gastrostomy fed, respiratory failure	*SLC25A2*	Candidate gene sequencing
S65	Generalised hypotonia and weakness, mild to moderate sensorineural deafness, bulbar dysfunction, gastrooesophageal reflux disease	Awaiting results	Recruited to 100,000GS
S66	Speech delay, four limb motor disorder, ataxia, behavioural problems, ventrolateral thalamic changes, Leigh-like	Not identified	This study
S67	Seizures, dystonia, congenital lactic acidosis	*RARS2*	Candidate gene sequencing
S68	Global developmental delay, ataxia, fatigue, myopathic EMG, hypermobility		To be recruited to 100,000 GS
S69	Infantile epileptic encephalopathy, infantile spasms, sensorineural deafness	Awaiting results	Recruited to

Patient	Clinical features	Candidate gene/s	Study identified/ recruited
			100,000GS
S70	Profound hypotonia with developmental delay, visual impairment, myopathic EMG		To be recruited to 100,000 GS
S71	Cataracts, developmental delay, hypertrophic cardiomyopathy, 3-methylglutaconic aciduria	*ATAD3A*	This study
S72	Vomiting, poor growth, developmental delay, infantile spasms, sensorineural deafness, pigmentary retinopathy		To be recruited to 100,000 GS
S73	Pancytopenia, seizures, hypotonia, immunodeficiency	Not identified	This study
S74	Infantile spasms, developmental delay	*AIG13*	This study

100,000 GS: 100,000 Genome Study