Persistent Ribonucleotides and Mitochondrial DNA Metabolism (in Normal and Disease States).

Chloe Frances Moss

University College London

and

The Francis Crick Institute

PhD Supervisor(s): Dr Ian Holt and Professor Anthony Schapira

A thesis submitted for the degree of

Doctor of Philosophy

University College London

October 2017

Declaration

I Chloe Moss 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.

Abstract

It has been known for over 40 years that the mammalian mitochondrial genome contains persistent sporadic ribonucleotides, for which there has been no molecular explanation to date. To better understand the significance of incorporated ribonucleotides in mitochondrial DNA I developed a deep sequencing method to determine the precise location of ribonucleotides in the mitochondrial DNA of a range of cell and tissue types. Using Next Generation Sequencing I have shown that the vast majority of ribosubstitution events in murine mitochondrial DNA from solid tissues are dominated by adenosine bases derived from adenosine triphosphate, the product of oxidative phosphorylation. In contrast, proliferating cells show a much lower rate of ribonucleotide incorporation. However, this can be manipulated by inducing cell cycle arrest, which leads to an increase in adenosine monophosphate incorporation rate, comparable to that of solid tissues. This study confirms that the specificity of the incorporated ribonucleotide base identity is directly influenced by the ratio of ribonucleotides to deoxyribonucleotides (NTP:dNTPs) within the mitochondrial compartment of the cell.

In a mouse model of Mpv17 deficiency there is a mitochondrial specific deoxynucleotide depletion which results in a significant increase in riboguanosine incorporation in mitochondrial DNA. The results described within this thesis demonstrate that elevated riboguanosine incorporation is associated with early-onset mitochondrial DNA depletion in liver and late-onset multiple deletions in brain of the Mpv17 ablated mice. These findings suggest that aberrant ribonucleotide incorporation is a primary mitochondrial DNA abnormality that can result in pathology.

Impact Statement

The main body of this thesis focuses on a novel advanced methodology to locate ribonucleotides in mtDNA using Next Generation Sequencing technologies. Moreover, the data indicate that this is not the only DNA modification which can be mapped to single nucleotide resolution using this technique, and it has been previously demonstrated that this technique is applicable to genomic DNA and not just mitochondrial DNA. Enabling the precise mapping of a vast array of DNA modifications *in vivo* facilitates a better understanding of epigenetics and mutagenesis among many other things. The versatility of this approach will be of great benefit to the scientific community.

Furthermore, better understanding of mitochondrial DNA modifications will help gain a greater perspective of mitochondrial DNA metabolism and repair, a subject which is often under contentious debate.

Improved understanding of mitochondrial DNA modifications and the role they play in diseased states will enable future advances in mitochondrial therapies. In many cases of mitochondrial disease, such as those described in this thesis associated with mutations in *MPV17*, the mechanism of disease is poorly understood which presents an obstacle in developing targeted therapies for patients. Understanding the fundamental cause and mechanism of mitochondrial dysfunction and the role of mitochondrial DNA damage will prove instrumental in progressing further on this front.

This work was an example of a successful collaboration predominantly between three groups in London and Edinburgh, across fields of expertise where information and know-how were exchanged to produce a novel and exciting project. Furthermore, such collaborations outside of the mitochondrial field have given the work extra scope to penetrate into the wider scientific community.

Acknowledgement

I am eternally indebted to the unrelenting support from my wonderful Mum, and my long-suffering boyfriend- Fred, both of whom rode the rollercoaster of this PhD alongside me. They were a never-ending source of love and humour and I couldn't have done it without either of them.

I also owe my thanks to my supervisor, Ian. I am forever grateful that he took a chance on a chemist with limited biology experience, and shared his passion and wealth of knowledge about the field. He was a patient and inspiring teacher and has left me with a great deal of skills and tools to continue my career in academia.

I am hugely grateful to Tony, who picked up a homeless PhD student, facilitating a smooth transition at what was a turbulent time, enabling me to complete my final year in a fun and stimulating environment and keeping my love for science alive right to the end. Thanks to my caretaker, Greg who, with a shared love of running, taught me that there's no lab disaster too big that it can't be cured by a long run around Regent's park!

And lastly to my fellow 'Mitogals'; Radha, Ilaria and Mara. Three wonderful friends for life who endured all the highs and lows of the lab with me every day for 4+ years. Here's to many more beers, PCRs and Ceilidhs (in that precise order)!

Table of Contents

Table of figures

List of tables

Abbreviations

1D-AGE neutral one dimensional agarose gel electrophoresis

2D-AGE neutral/neutral two-dimensional agarose gel electrophoresis

4-OHT Tamoxifen

8-oxodG 8-hydroxydeoxyguanine

^{32}P Phosphorus-32 – a radioactive isotope of phosphorus

7S RNA ssRNA from LSP to CI, often hybridized to mtDNA

7S DNA ssDNA from CI to ETAS, often forming D-loop

adPEO autosomal dominant progressive external opthalmoplegia

ADP adenosine diphosphate

AGS Aicardi–Goutières syndrome

ATP adenosine triphosphate

bp base pairs

BSA bovine serum albumin

CD circular dichroism

CTP cytosine triphosphate

COX cytochrome oxidase

°C degree Celsius

D-loop displacement loop

dATP deoxyadenosine triphosphate

dCTP deoxycytidine triphonsphate

dGTP deoxyguanosine triphonsphate

dTTP deoxythymidine triphosphate

DDM n-Dodecyl β-D-maltoside

DGUOK deoxyguanosine kinase

DMEM Dulbecco's modifid Eagle's media

DMSO dimethyl sulphoxide

DSB double strand break

dsDNA double stranded DNA

DTT dithiothreitol

EDTA Ethylenediaminetetraacetic acid

EM Electron Microscopy

EtBr ethidium bromide

FBS foetal bovine serum

g relative centrifugal force

GAPDH glyceraldehyde-3-phosphate dehydrogenase

GTP guanosine triphosphate

HR Homologous recombination

HSP1/HSP2 heavy strand mitochondrial promoter 1/2

H-strand heavy strand of mtDNA

IMM inner mitochondrial membrane

kB kilo bases

KO genetically engineered knockout

LSP light strand mitochondrial promoter

L-strand light strand of mtDNA

MBN mung bean nuclease

MDS Mitochondrial depletion syndrome

MELAS mitochondrial encephalomyopathy lactic acidosis and stroke-like episodes

MERRF myoclonus epilepsy and ragged-red fibres

MGME1 mitochondrial genome maintenance exonuclease 1

MMEJ Microhomology mediated end joining

mtDNA mitochondrial DNA

mtRIs mitochondrial replication intermediates

mTERF mitochondrial transcription termination factor

mtSSB single-stranded DNA binding protein

NADH Nicotinamide adenine dinucleotide

NCR non-coding region

NFW nuclease free water

NGS next generation sequencing

NHEJ Non-homologous end joining

nt nucleotide position

O_H canonical origin of heavy strand replication

O_L origin of light strand replication

Ori-b Additional origin of heavy strand replication, second to O_H

Ori-L origin of light strand replication

ori-z zone of initiation of mtDNA replication in major arc

OXPHOS oxidative phosphorylation

PAGE polyacrylamide gel electrophoresis

PBS phosphate buffered saline

PCR polymerase chain reaction

PEO progressive external opthalmoplegia

PGC1-α Peroxisome proliferator-activated receptor gamma coactivator 1-alpha

PK proteinase K

PMSF phenylmethanesulphonylfluoride

POLG mitochondrial DNA polymerase gamma

POLRMT mitochondrial RNA polymerase

qPCR quantitative PCR

rAMP ribo-adenosine 5'-monophosphate

rCMP ribo- cytidine 5'-monophosphate

rGMP ribo-guanosine 5'-monophosphate

rTMP ribo - thymidine 5'-monophosphate

rUMP ribo -uridine 5'-monophosphate

RER ribonucleotide excision repair

RITOLS RNA incorporation throughout the lagging strand

RNA/DNA hybrid duplex of RNA strand and complimentary DNA strand

RNA-DNA 5´ RNA – DNA 3´ junction

RNase ribonuclease

RNase H bacterial ribonuclease H

RNase H1 mammalian ribonuclease H2

RNase H2 mammalian ribonuclease H2

RNase HII bacterial ribonuclease HII

ROS reactive oxygen species

RR ribonucleotide reductase

SAP shrimp alkaline phosphatase

SCM Strand-coupled mechanism of DNA replication

SDM Strand-displacement mechanism of mtDNA replication

SDS sodium dodecyl sulfate

SSB single strand break

ssDNA single stranded DNA

ssRNA single-stranded RNA

T- -tamoxifen; control MEFs

T+ +tamoxifen; Rnaseh1$_{-/-}$ knockout MEFs

TALENs transcription activator-like effector nucleases

TAS termination associated sequences

TBE tris-borate EDTA

TCA tricarboxylic acid cycle

TE Tris-EDTA

TFAM mitochondrial transcription factor A

TK2 thymidine kinase 2

Twinkle T7 gp-4 like protein with intramitochondrial nucleoid localization

UCP-1 uncoupling protein 1

UTP uridine triphosphate

w/v weight/volume

WT wild-type

Chapter 1. Introduction

The presence of mitochondria in eukaryotic cells is generally accepted as the result of a symbiosis event rising from fusion of α-protobacteria and archaebacteria (Andersson et al., 1998, Gray, 2012, Lang et al., 1997, Gray et al., 2001). Since then, the genetic content of the mitochondrial ancestor has diminished over time owing to redundancy and transfer of information to the nucleus. Only a vestige remains on an extrachromosomal element- the mitochondrial genome, which encodes 13 essential proteins required for oxidative phosphorylation and the RNA elements required for their expression (22 tRNAs and 2 rRNAs).

Within the cell, mitochondria form an interconnected and dynamic network which is facilitated by the fusion and fission of organelles within the cell (Figure 1). The individual mitochondria themselves are characterised by an intricately folded double membrane where the proteins involved in the electron transport chain are embedded. The membrane encases numerous proteins involved in protein import and folding, as well providing as the sites for organelle interactions including the endoplasmic reticulum. The cristae structure is vital for mitochondrial function as it serves to increase the surface area of the organelle, maximising energy production.

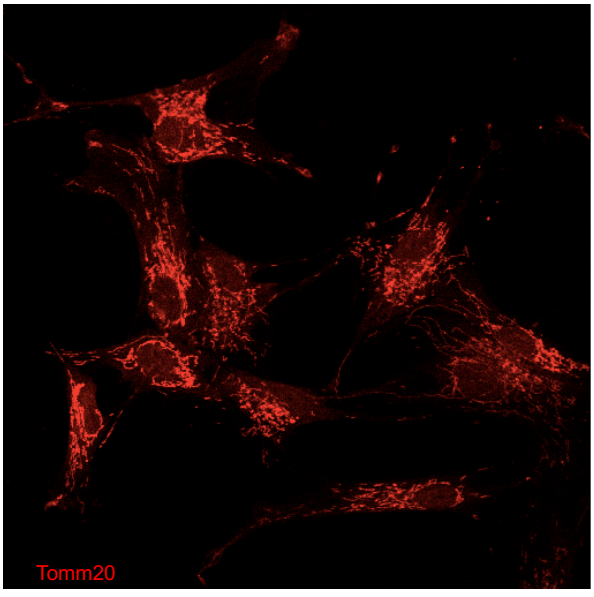


Figure 1: The interconnected mitochondrial matrix in Mouse Embryonic Fibroblasts (MEFs).

Confocal images from wild-type MEFs stained with fluorescent antibodies to mark the mitochondrial network (TOM20).

Mitochondria provide not only the proteins and machinery required for oxidative phosphorylation, but are also the hub for metabolism and are therefore essential for eukaryotic cell survival. They play an essential role in cell differentiation, signalling and autophagy.

1.1 The Role of Mitochondria in Cellular Respiration

Mitochondria are commonly referred to as the powerhouses of the cell and are responsible for generating almost all the cell's adenosine triphosphate (ATP) through phosphorylation of adenosine diphosphate (ADP).

1.1.1 The Tricarboxylic Acid Cycle

The primary series of reactions which ultimately release stored energy from the oxidation of carbohydrates, fats and proteins, is known as the tricarboxylic acid (TCA) cycle or the Krebs cycle. The TCA cycle also provides many of the precursors for essential amino acids and nucleotide synthesis.

This cycle, which takes place in the mitochondrial matrix in eukaryotic cells, utilises a two-carbon molecule in the form of acetyl-CoA, which is the product of catabolism of carbohydrates, fat and proteins. The acetyl-CoA enters the TCA cycle and the acetyl group is transferred to a four-carbon molecule; oxaloacetate, to form the six-carbon compound citrate. The citrate molecule goes through a series of reactions where it subsequently loses two carboxyl groups as carbon dioxide (CO_2) and reduces nicotinamide adenine dinucleotide, NAD^+, to NADH, ultimately regenerating a four-carbon oxaloacetate molecule. For every acetyl group transferred into the TCA cycle, three molecules of reduced NADH are produced.

The products of one turn of the cycle are ATP (or guanosine triphosphate, GTP), three molecules of NADH, one ubiquinol (QH_2, the reduced form of ubiquinone (also known as co-enzyme Q)) and two CO_2 molecules. The NADH and reduced flavin adenine dinucleotide ($FADH_2$) produced by the TCA cycle are then used by the

electron transport chain to produce ATP from ADP via oxidative phosphorylation (OXPHOS) (Figure 2).

1.1.2 The Electron Transport Chain

In eukaryotes, OXPHOS is carried out by five inner-membrane proteins within the mitochondria (four of which contain polypeptides encoded by the mitochondrial genome) and these are referred to as the electron transport chain since that is the role they play; facilitating the flow of electrons via various electron donors. This flow of electrons generates a pH gradient and electrical potential across the mitochondrial inner membrane. This potential energy is harnessed by another membrane-bound complex called ATP synthase (sometimes referred to as Complex V) which acts as a sophisticated proton channel enabling protons to flow back along this gradient. This flux of protons drives the phosphorylation of ADP to ATP in a process called chemiosmosis. This process is termed aerobic as it requires molecular oxygen.

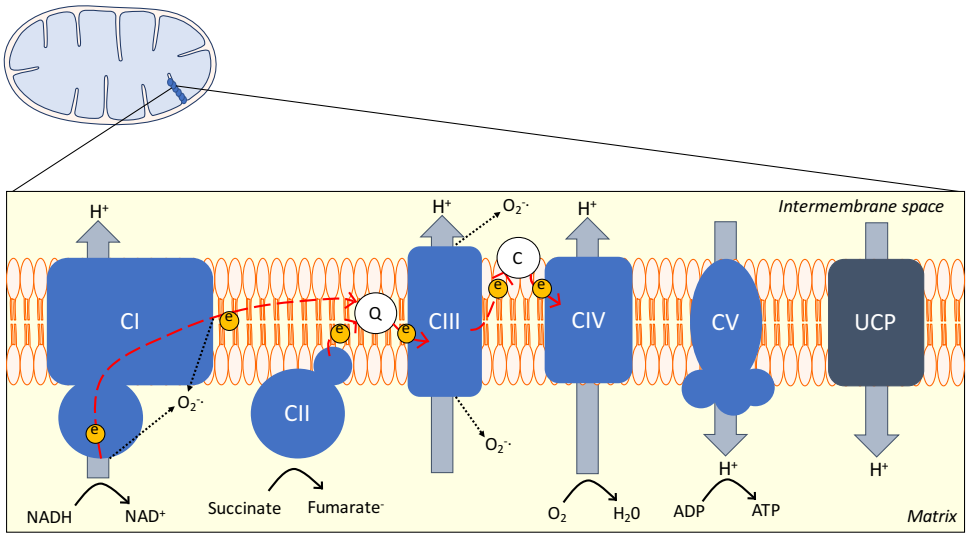


Figure 2: Schematic of the mammalian electron transport chain.

All 5 of the complexes, with the exception of Complex II, require polypeptides encoded by the mitochondrial genome to function and they function to facilitate electron transport across the membrane which ultimately produces ATP via oxidative phosphorylation. Superoxide $*O_2$ is produced as a by-product (at low levels) by complex I and II during cellular respiration.

1.1.3 Electron and Proton Leak

There are two important side-reactions that take place alongside the electron transport chain; electron and proton leak, both of which have an impact on the

efficiency of OXPHOS. Electrons can escape from the membrane-bound complexes and react undesirably with oxygen to form superoxide ($*O_2$), which is thought to be the major cause of oxidative damage within the mitochondria (Murphy, 2009, Sohal et al., 1995). Superoxide is a very reactive molecule but it is efficiently converted to hydrogen peroxide (H_2O_2) by superoxide dismutase (Mn-SOD in the matrix, Cu/Zn-SOD in the cytosol) and subsequently to oxygen and water by catalase or glutathione peroxidase (Brand et al., 2004).

Proton leak refers to the translocation of protons back across the membrane independent of ATP synthase, disrupting the proton gradient and reducing the capacity for ADP phosphorylation. Proton leak is the sum of two separate processes; basal proton leak and active proton leak. The former is unregulated and the latter is catalysed by specific mitochondrial inner membrane proteins and can be both inhibited and activated (Jastroch et al., 2010). The best-known example of active proton leak is facilitated by uncoupling-protein 1 (UCP1, Figure 2) which, for example, is activated in mammalian brown adipose as a mechanism of heat production.

1.1.4 Glycolysis

Under anaerobic conditions (the absence of free oxygen), glycolysis is an alternative, albeit less efficient, pathway that converts glucose into pyruvate, generating ATP. It is oxygen-independent, yet takes place in both the presence and the absence of oxygen in eukaryotes. Glycolysis occurs in the cytosol of the cell and produces numerous metabolites, many of which are subsequently used within the mitochondria.

1.2 The Mitochondrial Genome

Mitochondrial DNA (mtDNA) was discovered in the 1960s. It was firstly identified by electron microscopy (EM) in chick embryo mitochondria. MtDNA molecules were described as fibres with properties akin to nucleic acid (Nass and Nass, 1963). This discovery was corroborated by biochemical assays carried out in *Saccharomyces cerevisiae* which demonstrated the presence of significant amounts of DNA from purified mitochondria isolated by flotation in density gradients (Schatz et al., 1964).

Mammalian mtDNA is a compact 16 kilobase (kB) pair circular molecule with very few non-coding regions. The 13 proteins encoded by the mtDNA provide essential components of OXPHOS complexes I, III, IV and V and therefore all 37 mtDNA genes are critical for effective energy production. The two strands of mtDNA are termed heavy and light strand (H-strand and L-strand respectively) owing to different proportions of pyrimidine and purine nucleic acids on each strand and their buoyancies in a caesium chloride gradient.

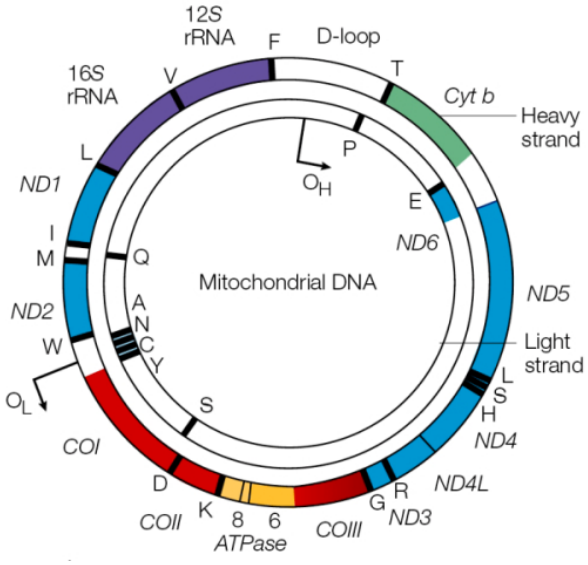


Figure 3: The mammalian mitochondrial genome.
[Reused with permission from Nature publishing group (Taylor and Turnbull, 2005)]

1.2.1 Nucleomitochondrial Interactions

Although the mitochondria have their own genome, functioning replisome and translation machinery, the mtDNA only codes for approximately only 1% of the total mitochondrial proteome. The other ~1,200 mitochondrial proteins are nuclear encoded, including all the proteins required for replication, transcription and translation. And their expression is dependent on many factors, not-least the effective crosstalk between the mitochondria and the nucleus.

Mitochondrial biogenesis and activity is not uniform across all cells and tissues and depends largely on the energetic and metabolic requirements of the individual cell or tissue which is ever-changing throughout the life of the cell. The mitochondrial content and activity of each cell is a dynamic feature, tightly controlled and

continually adjusted by mitochondria-nuclear crosstalk, also referred to as retrograde signalling. Changes in nuclear gene expressions can be triggered by several factors including mtDNA damage, reactive oxygen species (ROS), nutrient availability and fluctuating energy demands (Capps et al., 2003).

1.2.2 Regulatory Elements of MtDNA

MtDNA is approximately 16 kB in length (16,299 base pairs (bp) in mice and 16,569 bp in humans) and almost all the protein-coding genes are located on the heavy-strand sequence (the exception being the ND6 gene). Unlike the nuclear DNA, the mitochondrial genome has only one appreciable non-coding region (NCR); a short 1kB section of DNA which contains the initiation site(s) for leading strand replication (O_H), transcription and other regulatory elements (HSP1 and HSP2; heavy strand promoters, LSP; light strand promoter, conserved sequence block I-III (CSBI, CSBII and CSB III)). Interestingly, the NCR has been reported to have a significantly greater frequency of point mutations (in aged cells) than the rest of the mitochondrial genome (Michikawa et al., 1999).

1.2.3 The D-loop

Within this NCR is a stable triplex structure containing a short, additional stretch of DNA termed the mitochondrial D-loop (Nass, 1980). In mammals, the D-loop is approximately 500 bp in length and commonly referred to as 7S DNA based on its sedimentation properties. Its abundance is highly variable and its function is largely unknown. Interestingly, recent studies report the presence of an additional complementary RNA species, which has been christened the R-loop, whose presence appears to be linked to the segregation of mtDNA molecules within the organelle (Akman et al., 2016). This novel species is still not entirely understood, but may provide critical information about the replication-transcription switch in mitochondria.

1.2.4 MtDNA Topology and Packaging

The mtDNA molecule is a circular supercoiled plasmid-like entity. The majority of mtDNA molecules are clustered into compact nucleoprotein complexes referred to

as nucleoids. They are packaged by a variety of proteins, chiefly the widely-characterised mitochondrial Transcription Factor A, mitochondrial (TFAM), and tethered to the inner mitochondrial membrane (IMM). It is not known how nucleoids are utilised for replication or transcription, nor is it understood how nucleoids segregate within the mitochondria during fission.

There are more than two dozen distinct topoisomers identified, but it is not clear what roles they play (Kolesar et al., 2013). The different topoisomers include supercoiled circles, linear molecules, relaxed circles and higher-order catenated molecules among additional structures such as covalently closed circles which can be identified by 2D-AGE. Such structures are conserved across various mouse and human cell and tissue types, however the abundance of different topoisomers varies between tissues (Kolesar et al., 2013). This is in agreement with previous findings, reporting that mtDNA from human heart is highly catenated (Pohjoismäki et al., 2009). Again, the functional consequences of this are unclear but they are likely to provide an insight into the distinction between replication and transcription within the mitochondria.

Moreover, the nucleoids themselves can aggregate within the mitochondria and research has demonstrated that the initiation of mtDNA stress, elicited by TFAM deficiency, increases nucleoid aggregation and clustering and this phenomenon is postulated to be involved in priming the antiviral innate immune response (Reyes et al., 2015).

1.3 MtDNA Inheritance

MtDNA is significantly different from the DNA found within the nucleus for numerous reasons. The starkest being that mtDNA is maternally inherited. Shortly after fertilisation, paternal mitochondria, and the DNA contained within, are selectively destroyed and therefore not transmitted to the zygote and only the maternal mtDNA is passed onto the offspring. Maternal transmission of mtDNA is well established, however there are case studies which suggest that paternal mtDNA carry over is possible (Schwartz and Vissing 2002).

1.3.1 Heteroplasmy

Since each somatic cell can contain up to 1000's of mtDNA copies, a condition where there exists more than one variant of mtDNA co-exists is relatively common (Elliott et al., 2008, Payne et al., 2013). This is called heteroplasmy. In the case of mitochondrial disease caused by one or more mtDNA mutations, the severity of the pathology can be largely determined by the heteroplasmy level of the mutation(s) and the threshold above which the disease manifests is usually between 60-90%.

1.3.2 The Mitochondrial Bottleneck

One unique feature of mtDNA inheritance is the bottleneck phenomenon (Figure 4). Primary oocytes are derived from a dividing primordial germ cell containing a certain number of mitochondria harbouring either wild-type or mutant mtDNA. During mitosis of the primordial germ cell, the population of mitochondria are distributed between the two daughter cells seemingly without predisposition for the proportion of 'healthy' or 'mutant' mitochondria. This leads to heterogeneity amongst primordial oocytes, and depending on which is fertilised and eventually forms a foetus, this can lead to a variety of outcomes; in the worst instances, high mutant-load and perinatal death.

Although mtDNA does not follow the Mendelian rules of inheritance there are seemingly supplementary rules which work to encourage genetic drift within the population (Brown et al., 2001). The bottleneck hypothesis rationalises the dramatic reduction in the number of mtDNA molecules during germline transmission, which in turn explains the common incidence of significant changes in heteroplasmy levels throughout this time (Cree et al., 2008). However, recent studies report that the germline bottleneck is generated without an accompanying reduction in mtDNA content in mice (Cao et al., 2007, Cao et al., 2009).

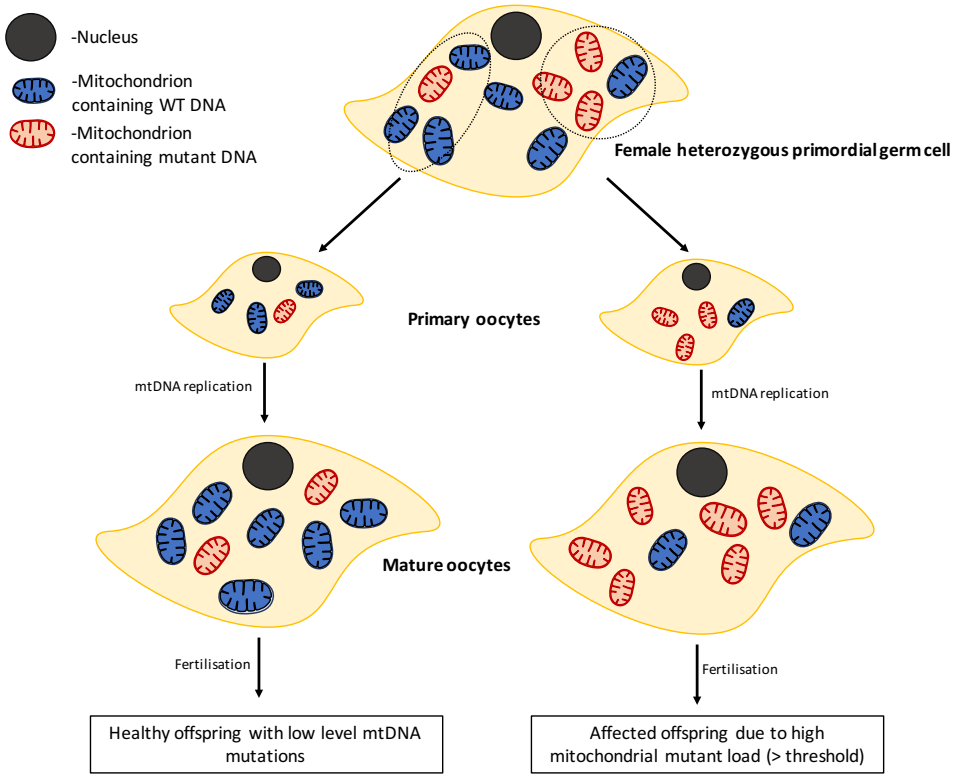


Figure 4: The mitochondrial bottleneck.

The question remains whether the nuclear background has an effect on mitochondrial DNA inheritance. Data derived from cytoplasmic hybrids, commonly referred to as cybrids (a cell derived from fusion of ρ^0 cells (cells lacking mtDNA) and an enucleated cell, mixing nuclear and mitochondrial genomes), has demonstrated that the nuclear background of a cell can influence the mtDNA heteroplasmy levels and their fluctuations over time (Dunbar et al., 1995). But it is unclear whether this mechanism, whatever it may be, is at play during germline transmission or not.

Despite recent advances and increasingly powerful sequencing methods, mitochondrial DNA inheritance is still poorly understood. It is not clear why, but mtDNA molecules containing deletions are almost never inherited; suggesting that there is some degree of selection against abhorrent mtDNA molecules (Chinnery et al., 2004, Stewart et al., 2008). It is for this reason that inheritance of mtDNA mutations, the origins of which are discussed later, is difficult to predict, as is the pathology of any mutations transmitted to the offspring, making the fundamental understanding of mtDNA replication and metabolism a vital piece of the puzzle.

1.4 Mitochondrial Biogenesis

Mitochondrial biogenesis is most commonly a response to the increased cellular energy demand of a cell whereby the cell increases its mitochondrial mass and mtDNA copy number to facilitate increased ATP production. Activation of biogenesis can be triggered by numerous external stimuli, including aerobic exercise.

1.4.1 PGC-1α

Peroxisome proliferator-activated receptor gamma coactivator 1-alpha (PGC-1α) is the master regulator of mitochondrial biogenesis and provides the link between external stimuli and mitochondrial biogenesis. It performs this task as a transcriptional co-activator which interacts with multiple nuclear transcription factors to enhance transcription of nuclear-encoded genes involved in energy metabolism. PGC-1α itself is activated by different stimuli including ROS and endurance exercise (Pilegaard et al., 2003).

1.4.2 Fusion and Fission

Biogenesis and mitochondrial degradation (mitophagy) are facilitated by the dynamic network of mitochondria which regularly undergo fusion (the joining together of two or more mitochondria) and fission (the division of a mitochondrion). Both processes are metabolically controlled by GTPases and the two processes oppose one another. The balance between fusion and fission is in flux, constantly reacting to the changing requirements of the cell (Mishra and Chan, 2016). Disruptions in both fusion and fission can hamper normal development and have been implicated in neurodegenerative diseases, such as Parkinson's (Narendra et al., 2008).

Fission is driven by the dynamin protein Drp1 which works by forming a spiral around the mitochondrion and constricting to cleave the inner and outer mitochondrial membranes and subsequently dividing the organelle. Fission is essential in dividing cells to populate the two daughter cells with sufficient mitochondrial mass. But fission isn't only active during the cell cycle. Fission can be activated by a multitude of external stimuli and an increase in fission, and therefore increased fragmentation of

the mitochondrial network, is associated with clearance of mitochondria by mitophagy.

The inverse process, mitochondrial fusion, is mediated by multiple proteins which are tethered to either the outer or inner mitochondrial membranes, most notably Mitofusins 1 and 2 (Mfn1 and Mfn2) and Opa1. Fusion can help combat cellular stress by combining multiple organelles if one or more contains damaged contents as a form of dilution (Schon and Gilkerson, 2010). Interestingly, nucleoids themselves do not appear to transfer DNA, therefore complementation of mitochondrial contents and functional mitochondrial proteins serves as a mechanism to compensate for the deleterious effects of mtDNA mutations within an individual organelle. Similarly, fusion of two or more organelles when one has become damaged by external stressors or environmental factors can enable the exchange of healthy lipids and proteins to mitigate the damage and works as a protective defence against cellular stress.

1.4.3 Mitophagy

Fusion is one way to combat low levels of damage within mitochondria, but this only works when the damage is below a certain threshold. Once this threshold is exceeded and the mitochondrion is beyond repair, it is destroyed. This autophagic clearance of mitochondria is termed mitophagy and is closely linked to the dynamic processes of fusion and fission within the cell (Egan et al., 2011). Elimination of extensively damaged mitochondria is essential to maintain a healthy mitochondrial network.

Mitophagy is relatively common in normal cells and there are numerous mechanisms of induction (see (Ding and Yin, 2012) for an extensive review). In fact, a study of cultured fibroblasts showed that 1 in every 5 daughter mitochondria are subsequently eliminated by mitophagy (Twig et al., 2008). Following fission of a damaged mitochondrion, damaged components or debris contained therein are segregated, resulting in two daughter mitochondria; one healthy and one containing damage. The latter is cleared via mitophagy to maintain a healthy population of mitochondria within the cell (Youle and van der Bliek, 2012).

1.5 Mitochondrial Transcription

Mitochondrial transcription is driven by the nuclear-encoded RNA polymerase (POLRMT) and a set of transcription factors (TFAM and transcription factor B2, mitochondrial (TFB2M)), all of which together can reconstitute transcription *in vitro* (Litonin et al., 2010), although there is an argument that TFB2M and POLRMT alone are sufficient (Shutt et al., 2010, Zollo et al., 2012, Lodeiro et al., 2012). POLRMT is a single-subunit, DNA-dependent protein with homology to the RNA polymerases of the bacteriophages T3 and T7 (Masters et al., 1987). However, unlike the bacteriophage RNA polymerases, POLRMT is unable to initiate RNA synthesis at mtDNA promoter sites alone, and instead requires additional factors (Wanrooij et al., 2008).

Transcription of the mtDNA encoded genes is initiated from promoters on each strand of DNA in two long polycistronic transcripts which are subsequently processed to form mature mRNA, tRNA and rRNAs by a handful of different RNases. There is one light-strand promoter (LSP) and two heavy-strand promoters (HSP1 and HSP2) found within the D-loop which have been identified using a variety of techniques including 5' mapping of primary mitochondrial transcripts and S1 nuclease protection assays (Montoya et al., 1983, Montoya et al., 1982, Yoza and Bogenhagen, 1984). The light strand transcript is that which is transcribed using the heavy strand DNA as a template, and vice versa for the heavy strand transcript.

The light strand is transcribed after initiation at LSP, forming one full length polycistronic transcript. The heavy strand transcript is also formed in a similar manner initiating at HSP. However, following the observation that heavy strand encoded rRNAs were synthesised at a higher rate than the heavy-strand mRNAs, Montoya *et al.* presented evidence for two distinct transcription initiation sites (Montoya et al., 1983). Another mechanism for modulating different RNA transcripts from the same strand is premature termination. This mechanism was revealed when it was shown that 16S rRNA molecules had imprecise 3'-termini, resulting from transcript termination at the adjacent $tRNA^{Leu(UUR)}$ gene (Kruse et al., 1989). The protein facilitating this process was identified as mitochondrial transcription termination factor 1 (mTERF1) which can bind and induce bending of the DNA

double helix (Shang and Clayton, 1994). mTERF1 is a member of the mTERF family of proteins, all of which are localised within the mitochondria, and are associated with not only termination of transcription, but also initiation of transcription and mtDNA replication (Hyvarinen et al., 2007, Roberti et al., 2009, Kruse et al., 1989). mTERF1 is the protein which binds just after the mitochondrial rRNAs (Fernandez-Silva et al., 1997) and has been associated with transcription termination. However, Terzioglu *et al.* demonstrated that the mTERF1 knockout mouse is viable, showing no change in mtRNA levels, indicating that the role of mTERF1 is not essential for mitochondrial function (Terzioglu et al., 2013).

Interestingly, the A3243G heteroplasmic mutation, located within human mTERF1's binding site, is the characteristic transition associated with mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes (MELAS) (Goto et al., 1990). *In vitro* studies demonstrate that this mutation disrupts the human mTERF binding site, impairing transcript termination (Shang and Clayton, 1994, Hess et al., 1991).

As the mitochondrial genome is deficient for introns it is believed that processing the polycistronic transcripts is a relatively simple process, involving few enzymes- not least the mitochondrial RNase P (Rossmanith and Karwan, 1998) and the zinc phosphodiesterase ELAC protein 2 (ELAC2). The organisation of the genome is such that almost all genes are separated by tRNA genes, therefore endonucleolytic excision of each tRNA facilitates the processing of the neighbouring mtRNAs. Akin to nuclear transcripts, all mtRNAs are polyadenylated by a polyA polymerase (mtPAP) with the exception of ND6 (the only heavy strand transcript) (Tomecki et al., 2004) (although they do not carry upstream polyA signals) (Taanman, 1999). Addition of CCA to the 3'- end of tRNAs is facilitated by TP(CTP):tRNA nucleotidyltransferase (Rossmanith et al., 1995).

1.5.1 Mitochondrial RNA Granules

There is still much more to be discovered about how immature mitochondrial transcripts are matured, modified, assembled and turned over within the organelle. More than ten years ago, researchers identified mitochondrial RNA granules (MRGs) by labelling the organelle with uridine analogue 5-bromouridine (BrU) which formed

punctate structures, not dissimilar to mtDNA nucleoids (Jourdain et al., 2013). Said structures colocalise with G-rich RNA sequence biding factor-1 (GRSF1) and the mitochondrial RNase P Complex both of which are associated with RNA processing, suggesting that the MRGs could be the sites of RNA maturation. Depletion of GRSF1 leads to an increase of immature RNA molecules and a decrease in mature RNAs (Jourdain et al., 2013, Jourdain et al., 2016) further supporting this conclusion.

1.6 Mitochondrial Translation

The 12S and 16S rRNA of the mitochondrial ribosome, together with $tRNA^{Val}$, are assembled with imported mitochondrial ribosomal proteins of the 39S mitochondrial large subunit and of the 28S small subunit, thus initiating the first steps of ribosome assembly (Richter-Dennerlein et al., 2015). Like the cytosolic ribosomes, mitochondrial ribosomes (mitoribosomes) are composed of a large and small subunit, but are characteristically more protein rich (and thus have an unusually low RNA content) compared to other ribosomal structures (Rorbach et al., 2016).

In 2009 the Nobel Prize in Chemistry was awarded to Venkatraman Ramakrishnan, Thomas A. Steitz and Ada E. Yonath for their work on elucidating the structure of the bacterial ribosome. It was in 2003 when Rajendra K. Agrawal produced the first 3-dimensional structure of the (bovine) mitochondrial 55S ribosome using cryo-EM. (Sharma et al., 2003). Following this, two separate groups began the race to produce the best crystal structure of the mitochondrial ribosome. In a race between the laboratories of Venkatraman Ramakrishnan (Cambridge, UK) and Nenad Ban (Zurich, Switzerland), there were rapid advancements in the mitochondrial ribosome structure field, resulting in near-atomic resolution of the mammalian mitochondrial ribosome (Amunts et al., 2014, Brown et al., 2014, De Silva et al., 2015, Fernandez-Silva et al., 1997, Greber et al., 2014a, Greber et al., 2014b).

Translation of the mitochondrial transcripts initiates with the mRNA binding to the 28S complex. Translation elongation then proceeds by cycles of aminoacyl-tRNAs binding, peptide bond formation, and displacement of deacylated tRNAs.

1.7 MtDNA Replication

Another distinguishing feature of mtDNA is its unorthodox mechanism of replication. Unlike nuclear DNA, mtDNA continues to replicate throughout the lifetime of the cell. The mechanism of mtDNA replication is a hotly debated subject and there are two main hypotheses; the strand displacement model (SDM) and the RITOLS/bootlace model (RNA Incorporated ThroughOut the Lagging Strand). For extensive reviews see the following; (Holt and Jacobs, 2014, Holt and Reyes, 2012, Clayton, 1982).

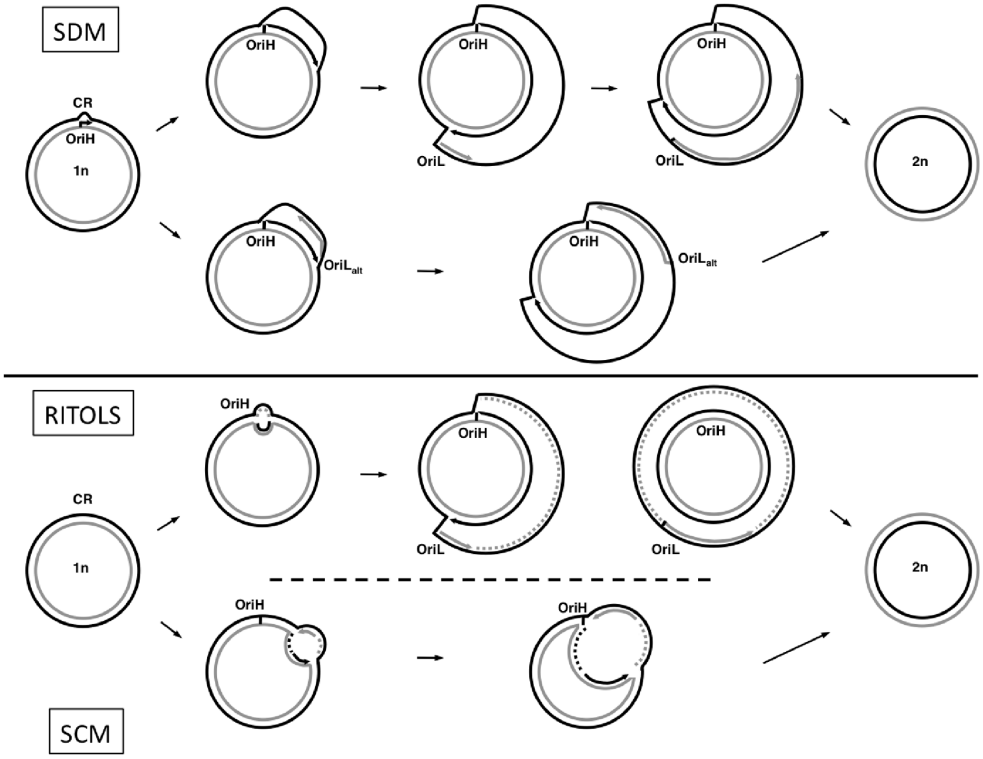


Figure 5: Mechanisms of mtDNA replication in vertebrate mitochondria:

[This image was originally published in PLOS One by Fonseca *et al.* and has been reproduced without modifications for this thesis]. The three predominant mechanisms of mtDNA replication; strand displacement model (SDM), Ribonucleotide Incorporation ThroughOut the Lagging Strand (RITOLS) and Strand-coupled mechanism (SCM). Where the solid lines denote DNA and the dashed lines represent RNA (Fonseca et al., 2014).

1.7.1 The Strand Displacement Model (SDM)

Both mechanisms are based on electron EM images from 1972 (Robberson et al., 1972) which show significant portions of single stranded DNA in actively replicating

mtDNA molecules. The two mechanisms agree that there are two dominant priming events, one on each strand. In the SDM model replication is initiated at the origin of heavy strand replication (O_H) in the major non-coding region (≈ 16,034 nucleotide position (nt) in mice). Leading strand synthesis proceeds for approximately 11 kb, displacing the lagging strand as it goes. This model proposes that the displaced lagging strand is coated with mitochondrial single-stranded binding protein (mtSSB) during this initial phase of replication, which is supported by Chip-Seq data from mtSSB pull-down (Miralles Fusté et al., 2014). Once the replication fork reaches a stem-loop structure, which defines the origin of light strand replication (Ori-L), DNA synthesis of the lagging strand can begin and two identical daughter molecules are formed. This model envisions continuous synthesis of both strands (Nass, 1980, Robberson et al., 1972, Clayton, 1982, Brown et al., 2005).

1.7.2 Ribonucleotide Incorporation Throughout the Lagging Strand (RITOLs)

In 2000, it was demonstrated that mtDNA replication intermediates (mtRIs) contained fully duplex structures based on 2D-AGE from highly purified organelles, which is discordant with the predicted single-stranded mtRIs of the previously described SDM (Yasukawa et al., 2006). These duplex structures were resistant to single-strand nuclease (following removal of any associated proteins by Proteinase K) which demonstrated that they could not be the mtSSB-bound intermediates of the SDM. Further interrogation of the mtRIs demonstrated that these structures contained significant portions of RNA/DNA hybrid. This has been attributed to the presence of long stretches of RNA hybridized to the lagging strand during leading strand synthesis (as opposed to mtSSB), the source of which is mitochondrial transcripts (mtRNA), also known as; bootlaces. This association of mtRNA with replicating DNA molecules has been corroborated by multiple methods including; sensitivity to Ribonuclease H1 (RNase H1), an enzyme which degrades the RNA portion of RNA/DNA hybrids, EM analysis and the capture of mtRIs with an antibody specific for RNA/DNA hybrids (Yasukawa et al., 2006, Reyes et al., 2013).

A common critique of this model is that the RNA species observed hybridised to the lagging strands of mtRIs is that this is merely an artefact, whereby transcripts within the mitochondria have bound to their complementary DNA sequence during

extraction. This was refuted by *in organello* pulse-chase experiments with both radioactive deoxynucleotide precursors and ribonucleotide precursors. This demonstrated that the observed hybrid species were in fact newly synthesised mtRIs which subsequently mature to full-length mtDNA molecules. Moreover, cross-linking of nucleic acids using psoralen prior to extraction dismissed the claim that the RITOLs intermediates were formed upon extraction.

An additional critique is the undeniable presence of vast amounts of mtSSB within the organelle, and therefore the question arises; if it is not bound to ssDNA during mtDNA replication then what is it doing? One possibility is that mtSSB is important for regulating and stabilising 7SDNA or RNA. In support of this replication-independent role for mtSSB, Mikhailov *et al.* showed that mtSSB is highly abundant in mtDNA preparations from *xenopus* oocytes, where mtDNA replication is not occurring (Mikhailov and Bogenhagen, 1996).

1.7.3 Ori-z and Strand-Coupled Bidirectional Replication

For many years, unidirectional, strand-asymmetric mode of replication was believed to be the sole primary mechanism of DNA replication in mammalian mitochondria. Using neutral/neutral 2D-AGE, however, it has been demonstrated that there are conventional double-stranded RIs in mitochondria indicative of coupled leading and lagging strand synthesis (θ replication), as well as the previously described strand-asymmetric mode (Holt et al., 2000). Initiation of this strand-coupled θ replication takes place downstream of O_H, outside of the NCR and there is evidence which suggests that θ replication is the predominant mechanism of DNA replication in mammalian mitochondria (Bowmaker et al., 2003).

Interestingly, attempts to locate and define an exact bidirectional origin of replication in mammalian mtDNA have proved difficult as overlapping fragments examined with 2D-AGE revealed what is described as an origin ‘zone’, rather than a discrete origin in both humans and mouse. The presence of different origins within different molecules has been reported to span across as much as 5 kB of the genome. Notably this window is smaller in mice than in humans (11,300- 15,500 and 10,500 – 16,000 respectively). This was demonstrated by revealing sets of overlapping fragments

associated with initiation arcs visualised by neutral/neutral 2D-AGE (Bowmaker et al., 2003).

In addition to this, cultured human cells which have undergone mtDNA depletion and recover show replication intermediates reminiscent of strand-coupled replication with initiation originating from within the NCR, distinct from O_H (Yasukawa et al., 2005). This second putative origin of replication, Ori-b, maps to approximately 15,600 nt in mice (Yasukawa et al., 2006). Not only does this provide additional evidence that mitochondria actively possess and use a conventional bidirectional origin and mode of replication, but also suggests that different modes of replication may be used under different circumstances.

1.8 Key Players in mtDNA Replication

1.8.1 Polymerase Gamma

Many of the proteins involved in DNA metabolism are shared between the nuclear and mitochondrial compartments, such as the primary mtDNA polymerase, Polymerase γ (Pol γ). For many years Pol γ was thought to be the sole polymerase in mitochondria, until recently where two additional polymerases have been reported in mitochondria; PrimPol and Polymerase β (Sykora et al., 2017, García-Gómez et al., 2013). (For a comprehensive view on mammalian mitochondrial polymerases, see (Krasich and Copeland, 2017)).

Pol γ is a heterotrimer, with a catalytic subunit encoded by the nuclear gene *POLG* and a dimeric accessory subunit encoded by *POLG2.* The holoenzyme has three activities; a 5'-3' DNA polymerase activity, a 3'-5' exonuclease activity and a 5'-dRP lyase activity associated with base excision repair. The dimeric accessory unit functions to help tight DNA binding and processivity (Lim et al., 1999). Its enzymatic activity and sensitivity to nucleotide analogues make it unique among classical eukaryotic polymerases (Ropp and Copeland, 1996).

Pol γ is essential for mtDNA replication and has a restrictive steric gate and highly specific active site which confers a high replication fidelity and low error rate, akin to other family A replicative polymerases (Copeland and Longley, 2003). Disruption to

the exonuclease domain of Pol γ leads to a dramatic increase in mtDNA mutation load (discussed in greater detail at the end of this chapter) and mutations in the POLG gene itself are associated with mitochondrial disease characterised by either deletions or depletion of mtDNA (Van Goethem et al., 2001). Although less common, mutations in POLG2 have been identified in cases of mitochondrial disease (Longley et al., 2006). Additionally, long term exposure to antiviral nucleoside analogue drugs (e.g. AZT) can lead to mitochondrial toxicity, and ultimately mitochondrial pathology due to Pol γ's heightened sensitivity to nucleotide analogues (Lewis and Dalakas, 1995).

Despite its high fidelity, Pol γ like almost all polymerases will (mis)incorporate ribonucleotides into DNA. Pol γ does in fact discriminate against ribonucleotides efficiently but with differential discrimination depending on the base identity. Subsequently, Pol γ is able to perform single-nucleotide reverse transcription from both DNA and RNA 3' termini although it's extension ability becomes significantly hindered by longer stretches of ribonucleotides (Kasiviswanathan and Copeland, 2011).

1.8.2 PrimPol

More recently, multiple studies have demonstrated the presence of an additional DNA polymerase in mammalian mitochondria. CCDC111, more commonly referred to as PrimPol, is a novel AEP-like (archaea and eukaryotes) polymerase-primase and is found in the nucleus, cytoplasm and mitochondria (García-Gómez et al., 2013).

PrimPol is unique in its ability to prime DNA synthesis. None of the other known human DNA polymerases are able to initiate DNA synthesis without an accompanying RNA polymerase. And not only is PrimPol able to initiate DNA synthesis using ribonucleotides, like a RNA primase, but PrimPol can utilise deoxyribonucleotides to initiate DNA synthesis (García-Gómez et al., 2013).

PrimPol is well documented as a translesion DNA synthesis (TLS) polymerase tailored to bypass the most common oxidative lesions in DNA, such as abasic sites and 8-oxoguanine (Mourón et al., 2013, Bianchi et al., 2013). Due to this ability,

however, PrimPol is error-prone with a preference to generate base insertions and deletions due to its lack of 3' to 5' exonuclease proofreading domain. It is believed that PrimPol's activity in the mitochondria is regulated by mtSSB to prevent the generation of indels (Guilliam et al., 2014) and is responsible for replication restart following damage and stalling or fork arrest. PrimPol knockout mice are viable but show impaired mtDNA replication, implicating it as a novel yet important player in mtDNA maintenance and metabolism (García-Gómez et al., 2013).

1.8.3 Mitochondrial Single Stranded Binding Protein

MtSSB is a heterotrimeric protein with a high affinity for single-stranded DNA. MtSSB is not like the nuclear single stranded binding protein (SSB) but more similar to eubacterial SSB, such as that from *E.coli* (Curth et al., 1994, Webster et al., 1997, Tiranti et al., 1993).

MtSSB is documented as an essential protein in many aspects of mtDNA replication and metabolism. MtSSB binds to single stranded DNA during replication to protect it from degradation by damaging agents such as free radicals within the mitochondria. MtSSB also interacts with other proteins within the mitochondria, such as Twinkle and multiple repair factors where mtSSB binding has a stimulatory effect (Korhonen et al., 2003). For example; in *Drosophila Melanogaster* when the mtSSB gene is disrupted there is mtDNA depletion and impaired respiration (Farr et al., 2004, Maier et al., 2001). MtSSB is also involved in the regulation of the D-loop whereby it plays an important role in the synthesis and abundance of 7S DNA (Ruhanen et al., 2010).

1.8.4 Twinkle

Twinkle is a hexameric helicase of the Rec-A superfamily and is partially homologous to the helicase domain of the T7 phage gp4 protein. Twinkle can catalyse 5' to 3' NTP-dependent unwinding of the mtDNA duplex *in vitro* and its activity is stimulated by mtSSB (Korhonen et al., 2003) supporting the idea that it is the mitochondrial replicative helicase. It is essential for mtDNA replication and is an important player in the reconstituted minimal replisome consisting of Pol γ, mtSSB and Twinkle which can effectively replicate the mitochondrial genome *in vitro* (Korhonen et al., 2004).

Knockdown via RNAi of Twinkle in human and insect cells causes mtDNA depletion and defects in Twinkle are associated with progressive external opthalmoplegia (PEO) with multiple mtDNA deletions in humans. Twinkle expression is directly linked to mtDNA copy number and over expression of Twinkle leads to an increase in copy number in transgenic mice (Tyynismaa et al., 2004).

Mice harbouring an autosomal dominant PEO (adPEO) mutation in Twinkle (a 13 amino-acid duplication in the linker region) accumulate large numbers of mtDNA deletions in the skeletal muscle and brain, and generally recapitulate the disease of the human patients, albeit significantly milder (Tyynismaa et al., 2005). The so-called 'deletor' mice do not exhibit a decreased lifespan however they do show characteristic histological features of late-onset mitochondrial myopathy, such as Cytochrome-*c*-oxidase (COX) negative muscle fibres (an indication of respiratory dysfunction). The model for deletion formation by adPEO mutations is Twinkle stalling or pausing during replication. Although the deletions form slowly *in vivo*, in young deletor mice there is already an accumulation of mtRIs indicative of stalling (Goffart et al., 2009). This is strong evidence to suggest that early problems with mtDNA replication results in the accumulation of deletions which can cause PEO.

1.8.5 Mitochondrial Genome Maintenance Exonuclease 1

Mitochondrial Genome Maintenance Exonuclease (MGME1) is a RecB-type exonuclease encoded by the orphan gene *C20orf72* in humans. Homozygous nonsense and missense mutations in the gene are characterised by common mitochondrial disease phenotypes such as PEO, mtDNA depletion and multiple mtDNA deletions in the muscle (Szczesny et al., 2013).

MGME1 cleaves single-stranded DNA and flap substrates and knockdown or complete knockout of the protein in both human cells and mice results in an accumulation of 7S DNA. It is generally accepted that MGME1 is required for effective mtDNA synthesis (Kornblum et al., 2013). RNase H1 is required during mtDNA replication for the removal of residual RNA primers. However, RNase H1 is unable to remove the last two ribonucleotides at the RNA-DNA junction and so it is believed that flap processing is invoked to efficiently removed these species. The

enzymes involved in this process in mitochondria are the helicase/nuclease DNA2, FEN1 and now MGME1 (Holt, 2009, Uhler et al., 2016).

Cells derived from patients with loss-of-function mutations exhibit an 11kB deleted linear mtDNA species which is also present in mice harbouring an exonuclease-deficient Pol γ (Uhler et al., 2016, Bailey et al., 2009, Nicholls et al., 2014). Moreover, *in vitro* reconstitution of MGEME1 activity shows that MGME1 is highly dependent on the 3'-5' exonuclease activity of Pol γ. This observation implicates flap processing and removal, and subsequent ligation in effective mtDNA replication and maintenance. However, it is yet to be proven that both linear species are formed via the same mechanism (Uhler et al., 2016).

1.8.6 Mitochondrial Topoisomerase I

Like a plasmid, mtDNA is mostly supercoiled and replication of the genetic information contained within requires relaxing of the circular duplex DNA, a process which relies on topoisomerase activity. Three putative topoisomerases have been identified in mitochondria (Top1mt, Top3α and Top2β), yet Top1mt is the most widely documented. Top1mt is a type 1B topoisomerase present in vertebrates and targeted exclusively to the mitochondrial compartment. Top1mt catalyses the transient cleavage and ligation of one strand of the mtDNA ahead of the replisome facilitating relaxation of the duplex supercoiled circular DNA. Despite this, Top1mt activity is not required for mitochondrial transcription (Zhang and Pommier, 2008).

MEFs from Top1mt deficient cells exhibit impaired mitochondrial function characterised by dysfunctional respiration, fatty acid oxidation, lipid peroxidation and induction of mitophagy. Interestingly, in cells lacking Top1mt there is an increase in ROS production and in the activation of DNA damage response pathways. For example; histone γH2AX, a marker for DSBs was elevated in Top1mt knockout MEFs (Douarre et al., 2012).

Studies which mapped the binding sites of Top1mt demonstrated an important role for Top1mt in the control region and suggested that Top1mt is involved in

degradation of specific mtDNA molecules (Dalla Rosa et al., 2014) and in D-loop maintenance (Zhang and Pommier, 2008).

1.8.7 Ribonuclease H1

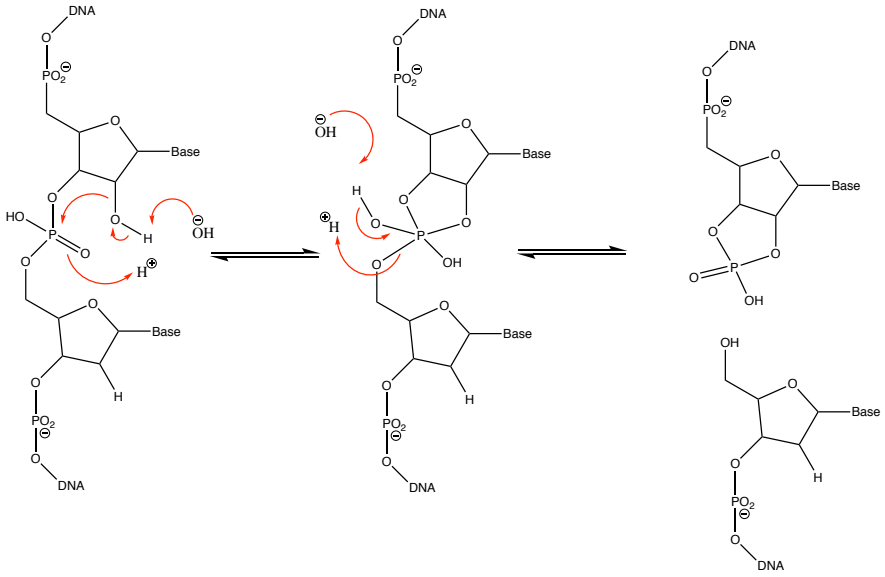
The mammalian RNase H1 is knockout-lethal. RNase H1 degrades the RNA portion of RNA/DNA hybrids (specifically stretches of RNA which are 4 consecutive ribonucleotide bases in length, or longer) in both the nuclear and mitochondrial compartments and it is essential for mtDNA replication (Cerritelli et al., 2003, Holmes et al., 2015). In mouse embryonic fibroblasts (MEFs) lacking RNase H1 there is ubiquitous primer retention at the origin(s) of replication which leads to replication stalling by Pol γ and subsequent double strand breaks (DSBs) at the origin(s) which results in mtDNA depletion and instability (Akman et al., 2016).

Mutations in the human *Rnaseh1* gene have recently been discovered which cause adult onset mitochondrial encephalomyopathy (Reyes et al., 2015). It is not clear what the mechanism of disease is but it may be linked to the role of the newly reported R-loop species and mtDNA segregation within the organelle (Akman et al., 2016). The human mutation and the mouse knockout model are strikingly different and this is likely because *Rnaseh1* ablation is incompatible with development beyond early embryogenesis (Cerritelli et al., 2003). This pathogenic mutation highlights the importance of RNase H1 in mtDNA metabolism and also the significance of the R-loop in mtDNA.

1.9 Ribonucleotides in DNA

Ribonucleotides are the most common non-canonical nucleotide incorporated into replicating DNA and all polymerases will (mis)incorporate ribonucleotides, but at different frequencies depending on their discrimination abilities. Ribonucleotides have an additional hydroxyl group compared to deoxyribonucleotides, and consequently embedded ribonucleotides in duplex DNA make the DNA more susceptible to damage via hydrolysis (Williams and Kunkel, 2014). In fact, a single ribonucleotide makes the DNA backbone 100,000 times more susceptible to hydrolysis. This is in part also due to the conformational changes associated with a distortion to the backbone when there is a single ribonucleotide present. The

distortion leads to puckering of the ribose sugar, making it more vulnerable to nucleophilic attack (Figure 6) (DeRose et al., 2012).



Scheme 1: Alkali hydrolysis of an embedded rNMP.

Nucleophilic attack of the ribonucleotide OH initiates the formation of a 2'3'-cyclic phosphate intermediate which is rapidly hydrolysed to the break the DNA backbone resulting in a 3'-phosphate and 5'-OH products via a SN2-like mechanism.

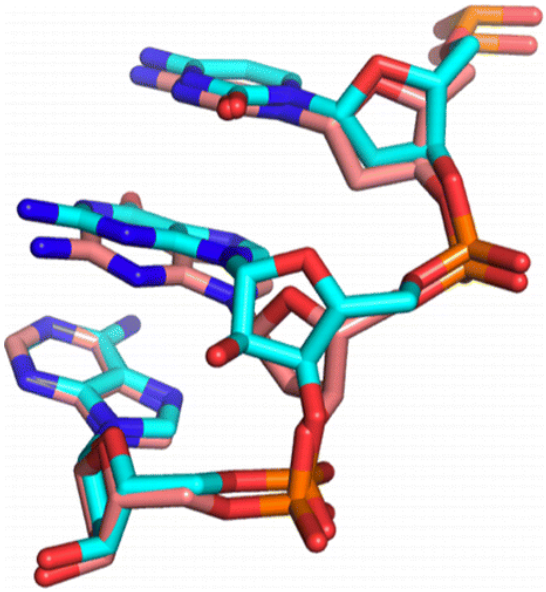


Figure 6: Solution Structure of the Dickerson DNA Dodecamer Containing a Single Ribonucleotide.

Reprinted with permission from DeRose *et al.* (DeRose et al., 2012). Copyright (2012) American Chemical Society.

1.9.1 Ribonuclease H2

Akin to RNase H1, Ribonuclease H2 (RNase H2) recognises and nicks duplex DNA at the sites of RNA/DNA hybrid, but specifically recognises single embedded ribonucleotides. This makes it an essential enzyme for maintaining genome stability and it is responsible for initiating ribonucleotide excision repair (RER) in the nucleus (Reijns et al., 2012, Sparks et al., 2012). Unlike RNase H1 which has a strong mitochondrial localisation pattern, RNase H2 has not been detected in mitochondria to date.

In humans, hypomorphic mutations in any of the genes which encode the three subunits of RNase H2 cause Aicardi-Goutières syndrome (AGS); a rare neuroinflammatory disorder. Interestingly, AGS can be caused my mutations in a variety of genes including *SAMHD1* which codes for a protein involved in nucleotide metabolism.

The RNase H2 knockout, like RNase H1, is embryonic lethal, but for different reasons. RNase H2 null mouse embryos accumulate ribonucleotides in their nuclear DNA which leads to genomic instability and a p53-mediated DNA damage response (Reijns et al., 2012, Nick McElhinny et al., 2010a). However, recent research indicates that the problems arise not from the persistent unrepaired ribonucleotides, but from aberrant RER processing in the absence of RNase H2, mediated by Topoisomerase I (TopI) (Huang et al., 2015, Kim et al., 2011, Huang et al., 2017). Top1 also possesses a ribonuclease activity and is able to recognise single ribonucleotides in duplex DNA. In the absence of RNase H2, Top1 binds to the embedded ribonucleotide, forming a cleavage complex at the site of ribonucleotide incorporation. The phosphotyrosyl bond is attacked by the 2'-OH and Top1 is released leaving 2',3'-cyclic phosphate ends (akin to those in iii). If there is a second Top1-cleavage complex adjacent to the nick, this enables dissociation of the short intervening DNA fragment (and excision of the ribonucleotide). Religation of the gap results in a small 2-5 base-pair deletion (Huang et al., 2015). Alternatively, following Top1 nicking additional cleavage on the opposite strand by Top1 can generate a double-strand break (DSB) (Huang et al., 2017). DSBs subsequently lead to genomic instability.

1.9.2 Persistent Ribonucleotides in mtDNA

Despite the vital importance of the mitochondrial genome and the considerable lack of non-coding regions, mtDNA is commonly reported to have a high mutation rate, as well as persistent (mis)incorporated ribonucleotides (Grossman et al., 1973).
Not only do ribonucleotides render the DNA more susceptible to hydrolysis, they also cause a distortion to the helix of the duplex structure (Figure 6)(Williams and Kunkel, 2014). It is unclear why ribonucleotides remain in the mitochondrial genome; whether it is due to a lack of repair mechanism or machinery, the redundancy of the multicopy genome or if there is a functional role for the ribonucleotides. Or a combination of two or more of these factors. There are numerous reports, a handful of which the reader is directed to, concerning the positive outcomes of retained ribonucleotides in DNA (Yao et al., 2013, Potenski and Klein, 2014, Dalgaard, 2012, Lujan et al., 2013).

1.9.3 Methods to Sequence Embedded Ribonucleotides

In 2015 multiple groups independently developed and published an array of next generation sequencing (NGS) approaches to identify the precise location of (mis)incorporated ribonucleotides in DNA (Figure 7) (Clausen et al., 2015, Ding et al., 2015, Koh et al., 2015, Keszthelyi et al., 2015). Of the four techniques detailed in Figure 7, three use alkali hydrolysis to fragment the DNA at the sites of ribonucleotides and one, EmRiboseq, uses RNase H2 *in vitro*. All techniques use the most advanced sequencing methods; either Illumina or Ion Torrent and claim to be able to provide single nucleotide resolution of embedded ribonucleotides.

Initially, at the time of their publication, all four techniques were focussed on the nuclear genome in yeast or bacteria. Moreover, in order to capture ample data for sequencing the library preparations were carried out in *Rnaseh2* null strains to increase the persistence of ribonucleotides within the DNA. These approaches proved to be valuable tools in identifying which DNA polymerases were acting on which strands/regions of DNA *in vivo* as well as locating origins of replication. It was only recently that attention moved to the mitochondrial genome. In 2017 Berglund *et al.* used the HydEn-seq method to locate and identify the sites of ribonucleotide incorporation in mtDNA from human cultured cells. The results indicated that there were base biases within the mtDNA, of which were attributable to the ratio of

nucleotide triphosphates (NTPs) to deoxyribonucleotides (dNTPs) within the mitochondrial compartment.

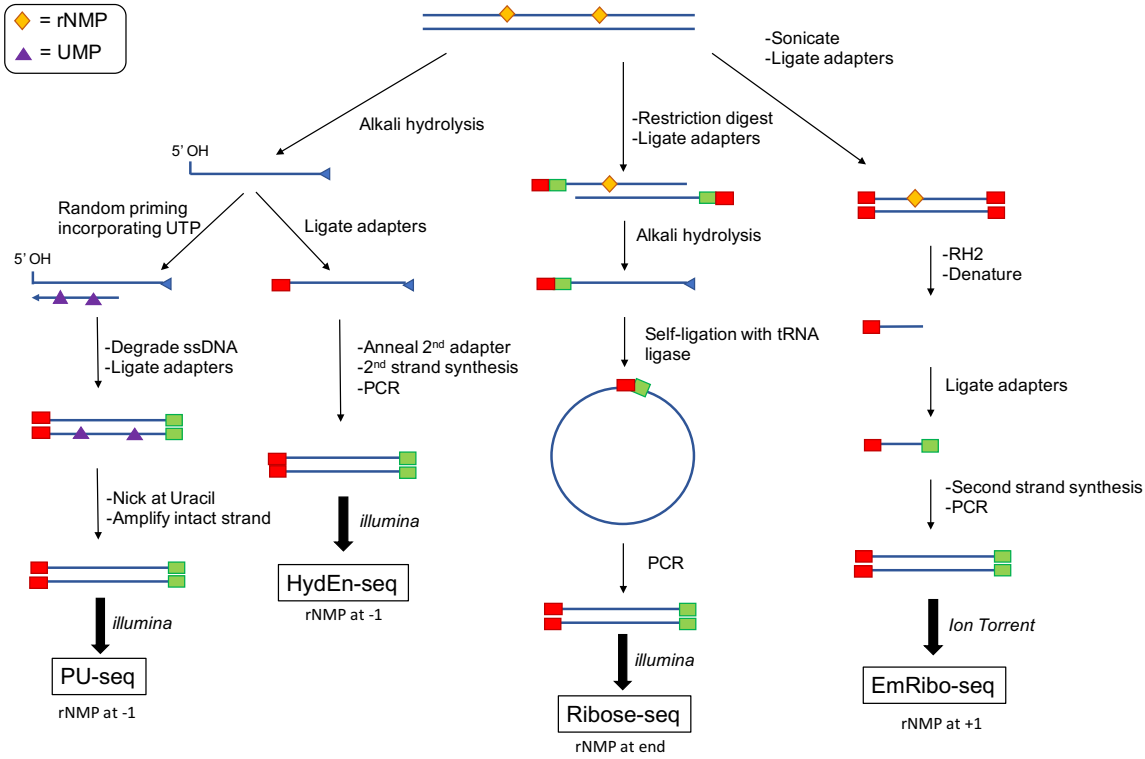


Figure 7: Techniques to sequence embedded ribonucleotides.
(Clausen et al., 2015, Ding et al., 2015, Koh et al., 2015, Keszthelyi et al., 2015)

1.10 Mitochondrial dNTP Metabolism

Mitochondria by nature require a distinct and separate mechanism to maintain sufficient dNTP pools for mtDNA synthesis in non-dividing cells. In proliferating cells, dNTPs are synthesised *de novo* in the cytosol and by two parallel salvage pathways acting in the cytosol and mitochondria. Deoxynucleosides are imported into the mitochondria and phosphorylated by Thymidine Kinase 2 (TK2) or Deoxyguanosine Kinase (DGUOK) (Figure 8). R2 and Thymidine Kinase 1 (TK1) are both only transiently present in proliferating cells modulated by the cell cycle. And when there is no longer a high demand for dNTPs for nuclear DNA synthesis, cytosolic dNTP synthesis is downregulated and the mitochondrial dNTP pools become largely reliant on the internal mitochondrial salvage pathway.

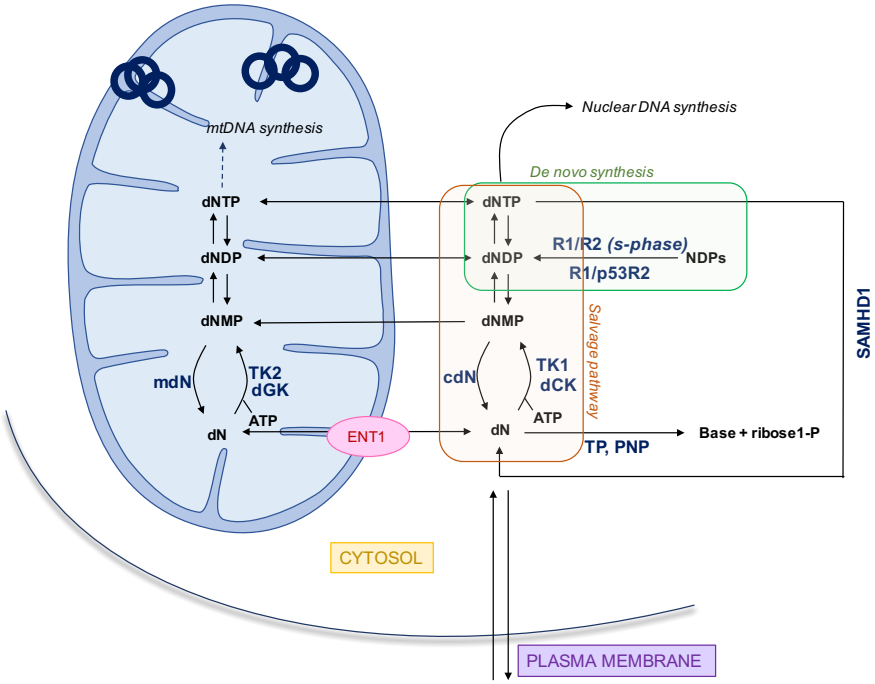


Figure 8: dNTP synthesis in actively dividing mammalian cells
Adapted from Rampazzo *et al.* (Rampazzo et al., 2010)

1.10.1 Thymidine Kinase 2

As aforementioned, TK2 is one of two deoxynucleoside kinases in the mitochondria, responsible for the first, and rate-limiting, phosphorylation step in the mitochondrial salvage pathway. TK2 specifically phosphorylates the pyrimidines thymidine, deoxycytidine and deoxyuridine (Roos et al., 2014, Saada‐Reisch, 2004). As many as 30 different mutations in the *TK2* gene are associated with mtDNA depletion syndromes (MDS). This is believed to be due to reduced activity of TK2, although the tissue-specific nature of the resulting phenotype is yet to be explained (Saada et al., 2001, Frangini et al., 2013). In addition to MDS, mutations in *TK2* have also been linked to PEO associated with multiple mtDNA deletions (Tyynismaa et al., 2012, Alston et al., 2013).

1.10.2 Deoxyguanosine Kinase

DGUOK is a deoxynucleoside kinase located within the mitochondria, responsible for phosphorylating the purine deoxyribonucleosides. Like TK2, autosomal recessive mutations in the *DGUOK* gene result in early-onset and often fatal MDS. The precise

mechanism of pathology is not fully understood but it is believed that the mutations result ultimately in imbalanced, depleted mitochondrial dNTP pools which impedes mtDNA replication. Additionally, there are documented cases of a missense mutation in *DGUOK* which leads to reduced activity and an accumulation of mtDNA deletions due to a defect in DGUOK activity (Ronchi et al., 2012).

1.10.3 Ribonucleotide Reductase

Ribonucleotide reductase (RR) is an enzyme which catalyses the reduction of ribonucleotide 5'-diphosphates to their corresponding 2'- deoxynucleotide. It is comprised of one large subunit RRM1 and two smaller subunits RRM2 and p53R2. p53R2 refers to the small subunit of RR and as the name suggests, has a p53 binding site and plays an important role in tumour suppression and DNA damage response. p53R2 is critical in dNTP synthesis outside of the S-phase and therefore plays a large role in providing dNTPs for DNA repair.

Mutations or complete ablation of p53R2 is associated with significant mtDNA depletion (Bourdon et al., 2007, Kollberg et al., 2009) and mtDNA deletions causing PEO (Tyynismaa et al., 2009). However, it is still not clear whether p53R2 affects mtDNA metabolism via perturbations in dNTP pools or in another manner as it has been shown that p53R2 positively correlates with mtDNA content, and even increasing p53R2 expression reduces ROS and protects mitochondrial membrane potential in human cancer cells (Wang et al., 2011). Despite this, no appreciable RR activity was detected in mitochondria, in contrast to earlier reports (Young et al., 1994).

1.10.4 MPV17

MPV17 is an inner mitochondrial membrane protein coded for by the *MPV17* gene located on chromosome 2p21-23 and it has been identified as a mutant gene associated with MDS (Spinazzola et al., 2006). Patients with mutations in *MPV17* exhibit numerous symptoms including liver failure and neurological impairments. *MPV17* mutations have also been associated with adult-onset neuropathy and leukoencephalopathy due to the accumulation of mtDNA deletions in skeletal muscle (Blakely et al., 2012).

The function of the protein is largely unknown but the human *MPV17* is the orthologue of the mouse kidney disease gene, *Mpv17*. The mouse MPV17 knockout is viable and as anticipated, exhibits late-onset proteinuric nephropathy as well as early-onset liver mtDNA depletion. These symptoms were associated with a significantly shorter lifespan (Dalla Rosa et al., 2016, Spinazzola et al., 2006, Viscomi et al., 2009).

It is likely that MPV17, absence of which exhibits similar features to the DGUOK knockout, is involved in mitochondrial dNTP metabolism. Both mouse liver mitochondria and quiescent patient cells have significantly reduced dNTP pools and mtDNA copy number, the latter of which is rescued by deoxynucleoside supplementation in the patient fibroblasts, thus indicating that the dNTP insufficiency associated with absence, or near-absence, of the protein in the mitochondria is the cause of the mtDNA depletion (Dalla Rosa et al., 2016).

S. *cerevisiae* express a protein, SYM1, which is the genetic orthologue of the mammalian *MPV17* gene. Akin to the MPV17 knockout models, the *SYM1* gene product is essential for effective OXPHOS, mitochondrial morphology and mtDNA stability under high-temperature and ethanol-dependent growth. Either deletion or disruption of the *SYM1* gene leads to an accumulation of mitochondrial respiratory-deficient ‘petite’ mutants (Dallabona et al., 2010).

1.10.5 dNTP Import

The mitochondrial and cytosolic dNTP pools are separated by the IMM which is impermeable to charged molecules, meaning that dNTPs have to be actively imported by dedicated transporters. dNTPs produced by the nucleotide *de novo* synthesis can be imported from the cytosol by deoxynucleotide transporters (PCN1 and PCN2). Additionally, the respective deoxynucleosides can enter the mitochondrial matrix through the Equilibrative Nucleoside Transporter 1 (ENT1), to then be phosphorylated by kinases of the mitochondrial salvage pathway (Figure 8). It is likely that there are additional nucleoside transporters, however their identity has so far eluded the field.

1.11 MtDNA Repair

For an extensive review please see (Kazak et al., 2012).

1.11.1 Mitochondria and Oxidative Damage

Superoxide, produced in the mitochondria as a by-product of OXPHOS (Figure 2), is quickly converted to H_2O_2 by SOD1 or 2, which does not react with DNA. However, via Fenton Chemistry (below) H_2O_2 can be converted to the highly reactive hydroxyl radical ($HO^{\bullet}$) which is known to cause DNA damage.

$$Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + HO^{\bullet} + OH^{-}$$

Attack on DNA by ROS can generate a plethora of DNA adducts and base modifications, of which 8-hydroxydeoxyguanine (8-oxodG) is one of the most abundant (Ames, 1989). What is more, because mitochondria are the source of the majority of cellular ROS, it has been observed that mtDNA contains a higher amount of steady-state amount of oxidative damage than nuclear DNA (Richter, 1995). 8-oxodG is particularly problematic because it is mutagenic due to its ability to promote mispairing. 8-oxodG is readily misincorporated opposite adenine (Pavlov et al., 1994), and so to avoid this, mitochondria possess a mitochondrial homologue of an 8-oxo-dGTPase, encoded by the *MTH1* gene, which hydrolyses free 8-oxodG moieties in the mitochondria before it can be incorporated into mtDNA (Kang et al., 1995).

However, when and if 8-oxodG is incorporated into mtDNA, the predominant pathway for repair is the mitochondrial base excision repair (BER) pathway which is initiated by oxoguanine DNA glycosylase (OGG1) and mice deficient for ogg1, have a 20-fold increase in 8-oxodG in the liver mtDNA compared to wild-type mice (de Souza-Pinto et al., 2001).

1.11.2 Base Excision Repair in Mitochondria

As eluded to above, mammalian mitochondria possess an active BER pathway, which is perhaps the best characterised repair pathway within the organelle. BER is

used to repair non-bulky lesions such as oxidised bases and abasic sites (Pinz and Bogenhagen, 1998, Szczesny et al., 2008). Like the nuclear BER pathway, the mechanism involves excision of the damaged base, strand cleavage, DNA end-processing, gap filling and ligation. BER can be divided into short-patch (SP) BER or long-patch (LP) BER. The former is utilised when nucleotide excision leaves a 3'-OH and a 5'-phosphate group for ligation. However, if a 5'- end is generated which is not conducive to ligation, LP-BER is used. Both have been reported in mammalian mitochondria, and are facilitated by membrane association of mtDNA (Boesch et al., 2010, Szczesny et al., 2008, Robertson et al., 2009).

Effective BER relies on efficient gap filling by a DNA polymerase. In the nucleus, sugar removal and gap filling is carried out by Pol β, which until recently was thought to be absent from mitochondria (Krasich and Copeland, 2017, Sykora et al., 2017). Nonetheless, Pol γ has been shown to possess intrinsic dRP lyase activity and can convert 5' dRP residues to a ligatable 5'-phosphate (Longley et al., 1998).

1.11.3 Single-strand Break Repair

Repair of single-strand breaks (SSBs) in mitochondria is largely similar to BER following detection of the SSB. The poly(ADP ribose) polymerase (PARP) family of proteins are responsible for detection in the nucleus, and PARP1 has been detected in mitochondria (Rossi et al., 2009).

SSBs can be caused by abortive DNA ligase activity which leaves a residual 5'AMP. The DNA strand-break repair protein aprataxin has been found to localise to mitochondria in human calls where it removes 5'-adenylate groups from DNA that arise from aborted ligation reactions. Depletion of aprataxin in human SH-SY5Y neuroblastoma cells and primary skeletal muscle myoblasts results in mitochondrial dysfunction, suggesting a direct role for the enzyme in mtDNA repair (Sykora et al., 2011).

1.11.4 Homologous Recombination in Mitochondria

Double strand breaks (DSBs) are much more problematic and present a greater potential problem to the cell, both in the nuclear and mitochondrial compartments. In

the nucleus, there are three predominant mechanisms for DSB repair; non-homologous end-joining (NHEJ), microhomology end-joining (MMEJ) and homologous recombination (HR).

Because the mitochondria possess a multicopy genome, turnover of mtDNA molecules is in itself a mechanism of damage processing; a process which cannot be afforded to the nuclear DNA. And thus, the mitochondria do not possess the same wealth of repair processes as the nucleus. Despite this, there is indirect evidence to suggest that mitochondria are able to repair DSBs (Bacman et al., 2009) as well as the argument that mammalian mitochondria possess all the necessary machinery for HR activity (Thyagarajan et al., 1996).

1.12 Mitochondrial Disease

Mitochondrial disorders are caused by dysfunctional mitochondria and can be triggered by either mutations within the mtDNA itself or in nuclear genes which code for mitochondrial proteins. Being the core of energy production in cells, the number of mitochondria varies between cell types, often relating to the energy demand of the tissues. For example, there is a high concentration of mitochondria in muscle fibres. As a consequence, mitochondrial disorders often manifest in high-energy demand tissues and common pathologies include blindness, deafness, neurological impairment, muscle weakness/lack of muscle co-ordination and heart disease. The term mitochondrial disease encompasses an extensive spectrum of disorders, many of which are still poorly understood.

Mitochondrial diseases are broadly divided into two factions; diseases which result from or exhibit mtDNA depletion (mtDNA depletion syndrome (MDS)), or mtDNA deletions (multiple deletions). MDS are typically early-onset and are often fatal. By contrast, multiple deletions tend to manifest in adults and the pathology is often less severe.

1.12.1 Diagnosing Mitochondrial Disease

Ragged red fibers are important biomarker for mitochondrial disease, in particular mitochondrial myopathies, but have also been identified in normal ageing (Rifai et al., 1995). The term refers to the appearance of a muscle biopsy which has been stained with Gomori trichrome and visualised under the microscope. Mitochondrial dysfunction in some instances can result in an accumulation of mitochondria located in the sub sarcolemma, that appear red once stained. This phenotype is often accompanied by muscle weakness and external opthalmoplegia.

Another characteristic feature of mitochondrial dysfunction is the presence of cytochrome oxidase (COX) deficient fibers (stained blue), which are indicative of respiratory chain deficiency (Johnson et al., 1983). They are identified by histochemical assay using a sequential cytochrome *c* oxidase-succinate dehydrogenase reaction, and are often found in a mosaic pattern within the muscle fibers (Murphy et al., 2012).

The above mentioned histochemical tools are extremely common assessments of mitochondrial disease alongside clinical assessments and physiological tests to assess oxidative capacity of patients. For a review of diagnosis of mitochondrial disease please refer to (Taylor et al., 2004).

1.12.2 MtDNA Deletions

In 1988 Holt *et al.* reported the presence of a sub-population of mtDNA molecules containing a deletion, alongside the wild-type molecules; a form of heteroplasmy. Said deletions were observed in biopsies from skeletal muscle but not blood from patients exhibiting mitochondrial disease (Holt et al., 1988). It is now well established that these deletions are pathogenic and contribute to the mitochondrial dysfunction (Pitceathly et al., 2012).

1.12.3 Disease-Causing Mutations in mtDNA

As aforementioned, the presence of mtDNA mutations can appear in just a fraction of the total mtDNA molecules in a cell, or in all molecules resulting in either heteroplasmy or homoplasmy. In the case of heteroplasmy there exists a sub-

population of mtDNA molecules carrying the mutation, which if increased past a certain, often variable, limit, causes the disease to manifest. When the mutation in question is a non-synonymous substitution, this is referred to as the threshold of disease. There is evidence from cell models that certain cell lines carrying mtDNA mutations will spontaneously segregate to homoplasmy; both 100% mutant or 100% wild-type (Dunbar et al., 1995, Holt et al., 1997). The mechanisms for segregation are still extremely unclear but are an important avenue for potential future therapeutic treatments.

The majority of diseases caused by mtDNA mutations are mutations contained within mitochondrial tRNAs. This includes the most common mtDNA mutation; A3245G in the $tRNA^{Leu}$, which causes MELAS, a condition which affects predominantly the brain and muscles and is often fatal. Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with another mtDNA $tRNA^{Lys}$ mutation at 8344 nt (Shoffner et al., 1990). All mitochondrial diseases caused by heteroplasmic mtDNA mutations are subject to mitochondrial inheritance down the maternal line, and thus predicting the severity of the resulting disease is extremely difficult.

1.12.4 Treating Mitochondrial Disease

Due to the plethora of multifactorial processes affecting the pathology of mitochondrial disease, the treatment of mitochondrial diseases is complex and specific for each individual case. Traditionally mitochondrial disorders have been treated with vitamins and supplements, with little proven benefit. However, more recently there are emerging therapies which show promise (for an in-depth review see (Nightingale et al., 2016)). One such method is the use of zinc finger nucleases or transcription activator-like effector nucleases (TALENs) directed to degrade specific mtDNA mutated molecules (Bacman et al., 2013, Minczuk et al., 2008, Reddy et al., 2015).

Another novel approach, which has attracted widespread media attention is the use of pronuclear transfer during *in vitro* fertilisation to prevent inheritance of mtDNA disease. This technique aims to produce a fertilised zygote where the maternal mitochondria have been replaced by healthy mitochondria from a donor egg.

Although promising, preliminary data suggests that it is not a guaranteed method of prevention (Hyslop et al., 2016).

1.13 The Role of mtDNA in Ageing

1.13.1 The Free Radical Theory of Ageing

There are numerous hypotheses regarding ageing and the role of mitochondria in the ageing process. The most famous theory is the free-radical theory of ageing, based on observations as long as 50 years ago that radiation of living things, and thus the subsequent production of free-radicals, shortens their lifespan (Harman, 1956). The theory, and its subsequent developments (Harman, 2009) postulates that animals age due to the accumulation of free-radical damage over time, and more specifically that the mitochondria are the hub for this process, producing high levels of free radicals and ROS. Despite its appeal this theory has many pitfalls. For example; experiments which have used genetic manipulations to increase antioxidant production have had no effect on lifespan and numerous species such as birds, bats and mole rats have a long lifespan along with high levels of oxidative damage even at young ages (Buffenstein et al., 2008). Although this theory still holds some weight, it is generally accepted that this is not the sole explanation for the ageing phenomenon (Muller et al., 2007).

The mtDNA continues to replicate throughout life unlike the nuclear DNA and consequently the mtDNA accumulates mutations and deletions as we grow old (Cortopassi et al., 1992). It is well established that somatic mtDNA mutations in human aging undergo clonal expansion and cause a mosaic respiratory chain dysfunction in different tissues (Krishnan et al., 2007, Larsson, 2010). It has generally been assumed that age-associated somatic mtDNA mutations are caused by accumulated damage during the aging process (Wallace, 2001). However, an alternative hypothesis is that most of the mutations are created as replication errors during embryogenesis and then undergo clonal expansion in adult life (Park and Larsson, 2011).

1.13.2 The Mutator Mouse

In 2004 and 2005 two separate groups developed a mouse model to test their hypothesis that the accumulation of mtDNA mutations is related to the ageing phenomenon (Trifunovic et al., 2004, Kujoth et al., 2007). The so-called 'mutator-mouse' carries a mutation in the PolgA sub-unit of the tDNA polymerase: Pol γ. The mutation in the exonucleolytic domain of PolgA cripples the proofreading capacity of the enzyme. The result is a dramatic increase in the number of mtDNA point mutations than seen in age-matched controls (although the extent of this increase is still under debate) and an accumulation of so-called linear deletions. The adult mutator mice exhibit a striking ageing phenotype. Although this work indicates that high levels of point mutations can cause a premature ageing phenotype this is very far from demonstrating that such point mutations are a significant contributor to ageing, not least because the mutation levels in the homozygous mice vastly exceed that observed in naturally aged mice or humans. This suggests that mtDNA mutations may play a role in ageing but are certainly not the sole cause.

1.14 Aims and Hypotheses

The aim of this PhD was to examine the phenomenon of ribonucleotide incorporation and retention within mammalian mtDNA; something of which is still poorly understood. It was intended to investigate ribonucleotide incorporation to better understand mtDNA synthesis, turnover and repair. After developing a mitochondrial-specific NGS approach to analyse ribonucleotides in mtDNA it was invoked as a tool to assess ribonucleotide incorporation in the context of mitochondrial disease and dysfunction.

Early hypotheses focused on ribonucleotides being increasingly incorporated as a function of time – ie a function of age. And that the ribonucleotides embedded within the mtDNA molecule could identify sites of replication initiation or pausing. As the project developed the former hypothesis was refuted and the latter was complicated by the high frequency of ribonucleotide retention.

Focus instead turned to the factors influencing ribonucleotide incorporation and the predominant hypothesis became the effect of changing mitochondrial nucleotide pools on ribonucleotide incorporation frequency. As it stands, the current hypothesis is that ribonucleotide incorporation is directly affected by the ratio of dNTPs to NTPs within the vicinity of the replisome. Moreover, that ribonucleotide incorporation beyond a certain threshold can become deleterious for the mitochondria.

Chapter 2. Materials & Methods

2.1 Primer and oligo sequences

nt (strand)	Sequence 5'-3'
300 (L)	ATTTCGTGCCAGCCACCGCGGTCATACGAT
T7 1000 (H)	TAATACGACTCACTATAGGGGTATGCTT
T7 14881 (L)	TAATACGACTCACTATAGGTCCCAGACATACTAGGAGACCCA
15800 (H)	AAGAACCAGATGTCTGATAAAGTTTC
15490 (H)	CTTGGGGAAAATAGTTTAATGTACG
T7 15750 (L)	TAATACGACTCACTATAGGATTAAACTTGGGGGTAGCTAAAC
16299 (H)	TTGTTAATGTTTATTGCGTAATAGAGT
340 (H)	GTTTGGGTTAATCGTATGACCG
trP1-top	CCTCTCTATGGGCAGTCGGTGAT-phosphorothioate-T
trp1-bottom	Phosphate-ATCACCGACTGCCCATAGAGAGG-dideoxyC
A-top	Phosphate-CTGAGTCGGAGACACGCAGGGATGAGATG-dideoxyG
A-bottom	Biotin-CCATCTCATCCCTGCGTGTCTCCGACTCAGNNNNNN-C3-phosphoramidite
Primer A	CCATCTCATCCCTGCGTGTCTCCGAC
Primer trP1	CCTCTCTATGGGCAGTCGGTGATT

Table 1: Primer sequences used to generate PCR products for riboprobe and DNA probe synthesis.

The T7 promoter sequence is underlined where applicable. Note that the phosphorothioate is used as an oligonucleotide modification to prevent nucleolytic degradation during library preparation (Shin et al., 2014).

2.2 Cell Culture

2.2.1 MEFs

Both primary and immortalized fibroblasts were cultured in Dulbecco's Modified Eagle's Medium (DMEM, Life Technologies) supplemented with 10% foetal bovine serum (FBS, Hyclone), 1% penicillin and streptomycin (PS, Life Technologies) at 37°C in a 5% CO_2 atmosphere.

2.2.2 Conditional RNase H1 Knockout MEFs

RNase H1 immortalised MEFs were grown under identical conditions as the wild-type MEFs. Conditional gene knockout was via the action of cre-recombinase on LoxP sites flanking exons V–VII of the Rnaseh1 gene using the drug Tamoxifen –

100 nM final concentration. Cells were typically harvested between 7 and 10 days of gene knockout.

2.2.3 Induction of Quiescence by Serum Starvation

Non-immortalised MEFs deficient for P53 were grown under standard conditions as detailed above. Quiescence was induced at ~ 80% confluence by serum starvation – a reduction from 10% to 0.1% FBS (dialysed serum). Quiescence was confirmed by arrest of growth and cell division and gross physical changes as observed by light microscopy. Cells under serum starvation were washed with phosphate-buffered saline (PBS) and the media changed daily.

2.3 Animals and Genotyping

CFW-Mpv17/J strain from Jackson laboratory (original stock number 002208) was backcrossed to C57Bl/6J background using the Marker-Assisted Accelerated Backcrossing (MAX-BAX®) provided by Charles River to generate 100% of C57Bl/6J within 10 generations. The congenic strain generated is identified as B6.CFW-Mpv17/J. Animals were genotyped by PCR, as described by Dalla Rosa *et al.* (Dalla Rosa et al., 2016).

2.4 Mitochondria Isolation from Mouse Liver

Fresh mouse livers (BL6/CD45.1) were immediately placed in cold phosphate-buffered saline on ice and added to a tenth volume of homogenization buffer (HB); 225 mM mannitol, 75 mM sucrose, 10 mM HEPES-NaOH, pH 7.8, 10 mM EDTA, 0.1% (w/v) fatty acid-free bovine serum albumin (BSA) and 357.5 μM β-mercaptoethanol. The tissue was washed with 5 volumes of HB and then added to 9 volumes of HB and homogenized using a motorized tight-fitting Dounce homogenizer. The homogenate was centrifuged at 600gmax for 10 min at 4ºC to pellet nuclei and intact cells then the homogenate/supernatant was collected and then centrifuged at 5,000 *gmax* for 10 min at 4ºC to pellet the mitochondria. The mitochondrial pellet was resuspended in 2 mL of HB and added to the sucrose gradient. The single-step sucrose gradient was prepared in a 25 x 89 mm tube with 17.5 mL of 1.5 M Sucrose (in HB) overlaid with an equal volume of 1 M sucrose. The sucrose gradients were centrifuged in a swing-out rotor at 40,000 *gmax* for 1h at 4ºC

after which the mitochondria resolve at the interface of the 1M and 1.5 M sucrose solutions.

The mitochondria were collected from the 1 M:1.5 M sucrose interface and transferred to a 15 mL falcon tube and added to 5 volumes of gradient buffer; 10 mM HEPES-NaOH, pH 7.8, 10 mM EDTA. The solution was mixed gently and centrifuged at 9,900 *gmax* for 10 min at 4ºC. The pellet was weighed and resuspended in 1.6 mL/g lysis buffer (LB); 20 mM HEPES-NaOH, pH 7.8, 75 mM NaCl, 50 mM EDTA and incubated with 10 µL Proteinase K (20 mg/mL, based on an average preparation of 10 adult mouse livers) on ice. After 10 min sodium dodecyl sulfate (SDS) was added to a final concentration of 0.25% and incubated at room temperature for 50 minutes. MtDNA was then isolated by Phenol-Chloroform extraction, described below.

2.5 Mitochondria Isolation from Mouse Embryonic Fibroblasts

40 X 20 cm^2 dishes were harvested at 80-90% confluency. Cells were washed with Phosphate-buffered saline (PBS) and detached from the plates with trypsin before being collected and centrifuged at 1,100 rpm for 5 minutes. The cell pellet was washed with PBS twice and weighed and then resuspended in 7.2 mL/g hypotonic buffer; 200 mM HEPES pH 8, 50 mM KCl, 15 mM $MgCl_2$. The suspension was left on ice for 10 min then homogenized with a tight fitting Dounce homogenizer with 20 strokes. After homogenization, the protocol is identical to that of the liver mitochondria isolation (see above).

2.6 Phenol-Chloroform DNA Extraction

An equal volume of phenol was added to the lysate and mixed thoroughly for 5 min then centrifuged at 5,000 *gmax* for 5 minutes. The upper aqueous phase was removed and added to a clean falcon tube and an equal volume of chloroform: isoamyl-alcohol 24:1, mixed and then centrifuged at 5,000 *gmax* for 5 minutes. The upper aqueous phase was removed and added to a clean falcon tube. 2 volumes of absolute ethanol were added and the solution was mixed and left on ice for 15 minutes before centrifuging at 12,000 *gmax* for 30 minutes. The supernatant was discarded then the pellet washed with 70% ethanol. All ethanol was removed and

the pellet was air dried and resuspended in tris-EDTA (TE) and the nucleic acid concentration determined by UV spectrometry.

2.7 Total DNA Extraction from Mouse Tissues

Fresh tissue was immediately placed in LB and homogenized with either a motorised tight-fitting Dounce homogenizer (liver and brain) or with a mechanical Ultra-Turrax high-speed homogenizer (heart and eyes (following removal of the optical lens)). Proteinase K was added (1 mg for brain, heart and eyes and 2 mg for liver) along with SDS to a final concentration of 0.25%. The tissue was incubated at 37 ºC on a shaker at 225 rpm for 2-3 hours. The DNA was extracted using the Phenol-Chloroform procedure outlined above.

2.8 Total DNA Extraction from Cell Culture

LB was added directly to the 90% confluent plate of cells and then left at room temperature for 10 minutes before collecting. 5 µL of Proteinase K (Sigma Aldrich, 20 mg/mL) was added per ~20 million cells along with SDS to a final concentration of 0.25% and incubated at 37 ºC for 2 hours. The DNA was extracted using the Phenol-Chloroform procedure outlined above.

2.9 PCR

PCR was performed with typically 100 ng of mtDNA or 500 ng total DNA using Ex Taq DNA polymerase according to the manufacturer's instructions (Takara).

2.10 MtDNA Copy Number Determination Using qPCR

MtDNA copy number was determined from total nucleic acid isolated from either cells or tissues. For each sample two PCR reactions were made with a) Nuclear primers (APP) or b) Mitochondrial primers (COXII). Each 25 µL reaction contained 25 ng total nucleic acid, 12.5 µL SYBR Green PCR master mix (applied Biosystems), 1.25 µL forward/reverse primer (30 µM APP, 10 µM COXII) and nuclease free water (NFW). The reactions were carried out in a 96-well plate and carried out according to (Schmittgen & Livak 2008) x. Changes in mtDNA amount were calculated using the $2^{-\Delta\Delta Ct}$ method and represented as fold changes relative to the indicated control.

2.11 Treatment of mtDNA with Nucleic Acid Modifying Enzymes

Typically, 500 ng – 1 µg of mtDNA or 5-6 µg of total DNA was used for each lane of a 1-Dimensional Agarose Gel Electrophoresis (1D-AGE). Restriction digestions were carried out in 100 µL of the appropriate reaction buffer with 10 units of enzyme according to the manufacturer's instructions (NEB). Restriction digests were incubated at 37 ºC for 2 hours and stopped by precipitation in ethanol (10 µL 3M NaAc and 220 µL absolute ethanol, kept at -20 ºC for at least 30 minutes and centrifuged at 12,000*gmax* for 30 minutes before air drying the pellet and re-suspending in TE).

RNase A (Qiagen) digestion was typically carried out at 37 ºC for 30 minutes under salt concentrations less than 100 mM to promote DNA-RNA hybrid hydrolysis activity as at NaCl concentrations higher than 0.3 M RNase A specifically cleaves single-stranded RNA.

Eco-RNase HI (Takara) is an RNA-specific endonuclease that hydrolyses the RNA portion of RNA-DNA hybrids. Typically, 1 µg of mtDNA was treated with 1 unit of RNase HI at 37 ºC for 90 minutes. One unit is the amount of the enzyme that produces 1 nmol of acid-soluble 3H in 20 minutes at 30℃ and pH7.7, with poly (rA)・poly (dT) as the substrate.

The single strand endonuclease used was Mung Bean nuclease (MBN) (Takara) which hydrolyses single-stranded DNA and RNA into 5'-phosphate and 3'-hydroxyl-containing products. MBN was used in the appropriate buffer with varying concentrations and amounts of time at 37 ºC.

Any sequential treatments were treated with either Eco-RNase HI or Eco-RNase HII (NEB) and then re-precipitated in ethanol, re-suspended and then treated with MBN. Reactions involving Eco-RNase HI or MBN were terminated with addition of SDS as per manufacturers guidelines.

2.12 1D-AGE

Where indicated DNA samples were denatured at 90 ˚C for 5 min then immediately put on ice before loading onto the gel. All gels were set with either 150 mL 1 X Tris-borate EDTA (TBE) and run with 1 X TBE running buffer or 1 X Tris-base, Acetic acid, EDTA (TAE). Typical run conditions were 250 mA for 4 hours, 0.8-1% agarose (UltraPureTM Agarose, Invitrogen).

2.13 Neutral-Neutral 2D-AGE

Prior to electrophoresis, mtDNA was treated with specific restriction enzyme(s) and precipitated with ethanol to create a desired fragment of interest as detailed above. Typically fragments sized 3-6 kb were run overnight at 20V, 0.4% agarose. Individual lanes were then excised and rotated 90°. The second dimension is run at 4 ºC, 1% agarose with 1 μL/100 mL EtBr, 280 mA for approximately 4hrs, or until sufficient separation has been achieved, which is visualised under UV light. Prior to Southern Blot (see below) the gel is flipped and washed in 25mM HCl for 15 minutes, before 2 X 20 minutes in 400 mM NaOH.

2.14 Southern Blot

1D-AGE: After electrophoresis, the gel was removed from the tank and inverted into a glass dish. If the DNA samples were denatured the gel was washed in 10 x SSC only. If the DNA samples were non-denatured the gel was washed in 400 mM NaOH for 15 min twice before removing all excess solution. Nylon membrane was cut to the shape of the gel and soaked in water briefly before being placed upon the gel. 5 sheets of stacked 3MM Whatman filter paper were placed on top as well as 10-12 cm of paper towels. A small weight was added to the blot to add pressure and the blot was left overnight. In the morning, the membrane was UV cross-linked, total energy 1,200 x 100 μJ/cm2.

2.15 Southern Hybridisation and Riboprobe Synthesis

The membrane was pre-hybridised with 10-15 mL of hybridisation solution; 2 x SSPE, 2% SDS, 6 x Dernahrdts reagent, 5% Dextran Sulfate, for 30 minutes at 55 ºC in a hybridisation oven.

To synthesise the riboprobe the desired PCR product was incubated with ATP, CTP, GTP and α-P32 UTP and the T7 enzyme at 20 ºC for 2 hours (Ambion Maxiscript T7 In vitro Transcription Kit). After this 15 µL DNase was added and left for 15 min at 37 ºC. Unincorporated bases were washed away on the Probe-Quant G50 micro-column after centrifuging for 1 min at 3000 *gmax*. The radiolabelled riboprobe was denatured at 90 ºC for 5 min then put on ice before adding to 10 mL of fresh hybridisation solution and then added onto the membrane. The membrane was incubated at 55 ºC overnight for 16-20 hours. The membrane was washed with 0.1 x SSPE, 0.5% SDS until there was no signal left in the wash solution as measured by the Geiger counter. The membrane was then exposed to film.

2.16 Mitochondrial NTP Quantification

2.16.1 Mitochondrial NTP/dNTP isolation

The mitochondrial dNTP and NTPs were quantified from crude mitochondrial pellets which were isolated by differential centrifugation of total homogenate either from cells or tissue, as described in (Spinazzola et al., 2006). The mitochondrial pellet was resuspended in 200 µL 0.5 N TCA and vortexed for 15 seconds. The protein was precipitated by centrifugation at 20,000 *gmax* for 5 minutes at 4 ºC. The supernatant was collected and the pH neutralised by adding 1.5 X volumes (300 µL) of 0.5 M trioctylamine in Freon (1,1,2- trichlorotrifluoroethane). The solution was centrifuged for 10 minutes at 4 ºC, 10,000 *gmax* and the upper-aqueous phase was collected. This extract containing the isolated dNTPs (and NTPs) was dried using a speed vacuum (RT, until completion ~2h) until no liquid remained and the pellet stored at -80 ºC.

2.16.2 NTP Quantification

Using extracts from mouse liver mitochondria, each dried pellet was resuspended in 20 μL TE and further diluted 1:5 before 1 μL was used per reaction. Each reaction mix contained 1μL α-P32 UTP and 2 μL T7 enzyme, 1 X T7 buffer with either i) all 3 remaining nucleotides (G, C, A) but no mitochondrial extracts, ii) C and A, iii) G and A, iv) G and C and v) all three nucleotides; G, C and A. After 2 hours incubation at 20 ºC the reaction mixes were spun through a Probe-Quant G50 micro-column (1 min at 3000 *gmax*) to get rid of unincorporated nucleotides. The resulting radiolabelled RNA products were run on 6% TBE gels at 100 V for 1 hour and the gel dried and visualised using X-ray film.

In order to quantify the relative amounts of NTPs from the extracts each sample was normalised to the lane containing all 4 nucleotides and mitochondrial extracts. The differences in labelling efficiency is then proportional to the internal concentration of the 'missing nucleotide' in the reaction which is compensated for by the mitochondrial extracts.

2.17 Image Quantification Using ImageJ

X-ray films were scanned and saved as .tiff files which were imported into ImageJ. In order to quantify the signal from the different lanes, the subtract background function was used. The signal intensity was measured from plotting the signal from each lane using identical lane markers and calculating the integral for each lane. In all cases, the measurements were normalised to the control lanes to account for variability between samples. Where appropriate standard deviations were calculated and this is shown in the error bars.

2.18 Illumina Truseq mtDNA Library Preparation

Isolated mtDNA samples (6 μg total nucleic acid in 50 μL) were fragmented using sonication were placed in microTUBEs AFA Fibre Pre-slit Snap-cap 6 x 16mm (520045) tubes and sheared using the following settings on the Covaris S220 using the Sonolite software:

Duty cycle: 10%
Intensity: 5
Cycles/burst: 200
Time (s): 120
Temperature (ºC): 4

Following this, the library was prepared according to the Illumina Truseq low sample library preparation protocol. DNA was then and subjected to 200 base-pair paired-end sequencing on an Illumina MiSeq.

2.18.1 Mutational Statistical Analysis

Data are expressed as the mean ± standard error of the mean (SEM). Group means were compared using parametric t-test or non-parametric Mann-Whitney test. One-way ANOVA was used to compare more than two independent groups. A P-value of <0.05 was considered to be statistically significant.

2.19 Modified EmRiboSeq Library Preparation

2.19.1 RNA Degradation and Shearing

Isolated mtDNA (20 μg) was resuspended in 50 μL 0.4 M NaCl solution and free RNA was removed by incubation with RNase A; 2.5 μg 1 hour at room temperature. Following this the mtDNA was sheared using the same covaris settings as above. Samples were then re-precipitated with ethanol and glycogen.

2.19.2 End Repair, dA Tailing and Adapter Ligation

DNA was resuspended in 85 μL with 10 μL of 10 x End Repair Buffer (NEB) and 5 μL of End Repair enzyme mix. Samples were incubated at 20 ºC for 1 hour. Samples were purified with 1.2 x volumes of AMPure XP magnetic beads, washed twice with 75% ethanol carried out on a magnetic stand (referred to as 'bead clean-up' henceforth) and eluted in 42 μL of NFW. For dA tailing the DNA solution was separated from the beads and added to 5 μL 10 x Klenow Buffer and 3 μL Klenow (NEB) for 1 hour at 37 ºC (Clark et al., 1987).

Following bead clean-up the library solution was separated from the beads and resuspended in 20 μL NFW. The solution was incubated overnight with 10 μL 5 x quick ligase buffer (NEB), 5 μL T4 quick ligase and 15 μL dstrP1 adapter at 16 ºC. The library was separated from the beads and resuspended in 195 μL NFW following a bead clean-up.

2.19.3 3' Blocking and Endonucleolytic Treatment

The 3' ends were blocked with addition of 2.5 μL 10 mM ddATP, 25 μL 2.5 mM $CoCl_2$ and 50 units of terminal transferase (and 25 μL 10 x terminal transferase reaction buffer (NEB)) for 2 hours at 37 ºC. Following a bead-clean up, the DNA was separated from the beads and was fragmented by Eco-RNase HII (25 units, NEB) in 100 μL with 10 μL 10x Thermopol buffer for 2 hours at 37 ºC. Fragmented DNA was separated from the beads and resuspended in 85 μL NFW after bead clean up.

2.19.4 Dephosphorylation and Second Adapter Ligation

The DNA solution is dephosphorylated by shrimp alkaline phosphatase (SAP), 5 units for 1 hour at 37 ºC with 10 μL 10x SAP reaction buffer. The SAP is inactivated by heating to 65 ºC for 15 minutes. After bead clean-up the solution is separated from the beads and resuspended in 25 μL NFW and heat denatured at 95 ºC for 3 minutes and then snap-cooled on ice. Immediately, 10 μL 5 x quick ligase buffer, 6 μL dsA adapter and 5 μL of T4 quick ligase (NEB) was added and the library solution incubated overnight at 16 ºC.

This time the bead-clean up used 1.8 x volumes of magnetic beads and the solution was eluted and separated in 40 μL NFW.

2.19.5 Single-stranded lLibrary Preparation and Second-strand Synthesis

Streptavidin beads (Dynabeads M-280) were prepared by transferring 20 μL to a new Eppendorf and using a magnetic stand to retain the beads, the supernatant was discarded and replaced with 1 mL of 1 x bind and wash buffer. Again, the supernatant was removed and replaced with 40 μL of 2 x bind and wash buffer. To this streptavidin bead solution, the 40 μL of DNA library was added and incubated on a

shaker for 15 minutes at room temperature. Using a magnetic stand to retain the beads the supernatant was removed and replaced with 50 μL saline sodium citrate (SSC) and incubated at room temperature for 5 minutes. This was repeated twice and then the supernatant replaced with 40 μL 150 mM NaOH. The solution was incubated on a shaker for 15 minutes. Using the magnetic stand, the eluate is removed and transferred to a fresh tube. This elution step is repeated and the eluates pooled resulting in 80 μL of 150 mM NaOH solution containing the DNA library. 120 μL NFW, 5 μL glycogen, 20 μL 3M NaAc and 550 μL 100% EtOH is added and the solution is re-precipitated.

The DNA library is resuspended in 9.5 μL NFW. 0.5 μL of Primer A, 10 μL of 2 x Phusion master mix (Life technologies) is added and the second strand is synthesised in a thermocycler:

98 ºC 1 minute
58 ºC 30 seconds
72 ºC 1 minute
4 ºC hold

To this solution 1.8 volumes of magnetic beads are used for a bead clean up and the solution is separated from the beads and eluted in 15 μL NFW. The final library solution is stored immediately at -80 ºC.

2.19.6 Quality Control of Final Library

Using 1 μL of the final library solution for each PCR mix the following PCR reactions were carried out:

10 μL final volume, using primers A and trp1 and either 16, 17 or 18 cycles:

i 98 ºC 10 seconds
ii 58 ºC 5 seconds
iii 72 ºC 1 minute
Step i-iii repeat
4 ºC hold

The entire 10 μL PCR mix is run on a 1.2% agarose 1 x TBE gel (with EtBr) with size markers to estimate the amount and size distribution of the DNA library.

2.19.7 Ion Torrent Sequencing

The final library is sequenced using Ion Proton technologies according to the original EmRiboSeq protocol from Ding *et al.*. Assuming the Bioanalyzer trace showed a quantifiable distribution of material between ~200 and 300 bp, which is confirmed by the above-described PCR, the library was then prepared for Ion PGM™ or Ion Proton™sequencing following the manufacturer's protocol. (The required input for Ion OneTouch™ emulsion PCR is 100 µL of 12 pM for the Ion Proton™ and 26 µL of 25 pM for the Ion PGM™).

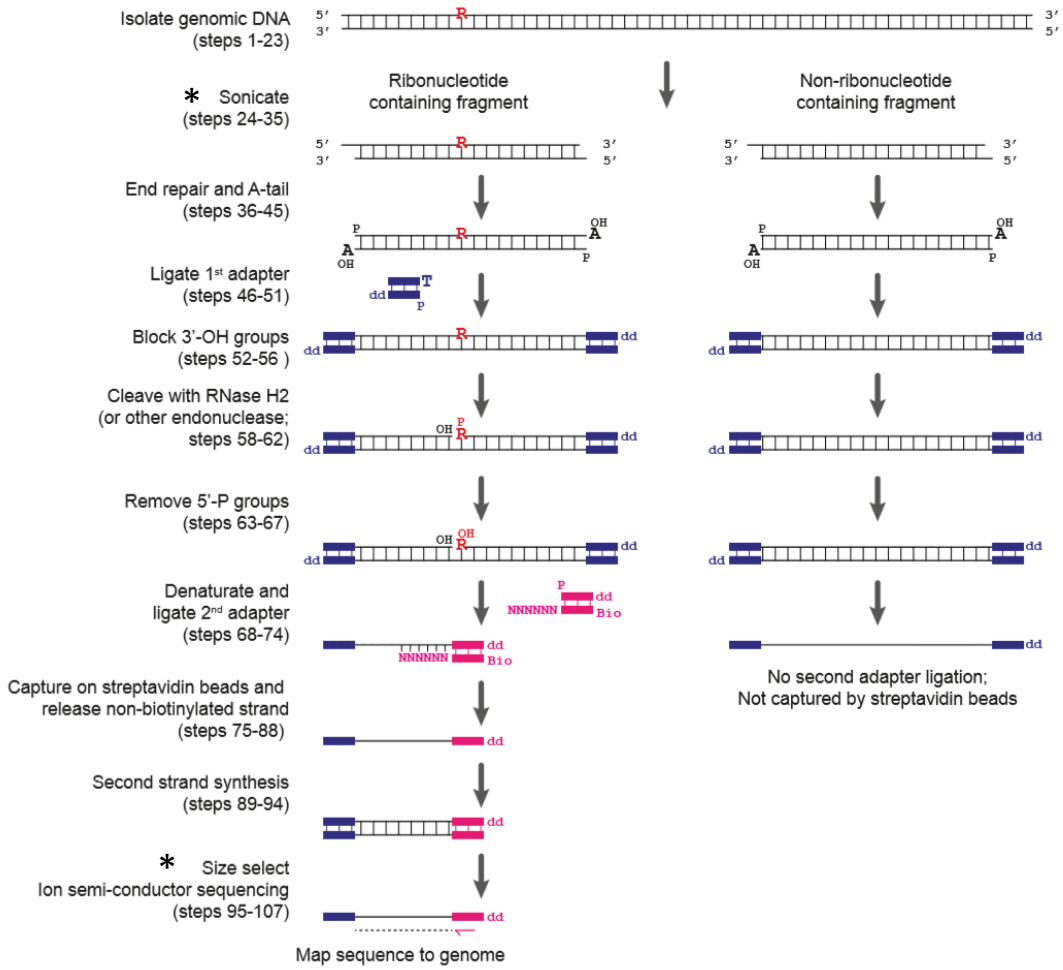


Figure 9: EmRiboseq

The EmRiboseq protocol as developed by Ding *et al.* (Ding et al., 2015). The * highlight the steps in the original protocol which are modified in the mitochondria-specific EmRiboseq protocol adapted and used within this thesis. It is important to emphasise that the site which is sequenced and mapped is that which is directly 5' to the embedded ribonucleotide (referred to as +1).

2.19.8 Primer-Directed EmRiboseq

The protocol was followed as detailed above with the exception of the endonucleolytic treatment (2.19.3). In place of Eco-RNase HII, the library was incubated with 1 μg RNase A, 30 units of Eco-RNase HI in 1x RNase HI buffer (Homemade; 50 mM Tris-HCl, 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, pH 8.3) for 2 hours at 37ºC.

2.20 EmRiboseq Data Analysis and Bioinformatics

2.20.1 Genomic Mapping Strategy

Mapping of reads largely as previously described (Ding et al., 2015). However, for the purposes of comparisons between chrM and the nuclear genome, and to look at the rDNA locus in more detail all libraries have been re-mapped. Bowtie2 paired end read mapping and Samtools alignment quality filtering. The reference genome for mapping was mm9 including all major chromosomes and random assembly fragments. The reference chrM sequence was aligned by blat to the reference genome requiring high identity matches (match-score > 100, identity > 95%). All identified matching segments (excepting chrM itself) were N-masked in the reference. The same filtering approach was applied to reference genome segments matching rDNA reference fragment (NCBI:). The rDNA reference fragment was included as an additional separate sequence in the Bowtie2 library.

2.20.2 Genomic Composition

Due to the repetitive nature of the genome and the necessity to filter on alignment quality, a substantial and compositionally biased portion of the genome cannot be mapped. Accounting for this, all segments of the reference genome that could be uniquely mapped (GRCmappability100 track from UCSC) were extracted (bedtools fastaextract) and their nucleotide and trinucleotide compositions calculated. Trinucleotide context was calculated for each identified ribonucleotide incorporation site, where the ribonucleotide was the central of the three nucleotides on the ribonucleotide-containing strand.

2.21 Western Blot Analysis from Cultured Cells

2.21.1 Antibodies

Antibody	Dilution	Secondary	Catalogue number
RNase H2A	1:200	rabbit	Collaborators - Jackson laboratory, Edinburgh
TOM20	1:500	rabbit	Santa Cruz 11415
Vinculin	1:4000	mouse	Abcam- ab130007
TFAM	1:1000	mouse	Collaborators
FEN1	1:2000	mouse	GenTex
Top1mt	1:2000	mouse	Collaborators – Pommier lab, NIH
Twinkle	1:500	rabbit	Sigma- HPA002532
GAPDH	1:50000	mouse	Sigma – G8795
NDUFB8	1:1000	mouse	Abcam ab110242
PrimPol	1:500	rabbit	Collaborators –Luis Blanco, Madrid
MGME1	1:250	rabbit	Atlas Antibodies – HPA040913
MPV17	1:500	rabbit	Proteintech 10310-1-AP

Table 2: Antibodies

2.21.2 Protein Isolation and Quantification

Typically, steady state protein levels were examined from one 15cm dish of attached cells, at no more than 80% confluence. Cells were collected by mild trypsin treatment to detach cells. Cells were spun at 1,100 *gmax* for 3 minutes and washed with cold PBS before being spun again. The resulting cell pellet was resuspended in 40 µL PBS + protein inhibitor cocktail (PIC) with 0.1% DDM and 50 units Benzonase (Millipore) and incubated for 20 minutes on ice. After this time, 1% SDS was added and the solution is vortexed regularly for 30 minutes on ice. The suspension was left at room temperature for 10 minutes prior to protein quantification determined using the DC^{TM} protein assay (Biorad).

2.21.3 PAGE and Membrane Transfer

Samples are diluted using PBS and 25 µg of total cellular protein was loaded per lane. 4-12% Bis-Tris gels were typically run at 120 V for 90 minutes, but this varied depending on the size of the protein of interest, with 1 x SDS PAGE running buffer.

To transfer the samples to the membrane, the membranes were first activated in methanol. Whatmann papers and blotting apparatus were submerged in 1 x transfer buffer (Tris-Glycine-SDS (0.25 M Tris, 1.92 M Glycine, pH. 8.3)), 20% methanol and the blotting apparatus set up as per the manufacturers guidelines. The transfer was left at 100 V for 90 minutes at 4 ºC.

2.21.4 Immunoblotting

The membrane was blocked with 5% Milk in 0.05% PBS-T at 4 ºC for at least 45 minutes. The membrane was incubated with the primary antibody overnight at 4 ºC in either PBS-T or milk depending on the antibody. The membrane was washed with 0.05% PBS-T and then incubated with the secondary antibody (HRP conjugated anti mouse and rabbit, Promega) for one hour (1:4,000 in PBS-T). The membrane was washed with 0.05% PBS-T and then imaged using luminescent reagent (ECL GE healthcare) and X-ray film.

2.22 Immunocytochemistry and Confocal Microscopy

Cells were plated and fixed at approximately 70% confluency using 2% neutral buffered formalin (NBF) (diluted in PBS). Following overnight incubation, cells were permeabilized in Triton X and washed with PBS extensively. The fixed cells were washed with blocking buffer (5% Goat Serum in PBS) and incubated with the primary antibodies (diluted in blocking buffer) overnight at 4 ºC.

The cells and primary antibodies are washed with 0.2% PBS-T to ensure that any unconjugated primary antibody has washed off and then incubated with the secondary antibody (1:1000 in 0.2% PBS-T) for one hour at room temperature. The slides were washed with PBS-T and oiled and inverted and imaged within 24 hours.

Chapter 3. Results I: Next Generation Sequencing of Ribonucleotides in Mammalian MtDNA

3.1 Background

All replicative polymerases incorporate ribonucleotides, however infrequently (Cerritelli and Crouch, 2016). Incorporated ribonucleotides in replicating DNA are known to serve as markers for the newly synthesised DNA strand and so can aid in quality control and repair systems such as mismatch repair which ensure faithful replication of the template DNA (Ghodgaonkar et al., 2013, Lujan et al., 2013, Kunkel and Erie, 2015).

In nuclear DNA, incorporated ribonucleotides are rapidly excised by an efficient RNase H2 mediated RER mechanism, and therefore embedded ribonucleotides are transient and an accumulation of unrepaired ribonucleotides leads to genomic instability (Reijns et al., 2012). Therefore, the presence of persistent, unrepaired ribonucleotides throughout the mtDNA is unique. And given the structural perturbations and chemical lability their presence confers to the DNA molecule it is not fully understood why they are tolerated. It is theorised that embedded ribonucleotides may be better tolerated by mtDNA because of its small genome size, the slow rate of DNA synthesis (Bogenhagen et al., 1979) or the bacteriophage-like enzymes involved in its replication (PEO1, POLG and POLRM) (Shutt and Gray, 2006).

In the last few years, multiple groups have developed various techniques to map embedded ribonucleotides to single resolution which has elucidated valuable information about polymerase behaviours and replication origins in eukaryotic DNA replication (Clausen et al., 2015, Ding et al., 2015, Koh et al., 2015, Keszthelyi et al., 2015). Nevertheless, such techniques have been neglected as a tool for understanding mtDNA replication and metabolism, until recently. Berglund *et al.* utilised one of the aforementioned techniques, HydEn-seq, to map alkali-sensitive sites in human mtDNA (Berglund et al., 2017). They reached the broad conclusion that nucleotide pools within the mitochondria are the predominant determinant of the identity and frequency of ribonucleotide incorporation in mtDNA. The results

presented in this chapter agree generally with those conclusions, suggesting that this is unequivocally the case.

In addition, Berglund *et al.* used two different cell models of perturbed mitochondrial dNTP pools to demonstrate that when dNTPs within the mitochondrial compartment are limiting, NTPs are utilised to maintain mtDNA replication. Their data support the conclusions that (mis)incorporation by Pol γ is the predominant source of ribonucleotides in mtDNA and that there is no evidence for a RER pathway in mitochondria. These statements are sound for the context in which ribonucleotide incorporation was analysed but here it is intended to provide a more comprehensive picture of ribonucleotide incorporation and its role in mtDNA metabolism and mitochondrial disease.

3.2 Southern Blot Analysis Reveals Persistent Ribonucleotides in Murine mtDNA

Isolated mtDNA can be fragmented by incubation *in vitro* with RNase H2, which nicks DNA 5' to any single embedded ribonucleotides in duplex DNA, due to the presence of persistent ribonucleotides throughout the mature mtDNA molecule. MtDNA can be fractionated by 1D-AGE and visualised using a radiolabelled riboprobe specific for the NCR containing region of the mtDNA. RNase HII-sensitive molecules can be identified as the smear of fragments which run below the full-length untreated molecule (Figure 10).

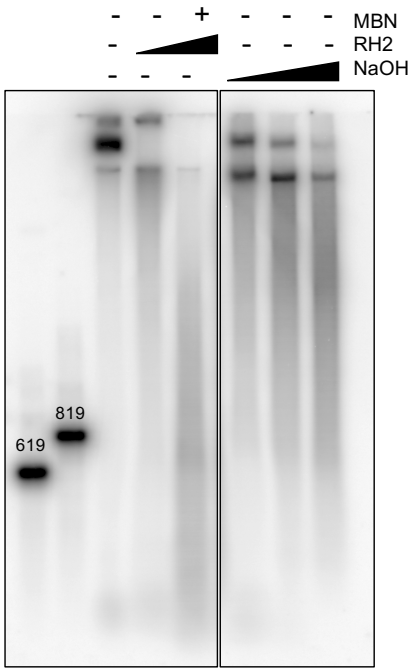


Figure 10: Southern blot analysis of mouse liver mtDNA treated with (+) or without (-) RNase HII.

Purified murine liver mtDNA was fractionated by 1D-AGE, after treatment with RNase H2 or a combination of RNase H2 and Mung Bean Nuclease (MBN), which cleaves single-stranded DNA and at nicks in duplex DNA, or denatured and hydrolysed with NaOH. The DNA was blot hybridized to probe to H-strand np 14,881-16,299. Size marker lanes with DNA fragments of 619 and 819 nucleotides are followed by mtDNA 1) uncut; 2) RNase H2 1:25 of 11.8 μM, 37°C, 2 h; 3) RNase H2 1:25, 2h MBN 5U 3h; 4) 100 mM NaOH; 5) 150 mM NaOH; 6) 200 mM NaOH 30 min.

3.3 Persistent Sites of Frequent Ribonucleotide Incorporation in Murine mtDNA

As demonstrated in Figure 10, mtDNA can be fragmented at sites of ribonucleotide retention using NaOH and/or RNase H2, and visualised via Southern Blot using a mtDNA-specific radiolabelled probe. Restriction digestion of mtDNA creates smaller fragments of DNA which can be interrogated more closely via 1D-AGE. MtDNA cut with a restriction enzyme and then incubated with Eco-RNase HII gives rise to a laddering effect (Figure 11). The mtDNA appears to be fragmented by RNase HII at persistent sites in the mtDNA of multiple molecules suggesting that there are sites within the mtDNA where there is frequently an embedded ribonucleotide. Not only

does this conclusively demonstrate the presence of sporadic ribonucleotides throughout many of the mtDNA molecules, but the laddering pattern suggests that there is a non-uniform distribution of ribonucleotides and that there are 'hot-spots' of persistent ribonucleotide incorporation throughout the genome.

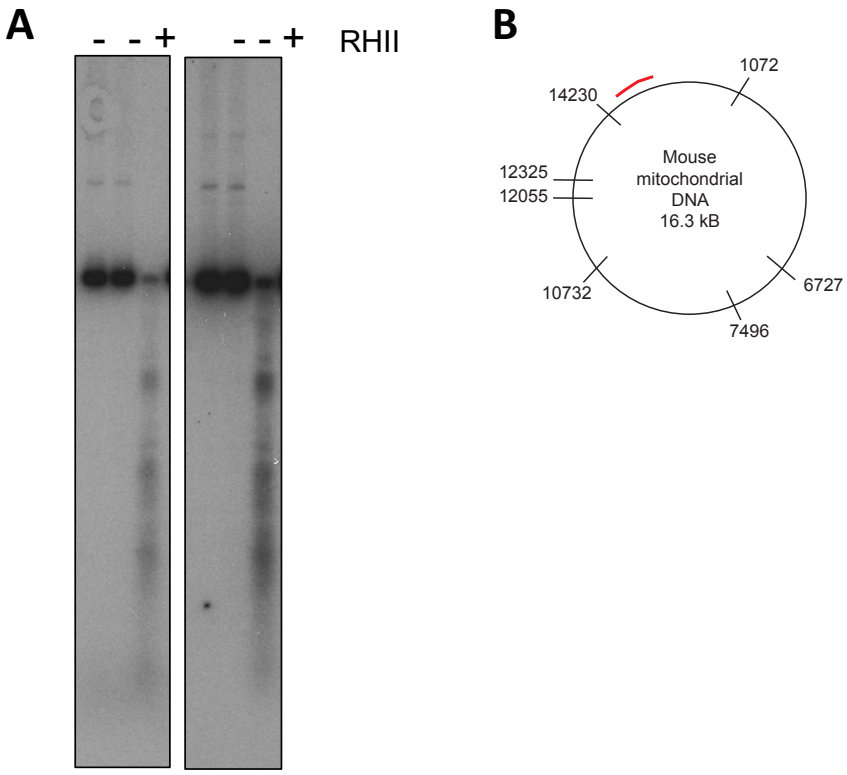


Figure 11: Southern blot analysis of mouse liver mtDNA treated with (+) or without (-) RNase HII after *BspHI* digestion gives rise to a laddering effect.

(A) The mtDNA was restriction digested with *BspHI*, and with (+) or without (-) Eco-RNase HII (2 units, 2 hours, 37 °C) heat denatured (3 minutes at 95 °C) and fractionated by 1% agarose gel electrophoresis prior to transfer to a nylon membrane and hybridization with a H-strand specific riboprobe (14881-15490 nt) to detect the 3.14 kilobase fragment spanning from 14230 to 1072 nt. (B) A diagrammatic restriction map of *BspHI* digested mtDNA along with the probe (red) used to detect the heavy strand of mtDNA.

A non-random pattern of ribonucleotide incorporation is indicative of a sequence bias in ribonucleotide incorporation and/or functional roles for particular loci of ribonucleotide incorporation, such as programmed pause sites or replication origins. In order to better examine the intricacies of ribonucleotide incorporation in mtDNA I employed a novel, tailored Next Generation Sequencing approach in order to gain qualitative detail about the sites of ribonucleotide incorporation at single nucleotide resolution.

3.4 Next Generation Sequencing Protocol Development and Troubleshooting

EmRiboseq is one of four published NGS protocols to map embedded ribonucleotides *in* vivo. The benefit of the EmRiboseq method, as developed by Ding *et al*., is the specificity for single embedded ribonucleotides conferred by using RNase HII (rather than alkaline hydrolysis which fragments abasic sites and long stretches of RNA, in addition to single ribonucleotides), as well as the ability to retain strand information. Prior to this study, the protocol had been adapted for nuclear DNA in *S.cerevisiae*, and therefore needed to be modified for mammalian mtDNA. Moreover, published data utilising EmRiboseq had been retrieved from *rnaseh2* deficient models (with elevated embedded ribonucleotides). For that reason, it was important to demonstrate that this method was both appropriate and applicable to control mouse tissues which are RNase H2, and therefore RER - proficient.

3.4.1 EmRiboseq Protocol Adaptations

To refine the procedure for use with mtDNA, the input DNA used was from sucrose-purified mitochondria. The original protocol was followed and extracts of the library were removed at different stages throughout to estimate the amounts of DNA in the library after each subsequent step and to identify any problematic points within the protocol (as summarised schematically in Figure 12B). Variations of the protocol included removing the beads from the reaction mix at each step, as it was possible that the beads were interfering with subsequent reactions by residually binding to the DNA libraries. Another consideration was the solution which is used to bind the DNA molecules to the beads; -polyethylene glycol (PEG).

It is well-known from anecdotal evidence and personal communication with other researchers that using magnetic beads to clean-up libraries can lead to substantial loss of material and this was confirmed by Southern-blot analysis of mtDNA fragments taken throughout the procedure which showed a progressive decrease in yield following bead clean-up(s) (Figure 12A). In preparations where the beads were retained in solution throughout (lanes 6 and 7) a large amount of material was lost and the resulting amount of DNA is below the limits of detection, demonstrating that

removal of the beads between steps was essential to retain the DNA under these conditions (Figure 12A). It is for this reason that the adapted EmRiboseq method as described in this thesis removes the beads before each subsequent enzymatic treatment, which results in consistently high yield library preparations. The other deviations from the original published protocol used in the final method in this thesis are when isolating the target DNA, the sonication settings and the omission of the size selection step (see methods).

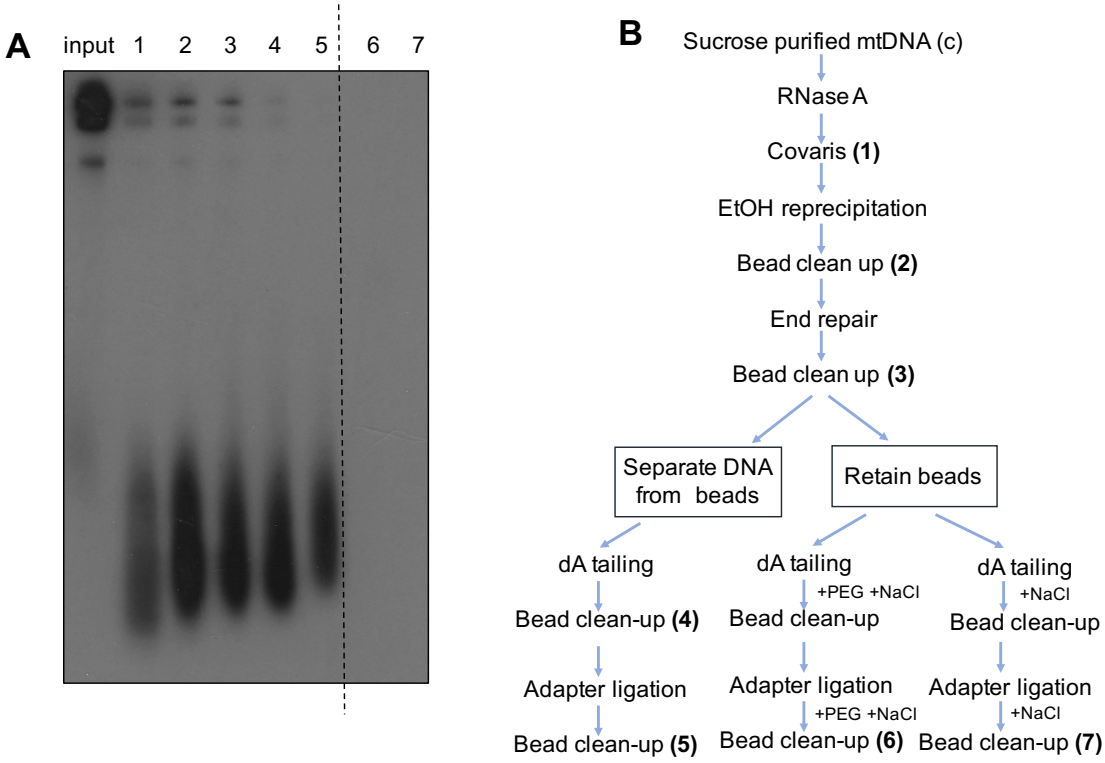


Figure 12: Separating beads from the DNA solution markedly reduces loss of material during library preparation.

(A) Southern blot (H-strand radiolabelled riboprobe nt 14881-16299) of progressive stages of the protocol where input is the untreated starting material – 16 kB closed-circular mtDNA. (B) Protocol variations of the initial phase of the EmRiboSeq protocol. The right-hand branch (1, 2, 3, 7) is identical to the original protocol. After Step (3) the sample was split and the beads were either retained or removed from the library using a magnetic stand. Removing the beads from the solution (lanes 4 and 5) resulted in higher DNA yield.

3.4.2 NUMTs

NUMTs are nuclear mtDNA sequences and are presumed to be the result of transfer of mitochondrial genomic information to the nucleus in evolving eukaryotes. In the

case of mitochondrial-EmRiboseq, the primary signal of interest is precise read end-mapping of mitochondrial sequences and therefore we used a global (rather than local) alignment strategy. In general, any preparation of sub-cellular compartment DNA is likely to be contaminated by nuclear DNA, so alignment would normally be to the whole nuclear and mitochondrial genome rather than just mitochondrial. However, these high identity mitochondrial segments integrated in the nuclear genome would either reduce mapping read coverage (through signal dilution) or reduce mapping quality score (and abundance) leading to erroneous filtering of reads. Masking of high identity chrM segments in the reference mm9 nuclear genome provided a solution to this problem. Figure 13 highlights regions of high identity matches to NUMTs throughout the mitochondrial genome and how this mapping strategy improves the mapping of mitochondrial reads. One negative from this method is that real nuclear reads which map to NUMTs in the library will be incorrectly attributed to mtDNA, but this is considered to be negligible owing to enrichment of mtDNA during DNA extraction, and procedural enrichment for ribonucleotide-containing fragments, of which is very low in wild-type nuclear DNA compared to mtDNA.

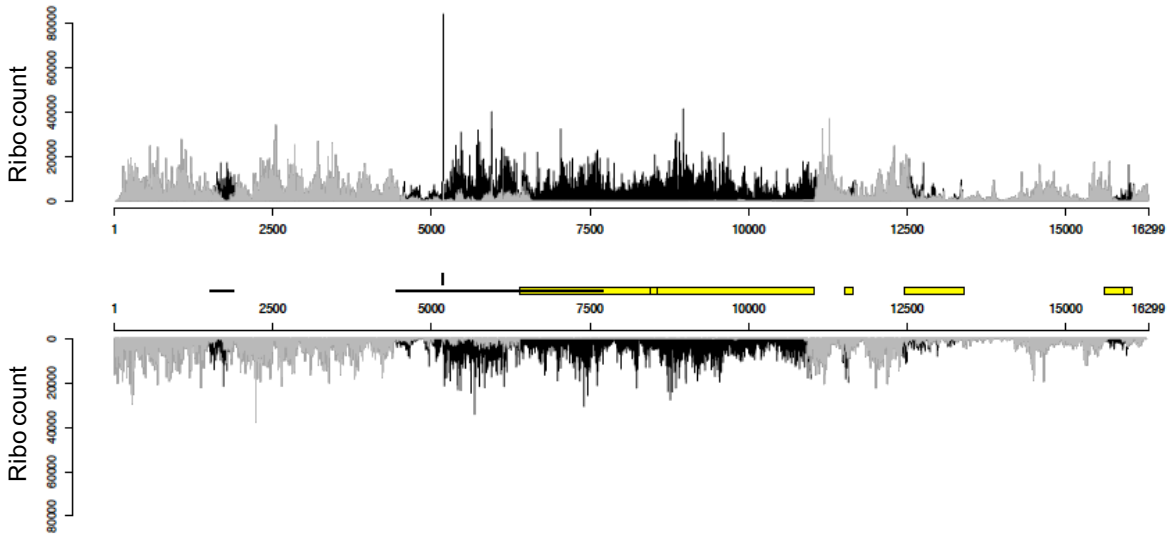


Figure 13: The effect of NUMTs on mitochondrial read mapping.

Sonicated wild-type murine liver data mapped to forward (top) and reverse (bottom) strands of chrM. Yellow segments show high identity matches to NUMTs from the *mus musculus* nuclear genome, deeper bars denote greater identity. The longest segment is 4648 nt and only shows one mismatch to the chrM reference. Grey bars show the results of standard mapping. Black bars show the difference when masking

the high identity NUMTs from the nuclear genome. [Bioinformatic analysis by Dr Martin Taylor]

3.4.3 Sequence Drop-out Analysis

It became clear from the read-mapping that there were significant regions of high and low coverage across the genome. To assess whether this was attributable to ribonucleotide incorporation frequencies or an artefact, wild-type mtDNA was prepared for EmRiboseq analysis with exclusion of the RNase H2 nicking step. Therefore, any residual signal could be attributed to either background nicks incurred during library preparation or residual unblocked ends from ineffective ligation (blocking). What became clear from this control-library analysis was that regions of low coverage within libraries were correlated with regions of low mapping quality. However, it is not the uniqueness of the sequence which was causing this phenomenon, but instead regional biases in read length. Consequently, shorter reads were less likely to map and thus producing a recurring pattern of peaks and troughs (Figure 14). Although why truncation occurs unevenly is not clear.

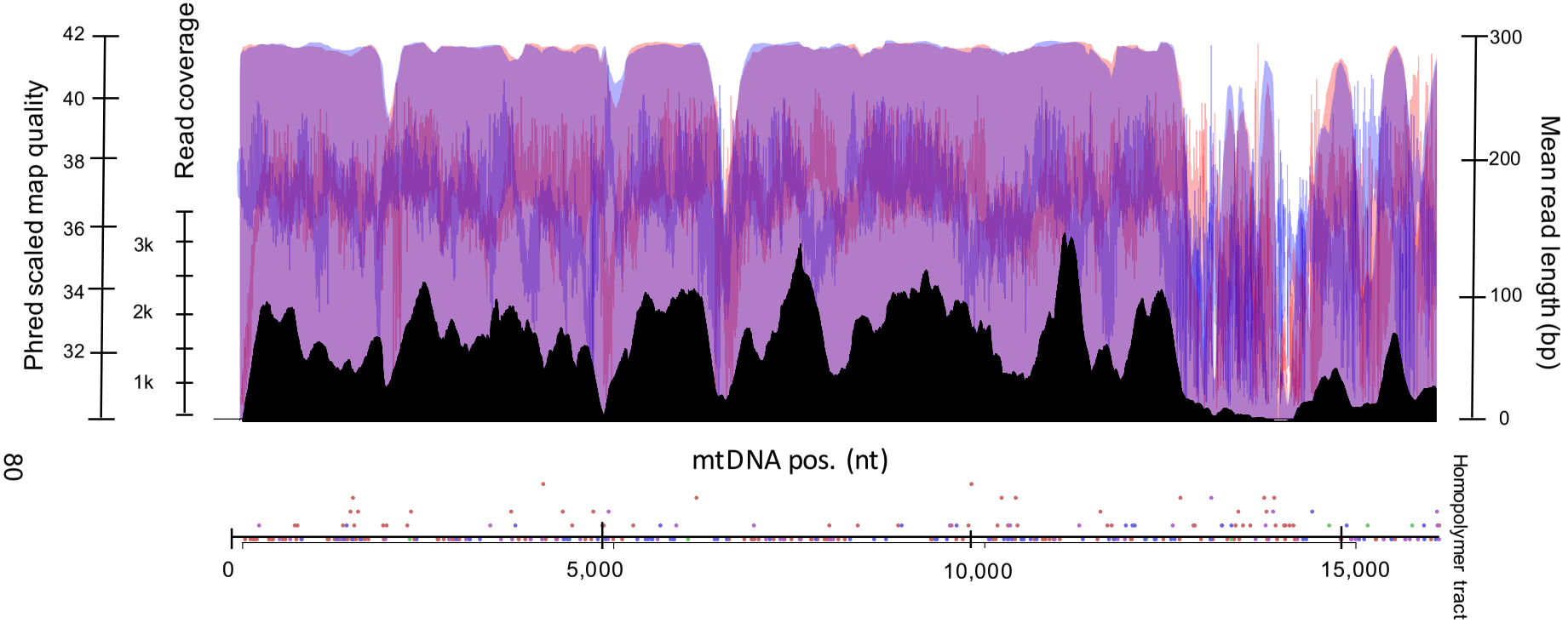


Figure 14: Sequence drop-out analysis of wildtype mouse liver mtDNA.

Overlay of mapping quality, read coverage and average read length of untreated mtDNA. Sequence coverage of the genome is shown in black (L-strand shown, but the effect is comparable for both strands), read quality in block purple, and the mean read length in red/blue where red is the L- strand and blue is the H- strand. [Bioinformatic analysis by Dr Martin Taylor]

3.5 Quantification of Ribonucleotides in MtDNA Using EmRiboseq

The EmRiboseq method alone is not absolutely quantitative because of the way in which the reads are captured; meaning that only molecules which are RNase HII-sensitive and are within the optimal range from the adapter terminus, are sequenced (Figure 9). Because the shearing stage of the protocol is, in theory, random, it is anticipated that the majority of embedded ribonucleotides throughout a population of mtDNA are identified which means that within a sample, relative abundances of different sites and or ribonucleotide base can be quantitatively compared but an absolute number of ribonucleotides per molecule is not possible.

In order to make this technique more quantitative I employed a method to normalise the reads to the total number of mtDNA molecules in the library using restriction digestion, a similar tactic to Ding *et al.* (Ding et al., 2015). Using this approach, it is also possible to quantify the background signal in the absence of RNase HII. As restriction digestion will cut all molecules with the correct restriction site, this provides an internal control. The enzyme used was DraI which recognises the site TTT|AAA and cuts nine times in *mus musculus* mtDNA (1478, 3043, 3896, 5276, 9820, 11011, 11653, 13198, 14177 nt). Three libraries were sequenced in parallel for mouse liver mtDNA (n=2); i) no treatment, ii) DraI and iii) DraI and RNase HII. When the read-ends were analysed by the flanking 3-mers, the ends mapping to TTT|AAA were quantified relative to the total number of ends. The value obtained from the untreated library (i) gives the background signal. Analysis of the 10 most frequent 6mers revealed that approximately 15% of reads from the (ii) DraI treated library, and 5% of the (iii) DraI and RNase HII treated library mapped to TTT|AAA sites (Figure 15). The high background signal is likely due to ineffective ligation (blocking) at the beginning of the library preparation, and/or breakage during subsequent steps following restriction digestion. It will also be affected by incomplete digestion, however, DraI incubation was optimised to ensure maximal digestion.

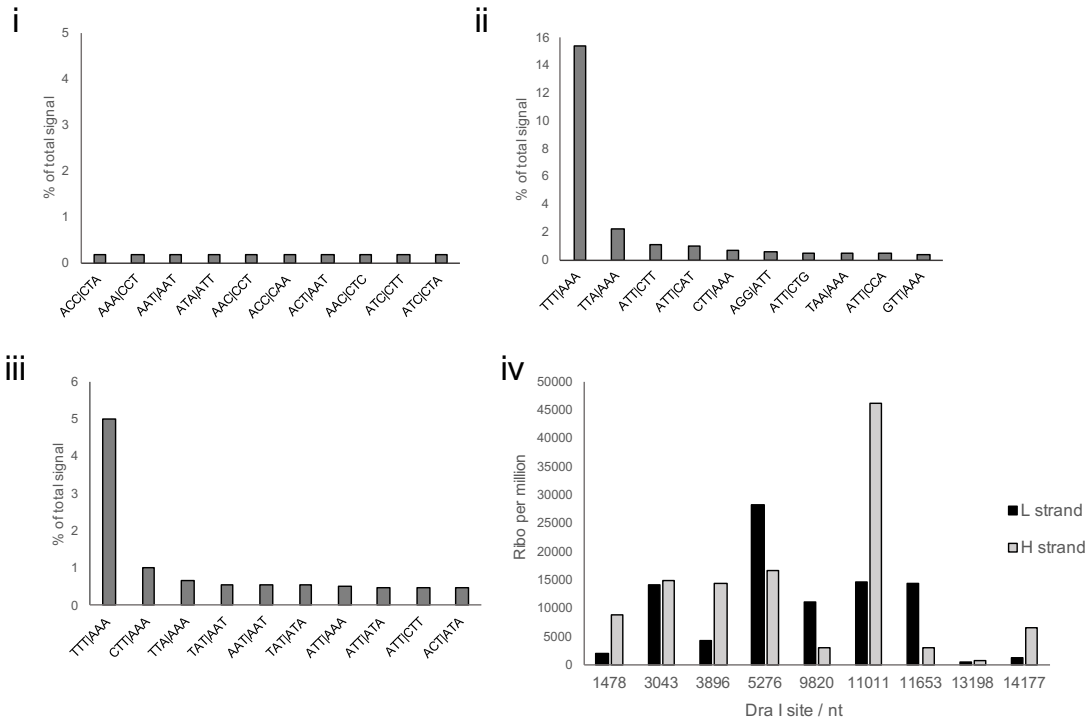


Figure 15: Quantification of the 10 most frequent 6mers in control library preparation (wild-type mouse liver, ChrM).

i) Wild-type liver mtDNA untreated during library preparation- all mapped reads in this instance are background signal, as shown in Figure 14. ii) Wild-type liver mtDNA incubated with DraI where the recognition site TTT|AAA is the most prominent site. iii) Wild-type liver mtDNA incubated with DraI and RNase HII. iv) total reads for each DraI site in mtDNA, separated by strand from ii.

When examining each DraI site from libraries which were treated with the restriction enzyme alone there is a huge amount of variation between different DraI sites across the genome and also between the same site on different strands (Figure 15 iv). This supports the evidence presented in Figure 14 that there is a great deal of coverage (read depth) bias of reads due to fragment size. For example, the most abundant site in Figure 15 iv) is reasonably proximal to the next site (11,011 and 11,653 nt respectively), and is therefore more likely to release fragment sizes which fall within the right size-range for the sequencing platform following DraI cutting. However, it is not clear why there are such big discrepancies for the same site between strands. Again, this could be attributed to size biases, but it would mean that there was frequently a sheared end either too far away from- or to near to- one side of the DraI cut site. Therefore, this could suggest that there are fragile sites within the genome where there is consistently a break during the shearing stage of the library

preparation. These sites could be attributed to sites of persistent ribonucleotides which persist at hot spots in the genome (Figure 11) and are structurally perturbed within the DNA double helix, and perhaps more likely to break during physical shearing (Chiu et al., 2014, DeRose et al., 2012, Evich et al., 2016). Significant under-representation of certain sites (For example; 13,198 nt) can be explained by the sequence drop-out analysis (Figure 14). Sites which fall within the trough between approximately 13,000 and 16,000 nt prove difficult to read and many reads are truncated and not aligned.

By using the restriction enzyme reads to quantify the number of mtDNA molecules within the library, a number of approximately 340 ribonucleotides per molecule was calculated. However, this is still preliminary and doesn't take into account the background signal, nor the regional bias and lost restriction-enzyme sites due to ribonucleotides. More development would need to be done to make the technique truly quantitative.

Reassuringly, in the library where there was no treatment, the trinucleotide rates across the genome (the frequency of all 64 possible 3mer sequences where the central base is the ribonucleotide along with its flanking bases) are approximately equal for each base, a phenomenon which was not observed in any other conditions (Figure 16). This indicates that the background has no base or triplet bias.

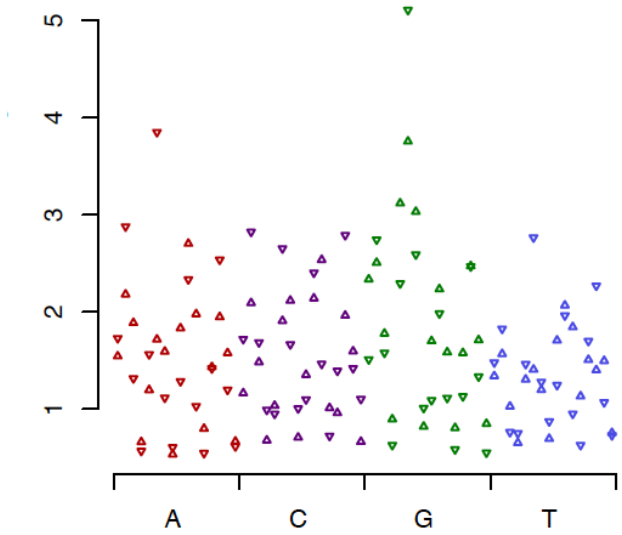


Figure 16: Trinucleotide rates of sequenced ends in mouse liver mtDNA untreated libraries.

Each triangle represents the rate of a trinucleotide sequence where the ribonucleotide is the middle base. Each of all the 64 trinucleotide possibilities is plotted as a function of its normalised signal in the library. Upward pointing triangles = L-strand, downward pointing triangles = H-strand. [Bioinformatic analysis by Dr Martin Taylor]

3.6 Equal Ribonucleotide Incorporation Frequencies in Both MtDNA Strands Revealed by EmRiboseq

Sequencing multiple biological replicates revealed that across all libraries there were near-identical frequencies of ribonucleotide incorporation between the strands of the mtDNA. Calculating ribonucleotide incorporation rate per trinucleotide separately for each strand demonstrated that the ribonucleotide incorporation frequencies are the same (a wild-type liver example is shown in Figure 16) and any differences can be attributed to differences in the sequence content between the two strands. This result strongly indicates that the same DNA polymerase is acting on both mtDNA strands. This is because equal relative rates of ribonucleotide incorporation on both strands uncovers information about the behaviour and propensity of the replicative polymerase to (mis)incorporate ribonucleotides. And this finding is concordant with all proposed mechanisms of mtDNA replication in mammals.

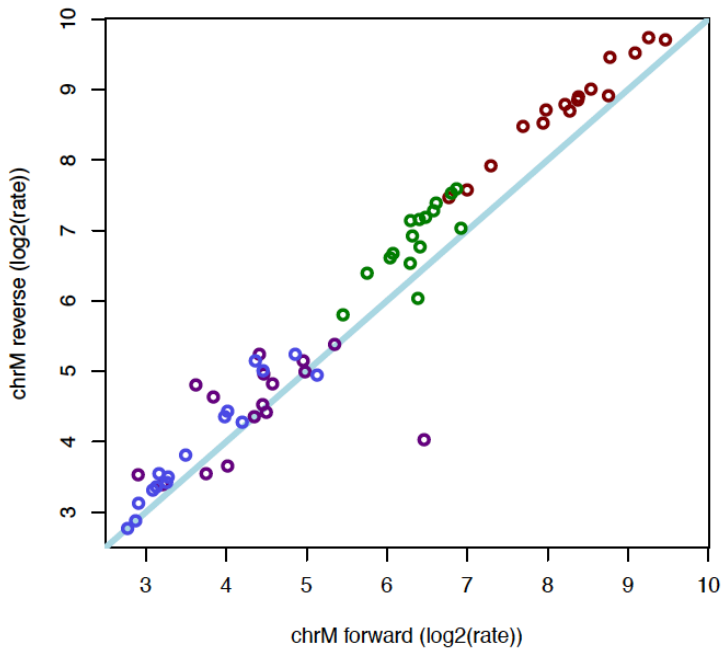


Figure 17: Ribonucleotide frequencies per trinucleotide match between the two mitochondrial (ChrM) strands.

Each site of the chrM genome was categorised onto one of 64 trinucleotide categories and the rate of ribonucleotide per trinucleotide category calculated separately for the forward (light) and reverse (heavy) strands. Each point represents one of the 64 trinucleotide categories (Red = rA, Green = rG, Blue = rU, Purple = rC). The diagonal corresponding to equal frequencies between strands is shown as a light blue solid line. The frequent outlier point (purple spot) corresponds to the trinucleotide at the Ori-L high frequency retained site. Excepting the Ori-L signal the two chrM strands show the same rate pattern. [Bioinformatic analysis by Dr Martin Taylor]

3.6.1 Polymerase γ, and not PrimPol, is Responsible for Ribonucleotide (Mis)incorporation in mtDNA in Cultured MEFs

For many years Pol γ was assumed to be the sole DNA polymerase in mammalian mitochondria. PrimPol is a X-family DNA polymerase which is also present in the mitochondria and is well-documented as a TLS polymerase (Bianchi et al., 2013, García-Gómez et al., 2013, Mourón et al., 2013). However, unlike Pol γ, PrimPol is not essential for mtDNA maintenance (García-Gómez et al., 2013). However, as a low-fidelity polymerase it is most likely to be involved in replication restart and fork rescue than responsible for synthesising large portions of mtDNA. In line with this, EmRiboseq analysis of wild-type and PrimPol knockout immortalised MEFs showed

no difference in trinucleotide ribonucleotide incorporation frequencies (Figure 18). In other words; there is no change in the observed patterns of ribonucleotide incorporation in the absence of PrimPol. From this it is inferred that under normal cell cultured conditions PrimPol is not incorporating the vast majority of ribonucleotides that are identified across the mitochondrial genome. That is not to say, that under conditions of increased oxidative stress where DNA damage and replication stalling is more likely, that PrimPol activity would be heightened. From this data, it is suggested that Pol γ, the mitochondrial replicative polymerase, is responsible for ribonucleotide incorporation in mitochondria and contribution by PrimPol is negligible.

If PrimPol were responsible for the embedded ribonucleotides there would not only be a difference in the abundance of ribonucleotides, but also the frequencies of the different trinucleotides. This is because each different polymerase has a different propensity for incorporating ribonucleotides; which is determined by its discrimination factor (DF); the ability of the polymerase to effectively discriminate between each deoxyribonucleotide base and its ribonucleotide equivalent.

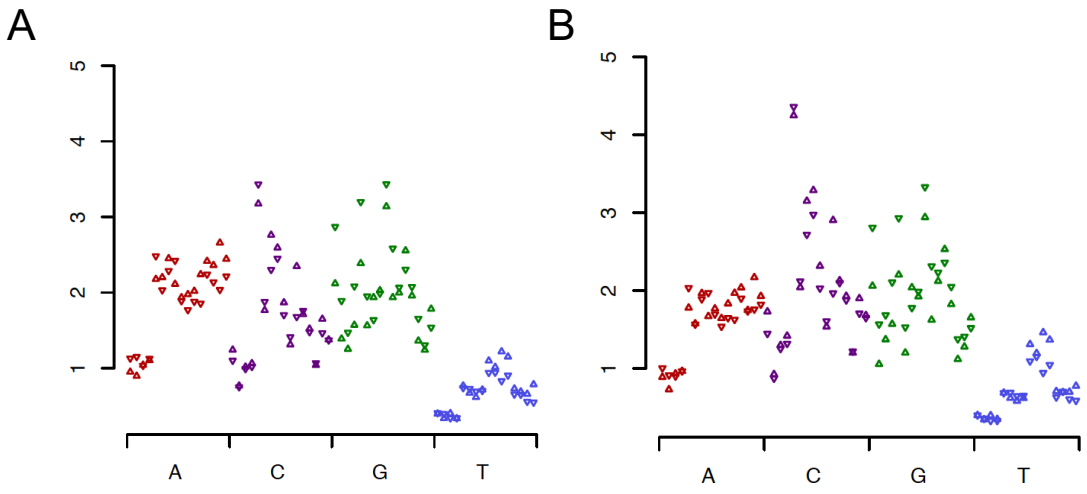


Figure 18: Trinucleotide frequencies of wildtype (A) and PrimPol knockout (B) immortalised MEFs.

Each triangle represents the rate of a trinucleotide sequence where the ribonucleotide is the middle base. Each of all the 64 trinucleotide possibilities is plotted as a function of its normalised signal in the library. Upward pointing triangles = L-strand, downward pointing triangles = H-strand. [Bioinformatic analysis by Dr Martin Taylor]

3.7 Base-Biases in Ribonucleotide Retention in Mouse MtDNA from Post-Mitotic Tissue

Despite the equality in ribonucleotide abundance between the two mtDNA strands, there was a striking bias among the four ribonucleotide bases. EmRiboseq analysis revealed that adenine mononucleoside phosphates (rAMPs) were the most frequently incorporated ribonucleotide in wild-type mouse liver mtDNA, where 83% (n=3) of all ribonucleotides were identified as rAMPs (Figure 19). This base-bias was replicated, although not to such a dramatic extent, in all post-mitotic tissues analysed (Figure 19). This is greater than the sequence identity of the mouse mitochondrial genome (ChrM), which is moderately AT rich.

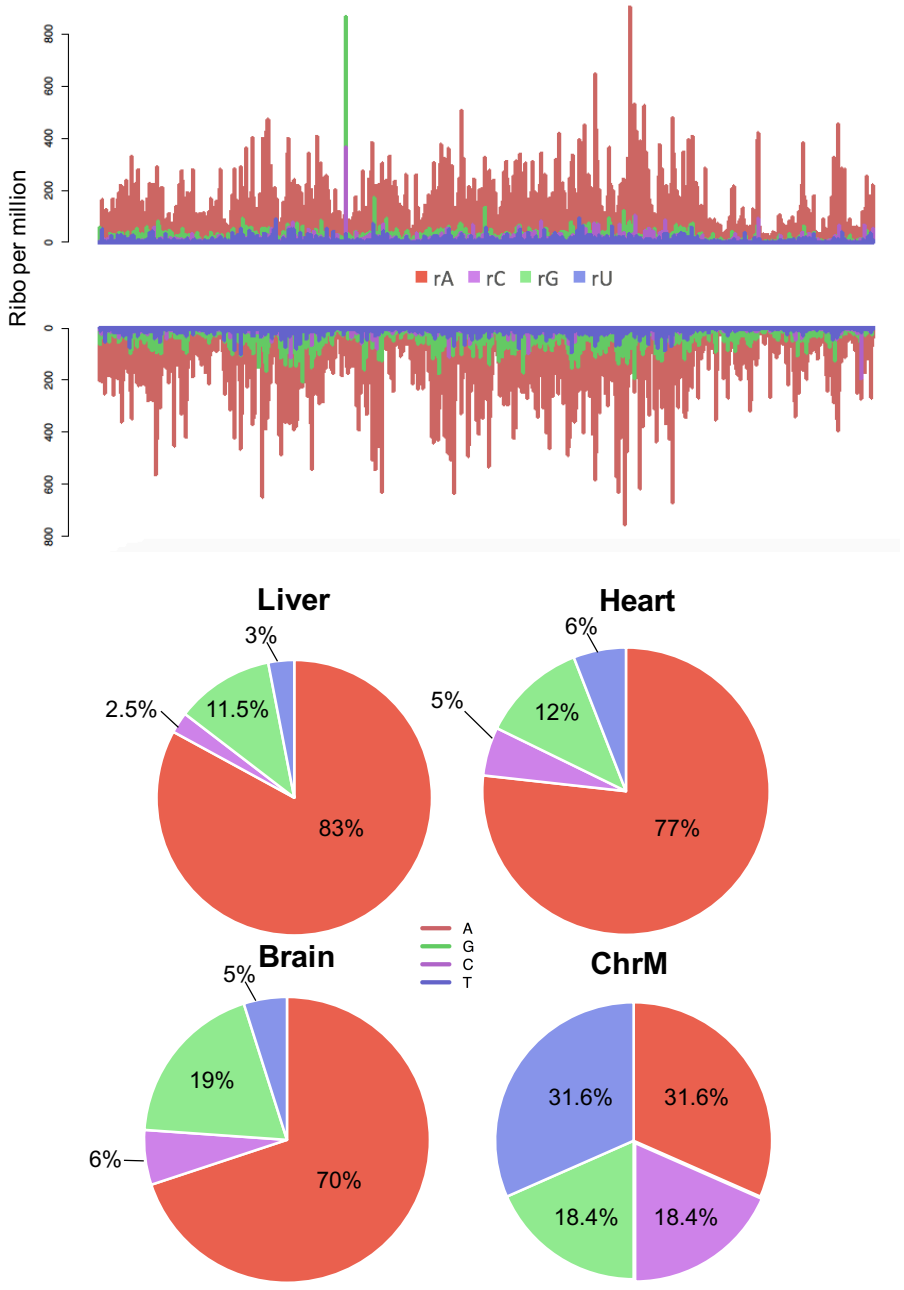


Figure 19: rAMP is the dominant ribonucleotide in mtDNA of murine solid tissues. (Top) The profile of ribonucleotide incorporation in wild-type mouse liver mtDNA. (Bottom) Base frequency in mtDNA EmRiboseq libraries from different tissues.

(Top) The frequency of ribonucleotides (y-axis) across the mitochondrial genome (x-axis) in mouse liver mtDNA The individual ribonucleotides are shown in different colours according to their base identity (rAMP – red, rCMP – purple, rGMP – green and rTMP/rUMP – blue). The strands are separated by light strand (positive) and heavy strand (negative). (Bottom) A pie chart representation of the proportion of each of the rNMPs in murine liver (n=3), brain (n=2) and heart (n=1) mtDNA and for comparison, the base content of the primary mtDNA sequence, chrM.

3.7.1 There is No Observed Base Bias in Ribonucleotides in Wild-Type Nuclear DNA

In contrast to the mtDNA, analysis of nuclear DNA reads present in the same EmRiboSeq libraries showed that there is no appreciable bias towards a particular rNMP in the nuclear genome of murine liver or that of the RER-proficient mouse embryonic fibroblasts (MEF) (Figure 20).

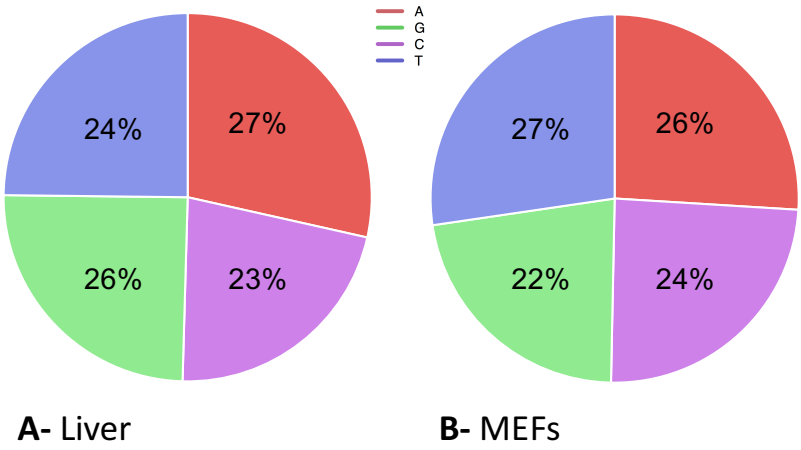


Figure 20: There is no skew to rAMP in murine nuclear DNA.
The proportions of the four rNMPs in the nuclear DNA of (A) murine liver and (B) mouse embryonic fibroblasts.

This relative equality in the mouse nuclear genome is of note, given there is not yet, to the best of my knowledge, any published data which examines the steady state level of nuclear ribonucleotides in a wild-type mammalian background (where RNase H2, and therefore the RNase H2-mediated RER pathway, is active). Analogous studies in *S.cerevisiae* strains deficient for RNase H2 found that rGMP and rCMP were incorporated more frequently than expected from the G+C content of the yeast nuclear genome (Koh et al., 2015). As evidenced from this study, this phenomenon is not mirrored in the wild-type mouse nuclear DNA. This is either reflective of the extremely low levels of retained ribonucleotides or background signal in the wild-type strain caused by random fragmentation of DNA during library preparation. Despite this, the data potentially highlights the differences in the DNA polymerases acting in

nuclear DNA replication between yeast and mouse, or the repair capabilities of the different organisms.

However, from this alone one cannot conclusively state that there is no skew in ribonucleotide incorporation in the nuclear genome as this observation could just be the natural background of the technique. This is supported by the fact that the observed base frequencies are roughly equal to one another but not concordant with the base bias of the nuclear genome itself (Figure 21).

Interestingly, there is a dramatic change in the base frequencies upon loss of RNase H2B in the nuclear genome, as would be expected from RNase H2's role in the nucleus (Figure 21) (Ding et al., 2015, Sparks et al., 2012, Williams et al., 2017). However, the role of RNase H2 in the nucleus is not the focus of this study and only limited conclusions can be drawn from this data with regard to the nuclear genome, as these data were from mitochondrially-enriched EmRiboseq libraries and it is not known if there are any intrinsic biases on any residual nuclear DNA in such preparations.

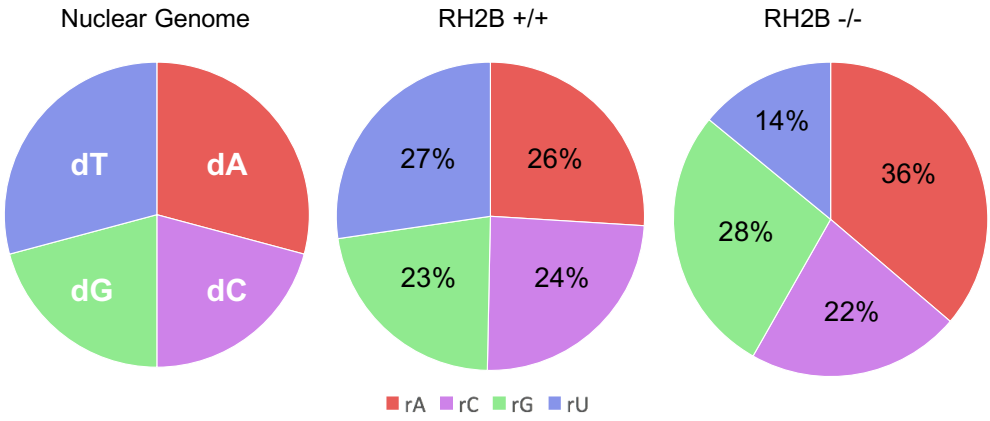


Figure 21: Ribonucleotide incorporation relative rates in nuclear DNA from RH2B proficient and deficient MEFs.

The proportion of ribonucleotides identified in nuclear DNA from p53 knockout MEFs with (+/+) and without (-/-) RNase H2B. There is an increase in the relative abundance of rAMPs upon loss of RNase H2B in the nuclear DNA.

3.7.2 Predictive Calculations of Ribonucleotide Incorporation Relative Rates in Mitochondria

The DF of a polymerase is the degree by which it is able to distinguish between the correct deoxyribonucleotide and its ribonucleotide equivalent to avoid ribonucleotide (mis)incorporation. The DF is an intrinsic feature of all DNA polymerases. As aforementioned, ribonucleotides are the most common non-canonical (mis)incorporated in DNA and yet their incorporation is unfavourable. Therefore, polymerases are adapted to evade ribonucleotide incorporation using steric gates to exclude the extra hydroxyl group which is present on the ribose sugar. But as evidenced by the requirement for RER and the presence of embedded ribonucleotides in DNA, the system is not perfect and ribonucleotides are frequently incorporated during DNA synthesis. And so, the DF of the active DNA polymerase influences the rate of ribonucleotide incorporation during DNA synthesis.

The observed base bias of embedded ribonucleotides in mtDNA (rA>rG>rC>rT/U – Figure 19) is consistent across sequenced mtDNA from wild-type mouse liver, brain and heart. This base hierarchy agrees with previous *in vitro* data from Kasiviswanathan *et al.* whereby the DF of Pol γ was determined using the human exonuclease-deficient (D198A/E200A) recombinant protein, suggesting that the DF of the mitochondrial DNA polymerase has a direct impact on ribonucleotide incorporation (Kasiviswanathan and Copeland, 2011).

Another important factor influencing ribonucleotide incorporation is substrate availability. During the S-phase of the cell cycle, intracellular dNTP concentrations are highest, and therefore the ratio of [NTPs]:[dNTPs] is at its lowest point. This ratio influences ribonucleotide incorporation rate as it determines the exposure of the DNA polymerase to the substrate for DNA synthesis at the replication fork. Consequently, if there is a decrease in dNTPs (and therefore an increase in the ratio of [NTPs]:[dNTPs]), ribonucleotide incorporation becomes more likely. dNTP pools fluctuate throughout the cell cycle according to demand, and the same can be said for mitochondrial concentrations which are largely dependent on dNTP import from the cytosol.

Using the DF of Pol γ in combination with mitochondrial dNTP and NTP quantifications from Wheeler *et al*. (measured in *Rattus norvegicus*) it is possible to create a formula to justify the observed frequencies of ribonucleotide incorporation in murine liver mtDNA (Table 3). These calculations are based on the assumption that ribonucleotide incorporation is a direct result of both polymerase error and substrate availability.

The predictive values shown in Figure 22 consider both the fidelity of the polymerase and the precursor mitochondrial pools, together with the base-content of the mouse mitochondrial genome. The values are very similar to the real base-biases from EmRiboseq, validating the conclusion that the identity and frequency of ribonucleotide incorporation in mtDNA is determined by the mitochondrial [NTP]:[dNTP] and the DF of Pol γ. The formulae used to calculate the hypothetical incorporation frequencies of each ribonucleotide base (x) per mtDNA molecule are shown below:

ii) If the DF were the sole determinant of ribonucleotide incorporation:

$$= \frac{No.of\ x\ bases\ in\ a\ mtDNA\ molecule}{DF\ of\ Pol\,\gamma}$$

iii) If the [NTP]:[dNTP] ratios were the sole determinant of ribonucleotide incorporation:

$$= No\ of\ x\ bases\ in\ mtDNA\ molecule \times [NTP]:[dNTP]$$

iv) a composite of ii) and iii):

$$= \frac{[NTP]:[dNTP]}{DF} \times no.of\ x\ bases\ in\ mtDNA\ molecule$$

The absolute values obtained are hypothetical and were used to estimate the relative proportions of the ribonucleotide bases throughout the genome and so are compared as a fraction of the total predicted ribonucleotides in Figure 22. It is important to note

that the DF values are obtained from an exonuclease deficient Pol γ. This is in accordance with the conclusions drawn from Berglund *et al.* who found that there was no change in ribonucleotide incorporation in the mutator mouse background (exonuclease-deficient Pol γ), demonstrating that Pol γ does not repair (mis)incorporated ribonucleotides by proof-reading (Berglund et al., 2017).

A

	DF
dA/rA	9,300
dC/rC	6,600
dG/rG	1,100
dT/rU	77,000

B

rN	Approx. conc nmol/mg	dN	Approx. conc pmol/mg	*rN/dN
rA	4.78 ± 2.93	dA	2.4 ± 1.88	1,800
rC	0.05 ± 0.05	dC	6.35 ± 5.43	8.49
rG	0.49 ± 0.36	dG	18.7 ± 14.9	25.3
rU	0.11 ± 0.06	dT	1.49 ± 1.85	50.9

* Ratios calculated using the maximum rN and dN values

Table 3: A) The DF of Pol γ and B) the [NTP]:[dNTP] ratios within mitochondria

(A) The human (exonuclease deficient) Pol γ DF as reported by Kasiviswanathan *et al.* (B) The mitochondrial NTP and dNTP concentrations in rat liver reported by Wheeler *et al.* and the calculated [NTP]:[dNTP] ratios.

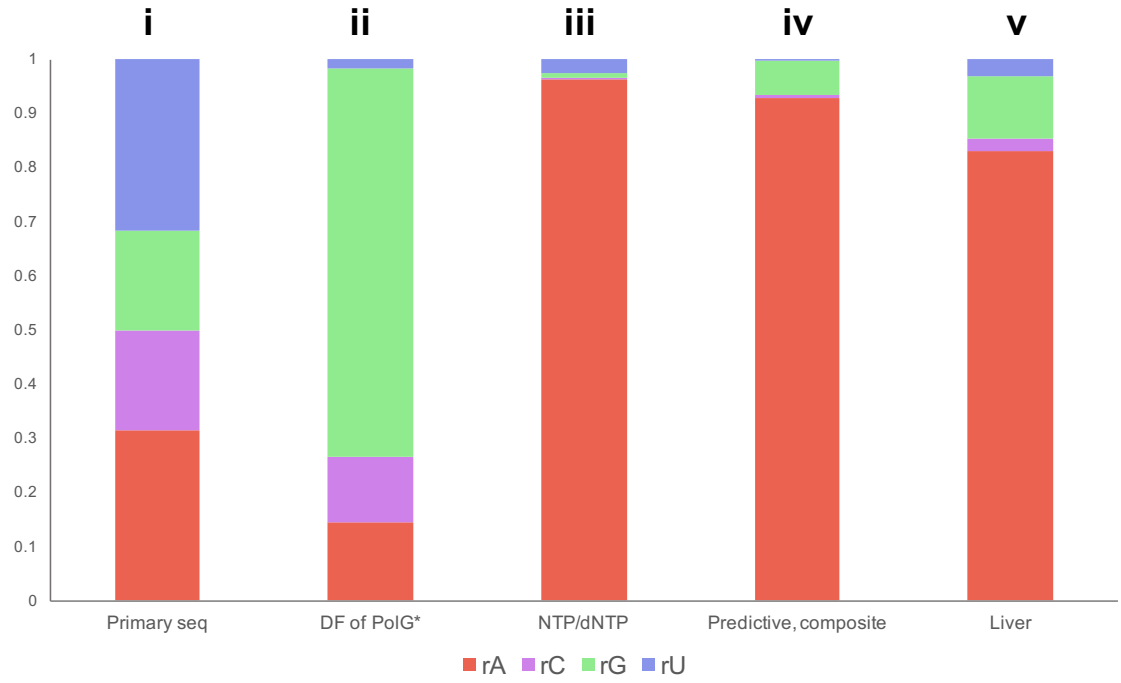


Figure 22: Predicted (calculated, see Table 3) proportions of individual base ribonucleotide incorporation versus EmRiboseq data from mouse liver for each individual base.

Data expressed as fraction of total ribonucleotides. i) the proportional rNMPs values predicted from the primary mtDNA sequence, represented as a stacked bar chart, alongside the expected proportion of each rNMP if the sole determinant were either (ii) the DF of the mitochondrial DNA polymerase (Pol γ) (derived from (Kasiviswanathan and Copeland, 2011), (iii) the [rNTP]:[dNTP] ratio in mitochondria (derived from {(Wheeler and Mathews, 2011), Table 3B}), both adjusted for the mtDNA sequence; or if (iv) rNMP incorporation was a function of both the Pol γ DF and the [NTP]:[dNTP] ratio (composite) which is remarkably similar to the observed EmRiboseq values for liver (v).

3.8 Flanking Deoxynucleotides Influence rNMP Incorporation

A ribonucleotide may be more or less likely to be incorporated in DNA depending on the preceding dNMP or the following dNTP and this trinucleotide context was shown to be influential in determining the overall ribosubstitution rate *in vitro* (Kasiviswanathan and Copeland, 2011). Moreover, analyzing trinucleotide frequencies can be informative when assessing the activities of different DNA polymerases, as shown in Figure 17.

Trinucleotide context analysis of the EmRiboSeq datasets indicated some biases for and against particular flanking dNMPs leading to the following conclusions: i) ATP is more often incorporated after dCMP or dTMP residues; ii) there is a general bias against incorporating the same base after a rNMP (e.g. AaA, TcC, TgG, TuT); and iii) triplet symmetry is favoured in a number of cases (CaC, TaT, AcA, AuA and GcG) (Figure 23). Although rUMP was most likely to be followed by dAMP, this should be seen in the context of the very low numbers of rUMPs in murine liver mtDNA; whereas the triplets CaC, TaC and TaT that were 1.6-1.8 times more frequent than expected represent hundreds more sites of ribosubstitution because of the high levels of rAMPs in the mtDNA. Although repair systems may contribute to the observed biases, no mechanisms have been identified that allow targeted removal of embedded ribonucleotides from mtDNA; therefore, these features are attributed to Pol γ.

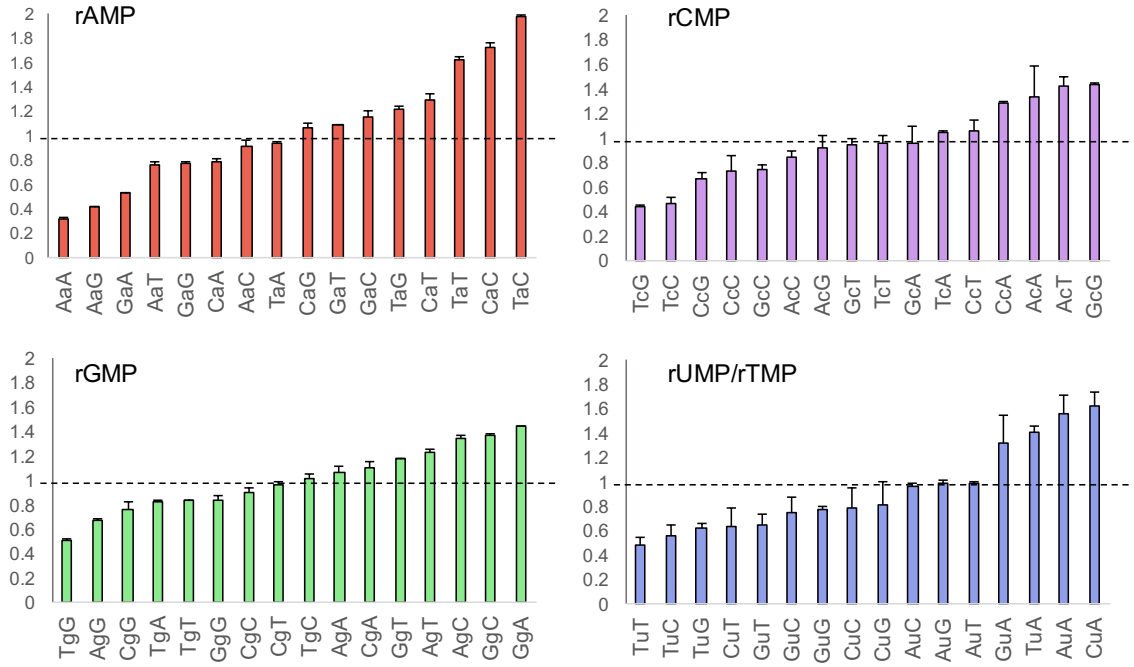


Figure 23: Trinucleotide context analysis to assess the frequency of dNMPs flanking each embedded ribonucleotide for murine liver mtDNA (n=3).

The frequency of each ribonucleotide (rNMP) with respect to all the possible flanking dNMPs (64 triplets) was normalized for the occurrence of each triplet sequence in the murine mitochondrial genome and expressed relative to the mean (set to 1) of all 16 possible triplets for each rNMP, where the lower case central base represents the ribonucleotide. The range from least to most frequent triplet was almost five-fold for rAMP and rUMP, but less than three-fold for rGMP. Error bars represent the standard deviation of 3 biological replicates.

3.9 There is No Evidence for Classical Ribonucleotide Excision Repair in Mitochondria

It has long been assumed that mitochondria lack the capacity for RER due to the absence of RNase H2 (Reijns et al., 2012). However, many of the other enzymes involved in nuclear RER are present in the mitochondrial compartment; such as FEN1 and Lig3. It has also been postulated that the mitochondrial DNA topoisomerase, Top1mt, can recognise embedded single ribonucleotides and such activity is well documented for Top1 in the nuclear compartment in the absence of RNase H2 (Huang et al., 2015, Huang et al., 2017, Kim et al., 2011, Williams et al., 2013). Therefore, mammalian mitochondria theoretically possess the ability to excise ribonucleotides. Before dismissing this possibility, it was important to verify that RNase H2 was not present in mitochondria. This was done using immunocytochemistry analysis of protein localisation and cell fractionations (Figure 24) of cells deficient for RNase H2B (gifted from the laboratory of Dr Andrew Jackson, Edinburgh University).

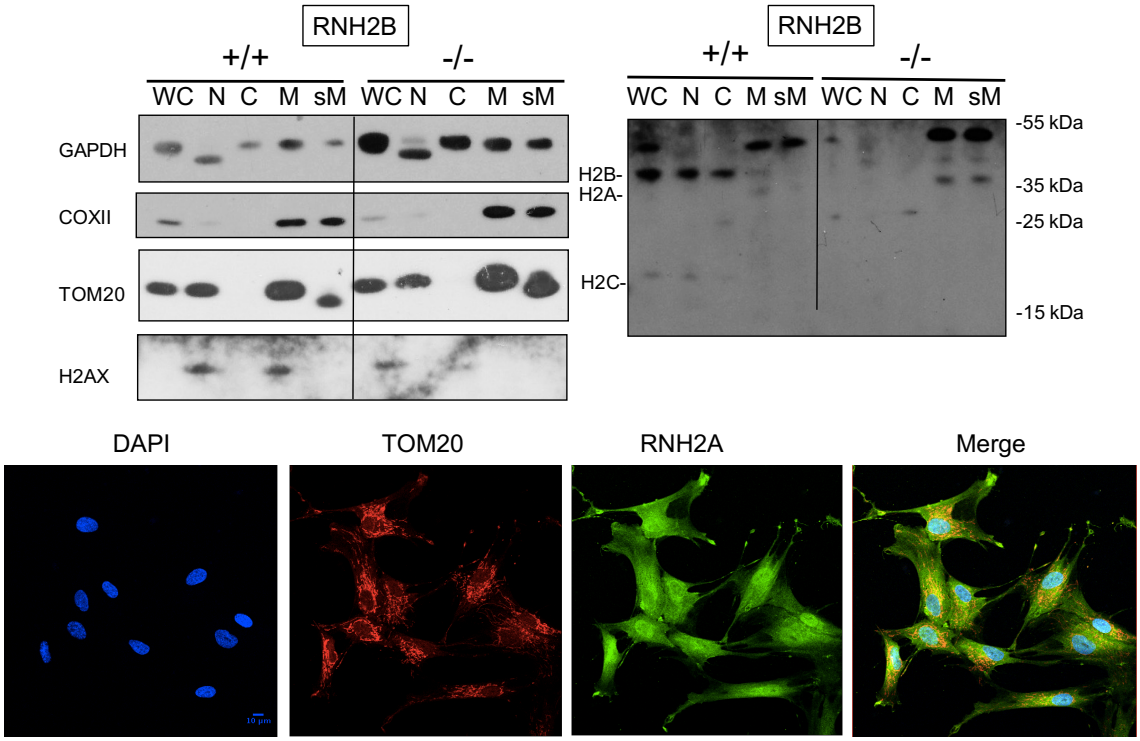


Figure 24: (Top) cell fractionation of RNH2B +/+ and -/- MEFs (p53 -/-) and immunoblotting. (Bottom) immunocytochemistry of wild-type MEFs.

(Top) Cellular fractionation of p53 deficient MEFs where WC = whole cell lysate, N = nuclear, C = cytoplasm M= mitochondria, sM = shaved mitochondria (+ trypsin).

Cells were prepared via sucrose gradient purification and samples extracted at different stages throughout the procedure (see methods). H2AX serves as a control for the nuclear compartment, TOM20 for the outer mitochondrial membrane and COXII inside the mitochondria. No clear RNase H2 signal was observed inside the mitochondrial fraction. The sizes of the RNase H2 subunits are; RNH2A 33.5 kDa, RNH2B 34.7 kDa, RNH2C 17.8 kDa. (Bottom) confocal images from wild-type MEFs stained with fluorescent antibodies to mark the mitochondrial network (TOM20) and the A subunit of RNase H2. From the merged image, there appears to be no co-localisation of TOM20 and RNH2A under standard conditions. [This experiment was done in collaboration with Dr Romina Durigon who performed the immunoblotting and mitochondrial shaving.]

These data are not conclusive as the cells were assessed under standard proliferating cell culture conditions, which as discussed in greater detail later in the thesis, involves little ribonucleotide incorporation in mtDNA, thus not invoking a high demand for RER. So there remains the possibility, that RNase H2 is recruited to the mitochondria under certain conditions.

Despite this, when looking at the RNase H2B knockout cells in quiescence, where there is an accompanying decrease in cytosolic dNTP production, in the knockout there is both a concomitant decrease in proteins associated with mtDNA maintenance (for example; the alternative splice variant of Twinkle; Twinky (Spelbrink et al., 2001), which is absent in the knockout cells), and an increase in top1mt. This suggests that in the absence of RNase H2B there is an alteration in mtDNA metabolism which is likely a downstream effect of increased nuclear genomic instability. Whether this is related to RER in mitochondria is not clear, as many of the proteins, such as Fen1, also have a role in the nucleus. Thus, it may be that altered expression of nuclear DNA factors in response to RNase H2B deficiency includes factors that are shared with the mitochondria, and that the remodelled gene expression conflicts with the usual changes in mtDNA metabolism that accompany exit from the cell cycle. In this scenario, RNase H2B has no direct role in mitochondria, but instead has downstream effects on the organelle owing to increased nuclear genomic instability. In any case, there is no evidence of an increase in ribonucleotide incorporation, nor patterns of ribonucleotide incorporation in mtDNA in RNase H2B deficient cells (Figure 25 and Figure 26B).

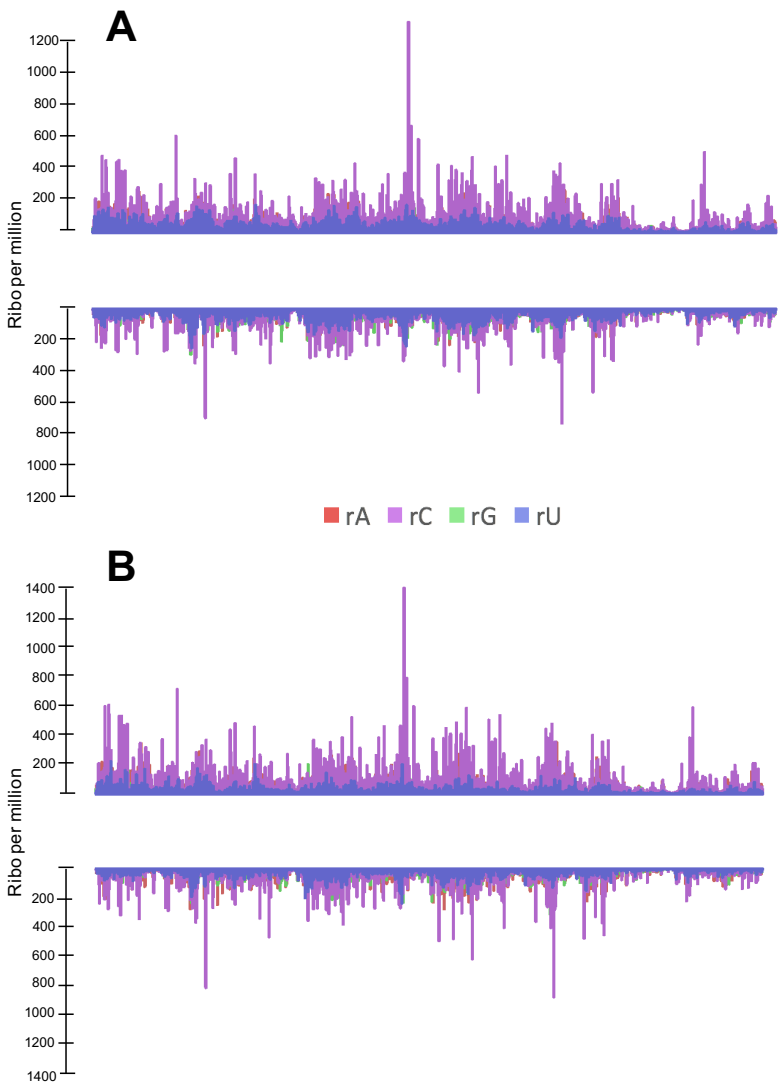


Figure 25: Ribonucleotide profile of A) RNase H2B wild-type and B) RNase H2B knockout MEFs in mtDNA identified by EmRiboseq. Ablation of RNase H2B has no effect on ribosubstitution in mtDNA.

The frequency of ribonucleotide incorporation in mtDNA is plotted as ribonucleotide per million along the y-axis and the genome position 1 to 16,299 nt on the x-axis. The base-identity of each ribonucleotide is identified by its colour as highlighted in the figure legend. [Bioinformatic analysis by Dr Martin Taylor]

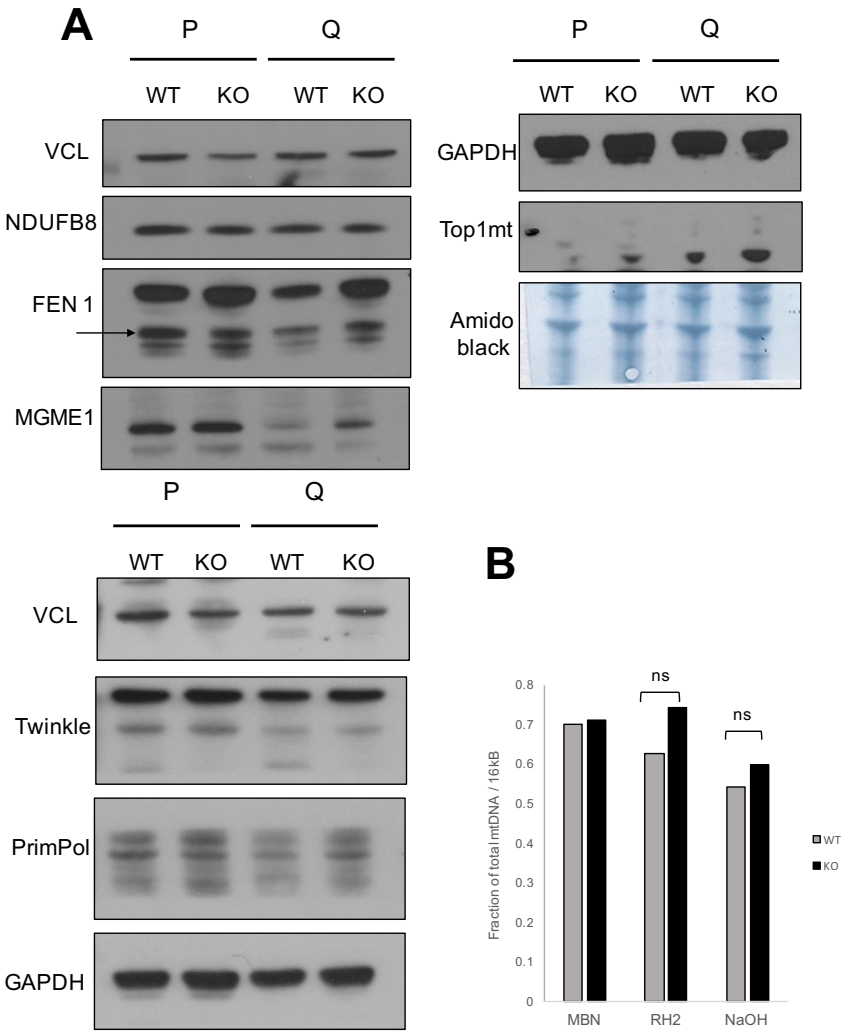


Figure 26: A) Steady state level of mtDNA maintenance proteins in RNase H2B wild-type (WT) and knockout (KO) MEFs. B) Quantification on Southern blot analysis of mtDNA from RNase H2B wild-type and knockout MEFs.

(A) Western blot of steady state level of proteins from cultured p53 deficient MEFs with (WT) or without RNase H2B (KO). (**B**) Quantification from Southern blot analysis of isolated mtDNA from WT and KO MEFs treated with RNase H2 and alkali *in vitro*. The fraction of resistant 16 kB molecules are shown. Proliferating (P), quiescence (10 days serum restriction, Q)

3.10 Results I: Concluding Remarks

The mammalian mtDNA molecule contains persistent ribonucleotides. This is a consequence of the absence of a canonical RER mechanism within the mitochondria. Using a ribonucleotide directed NGS method, EmRiboseq, I mapped the sites of embedded ribonucleotides in mouse mtDNA. It was demonstrated that there is

significant regional and base-bias in ribonucleotide incorporation. Moreover, ribonucleotide incorporation is largely attributable to the mitochondrial compartmental ratio of NTP:dNTPs, the DF of the mitochondrial replicative polymerase; Pol γ, and the base content of the genome itself.

Concordant with the conclusion that mitochondria do not possess the ability to excise ribonucleotides, I found no changes in ribonucleotide incorporation rate in mtDNA from RNase H2B ablated MEFs, nor any RNase H2 in mitochondria from immunoblotting and immunocytochemistry experiments.

Chapter 4. Results II: The Role of dNTP Metabolism in MtDNA Maintenance

4.1 Background

MtDNA depletion syndromes (MDS) are a branch of mitochondrial disease and these severe autosomal recessive disorders have a broad spectrum of symptoms and are typically characterised by a decreased mtDNA copy number in clinically affected tissues. MDS are categorised by the affected tissue and are most commonly identified as hepatocerebral, encephalomyopathic, myopathic or neurogastrointestinal. It is proposed that a reduction in mtDNA leads to reduced synthesis of respiratory chain subunits which results in insufficient energy production in affected tissues. This can lead to organ failure which is why such conditions are often fatal.

MDS are commonly caused by mutations in genes associated with mtDNA maintenance; either the synthesis of dNTPs or mtDNA replication. So far, there are at least 12 different nuclear genes known to be directly linked to mtDNA maintenance and MDS. Hepatocerebral forms of MDS (mtDNA depletion in the liver and brain) have been associated with mutations in the *POLG, PEO1* (Twinkle), *DGUOK* and *MPV17* genes (Uusimaa et al., 2014).

Besides depletion of mtDNA, aberrant mtDNA maintenance can lead to accumulation of tissue-specific mtDNA deletions. Both depletion and deletions in mtDNA lead to severe mitochondrial dysfunction, yet unlike depletion which manifests in infancy, mtDNA deletions tend to develop later in life, causing adult-onset phenotypes typically in skeletal muscle or brain (Nishino et al., 1999, Spelbrink et al., 2001, Van Goethem et al., 2001, Ronchi et al., 2012, Tyynismaa et al., 2012, Alston et al., 2013).

This chapter focuses on a specific mouse model of mitochondrial disease, which is the *Mpv17* knockout mouse. This model is used to examine the role of mitochondrial dNTP metabolism in mtDNA integrity with a focus on ribonucleotide incorporation.

Using the EmRiboseq technique this chapter aims to uncover how perturbation of dNTP pools in mitochondria influences ribonucleotide incorporation in mtDNA.

4.1.1 MPV17

In humans, MPV17 is a protein encoded by the *MPV17* gene comprising of 176 amino acids, located at p21-23 on chromosome 2. It is ubiquitously expressed and the mouse homologue is the kidney disease gene *Mpv17*. MPV17 belongs to a family of membrane proteins in mammals (others include; PXMP2, MPV17, MP-L, and MPV17L2) and is the genetic orthologue of the SYM1 protein found in yeast.

In 2006 Spinazzola *et al.* demonstrated that the human MPV17 protein was imported to the mitochondria and tightly imbedded within the mitochondrial inner membrane (Spinazzola et al., 2006). In accordance with this conclusion, the protein contains four hydrophobic residues, indicative of transmembrane domains (Spinazzola et al., 2006). The same is true for the orthologue protein SYM1 (Trott and Morano, 2004). The protein itself is 20 kDa and has an amino acid sequence akin to a typical channel protein. In support of this, using protein secondary structure predictions Antonenkov *et al.* identified five α-helical regions in the MPV17 sequence which are long enough to penetrate a lipid bilayer. These *in silico* calculations were supported by *in vitro* experiments with recombinant MPV17 whose circular dichroism (CD) spectra demonstrated that the α-helices were preserved (Antonenkov et al., 2015). Despite ample evidence that MPV17 is a mitochondrial IMM protein, the function of the protein is still unknown.

4.2 Ribonucleotide Incorporation in Different Post-Mitotic Tissues

It was demonstrated in the previous chapter that ribonucleotide incorporation in mtDNA is largely influenced by the mitochondrial concentrations of dNTPs and NTPs. And subsequently, in mtDNA from solid tissues, rAMP is the most common (mis)incorporated ribonucleotide due to high ATP concentrations within the organelle (Figure 19). Mitochondrial ATP concentrations are dependent on oxidative phosphorylation and ultimately the energy demand of the cell, which varies between different cell and tissue types. The same can be said for the other nucleotides which

serve as metabolites for other reactions within the organelle, for example GTP which is also used to drive key reactions such as ribosome assembly (De Silva et al., 2015). Therefore, a pertinent question is whether these differences in energy demand and metabolism in various tissues have an impact on the ribonucleotide incorporation patterns in mtDNA.

To address this question, the gross levels of embedded ribonucleotides in mtDNA from various tissues were analysed using Southern Blot analysis of mtDNA molecules incubated with alkali *in vitro.* As ribonucleotides are alkali sensitive, fragmentation of the DNA is indicative of embedded ribonucleotides. The smaller fragments indicate more frequent alkali-sensitive sites. When examining the gross ribonucleotide levels in mtDNA from different murine tissues using 1D-AGE, there were only minor differences (on average, across the tissues examined, approximately 70-80% of mtDNA molecules were sensitive to alkali hydrolysis- Figure 27B). It does appear that mtDNA from skeletal muscle is more sensitive to alkali hydrolysis than the other tissues, yet after 200 mM NaOH for one hour, a comparable fraction of sensitive and insensitive molecules remain.

Despite only mild differences in alkali sensitivity, there were marked differences in the topology of the mtDNA molecules from different tissues (Figure 26A). It is not clear from this technique alone what these differences represent. MtDNA topology could be linked to the transcription and replication activity within the organelle as strand separation alters the supercoiled structure of mtDNA. It is also worth noting that there were surprisingly few multimeric species in heart mtDNA, which is known to be highly catenated in humans (Pohjoismäki et al., 2009). However, it is important not to over-interpret the importance of these species as their presence is variable, and could be attributed to technical elements of the southern blot procedure. In order to evaluate these species further, different complementary approaches would have to be employed, such as; gel extraction of the multimeric species prior to subsequent analysis.

The topological differences could also be linked to segregation (Akman et al., 2016) or copy number (Figure 27). Accordingly, qPCR analysis of mtDNA copy number shows significant variation among tissues. MtDNA copy number in cells from heart and muscle is the highest (as would be expected from high energy demand tissues), whereas copy number in total brain as well as isolated eyes and hippocampus is lower compared to the control tissue, liver (Figure 27C). This indicates that a low copy number correlates

with a greater fraction of multi-catenated, higher order structures, which could be indicative of replication rates in different cell types, where more open mtDNA structures facilitate high mtDNA replication. Nonetheless, there is no evidence for significant differences in ribonucleotide incorporation in post-mitotic tissues, refuting the hypothesis that differences in energy demand or metabolism have a significant impact on the frequency of ribonucleotide incorporation in mtDNA.

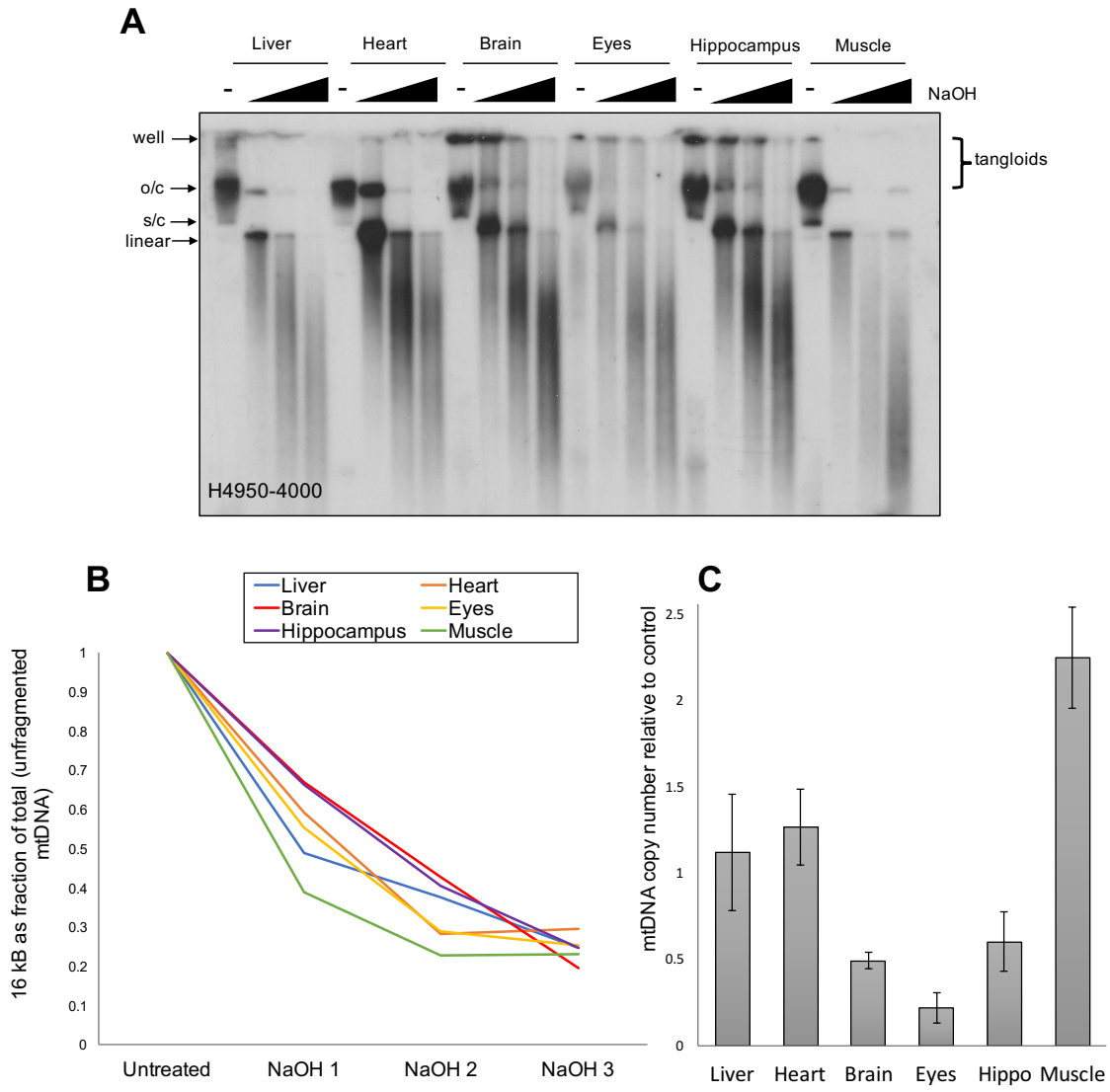


Figure 27: Alkali fragmentation of mtDNA from solid murine tissues.

(A) 1D-AGE of alkali treated mtDNA from solid tissues. Lane 1: untreated, lane 2: 50 mM NaOH 1h, lane 3: 100 mM NaOH 1h, lane 4: 200 mM NaOH for 1h (all incubated at room temperature). (B) Gel quantification shows little difference in ribonucleotide content in mtDNA from the solid tissues examined- evidenced by similar extent of fragmentation (~80% of mtDNA molecules are NaOH sensitive at the highest NaOH concentration used). (C): Copy number variation amongst post-mitotic tissue relative to liver mtDNA (n=4).

4.3 Ribonucleotide Incorporation Rate is Lower in mtDNA from Proliferating Cells

Although there were no gross differences in ribonucleotide levels in mtDNA across different tissues, the question remained whether there would be a difference in proliferating cells where dNTP concentrations are much higher due to the high demand for dNTPs to replicate the nuclear DNA. Initially this was examined by comparing the extent of fragmentation induced by *in vitro* alkali treatment of mtDNA in cultured cells and post-mitotic tissue. Figure 28 shows that there are significantly more alkali sensitive sites in mtDNA from liver compared with cultured cells (Figure 28).

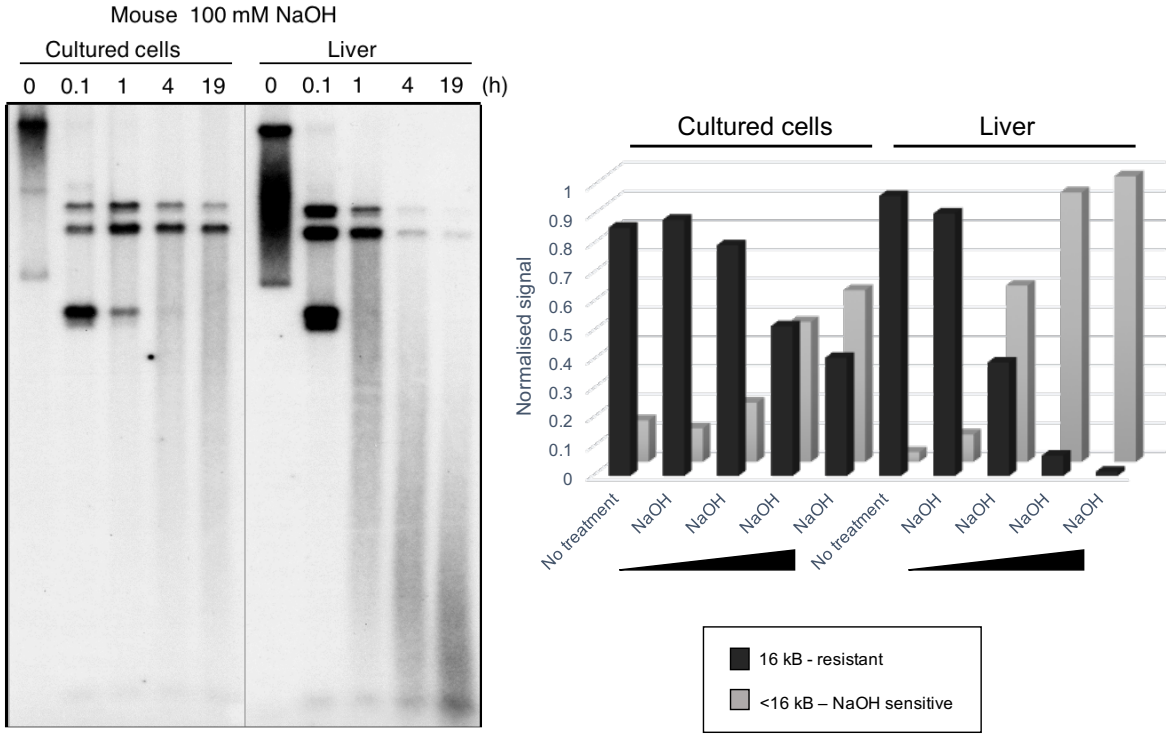


Figure 28: MtDNA of cultured cells is less alkali labile than that of solid tissues.
[Experimental work and image courtesy of Dr Takehiro Yasukawa, analysis carried out by myself]. Riboprobe H15,007-15,805 nt.

The difference in alkali sensitivity between cells and tissue indicates that there are fewer ribonucleotides in mtDNA from cultured cells. This was corroborated by EmRiboseq analysis of cultured cells which revealed a dramatic difference in ribonucleotide distribution and base-specificity in mtDNA between cultured cells and

tissue, in this case MEFs and liver (Figure 29). This supports the conclusion that ribonucleotide incorporation rate in mtDNA is largely determined by the [NTP]:[dNTP] ratio within the organelle.

EmRiboseq showed that rAMP incorporation was dramatically higher in liver compared to cultured MEFs (from 40% in MEFs (n=3) and 83% in liver, (n=3)) Moreover, base-distribution of ribonucleotides in liver mtDNA skewed away from equality compared to cells (Figure 29).

In (the majority of) solid tissues, most cells are no longer dividing and so there is minimal DNA synthesis taking place within the nucleus. Although mtDNA replication is not constrained to the cell cycle, the switch from mitotic to post-mitotic status, and the arrest of nuclear DNA replication, has a dramatic effect on the metabolism of the precursors of DNA synthesis. I have demonstrated that mtDNA from proliferating mouse cells has significantly fewer ribonucleotides than mtDNA from solid tissues (Figure 28) and it can be inferred that rAMPs account for much of the difference, as rAMPs form a much smaller proportion of embedded ribonucleotides in proliferating cells than solid tissue (Figure 29). The difference can, in part, be attributed to the drastic reduction in the total cellular pool of deoxyribonucleotides upon exit from the cell cycle, which in turn has a downstream impact on the size of the pools within the mitochondrial compartment (Ferraro et al., 2005, Reichard, 1988).

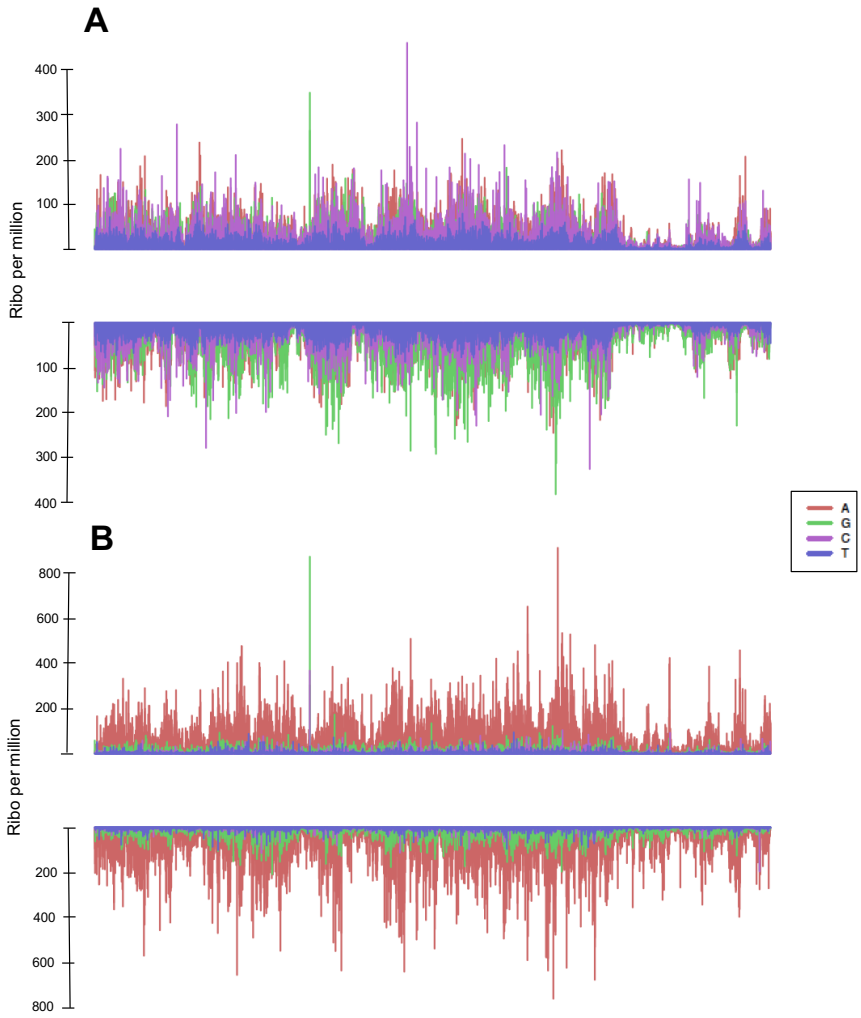


Figure 29: The sites of ribonucleotides in mouse mtDNA from immortalised MEFs (A) and liver (B).

The ribonucleotide profile of incorporation across the genome from 1-16,299 nt on the x-axis and ribonucleotide (ribo) per million on the y-axis. The individual ribonucleotides are shown in different colours according to their base identity (rAMP – red, rCMP – purple, rGMP – green and rTMP/rUMP – blue). The strands are separated by light strand (positive) and heavy strand (negative). [Bioinformatic analysis carried out by Dr Martin Taylor]

4.4 Cell Cycle Exit Results in a Dramatic Increase in rAMP Incorporation in MtDNA

To investigate this phenomenon and further test the hypothesis that the ribonucleotide incorporation rate in mtDNA is directly affected by the mitochondrial dNTP pools, we induced cell-cycle arrest in a cell culture model. Immortalised MEFs

deficient for p53 were grown under standard cell culture conditions and quiescence was induced by serum starvation for 10 days which causes growth arrest and exit from the cell cycle (see methods). MEFs deficient for p53 were used in place of immortalised wild-type MEFs as the latter showed no signs of quiescence under severe serum starvation whereas the p53 -/- MEFs stopped growing and their morphology changed dramatically, as routinely observed when quiescence was induced in primary human fibroblasts by serum restriction (Figure 30).

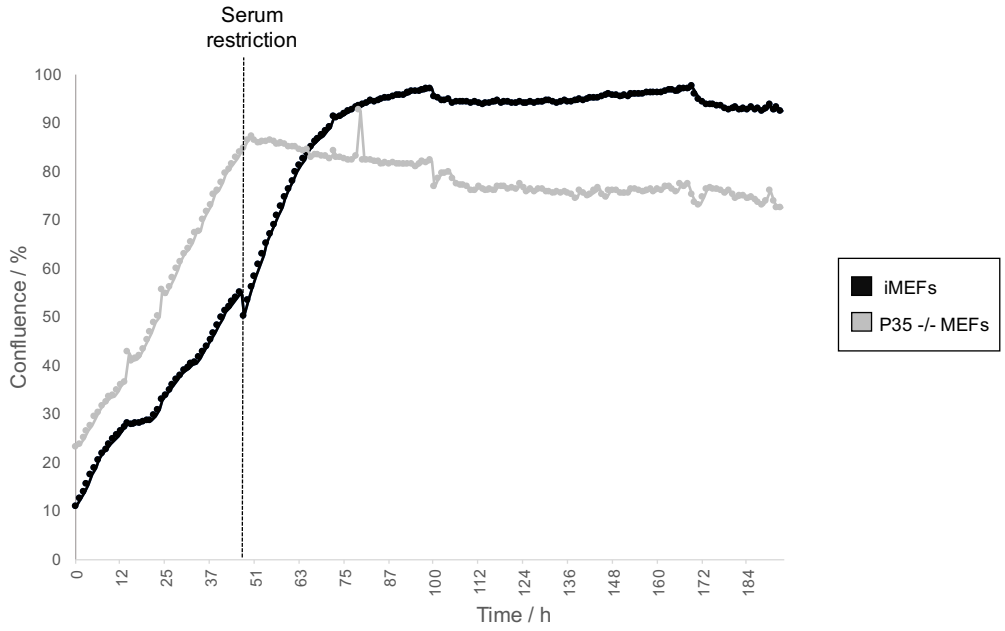


Figure 30: Growth arrest of p53 -/- MEFs with serum starvation.

Cells were plated and cultured in high glucose and 10% serum. Serum was reduced to 0.1% at 48hrs and the growth of the cells was monitored using an incucyte (Essen Instruments). p53 -/- MEFs stop growing almost immediately, whereas immortalized MEFs continue to proliferate until confluence (and can be re-plated and will continue to grow in low serum- data not shown).

EmRiboSeq analysis of quiescent cells (10 days serum-starvation) revealed a dramatic shift in the ribonucleotide incorporation profile, akin to that of post-mitotic cells (Figure 31). This dramatic shift in the ribonucleotide distribution is concordant with the hypothesis that the changes in the nucleotide pools has a direct impact on the mtDNA.

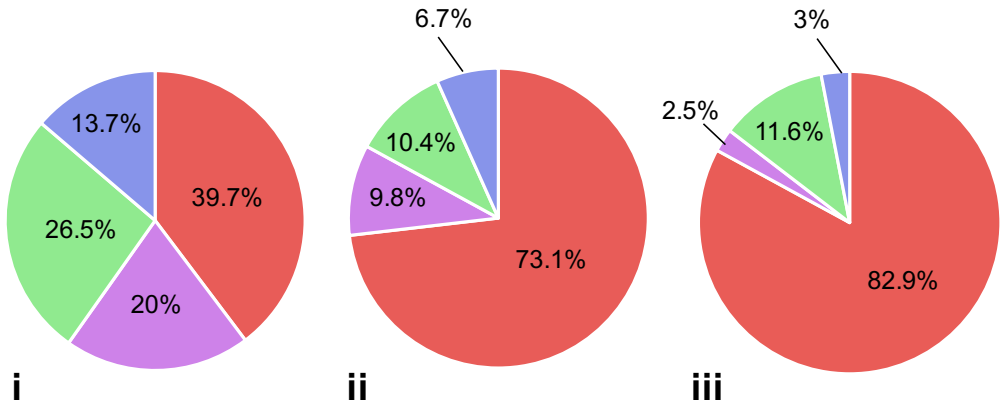


Figure 31: Exit from the cell cycle results in an increase in rAMPs in mtDNA of MEFs approaching that of solid tissues.

Pie charts depicting the proportion of identified embedded ribonucleotides in mtDNA from i) proliferating MEFs, ii) quiescent MEFs (induced by serum starvation – 10days) and iii) mouse liver.

Although this result demonstrates the impact of changing nucleotide pools on ribonucleotide (mis)incorporation rate, it is not clear why there is a substantial disproportional bias in rAMPs in quiescence which isn't observed in proliferating cells.

4.5 Complex I Deficiency and Ribonucleotide Incorporation

As dNTP pools are not the sole determinant of the NTP:dNTP ratios, it was important to investigate the effect of differential ATP production on ribonucleotide incorporation. In order to do this, I used a mouse model, which is deficient for the *Ndufs4* gene. This gene encodes for an 18 kDa subunit of the 45-protein Complex I and its knockout results in severe OXPHOS dysfunction, ultimately leading to premature death. Prior to death the mice are significantly smaller than their wild-type counterparts, exhibit lethargy, blindness, loss of motor control and elevated serum lactate (Kruse et al., 2008).

EmRiboseq analysis of wild-type and *Ndufs4* knockout mouse brain mtDNA revealed no significant difference in the proportion of embedded rAMPs, which suggests that the ATP levels are maintained in the mitochondria of the complex I deficient mice, perhaps via elevated glycolysis or that the concentration of ATP at the replication fork is not a direct function of the total amount of ATP in the organelle (Figure 32).

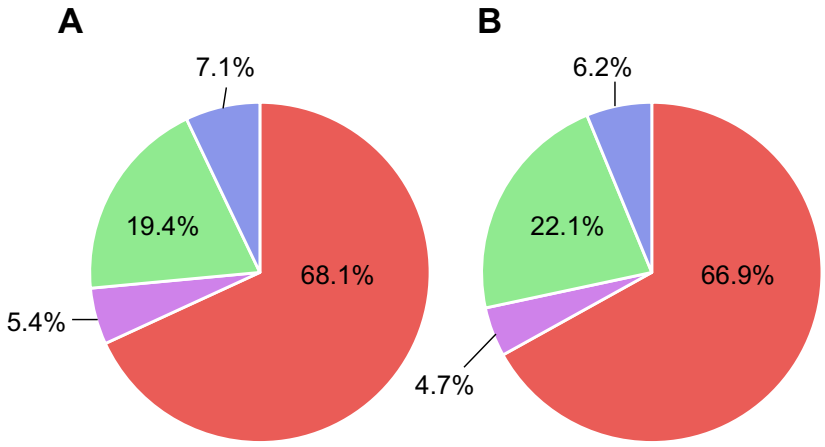


Figure 32: There are no significant differences in ribonucleotide incorporation in Complex I deficient mouse brains compared to controls.

(A) Wild-type mouse brain and (B) Ndufs4 knockout brain mtDNA. The proportional base bias of incorporated ribonucleotides identified by EmRiboseq. [*Ndufs4* knockout mice were gifted from the lab of Professor Jan Smeitink, Nijmegen].

4.6 Mpv17 Deficiency and Mitochondrial dNTP Pools

In proliferating cells, dNTPs are transported into the mitochondria from the cytosol, generated either by *de novo* synthesis or the cytosolic salvage pathway. But in non-dividing cells there is a dramatic reduction in cytosolic synthesis of dNTPs, as there is no longer a demand from nuclear DNA replication, and the mitochondria need to supplement the limited supply of dNTPs from the cytosol by activating their internal salvage pathway (Figure 8) to provide sufficient dNTPs for mtDNA synthesis. This is why the proteins involved in the mitochondrial salvage pathway are so crucial in maintaining mtDNA integrity and mitochondrial function.

The human *MPV17* gene, located on chromosome 2p21-23, is comprised of eight exons and mutations in *MPV17* are known to be linked to two types of common mtDNA abnormalities associated with mitochondrial dysfunction; mtDNA depletion and mtDNA deletions. To date, MDS caused by *MPV17* mutations has been reported in 32 patients with the clinical manifestations including early progressive liver failure, neurological abnormalities, hypoglycaemia and raised blood lactate (Uusimaa et al., 2014). Although the exact function of MPV17 is still unknown, it has been shown that the protein is involved in the maintenance of mitochondrial dNTP pools (Dalla Rosa

et al., 2016). Ablation of *Mpv17* in the mouse leads to perturbations in the mitochondrial dNTP pools and decreased mtDNA copy number in the liver (Figure 33), recapitulating the phenotype of the human patients with MPV17 associated MDS (Dalla Rosa et al., 2016). Decreased dNTP pools are associated with slow mtDNA replication, as shown by the increase in replication intermediates in the Mpv17 mouse knockout liver (Figure 33C). Maintenance of the dNTP pools within the mitochondria is critical for effective mtDNA synthesis and so perturbations in the pools can have a direct impact on mtDNA, both quantitatively and qualitatively.

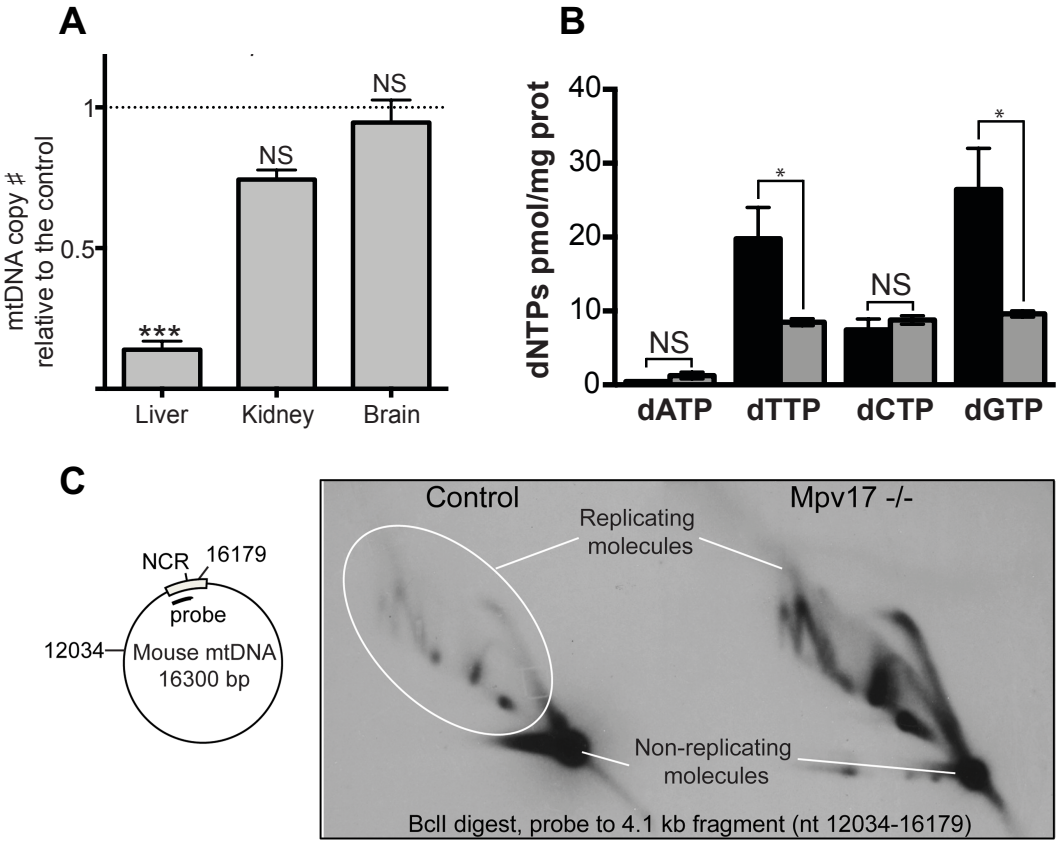


Figure 33: A) MtDNA copy number in *Mpv17* ablated mice relative to controls. B) Mitochondrial dNTP pools in *Mpv17* ablated mouse livers compared to controls. C) *Mpv17* ablation results in a marked increase of mtDNA replication intermediates.

[Figure re-used with permission from the authors (Dalla Rosa et al., 2016). Experimental work and analysis was carried out by Drs Dalla Rosa, Spinazzola and Holt]. **A**) qPCR analysis of mtDNA copy number of liver, kidney and brain in Mpv17 knockout mice relative to the control. **B**) Quantification of mitochondrial dNTPs from wild-type (black) and knockout mice (grey) livers **C**) Analysis of mtDNA replication intermediates in the liver of WT and Mpv17-/- mice. MtDNAs from six *Mpv17* knockout mice and six wild-type livers were isolated, digested with *BclI* and fractionated by 2D-AGE and blot hybridized to a probe to the major non-coding region (NCR) of the murine mitochondrial genome as indicated schematically.

Because of the significant decrease in dNTP pools, the Mpv17 mouse provides an excellent model with which to study the effect of altered NTP: dNTP pools on mtDNA metabolism and replication. Moreover, this is an opportunity to assess if changes in ribonucleotide incorporation might have implications for mitochondrial disease.

4.7 MtDNA Mutational Load in MPV17 Deficient Tissues

Cellular dNTP pool sizes are regulated by multiple mechanisms, both allosteric and genetic, and perturbations in the dNTP pools are known to have mutagenic consequences (Mathews and Song, 2007, Mathews, 2014). Thanks to *in vitro* studies the predominant mechanisms of mutagenesis have been identified as i) misinsertion (i.e. incorrect base pairing formed due to an excess or deficiency of a specific nucleotide base), or ii) inhibition of proof-reading (next-nucleotide effect) whereby dNTP excess encourages rapid DNA extension, before the polymerase can correct the error (Song et al., 2005, Mathews and Song, 2007).

There are multiple instances of nuclear-encoded proteins which when silenced or knocked-out have an impact on mitochondrial dNTP pools (Kiliç et al., Leshinsky-Silver et al., 2011, Ronchi et al., 2012). In the case of loss of function mutations in the gene encoding thymidine phosphorylase (TYMP), the changes in nucleotide pools cause an increase in mtDNA mutations (Nishigaki et al., 2003), a phenomenon which has been well-established *in vitro* for other cases of nucleotide pool asymmetry (Marchler-Bauer et al., 2005). It is therefore important to establish if there is an increase in mtDNA mutations in the absence of Mpv17, and whether this is contributing to the impaired OXPHOS capacity.

To determine the effect of the reduced dNTP pools on mtDNA fidelity, deep sequencing of purified mtDNA from the livers and brains of three pairs of wild-type and Mpv17 deficient mice was carried out using the illumina Truseq technology and single nucleotide polymorphisms (SNPs) were quantified as detailed in the methods. In mouse liver mtDNA, where there is a perturbation of dNTP pools and a reduced copy number, there was no significant increase in the mutational load (Table 4). Brain mtDNA also showed no differences, although in brain the frequency of

mutations in mtDNA were an order of magnitude greater than in liver. Irrespective of genotype, adenine was the most commonly mutated base, and guanine the least. The mutational base-bias was remarkably consistent amongst both tissues and genotypes.

Liver

Run	Genotype	Total bases	Total ML (e-4)	A	C	G	T
1	WT	8.76E+07	4.3	32.3%	24.9%	19.1%	26.0%
1	KO	2.50E+07	3.3	30.3%	26.1%	20.3%	23.3%
1	WT	1.13E+08	3.4	32.4%	22.9%	18.2%	25.3%
1	KO	5.97E+06	5.8	38.3%	20.5%	14.1%	26.4%
2	WT	1.82E+08	9.9	32.2%	26.4%	16.0%	23.8%
2	KO	1.16E+07	22.1	35.0%	22.9%	15.4%	25.6%

Brain

Run	Genotype	Total bases	Total ML (e-3)	A	C	G	T
3	WT	6.61E+08	1.26	30.3%	27.2%	16.3%	25.8%
3	WT	1.51E+08	5.51	30.5%	27.4%	16.3%	26.0%
3	WT	3.19E+08	2.60	30.4%	27.4%	16.3%	25.9%
3	KO	3.76E+08	2.21	30.4%	27.3%	16.3%	25.9%
3	KO	1.02E+08	8.16	30.4%	27.3%	16.3%	25.9%
3	KO	2.73E+08	3.04	30.5%	27.4%	16.3%	25.9%

Table 4: Mutational load in Mpv17 wild-type and knockout 2-month-old mouse mtDNA

It is worth noting that this mutational analysis revealed a dramatic difference in the mutational rates in the two tissues examined. Regardless of genotype, the frequency of mutations in mtDNA were an order of magnitude greater in the brains than in the livers of mice. [Bioinformatic analysis carried out by Lilian Hunt]

This shows that the dramatic reduction in dNTP pools in *Mpv1* knockout mice does not cause an increase in mutational load in the samples examined. This is true for both liver, where there is an observed reduction in copy number and dNTP pools, and brain, where copy number and dNTP levels are normal (Figure 37). However, this does not exclude a possible increase in base mis-incorporation by Pol γ in the Mpv17 knockout. It is possible that repair, either by Pol γ itself or an independent base-excision repair pathway, or mtDNA turnover is sufficient to prevent mutated copies of mtDNA persisting. However, there is no evidence for this in other mtDNA mutator models, nor would it be expected to have any pathophysiological

consequences. Another explanation for the absence of mutations in the knockout is that the base-specific reduction in dNTP pools in the mitochondria results in a roughly equimolar base distribution (Figure 33), opposed to the asymmetric pools in controls, which although reduced, may facilitate correct base pairing and faithful mtDNA replication which in turn explains the slow replication phenotype. This proposed mechanism is the opposite of the next nucleotide effect whereby saturation of the replication fork and the replicative polymerase causes rapid DNA extension without proof-reading. In this instance, reduced dNTP concentrations at the fork and slowed rate of replisome progression facilitates proof-reading and faithful DNA synthesis.

4.8 1D-AGE Comparison of MPV17 Wild-type and Knockout mtDNA from Liver

Incorrect base pairing is only one way that a DNA polymerase can create errors. Another is incorporating the wrong sugar. Ribonucleotide (mis)incorporation is the most common type of polymerase error and ribonucleotides are in vast excess compared to their deoxyribonucleotide counterparts, a phenomenon which is exaggerated in the Mpv17 knockout liver where there is dGTP and dTTP pool depletion. Mpv17 deficiency was therefore predicted to cause a concomitant increase in ribonucleotide incorporation in mtDNA. To assess ribonucleotide incorporation frequencies in the absence of Mpv17, Southern blot analysis was used to compare the alkali sensitivity of mtDNA isolated from the livers of wild-type and knockout mice (Figure 34).

As shown in Figure 34, there is a subtle difference in alkali sensitivity which is evident from the gel quantification, suggesting that there is a mild increase in alkali-sensitive sites in mtDNA from Mpv17 deficient livers, inferred to be due to increased ribonucleotide incorporation in the Mpv17 knockout mtDNA. In addition to this, there are disparities in the mtDNA topology between the wild-type and knockout samples. The mtDNA from the knockout contains a greater proportion of multimeric species which reside in the well of the gel, similar to those observed in brain mtDNA (Figure 27). Interestingly said species are alkali sensitive which indicates that they contain significant numbers of alkali-sensitive sites. Yet, as aforementioned it is difficult to

make conclusive statements about these well-residing species without additional experiments.

Notably, the variation between the two probes in Figure 34 indicates that the alkali-sensitive sites are not uniform across the genome. In order to better understand these differences, it seemed logical to turn to a qualitative approach with the detail conferred by NGS. EmRiboseq analysis enables base-specific analysis of ribonucleotide incorporation which is pertinent in the Mpv17 model, where there is base-specific dNTP depletion.

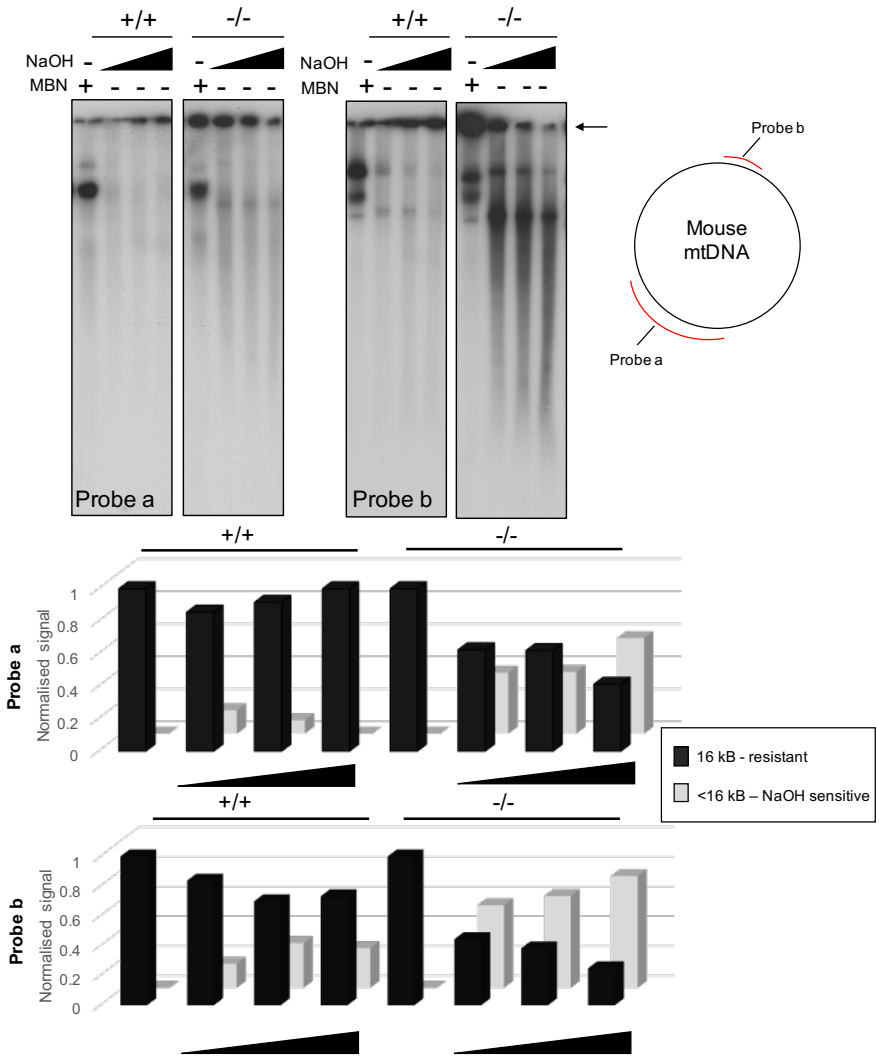


Figure 34: 1D-AGE of mtDNA from wild-type and MPV17 knockout liver treated with NaOH.

Isolated purified mtDNA was treated with MBN (5U, 2') as a control for single-stranded DNA breaks or increasing amounts of NaOH (100mM 30', 200mM 30', 200 mM 1h) and fractionated by 1D-AGE. The DNA was blot hybridized and probed for

two distinct and separate regions of double-stranded mtDNA; **probe a**: 4000-8120 nt and **probe b**: 14881-16299 nt (D-loop containing region). The arrow indicates the well of the gel. The lower panel is the normalised quantification of the southern blots where the signal has been divided into insensitive (16 kB) and sensitive (<16 kB) species to more easily compare the two samples.

4.9 There is a Marked Increase in rGMP Incorporation in MPV17 Deficient Liver mtDNA

EmRiboseq was used to gain a qualitative insight into differences in ribonucleotide incorporation in the absence of Mpv17. Although there is only a mild increase in ribonucleotide abundance in the absence of Mpv17 (Figure 34), EmRiboseq analysis of wild-type and Mpv17 knockout liver mtDNA revealed a dramatic shift in the profile of ribonucleotide incorporation in the absence of Mpv17 (Figure 35). In Mpv17 deficient liver mtDNA there is a significant increase in the proportion of rGMPs embedded throughout the mtDNA molecule compared with the wild-type equivalent. The observation of increased rGMP is specific to the mtDNA as the nuclear DNA showed no appreciable difference between the knockout and wild-type (Figure 36). This is in accordance with the role of Mpv17 in mitochondrial dNTP metabolism, specifically.

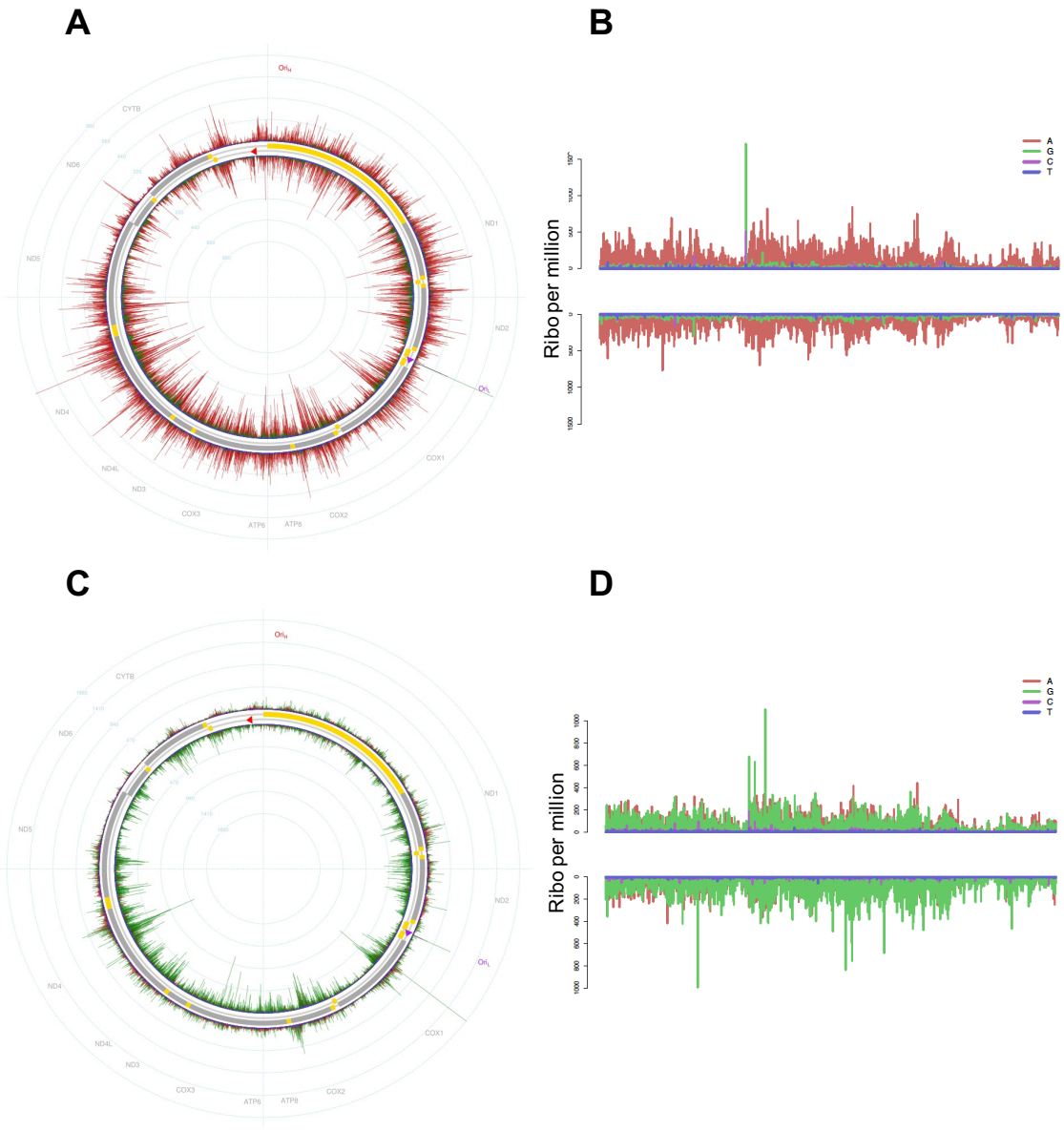


Figure 35: Ribonucleotide incorporation profile in mtDNA from wild-type (WT) and MPV17 knockout (KO) liver.

(**A and C**) Circos plots of the annotated mtDNA genome in Mpv17 wild-type (WT, **A**) and knockout (KO, **C**) mouse liver. (**B and D**) The ribonucleotide profile of incorporation across the genome from 1-16,299 nt on the x-axis and ribonucleotide (ribo) per million on the y-axis for wild-type (**B**) and knockout (**D**) mouse liver mtDNA. Mitochondrial tRNAs and rRNAs are represented by yellow in A and C. [Bioinformatic analysis carried out by Dr Martin Taylor]

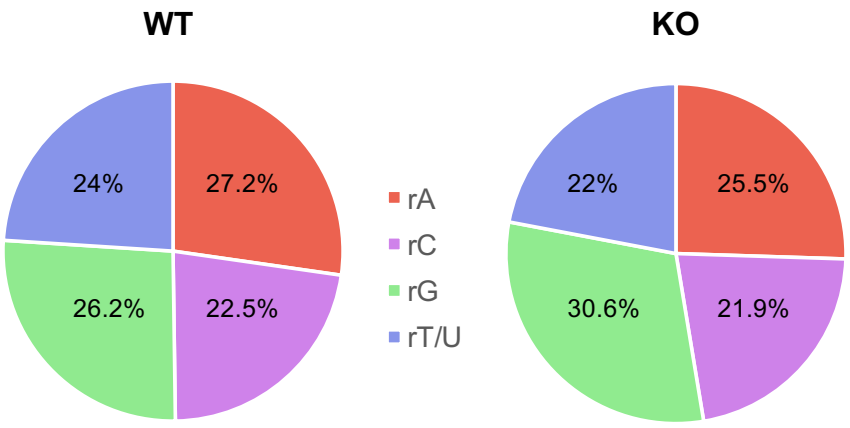


Figure 36: Proportional base distribution of ribonucleotides in Mpv17 wild-type (WT) and knockout (KO) liver nuclear DNA as identified by EmRiboseq reveals no significant differences.

4.10 There is a Marked Increase in rGMP Incorporation in Mpv17 Deficient Tissues Where there is no MtDNA Depletion and Normal dGTP Levels

Analogous to patients with *MPV17* mutations, the knockout mice show liver-specific mtDNA depletion and nucleotide imbalances (Figure 33). In addition to the liver phenotype, patients carrying mutations in *MPV17* also have neurological dysfunction. Thus, it was important to examine the effect of *MPV17* ablation on mtDNA metabolism in the brain.

Both brain and heart from Mpv17-/- mice have normal mtDNA copy number and mitochondrial dNTP concentrations (Figure 37). Therefore, with respect to ribonucleotide incorporation, the prediction would be that there is no change in the ribonucleotide content of mtDNA in unaffected tissues from the knockout mice, in accordance with the conclusion that the ribonucleotide incorporation rate is dependent on NTP:dNTP ratios within the mitochondria. However, EmRiboseq analysis of heart mtDNA and brain mtDNA from wild-type and Mpv17 knockout material revealed a comparable shift to rGMP incorporation in the absence of Mpv17 (Figure 38). The increase of rGMP in heart and brain mtDNA was not as dramatic as the liver but it was substantial nonetheless, further implicating Mpv17 in dNTP

metabolism, especially with respect to guanine, and demonstrating that its function has a very tangible downstream effect on mtDNA content. Therefore, this result demonstrates that the observed increase rGMP incorporation is not simply a consequence of altered GTP:dGTP ratios.

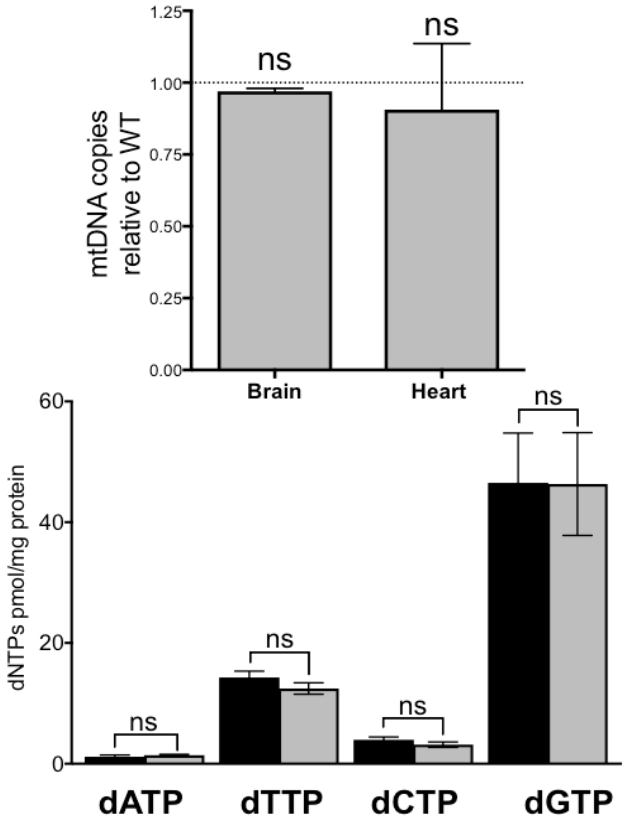


Figure 37: (Top) MtDNA copy number in brain and heart and (bottom) dNTP levels in brain.

The mtDNA copy number from Mpv17 -/- mice is equivalent to that of the wild-type mice in brain and heart tissues and there is no observed nucleotide pool depletion in the brain mitochondria. [Experimental work and analysis carried out by Dr Ilaria Dalla Rosa].

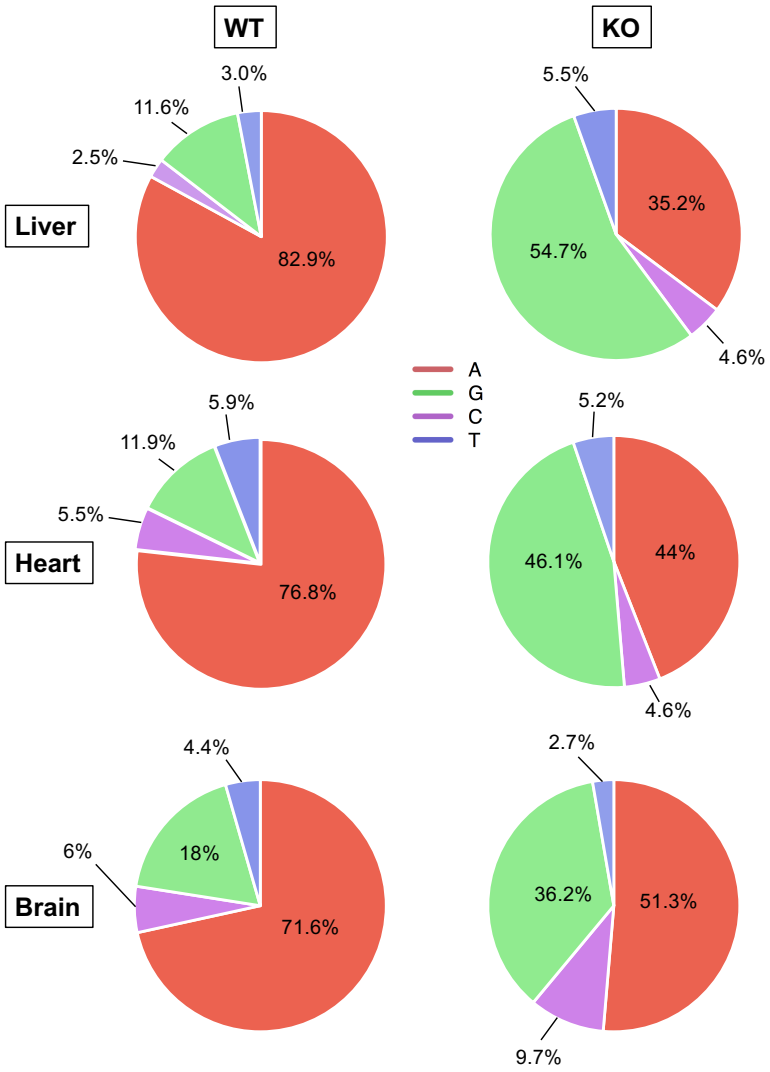


Figure 38: Relative frequency of incorporated ribonucleotides of liver, brain and heart mtDNA of Mpv17 ablated mice.

The relative abundance of each embedded ribonucleotide base in mtDNA of 2-3 months of age determined by EmRiboSeq represented as pie charts; rA – rAMP, rC – rCMP, rG – rGMP, rU – rUMP/rTMP. Percentages were calculated as the fraction of the total number of identified ribonucleotide sites across the mitochondrial genome, and not normalised for base content i.e. frequency of each base within the mtDNA sequence.

Furthermore, Southern blot analysis of wild-type and knockout brain mtDNA revealed a mild increase in alkali sensitivity, suggesting that, like the liver, in the brain there is an increase in ribonucleotide incorporation in the absence of Mpv17 (Figure 39).

2D-AGE analysis of brain material was attempted but using this technique I am unable to make any statements about the abundance of mtRIs as the abundance of such is so low in wild-type steady-state brain mtDNA, unlike the liver.

4.10.1 There are Multiple Deletions in the Mpv17 Knockout Mouse Brain

It has been demonstrated that even in the absence of dNTP or mtDNA depletion there is an increase in rGMP incorporation in the brain material of Mpv17 knockout mice (Figure 38), without an increase in mutation rate (Table 4). It was hypothesized that in the liver, the reduced dNTPs were an adaptation to avoid mutagenic mtDNA replication as equimolarity promotes faithful replication (Song et al., 2005). Thus, it is possible that the dNTP depletion is a downstream feedback effect of increased rGMP incorporation, of which is the primary phenomenon in Mpv17 deficiency.

If increased rGMP incorporation precedes dGTP and dTTP depletion then it is likely that rGMP incorporation causes other problems of mtDNA replication which ultimately result in mtDNA depletion, as in the liver. Total DNA from brains from old (>12 months) Mpv17 knockout mice was analysed using 1D-AGE which revealed the presence of late-onset mtDNA deletions (Figure 39). It is evident that the deletions shown in Figure 39B are not as abundant as other examples of MDS. Long-range PCR was invoked to complement the 1D-AGE analysis but I found it was so sensitive (polymerase stalling could be a result of the embedded ribonucleotides in mtDNA) that ‘deletions’, or more accurately; stalled intermediates, were present even in controls. So, it is important to utilise a NGS method to corrobate the 1D-AGE results.

Deletions are synonymous with replication stalling which is extremely likely if the replisome is encountering more ribonucleotides. Kasiviswanathan and colleagues demonstrated *in vitro* how extension from a 3’ rG terminus (as opposed to a dG terminus) reduces the catalytic efficiency of Pol γ by up to five-fold. Moreover, increasing number of contiguous ribonucleotides drastically hinders Pol γ’s ability to extend from either a 3’ deoxyribonucleotide or ribonucleotide, which *in vivo* would result in reduced replication rate and ultimately, fork stalling (Kasiviswanathan and Copeland, 2011). Therefore, increased rGMP incorporation in the absence of Mpv17

is expected to impede mtDNA replication and thereby offers a credible explanation of the multiple deletions and depletion observed in the brain and liver of the knockout mice.

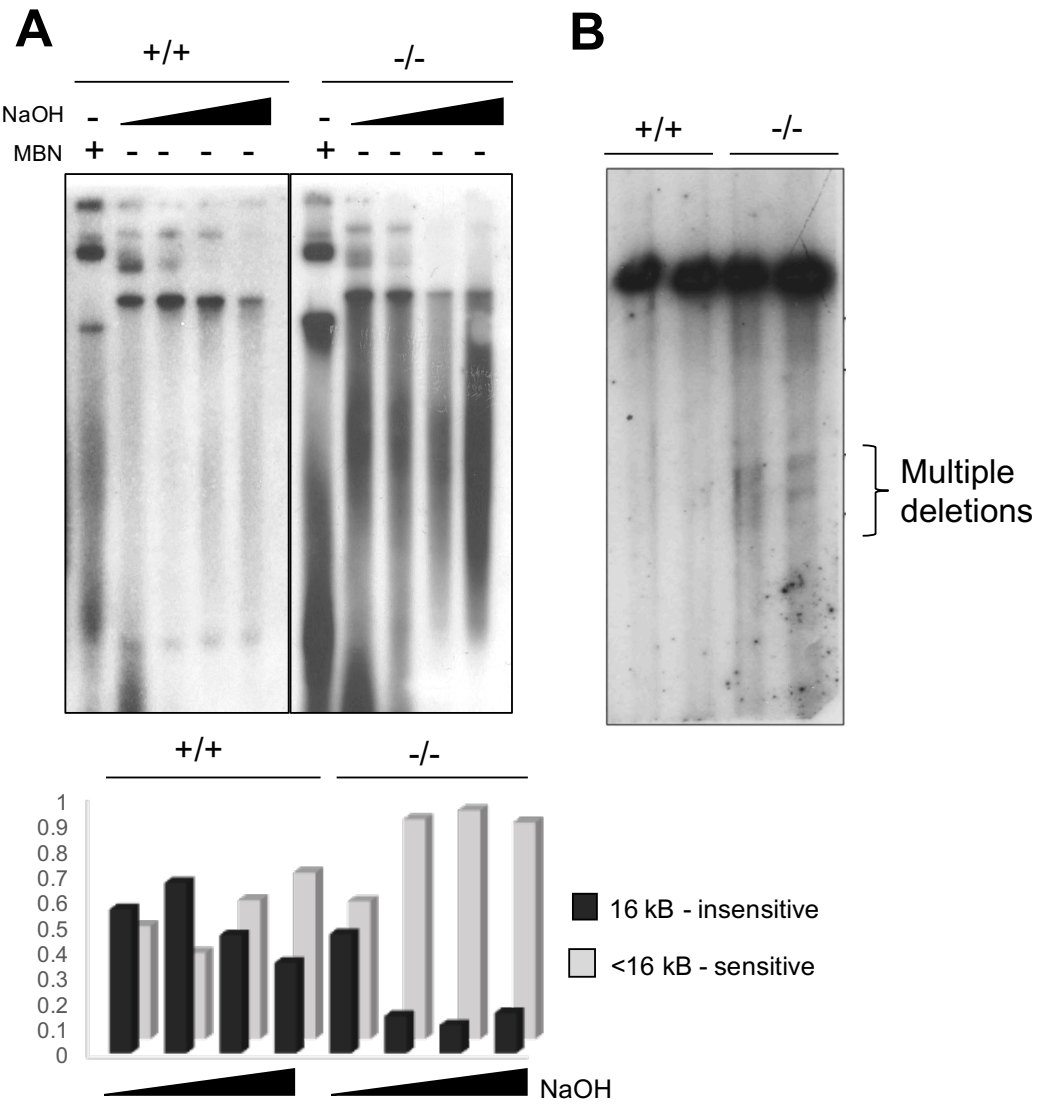


Figure 39: 1D-AGE of mtDNA from wild-type and Mpv17 knockout 12 month-old brain treated with alkali.

(A) Isolated purified mtDNA (2-3 month brain) was treated with MBN (5U, 2') as a control for single-stranded DNA breaks or increasing amounts of NaOH (50mM 1 hr, 100mM 1hr, 200 mM 1hr and 500 mM 1hr) and fractionated by 1D-AGE. The DNA was blot hybridized and probed for the H-strand 14881-16120 nt (D-loop containing region). The lower panel is the normalised quantification of the southern blots where the signal has been divided into insensitive (16 kB) and sensitive (<16 kB) species to more easily compare the two samples. (**B**) Southern blot of brain DNA of mice aged 1 year hybridized with a probe to np 14881-16299 (heavy strand) in Mpv17 wild-type and knockout mice. [Panel B was carried out by Dr Ilaria Dalla Rosa].

It is important to note that the brain is an extremely heterogeneous tissue, containing many different cell types. Therefore, it is difficult to make conclusions about any specific cell type. That being said, all analysis and comparisons made from brain material in this study were carried out on total brain homogenate, so any differences observed are real within the context of all cells contained within the brain.

4.11 Mpv17 Deficiency and Mitochondrial NTP pools

Mass spectrometry is commonly used to identify and quantify dNTPs and NTPs, however the technique relies on expert analysis and even then, it can be difficult to discriminate between various metabolites. Having experienced first-hand the difficulty of working with dNTPs (as they degrade very easily and quickly), I developed an indirect method to measure relative NTP concentrations within the mitochondria. Using mitochondrial dNTP extracts which were analysed in parallel for dNTP quantification (see Methods), the NTPs were quantified using a radiolabelling experiment. In this approach, an RNA oligo is synthesised using commercially available nucleotides, radiolabelled UTP (α-P^{32}) and the isolated mitochondrial nucleotides, whereby a DNA oligo with T7 promoter sequence is transcribed. In order to measure the relative amounts of NTPs, in each individual reaction mixture one of the four commercial nucleotides is omitted and therefore there is a requirement for that omitted ribonucleotide from the mitochondrial extract to synthesise the radiolabelled RNA oligo. The amount of radioactive labelling can then be used as a read-out for the availability of each base within the mitochondrial extracts to compensate for the absence of each base within the reaction mix. Whilst this approach cannot give an exact concentration, it is extremely informative about the relative abundance of ribonucleotides.

This approach revealed that there is no significant difference in the relative mitochondrial ribonucleotide amounts in wild-type and knockout liver and kidney (Figure 40). Given that all tissues of the knockout animal examined to date have elevated embedded rGMPs, one can infer that the increase in mitochondrial GTP concentrations is not the primary cause of the high frequency of embedded rGMPs in the absence of Mpv17.

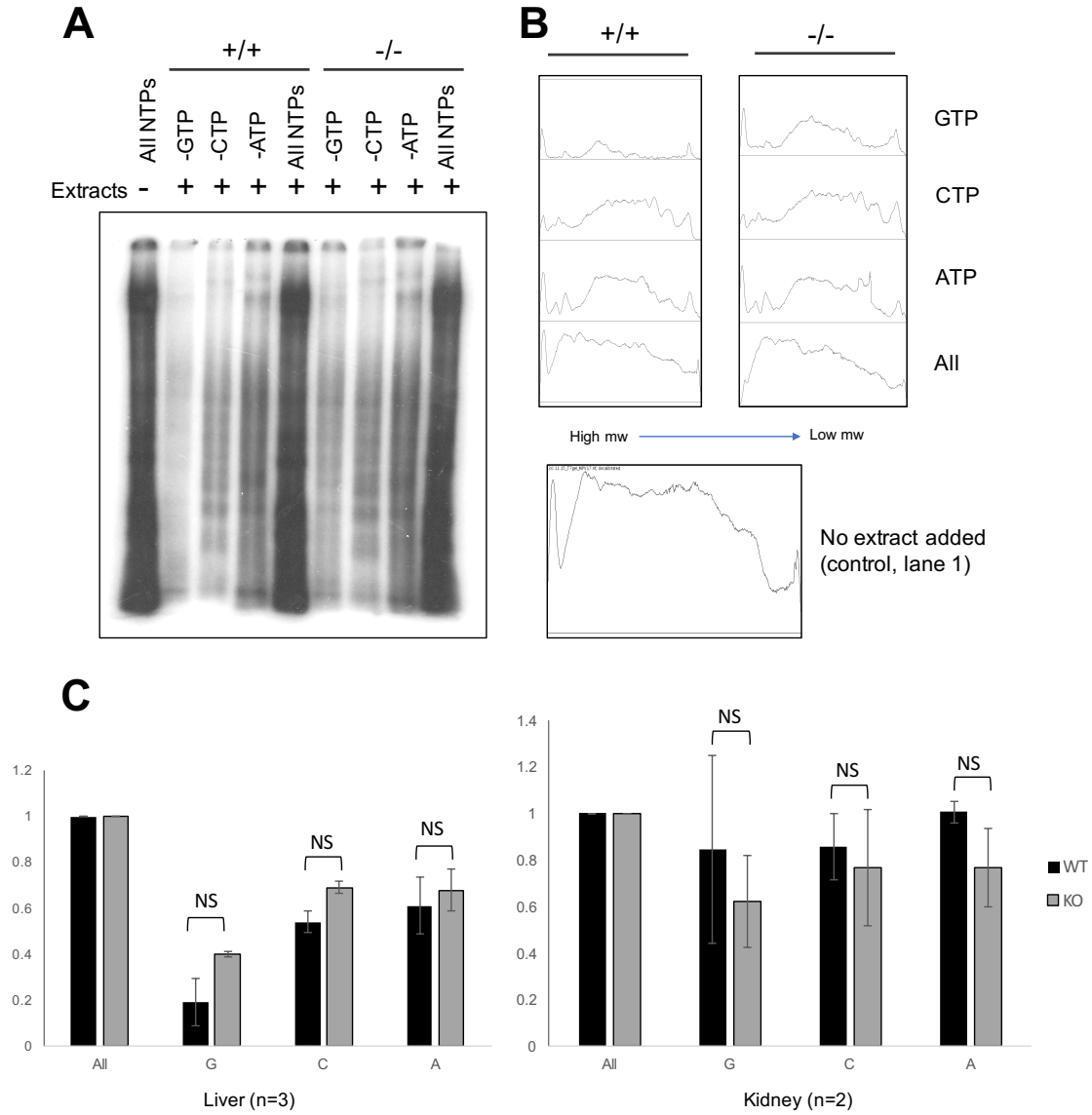


Figure 40: Mitochondrial NTP quantification of Mpv17 knockout and wild-type liver and kidney.

(**A**) A representative gel of radiolabelled RNA using mitochondrial extracts in combination with commercially available nucleotides. (**B**) Gel quantification of signal intensity, the integral of each gel was used to quantify the total signal. (**C**) Relative mitochondrial NTP levels from Mpv17 wild-type (WT) and knockout (KO) mouse liver and kidneys. (None of the differences achieved statistical significance).

4.12 MPV17 is Upregulated in Quiescence

As with many mitochondrial disorders, there are tissue-specific phenotypes, yet the precise rationale for why some tissues are so severely affected whilst others remain

unaffected is not known. Mpv17 is ubiquitously expressed so there is little evidence to suggest that it has a tissue-specific role (Spinazzola et al., 2006).

Cultured patient fibroblasts with mutations in MPV17 show no depletion, nor OXPHOS deficiency until quiescence, where they undergo depletion (Dalla Rosa et al., 2016). Immunoblotting confirmed that Mpv17 expression increases dramatically in immortalised MEFs under serum starvation (Figure 41). The same trend has been shown in human fibroblasts (Dalla Rosa et al., 2016), suggesting a role for Mpv17 in the mitochondrial dNTP salvage pathway. Serum starvation is associated with a dramatic change in ribonucleotide incorporation rate (Figure 29) and this is accompanied by an increase in the steady state levels of the mitochondrial topoisomerase (Top1mt). An outstanding question is whether top1mt, like its nuclear equivalent top1, is able to recognise and excise embedded ribonucleotides (Williams et al., 2013), which could explain its increased expression in line with increased rNMP incorporation and Mpv17 expression.

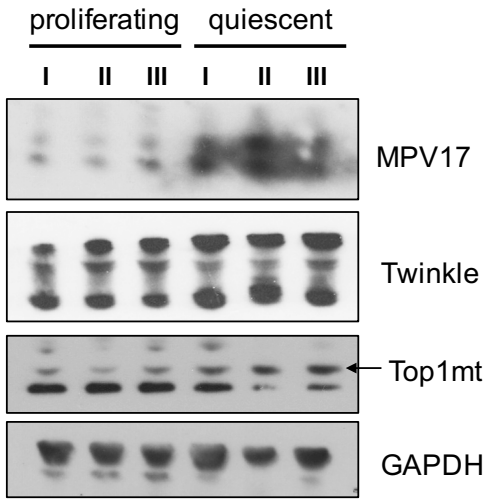


Figure 41: Steady state level of mitochondrial proteins in proliferating and quiescent MEFs

Total cellular proteins were analysed in triplicate from proliferating and quiescent (serum starvation, 10 days) immortalised MEFs.

4.13 Results II: Concluding Remarks

All DNA polymerases (mis)incorporate ribonucleotides despite their preference for deoxyribonucleotides, and analysis of cultured cells indicates that mammalian mtDNA tolerates such replication errors. Here, I've shown that the DNA in mitochondria of solid tissues contains many more embedded ribonucleotides than that of cultured cells, consistent with the former's high ratio of ribonucleotide to deoxynucleotide triphosphates, and that rAMPs are the predominant events. This was corroborated by analysis of quiescent cells which show a dramatic increase in rAMP incorporation, akin to mtDNA from solid tissues. The pattern of ribonucleotide incorporation in mtDNA changes in a mouse model of Mpv17 deficiency, as rGMPs increase markedly, in some cases becoming the major embedded ribonucleotides of mtDNA. However, while mitochondrial dGTP is reduced in the knockout liver, the brain shows no change in the overall dGTP pool, leading me to infer that Mpv17 determines the local concentration or quality of dGTP. Embedded rGMPs are expected to impede DNA replication, and elevated rGMP incorporation is associated with early-onset mtDNA depletion in liver and late-onset multiple deletions in brain of the Mpv17 ablated mice. These findings suggest aberrant ribonucleotide incorporation is a primary mtDNA abnormality that can result in pathology.

Chapter 5. Results III: RNase H1 and Persistent RNA in MtDNA

5.1 Background

In addition to single sporadic ribonucleotides throughout the mitochondrial genome, stretches of RNA are also embedded within the DNA duplex and are known to persist in mtDNA (Brown et al., 2008). These RNA patches can be attributed to residual primers or replication re-start upstream of a DNA lesion. In nuclear and mtDNA, persistence of long stretches of RNA/DNA hybrid impede DNA replication and when left unprocessed lead to genomic instability (Williams et al., 2016, Holmes et al., 2015).

In mammals, RNA stretches of 4 or more ribonucleotides hybridised to DNA in mitochondria and the nucleus are processed by RNase H1. RNases H are part of the nucleotidyl-transferase superfamily and endo-nucleolytically cleave the RNA portion in RNA/DNA hybrids ((Cerritelli and Crouch, 2009, Pallan and Egli, 2008) and the references therein). RNase H activity was first reported in 1969 from calf thymus (Stein and Hausen, 1969) and following years of research the family of enzymes has been divided into two classes depending on their amino acid sequence (Ohtani et al., 1999). Type 1 RNases H (RNase H1/I) recognise longer stretches of RNA/DNA hybrids (> 4 ribonucleotides). In mammals RNase H1 is present in both nuclear and mitochondrial compartments, and in the latter it is essential for mtDNA replication during embryonic development (Cerritelli et al., 2003). Type two RNases H (RNase H2/II/III), as aforementioned, are able to recognise single embedded ribonucleotides, and in mammals, RNase H2 is responsible for initiating RER in the nucleus.

In eukaryotic RNases H1, the N-terminal domain contains the hybrid binding domain (HBD) which is critical for its activity, alongside its catalytic domain and connecting domain (Nowotny et al., 2008). Many, but not all, eukaryotic RNases H1 have a mitochondrial targeting sequence (MTS) upstream of the HBD. RNases H1 are well conserved in eukaryotes, and in both prokaryotes and archaea the domains are maintained but the protein is not required for viability in prokaryotes or unicellular eukaryotes (Kochiwa et al., 2007). In *S.cerevisiae*, RNase H1 can be deleted

showing only modest increases in sensitivity to DNA-damaging agents, suggesting the role of RNase HI is dispensable in *S.cerevisiae* (Arudchandran et al., 2000) and there is no evidence to suggest that it is localised to mitochondria in *S.cerevisiae* to the best of my knowledge (Frank et al., 1999). This ability to survive without RNase HI could be attributed to the ability of *S.cerevisiae* to function without mtDNA (Chandel and Schumacker, 1999). By contrast, ablation of *Rnaseh1* in mice results in embryonic lethality at 8.5 days post-coitus, due to massive mtDNA depletion (Cerritelli et al., 2003). Unlike nuclear DNA, mtDNA replication is initiated after implantation following multiple cell divisions where the maternal mtDNA is diluted (El Shourbagy et al., 2006, Houghton, 2006). In the case of the RNase H1 knockout, the mitochondria are unable to maintain replication of their mtDNA due to retained primers which impede replication, and this ultimately leads to dysfunctional mitochondria.

5.2 Pathological Mutations in *RNASEH1*

The significance of RNase H1 in mtDNA metabolism has been highlighted by recent reports of pathological mutations in *RNASEH1* which cause adult-onset neuromuscular disease (Reyes et al., 2015). The features of the RNase H1-associated pathology are characteristic of mitochondrial disease. The precise missense mutation in *RNASEH1*; c.424G > A; p.Val142Ile (hereafter referred to as V142I) is reported to have a reduced catalytic activity (Reyes et al., 2015) and V124I derived fibroblasts show mtDNA aggregation, reduced mitochondrial translation and impaired respiration (Akman et al., 2016). Contrary to the mouse RNase H1 knockout cells (Holmes et al., 2015), there is no evidence for primer retention nor mtDNA depletion in patient-derived cells, but the V124I fibroblasts instead show an increase in 7S DNA (Reyes et al., 2015) and a corresponding depletion of the R-loop (Akman et al., 2016). These observations strongly implicate RNase H1 as an essential player in mtDNA replication and metabolism.

5.3 Primer Retention in the Absence of RNase H1 in MEFs

The model used to examine mtDNA metabolism in the absence of RNase H1 is a conditional knockout cell line engineered from MEFs (∆RH1 MEFs). MEFs were genetically modified to carry a single copy of RNase H1 that is ablated when Cre-

recombinase is induced by the addition of 4-hydroxy-Tamoxifen (4-OHT). Upon loss of RNase H1, there is an accumulation of mtDNA primers within mtDNA (Holmes et al., 2015). 1D-AGE analysis of mtDNA from 4-OHT+ ΔRH1 MEFs revealed prominent ends at four discrete sites, three of which mapped to previously defined replication initiation sites in the mitochondrial genome; O_H, Ori-b and Ori-L. The fourth at LSP is on the L-strand and is proposed to act as the initiation point for synthesis of the minor arc portion of the L-strand. To date, Southern blot analysis has been used to map these sites using *in vitro* Eco-RNase H1 digestion (Figure 42) (Holmes et al., 2015). Analysis of the 4-OHT+ ΔRH1 MEFs demonstrated that there is a significant amount of replication initiating with primer synthesis at LSP and transition to DNA synthesis at O_H and Ori-b. As both of these primer species accumulate in the knockout it was concluded that RNase H1 is absolutely required for their processing, which under normal conditions it does very efficiently, as they are undetectable at steady state in cells.

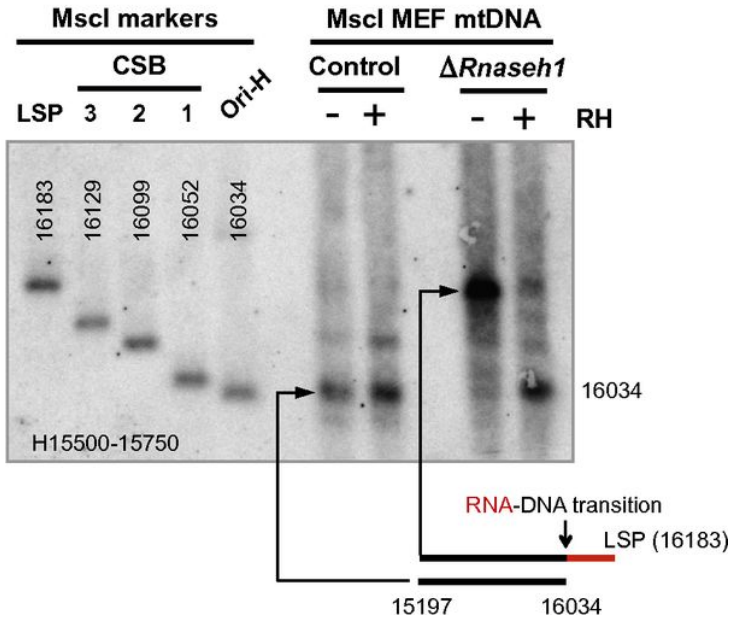


Figure 42: Loss of Rnaseh1 reveals one prominent origin of replication in the NCR and a primer starting at the light strand promoter.

[This image has been re-used with permission from the authors. J. Bradley Holmes *et al.* PNAS 2015;112:9334-9339, Copyright (2015) National Academy of Sciences. (Holmes et al., 2015)]. *MscI* digested and denatured mtDNAs were fractionated by 1D-AGE and hybridized to riboprobe H15500-15750. Where indicated, samples were treated with (+) or without (−) Eco-RNase HI, before denaturation. DNA from mitochondria of MEFs treated without or with 100 nM 4-OHT for 8 days. Interpretations of the RNase HI sensitive species appear below; red lines, RNA; blue lines, DNA.

5.4 The Most Frequent Site of Ribonucleotide Incorporation in Mouse MtDNA Maps to Ori-L

The most abundant primer identified by Holmes *et al.* was at Ori-L, which was present in both wild-type and 4-OHT+ ΔRH1 MEFs samples) (Holmes et al., 2015). In agreement with this finding, EmRiboseq analysis of wild-type mouse mtDNA revealed that the highest site of ribonucleotide incorporation was a single rGMP at 5188 np on the light strand (Figure 43), which is the initiation site for lagging stand replication (Brennicke and Clayton, 1981). This high rGMP peak is the exception to most other sites which are dominated by rAMPs (Figure 19), and is at least an order of magnitude more abundant than any other rGMP site in mouse liver mtDNA (Figure 43A). The prominence of this peak was such that it was strikingly present in all RNase H2-treated libraries.

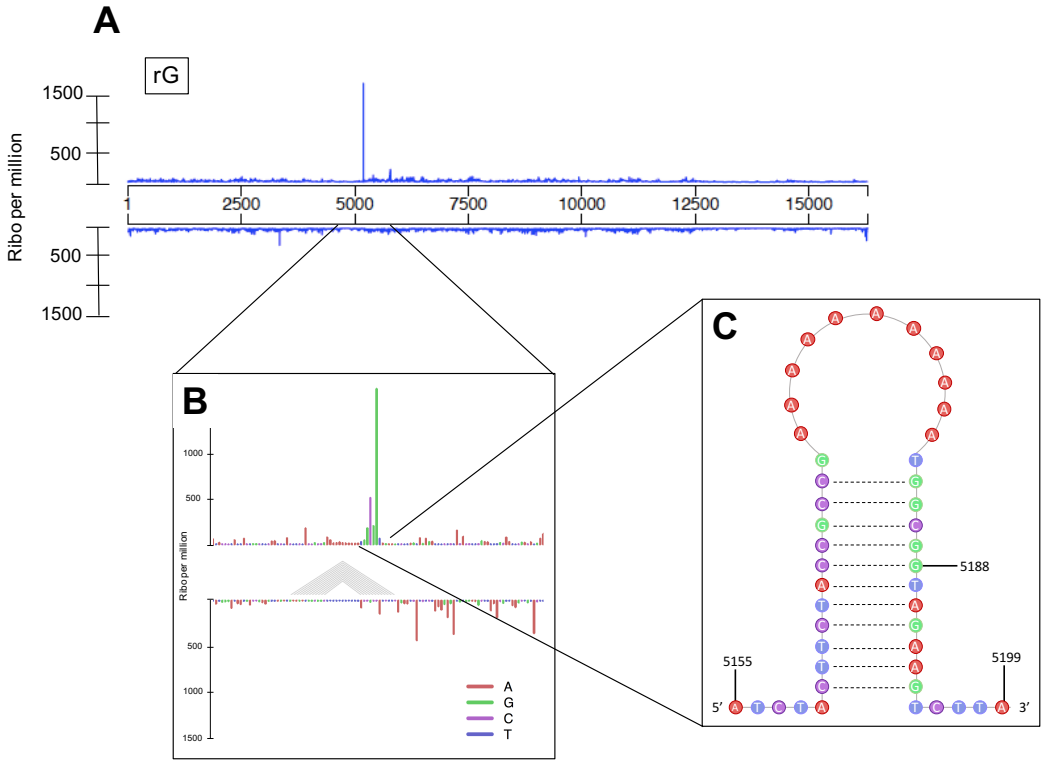


Figure 43: (A) Isolated profile of rGMP incorporation in wild-type mouse liver mtDNA (B) Zoom in on Ori-L region that comprises a conserved stem-loop and poly-T tract used to template light strand RNA primer synthesis. (C) The predicted stem-loop of Ori-L, the origin of light-strand replication.

(A) Based on EmRiboSeq, the single most frequent ribosubstitution of murine mtDNA is a guanosine at np 5,188, which confirms the location of Ori-L, and validates the technique. The example shown is murine liver mtDNA. (C) The position of the

prominent rNMP at np 5188 mapped to the predicted stem-loop structure for Ori-L (Bogenhagen et al., 1979).

5.4.1 Ori-L is Not Processed by RNase H1 *in vivo*

The Ori-L peak is of note, as it is present in high abundance in the wild-type (4OHT-) MEFs and wild-type solid tissues, where RNase H1 is present and active. This is concordant with previous reports that the primer at Ori-L is present in controls, but is more abundant upon ablation of RNase H1 (Holmes et al., 2015). Therefore, unlike O_H and Ori-b, Ori-L is not rapidly processed *in vivo.* Thus, RNase H1 either spares some primers at Ori-L or at this position they are processed less rapidly than those in the NCR, which could be due to association of the protein with the IMM or the replisome.

Due to its abundance across all libraries, it is likely that there is a persistent primer at Ori-L in the majority of wild-type mtDNA molecules, although using this technique alone that conclusion cannot be drawn. This is a logical conclusion based on the fact that the strand-asynchronous mechanisms (both the SDM and RITOLS versions) of mtDNA replication invoke the use Ori-L as the major site of initiation of second strand DNA synthesis, and therefore any mtDNA molecules synthesised via either mechanism would retain a primer at that site, if indeed the RNA is not processed by RNase H1.

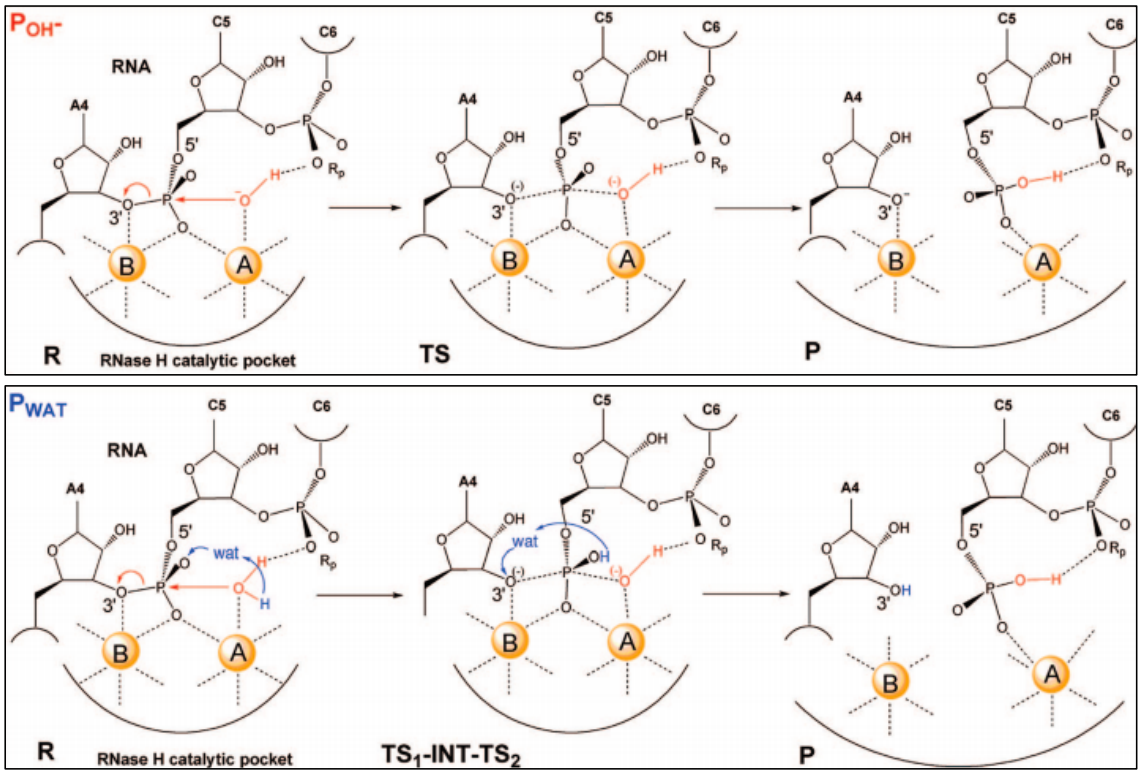
5.5 Primer-directed EmRiboseq Protocol Development

As shown in chapter 3, EmRiboseq is a powerful tool in mapping single embedded ribonucleotides. RNase H2 is also able to nick at sites of longer RNA/DNA hybrids, however discriminating between such sites among single embedded ribonucleotides using EmRiboseq is a challenge. In order to map RNA/DNA hybrids >3 nucleotides, RNase H1 was employed in place of RNase H2 in a primer-directed, modified-EmRiboseq approach.

5.5.1 The Mechanism of RNase H1

RNase H1 activity is not processive and appears to bind and release each ribonucleotide portion of an RNA/DNA hybrid following cleavage. RNase H1 is magnesium-dependent and quantum mechanics and molecular simulations predict

that cleavage of the RNA phosphate backbone occurs in a two-step, two-metal ion catalytic mechanism (Scheme 2) (Rosta et al., 2011).



Scheme 2: Chemical mechanisms for phosphodiester bond hydrolysis catalysed by RNase H.

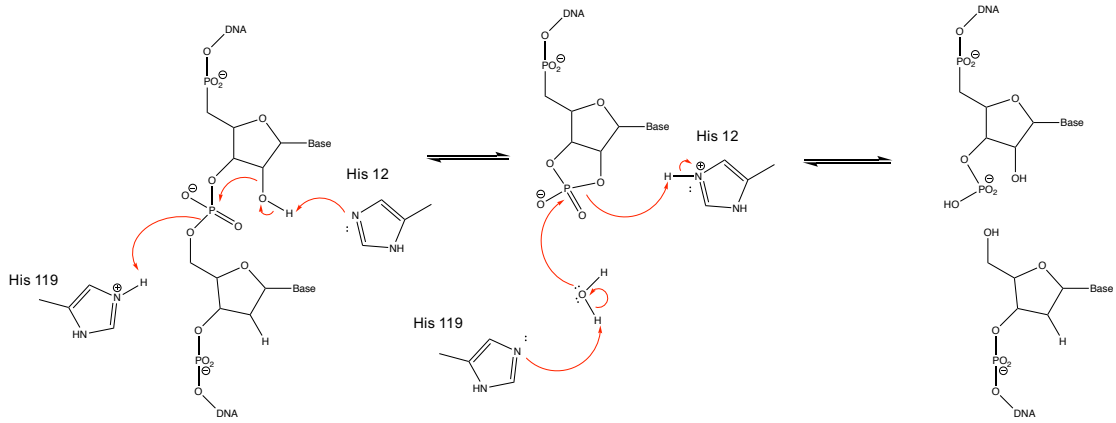
[Reprinted with permission from (De Vivo et al., 2008). Copyright 2008 American Chemical Society.] Two nucleophilic agents are considered: water (WAT pathway, bottom) and hydroxide ion (OH^- pathway, top). An in-line SN2-like nucleophilic attack on the scissile phosphorus atom by the nucleophilic group is followed by the inversion of the phosphate stereo configuration and formation of the 5′-phosphate and 3′-hydroxy function of the RNA strand.

RNase H1 is thought to act mechanistically in a similar manner to RNase H2; nicking 5’ to the RNA site and leaving a terminal 3’-hydroxyl and a 5’-phosphate. This is based on the crystal structure of *Bacillus halodurans* RNase H with the RNA/DNA hybrid substrate, and accompanying molecular dynamics simulations. (De Vivo et al., 2008). However, RNase H1 alone does not remove all the ribonucleotides in a stretch of RNA/DNA hybrid. Therefore, products of RNase H1 hydrolysis would leave residual terminal ribonucleotide(s) at the 5’ end which are likely to hinder subsequent ligation steps within the EmRiboseq protocol. It is for this reason RNase A was used in parallel, to remove the terminal ribonucleotide(s) which would otherwise impede

ligation. RNase A, unlike the RNases H, leaves a terminal 3'-phosphate and a 5'-hydroxyl (Scheme 3).

5.5.2 RNase A has RNA/DNA Hybrid Activity *in vitro*

RNase A is best known for its ability to hydrolyse single stranded RNA. Yet at low salt concentrations (<100 mM NaCl) it is also capable of hydrolysing RNA/DNA hybrids. In order to evaluate its ability to hydrolyse embedded ribonucleotides, isolated mtDNA was incubated with increasing concentrations of RNase A under low salt conditions. Southern blot analysis of isolated mtDNA molecules (containing persistent embedded ribonucleotides) treated with RNase A *in vitro,* induced fragmentation (Figure 44), confirming RNase A's hybrid activity.



Scheme 3: A schematic diagram of RNase A hydrolysis of a single ribonucleotide embedded in the DNA duplex under low salt conditions.

In low salt (<0.1M) the His12 residue of RNase A is sufficiently nucleophilic to attack the labile hydroxyl group of the nucleotide and induce hydrolysis at the 3' end of the sugar in an RNA/DNA hybrid. The intermediate is a 2'3'- cyclic nucleotide intermediate. The product is a 3' ribonucleotide with a phosphate group.

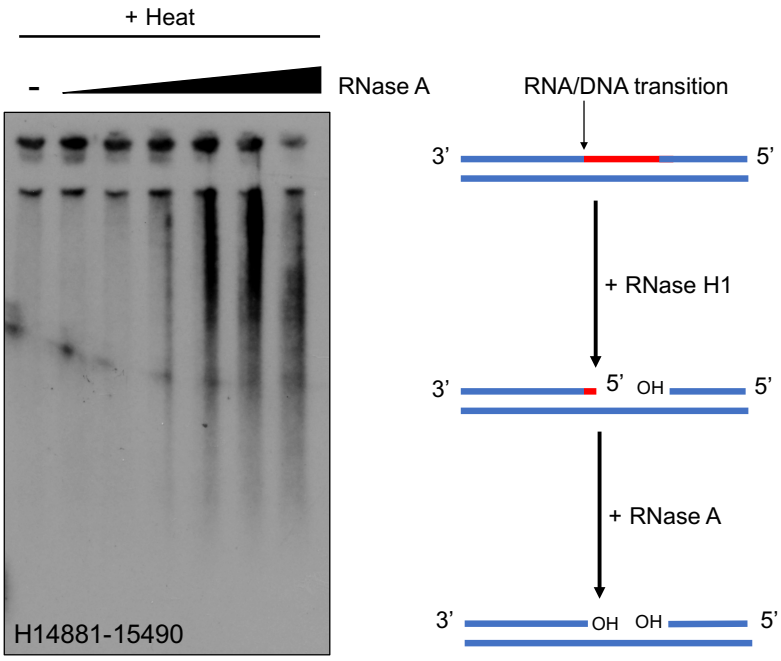


Figure 44: (Left) RNase A hydrolysis of murine liver mtDNA under low salt conditions. (Right) The mechanism of primer-directed EmRiboseq.

(Left) Mouse liver mtDNA was incubated with RNase A (40, 80, 120, 160, 200, 240, 280 ng for 1 hour at 37°C in 50 nM NaCl), heat denatured for 3 minutes at 95 °C and fractionated by neutral 1D-AGE. The DNA was blot hybridized to probe to H-strand nt 14,881-15490. (Right) The proposed mechanism of primer-directed EmRiboseq using RNase H1 and RNase A, resulting in a 3'- and 5'- free hydroxyl.

5.6 Primer Directed EmRiboseq of 4OHT+ ΔRH1 MEFs

5.6.1 Comparison of EmRiboseq Approaches

To confirm that the chemistry behind the protocol alterations was as anticipated, individual peaks were examined to highlight the similarities and differences between the two EmRiboseq approaches (Figure 45). This analysis indicated that the most prominent peaks were largely shared by both treatments and thus are attributable to RNA patches rather than single ribonucleotides. As expected, the peak heights relative to the background are significantly enhanced in the primer-directed libraries due to specificity for >3 contiguous ribonucleotides (Figure 45C and D). The analysis also revealed that the RNase H1 and RNase A combination was effective in removing the terminal ribonucleotide and moreover suggests that RNase H2 is removing all the RNA moieties in the RNA patches, as well as single ribonucleotides.

Whether that is via processive hydrolysis, or repetitive binding and release of the DNA molecule is not known.

As shown in Figure 44, RNase A is also able to hydrolyse single embedded ribonucleotides, and so it might have been expected to produce a ribonucleotide profile broadly similar to RNase H2 EmRiboseq. However, in the primer-directed protocol, the amount of RNase A is very low. It is possible that in the primer-directed EmRiboseq approach RNase A is recognising single ribonucleotides, like RNase H2, but likely has a greater preference for the terminal ribonucleotides left by RNase H1 (which are much more structurally and chemically similar to RNase A's primary substrate; RNA). Another possibility is that RNase A is not doing anything, which cannot be dismissed until libraries treated by RNase H1 alone are investigated.

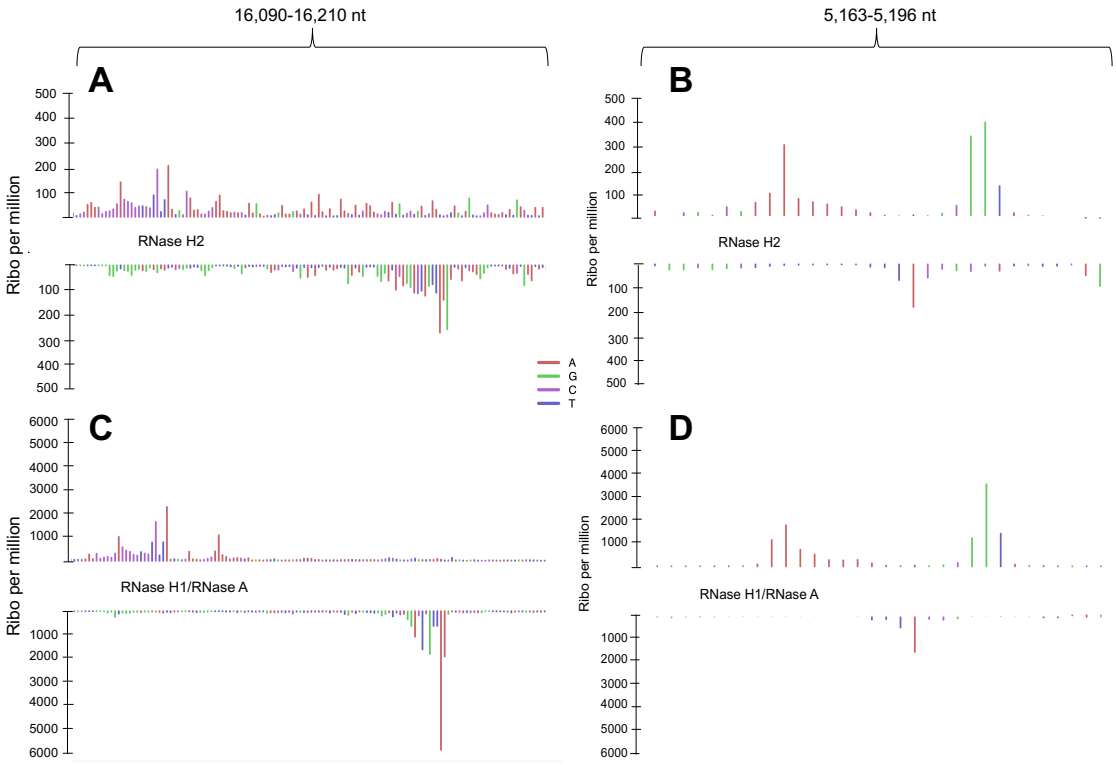


Figure 45: Zoom in of peaks from 4OHT+ libraries prepared with RNase H2 (A and B) or RNase H1 and RNase A (C and D).

(L-strand +, H-strand -) All profiles are of 4OHT+ ΔRH1 MEFs mtDNA. Take note of variations in the scale of the y-axis, which indicate ribonucleotide per million reads. (**A**) 16,090-16,210, RNase H2 EmRiboseq. (**B**) 5,163-5,196 nt RNase H2 EmRiboseq. (**C**) 16,090-16,210, RNase H1 and RNase A primer-directed EmRiboseq. (**D**) 5,163-5,196 nt, RNase H1 and RNase A primer-directed EmRiboseq. [Bioinformatic analysis carried out by Dr Martin Taylor]

5.6.2 Primer Directed EmRiboseq Reveals Persistent Primers in mtDNA from 4OHT+ ΔRH1 MEFs

By comparing libraries created from the same mtDNA samples treated with either RNase H2 (original EmRiboseq protocol) or RNase H1 with RNase A (primer-directed EmRiboseq approach) in parallel, it is clear that the two approaches are distinct and identify different types of ribonucleotide incorporation in mtDNA (Figure 45 and Figure 46). As expected, the modified primer-directed approach (Figure 46C and D) yielded fewer overall sites of ribonucleotide incorporation, as RNase H1 will only hydrolyse a patch of four or more successive ribonucleotides. Just three sites of persistent RNA in control (4OHT-) MEFs were prominent (Figure 46C). Upon loss

of RNase H1 (4OHT+) the relative peak height increased at the same three sites and yielded several additional sites (Figure 46D).

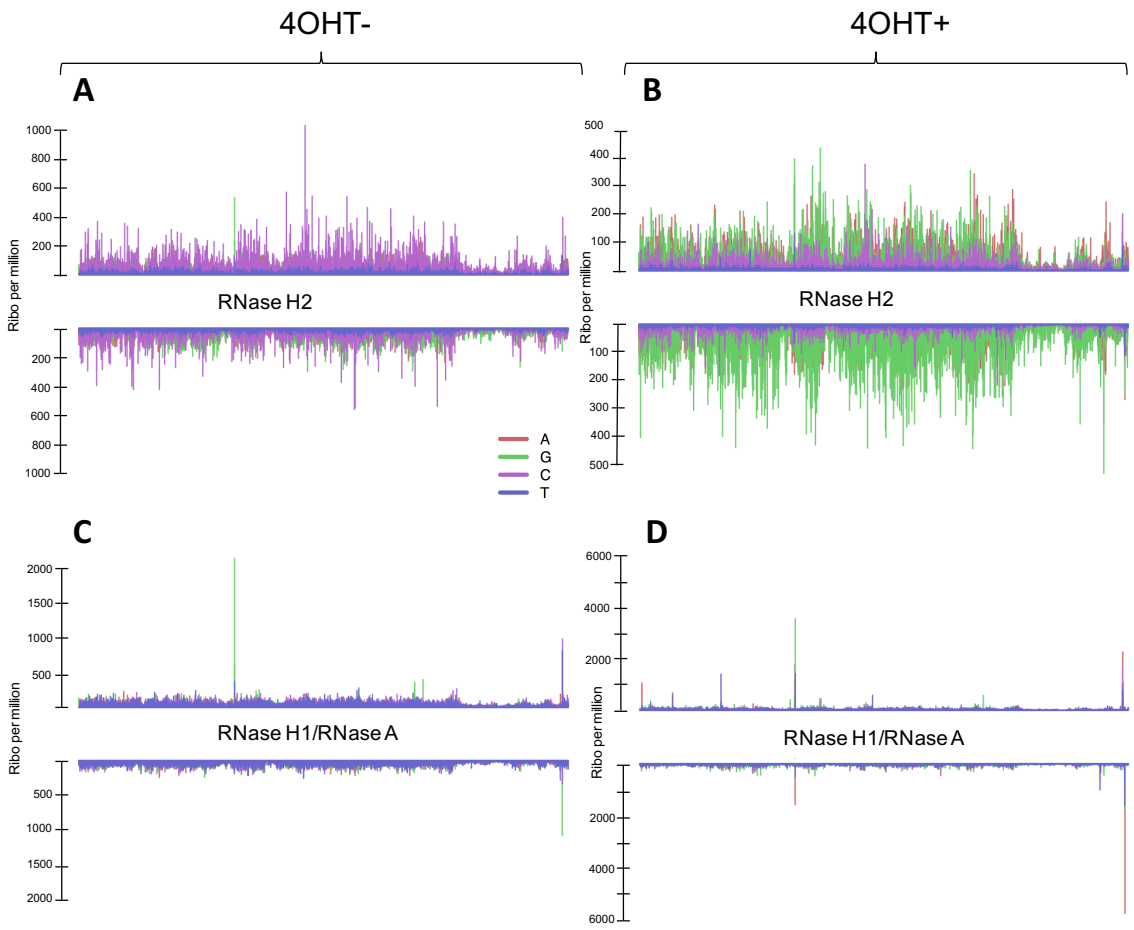


Figure 46: EmRiboseq of 4OHT- and 4OHT+ ΔRH1 MEFs.

Isolated mtDNA molecules from 4OHT- (A and C) and 4OHT+ ΔRH1 MEFs (B and D) were treated with RNase H2 or Eco-RNase HI and RNase A *in vitro* and the ends mapped. **(A)** 4-OHT- MEFs treated with RNase H and RNase A reveal the major sites of primer retention *in vivo*. **(B)** 4-OHT+ MEFs (ΔRH1) show the same peaks as in C, in addition to several other candidate primer sites. Take note of variations in the scale of the y-axis. [Bioinformatic analysis carried out by Dr Martin Taylor]

Recapitulating the Southern Blot analysis, comparisons of 4OHT- and 4OHT+ ΔRH1 MEFs identified sites of persistent RNA only found in mtDNA from RNase H1 deficient MEFs, and absent in the wild-type (Figure 47).

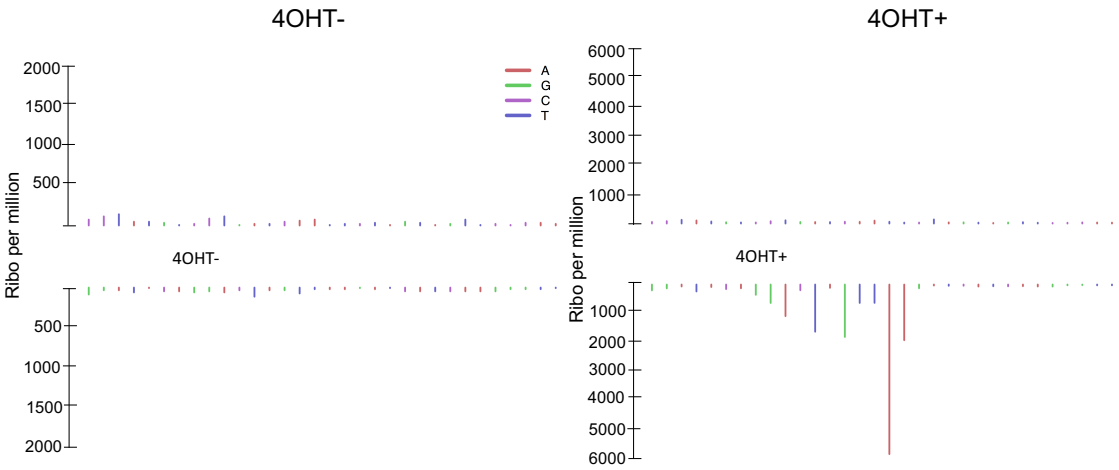
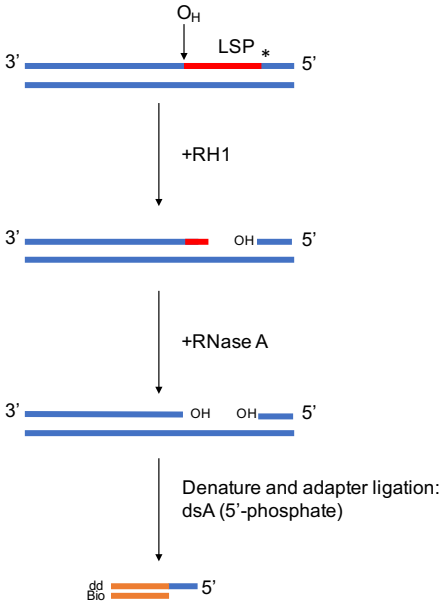


Figure 47: Ablation of RNase H1 in MEFs generates novel primers in mtDNA

Zoom of the region 16,170-16,200 nt in ΔRH1 MEFs in the absence of (left) and presence of (right) 4OHT+. Primer-directed EmRiboseq libraries prepared with RNase H1 and RNase A. Upon loss of RNase H1, there is a peak at 16,186 nt on the heavy strand (negative).

5.6.3 LSP

Ori-L is striking because the peak is present in all libraries, either RNase H2 or RNase H1/RNase A treated. Despite this, in RNase H1/RNase A treated 4OHT+ ΔRH1 libraries, Ori-L is not the highest peak (Figure 46D). The highest peak identified by primer-directed EmRiboseq was at 16186 np on the heavy strand (Figure 48C). This is 3 nucleotides upstream from LSP (Chang and Clayton, 1986), which plays a critical role in mtDNA replication initiation. LSP is where the light-strand transcript originates from and it provides the primer for the initiation of heavy strand replication at O_H or ori-b (Chang and Clayton, 1985). This is in agreement with conclusions made from Holmes *et al.* that the majority of mtDNA replication is initiated at LSP (Holmes et al., 2015). This result establishes that the 3' RNA/DNA transition site has been mapped in these libraries, which is concordant with the chemistry of RNase H1 and RNase A hydrolysis (Scheme 4).



Scheme 4: RNase H1 and RNase A hydrolysis at LSP maps the 3' DNA/RNA junction (*)

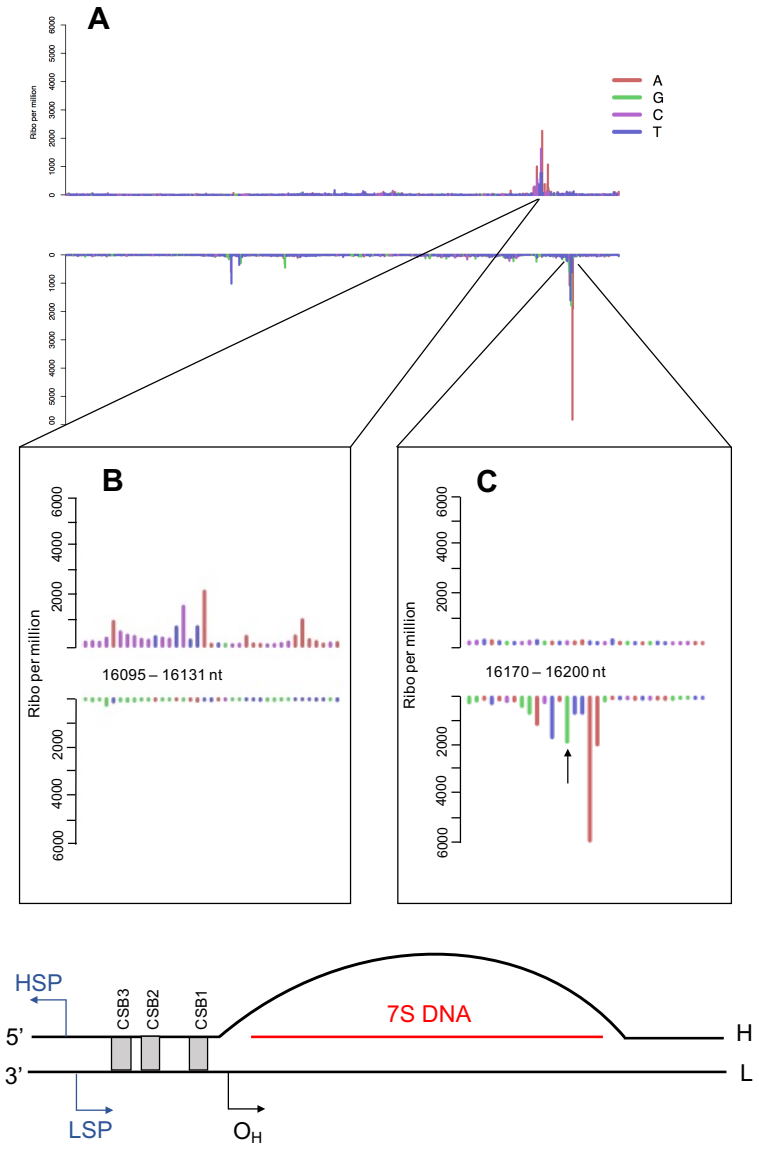


Figure 48: Additional RNA/DNA hybrid sites identified by primer-directed EmRiboseq of 4-OHT+ ΔRH1 MEFs.

(A) Zoom of the region 15,000 to 16,299 nt encompassing the NCR. (**B**) 16,095 – 16,113 nt reveals a peak at 16,112 nt on the L-strand. (**C**) 16,170-16,200 nt highlights the peak at 16,186 nt on the H-strand which is the most abundant peak detected in the library. LSP is indicated by the arrow (16,183 nt). (Bottom) A diagram of the D-loop in mouse mtDNA.

With the exception of Ori-L, the most populous sites of RNA mapped by primer-directed EmRiboseq in the 4OHT+ ΔRH1 MEFs mtDNA were all located within the NCR, where all the regulatory elements are found (Table 5).

A

Peak height / Ribo per million	Strand	Position / np	Relevance
8362	Heavy	16186	LSP
5106	Light	5188	Ori-L
3248	Light	16112	CSB III
2728	Heavy	16187	LSP
2567	Heavy	16183	LSP
2557	Light	5174	Ori-L
2342	Light	16109	CSB II/III
2308	Heavy	16181	LSP
2283	Heavy	5183	Ori-L
2038	Light	5189	Ori-L

B

	Site
Total genome	16,299
D-loop	15,423-16,299
HSP 1	16,282
LSP	16,183
CSB III	16,114-16,131
CSB II	16,089-16,104
CSB I	16,035-16,058
O_H	16,034
Ori-b	~15,600
Ori-L	5,188

Table 5: (A) The highest peaks in 4-OHT+ ΔRH1 MEFs primer-directed EmRiboseq libraries. (B) Nucleotide positions as published at www.ncbi.nlm.nih.gov.

5.7 There is an Increase in rGMP Incorporation in ΔRH1 MEFs Accompanied by Reduced mtDNA Copy Number

The pathogenic associations of increased rGMP incorporation in mtDNA with mitochondrial dysfunction were discussed in the previous chapter. It is interesting to find an increase in the relative abundance of rGMP in the 4OHT+ ΔRH1 MEFs (Figure 46 and Figure 49) identified by EmRiboseq. This is further evidence that RNase H1 has a significant role in mtDNA metabolism, and not simply just primer processing (Akman et al., 2016). It also implicates guanosine metabolism as a critical feature of mitochondrial function. It is not yet known whether RNase H1 has an impact on mitochondrial nucleotide pools. Investigating this further will provide critical insight to this observation.

Moreover, when 4OHT is added to the conditional knockout MEFs and RNase H1 is excised, there is a corresponding reduction in mtDNA copy number (Figure 49C). This is likely due to the accumulation of stalled replication intermediates due to loss of RNase H1 (Holmes et al., 2015). Nonetheless, there is no obvious rationale for why rGMP incorporation is increased.

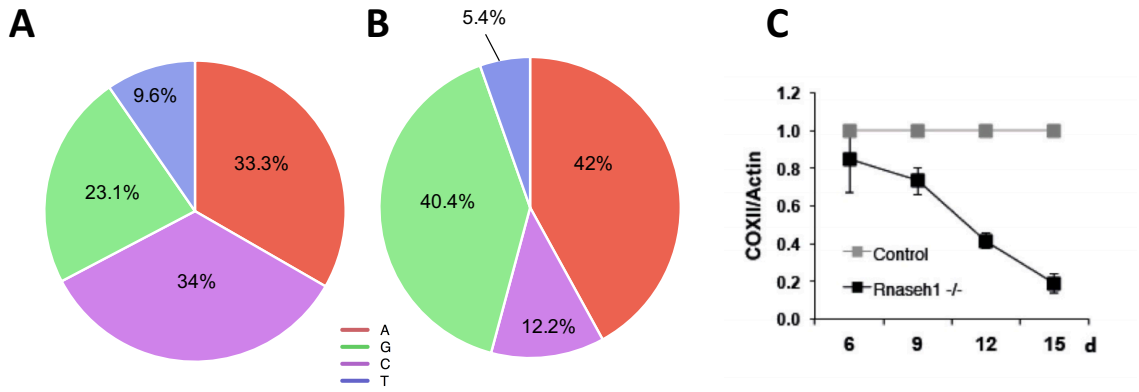


Figure 49: Proportion of incorporated ribonucleotide bases in A) 4OHT- and B) 4OHT+ ΔRH1 MEFs. C) MtDNA copy number upon addition of 4OHT compared to the controls (4OHT-).

EmRiboseq libraries of RNase H2-treated mtDNA from A) 4OHT- and B) 4OHT+ ΔRH1 MEFs as shown in Figure 46A and B respectively. Proportional base contributions were calculated from total mitochondrial ribonucleotides. C) [This image has been re-used with permission from the authors. J. Bradley Holmes *et al.* PNAS 2015;112:9334-9339, (Holmes et al., 2015)].

5.8 Results III: Concluding Remarks

RNase H1 is critical for mtDNA replication, and is therefore knockout lethal. Conditional knockout of RNase H1 from MEFs results in the accumulation of retained primers in mtDNA. This leads to mtDNA depletion due to replicative stalling and ultimately gross mitochondrial dysfunction. The conditional RNase H1 system has provided an excellent model with which to examine mtDNA replication *in* vitro. By modifying the original EmRiboseq protocol I have mapped the sites of retained primers in ΔRH1 MEFs. Concordant with reports from Holmes *et al.* the data concludes that the majority of mtDNA synthesis is initiated at LSP. In order to map the precise mtDNA origins in the future (as opposed to the 3' DNA/RNA junction), this technique would need to be modified to ligate adapters using a 3'-phosphate biotinylated adapter, as opposed to a 5'- one. Nonetheless, this is an early demonstration that the EmRiboseq is an adaptable robust NGS protocol for mapping sites of DNA modifications *in vivo.*

Chapter 6. Discussion

Over the past 20 years, NGS methods have become ubiquitous in everyday research and thanks to rapid developments in efficiency and sequencing power, massive parallel sequencing has become an accessible and affordable tool for analysing DNA mutations and modifications *in vivo.*

6.1 HydEn-Seq and EmRiboseq

In 2015, four groups published four distinct methods for mapping ribonucleotide incorporation, all but one omitting any mitochondrial data from the final publication (Clausen et al., 2015, Ding et al., 2015, Koh et al., 2015, Keszthelyi et al., 2015). It wasn't until 2017, when Berglund *et al.* applied the HydEn-seq method to isolated mtDNA, that ribonucleotide incorporation in mtDNA was extensively examined, 44 years after the initial discovery that mtDNA contains persistent ribonucleotides (Berglund et al., 2017, Grossman et al., 1973). They concluded that the identity and frequency of ribonucleotide incorporation in mtDNA was dictated by the compartmental nucleotide pools. This conclusion is broadly supported by this study.

Using an *in vitro* model of ribonucleotide incorporation on a primed DNA template, where the exonuclease activity was inactivated by a D274A substitution in the second exonuclease motif in the POLGA subunit, Berglund *et al.* showed that the exonuclease activity of Pol γ has no discernible effect on ribonucleotide incorporation *in vivo* (Berglund et al., 2017). In this thesis, I have found no evidence for an active RER pathway in the mitochondria in accordance with this conclusion.

6.1.1 Limitations of HydEn-Seq

Berglund *et al.* looked at ribonucleotide incorporation in mtDNA from cultured human cells. As has been demonstrated here, there is minimal ribonucleotide incorporation in proliferating cells, the profile of which is very different compared to that of solid tissues. A critique of the study is that neglecting the analysis of solid tissues was an oversight, given the significance of mtDNA metabolism following exit from the cell cycle. Considering that many mitochondrial diseases manifest in a tissue-specific manner, it would have been beneficial to examine various tissues, as has been done

in this study. Moreover, using HeLa cells as a control is dubious considering that they are cancerous, as opposed to true controls.

Akin to all of the aforementioned methods for mapping ribonucleotides, including EmRIboseq, the efficiency of HydEn-seq is hampered by the requirement for adapter ligation. Adapters are also often associated with sequence bias of adapter capture, of which is evident in the study from Berglund *et al.* where HydEn-seq libraries displayed a 14-fold higher coverage on the L-strand relative to the H-strand (Berglund et al., 2017).

An additional pitfall of HydEn-seq is the inability to discriminate between alkali-sensitive sites, such as longer stretches of RNA and abasic sites, and single embedded ribonucleotides.

6.1.2 Limitations of EmRiboseq

A strength of the EmRiboseq approach is the substrate-specificity conferred by using RNase H2, in place of alkali hydrolysis. However, as with the other approaches, total quantification is difficult (3.5). One reason is the sequencing technology (Ion Torrent™), which has an optimal fragment size; normally in the range of 100 to 400 base pairs. This means that any DNA fragments which fall outside of this window are unlikely to be sequenced.

A critical problem, which has hampered progress during this project is the recurring trough across the genome due to truncation of reads (Figure 14). This problem has no obvious solution at this stage. To develop this approach further, it is absolutely critical to address this issue in order to make any valid conclusions about ribonucleotide incorporation *in vivo*.

The other limitations arise from the EmRiboseq protocol itself (Ding et al., 2015). Having to ligate adapters at two separate stages in the protocol is problematic for many reasons. The primary being the low efficiency of ligation reactions which results in many molecules not becoming adapter-ligated and subsequently not sequenced. Equally, one end of a molecule may be adapter ligated in the first

instance (and the other end escape ligation), and therefore remain ligatable for second adapter ligation where ligation should be, in theory, specific for RNase H2-sensitive molecules. This phenomenon would give rise to false positives; an effect which is hopefully eliminated in background screening (3.4.3). Another problem is that ribonucleotides in close proximity to one another are not sequenced as outlined schematically in Figure 50. The best way to overcome this latter problem is to sequence as many technical (and biological, if possible) replicates as possible in order to get a powerful average read of ribonucleotides across the genome in all samples.

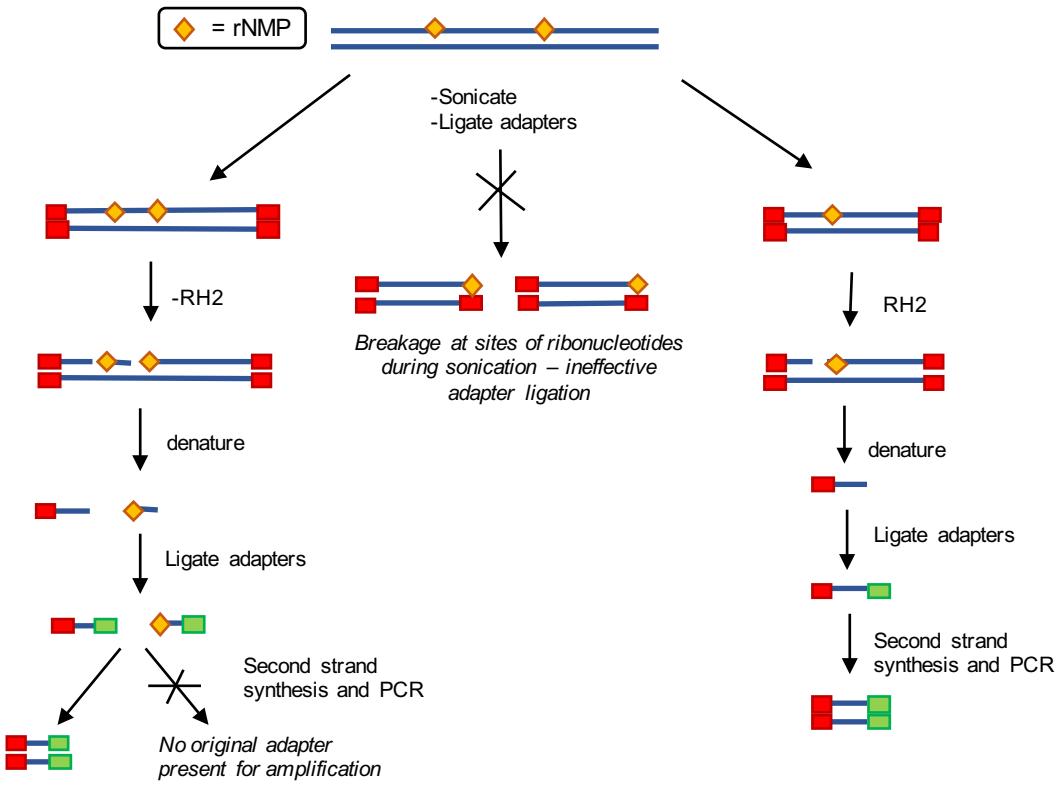


Figure 50: Limitations of the EmRiboseq approach.

Examples of problematic instances which could feasibly occur in any given library preparation including proximal ribonucleotides not being sequenced and DNA shearing causing breaks at sites of ribonucleotides.

6.1.3 Potential Alternatives to Ion Torrent Sequencing

Many of the limitations of the EmRiboseq approach are associated with the requirements of the sequencing technology. Something which is also encountered

with Illumina sequencing. That is, the requirement for physical fragmentation of the DNA libraries, adapter ligations and amplifications (albeit limited). In the last decade there have been rapid advances in sequencing technologies. One such emerging technology is nanopore sequencing; a technique which shows promise in being able to identify not only the four canonical DNA bases but other modified bases, based on their chemical properties. This would be perfect for the application of identifying ribonucleotides whilst avoiding the aforementioned problematic requirements. However, currently this technology is still in its infancy and there are still relatively high error rates in reading DNA moieties.

By contrast, one of the earliest DNA sequencing technologies; PacBio sequencing, is able to read much longer reads which would overcome what I believe to be the biggest problem with quantification; and that is losing reads due to fragment sizes. Although this technique has a higher error rate than the more sensitive Ion Torrent and Illumina technologies; this could be beneficial for the applications described here, or even in parallel.

6.1.4 Mapping Primers using EmRiboseq

In Chapter 5 I modified the EmRiboseq protocol in order to map origins *in vivo*. This was facilitated by conditional ΔRH1MEFs that enable the capture of RNA/DNA hybrids which would otherwise be rapidly processed in control cells. This demonstrated the versatility of the technique and using this approach it was possible to map persistent primers *in vivo* to single nucleotide resolution (Figure 43). The major limitation of this was that the combination of RNase H1 and RNase A meant that the 3' DNA/RNA junctions were mapped and not the precise replication origin(s). One way to combat this is to re-design the primers used in the EmRiboseq protocol so that there is a 5'-phosphate adapter to map the 5' DNA/RNA junctions. Another possibility is to incubate the libraries with a single-strand nuclease in addition to RNase H1 and RNase A, which would enable identification of replication origins (by creating exposed ends on the opposite site).

An additional caveat of this analysis was that, like the RNase H2 approach, the reads were subject to sequence coverage bias. This was more acutely palpable in the

ΔRH1 libraries as many of the sites of interest fall inside the NCR, and therefore the trough of reads (Figure 14). One way to help improve this is to normalise the data in its entirety to the background signal.

6.1.5 Future Applications for Modified EmRiboseq

One of the greatest advantages of the EmRiboseq method is its adaptability. The critical stage in the protocol which makes it specific for ribonucleotides is the endonucleolytic treatment with RNase H2. Yet, as demonstrated this can be adapted and modified to target other types of DNA modifications. In theory, a multitude of DNA modifications can be sequenced via this method, with the pre-requisite that there is a treatment or enzyme which creates a ligatable 3'-OH end at the site of interest. For example, the DNA glycosylase OGG1, is able to excise 8-oxodG residues by attacking the *N*-glycosylic bond, catalysing base removal and subsequent excision through hydrolysis via β-elimination and cleavage of the DNA phosphate backbone resulting in 3'-phospho-α,β-unsaturated aldehyde (3,4-hydroxy 2-pentenal) and 5-phosphate termini (Hill et al., 2001, Sun et al., 1995). Harnessing this, along with the EmRiboseq method, would provide a powerful approach to mapping oxidised bases *in vivo*.

6.2 Mitochondrial Evolution and rAMP Tolerance

The most popular theory of the origin of mitochondria and their presence in eukaryotic cells is the endosymbiosis theory which states that mitochondria originated from the fusion of α-protobacteria and archaebacteria (Andersson et al., 1998, Gray, 2012, Lang et al., 1997, Gray et al., 2001). And thus, mitochondria are very similar to their bacterial ancestors. One key example being that mtDNA replicates in a unique manner, reminiscent of bacterial replication (Lovett et al., 1974). In line with this, many of the proteins which play a role in mtDNA replication, such as TWINKLE and POLRMT, share sequence homology with T7 bacteriophage proteins (Ringel et al., 2011, Milenkovic et al., 2013).

Mitochondria have evolved as the energy-providing hub of the eukaryotic cell for approximately two billion years and have developed tools and mechanisms to preserve the integrity of the DNA contained within. One such protective mechanism

is the packaging of mtDNA by protective proteins such as TFAM. TFAM is the mtDNA packaging protein and is thought to protect the DNA from ROS and other damaging agents, both endogenous and exogenous. Moreover, there are extensive repair pathways within the mitochondria (see (Kazak et al., 2012) for a comprehensive review). There is strong evidence for base-excision repair in mammalian mitochondria, as well as a multitude of proteins which localise to the organelle and are involved in processing damaged bases and nucleotides, such as OGG1 (Furihata, 2015, Ruchko et al., 2011). However, one critical repair pathway has eluded the mitochondria; RER. The absence of RNase H2 within the mitochondria is unexpected given that the vast majority of bacteria have an extensive set of Ribonucleases H proteins; RNase HI, RNase HII and RNase HIII. And thus, the persistence of ribonucleotides within the DNA is unique to mitochondria.

As has been outlined throughout this thesis, ribonucleotides embedded into duplex DNA confer susceptibility to hydrolysis, strand nicking and double strand breaks and cause perturbations to the double helix itself (DeRose et al., 2012). Ribonucleotides far outnumber deoxyribonucleotides within the cell and what is more, ribonucleotides and deoxyribonucleotides differ only by one atom so it is therefore a challenge for DNA polymerases to discriminate between the two. DNA polymerases use steric gates to obstruct ribonucleotide incorporation, however they have a lower limit and all polymerases incorporate ribonucleotides, however infrequently (Table 6). Evidently, ribonucleotide (mis)incorporation is an unfavourable consequence of imbalanced NTP:dNTPs, but how a system reacts to this inevitability is not necessarily the same.

In the nucleus, there is a proficient RER pathway which is RNase H2-mediated and efficiently excises embedded ribonucleotides from the newly synthesised DNA strand. This is a necessity in the nuclear DNA where double-strand breaks can cause genome instability if left unrepaired, or if erroneously repaired. However, in the mitochondria where there is a multicopy genome, damaged molecules can be turned over and therefore ribonucleotide (mis)incorporation is not such a critical state of affairs. What is more, there is a growing body of evidence to suggest that the genomic instability arising from ribonucleotide (mis)incorporation in nuclear DNA (in RNase H2-deficient models) is not from the induced damage susceptibility of the

DNA, but a consequence of aberrant repair and replisome stalling (Williams et al., 2013, Kim et al., 2011, Huang et al., 2017). Therefore, it is not hard to envisage an alternative system for tolerating ribonucleotide incorporation; which is the evolution of a multicopy genome along with a dedicated replicative DNA polymerase which is able to proficiently reverse transcribe embedded ribonucleotides (Table 6) (Kasiviswanathan and Copeland, 2011), conferring a semblance of tolerance of ribonucleotides in mtDNA.

Compared to other mammalian replicative polymerases, Pol γ has a high DF against ribonucleotides, especially ATP (Table 6). This is due to the steric gate at residue E895, which functions to minimise ribonucleotide incorporation. It would be interesting to explore whether mutating this residue, and disrupting the steric gate, would result in a gross increase in ribonucleotide incorporation in mtDNA, and whether this would lead to an increase in replication stalling, deletions, depletion and ultimately OXPHOS insufficiency as in the Mpv17 knockout mouse (Figure 51). However, it is likely that this residue is also confers fidelity in correct-base pairing to the polymerase, therefore mutating said residue could also increase mtDNA point mutations.

Polymerase	Family	Function	Sugar Discrimination	rNMP extension	rNMP bypass
Human Pol γ	A	mtDNA replication	11,000-77,000	7-33%	51%
S.Cerevisiae Pol α	B	Genome replication	3,000-10,000	100%	50-75%
S.Cerevisiae Pol δ	B	Genome replication, NER, BER, MMR, DSBR	10,000-1,700	/	7.4-69%
S.Cerevisiae Pol ε	B	Genome replication, NER, BER, MMR, DSBR	500-6,700	90%	32-85%
S.Cerevisiae Pol ζ	B	TLS, DSBR, ICLR, SHM	470-4,700	/	80-95%
Human Pol δ	B	Genome replication, NER, BER, MMR, DSBR	2,000	/	78-85%
Human Pol ε	B	Genome replication, NER, BER, MMR, DSBR	210-5,300	37-52%	66%
Human Pol β	X	BER, MMR	200-8,200	100%	/
Human Pol λ	X	BER, NHEJ, VDJR	3,000-50,000	70%	/
Human Pol μ	X	NHEJ, VDJR	1.4-11	25%	/
Human Pol ι	Y	TLS, BER, SHM	1,100	100%	/

Table 6: Base and sugar fidelity of several DNA polymerases from the four eukaryotic families.

Abbreviations: NER, nucleotide excision repair; BER, base excision repair; MMR, mismatch repair; DSBR, double strand break repair; TLS, translesion synthesis;

ICLR, interstrand cross-link repair; SHM, somatic hypermutation; NHEJ, non-homologous end-joining; VDJR, VDJ recombination. Data adapted from Cerritelli *et al.* and the references therein (Cerritelli and Crouch, 2016).

6.3 Predictive Ribonucleotide Incorporation Frequencies

As shown in chapter 3, rAMP is the dominating ribonucleotide embedded in mtDNA from post-mitotic tissue (Figure 19). This is largely unsurprising due to the high compartmental concentration of ATP due to OXPHOS. The proportion of the four different ribonucleotides in mouse liver mtDNA is largely in agreement with predictive frequencies (Figure 22); discrepancies of which can be rationalised by the sequencing limitations and using *in vitro* values for Pol γ. In order to maintain mtDNA copy number, Pol γ must have to frequently encounter embedded rAMPs and reverse transcribe them, correctly, to create a newly synthesised daughter molecule. Data from Kasiviswanathan *et al.* indicate that this is indeed the case, and that Pol γ is capable of effective reverse transcription (Table 6); likely an activity which has facilitated the evolution of a DNA-containing organelle with high compartmental ATP concentrations (Kasiviswanathan et al., 2012). Moreover, mtDNA replication is far slower than nuclear DNA replication (Bogenhagen and Clayton, 1977), with unique mechanism(s) of replication. Perhaps these two observations are additional adaptations to ensure faithful replication and minimise ribonucleotide incorporation.

The EmRiboseq analysis combined with restriction enzyme digestion estimates that there are approximately 340 ribonucleotides per mtDNA molecule in mouse liver. Berglund *et al.* postulated that there were approximately 36 ribonucleotides per mtDNA molecule from HeLa cells (Berglund et al., 2017), which is in the same range of that predicted by Grossman *et al.* (Grossman et al., 1973). This equates to a difference of an order of magnitude between cells and liver mtDNA. Despite demonstrating that there are far more ribonucleotides in mtDNA from solid tissue (Figure 28), I believe that the figure obtained from liver mtDNA is an overestimate attributable to the number of background reads within the EmRiboseq library.

The predictive calculations, although accurate relative to the proportion of the different bases, estimates a total of 2,148 ribonucleotides per mtDNA molecule. This

is significantly higher than the calculated value of 340 from EmRiboseq and therefore suggests that it is not just the NTP:dNTP ratio within the mitochondria and the DF of Pol γ which determine ribonucleotide incorporation. There are two main possibilities as to why ribonucleotide incorporation is less frequent than predicted; i) dNTPs are concentrated in the vicinity of the replisome, skewing the NTP:dNTP ratio; or ii) the mitochondria possess some ability to excise and repair ribonucleotides, or clear molecules which exceed a certain threshold.

6.4 Potential Advantages of Ribonucleotides in MtDNA

One of my arguments for the absence of RER within the mitochondria is the multicopy genome which means that damaged or compromised mtDNA molecules can be turned over. The caveat of this is that there are multiple repair pathways active in mitochondria. This suggests that mtDNA turnover is not a primary mechanism of damage tolerance. It is therefore possible that the persistence of ribonucleotides in mtDNA confers an advantage to the organelle.

6.4.1 Mismatch Repair

In many systems mismatch repair (MMR) is used as a mechanism to correctly repair lesions or mis-insertions during replication (Nick McElhinny et al., 2010b, Yao et al., 2013). During MMR the newly synthesised DNA strand is marked by embedded ribonucleotides, to facilitate identification of the template strand for repair. However, such a mechanism in mitochondria would be futile as there are persistent ribonucleotides on both the template and newly synthesised strand.

6.4.2 Programmed Pause Sites

An intriguing finding presented within this thesis is the presence of hot-spots of ribonucleotide incorporation (Figure 11). There are sites across the genome where there is frequently a ribonucleotide in many molecules, aside from the highly documented replication origins, which are not easily explained by polymerase behaviour nor sequence context. This suggests that there are sites of persistent ribonucleotide incorporation in mtDNA. These could represent programmed pause sites, as the replisome is known to pause at sites of embedded ribonucleotides (Yao et al., 2013). Programmed pause sites could serve as a mechanism to control

replication rate, invoke repair processes in the case of DNA damage, or even to prevent collision of replication and transcription machinery (Rothstein et al., 2000). From this perspective, ribonucleotide incorporation in mtDNA can be seen as advantageous.

6.4.3 Maintaining a Master Copy of mtDNA

As above stated, an argument for the persistence of ribonucleotides in mtDNA is that the genome is multicopy, therefore compromising the integrity of the molecules does not matter as long as there is a master copy (ribonucleotide-sparse or -free) to preserve future copies of mtDNA. Having a multicopy genome means that the mitochondria have the capacity to turnover damaged molecules and replace them with newly synthesised ones, independent from the cell cycle. However, as has been demonstrated within this thesis, ribonucleotide incorporation is almost unavoidable and with no evidence for a RER mechanism, all molecules which have been actively replicated should, in theory, contain persistent ribonucleotides. However, this is not the case. From Southern blot analysis when mtDNA is incubated with either NaOH or RNase HII *in vitro* less than 100% of mtDNA molecules are alkali-sensitive and as such, there are mtDNA molecules which appear to be resistant to *in vitro* fragmentation (Figure 27). This suggests that there are ribonucleotide-free mtDNA molecules existing within the population.

One way to retain a master copy of mtDNA is to maintain one or more ribonucleotide free molecules of mtDNA which would serve as the pristine template. Preserving said copies would mean they would likely be tightly packaged (to avoid oxidative damage from ROS and other agents) and perhaps physically segregated from other actively replicating or transcribing molecules. This model would infer that mtDNA from primordial germ cells (PGCs) contain relatively few ribonucleotides, compared to other cell types.

6.5 Ribonucleotides and the Threshold Effect

A characteristic feature of mtDNA is the phenomenon of heteroplasmy. Almost all of the mitochondrial genome contains coding DNA and therefore mtDNA mutations are highly likely to occur in protein-coding genes. An expansion of mutated molecules

within a population of mtDNA can have detrimental effects if; i) it is a nonsynonymous mtDNA mutation and, ii) the fraction of mtDNA molecules bearing the mutation exceeds a certain threshold. This threshold effect is vital to the pathology of mitochondrial disorders and is a key element of mtDNA inheritance and disease. It is not purely an arbitrary level above which dysfunction takes place, but a critical limit above which the healthy molecules cannot compensate for the presence of the mutated molecules.

As shown from the Mpv17 knockout mice, there is a substantial increase in replication stalling, and subsequently an accumulation of deletions and depletion, accompanied with an increase rGMP incorporation rate. Therefore, it could be argued that there is also a threshold effect with respect to rNMP levels in mtDNA. This threshold appears to be more qualitative than quantitative, as rAMPs are naturally abundant in mtDNA from post mitotic tissues. In order to examine this further it would be beneficial to have a way to directly increase ribonucleotide incorporation in mtDNA, for example by either mutating the aforementioned E895 residue of Pol γ or increasing the mitochondrial rNTP concentrations. Through this, the effect of increasing rNMP incorporation and subsequent replication stalling and deletion formation could be investigated.

6.6 MPV17 and Embedded Ribonucleotides in mtDNA

MtDNA deletions are a common underlying cause of mitochondrial dysfunction and are commonly found in aged tissues (Crott et al., 2005, Herbst et al., 2007, Yu-Wai-Man et al., 2010). Deletions have often been assumed to be distinct from the phenomenon of mtDNA depletion. However, the data contained within Chapter 4 suggest that the two phenomena are on the same spectrum of mitochondrial dysfunction. In the Mpv17 knockout mouse model, it appears that mtDNA depletion is an adaptive response to the accumulation of deleterious deleted molecules within the mitochondria, of which is tightly linked to rGMP levels in mtDNA.

6.6.1 rGMP and Mitochondrial Disease

Riboguanosine is the most commonly incorporated rNMP by DNA polymerases *in vitro* (Clausen et al., 2013, Evich et al., 2016) and Pol γ is poorest at discriminating

between GTP and dGTP (Kasiviswanathan and Copeland, 2011). However, as demonstrated, GTP is not the most commonly incorporated ribonucleotide *in vivo* in mtDNA (Figure 19). The reason it is not as populous as rAMP is likely due to the GTP:dGTP ratio which is regulated by a multitude of proteins, including Mpv17, although how exactly remains a mystery.

Ubiquitous expression of the protein (Spinazzola et al., 2006) and the normal levels of dGTP in the brain of the Mpv17 knockout mouse argue against the protein being directly involved in dNTP synthesis or transportation (Figure 37). Nevertheless, the presence of elevated rGMPs within the mtDNA in all tissues analysed, independent of mtDNA copy number, indicate that guanosine metabolism is critical to the disorder. One possibility, is that Mpv17 serves to provide a supply of pristine, undamaged dGTPs to the replisome in mitochondria to facilitate faithful mtDNA replication.

The link between guanosine metabolism and Mpv17 is not novel. When Mpv17 is depleted in zebrafish, the fish lose their stripes, which are composed of crystalline guanosine (although there is no accompanying mtDNA depletion) (Krauss et al., 2013). It is conceivable that in the zebrafish there is a naturally high abundance of guanosine in order to maintain their characteristic stripes, and also providing the mitochondria with a sufficient source of dGTP pre-cursors for mtDNA replication. However, when Mpv17 is depleted, there is no longer sufficient dGTP for mtDNA replication and so the majority of guanosine is shunted to the mitochondria to maintain normal mtDNA copy numbers whilst sacrificing the stripes. This hypothesis places the emphasis on dGTP depletion being the underlying cause of the pathology.

Despite evidence that rGMP incorporation is critical to the Mpv17 phenotype, a direct causative link between increased rGMP incorporation and slow replication and reduced copy number cannot be drawn without evidence that embedded rGMPs pose a problem to the replicative polymerase. Yao and colleagues demonstrated *in vitro* that increased rNTP concentrations at the replisome (which would be the relative effect of dNTP depletion observed in the Mpv17 knockout liver), not only slows the replisome, but also that GTP is the most detrimental nucleotide impeding DNA synthesis (Yao et al., 2013).

Moreover, there are ample examples in the literature demonstrating that a single embedded rGMP causes significant local torsional changes to the duplex DNA (whereby the lesion locally perturbs the structure asymmetrically on the 3' side), without appreciably affecting duplex stability (Chiu et al., 2014, DeRose et al., 2012, Evich et al., 2016). Furthermore, the X-ray crystal structure of a duplex containing a single rGMP moiety suggests that there is a global conformational change from B-DNA to A-DNA with all sugars in the C3'-*endo* conformation (Ban et al., 1994) however, more recent studies have disputed this (Evich et al., 2016). Moreover, rGMP creates a greater conformational change compared to rAMP according to crystallography data (Egli et al., 1993). Therefore, it is very plausible that increased rGMP incorporation and torsional stress to the mtDNA molecule causes replication stress and subsequent stalling as observed in the absence of Mpv17, more-so than the single rAMPs observed in controls.

6.6.2 A Proposed Disease Model

Fibroblasts with mutations in genes encoding DGUOK, akin to MPV17 patient-derived cells, have decreased mitochondrial dGTP pool size and subsequently increased rGMP in mtDNA was reported from patient cells harbouring pathogenic *DGUOK* mutations (Berglund et al., 2017). Furthermore, patient fibroblasts with mutations in the gene encoding TK2 exhibited increased rCMP incorporation in mtDNA, concordant with a depleted mitochondrial dCTP pool (Berglund et al., 2017). It would be interesting to further explore the TK2 model to see whether increased rCMP incorporation had the same pathogenic consequences as the increased rGMP in the Mpv17 knockout mice. In order to do this, I would suggest analysing mtDNA in tissues from the TK2 knockout mice to see if there are any accompanying deletions, in addition to mitochondrial dysfunction and depletion, and whether mtDNA deletions precede the mtDNA depletion (Paredes et al., 2013).

High levels of rGMPs are a ubiquitous feature of Mpv17 deficiency and precede any consequent phenotype. I propose that Mpv17 has a role in providing sufficient amounts of dGTP to the replisome. In the absence of Mpv17 there is an accumulation of rGMP throughout the mtDNA which in turn causes replication stalling which leads to late-onset mtDNA deletions in the brain. In order to clear the mitochondria of

harmful mtDNA deleted molecules, the mtDNA is rapidly turned over which results in mtDNA depletion. It is unclear whether the reduction of mtDNA precedes the depletion of dGTP pools, which is proposed to be a mechanism to facilitate faithful mtDNA replication in the absence of Mpv17 (this would have to be approached by looking at more Mpv17 knockout mice at different stages of development). The latter stages of the disease progression are as seen in the liver where there is significant mtDNA depletion and dGTP (and dTTP) depletion, and ultimately OXPHOS deficiency.

MtDNA is thought to be turned over regularly, with a half-life of the order of months (Poovathingal et al., 2012). Furthermore, turnover rate across different tissues has been reported to be different in rat (Gross et al., 1969), indicating that turnover may be modified as a result of increased metabolism, mutation or energy demands of cells. Turnover of damaged mtDNA molecules is a key mechanism of coping with mtDNA damage and mitochondrial stress. In this instance, it is proposed that mtDNA degradation is upregulated to avoid the accumulation of deleterious mtDNA molecules.

To confirm my conclusions that the rGMPs in mtDNA are pathogenic, it would be highly beneficial to analyse the mtDNA from different stages in the DGUOK knockout mouse to confirm that the rGMP is deleterious, and that it is not an Mpv17-specific pathology which has been identified.

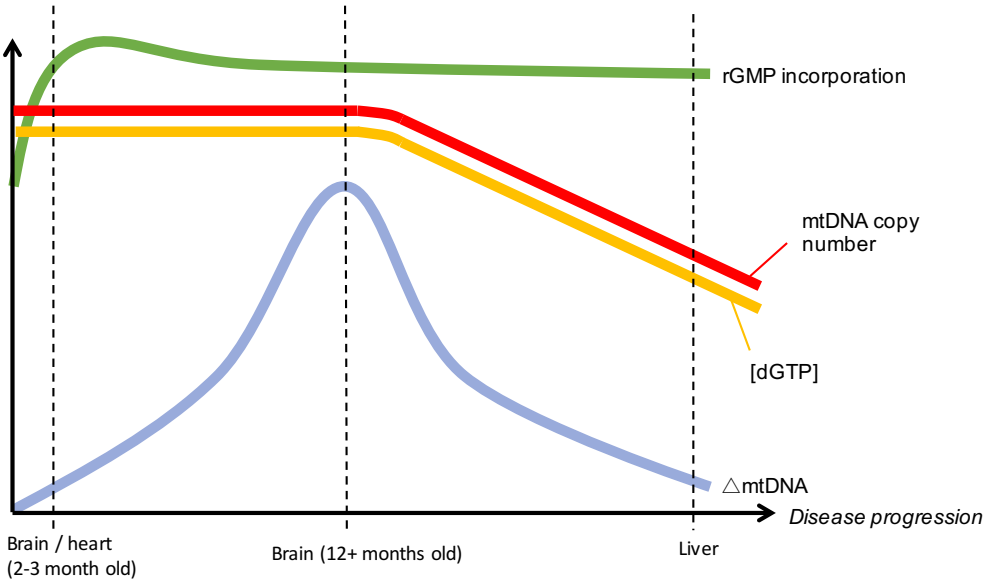


Figure 51: A model of mitochondrial disease for Mpv17 deficiency in mice.

6.7 Ribonucleotide Retention Could be a Cause of mtDNA Deletions

It is highly likely that embedded ribonucleotides in the mtDNA duplex cause considerable difficulties for the replicative polymerase, due to torsional conformational distortions (Chiu et al., 2014, DeRose et al., 2012, Evich et al., 2016). As has been inferred from the Mpv17 knockout mice, rGMPs impede mtDNA replication and deletions were observed in the mtDNA from Mpv17 knockout mice where there was an increase in rGMPs.

For a long time the precise mechanism of mtDNA deletion formation was not known, despite the prevalence of mtDNA deletions in human patients with mitochondrial disease. Recently, analysis of single molecule mtDNA replication using DNA combing was used to investigate the cause of the common deletion; a 4,977-bp region flanked by 13-bp repeats (Phillips et al., 2017). Phillips *et al.* showed that the common deletion is formed as a consequence of frequent fork stalling, a process which is mediated by the mitochondrial replisome, but independent of canonical DSB repair (Phillips et al., 2017). It is therefore conceivable that an increase in ribonculeotides in mtDNA leads to an increase in replisome stalling and erroneous repair which results in deletions. The nature and size of the resulting deletion is

subject to location and the presence of proximal repetitive sequences. Therefore, deletion formation is possible at any time where fork stalling occurs in such regions, such as the common deletion, but is more likely when there is an increase in mtDNA lesions, such as ribonucleotides.

6.8 Ribonucleotides in mtDNA in Cancer and Ageing

A profound difference in ribonucleotide incorporation profiles of mtDNA was found when comparing proliferating and quiescent/post-mitotic cells (Figure 31). The dramatic difference in ribonucleotide incorporation can be largely explained by the increased mitochondrial dNTP pools when nuclear DNA synthesis is taking place, however, this does not explain the under-representation of rAMP in cycling cells. Akin to quiescent cells, in mtDNA from solid tissues the vast majority of embedded ribonucleotides are rAMPs, and this can be explained by the compartmental ratios of NTP:dNTPs in the mitochondria. Yet, mtDNA from proliferating cells have relatively few rAMPs given the high ATP:dATP within the organelle (Figure 29). This hints at a possibility that the mitochondria are in some way able to evade ribonucleotide incorporation in proliferating cells via an unknown mechanism.

The dramatic difference between proliferating and non-proliferating cells is significant for two reasons. Firstly, because as has been highlighted, ribonucleotides impair the integrity of the DNA duplex and therefore an increase in ribonucleotide content is associated with increased risk of double-strand breaks, deletions and depletion. And therefore, an increase in ribonucleotide incorporation, such as with exit from the cell cycle, must therefore be associated with increased mtDNA damage or damage-susceptibility, such as ss- or ds- breaks. Secondly, if ribonucleotide incorporation is directly influenced by the metabolic state of the cell, the phenomenon of reduced ribonucleotide incorporation could be a critical feature of proliferating cells, utilised to maintain mtDNA integrity.

The Warburg effect is a widely documented phenomenon of cancer cells and describes the observation that cancerous cells are largely glycolytic, even in the presence of functioning mitochondria and high energy demands (Liberti and Locasale). There are many hypotheses as to why this is advantageous for the cell,

and it is likely that this is a process which is beneficial on many fronts. One such advantage could be that by eschewing ATP production from the mitochondria to glycolysis there is a reduction in both mitochondrial ROS and ATP. If this state mimics that of proliferating cells in culture then the mtDNA from cancer cells will have fewer embedded ribonucleotides, and thus preserving the integrity of the mtDNA for future cell divisions. It would be an exciting avenue of further research to explore this hypothesis and it could ultimately serve as a biomarker for malignant cells *in vivo*.

Furthermore, an energetic shift from OXPHOS to glycolysis in cancer cells could be explained by the need for other biosynthetic processes. Glycolysis also produces glucose-6-phosphate, ammonia and aspartate which are required for the synthesis of purine and pyrimidine nucleotides (Jose et al., 2011). An increase in nucleotide synthesis would provide additional dNTPs for actively dividing cells, as well as reducing the mitochondrial NTP:dNTP ratio, thus reducing ribonucleotide incorporation rate in mtDNA.

However, in recent years biogenetic studies of tumour metabolism have shown that the Warburg effect is a blanket statement which is not applicable to all tumour types, and there are multiple cases of tumours which are reliant on OXPHOS (Griguer et al., 2005). Tumours are varied and heterogeneous by nature and their metabolic apparatus and modalities of energy substrate utilisation can differ significantly depending on the stage of tumorigenesis, tissue type, oncogene activation and microenvironmental substrate conditions (Garber, 2006).

6.8.1 MtDNA as a Biological Clock

The idea that cells are able to prevent excessive ribonucleotide incorporation by eschewing ATP production, or by an alternative means, is indicative of a preventative mechanism against compromising mtDNA quality. One could speculate that embedded ribonucleotides act as a biological clock. Accumulation of which results in compromised energy production and maybe even induction of mitophagy or autophagy. However, I found no evidence for increased ribonucleotide incorporation in aged mice, therefore, akin to the Mpv17 model, this threshold could be a qualitative shift in ribonucleotides. So, an increase in certain ribonucleotide bases, rather than a gross overall increase.

6.9 A Proposed Mechanism for Mitochondrial Ribonucleotide Excision Repair in Mammals

Mitochondria seem to possess a non-traditional mechanism of surviving the inevitability of ribonucleotide (mis)incorporation during DNA replication. Rather than adopting a classical RER mechanism, it appears that mitochondria have evolved a tolerance of embedded rAMPs within the mtDNA (6.2). Despite lacking evidence for a canonical RER mechanism, almost all of the proteins associated with RER in the nucleus are present within the mitochondrial compartment (with the exception of the aforementioned RNase H2 (Figure 24)). In the absence of RNase H2 in the nucleus, Top1 is able to recognise single embedded ribonucleotides and initiate an alternative RER pathway following incision 3' to the ribonucleotide moiety (Huang et al., 2017). Top1mt is the only DNA topoisomerase specific for mtDNA in vertebrates, but there is also evidence for the presence of Top2α and Top2β active in murine mitochondria (Douarre et al., 2012). If Top1mt is able to recognise and nick single embedded ribonucleotides, then it could be argued that mitochondria possess, in theory, all the necessary tools to carry out topoisomerase mediated- RER.

ChIP-seq analysis of a poisonous Top1mt (Top1mt*), which carries mutations stabilising the DNA-bound complex of Top1mt, revealed the profile of Top1mt binding in MEFs. Top1mt binding hot-spots do not, however, seem to correlate with regions of high ribonucleotide incorporation in mtDNA. Moreover, an increased abundance of Top1mt* in the NCR and rRNA regions suggests that Top1mt is chiefly involved in releasing torsional stresses at regions of high transcription (Dalla Rosa et al., 2014). This implies that Top1mt is not actively binding at sites of ribonucleotide incorporation, however this may be because the steady state levels of ribonucleotides in mtDNA from cultured cells is very low, and below the threshold above which they become problematic. *In vitro* analysis of Top1mt would be required to first of all demonstrate whether it has ribonucleotide binding and nicking activity akin to Top1.

6.10 Concluding Remarks and Future Prospects

Within this thesis I have adapted and developed the EmRiboseq protocol published by Ding *et al.* (Ding et al., 2015) for use with mtDNA and using this revealed the profile of ribonucleotide incorporation in mouse mtDNA *in vivo.* The data showed that the vast majority of ribonucleotides in mtDNA of solid tissues are rAMPs, and they are present throughout the genome, but they are mostly benign.

There are far fewer ribonucleotides in proliferating cells, in part due to the higher concentration of dNTPs. Nonetheless, an explanation for the under-representation of rAMPs in mtDNA from proliferating cells is wanting, hinting that there is more to be uncovered in the mechanism of ribonucleotide incorporation and metabolism in mtDNA.

The striking changes in ribonucleotide incorporation in mtDNA associated with Mpv17 deficiency raise the question of whether this phenomenon could explain other currently obscure causes of mtDNA disease. In light of the data presented within this thesis, many aspects of mtDNA disorders should be reconsidered in light of the fact that a striking abnormality, embedded ribonucleotides, has hitherto been largely overlooked.

By manipulating the chemistry of the EmRiboseq protocol, the method was adapted to map another form of RNA incorporation; primers. This was a preliminary example of how adaptable the EmRiboseq technique is and demonstrates its potential to map a multitude of DNA modifications *in vivo*. Regarding mtDNA, this could massively benefit the understanding of mtDNA metabolism with respect to damage, repair, ageing and disease. I hope that EmRiboseq will continue to be adapted and tailored to such avenues.

Reference List

AKMAN, G., DESAI, R., BAILEY, L. J., YASUKAWA, T., DALLA ROSA, I., DURIGON, R., HOLMES, J. B., MOSS, C. F., MENNUNI, M., HOULDEN, H., CROUCH, R. J., HANNA, M. G., PITCEATHLY, R. D. S., SPINAZZOLA, A. & HOLT, I. J. 2016. Pathological ribonuclease H1 causes R-loop depletion and aberrant DNA segregation in mitochondria. *Proceedings of the National Academy of Sciences,* 113**,** E4276-E4285.

ALSTON, C. L., SCHAEFER, A. M., RAMAN, P., SOLAROLI, N., KRISHNAN, K. J., BLAKELY, E. L., HE, L., CRAIG, K., ROBERTS, M., VYAS, A., NIXON, J., HORVATH, R., TURNBULL, D. M., KARLSSON, A., GORMAN, G. S. & TAYLOR, R. W. 2013. Late-onset respiratory failure due to TK2 mutations causing multiple mtDNA deletions. *Neurology,* 81**,** 2051-3.

AMES, B. N. 1989. Endogenous Oxidative DNA Damage, Aging, and Cancer. *Free Radical Research Communications,* 7**,** 121-128.

AMUNTS, A., BROWN, A., BAI, X. C., LLACER, J. L., HUSSAIN, T., EMSLEY, P., LONG, F., MURSHUDOV, G., SCHERES, S. H. W. & RAMAKRISHNAN, V. 2014. Structure of the yeast mitochondrial large ribosomal subunit. *Science,* 343**,** 1485-1489.

ANDERSSON, S. G., ZOMORODIPOUR, A., ANDERSSON, J. O., SICHERITZ-PONTEN, T., ALSMARK, U. C., PODOWSKI, R. M., NASLUND, A. K., ERIKSSON, A. S., WINKLER, H. H. & KURLAND, C. G. 1998. The genome sequence of Rickettsia prowazekii and the origin of mitochondria. *Nature,* 396**,** 133-40.

ANTONENKOV, V. D., ISOMURSU, A., MENNERICH, D., VAPOLA, M. H., WEIHER, H., KIETZMANN, T. & HILTUNEN, J. K. 2015. The Human Mitochondrial DNA Depletion Syndrome Gene MPV17 Encodes a Non-selective Channel That Modulates Membrane Potential. *The Journal of Biological Chemistry,* 290**,** 13840-13861.

ARUDCHANDRAN, A., CERRITELLI, S., NARIMATSU, S., ITAYA, M., SHIN, D. Y., SHIMADA, Y. & CROUCH, R. J. 2000. The absence of ribonuclease H1 or H2 alters the sensitivity of Saccharomyces cerevisiae to hydroxyurea, caffeine and ethyl methanesulphonate: implications for roles of RNases H in DNA replication and repair. *Genes Cells,* 5**,** 789-802.

BACMAN, S. R., WILLIAMS, S. L. & MORAES, C. T. 2009. Intra- and inter-molecular recombination of mitochondrial DNA after in vivo induction of multiple double-strand breaks. *Nucleic Acids Res,* 37**,** 4218-26.

BACMAN, S. R., WILLIAMS, S. L., PINTO, M., PERALTA, S. & MORAES, C. T. 2013. Specific elimination of mutant mitochondrial genomes in patient-derived cells by mitoTALENs. *Nat Med,* 19**,** 1111-3.

BAILEY, L. J., CLUETT, T. J., REYES, A., PROLLA, T. A., POULTON, J., LEEUWENBURGH, C. & HOLT, I. J. 2009. Mice expressing an error-prone DNA polymerase in mitochondria display elevated replication pausing and chromosomal breakage at fragile sites of mitochondrial DNA. *Nucleic acids research,* 37**,** 2327-35.

BAN, C., RAMAKRISHNAN, B. & SUNDARALINGAM, M. 1994. A single 2'-hydroxyl group converts B-DNA to A-DNA. Crystal structure of the DNA-RNA chimeric decamer duplex d(CCGGC)r(G)d(CCGG) with a novel intermolecular G-C base-paired quadruplet. *J Mol Biol,* 236**,** 275-85.

BERGLUND, A.-K., NAVARRETE, C., ENGQVIST, M. K. M., HOBERG, E., SZILAGYI, Z., TAYLOR, R. W., GUSTAFSSON, C. M., FALKENBERG, M. & CLAUSEN, A. R. 2017. Nucleotide pools dictate the identity and frequency of ribonucleotide incorporation in mitochondrial DNA. *PLOS Genetics,* 13**,** e1006628-e1006628.

BIANCHI, J., RUDD, S. G., JOZWIAKOWSKI, S. K., BAILEY, L. J., SOURA, V., TAYLOR, E., STEVANOVIC, I., GREEN, A. J., STRACKER, T. H., LINDSAY, H. D. & DOHERTY, A. J. 2013. PrimPol bypasses UV photoproducts during eukaryotic chromosomal DNA replication. *Molecular Cell,* 52**,** 566-573.

BLAKELY, E. L., BUTTERWORTH, A., HADDEN, R. D., BODI, I., HE, L., MCFARLAND, R. & TAYLOR, R. W. 2012. MPV17 mutation causes neuropathy and leukoencephalopathy with multiple mtDNA deletions in muscle. *Neuromuscul Disord,* 22**,** 587-91.

BOESCH, P., IBRAHIM, N., DIETRICH, A. & LIGHTOWLERS, R. N. 2010. Membrane association of mitochondrial DNA facilitates base excision repair in mammalian mitochondria. *Nucleic Acids Res,* 38**,** 1478-88.

BOGENHAGEN, D. & CLAYTON, D. A. 1977. Mouse L cell mitochondrial DNA molecules are selected randomly for replication throughout the cell cycle. *Cell,* 11**,** 719-27.

BOGENHAGEN, D., GILLUM, A. M., MARTENS, P. A. & CLAYTON, D. A. 1979. Replication of Mouse L-cell Mitochondrial DNA. *Cold Spring Harbor Symposia on Quantitative Biology,* 43**,** 253-262.

BOURDON, A., MINAI, L., SERRE, V. R., JAIS, J.-P., SARZI, E., AUBERT, S., CHR?TIEN, D., DE LONLAY, P., PAQUIS-FLUCKLINGER, V. R., ARAKAWA, H., NAKAMURA, Y., MUNNICH, A. & R?TIG, A. S. 2007. Mutation of RRM2B, encoding p53-controlled ribonucleotide reductase (p53R2), causes severe mitochondrial DNA depletion. *Nature Genetics,* 39**,** 776-780.

BOWMAKER, M., YANG, M. Y., YASUKAWA, T., REYES, A., JACOBS, H. T., HUBERMAN, J. A. & HOLT, I. J. 2003. Mammalian mitochondrial DNA replicates bidirectionally from an initiation zone. *The Journal of Biological Chemistry,* 278**,** 50961-50969.

BRAND, M. D., AFFOURTIT, C., ESTEVES, T. C., GREEN, K., LAMBERT, A. J., MIWA, S., PAKAY, J. L. & PARKER, N. 2004. Mitochondrial superoxide: production, biological effects, and activation of uncoupling proteins. *Free Radical Biology and Medicine,* 37**,** 755-767.

BRENNICKE, A. & CLAYTON, D. A. 1981. Nucleotide assignment of alkali-sensitive sites in mouse mitochondrial DNA. *J Biol Chem,* 256**,** 10613-7.

BROWN, A., AMUNTS, A., BAI, X. C., SUGIMOTO, Y., EDWARDS, P. C., MURSHUDOV, G., SCHERES, S. H. W. & RAMAKRISHNAN, V. 2014. Structure of the large ribosomal subunit from human mitochondria. *Science,* 346**,** 718-722.

BROWN, D. T., SAMUELS, D. C., MICHAEL, E. M., TURNBULL, D. M. & CHINNERY, P. F. 2001. Random Genetic Drift Determines the Level of Mutant mtDNA in Human Primary Oocytes. *The American Journal of Human Genetics,* 68**,** 533-536.

BROWN, T. A., CECCONI, C., TKACHUK, A. N., BUSTAMANTE, C. & CLAYTON, D. A. 2005. Replication of mitochondrial DNA occurs by strand displacement with alternative light-strand origins, not via a strand-coupled mechanism. *Genes & development,* 19**,** 2466-76.

BROWN, T. A., TKACHUK, A. N. & CLAYTON, D. A. 2008. Native R-loops Persist throughout the Mouse Mitochondrial DNA Genome. *Journal of Biological Chemistry,* 283**,** 36743-36751.

BUFFENSTEIN, R., EDREY, Y. H., YANG, T. & MELE, J. 2008. The oxidative stress theory of aging: embattled or invincible? Insights from non-traditional model organisms. *Age (Dordrecht, Netherlands),* 30**,** 99-109.

CAO, L., SHITARA, H., HORII, T., NAGAO, Y., IMAI, H., ABE, K., HARA, T., HAYASHI, J.-I. & YONEKAWA, H. 2007. The mitochondrial bottleneck occurs without reduction of mtDNA content in female mouse germ cells. *Nature Genetics,* 39**,** 386-390.

CAO, L., SHITARA, H., SUGIMOTO, M., HAYASHI, J.-I., ABE, K. & YONEKAWA, H. 2009. New Evidence Confirms That the Mitochondrial Bottleneck Is Generated without Reduction of Mitochondrial DNA Content in Early Primordial Germ Cells of Mice. *PLoS Genetics,* 5**,** e1000756-e1000756.

CAPPS, G. J., SAMUELS, D. C. & CHINNERY, P. F. 2003. A model of the nuclear control of mitochondrial DNA replication. *Journal of theoretical biology,* 221**,** 565-583.

CERRITELLI, S. M. & CROUCH, R. J. 2009. Ribonuclease H: the enzymes in eukaryotes. *Febs j,* 276**,** 1494-505.

CERRITELLI, S. M. & CROUCH, R. J. 2016. The Balancing Act of Ribonucleotides in DNA. *Trends in Biochemical Sciences,* 41**,** 434-445.

CERRITELLI, S. M., FROLOVA, E. G., FENG, C., GRINBERG, A., LOVE, P. E. & CROUCH, R. J. 2003. Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. *Mol Cell,* 11.

CHANDEL, N. S. & SCHUMACKER, P. T. 1999. Cells depleted of mitochondrial DNA (ρ0) yield insight into physiological mechanisms. *FEBS Letters,* 454**,** 173-176.

CHANG, D. D. & CLAYTON, D. A. 1985. Priming of human mitochondrial DNA replication occurs at the light-strand promoter. *Proc Natl Acad Sci U S A,* 82**,** 351-5.

CHANG, D. D. & CLAYTON, D. A. 1986. Precise assignment of the light-strand promoter of mouse mitochondrial DNA: a functional promoter consists of multiple upstream domains. *Molecular and Cellular Biology,* 6**,** 3253-3261.

CHINNERY, P. F., DIMAURO, S., SHANSKE, S., SCHON, E. A., ZEVIANI, M., MARIOTTI, C., CARRARA, F., LOMBES, A., LAFORET, P., OGIER, H., JAKSCH, M., LOCHMULLER, H., HORVATH, R., DESCHAUER, M., THORBURN, D. R., BINDOFF, L. A., POULTON, J., TAYLOR, R. W., MATTHEWS, J. N. & TURNBULL, D. M. 2004. Risk of developing a mitochondrial DNA deletion disorder. *Lancet,* 364**,** 592-6.

CHIU, H.-C., KOH, K. D., EVICH, M., LESIAK, A. L., GERMANN, M. W., BONGIORNO, A., RIEDO, E. & STORICI, F. 2014. RNA intrusions change DNA elastic properties and structure. *Nanoscale,* 6**,** 10009-17.

CLARK, J. M., JOYCE, C. M. & BEARDSLEY, G. P. 1987. Novel blunt-end addition reactions catalyzed by DNA polymerase I of Escherichia coli. *J Mol Biol,* 198**,** 123-7.

CLAUSEN, A. R., LUJAN, S. A., BURKHOLDER, A. B., OREBAUGH, C. D., WILLIAMS, J. S., CLAUSEN, M. F., MALC, E. P., MIECZKOWSKI, P. A., FARGO, D. C., SMITH, D. J. & KUNKEL, T. A. 2015. Tracking replication enzymology in vivo by genome-wide mapping of ribonucleotide incorporation. *Nature Structural & Molecular Biology,* 22**,** 185-191.

CLAUSEN, A. R., MURRAY, M. S., PASSER, A. R., PEDERSEN, L. C. & KUNKEL, T. A. 2013. Structure-function analysis of ribonucleotide bypass by B family DNA replicases. *Proceedings of the National Academy of Sciences of the United States of America,* 110**,** 16802-7.

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

COPELAND, W. C. & LONGLEY, M. J. 2003. DNA polymerase gamma in mitochondrial DNA replication and repair. *TheScientificWorldJournal,* 3**,** 34-44.

CORTOPASSI, G. A., SHIBATA, D., SOONG, N. W. & ARNHEIM, N. 1992. A pattern of accumulation of a somatic deletion of mitochondrial DNA in aging human tissues.

Proceedings of the National Academy of Sciences of the United States of America, 89**,** 7370-4.

CREE, L. M., SAMUELS, D. C., DE SOUSA LOPES, S. C., RAJASIMHA, H. K., WONNAPINIJ, P., MANN, J. R., DAHL, H.-H. M. & CHINNERY, P. F. 2008. A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes. *Nature Genetics,* 40**,** 249-254.

CROTT, J. W., CHOI, S. W., BRANDA, R. F. & MASON, J. B. 2005. Accumulation of mitochondrial DNA deletions is age, tissue and folate-dependent in rats. *Mutat Res,* 570**,** 63-70.

CURTH, U., URBANKE, C., GREIPEL, J., GERBERDING, H., TIRANTI, V. & ZEVIANI, M. 1994. Single-stranded-DNA-binding proteins from human mitochondria and Escherichia coli have analogous physicochemical properties. *Eur J Biochem,* 221**,** 435-43.

DALGAARD, J. Z. 2012. Causes and consequences of ribonucleotide incorporation into nuclear DNA. *Trends in genetics : TIG,* 28**,** 592-7.

DALLA ROSA, I., CÁMARA, Y., DURIGON, R., MOSS, C. F., VIDONI, S., AKMAN, G., HUNT, L., JOHNSON, M. A., GROCOTT, S., WANG, L., THORBURN, D. R., HIRANO, M., POULTON, J., TAYLOR, R. W., ELGAR, G., MARTÍ, R., VOSHOL, P., HOLT, I. J. & SPINAZZOLA, A. 2016. MPV17 Loss Causes Deoxynucleotide Insufficiency and Slow DNA Replication in Mitochondria. *PLOS Genetics,* 12**,** e1005779-e1005779.

DALLA ROSA, I., HUANG, S. Y. N., AGAMA, K., KHIATI, S., ZHANG, H. & POMMIER, Y. 2014. Mapping Topoisomerase Sites in Mitochondrial DNA with a Poisonous Mitochondrial Topoisomerase I (Top1mt). *Journal of Biological Chemistry,* 289**,** 18595-18602.

DALLABONA, C., MARSANO, R. M., ARZUFFI, P., GHEZZI, D., MANCINI, P., ZEVIANI, M., FERRERO, I. & DONNINI, C. 2010. Sym1, the yeast ortholog of the MPV17 human disease protein, is a stress-induced bioenergetic and morphogenetic mitochondrial modulator. *Hum Mol Genet,* 19**,** 1098-107.

DE SILVA, D., TU, Y.-T., AMUNTS, A., FONTANESI, F. & BARRIENTOS, A. 2015. Mitochondrial ribosome assembly in health and disease. *Cell Cycle,* 14**,** 2226-2250.

DE SOUZA-PINTO, N. C., EIDE, L., HOGUE, B. A., THYBO, T., STEVNSNER, T., SEEBERG, E., KLUNGLAND, A. & BOHR, V. A. 2001. Repair of 8-oxodeoxyguanosine lesions in mitochondrial dna depends on the oxoguanine dna glycosylase (OGG1) gene and 8-oxoguanine accumulates in the mitochondrial dna of OGG1-defective mice. *Cancer Res,* 61**,** 5378-81.

DE VIVO, M., DAL PERARO, M. & KLEIN, M. L. 2008. Phosphodiester cleavage in ribonuclease H occurs via an associative two-metal-aided catalytic mechanism. *J Am Chem Soc,* 130**,** 10955-62.

DEROSE, E. F., PERERA, L., MURRAY, M. S., KUNKEL, T. A. & LONDON, R. E. 2012. Solution Structure of the Dickerson DNA Dodecamer Containing a Single Ribonucleotide. *Biochemistry,* 51**,** 2407-2416.

DING, J., TAYLOR, M. S., JACKSON, A. P. & REIJNS, M. A. M. 2015. Genome-wide mapping of embedded ribonucleotides and other noncanonical nucleotides using emRiboSeq and EndoSeq. *Nature Protocols,* 10**,** 1433-1444.

DING, W.-X. & YIN, X.-M. 2012. Mitophagy: mechanisms, pathophysiological roles, and analysis. *Biological chemistry,* 393**,** 547-564.

DOUARRE, C., SOURBIER, C., DALLA ROSA, I., BRATA DAS, B., REDON, C. E., ZHANG, H., NECKERS, L., POMMIER, Y., BONAWITZ, N. D., CLAYTON, D. A., SHADEL, G. S., BAILEY, C. M., ANDERSON, K. S., CHAMPOUX, J. J., POMMIER, Y., ZHANG, H., BARCELO, J. M., LEE, B., KOHLHAGEN, G., ZIMONJIC, D. B., WANG, Y., LYU, Y. L., WANG, J. C., LOW, R. L., ORTON, S.,

FRIEDMAN, D. B., ZHANG, H., MENG, L. H., ZIMONJIC, D. B., POPESCU, N. C., POMMIER, Y., ZHANG, H., POMMIER, Y., SAS, K., ROBOTKA, H., TOLDI, J., VECSEI, L., WALLACE, D. C., WARBURG, O., KANG, D., HAMASAKI, N., SCHAPIRA, A. H., SCARPULLA, R. C., LEE, H. C., WEI, Y. H., LI, F., WANG, Y., ZELLER, K. I., POTTER, J. J., WONSEY, D. R., KELLY, D. P., SCARPULLA, R. C., BUTOW, R. A., AVADHANI, N. G., EROL, A., LEE, H. C., YIN, P. H., CHI, C. W., WEI, Y. H., FU, X., WAN, S., LYU, Y. L., LIU, L. F., QI, H., KLUZA, J., MARCHETTI, P., GALLEGO, M. A., LANCEL, S., FOURNIER, C., GONG, B., CHEN, Q., ALMASAN, A., BRAND, M. D., NICHOLLS, D. G., TONDERA, D., GRANDEMANGE, S., JOURDAIN, A., KARBOWSKI, M., MATTENBERGER, Y., CRUZ, S. D., PARONE, P. A., GONZALO, P., BIENVENUT, W. V., TONDERA, D., EATON, J. S., LIN, Z. P., SARTORELLI, A. C., BONAWITZ, N. D., SHADEL, G. S., VIRBASIUS, J. V., SCARPULLA, R. C., SCARPULLA, R. C., ZOPPOLI, G., DOUARRE, C., ROSA, I. D., LIU, H., REINHOLD, W., PATEL, A. J., LAURITZEN, I., LAZDUNSKI, M., et al. 2012. Mitochondrial Topoisomerase I is Critical for Mitochondrial Integrity and Cellular Energy Metabolism. *PLoS ONE,* 7**,** e41094-e41094.

DUNBAR, D. R., MOONIE, P. A., JACOBS, H. T. & HOLT, I. J. 1995. Different cellular backgrounds confer a marked advantage to either mutant or wild-type mitochondrial genomes. *Proc Natl Acad Sci U S A,* 92**,** 6562-6.

EGAN, D. F., SHACKELFORD, D. B., MIHAYLOVA, M. M., GELINO, S., KOHNZ, R. A., MAIR, W., VASQUEZ, D. S., JOSHI, A., GWINN, D. M., TAYLOR, R., ASARA, J. M., FITZPATRICK, J., DILLIN, A., VIOLLET, B., KUNDU, M., HANSEN, M. & SHAW, R. J. 2011. Phosphorylation of ULK1 (hATG1) by AMP-Activated Protein Kinase Connects Energy Sensing to Mitophagy. *Science,* 331**,** 456.

EGLI, M., USMAN, N. & RICH, A. 1993. Conformational influence of the ribose 2'-hydroxyl group: crystal structures of DNA-RNA chimeric duplexes. *Biochemistry,* 32**,** 3221-37.

EL SHOURBAGY, S. H., SPIKINGS, E. C., FREITAS, M. & ST JOHN, J. C. 2006. Mitochondria directly influence fertilisation outcome in the pig. *Reproduction,* 131**,** 233-45.

ELLIOTT, H. R., SAMUELS, D. C., EDEN, J. A., RELTON, C. L. & CHINNERY, P. F. 2008. Pathogenic Mitochondrial DNA Mutations Are Common in the General Population. *The American Journal of Human Genetics,* 83**,** 254-260.

EVICH, M., SPRING-CONNELL, A. M., STORICI, F. & GERMANN, M. W. 2016. Structural Impact of Single Ribonucleotide Residues in DNA. *Chembiochem,* 17**,** 1968-1977.

FARR, C. L., MATSUSHIMA, Y., LAGINA, A. T., 3RD, LUO, N. & KAGUNI, L. S. 2004. Physiological and biochemical defects in functional interactions of mitochondrial DNA polymerase and DNA-binding mutants of single-stranded DNA-binding protein. *J Biol Chem,* 279**,** 17047-53.

FERNANDEZ-SILVA, P., MARTINEZ-AZORIN, F., MICOL, V. & ATTARDI, G. 1997. The human mitochondrial transcription termination factor (mTERF) is a multizipper protein but binds to DNA as a monomer, with evidence pointing to intramolecular leucine zipper interactions. *Embo j,* 16**,** 1066-79.

FERRARO, P., PONTARIN, G., CROCCO, L., FABRIS, S., REICHARD, P. & BIANCHI, V. 2005. Mitochondrial Deoxynucleotide Pools in Quiescent Fibroblasts: A POSSIBLE MODEL FOR MITOCHONDRIAL NEUROGASTROINTESTINAL ENCEPHALOMYOPATHY (MNGIE). *Journal of Biological Chemistry,* 280**,** 24472-24480.

FONSECA, M. M., HARRIS, D. J., POSADA, D., DONATH, A. & MIDDENDORF, M. 2014. The Inversion of the Control Region in Three Mitogenomes Provides

Further Evidence for an Asymmetric Model of Vertebrate mtDNA Replication. *PLoS ONE,* 9**,** e106654-e106654.

FRANGINI, M., FRANZOLIN, E., CHEMELLO, F., LAVEDER, P., ROMUALDI, C., BIANCHI, V. & RAMPAZZO, C. 2013. Synthesis of Mitochondrial DNA Precursors during Myogenesis, an Analysis in Purified C2C12 Myotubes. *The Journal of Biological Chemistry,* 288**,** 5624-5635.

FRANK, P., BRAUNSHOFER-REITER, C., KARWAN, A., GRIMM, R. & WINTERSBERGER, U. 1999. Purification of Saccharomyces cerevisiae RNase H(70) and identification of the corresponding gene. *FEBS Letters,* 450**,** 251-256.

FURIHATA, C. 2015. An active alternative splicing isoform of human mitochondrial 8-oxoguanine DNA glycosylase (OGG1). *Genes and Environment,* 37**,** 21.

GARBER, K. 2006. Energy Deregulation: Licensing Tumors to Grow. *Science,* 312**,** 1158.

GARCÍA-GÓMEZ, S., REYES, A., MARTÍNEZ-JIMÉNEZ, M. I., CHOCRÓN, E. S., MOURÓN, S., TERRADOS, G., POWELL, C., SALIDO, E., MÉNDEZ, J., HOLT, I. J. & BLANCO, L. 2013. PrimPol, an archaic primase/polymerase operating in human cells. *Molecular Cell,* 52**,** 541-553.

GHODGAONKAR, MEDINI M., LAZZARO, F., OLIVERA-PIMENTEL, M., ARTOLA-BORÁN, M., CEJKA, P., REIJNS, MARTIN A., JACKSON, ANDREW P., PLEVANI, P., MUZI-FALCONI, M. & JIRICNY, J. 2013. Ribonucleotides Misincorporated into DNA Act as Strand-Discrimination Signals in Eukaryotic Mismatch Repair. *Molecular Cell,* 50**,** 323-332.

GOFFART, S., COOPER, H. M., TYYNISMAA, H., WANROOIJ, S., SUOMALAINEN, A. & SPELBRINK, J. N. 2009. Twinkle mutations associated with autosomal dominant progressive external ophthalmoplegia lead to impaired helicase function and in vivo mtDNA replication stalling. *Hum Mol Genet,* 18**,** 328-40.

GOTO, Y., NONAKA, I. & HORAI, S. 1990. A mutation in the tRNA(Leu)(UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature,* 348**,** 651-3.

GRAY, M. W. 2012. Mitochondrial Evolution. *Cold Spring Harbor Perspectives in Biology,* 4**,** a011403.

GRAY, M. W., BURGER, G. & LANG, B. F. 2001. The origin and early evolution of mitochondria. *Genome Biology,* 2**,** reviews1018.1-reviews1018.5.

GREBER, B. J., BOEHRINGER, D., LEIBUNDGUT, M., BIERI, P., LEITNER, A., SCHMITZ, N., AEBERSOLD, R. & BAN, N. 2014a. The complete structure of the large subunit of the mammalian mitochondrial ribosome. *Nature,* 515**,** 283-6.

GREBER, B. J., BOEHRINGER, D., LEITNER, A., BIERI, P., VOIGTS-HOFFMANN, F., ERZBERGER, J. P., LEIBUNDGUT, M., AEBERSOLD, R. & BAN, N. 2014b. Architecture of the large subunit of the mammalian mitochondrial ribosome. *Nature,* 505**,** 515-9.

GRIGUER, C. E., OLIVA, C. R. & GILLESPIE, G. Y. 2005. Glucose metabolism heterogeneity in human and mouse malignant glioma cell lines. *J Neurooncol,* 74**,** 123-33.

GROSS, N. J., GETZ, G. S. & RABINOWITZ, M. 1969. Apparent Turnover of Mitochondrial Deoxyribonucleic Acid and Mitochondrial Phospholipids in the Tissues of the Rat. *Journal of Biological Chemistry,* 244**,** 1552-1562.

GROSSMAN, L. I., WATSON, R. & VINOGRAD, J. 1973. The presence of ribonucleotides in mature closed-circular mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America,* 70**,** 3339-43.

GUILLIAM, T. A., JOZWIAKOWSKI, S. K., EHLINGER, A., BARNES, R. P., RUDD, S. G., BAILEY, L. J., SKEHEL, J. M., ECKERT, K. A., CHAZIN, W. J. & DOHERTY, A. J. 2014. Human PrimPol is a highly error-prone polymerase regulated by single-stranded DNA binding proteins. *Nucleic Acids Research*.

HARMAN, D. 1956. Aging: a theory based on free radical and radiation chemistry. *Journal of gerontology,* 11**,** 298-300.

HARMAN, D. 2009. Origin and evolution of the free radical theory of aging: a brief personal history, 1954–2009. *Biogerontology,* 10**,** 773-81.

HERBST, A., PAK, J. W., MCKENZIE, D., BUA, E., BASSIOUNI, M. & AIKEN, J. M. 2007. Accumulation of Mitochondrial DNA Deletion Mutations in Aged Muscle Fibers: Evidence for a Causal Role in Muscle Fiber Loss. *The journals of gerontology. Series A, Biological sciences and medical sciences,* 62**,** 235-245.

HESS, J. F., PARISI, M. A., BENNETT, J. L. & CLAYTON, D. A. 1991. Impairment of mitochondrial transcription termination by a point mutation associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature,* 351**,** 236-239.

HILL, J. W., HAZRA, T. K., IZUMI, T. & MITRA, S. 2001. Stimulation of human 8-oxoguanine-DNA glycosylase by AP-endonuclease: potential coordination of the initial steps in base excision repair. *Nucleic Acids Research,* 29**,** 430-438.

HOLMES, J. B., AKMAN, G., WOOD, S. R., SAKHUJA, K., CERRITELLI, S. M., MOSS, C., BOWMAKER, M. R., JACOBS, H. T., CROUCH, R. J. & HOLT, I. J. 2015. Primer retention owing to the absence of RNase H1 is catastrophic for mitochondrial DNA replication. *Proceedings of the National Academy of Sciences of the United States of America,* 112.

HOLT, I. J. 2009. Mitochondrial DNA replication and repair: all a flap. *Trends in Biochemical Sciences,* 34**,** 358-365.

HOLT, I. J., DUNBAR, D. R. & JACOBS, H. T. 1997. Behaviour of a Population of Partially Duplicated Mitochondrial DNA Molecules in Cell Culture: Segregation, Maintenance and Recombination Dependent Upon Nuclear Background. *Human Molecular Genetics,* 6**,** 1251-1260.

HOLT, I. J., HARDING, A. E. & MORGAN-HUGHES, J. A. 1988. Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. *Nature,* 331**,** 717-9.

HOLT, I. J. & JACOBS, H. T. 2014. Unique features of DNA replication in mitochondria: a functional and evolutionary perspective. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology,* 36**,** 1024-1031.

HOLT, I. J., LORIMER, H. E. & JACOBS, H. T. 2000. Coupled Leading- and Lagging-Strand Synthesis of Mammalian Mitochondrial DNA. *Cell,* 100**,** 515-524.

HOLT, I. J. & REYES, A. 2012. Human mitochondrial DNA replication. *Cold Spring Harbor Perspectives in Biology,* 4.

HOUGHTON, F. D. 2006. Energy metabolism of the inner cell mass and trophectoderm of the mouse blastocyst. *Differentiation,* 74**,** 11-8.

HUANG, S.-Y. N., GHOSH, S. & POMMIER, Y. 2015. Topoisomerase I Alone Is Sufficient to Produce Short DNA Deletions and Can Also Reverse Nicks at Ribonucleotide Sites. *Journal of Biological Chemistry,* 290**,** 14068-14076.

HUANG, S.-Y. N., WILLIAMS, J. S., ARANA, M. E., KUNKEL, T. A. & POMMIER, Y. 2017. Topoisomerase I-mediated cleavage at unrepaired ribonucleotides generates DNA double- strand breaks. *The EMBO Journal,* 36**,** 361-373.

HYSLOP, L. A., BLAKELEY, P., CRAVEN, L., RICHARDSON, J., FOGARTY, N. M. E., FRAGOULI, E., LAMB, M., WAMAITHA, S. E., PRATHALINGAM, N., ZHANG, Q., O'KEEFE, H., TAKEDA, Y., ARIZZI, L., ALFARAWATI, S., TUPPEN, H. A., IRVING, L., KALLEAS, D., CHOUDHARY, M., WELLS, D., MURDOCH, A. P., TURNBULL, D. M., NIAKAN, K. K. & HERBERT, M. 2016. Towards clinical application of pronuclear transfer to prevent mitochondrial DNA disease. *Nature,* 534**,** 383-386.

HYVARINEN, A. K., POHJOISMAKI, J. L., REYES, A., WANROOIJ, S., YASUKAWA, T., KARHUNEN, P. J., SPELBRINK, J. N., HOLT, I. J. & JACOBS, H. T. 2007.

The mitochondrial transcription termination factor mTERF modulates replication pausing in human mitochondrial DNA. *Nucleic Acids Res,* 35**,** 6458-74.

JASTROCH, M., DIVAKARUNI, A. S., MOOKERJEE, S., TREBERG, J. R. & BRAND, M. D. 2010. Mitochondrial proton and electron leaks. *Essays in biochemistry,* 47**,** 53-67.

JOHNSON, M. A., TURNBULL, D. M., DICK, D. J. & SHERRATT, H. S. 1983. A partial deficiency of cytochrome c oxidase in chronic progressive external ophthalmoplegia. *J Neurol Sci,* 60**,** 31-53.

JOSE, C., BELLANCE, N. & ROSSIGNOL, R. 2011. Choosing between glycolysis and oxidative phosphorylation: A tumor's dilemma? *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* 1807**,** 552-561.

JOURDAIN, A. A., BOEHM, E., MAUNDRELL, K. & MARTINOU, J.-C. 2016. Mitochondrial RNA granules: Compartmentalizing mitochondrial gene expression. *The Journal of cell biology,* 212**,** 611-614.

JOURDAIN, ALEXIS A., KOPPEN, M., WYDRO, M., RODLEY, CHRIS D., LIGHTOWLERS, ROBERT N., CHRZANOWSKA-LIGHTOWLERS, ZOFIA M. & MARTINOU, J.-C. 2013. GRSF1 Regulates RNA Processing in Mitochondrial RNA Granules. *Cell Metabolism,* 17**,** 399-410.

KANG, D., NISHIDA, J.-I., IYAMA, A., NAKABEPPU, Y., FURUICHI, M., FUJIWARA, T., SEKIGUCHI, M. & TAKESHIGE, K. 1995. Intracellular Localization of 8-Oxo-dGTPase in Human Cells, with Special Reference to the Role of the Enzyme in Mitochondria. *Journal of Biological Chemistry,* 270**,** 14659-14665.

KASIVISWANATHAN, R., COLLINS, T. R. L. & COPELAND, W. C. 2012. The interface of transcription and DNA replication in the mitochondria. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms,* 1819**,** 970-978.

KASIVISWANATHAN, R. & COPELAND, W. C. 2011. Ribonucleotide discrimination and reverse transcription by the human mitochondrial DNA polymerase. *The Journal of biological chemistry,* 286**,** 31490-500.

KAZAK, L., REYES, A. & HOLT, I. J. 2012. Minimizing the damage: repair pathways keep mitochondrial DNA intact. *Nature Reviews. Molecular Cell Biology,* 13**,** 659-671.

KESZTHELYI, A., DAIGAKU, Y., PTASIŃSKA, K., MIYABE, I. & CARR, A. M. 2015. Mapping ribonucleotides in genomic DNA and exploring replication dynamics by polymerase usage sequencing (Pu-seq). *Nature Protocols,* 10**,** 1786-1801.

KILIÇ, M., SIVRI, H. S., DURSUN, A., TOKATLI, A., DE MEIRLEIR, L., SENECA, S., AKÇÖREN, Z., YIĞIT, S., TOPALOĞLU, H. & COŞKUN, T. A novel mutation in the DGUOK gene in a Turkish newborn with mitochondrial depletion syndrome. *The Turkish journal of pediatrics,* 53**,** 79-82.

KIM, N., HUANG, S.-Y. N., WILLIAMS, J. S., LI, Y. C., CLARK, A. B., CHO, J.-E., KUNKEL, T. A., POMMIER, Y. & JINKS-ROBERTSON, S. 2011. Mutagenic Processing of Ribonucleotides in DNA by Yeast Topoisomerase I. *Science,* 332**,** 1561.

KOCHIWA, H., TOMITA, M. & KANAI, A. 2007. Evolution of ribonuclease H genes in prokaryotes to avoid inheritance of redundant genes. *BMC Evolutionary Biology,* 7**,** 128.

KOH, K. D., BALACHANDER, S., HESSELBERTH, J. R. & STORICI, F. 2015. Ribose-seq: global mapping of ribonucleotides embedded in genomic DNA. *Nature Methods,* 12**,** 251-257.

KOLESAR, J. E., WANG, C. Y., TAGUCHI, Y. V., CHOU, S. H. & KAUFMAN, B. A. 2013. Two-dimensional intact mitochondrial DNA agarose electrophoresis reveals the structural complexity of the mammalian mitochondrial genome. *Nucleic Acids Res,* 41**,** e58.

KOLLBERG, G., DARIN, N., BENAN, K., MOSLEMI, A.-R., LINDAL, S., TULINIUS, M. R., OLDFORS, A. & HOLME, E. 2009. A novel homozygous RRM2B missense mutation in association with severe mtDNA depletion. *Neuromuscular Disorders,* 19**,** 147-150.

KORHONEN, J. A., GASPARI, M. & FALKENBERG, M. 2003. TWINKLE Has 5′ → 3′ DNA Helicase Activity and Is Specifically Stimulated by Mitochondrial Single-stranded DNA-binding Protein. *Journal of Biological Chemistry,* 278**,** 48627-48632.

KORHONEN, J. A., PHAM, X. H., PELLEGRINI, M. & FALKENBERG, M. 2004. Reconstitution of a minimal mtDNA replisome in vitro. *The EMBO Journal,* 23**,** 2423-2429.

KORNBLUM, C., NICHOLLS, T. J., HAACK, T. B., SCHÖLER, S., PEEVA, V., DANHAUSER, K., HALLMANN, K., ZSURKA, G., RORBACH, J., IUSO, A., WIELAND, T., SCIACCO, M., RONCHI, D., COMI, G. P., MOGGIO, M., QUINZII, C. M., DIMAURO, S., CALVO, S. E., MOOTHA, V. K., KLOPSTOCK, T., STROM, T. M., MEITINGER, T., MINCZUK, M., KUNZ, W. S. & PROKISCH, H. 2013. Loss-of-function mutations in MGME1 impair mtDNA replication and cause multisystemic mitochondrial disease. *Nature GeNetics,* 2.

KRASICH, R. & COPELAND, W. C. 2017. DNA polymerases in the mitochondria: A critical review of the evidence. *Front Biosci (Landmark Ed),* 22**,** 692-709.

KRAUSS, J., ASTRINIDES, P., FROHNHÖFER, H. G., WALDERICH, B. & NÜSSLEIN-VOLHARD, C. 2013. transparent; a gene affecting stripe formation in Zebrafish, encodes the mitochondrial protein Mpv17 that is required for iridophore survival. *Biology Open*.

KRISHNAN, K. J., GREAVES, L. C., REEVE, A. K. & TURNBULL, D. 2007. The ageing mitochondrial genome. *Nucleic Acids Res,* 35**,** 7399-405.

KRUSE, B., NARASIMHAN, N. & ATTARDI, G. 1989. Termination of transcription in human mitochondria: Identification and purification of a DNA binding protein factor that promotes termination. *Cell,* 58**,** 391-397.

KRUSE, S. E., WATT, W. C., MARCINEK, D. J., KAPUR, R. P., SCHENKMAN, K. A. & PALMITER, R. D. 2008. Mice with Mitochondrial Complex I Deficiency Develop a Fatal Encephalomyopathy. *Cell Metabolism,* 7**,** 312-320.

KUJOTH, G. C., BRADSHAW, P. C., HAROON, S., PROLLA, T. A. & WALLACHER-SCHOLZ, B. 2007. The Role of Mitochondrial DNA Mutations in Mammalian Aging. *PLoS Genetics,* 3**,** e24-e24.

KUNKEL, T. A. & ERIE, D. A. 2015. Eukaryotic Mismatch Repair in Relation to DNA Replication. *Annual Review of Genetics,* 49**,** 291-313.

LANG, B. F., BURGER, G., O'KELLY, C. J., CEDERGREN, R., GOLDING, G. B., LEMIEUX, C., SANKOFF, D., TURMEL, M. & GRAY, M. W. 1997. An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature,* 387**,** 493-7.

LARSSON, N. G. 2010. Somatic mitochondrial DNA mutations in mammalian aging. *Annu Rev Biochem,* 79**,** 683-706.

LESHINSKY-SILVER, E., MALINGER, G., BEN-SIRA, L., KIDRON, D., COHEN, S., INBAR, S., BEZALELI, T., LEVINE, A., VINKLER, C., LEV, D. & LERMAN-SAGIE, T. 2011. A large homozygous deletion in the SAMHD1 gene causes atypical Aicardi?Gouti?res syndrome associated with mtDNA deletions. *European Journal of Human Genetics,* 19**,** 287-292.

LEWIS, W. & DALAKAS, M. C. 1995. Mitochondrial toxicity of antiviral drugs. *Nat Med,* 1**,** 417-22.

LIBERTI, M. V. & LOCASALE, J. W. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends in Biochemical Sciences,* 41**,** 211-218.

LIM, S. E., LONGLEY, M. J. & COPELAND, W. C. 1999. The Mitochondrial p55 Accessory Subunit of Human DNA Polymerase γ Enhances DNA Binding, Promotes Processive DNA Synthesis, and Confers N-Ethylmaleimide Resistance. *Journal of Biological Chemistry,* 274**,** 38197-38203.

LITONIN, D., SOLOGUB, M., SHI, Y., SAVKINA, M., ANIKIN, M., FALKENBERG, M., GUSTAFSSON, C. M. & TEMIAKOV, D. 2010. Human Mitochondrial Transcription Revisited ONLY TFAM AND TFB2M ARE REQUIRED FOR TRANSCRIPTION OF THE MITOCHONDRIAL GENES IN VITRO * □ S.

LODEIRO, M. F., UCHIDA, A., BESTWICK, M., MOUSTAFA, I. M., ARNOLD, J. J., SHADEL, G. S. & CAMERON, C. E. 2012. Transcription from the second heavy-strand promoter of human mtDNA is repressed by transcription factor A in vitro. *Proceedings of the National Academy of Sciences of the United States of America,* 109**,** 6513-8.

LONGLEY, M. J., CLARK, S., YU WAI MAN, C., HUDSON, G., DURHAM, S. E., TAYLOR, R. W., NIGHTINGALE, S., TURNBULL, D. M., COPELAND, W. C. & CHINNERY, P. F. 2006. Mutant POLG2 disrupts DNA polymerase gamma subunits and causes progressive external ophthalmoplegia. *Am J Hum Genet,* 78**,** 1026-34.

LONGLEY, M. J., PRASAD, R., SRIVASTAVA, D. K., WILSON, S. H. & COPELAND, W. C. 1998. Identification of 5'-deoxyribose phosphate lyase activity in human DNA polymerase gamma and its role in mitochondrial base excision repair in vitro. *Proc Natl Acad Sci U S A,* 95**,** 12244-8.

LOVETT, M. A., KATZ, L. & HELINSKI, D. R. 1974. Unidirectional replication of plasmid ColE1 DNA. *Nature,* 251**,** 337-340.

LUJAN, S. A., WILLIAMS, J. S., CLAUSEN, A. R., CLARK, A. B. & KUNKEL, T. A. 2013. Ribonucleotides are signals for mismatch repair of leading-strand replication errors. *Molecular cell,* 50**,** 437-43.

MAIER, D., FARR, C. L., POECK, B., ALAHARI, A., VOGEL, M., FISCHER, S., KAGUNI, L. S. & SCHNEUWLY, S. 2001. Mitochondrial single-stranded DNA-binding protein is required for mitochondrial DNA replication and development in Drosophila melanogaster. *Mol Biol Cell,* 12**,** 821-30.

MARCHLER-BAUER, A., ANDERSON, J. B., CHERUKURI, P. F., DEWEESE-SCOTT, C., GEER, L. Y., GWADZ, M., HE, S., HURWITZ, D. I., JACKSON, J. D., KE, Z., LANCZYCKI, C. J., LIEBERT, C. A., LIU, C., LU, F., MARCHLER, G. H., MULLOKANDOV, M., SHOEMAKER, B. A., SIMONYAN, V., SONG, J. S., THIESSEN, P. A., YAMASHITA, R. A., YIN, J. J., ZHANG, D. & BRYANT, S. H. 2005. CDD: a Conserved Domain Database for protein classification. *Nucleic Acids Res,* 33.

MASTERS, B. S., STOHL, L. L. & CLAYTON, D. A. 1987. Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7. *Cell,* 51**,** 89-99.

MATHEWS, C. K. 2014. Deoxyribonucleotides as genetic and metabolic regulators. *The FASEB Journal,* 28**,** 3832-3840.

MATHEWS, C. K. & SONG, S. 2007. Maintaining precursor pools for mitochondrial DNA replication. *The FASEB Journal,* 21**,** 2294-2303.

MICHIKAWA, Y., MAZZUCCHELLI, F., BRESOLIN, N., SCARLATO, G. & ATTARDI, G. 1999. Aging-dependent large accumulation of point mutations in the human mtDNA control region for replication. *Science,* 286**,** 774-9.

MIKHAILOV, V. S. & BOGENHAGEN, D. F. 1996. Effects of Xenopus laevis Mitochondrial Single-stranded DNA-binding Protein on Primer-Template Binding and 3′ →5′ Exonuclease Activity of DNA Polymerase γ. *Journal of Biological Chemistry,* 271**,** 18939-18946.

MILENKOVIC, D., MATIC, S., KÜHL, I., RUZZENENTE, B., FREYER, C., JEMT, E., PARK, C. B., FALKENBERG, M. & LARSSON, N.-G. 2013. TWINKLE is an essential mitochondrial helicase required for synthesis of nascent D-loop strands and complete mtDNA replication. *Human molecular genetics,* 22**,** 1983-93.

MINCZUK, M., PAPWORTH, M. A., MILLER, J. C., MURPHY, M. P. & KLUG, A. 2008. Development of a single-chain, quasi-dimeric zinc-finger nuclease for the selective degradation of mutated human mitochondrial DNA. *Nucleic Acids Res,* 36**,** 3926-38.

MIRALLES FUSTÉ, J., SHI, Y., WANROOIJ, S., ZHU, X., JEMT, E., PERSSON, O., SABOURI, N., GUSTAFSSON, C. M. & FALKENBERG, M. 2014. In vivo occupancy of mitochondrial single-stranded DNA binding protein supports the strand displacement mode of DNA replication. *PLoS genetics,* 10**,** e1004832-e1004832.

MISHRA, P. & CHAN, D. C. 2016. Metabolic regulation of mitochondrial dynamics. *The Journal of Cell Biology,* 212**,** 379-387.

MONTOYA, J., CHRISTIANSON, T., LEVENS, D., RABINOWITZ, M. & ATTARDI, G. 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**,** 7195-9.

MONTOYA, J., GAINES, G. L. & ATTARDI, G. 1983. The pattern of transcription of the human mitochondrial rRNA genes reveals two overlapping transcription units. *Cell,* 34**,** 151-9.

MOURÓN, S., RODRIGUEZ-ACEBES, S., MARTÍNEZ-JIMÉNEZ, M. I., GARCÍA-GÓMEZ, S., CHOCRÓN, S., BLANCO, L. & MÉNDEZ, J. 2013. Repriming of DNA synthesis at stalled replication forks by human PrimPol. *Nature Structural & Molecular Biology,* 20**,** 1383-1389.

MULLER, F. L., LUSTGARTEN, M. S., JANG, Y., RICHARDSON, A. & VAN REMMEN, H. 2007. Trends in oxidative aging theories. *Free radical biology & medicine,* 43**,** 477-503.

MURPHY, J. L., RATNAIKE, T. E., SHANG, E., FALKOUS, G., BLAKELY, E. L., ALSTON, C. L., TAIVASSALO, T., HALLER, R. G., TAYLOR, R. W. & TURNBULL, D. M. 2012. Cytochrome c oxidase-intermediate fibres: Importance in understanding the pathogenesis and treatment of mitochondrial myopathy. *Neuromuscular Disorders,* 22**,** 690-698.

MURPHY, MICHAEL P. 2009. How mitochondria produce reactive oxygen species. *Biochemical Journal,* 417**,** 1-13.

NARENDRA, D., TANAKA, A., SUEN, D.-F. & YOULE, R. J. 2008. Parkin is recruited selectively to impaired mitochondria and promotes their autophagy. *The Journal of Cell Biology,* 183**,** 795.

NASS, M. M. & NASS, S. 1963. INTRAMITOCHONDRIAL FIBERS WITH DNA CHARACTERISTICS. I. FIXATION AND ELECTRON STAINING REACTIONS. *J Cell Biol,* 19**,** 593-611.

NASS, M. M. K. 1980. Analysis of the two heavy and light strand origins and the direction of replication of mitochondrial DNA within a detailed physical map of this genome in transformed hamster cells. *Journal of Molecular Biology,* 140**,** 231-256.

NICHOLLS, T. J., ZSURKA, G., PEEVA, V., SCHÖLER, S., SZCZESNY, R. J., CYSEWSKI, D., REYES, A., KORNBLUM, C., SCIACCO, M., MOGGIO, M., DZIEMBOWSKI, A., KUNZ, W. S. & MINCZUK, M. 2014. Linear mtDNA fragments and unusual mtDNA rearrangements associated with pathological deficiency of MGME1 exonuclease. *Human Molecular Genetics,* 23**,** 6147-6162.

NICK MCELHINNY, S. A., KUMAR, D., CLARK, A. B., WATT, D. L., WATTS, B. E., LUNDSTROM, E. B., JOHANSSON, E., CHABES, A. & KUNKEL, T. A. 2010a.

Genome instability due to ribonucleotide incorporation into DNA. *Nat Chem Biol,* 6**,** 774-81.

NICK MCELHINNY, S. A., WATTS, B. E., KUMAR, D., WATT, D. L., LUNDSTRÖM, E.-B., BURGERS, P. M. J., JOHANSSON, E., CHABES, A. & KUNKEL, T. A. 2010b. Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proceedings of the National Academy of Sciences,* 107**,** 4949-4954.

NIGHTINGALE, H., PFEFFER, G., BARGIELA, D., HORVATH, R. & CHINNERY, P. F. 2016. Emerging therapies for mitochondrial disorders. *Brain,* 139**,** 1633-1648.

NISHIGAKI, Y., MARTÍ, R., COPELAND, W. C. & HIRANO, M. 2003. Site-specific somatic mitochondrial DNA point mutations in patients with thymidine phosphorylase deficiency. *The Journal of clinical investigation,* 111**,** 1913-21.

NISHINO, I., SPINAZZOLA, A. & HIRANO, M. 1999. Thymidine phosphorylase gene mutations in MNGIE, a human mitochondrial disorder. *Science (New York, N.Y.),* 283**,** 689-92.

NOWOTNY, M., CERRITELLI, S. M., GHIRLANDO, R., GAIDAMAKOV, S. A., CROUCH, R. J. & YANG, W. 2008. Specific recognition of RNA/DNA hybrid and enhancement of human RNase H1 activity by HBD. *Embo j,* 27**,** 1172-81.

OHTANI, N., HARUKI, M., MORIKAWA, M. & KANAYA, S. 1999. Molecular diversities of RNases H. *J Biosci Bioeng,* 88**,** 12-9.

PALLAN, P. S. & EGLI, M. 2008. Insights into RNA/DNA hybrid recognition and processing by RNase H from the crystal structure of a non-specific enzyme-dsDNA complex. *Cell cycle (Georgetown, Tex.),* 7**,** 2562-2569.

PAREDES, J. A., ZHOU, X., HÖGLUND, S. & KARLSSON, A. 2013. Gene Expression Deregulation in Postnatal Skeletal Muscle of TK2 Deficient Mice Reveals a Lower Pool of Proliferating Myogenic Progenitor Cells. *PLOS ONE,* 8**,** e53698.

PARK, C. B. & LARSSON, N.-G. 2011. Mitochondrial DNA mutations in disease and aging. *The Journal of Cell Biology,* 193.

PAVLOV, Y. I., MINNICK, D. T., IZUTA, S. & KUNKEL, T. A. 1994. DNA replication fidelity with 8-oxodeoxyguanosine triphosphate. *Biochemistry,* 33**,** 4695-701.

PAYNE, B. A. I., WILSON, I. J., YU-WAI-MAN, P., COXHEAD, J., DEEHAN, D., HORVATH, R., TAYLOR, R. W., SAMUELS, D. C., SANTIBANEZ-KOREF, M. & CHINNERY, P. F. 2013. Universal heteroplasmy of human mitochondrial DNA. *Human Molecular Genetics,* 22**,** 384-390.

PHILLIPS, A. F., MILLET, A. R., TIGANO, M., DUBOIS, S. M., CRIMMINS, H., BABIN, L., CHARPENTIER, M., PIGANEAU, M., BRUNET, E. & SFEIR, A. 2017. Single-Molecule Analysis of mtDNA Replication Uncovers the Basis of the Common Deletion. *Molecular Cell,* 65**,** 527-538.e6.

PILEGAARD, H., SALTIN, B. & NEUFER, P. D. 2003. Exercise induces transient transcriptional activation of the PGC-1α gene in human skeletal muscle. *The Journal of Physiology,* 546**,** 851-858.

PINZ, K. G. & BOGENHAGEN, D. F. 1998. Efficient repair of abasic sites in DNA by mitochondrial enzymes. *Mol Cell Biol,* 18**,** 1257-65.

PITCEATHLY, R. D. S., RAHMAN, S. & HANNA, M. G. 2012. Single deletions in mitochondrial DNA – Molecular mechanisms and disease phenotypes in clinical practice. *Neuromuscular Disorders,* 22**,** 577-586.

POHJOISMÄKI, J. L. O., GOFFART, S., TYYNISMAA, H., WILLCOX, S., IDE, T., KANG, D., SUOMALAINEN, A., KARHUNEN, P. J., GRIFFITH, J. D., HOLT, I. J. & JACOBS, H. T. 2009. Human heart mitochondrial DNA is organized in complex catenated networks containing abundant four-way junctions and replication forks. *The Journal of Biological Chemistry,* 284**,** 21446-21457.

POOVATHINGAL, S. K., GRUBER, J., LAKSHMANAN, L., HALLIWELL, B. & GUNAWAN, R. 2012. Is mitochondrial DNA turnover slower than commonly assumed? *Biogerontology,* 13**,** 557-64.

POTENSKI, C. J. & KLEIN, H. L. 2014. How the misincorporation of ribonucleotides into genomic DNA can be both harmful and helpful to cells. *Nucleic acids research,* 42**,** 10226-34.

RAMPAZZO, C., MIAZZI, C., FRANZOLIN, E., PONTARIN, G., FERRARO, P., FRANGINI, M., REICHARD, P. & BIANCHI, V. 2010. Regulation by degradation, a cellular defense against deoxyribonucleotide pool imbalances. *Mutation research,* 703**,** 2-10.

REDDY, P., OCAMPO, A., SUZUKI, K., LUO, J., BACMAN, SANDRA R., WILLIAMS, SION L., SUGAWARA, A., OKAMURA, D., TSUNEKAWA, Y., WU, J., LAM, D., XIONG, X., MONTSERRAT, N., ESTEBAN, CONCEPCION R., LIU, G.-H., SANCHO-MARTINEZ, I., MANAU, D., CIVICO, S., CARDELLACH, F., DEL MAR O'CALLAGHAN, M., CAMPISTOL, J., ZHAO, H., CAMPISTOL, JOSEP M., MORAES, CARLOS T. & IZPISUA BELMONTE, JUAN C. 2015. Selective Elimination of Mitochondrial Mutations in the Germline by Genome Editing. *Cell,* 161**,** 459-469.

REICHARD, P. 1988. Interactions between deoxyribonucleotide and DNA synthesis. *Annu Rev Biochem,* 57**,** 349-74.

REIJNS, M. A. M., RABE, B., RIGBY, R. E., MILL, P., ASTELL, K. R., LETTICE, L. A., BOYLE, S., LEITCH, A., KEIGHREN, M., KILANOWSKI, F., DEVENNEY, P. S., SEXTON, D., GRIMES, G., HOLT, I. J., HILL, R. E., TAYLOR, M. S., LAWSON, K. A., DORIN, J. R. & JACKSON, A. P. 2012. Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. *Cell,* 149**,** 1008-22.

REYES, A., KAZAK, L., WOOD, S. R., YASUKAWA, T., JACOBS, H. T. & HOLT, I. J. 2013. Mitochondrial DNA replication proceeds via a 'bootlace' mechanism involving the incorporation of processed transcripts. *Nucleic Acids Research,* 41**,** 5837-5850.

REYES, A., MELCHIONDA, L., NASCA, A., CARRARA, F., LAMANTEA, E., ZANOLINI, A., LAMPERTI, C., FANG, M., ZHANG, J., RONCHI, D., BONATO, S., FAGIOLARI, G., MOGGIO, M., GHEZZI, D. & ZEVIANI, M. 2015. RNASEH1 Mutations Impair mtDNA Replication and Cause Adult-Onset Mitochondrial Encephalomyopathy. *American Journal of Human Genetics,* 97**,** 186-193.

RICHTER, C. 1995. Oxidative damage to mitochondrial DNA and its relationship to ageing. *The International Journal of Biochemistry & Cell Biology,* 27**,** 647-653.

RICHTER-DENNERLEIN, R., DENNERLEIN, S. & REHLING, P. 2015. Integrating mitochondrial translation into the cellular context. *Nat Rev Mol Cell Biol,* 16**,** 586-592.

RIFAI, Z., WELLE, S., KAMP, C. & THORNTON, C. A. 1995. Ragged red fibers in normal aging and inflammatory myopathy. *Ann Neurol,* 37**,** 24-9.

RINGEL, R., SOLOGUB, M., MOROZOV, Y. I., LITONIN, D., CRAMER, P. & TEMIAKOV, D. 2011. Structure of human mitochondrial RNA polymerase. *Nature,* 478**,** 269-273.

ROBBERSON, D. L., KASAMATSU, H. & VINOGRAD, J. 1972. Replication of mitochondrial DNA. Circular replicative intermediates in mouse L cells. *Proceedings of the National Academy of Sciences of the United States of America,* 69**,** 737-41.

ROBERTI, M., POLOSA, P. L., BRUNI, F., MANZARI, C., DECEGLIE, S., GADALETA, M. N. & CANTATORE, P. 2009. The MTERF family proteins: Mitochondrial transcription regulators and beyond. *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* 1787**,** 303-311.

ROBERTSON, A. B., KLUNGLAND, A., ROGNES, T. & LEIROS, I. 2009. DNA repair in mammalian cells: Base excision repair: the long and short of it. *Cell Mol Life Sci,* 66**,** 981-93.

RONCHI, D., GARONE, C., BORDONI, A., GUTIERREZ RIOS, P., CALVO, S. E., RIPOLONE, M., RANIERI, M., RIZZUTI, M., VILLA, L., MAGRI, F., CORTI, S., BRESOLIN, N., MOOTHA, V. K., MOGGIO, M., DIMAURO, S., COMI, G. P. & SCIACCO, M. 2012. Next-generation sequencing reveals DGUOK mutations in adult patients with mitochondrial DNA multiple deletions. *Brain : a journal of neurology,* 135**,** 3404-15.

ROOS, S., LINDGREN, U., EHRSTEDT, C., MOSLEMI, A. R. & OLDFORS, A. 2014. Mitochondrial DNA depletion in single fibers in a patient with novel TK2 mutations. *Neuromuscular Disorders,* 24**,** 713-720.

ROPP, P. A. & COPELAND, W. C. 1996. Cloning and Characterization of the Human Mitochondrial DNA Polymerase, DNA Polymerase γ. *Genomics,* 36**,** 449-458.

RORBACH, J., GAO, F., POWELL, C. A., D'SOUZA, A., LIGHTOWLERS, R. N., MINCZUK, M. & CHRZANOWSKA-LIGHTOWLERS, Z. M. 2016. Human mitochondrial ribosomes can switch their structural RNA composition. *Proceedings of the National Academy of Sciences of the United States of America,* 113**,** 12198-12201.

ROSSI, M. N., CARBONE, M., MOSTOCOTTO, C., MANCONE, C., TRIPODI, M., MAIONE, R. & AMATI, P. 2009. Mitochondrial localization of PARP-1 requires interaction with mitofilin and is involved in the maintenance of mitochondrial DNA integrity. *J Biol Chem,* 284**,** 31616-24.

ROSSMANITH, W. & KARWAN, R. M. 1998. Characterization of human mitochondrial RNase P: novel aspects in tRNA processing. *Biochem Biophys Res Commun,* 247**,** 234-41.

ROSSMANITH, W., TULLO, A., POTUSCHAK, T., KARWAN, R. & SBIS, E. 1995. Human Mitochondrial tRNA Processing. *Journal of Biological Chemistry,* 270**,** 12885-12891.

ROSTA, E., NOWOTNY, M., YANG, W. & HUMMER, G. 2011. Catalytic Mechanism of RNA Backbone Cleavage by Ribonuclease H from Quantum Mechanics/Molecular Mechanics Simulations. *Journal of the American Chemical Society,* 133**,** 8934-8941.

ROTHSTEIN, R., MICHEL, B. & GANGLOFF, S. 2000. Replication fork pausing and recombination or “gimme a break”. *Genes & Development,* 14**,** 1-10.

RUCHKO, M. V., GORODNYA, O. M., ZULETA, A., PASTUKH, V. M. & GILLESPIE, M. N. 2011. The DNA Glycosylase, Ogg1, Defends Against Oxidant-induced mtDNA Damage and Apoptosis in Pulmonary Artery Endothelial Cells. *Free radical biology & medicine,* 50**,** 1107-1113.

RUHANEN, H., BORRIE, S., SZABADKAI, G., TYYNISMAA, H., JONES, A. W. E., KANG, D., TAANMAN, J.-W. & YASUKAWA, T. 2010. Mitochondrial single-stranded DNA binding protein is required for maintenance of mitochondrial DNA and 7S DNA but is not required for mitochondrial nucleoid organisation. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1803**,** 931-939.

SAADA, A., SHAAG, A., MANDEL, H., NEVO, Y., ERIKSSON, S. & ELPELEG, O. 2001. Mutant mitochondrial thymidine kinase in mitochondrial DNA depletion myopathy. *Nature Genetics,* 29**,** 342-344.

SAADA ‐ REISCH, A. 2004. Deoxyribonucleoside Kinases in Mitochondrial DNA Depletion. *Nucleosides, Nucleotides and Nucleic Acids,* 23**,** 1205-1215.

SCHATZ, G., HASLBRUNNER, E. & TUPPY, H. 1964. Deoxyribonucleic acid associated with yeast mitochondria. *Biochemical and Biophysical Research Communications,* 15**,** 127-132.

SCHON, E. A. & GILKERSON, R. W. 2010. Functional complementation of mitochondrial DNAs: mobilizing mitochondrial genetics against dysfunction. *Biochim Biophys Acta,* 1800**,** 245-9.

SCHWARTZ , M. & VISSING , J. 2002. Paternal Inheritance of Mitochondrial DNA. *New England Journal of Medicine,* 347**,** 576-580.

SHANG, J. & CLAYTON, D. A. 1994. Human mitochondrial transcription termination exhibits RNA polymerase independence and biased bipolarity in vitro. *Journal of Biological Chemistry,* 269**,** 29112-29120.

SHARMA, M. R., KOC, E. C., DATTA, P. P., BOOTH, T. M., SPREMULLI, L. L. & AGRAWAL, R. K. 2003. Structure of the mammalian mitochondrial ribosome reveals an expanded functional role for its component proteins. *Cell,* 115**,** 97-108.

SHIN, J., MING, G. L. & SONG, H. 2014. Decoding neural transcriptomes and epigenomes via high-throughput sequencing. *Nat Neurosci,* 17**,** 1463-75.

SHOFFNER, J. M., LOTT, M. T., LEZZA, A. M. S., SEIBEL, P., BALLINGER, S. W. & WALLACE, D. C. 1990. Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial DNA tRNALys mutation. *Cell,* 61**,** 931-937.

SHUTT, T. E. & GRAY, M. W. 2006. Bacteriophage origins of mitochondrial replication and transcription proteins. *Trends Genet,* 22**,** 90-5.

SHUTT, T. E., LODEIRO, M. F., COTNEY, J., CAMERON, C. E. & SHADEL, G. S. 2010. Core human mitochondrial transcription apparatus is a regulated two-component system in vitro. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 12133-8.

SOHAL, R. S., SOHAL, B. H. & ORR, W. C. 1995. Mitochondrial superoxide and hydrogen peroxide generation, protein oxidative damage, and longevity in different species of flies. *Free Radic Biol Med,* 19**,** 499-504.

SONG, S., PURSELL, Z. F., COPELAND, W. C., LONGLEY, M. J., KUNKEL, T. A. & MATHEWS, C. K. 2005. DNA precursor asymmetries in mammalian tissue mitochondria and possible contribution to mutagenesis through reduced replication fidelity. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 4990-5.

SPARKS, J. L., CHON, H., CERRITELLI, S. M., KUNKEL, T. A., JOHANSSON, E., CROUCH, R. J. & BURGERS, P. M. 2012. RNase H2-initiated ribonucleotide excision repair. *Molecular cell,* 47**,** 980-6.

SPELBRINK, J. N., LI, F. Y., TIRANTI, V., NIKALI, K., YUAN, Q. P., TARIQ, M., WANROOIJ, S., GARRIDO, N., COMI, G., MORANDI, L., SANTORO, L., TOSCANO, A., FABRIZI, G. M., SOMER, H., CROXEN, R., BEESON, D., POULTON, J., SUOMALAINEN, A., JACOBS, H. T., ZEVIANI, M. & LARSSON, C. 2001. Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria. *Nature genetics,* 28**,** 223-31.

SPINAZZOLA, A., VISCOMI, C., FERNANDEZ-VIZARRA, E., CARRARA, F., D'ADAMO, P., CALVO, S., MARSANO, R. M., DONNINI, C., WEIHER, H., STRISCIUGLIO, P., PARINI, R., SARZI, E., CHAN, A., DIMAURO, S., R?TIG, A., GASPARINI, P., FERRERO, I., MOOTHA, V. K., TIRANTI, V. & ZEVIANI, M. 2006. MPV17 encodes an inner mitochondrial membrane protein and is mutated in infantile hepatic mitochondrial DNA depletion. *Nature Genetics,* 38**,** 570-575.

STEIN, H. & HAUSEN, P. 1969. Enzyme from calf thymus degrading the RNA moiety of DNA-RNA Hybrids: effect on DNA-dependent RNA polymerase. *Science,* 166.

STEWART, J. B., FREYER, C., ELSON, J. L., WREDENBERG, A., CANSU, Z., TRIFUNOVIC, A. & LARSSON, N.-G. 2008. Strong Purifying Selection in Transmission of Mammalian Mitochondrial DNA. *PLOS Biology,* 6**,** e10.

SUN, B., LATHAM, K. A., DODSON, M. L. & LLOYD, R. S. 1995. Studies on the catalytic mechanism of five DNA glycosylases. Probing for enzyme-DNA imino intermediates. *J Biol Chem,* 270**,** 19501-8.

SYKORA, P., CROTEAU, D. L., BOHR, V. A. & WILSON, D. M., 3RD 2011. Aprataxin localizes to mitochondria and preserves mitochondrial function. *Proc Natl Acad Sci U S A,* 108**,** 7437-42.

SYKORA, P., KANNO, S., AKBARI, M., KULIKOWICZ, T., BAPTISTE, B. A., LEANDRO, G. S., LU, H., TIAN, J., MAY, A., BECKER, K. A., CROTEAU, D. L., WILSON, D. M., SOBOL, R. W., YASUI, A. & BOHR, V. A. 2017. DNA polymerase beta participates in mitochondrial DNA repair. *Molecular and Cellular Biology*.

SZCZESNY, B., TANN, A. W., LONGLEY, M. J., COPELAND, W. C. & MITRA, S. 2008. Long patch base excision repair in mammalian mitochondrial genomes. *J Biol Chem,* 283**,** 26349-56.

SZCZESNY, R. J., HEJNOWICZ, M. S., STECZKIEWICZ, K., MUSZEWSKA, A., BOROWSKI, L. S., GINALSKI, K. & DZIEMBOWSKI, A. 2013. Identification of a novel human mitochondrial endo-/exonuclease Ddk1/c20orf72 necessary for maintenance of proper 7S DNA levels. *Nucleic acids research,* 41**,** 3144-61.

TAANMAN, J.-W. 1999. The mitochondrial genome: structure, transcription, translation and replication. *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* 1410**,** 103-123.

TAYLOR, R. W., SCHAEFER, A. M., BARRON, M. J., MCFARLAND, R. & TURNBULL, D. M. 2004. The diagnosis of mitochondrial muscle disease. *Neuromuscul Disord,* 14**,** 237-45.

TAYLOR, R. W. & TURNBULL, D. M. 2005. Mitochondrial DNA mutations in human disease. *Nature reviews. Genetics,* 6**,** 389-402.

TERZIOGLU, M., RUZZENENTE, B., HARMEL, J., MOURIER, A., JEMT, E., LÓPEZ, MARCELA D., KUKAT, C., STEWART, JAMES B., WIBOM, R., MEHARG, C., HABERMANN, B., FALKENBERG, M., GUSTAFSSON, CLAES M., PARK, CHAN B. & LARSSON, N.-G. 2013. MTERF1 Binds mtDNA to Prevent Transcriptional Interference at the Light-Strand Promoter but Is Dispensable for rRNA Gene Transcription Regulation. *Cell Metabolism,* 17**,** 618-626.

THYAGARAJAN, B., PADUA, R. A. & CAMPBELL, C. 1996. Mammalian mitochondria possess homologous DNA recombination activity. *J Biol Chem,* 271**,** 27536-43.

TIRANTI, V., ROCCHI, M., DIDONATO, S. & ZEVIANI, M. 1993. Cloning of human and rat cDNAs encoding the mitochondrial single-stranded DNA-binding protein (SSB). *Gene,* 126**,** 219-25.

TOMECKI, R., DMOCHOWSKA, A., GEWARTOWSKI, K., DZIEMBOWSKI, A. & STEPIEN, P. P. 2004. Identification of a novel human nuclear-encoded mitochondrial poly(A) polymerase. *Nucleic Acids Research,* 32**,** 6001-6014.

TRIFUNOVIC, A., WREDENBERG, A., FALKENBERG, M., SPELBRINK, J. N., ROVIO, A. T., BRUDER, C. E., BOHLOOLY-Y, M., GIDLÖF, S., OLDFORS, A., WIBOM, R., TÖRNELL, J., JACOBS, H. T. & LARSSON, N.-G. 2004. Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature,* 429**,** 417-23.

TROTT, A. & MORANO, K. A. 2004. SYM1 is the stress-induced Saccharomyces cerevisiae ortholog of the mammalian kidney disease gene Mpv17 and is required for ethanol metabolism and tolerance during heat shock. *Eukaryot Cell,* 3**,** 620-31.

TWIG, G., ELORZA, A., MOLINA, A. J., MOHAMED, H., WIKSTROM, J. D., WALZER, G., STILES, L., HAIGH, S. E., KATZ, S., LAS, G., ALROY, J., WU, M., PY, B. F., YUAN, J., DEENEY, J. T., CORKEY, B. E. & SHIRIHAI, O. S. 2008. Fission and selective fusion govern mitochondrial segregation and elimination by autophagy. *Embo j,* 27**,** 433-46.

TYYNISMAA, H., MJOSUND, K. P., WANROOIJ, S., LAPPALAINEN, I., YLIKALLIO, E., JALANKO, A., SPELBRINK, J. N., PAETAU, A. & SUOMALAINEN, A. 2005. Mutant mitochondrial helicase Twinkle causes multiple mtDNA deletions and a late-onset mitochondrial disease in mice. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 17687-17692.

TYYNISMAA, H., SEMBONGI, H., BOKORI-BROWN, M., GRANYCOME, C., ASHLEY, N., POULTON, J., JALANKO, A., SPELBRINK, J. N., HOLT, I. J. & SUOMALAINEN, A. 2004. Twinkle helicase is essential for mtDNA maintenance and regulates mtDNA copy number. *Human Molecular Genetics,* 13**,** 3219-3227.

TYYNISMAA, H., SUN, R., AHOLA-ERKKILA, S., ALMUSA, H., POYHONEN, R., KORPELA, M., HONKANIEMI, J., ISOHANNI, P., PAETAU, A., WANG, L. & SUOMALAINEN, A. 2012. Thymidine kinase 2 mutations in autosomal recessive progressive external ophthalmoplegia with multiple mitochondrial DNA deletions. *Human Molecular Genetics,* 21**,** 66-75.

TYYNISMAA, H., YLIKALLIO, E., PATEL, M., MOLNAR, M. J., HALLER, R. G. & SUOMALAINEN, A. 2009. A heterozygous truncating mutation in RRM2B causes autosomal-dominant progressive external ophthalmoplegia with multiple mtDNA deletions. *American journal of human genetics,* 85**,** 290-5.

UHLER, J. P., THÖRN, C., NICHOLLS, T. J., MATIC, S., MILENKOVIC, D., GUSTAFSSON, C. M. & FALKENBERG, M. 2016. MGME1 processes flaps into ligatable nicks in concert with DNA polymerase γ during mtDNA replication. *Nucleic acids research,* 44**,** 5861-71.

UUSIMAA, J., EVANS, J., SMITH, C., BUTTERWORTH, A., CRAIG, K., ASHLEY, N., LIAO, C., CARVER, J., DIOT, A., MACLEOD, L., HARGREAVES, I., AL-HUSSAINI, A., FAQEIH, E., ASERY, A., AL BALWI, M., EYAID, W., AL-SUNAID, A., KELLY, D., VAN MOURIK, I., BALL, S., JARVIS, J., MULAY, A., HADZIC, N., SAMYN, M., BAKER, A., RAHMAN, S., STEWART, H., MORRIS, A. A. M., SELLER, A., FRATTER, C., TAYLOR, R. W. & POULTON, J. 2014. Clinical, biochemical, cellular and molecular characterization of mitochondrial DNA depletion syndrome due to novel mutations in the MPV17 gene. *European Journal of Human Genetics,* 22**,** 184-191.

VAN GOETHEM, G., DERMAUT, B., LÖFGREN, A., MARTIN, J. J. & VAN BROECKHOVEN, C. 2001. Mutation of POLG is associated with progressive external ophthalmoplegia characterized by mtDNA deletions. *Nature genetics,* 28**,** 211-2.

VISCOMI, C., SPINAZZOLA, A., MAGGIONI, M., FERNANDEZ-VIZARRA, E., MASSA, V., PAGANO, C., VETTOR, R., MORA, M. & ZEVIANI, M. 2009. Early-onset liver mtDNA depletion and late-onset proteinuric nephropathy in Mpv17 knockout mice. *Hum Mol Genet,* 18**,** 12-26.

WALLACE, D. C. 2001. A mitochondrial paradigm for degenerative diseases and ageing. *Novartis Found Symp,* 235**,** 247-63; discussion 263-6.

WANG, X., LIU, X., XUE, L., ZHANG, K., KUO, M.-L., HU, S., ZHOU, B., ANN, D., ZHANG, S. & YEN, Y. 2011. Ribonucleotide Reductase Subunit p53R2 Regulates Mitochondria Homeostasis and Function in KB and PC-3 Cancer Cells. *Biochemical and biophysical research communications,* 410**,** 102-107.

WANROOIJ, S., FUSTÉ, J. M., FARGE, G., SHI, Y., GUSTAFSSON, C. M. & FALKENBERG, M. 2008. Human mitochondrial RNA polymerase primes lagging-strand DNA synthesis in vitro. *Proceedings of the National Academy of Sciences,* 105**,** 11122-11127.

WEBSTER, G., GENSCHEL, J., CURTH, U., URBANKE, C., KANG, C. & HILGENFELD, R. 1997. A common core for binding single-stranded DNA: structural comparison of the single-stranded DNA-binding proteins (SSB) from E. coli and human mitochondria. *FEBS Letters,* 411**,** 313-316.

WILLIAMS, J. S., GEHLE, D. B. & KUNKEL, T. A. 2017. The role of RNase H2 in processing ribonucleotides incorporated during DNA replication.

WILLIAMS, J. S. & KUNKEL, T. A. 2014. Ribonucleotides in DNA: origins, repair and consequences. *DNA repair,* 19**,** 27-37.

WILLIAMS, J. S., LUJAN, S. A. & KUNKEL, T. A. 2016. Processing ribonucleotides incorporated during eukaryotic DNA replication. *Nat Rev Mol Cell Biol,* 17**,** 350-363.

WILLIAMS, JESSICA S., SMITH, DANA J., MARJAVAARA, L., LUJAN, SCOTT A., CHABES, A. & KUNKEL, THOMAS A. 2013. Topoisomerase 1-Mediated Removal of Ribonucleotides from Nascent Leading-Strand DNA. *Molecular Cell,* 49**,** 1010-1015.

YAO, N. Y., SCHROEDER, J. W., YURIEVA, O., SIMMONS, L. A. & O'DONNELL, M. E. 2013. Cost of rNTP/dNTP pool imbalance at the replication fork. *Proceedings of the National Academy of Sciences of the United States of America,* 110**,** 12942-7.

YASUKAWA, T., REYES, A., CLUETT, T. J., YANG, M.-Y., BOWMAKER, M., JACOBS, H. T. & HOLT, I. J. 2006. Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand. *The EMBO journal,* 25**,** 5358-5371.

YASUKAWA, T., YANG, M.-Y., JACOBS, H. T. & HOLT, I. J. 2005. A bidirectional origin of replication maps to the major noncoding region of human mitochondrial DNA. *Molecular Cell,* 18**,** 651-662.

YOULE, R. J. & VAN DER BLIEK, A. M. 2012. Mitochondrial Fission, Fusion, and Stress. *Science,* 337**,** 1062.

YOUNG, P., LEEDS, J. M., SLABAUGH, M. B. & MATHEWS, C. K. 1994. Ribonucleotide Reductase: Evidence for Specific Association with HeLa-Cell Mitochondria. *Biochemical and Biophysical Research Communications,* 203**,** 46-52.

YOZA, B. K. & BOGENHAGEN, D. F. 1984. Identification and in vitro capping of a primary transcript of human mitochondrial DNA. *The Journal of biological chemistry,* 259**,** 3909-15.

YU-WAI-MAN, P., LAI-CHEONG, J., BORTHWICK, G. M., HE, L., TAYLOR, G. A., GREAVES, L. C., TAYLOR, R. W., GRIFFITHS, P. G. & TURNBULL, D. M. 2010. Somatic Mitochondrial DNA Deletions Accumulate to High Levels in Aging Human Extraocular Muscles. *Investigative Ophthalmology & Visual Science,* 51**,** 3347-3353.

ZHANG, H. & POMMIER, Y. 2008. Mitochondrial topoisomerase I sites in the regulatory D-loop region of mitochondrial DNA. *Biochemistry,* 47**,** 11196-203.

ZOLLO, O., TIRANTI, V. & SONDHEIMER, N. 2012. Transcriptional requirements of the distal heavy-strand promoter of mtDNA. *Proceedings of the National Academy of Sciences of the United States of America,* 109**,** 6508-12.