Mechanisms Involved in the Accumulation of Mitochondrial DNA Defects Following Anti-HIV Therapy

Kristian Gardner MSc

Thesis submitted to Newcastle University in candidature for the Degree of Doctor of Philosophy

Wellcome Trust Centre for Mitochondrial Research

Institute of Genetic Medicine

Faculty of Medical Sciences

Newcastle University

November 2014

Abstract

HIV-infected patients present with a range of pathologies that have been associated with mitochondrial toxicity, and more specifically, induced through exposure to the main class of antiretrovirals used in anti-HIV therapy (HAART), nucleoside analogue reverse transcriptase inhibitors, or NRTIs. It has recently been found that such patients have an excess of mitochondrial DNA mutations when exposed to NRTIs; however, the underlying pathophysiological process of this remains undetermined.

To elaborate upon this further, a range of molecular approaches were developed to study the behaviour of mitochondrial DNA mutations, both large scale deletions and single point mutations, in both tissue samples and *in vitro*. High throughput, ultra-deep DNA sequencing technology was also utilised to characterise mitochondrial DNA mutation burden in detail.

Here I present data from physiological samples of HIV-infected individuals receiving NRTIs indicating an increased level of point mutation heteroplasmy and the level of the common mitochondrial deletion, compared to HIV-infected individuals who are NRTI treatment naïve. The mechanisms of this mutation level increase are elucidated through *in vitro* studies indicating that deletions accumulate through a size dependent mechanism during exposure and a point mutation level increase is mediated through a bottleneck mechanism. The possibility of NRTI *de novo* mutagenesis is refuted through the use of next generation sequencing technologies; the data also further supports a bottleneck mechanism of increased point mutation level. I also refute the idea that compounds associated with mitochondrial biogenesis can reduce mitochondrial depletion during NRTI exposure and that there appears to be no genetic predisposition to pathologies in sub-Saharan African individuals.

I therefore conclude that there is an accelerated clonal expansion of pre-existing mitochondrial DNA mutations through NRTI exposure, which is mediated by a size dependent replication bias for deletions, and a molecular bottleneck effect which accelerates drift for point mutations. These data suggest an acceleration of normal cellular ageing through mitochondrial mechanisms in HIV-infected individuals.

Acknowledgements

First and foremost, I owe my sincere gratitude to my supervisors, Professor Patrick Chinnery and Dr Brendan Payne, for providing me with the opportunity to undertake this PhD project. I am deeply indebted to both of them for their constant support, guidance and inspiration over the past three years. It is with great pleasure that I can say I was a part of their group during my PhD studentship.

I would like to thank Dr Angela Pyle for her caring support throughout my PhD, whether it was tolerating my complaining first thing on a morning or aiding me with a laboratory technique, she always gave what little time she had 'spare' to help. I would also like to extend thanks to Drs Gavin Hudson and Vasileios Floros, for their help, advice and thought provoking discussions, albeit if the latter was not always about the world of mitochondria (!). I would also like to thank Dr Jonathan Coxhead for next generation sequencing support and guidance.

I would like to thank all PFC group members who helped me develop the techniques required throughout my PhD, and for making the laboratory a pleasurable place to work. At this point, I would like to give thanks to Dr Jennifer Duff and Rafiqul Hussain, for ensuring the lab operated smoothly and aiding me wherever possible.

I would like to express gratitude to my collaborators (in no particular order); Professor Carlos Moraes, for the 'trans-mitochondrial cybrid' cell line; Dr Gabrielle Saretzki, for the 'aged fibroblast' cell line; Dr Joep van Oosterhout, for the 'd4T Malawian' samples; Professor Jo Poulton, for the 'AZT mouse' samples; and Dr Peter Hall from AstraZeneca, for insights into the pharmaceutical world and co-authoring on my review paper.

My final and most important thanks go out to my mother, father, brother and Claudia, for their constant love, understanding and support throughout my studies.

Author Declaration

This thesis is submitted for the degree of Doctor of Philosophy, in the University of Newcastle upon Tyne, UK. The research detailed within was performed in the Institute of Genetic Medicine under the supervision of Professor Patrick Chinnery and Dr Brendan Payne, between September 2011 and September 2014, and is solely my work, unless otherwise stated and the appropriate acknowledgment provided.

I certify that none of the material offered within has been previously submitted for a degree or any other academic qualification by either me or anyone else. This copy has been supplied in the understanding that it is subject to copyright legislation and no quotation may be made from the thesis without correct acknowledgement.

Kristian Gardner

Table of Contents

List of Figures

List of Tables

List of Publications

1. **Gardner, K.,** Payne, B.A., Horvath, R. and Chinnery, P.F. (2014) 'Use of stereotypical mutational motifs to define resolution limits for the ultra-deep resequencing of mitochondrial DNA', *European J Human Genetics*.

2. **Gardner, K.,** Hall, P.A., Chinnery, P.F. and Payne, B.A. (2013) 'HIV Treatment and Associated Mitochondrial Pathology: Review of 25 Years of in Vitro, Animal, and Human Studies', *Toxicologic Pathology; 42(5):811-822*

3. Van Oosterhout, J.J., **Gardner, K.,** Mallewa, J., Kaunda, S., Kampira, E., Payne, B., Heyderman, R.S. and Chinnery, P. (2013) 'Severe toxicity and polymerase-gamma gene abnormalities in Malawian adults on stavudine-based antiretroviral therapy', *Pharmacogenet Genomics, 23(11), pp. 624-6.*

4. Payne, B.A., **Gardner, K.** and Chinnery, P.F. (2014) 'Mitochondrial DNA mutations in ageing and disease: implications for HIV?' *Antiviral Therapy*.

5. Payne, B.A., **Gardner, K.**, Coxhead, J., and Chinnery, PF. (2014) 'Deep re-sequencing of mitochondrial DNA'. *MiMB Chapter; (In Press).*

List of Abbreviations

3TC	Lamivudine
ABC	Abacavir
AICAR	5-Aminoimidazole-4-carboxamide ribonucleotide
AIDS	Acquired immunodeficiency syndrome
AMP	Adenosine monophosphate
AMPK	Adenosine monophosphate kinase
ATP	Adenosine triphosphate
AZT	Zidovudine
BAM	Binary alignment file
BWA	Burrows wheeler alignment
CCR5	C-C chemokine receptor type 5
CD4	Cluster of differentiation 4
COX	Cytochrome c oxidase
COX/SDH	Cytochrome c oxidase/succinate dehydrogenase
Ct	Cycle threshold
CXCR4	C-X-C chemokine receptor type 4
d4T	Stavudine
ddC	Di-deoxycytidine
ddI	Di-deoxyinosine
D-Loop	Mitochondrial displacement loop
DMEM	Dulbecco's Modified Eagle's medium
DMSO	Dimethyl Sulfoxide
dNTPs	Dinucleotide triphosphates

ERRs	Oestrogen related receptors
FAO	Fatty acid oxidation
FCS	Foetal calf serum
gp120	Glycoprotein 120
H&E	Haematoxylin-eosin
HAART	Highly active antiretroviral therapy
HBSS	Hanks balanced salt solution
HEK 293	Human Embryonic Kidney 293
HIV	Human immunodeficiency virus
kDa	Kilo-Daltons
K-S	Kolmogorov Simonov
KSS	Kearns Sayers syndrome
LB	Lysogeny broth
LD	Lipodystrophy
LHON	Leber Hereditary optic neuropathy
LMD	Laser microdissection
MAF	Minimum allele frequency
MELAS	Mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes
MEM	Minimum exact matches
mtDNA	Mitochondrial DNA
MT-ND1	Mitochondrial NADH dehydrogenase 1
MT-ND4	Mitochondrial NADH dehydrogenase 4
mtSSBP	Mitochondrial single-stranded DNA-binding protein
nDNA	Nuclear DNA

NGS	Next generation sequencing
NNRTI	Non-nucleoside analog reverse transcriptase inhibitor
NRF 1/2	Nuclear respiratory factors 1/2
NRTI	Nucleoside analog reverse transcriptase inhibitor
OH	Origin of heavy strand replication
OL	Origin of light strand replication
OXPHOS	Oxidative phosphorylation
PBMCs	Peripheral blood mononuclear cells
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEN	Polyethylene naphthalate
PGC-1α	Peroxisome proliferator activated receptor γ co-activator-1α
PI	Protease inhibitor
PN	Peripheral neuropathy
POLG	Polymerase γ encoding gene
PPAR	Peroxisome proliferator-activated receptor
qPCR	Quantitative polymerase chain reaction
rCRS	Revised Cambridge reference sequence
ROS	Reactive oxygen species
RRF	Ragged red fibres
rRNA	Ribosomal RNA
RSRS	Reconstructed sapien reference sequence
RT	Reverse transcriptase
S.D	Standard deviation
SAM	Sequence alignment mapping

SDH	Succinate dehydrogenase
SIRT1	Sirtuin 1
SNP	Single nucleotide polymorphism
SQ	Starting quantity
TAE	Tris-acetate-EDTA
TDF	Tenofovir disoproxil fumarate
TFAM	Mitochondrial transcription factor A
TNF	Tumour necrosis factor
tRNA	Transfer RNA
v/v	Volume per volume
VCF	Variant calling format
w/v	Weight per volume
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
β2M	Beta-2-microglobulin
Δ	Change/variation

Chapter 1. Introduction

Table of Contents

1.1 Mitochondria

1.1.1 Origin

It is believed that mitochondria arose from an alpha-proteobacterium enclosed into a eukaryotic progenitor (Lane and Martin, 2010). This theory has come to be known as the endosymbiotic theory and was proposed in 1970's by *Dr Lynn Margulis*' observations that mitochondrial components were highly similar to those found in bacteria (Margulis, 1975). Whilst mammalian mitochondria have retained only a small amount of bacterial features, it is estimated that the 'current' mitochondria are merely evolutionary derivatives of the original endosymbiont, after encapsulation by the eukaryote, which accounts for distinct differences between the two (Gabaldon and Huynen, 2004).

1.1.2 Structure

The similarity retained between mitochondria and their bacterial ancestors, is evidenced by each mitochondrion comprising of an outer and inner structural membrane, an intermembrane space/matrix and its own 'circular' DNA, mtDNA (mitochondrial DNA), which is distinct from nDNA (nuclear DNA).

The outer mitochondrial membrane is a phospholipid bilayer, comprising channel forming proteins known as porins, that form mitochondrial conduits to allow the transportation, or non-specific diffusion, of small molecules (<10kDa). Due to the increased level of enzymatic activity and protein transportation, the proportion of proteins present in the outer membrane compared to the inner membrane is approximately four times as great (in relation to the ratio of proteins: lipids).

The inner membrane, lacks many of the proteins present in the outer membrane, namely the porin family; however, it is abundant in cardiolipins, which are required for making the inner membrane impermeable to solutes and ions (Patil and Greenberg, 2013). Therefore, molecules present in the intermembrane space need to be actively transported out by a specific system. This system of transportation is driven by membrane electrical potential (Frazier *et al.*, 2003).

The mtDNA in humans, has reduced in size to 16,569bp in length, proposed to be a result of the millions of years of evolution, to become almost exclusively exonic in nature, with only one region known to be non-coding (see Figure 1.1).

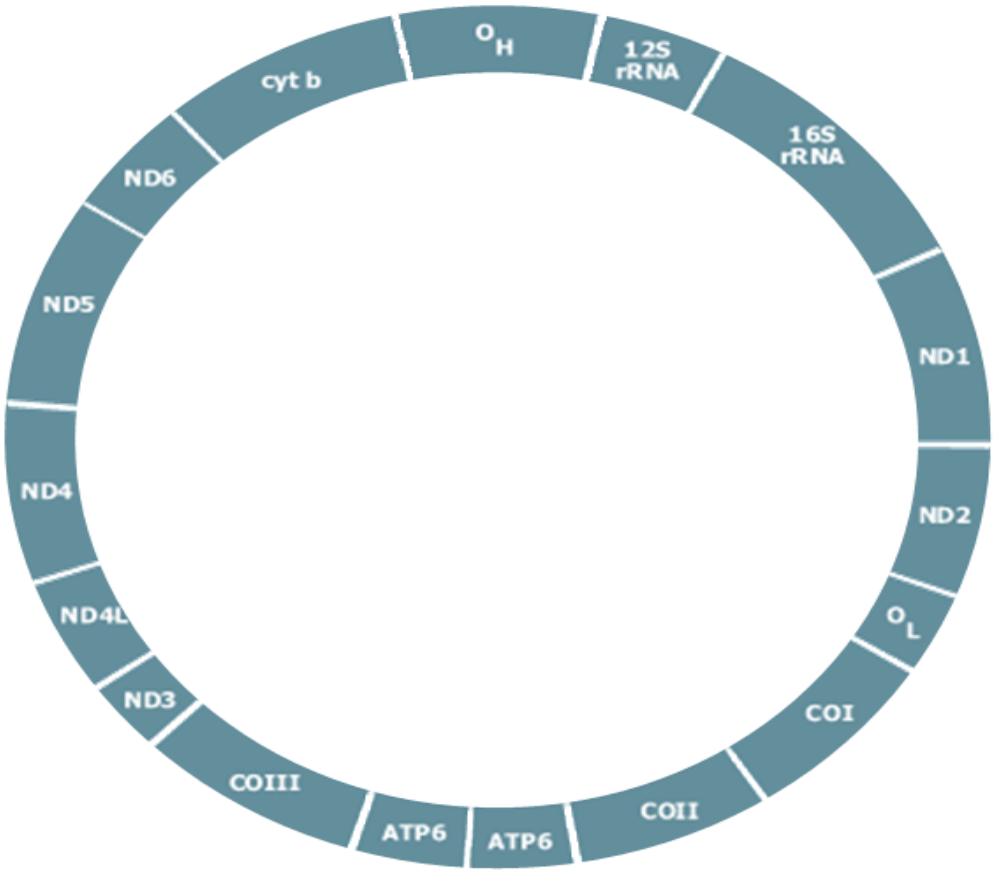

Figure 1.1 Structure of mitochondrial DNA (mtDNA) highlighting the highly exonic nature of the species and the individual genes which are encoded. The d-loop (displacement loop) is the only non-coding region with the origin of replication of the 'heavy' strand (O_H) of the mtDNA. Image adapted from https://www.nfstc.org.

Human mtDNA is comprised of 37 genes, of which 22 encode mitochondrial transfer RNAs (tRNAs), 2 genes encode mitochondrial ribosomal RNAs (rRNAs) and 13 encode polypeptides of the electron transport chain. However, more than 80 subunits of the electron transport chain are encoded by the nuclear DNA (nDNA). Typically, each cell will contain hundreds to thousands of copies of the mitochondrial genome (White, 2001).

1.1.3 Respiratory chain

Mitochondria are responsible for encoding the functional subunits of the respiratory chain. The respiratory chain is located within the inner membrane of mitochondria and is made up of four protein complexes (NADH, succinate dehydrogenase, ubiquinol cytochrome c oxidoreducatase, and cytochrome c oxidase) and ATP synthase regarded as the fifth complex. The respiratory chain also requires two electron carriers to act as intermediate units, ubiquinone (also known as coenzyme Q10) and cytochrome c.

During the process of respiration and the drive of the respiratory chain, there is a generation of superoxide of approximately 1-2% of all the electrons transported along the respiratory chain. The superoxide generation has been strongly linked to the processes of complex I (Hirst *et al.*, 2008). This is due to such a low redox potential required for one electron reduction of di-oxygen to superoxide. Although the matrix of the mitochondrion has limited mechanisms to protect both DNA and structural proteins, against oxidative damage, such mechanisms can only work at a finite rate and therefore, a certain level of damage is inevitable. However, oxidative damage can start to disrupt the respiratory chain, especially in older age when mitochondrial activity decreases, at any of the complexes (I-IV) and the build-up of ROS (reactive oxygen species) is felt to be a major contributor to mitochondrial damage in general.

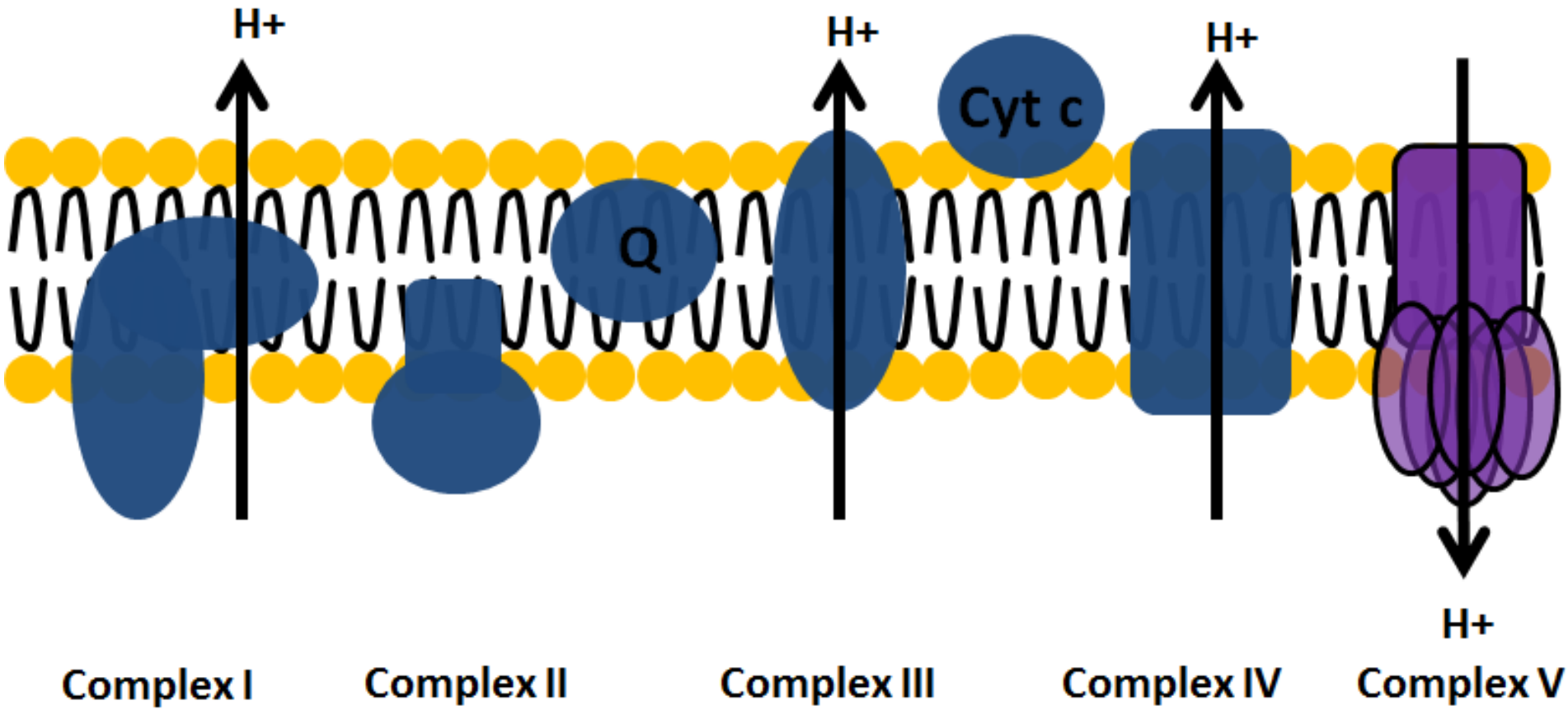

Figure 1.2 An illustrative schematic of the respiratory system found within the inner membrane of the mitochondria comprised of: complex I, NADH; complex II, succinate dehydrogenase; Q, coenzyme Q10; complex III, Ubiquinol cytochrome c oxidoreducatase; cyt c, cytochrome c; complex IV, cytochrome c oxidase and complex V, ATP synthase. H+ indicates hydrogen with arrows indicating flow of the ion(s).

1.2 Mitochondrial Genetics

1.2.1 Mitochondrial DNA inheritance

The standard model that describes mtDNA inheritance is that it is transmitted strictly though the maternal line (Giles *et al.*, 1980), although there are claims of 1-2% of paternal inheritance/leakage present in mice (Gyllensten *et al.*, 1991) with controversial claims that it may be present in humans also (Schwartz and Vissing, 2002).

1.2.2 Mitochondrial DNA nucleoid structure

Mitochondrial DNA was originally thought to be naked with histones acting as a protective coating, in a similar manner to the nDNA. However, it is believed that approximately 5-7 mtDNA molecules are compacted into nucleoprotein complexes for protection (nucleoids) and are estimated to be 70nm in size (Nass, 1969; Iborra *et al.*, 2004).

The protein structure of nucleoids has been well studied and is strongly implicated to be comprised of TFAM (mitochondrial transcription factor A) and mitochondrial single-stranded DNA-binding protein (mtSSBP) along with a range of other mitochondrial metabolic proteins both well and not so well studied, such as: mtDNA polymerase (pol γ), T7-like helicase, mtRNA polymerase, suv3-like helicase and DEAD box protein 28, as described by *Wang and Bogenhagen* (Y. Wang and Bogenhagen, 2006).

However, it has recently been suggested that each nucleoid consists of a single mtDNA molecule and is mainly comprised of TFAM, suggesting that TFAM is the main factor in packaging and organising mtDNA into nucleoids. This observation may potentially have further implications for our understanding of the mechanisms of segregation and transmission of mtDNA disease (later discussed in Sections 1.2.10 and 1.3) (Kukat *et al.*, 2011).

1.2.3 Replication and maintenance

Mitochondrial DNA is replicated by the enzyme polymerase gamma (pol γ). Pol γ is encoded by the nuclear genes *POLG1* and *POLG2*, the former encodes the catalytic subunit of pol γ and the latter is responsible for the coding of a 55kDa subunit; they have chromosome loci of 15q25 and 17q21, respectively.

Pol γ is thought to be solely responsible for all of the mtDNA synthetic reactions as well as the repair of the DNA. However, its complex role in fully maintaining the DNA is

not fully understood. Pol γ is unique among the cellular replicative DNA polymerases because it is potentially highly susceptible to inhibition by particular compounds such as ethidium bromide. It is well accepted that mtDNA replication does not coincide with cellular replication and is known to occur completely independently of nuclear replication (Bogenhagen and Clayton, 1977).

The mtDNA helicase, twinkle, interacts with a 13-15kDa protein, mitochondrial single-stranded DNA-binding protein (mtSSBP) to achieve a destabilization of the mtDNA helix during replication. Twinkle is further stimulated by mtSSBP, which results in an increased fidelity of pol γ and increased processivity (Kaguni, 2004).

The essential component of the mitochondrial nucleoid structure, TFAM, is necessary for the initiation of transcription and replication of mtDNA. It has also been suggested that TFAM is not only necessary for providing structure, but also regulates protein binding at the *cis*-regulatory displacement loop (D-loop; a non-coding region) of the mtDNA (Ghivizzani *et al.*, 1994). The level of TFAM has been found to be directly proportional to the amount of mtDNA present within a sample, further providing support for its regulatory mechanism (Ekstrand *et al.*, 2004).

Currently there are two models that explain the mechanism behind the replication of mtDNA; the asynchronous displacements mechanism, and the leading-lagging strand synthesis mechanism (Holt *et al.*, 2000; Yang *et al.*, 2002; Clayton, 2003). According to the former mechanism, the heavy strand (O_H) is synthesized first until the light strand origin is reached (O_L), then the light stranding is replicated. The latter mechanism suggests that there's a coupled unidirectional synthesis of the mtDNA strands. Neither theories have been proven to significantly out way the likelihood of the other, with a suggestion that both mechanisms may occur in a cell (Clay Montier *et al.*, 2009).

1.2.4 Mitochondrial translation

There's an incomplete understanding and an underdeveloped model for studying the mitochondrial translation system. All of the mitochondrial ribosomal proteins are nuclear encoded, transcribed and translated. They're are then transported to the mitochondria where they form complexes with the mitochondrial encoded rRNAs 12s and 16s (O'Brien, 2003).

13 polypeptide proteins that make up the respiratory chain are synthesised within the mammalian mitochondria. Translation is initiated by mitochondrial initiation factor 2

($IF2_{mt}$) and mitochondrial initiation factor 3 ($IF3_{mt}$). Together, these factors promote the binding of fMet-tRNA to the small ribosomal subunit, 28s, and the dissociation of the large ribosome 55s. This consequently leads to the release of 39s subunit and the creation of the 28s-$IF3_{mt}$ complex (Christian and Spremulli, 2009). The elongation phase of the translation then occurs, until termination through one of the four mitochondrial stop codons (UAA, UAG, AGA or AGG) is encountered. Following successful termination of translation and the hydrolysis of the new protein, the ribosome is recycled through one of two recycling factors: RRFmt, or EF-G2. The recycling process has been found by the two groups to be essential for normal cell viability (Rorbach *et al.*, 2008; Tsuboi *et al.*, 2009).

1.2.5 Mitochondrial replication regulation

The replication of mtDNA is a multi-faceted system, implicated in mitochondrial biogenesis. Peroxisome proliferator activated receptor γ co-activator-1α (PGC-1α) is known to play a key role in the process of mitochondrial biogenesis (Jornayvaz and Shulman, 2010). Key interactions that PGC-1α has which contribute towards biogenesis include NRF1/2 (nuclear respiratory factors), PPAR α, β, δ and γ (peroxisome proliferator-activated receptor), TFAM, thyroid hormone, oestrogen related receptors and glucocorticoids (Jornayvaz and Shulman, 2010).

A major regulator of mitochondrial biogenesis is through activation of NRF1/2 is due to their ability to increase the expression of TFAM. TFAM is known to drive transcription and replication of mtDNA (as previously discussed 1.2.3). NRF1/2 has also been found to increase the up-regulation of other mitochondrial enzymes required for mtDNA transcription (Virbasius and Scarpulla, 1994). PGC-1α activates the transcriptional function of NRF-1 on the promoter of TFAM, as well as promoting NRF-2 binding to the COX IV promoter as the mechanisms to induce biogenesis (Wu *et al.*, 1999; Baar *et al.*, 2002).

In addition to NRF1/2, PGC-1α interacts with and co-activates the PPARs. The PPAR family plays an extensive role in the expression of proteins involved in the intra- and extra-mitochondrial fatty acid oxidation and transportation (FAO) system. PPAR-α and δ enhance fatty acid transporters; PPAR-β has been strongly suggested to play a role in glucose metabolism, along with PPAR-γ (Kota *et al.*, 2005).

PGC-1α activates the other factors such as thyroid hormone, glucocorticoids and oestrogen related receptors (ERRs). EERα and γ target a common set of promoters that regulate proteins involved in the uptake of substances and ATP generation (Dufour *et al.*, 2007), as well as mitochondrial ribosomal machinery (Giguere, 2008). It has been found that by down regulating or knocking out the expression of the EERs, mitochondrial content also decreases, further supporting the important role of EERs in mitochondrial biogenesis and content control (Schreiber *et al.*, 2004).

1.2.6 Mitochondrial dynamics

Mitochondria are subject to constant morphological changes through the combined actions of fusion, fission and movement along cytoskeletal tracks (Alexander M. van der Bliek *et al.*, 2013). The key proteins involved in mitochondrial dynamics are MFN1, MFN2, OPA1, FIS1, MTP18 and DRP1, with the first three mediating mitochondrial fusion and the latter three mediating mitochondrial fission (Mishra and Chan, 2014).

A mutation or repression in the genes, OPA1, MFN1 or MFN2 can lead to mitochondrial fragmentation or a reduction in the amount of mitochondrial filaments (Bach *et al.*, 2003; Hsiuchen Chen *et al.*, 2003; Arnoult *et al.*, 2005). This can have detrimental effects on mitochondrial ATP production due to the link between mitochondrial fusion, especially MFN2, and high energy demanding cells and the loss of OPA1 resulting in a release of cytochrome c and increased apoptosis (Bach *et al.*, 2003; Olichon *et al.*, 2003). Conversely, an increased expression or gain in function of these genes, has been found to result in increased mitochondrial filament or network length (Santel and Fuller, 2001).

A mutation or repression in the genes, FIS1, MTP18 and DRP1, can result in an elongation of the mitochondrial with increased network size and a reduction in the amount of mitophagy (AM van der Bliek *et al.*, 1993; Yoon *et al.*, 2003; Stojanovski *et al.*, 2004). An increased expression of these genes or an overexpression of proteins may result in mitochondrial network fragmentation (Yoon *et al.*, 2003).

Mitochondrial dynamics is a complex system that is also responsible for maintenance of mitochondrial function through a separation of dysfunctional mitochondria from the network, which are targeted for degradation through mitophagy (Ashrafi and Schwarz, 2013).

1.2.7 Mitochondrial copy number regulation

The process of mtDNA degradation and synthesis is crucial in the maintenance and homeostasis of normal cellular function. The turnover process itself, does not have a well-defined mechanism for regulation of when to proliferate or degrade mtDNA. Although it may appear somewhat stochastic, the most widely accepted model of mtDNA copy number regulation is the copy number threshold hypothesis (Clay Montier *et al.*, 2009).

In the copy number threshold model, when mtDNA copy number decreases and reaches a defined lower threshold, unknown factors trigger the up-regulation of mtDNA replication and therefore, restore the copy number back to or within the 'normal' level. Conversely, when the copy number reaches the 'high' threshold, unknown factors come into play to reduce the amount of turnover of mtDNA and trigger mtDNA degradation, to bring the number back down (Clay Montier *et al.*, 2009).

Although the above mechanism appears somewhat vague, other hypotheses have been proposed but have less supporting evidence to reinforce their existence. The most recent study highlights that copy number, specifically in PBMCs (peripheral blood mononuclear cells), is regulated by TFAM expression and is determined by the level of oxidative protection that is required; When there is a build-up of oxidative stress through reactive oxidative species (ROS), copy number is up-regulated, and vice versa for the reverse process. (Chakrabarty *et al.*, 2014).

1.2.8 Mitochondrial DNA mutations and heteroplasmy

The first mtDNA point mutation that was discovered was the m.11778G>A mutation associated with Leber's Hereditary Optic Neuropathy (LHON) (DC Wallace *et al.*, 1988). This was then followed by the first mtDNA deletion, which was associated with myopathy (Holt *et al.*, 1988).

The replication of the mitochondrial genome can result in mutations in the form of single point mutations and small- or large-scale deletions. Some of these mutations have naturally been found to be deleterious and cause disease. mtDNA deletions are thought to be caused by one of four mechanisms; double strand breaks; slippage of pol γ; insufficient repair and protection mechanisms (twinkle/histones) and homologous recombination (Meissner *et al.*, 2008). MtDNA mutations induced through oxidative stress is the result of ROS; a by-product of the electron transport chain. Of the ROS, the

reactive hydroxyl group reacts with mtDNA by adding double bonds in DNA bases. This is performed by the removal of a hydrogen atom from the methyl group and results in the modification of the base position, expressed as a mutation (Cooke *et al.*, 2003).

Due to the polyploid nature of mitochondrial DNA and a large number of mtDNA copies present in a cell (up to thousands per cell), if one copy develops a mutation, this generates two populations, a mutant and a wild type. This state is known as heteroplasmy and the level of mutation expressed as a percentage.

If a deletion mutation causes the removal of parts of the genes that encode for the OXPHOS system, such as the 4977bp 'common deletion' that resides in the 'major arc' region where deletions are commonly found, this can result in respiratory dysfunction. The mitochondrial common deletion resides between two 13-bp direct repeat potions at m. 13447-13459 and m.8470-8482. It has gained great interest as it is the cause of several sporadic diseases, including Pearson's disease and Kearns-Sayer syndrome. It therefore has become known as the 'common' deletion and used as a tissue biomarker of ageing (Taylor and Turnbull, 2005).

Point mutations can also cause significant functional effects, for example, m.3243A>G, can cause a severe mitochondrial disease known as MELAS (mitochondrial encephalomyopathy, lactic acidosis, stroke-like episodes) (Taylor and Turnbull, 2005; Meissner *et al.*, 2008). Previous studies have found that there are distinct different types of mutations in different tissues that have been associated with ageing. In post mitotic tissues the predominant mutated species are large scale mtDNA deletions (Bender *et al.*, 2006), whereas in mitotic cells, mtDNA point mutations have been commonly identified, with the d-loop associated as a potential mutation hot-spot (Shin *et al.*, 2004; De Alwis *et al.*, 2009; Laura C. Greaves *et al.*, 2010; Kassem *et al.*, 2011).

If the level of a mutation that causes a deleterious effect exceeds a threshold, it can cause defects that can compromise a whole organ, if the cells affected are part of a network, such as those present in the central nervous system. The process of accumulating mitochondrial DNA mutations to the high threshold level, usually deemed 65-70% heteroplasmy (Durham *et al.*, 2005), is known as clonal expansion and the general principle is shown in Figure 1.3.

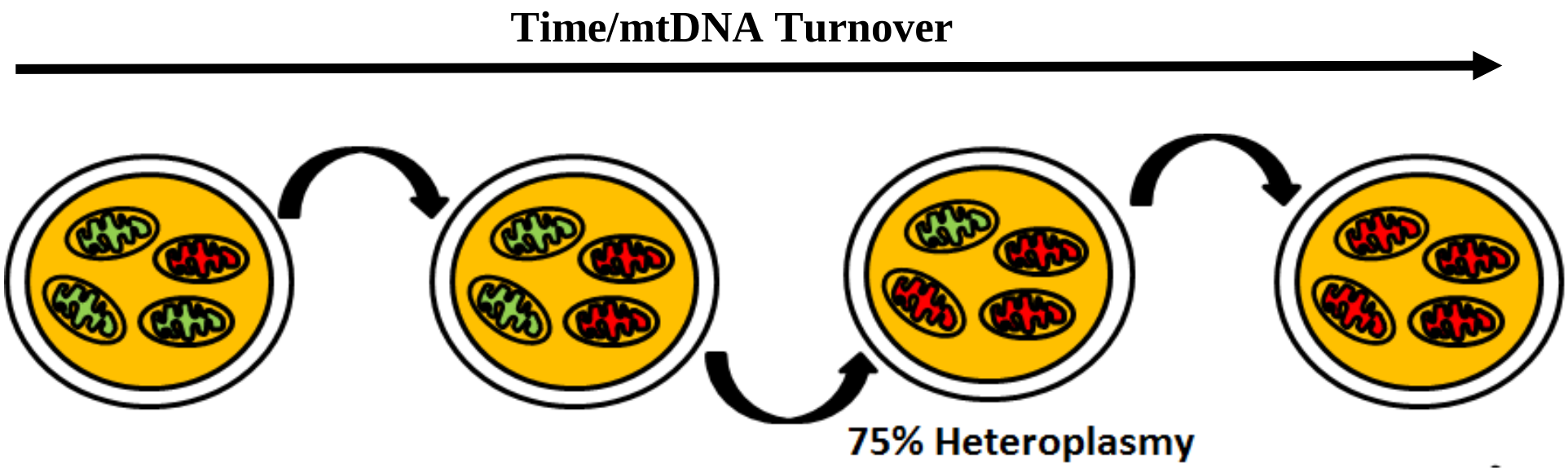

Figure 1.3 A schematic illustrating a basic representation of the process of clonal expansion of an mtDNA mutation through mtDNA turnover, exceeding the threshold level which result in dysfunction, at approximately 75% heteroplasmy. Green indicates normal mitochondria and red indicates mutated. Mitochondria.

1.2.9 Clonal expansion

The mitochondrial genome accumulates acquired (somatic) mutations throughout life, some expanding to high levels, and others not. Unlike the nuclear DNA, the mtDNA has an independent replication to the cell cycle, this is known as relaxed replication (see Section 1.2.2) (Elson *et al.*, 2001).

It was originally thought that mutations may accumulate within single cells due to a replicative advantage on the theory that smaller deleted or 'mutant' molecules will replicate quicker than a whole, wild type mtDNA molecule (in the case of deletions). However, it should be noted that this hypothesis, is not widely accepted, due to the fact that replication of the mitochondrial genome is not subject to a rate-limiting factor and the 'faster replication' of deleted molecules should in theory, not influence the ratio between deleted and wild type molecules over time, i.e. the half-life greatly exceeds replication time of mtDNA. Furthermore, a size dependant mechanism does not explain the increase of heteroplasmy of point mutations, as seen within patients with inherited mitochondrial disease (Elson *et al.*, 2001).

Due to the dynamic nature of mitochondria; frequently undergoing rounds of replication alongside fusion and fission, it raises the possibility that clonal expansion may be caused by random genetic drift. This has been previously simulated using a relaxed replication model of mitochondrial genome replication (Elson *et al.*, 2001). This study showed that relaxed replication can cause a mutation to clonally expand and reach high levels within an individual post-mitotic cell without any 'replicative advantage'. It also showed predictions of the levels of COX-negative cells (see Section 1.3.3) in an aged individual that are in concordance with those observed using histochemical techniques.

1.2.10 Transmission of mtDNA mutations

The inheritance or transmission of mtDNA mutations is complex. All homoplasmic mutations (i.e. 100% mutant) are transmitted to all offspring through the maternal line. The transmission of a homoplasmic mutation does not directly relate to penetrance of disease in the offspring. For instance, the transmission of homoplasmic mutations associated with Leber hereditary optic neuropathy (LHON), has been found to cause disease in 50% of males but only 10% of females (Man *et al.*, 2003).

The transmission of heteroplasmic mtDNA mutations, especially point mutations, is more complex than that of homoplasmic. The regulator of transmission is hypothesised

to be one of two mechanisms: the mitochondrial bottleneck and mitochondrial segregation. There are three proposed explanations or mechanisms for how it regulates the segregation of mtDNA heteroplasmic mutations. The first mechanism describes the process as a preferential amplification of clustered populations containing identical/highly similar sequences (L. Cao *et al.*, 2009). The second proposed theory, albeit the most recent, claims that there's an amplification of a sub population of mitochondrial genomes in folliculogenesis (Wai *et al.*, 2008). The final mechanistic description postulates that a decline in mtDNA copy number creates a rapid segregation, and allows for only a sub-population of mitochondrial genomes to repopulate during embryogenesis (Cree *et al.*, 2008).

In all of the above models, the result is that a subpopulation of maternal mtDNA generates the offspring mtDNA, and hence the heteroplasmy levels of the offspring mtDNA may be significantly skewed from the maternal.

1.2.11 Mitochondrial haplogroups

There is known to be substantial homoplasmic mtDNA sequence variations between individuals across the global population. This variation arose through ancient mtDNA polymorphisms accumulating along radiating maternal lineages as people migrated and populated the world. When people populated and colonised an area, the polymorphisms became enriched and generated a distinct 'tree' with geographical region specific mtDNA sequences (D. C. Wallace and Chalkia, 2013). The mtDNA sequence specific tree makes up what is known as mitochondrial haplogroups, and is made up of macro-lineages or macro-haplogroups, totalling 18 distinct lineages, 9 of which are localised to European individuals (Torroni and Wallace, 1994).

However, there is emerging evidences that indicate there is also transmission of low level heteroplasmies throughout sub-populations. A recent study highlighted transmission of very low level mtDNA variant heteroplasmies in centenarians to their offspring, indicating lineage inheritance of low level variants as well as the homoplasmic changes (Giuliani *et al.*, 2014). The presence of low-level variants gives rise to a range of implications, especially for the understanding of mitochondrial disease and ageing associated mutations. Such implications include the clonal expansions throughout life contributing to ageing and the inheritance of low level pathogenic mutations contributing towards disease.

1.3 Mitochondrial Disorders

1.3.1 An overview of mitochondrial disorders

Mitochondrial disorders are phenotypically and genotypically heterogeneous and arise as a result of the dysfunction of the mitochondrial respiratory pathway. In approximately 15% of cases, disorders arise due to mutations in the mtDNA (Dimauro and Davidzon, 2005), with the rest induced through nuclear genes encoding for mitochondrial proteins.

1.3.2 Mitochondrial pathologies

Mitochondrial disorders can range from affecting a single organ, such as the eye in Leber Hereditary optic neuropathy (LHON) (Yu-Wai-Man *et al.*, 2009) to multisystem diseases such as those seen in mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes (MELAS) (Pavlakis *et al.*, 1984). Mutations in the mitochondrial polymerase γ gene (*POLG*) result in pleiotropy: some mutations would cause ocular weakness, whereas others would cause ataxia and muscle weakness.

1.3.3 Diagnosing mitochondrial pathology

There are a range of diagnostic tools used routinely in clinical and research practices to detect and define mitochondrial pathology. The techniques include histopathological staining and more recently, the use of molecular techniques to define pathology in both tissue specimens and DNA extracts.

Basic haematoxylin-eosin (H&E) stains do not really give any clues to the presence of mitochondrial damage. One of the stereotypical features of mitochondrial myopathy is the presence of ragged red fibres (RRF). RRF are characteristic of mitochondrial damage in muscle tissue and the accumulation of diseased mitochondria is the underling explanation for the RRF phenomenon. RRF are usually detected by the use of a modified Gomori trichrome stain and visualised as red patches around the edges of a cell on a muscle section. Frequently the implementation of COX/SDH (cytochrome c oxidase/succinate dehydrogenase) staining is also used to define mitochondrial defects by detecting a biochemical defect displayed through cellular dysfunction.

COX contributes to accepting electrons from cytochrome c and generating a mitochondrial membrane electrochemical potential. Such activity is histochemically detected through the use of an electron donor as a substitute of cytochrome c. When oxidation occurs, there is a formation of a brown pigment, and thus when a brown stain

is present on the cell of the muscle Section, it indicates COX activity is present. COX is encoded by both nuclear and mitochondrial DNA. If a mitochondrial DNA defect is present within a cell then COX activity in that cell will be lost, assuming there's a sufficient heteroplasmy level to cause a defect. However, this does not conclude whether it is mitochondrial, nuclear or both; therefore, staining for SDH (mitochondrial complex II), which is encoded exclusively by the nDNA, will elucidate upon this issue as a blue counter stain will highlight the cells which contain mtDNA defects. SDH catalyses the oxidation of succinate into fumaric acid, a histochemical demonstration of this enzyme is performed by utilising a tetrazolium compound, which in the presence of active SDH, will be reduced to a formazan dye and produce a blue colouration (Gardner *et al.*, 2013). This technique has been found useful due to claims that over 89% of patients expressing mitochondrial DNA mutations (specifically deletions) will have abnormal COX/SDH staining (Lamont *et al.*, 1998).

Microscopy, both light and electron/ultrastructure microscopy are important tools in the evaluation of mitochondrial pathology. Thirty to 44% of individuals with mitochondrial pathology will demonstrate mitochondrial morphological changes detectable with the use of electron microscopy; however, the technique is not widely available in all diagnostic settings (Kyriacou *et al.*, 1999).

More recently, advancements in technology has enabled more specific analyses and determination of the severity of mitochondrial pathology to be defined. The use of biochemical analysis can be used to determine the exact subunit(s) at fault in a respiratory disorder. The molecular analyses of mitochondrial copy number and major arc deletion quantification (by qPCR and LR-PCR if they're rearrangements) can be used to detect mitochondrial depletion and deletion disorders. Finally the ability to sequence the entire mitochondrial genome can give clear indications to specific point mutations to entire genome rearrangements. If there is a deletion mutation detected, sequencing can also be performed to determine genes affected (mitochondrial or nuclear). All of these can be used to more precisely diagnose mitochondrial pathology and any associated disorder, which gives a valid indication of the most appropriate treatment procedure (Rotig *et al.*, 2004; Phillips *et al.*, 2014).

1.3.4 Treatments for mitochondrial disorders

There is no definitive means of halting mitochondrial disorder progression and the ability to clinically define therapeutic responsiveness and optimal drug dosing is also

rather limited. However, there is a range of treatment options that have been suggested to be beneficial in the treatment of mitochondrial disorders.

The basic clinical goal in the treatment of mitochondrial disorders consists of increasing energy and promoting mitochondrial biogenesis. Increasing energy is usually induced through increasing the level of ATP production and reducing the level of free radical production (Avula *et al.*, 2014). Supplements such as co-enzyme Q10, L-carnitine and riboflavin are among the few used to increase ATP production through aid to the respiratory chain (Parikh *et al.*, 2009).

Co-enzyme Q10 was initially debated as being an ineffective method for treating or aiding with mitochondrial disorders (Matthews *et al.*, 1993). However some of the most recent studies suggest that regular oral supplementation can be beneficial as an anti-oxidant aid, providing symptomatic relief (Garrido-Maraver *et al.*, 2014). Co-supplementation with riboflavin has recently been suggested as more beneficial with positive benefits as prophylactic treatment in mitochondrial driven disorders, including unusual mitochondrial pathologies such as migraines (Markley, 2012). L-carnitine is more tissue specific with the most recent support as a beneficial compound in muscular disorders (D'Antona *et al.*, 2014).

A set of 'emerging therapies' are currently being tested to assess whether or not they have ameliorative effects in patients with mitochondrial disorders. These therapies are compounds known to have an involvement in activating intermediates of the pathway of mitochondrial biogenesis (for pathways, see Section 1.2.5). Sirtuins are among the most popular; however, there are no clinical trials as of yet to suggest such drugs are valid in the treatment of mitochondrial disorders. There are a group of naturally occurring 'supplements' such as resveratrol and AICAR (5-Aminoimidazole-4-carboxamide ribonucleotide) that act as up-regulators of co-factors of mitochondrial biogenesis and have been implicated as having a valid use in mitochondrial disorder treatment, although no conclusive assessment of their function has truly been determined in long term *in vivo* exposure (Parikh *et al.*, 2009).

1.4 Mitochondrial Involvement in Ageing

The connection between mitochondria and ageing is known as the mitochondrial theory of ageing. This theory incorporates the production of ROS, a by-product of the OXPHOS system, as a mediator of mtDNA mutations in the genome. This damage then leads to potential respiratory chain uncoupling and further ROS production and ultimately, a viscous cycle that increases with age.

Effects of mitochondrial ageing is most evident in a tissue that is non replicative (post-mitotic), such tissues include brain and skeletal muscle. The role of mitochondrial abnormalities in the aetiology of ageing has been characterised and comprises of a variety of features. One of these features is the decline in activity of the partially mitochondrial-encoded complex IV (COX). Somatic mtDNA defects accumulate during normal ageing and lead to cellular defects of COX, which is an essential component of the respiratory chain and therefore, a critical component of ATP production (Bua *et al.*, 2006). These mutations collect in individual cells, ultimately causing a cellular biochemical defect.

1.4.1 Age-associated disease and mitochondrial mutations

There has been a wide range of studies reporting that there are links, both directly and indirectly, between mitochondrial DNA mutations and age associated diseases such as: neurodegenerative disease, diabetes and cancer. Research suggests that as age increases, mtDNA mutations accumulate within individual cells and tissues.

Mitochondrial DNA mutations that are known to cause mitochondrial disease, have been shown to also accumulate with age (Corral-Debrinski *et al.*, 1992). The first association of mitochondrial deletions and age was first found when studying the common deletion (m.8470 - m.13447) in tissue. This common deletion was found to be present in a number of tissues (skeletal muscle, brain, liver) in aged individuals (Brierley *et al.*, 1998). Point mutations have also been shown to accumulate with age and specifically, the m.3243A>G mutation causing the mitochondrial disease MELAS. This disease and the associated mutation have only been found in adult patients (Michikawa *et al.*, 1999).

Arguably, one of the best described examples of age associated mtDNA point mutations, are those found in human colonic crypts. By the age of 80, approximately 15% of all crypts contain a respiratory defect. This defect has been shown to be caused by the

clonal expansion of point mutations (Taylor *et al.*, 2003). In general, mtDNA mutations accumulate in a range of tissues with age and are most commonly found in the colon, muscle and brain.

Finally, the accumulation of mtDNA mutations has also been strongly implicated in neurodegenerative disorders, such as with Tau pathology and specifically, Alzheimer's disease (Corral-Debrinski *et al.*, 1994; Hirano *et al.*, 1997; Bonilla *et al.*, 1999; de la Monte *et al.*, 2000; Coskun *et al.*, 2004; Swerdlow and Khan, 2004; J. Wang *et al.*, 2005). The accumulation of mtDNA mutations through age in neurons is debated to being either an ageing phenomenon, or a disease specific phenomenon. Normal ageing brains present with increased mtDNA mutations, but this level is significantly increased in Alzheimer's brains. The cause and effect connection is therefore yet to be made.

1.4.2 Mechanisms of accumulation of mutations through age

It is well established that mtDNA mutations undergo a process of clonal expansions and cause a respiratory chain deficiency mosaic in a multitude of tissues (Krishnan *et al.*, 2007). However, it is essential to reinforce that this is a clonal expansion of many mutations through the ageing process, across many tissues, rather than one monoclonal expansion in a tissue specific manner, which is usually described in mitochondrial disorders.

The origin of these mutations is debated; one could also argue there's an accumulation of mtDNA mutations from ROS production (by-product of the OXPHOS system) due to the lack of histones protecting the mtDNA (Harman, 1956; Richter C, 1995). However, replication errors may also be the cause of mutations through the pol γ lacking proofreading. More recently however, there's an argument for the accumulation of somatic mutations through age from clonal expansions of very low level mutations of pre-existing mutations created through replication errors during embryogenesis, or transmitted from the mother (D. C. Wallace, 2001; Larsson, 2010).

The mechanism (which occurs in post mitotic tissue) that has been suggested states that the mtDNA genomes which harbour deleterious mutations, have a replicative advantage over wild type mtDNA and thus replicate by clonal expansion (see Section 1.2.9). This is specifically important if the mutation resides in the D-Loop, the site of replication regulation. There have been studies quoting mtDNA region specific mutations that are associated with age. Amongst these mutations, point mutations at positions m.189 and

m.414 are known to be associated in aged individuals and have been proven not to be inherited (Michikawa *et al.*, 1999).

1.4.3 Polymerase γ mouse

As stated above, the mitochondrial theory of ageing proposes that the accumulation of mtDNA mutations lead to ageing and age associated disease. However, until the advent of a transgenic mouse model that expressed a proofreading deficient pol γ, the supporting evidence was limited.

The mice that express these pol γ enzymes with a lack of proofreading activity are born phenotypically normal; however, they accumulate mtDNA mutations rapidly throughout age until ultimately expressing a premature ageing phenotype, features of which include the following: weight loss, lipoatrophy, anaemia, reduced fertility and early death (Trifunovic *et al.*, 2004).

Despite the claim of these mice being the evidence of the mitochondrial theory of ageing, it has been quoted that these mice accumulate mitochondrial mutations at a rate that is 10-fold greater than that of normal ageing mice and that actually mtDNA deletions are driving the ageing phenotype (Vermulst *et al.*, 2007); this has led to disputes over whether this mouse and accumulation of point mutations is functionally relevant to normal human ageing. Although it is not conclusive, these mice do show a premature ageing phenotype and is fully consistent with the idea mtDNA point mutations create amino acid substitutions of the respiratory chain subunits; thus, fully in support of the theory of mitochondrial ageing (Khrapko and Vijg, 2007; Edgar *et al.*, 2009).

1.4.4 Pathophysiological effects of mtDNA mutations

The link between the pathophysiological effects on ageing of mtDNA mutations, have to a large extent been dependent on the use of mouse model described in Section 1.4.3. It is key to note that there is no perfect mechanism of transfecting mtDNA mutations in animal models and therefore, difficult to study naturally occurring mtDNA mutations. The downstream effects of mtDNA mutations have been identified as a result of reduced oxidative phosphorylation. Contribution towards ageing is likely driven through a number of mechanisms that result from a reduction in cellular respiration.

Human patients with mitochondrial diseases and mtDNA mutations have been reported to often have extensive cell death in the affected tissues, which is generally attributed to

an increased level of apoptosis from a reduction in mitochondrial proteins and the oxidative capacity of the tissue/cell (Oldfors *et al.*, 1990). However, most recent evidence also suggests that the accumulation of low level mtDNA mutations that are inherited can affect the lifespan of a subject (Ross *et al.*, 2014).

The abnormalities found within ageing human skeletal muscle has been attributed towards the accumulation of the large scale common mtDNA deletion, which is well characterised in low mitotic tissues, such as muscle and brain (Norman Arnheim and Cortopassi, 1992b; Lezza *et al.*, 1994). There is one study that examined the effect of mitochondrial dysfunction and cellular phenotypes in replicative tissues, resulting from age associated mtDNA mutations (Nooteboom *et al.*, 2010). It was found that this tissue had increased level of apoptosis and attenuated cellular proliferation, which lead to decreased future cell populations.

Whether mtDNA mutations are causally related to ageing and mitochondrial intrinsic ageing is debated, but supporting evidence continues to grow which aids the argument (Larsson, 2010). Although, current studies do not rule out the fact that mitochondrial damage and mtDNA mutations may be a biomarker of ageing (Khrapko and Vijg, 2007).

1.5 Studying Mitochondrial DNA *in vitro*

The key type of somatic mitochondrial DNA mutations in post-mitotic cells are large scale deletions. Whilst studying mitochondrial copy number in cells with a deletion is a reasonably well-accepted method using biopsy samples, it is almost impossible to successfully maintain primary cells that contain an mtDNA mutation, in a cell culture model. This is because when the primary cell line with a single deletion mutation is established; those cells which, by chance, contain higher levels of the deletion are rapidly out-competed by those cells with lower levels. Thus, the deletion becomes quickly lost from the culture system and it ultimately all reverts to wild-type mtDNA. This is because the deletion conveys an OXPHOS defect to the cell which means it replicates less well and highlights the distinct difference between post-mitotic cells mutation accumulation and mitotic cell mutation accumulation, i.e. post-mitotic cells have no means of losing deleted mtDNA.

1.5.1 Common cell lines used in mitochondrial research

There are a range of cell lines routinely used in mitochondrial research such as myoblasts, fibroblasts, HEK 293, cardiomyocytes and stem cells, to name a few. However, when studying mitochondrial DNA point mutations, myoblasts and fibroblasts are often popular cells to use *in vitro*, due to their ability to maintain both homoplasmic and heteroplasmic mutations, *in vitro* throughout passage. However, a relatively invasive clinical procedure is required to generate a primary cell line from a patient harbouring a mitochondrial mutation of interest. There are a range of studies successfully using such cell lines during *in vitro* studies for analysis of mutation involvement within mitochondrial disorders (Nishigaki *et al.*, 2003; Grazina *et al.*, 2007; Saretzki, 2009; Cox *et al.*, 2012; De la Mata *et al.*, 2012; Garrido-Maraver *et al.*, 2012).

1.5.2 Trans-mitochondrial cybrids

A trans-mitochondrial cybrid cell line is a cell line that has been derived from two separate cell lines, both with specific features of interest and fused together to form a new cell. These cells are specifically of interest due to their ability to maintain an mtDNA heteroplasmic mutation *in vitro*.

As depicted in Figure 1.4, to create a trans-mitochondrial cell line to use *in vitro*, an immortalised cell line nuclear background (such as an osteosarcoma) is rendered rho_0 (completely deficient of mtDNA by the use of a chemical substrate) and the cytoplasm of a primary cell line which contains a particular mtDNA genotype of interest is added.

When fused together, they give rise to a new cell with the nDNA background of the immortal cell with the mtDNA of interest. Such cell lines are potentially therefore, of great use in mtDNA mutation experiments and some cell lines have been quoted to maintain their mtDNA heteroplasmy status to a great number of passage due to their immortal nature (Diaz *et al.*, 2002).

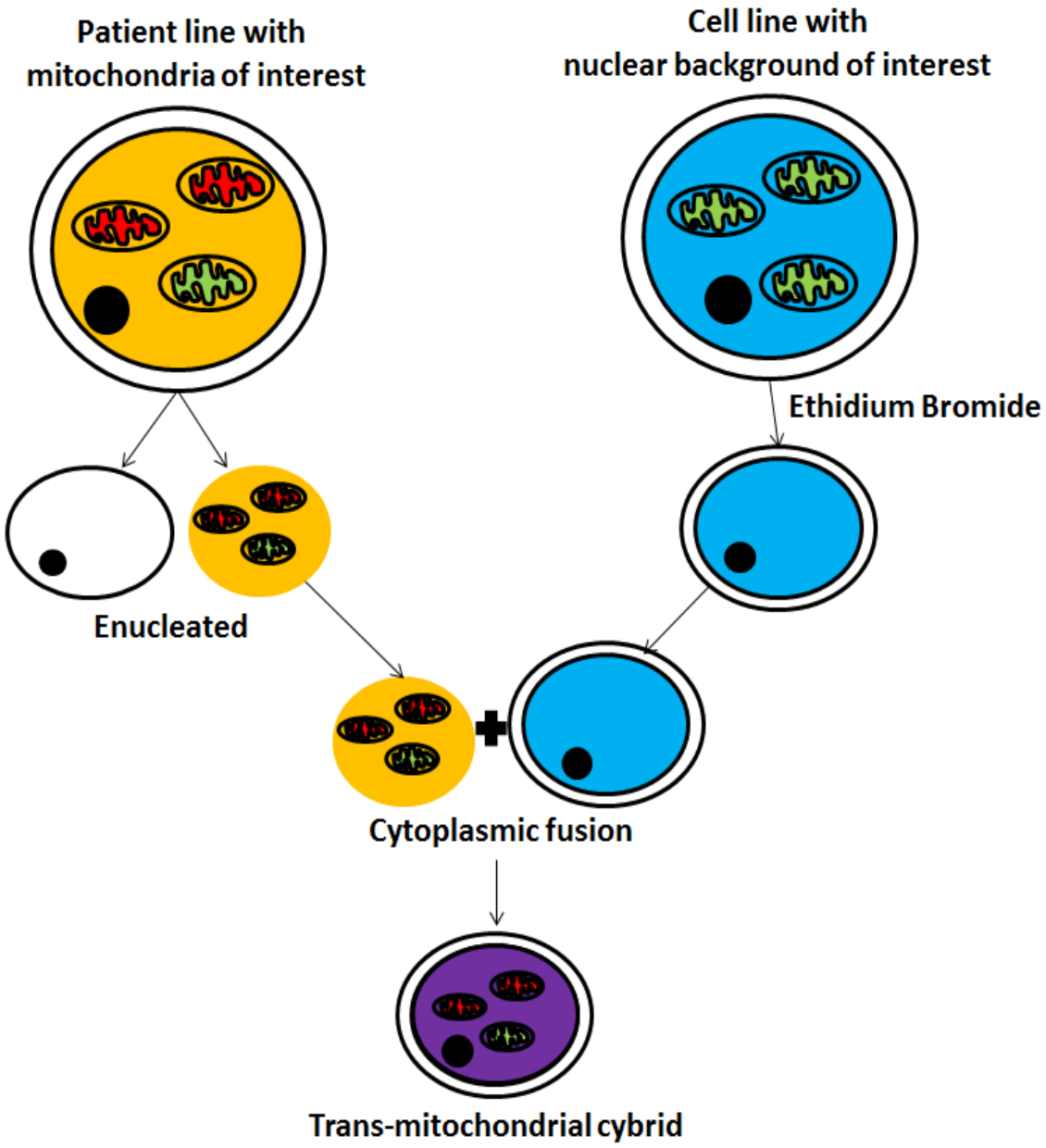

Figure 1.4 An overview of generating a trans-mitochondria cybrid cell line for *in vitro* studying of mtDNA heteroplasmic mutation(s). The patient cell line with mitochondria of interest is enucleated and fused with a mitochondrial deficient cell (ethidium bromide) with a nucelar background of interest i.e immortal. It is not clear however, why these cells are able to maintain their heteroplasmy.

1.6 Next Generation Sequencing of Mitochondrial DNA

Next generation sequencing (NGS), also known as high-throughput DNA sequencing or massively-parallel sequencing, is increasingly becoming a prolific research tool in the scientific community. The technology and the analysis tools are developing at a very rapid pace, and the proven reliability of the technique and increased understanding in how to interpret the data, has seen the application move into the clinical diagnostic community.

1.6.1 NGS Platforms and sequencing applications

There have been three prevailing sequencing platforms on the market: Illumina's platforms, Roche's 454 and Applied Biosystems' SOLiD systems; mtDNA sequencing has been successfully performed using all three (Craven *et al.*, 2010; Payne *et al.*, 2013) with none appearing to be clearly superior in detecting very low level mutations through a number of different approaches (Brodin *et al.*, 2013; Payne *et al.*, 2013; Gardner *et al.*, 2014). However, despite these claims, there has been no conclusive head-to-head comparison in the field of mitochondrial research.

Each sequencing platform has its own specific sequencing chemistry and therefore, its own sample sequencing workflow. However, a generic overview (see Figure 1.5) consists of initial amplification/enrichment of target regions, pooling the samples for sequencing and processing with the platform specific preparation kit. Finally, the products are run on the platform machine, producing data output as raw sequencing products (fasta, fastq).

NGS has been used in the detection of variants in a number of different mitochondrial disorders using DNA enrichment techniques such as long range PCR and chip-based capture methods (Calvo *et al.*, 2012; Elo *et al.*, 2012; Graham, 2012; Haack *et al.*, 2012). It is essential to enrich mtDNA during the process of NGS as it comprises only 1-2% of the total amount of DNA content. It has been more recently found that NGS of mtDNA is efficiently enriched using long range PCR and nuclear genes using RainDance emulsion PCR (Dames *et al.*, 2013).

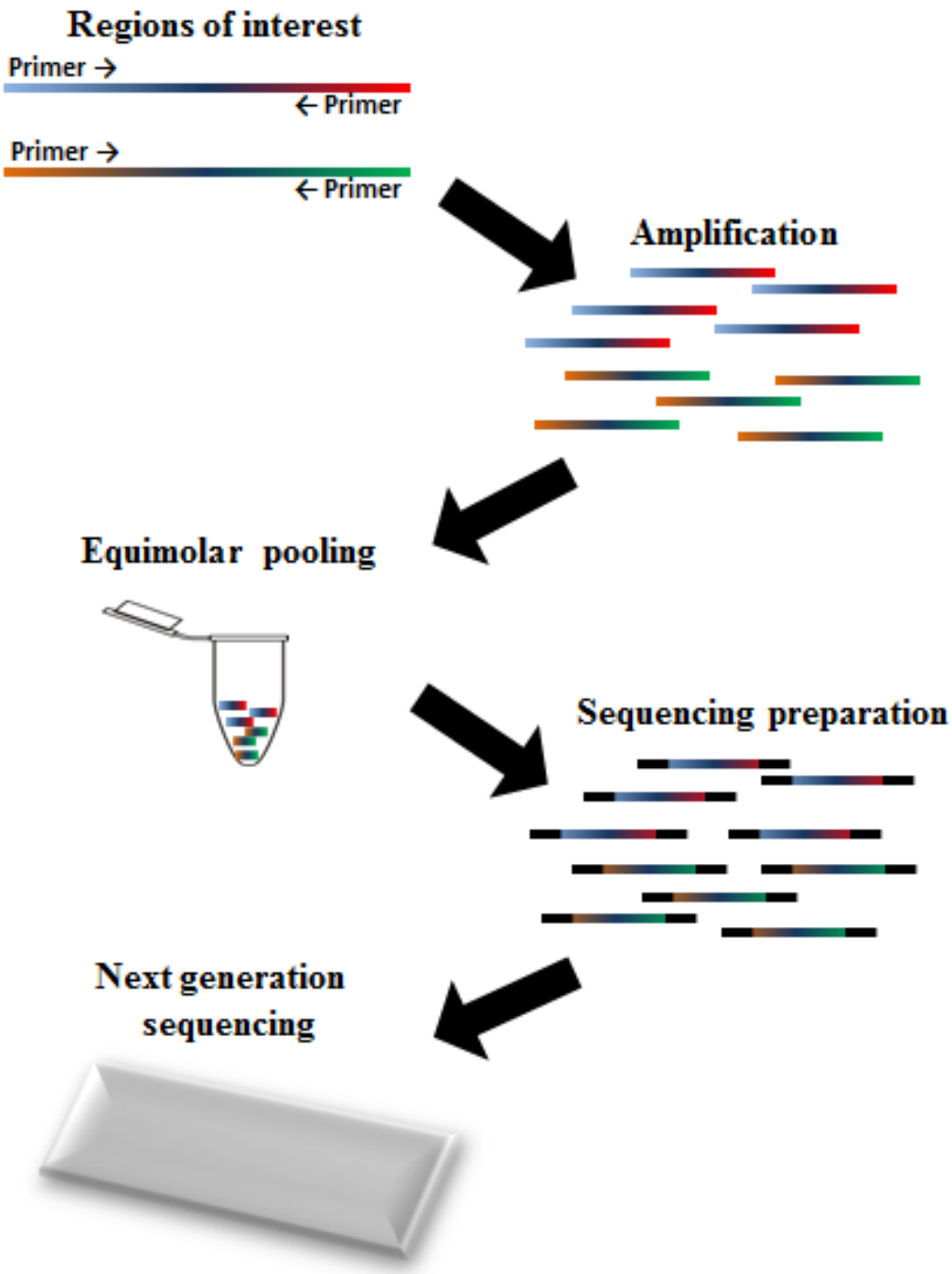

Figure 1.5 The overview of next generation sequencing methodology. The region(s) of interest are firstly amplified, then pooled and processed using the platform specific sequencing chemistry. Finally the products are sequenced on the platform.

1.6.2 Sequence quality

The quality of NGS data is typically based upon sequenced read information and Phred scores. Phred scores were originally developed to contribute towards the automation of DNA sequencing generated in the Human Genome Project. It is a mathematical derivation of base call quality when characterizing the DNA code sequenced (Ewing *et al.*, 1998).

Phred scores are logarithmically linked to the error probabilities of base call accuracy. Therefore, Phred scores are used as a determinant of sequencing accuracy and the quality-based sequence consensus of the data. Ewing and Green 1998, derived and implemented the equation to calculate Phred score, in their original sequencing quality analysis package, *Phrap* (Ewing and Green, 1998):

$$Q = -10 \cdot \log 10 \ (P)$$

Where; P = error probability of a given base call; Q = Phred quality score.

Phred values are calculated on a scale of 1-99 and are typically quoted in values of 10, where 10 is equivalent to a 1 in 10 chance that the base call is an error, therefore 90% accuracy. 20 is equivalent to a 1 in 100 chance of a base call error, so 99% accuracy, and so on. Standard NGS analysis pipelines generally use a Phred 30 as the minimum for base call error thresholds. Table 1.1 highlights Phred scores in 10's, along with their corresponding error rates and base call accuracy.

Phred score	Base call error	Accuracy
10	1 in 10	90%
20	1 in 100	99%
30	1 in 1000	99.9%
40	1 in 10000	99.99%
50	1 in 100000	99.999%

Table 1.1 Phred scores and the corresponidng base call error rate and accuracy

Other considerations of sequencing data quality include the base quality breakdown per read, which inherently has implications on the mapping quality. Mapping quality scores allow for assessment of high quality reads which align to the reference/consensus sequence, allowing for a more accurate base call downstream (mutant vs. reference).

1.6.3 Mitochondrial DNA-specific sequencing considerations

Due to the inherent differences between mtDNA and nuclear DNA, the application of NGS and data interpretation does not always align between the two.

When designing the long range PCR primers to amplify and enrich the mtDNA for sequencing, they should ideally lie in a relatively conserved region of the genome and generate an amplicon that is no smaller than ~2kb in size, to avoid an increased possibility of amplifying nuclear homologous regions, Numts (Hazkani-Covo *et al.*, 2010). Additional checks of primer specificity should be performed by the use of an *in silico* PCR system and/or negative amplification of an amplicon when using Rho-0 (mitochondrial deficient) DNA, as previously described (Payne *et al.*, 2011).

A consideration that should be made when analysing mitochondrial DNA is the reference sequence used in alignment. Although rCRS has become the standard, a debate has recently arisen as to whether or not the rCRS should be replaced with the Reconstructed Sapiens Reference Sequence (RSRS) on the grounds of ancestral similarities (Behar *et al.*, 2012). Although there's no real advantage of using either of the sequences, and manipulation of the reference sequence may lead to innumerable mistakes in NGS data interpretation. Potentially the most logical approach would be to choose the sequence that would best match the ancestral background of the samples in question (i.e. RSRS for African samples and rCRS for European samples).

An additional reference sequence related consideration is the artificially inserted 'N' at mitochondrial position 3107, found in the revised Cambridge reference sequence (rCRS) (Andrews *et al.*, 1999). If this is overlooked, the alignment will be incorrect and therefore the variant calling tool will falsely call variants at and around this position. The two approaches to resolve this issue is to leave the 'N' in the reference sequence and when 3107 is called as a variant, remove it in the post analyses; alternatively, position 3107 can be removed from the rCRS prior to sequence alignment, and the position of variants called after 3107 is adjusted by one (F. Ye *et al.*, 2014).

As previously described (see Section 1.6.1) there are a number of different sequencing platforms; however, one platform has been the predominant market leader from the outset, Illumina (F. Ye *et al.*, 2014). Due to the short read nature of the Illumina sequencing chemistry, a very important and often overlooked consideration that should be taken, is the effect of sequencing strand bias. This can be defined as a difference in the inferred genotype from the positive sequencing strand to the negative; with one reading homozygous and the other heterozygous (Yan Guo *et al.*, 2012). This is especially important when calling very low level variants. Therefore, careful implementation of a strand call filter should be taken to minimise false-positive calls being accepted in the variant analysis stage. Although, it is recognised that variant calling software may innately apply some form of weak strand bias filter in the default parameters. It is worth noting that it is presumed that 'strand biased variants' are indeed errors and not true variants.

1.6.4 Low level heteroplasmy detection

An area of mtDNA analysis where NGS has the greatest potential to supplant previous technology is in the detection of very low level variants. Studying the presence of very low level variants, below the level of ~1%, is likely to yield important insights into somatic mtDNA mutations; however, there's no defined 'gold standard' methodology to successfully do so and consequently, there are continual efforts to determine the lower resolution limits of NGS and the most appropriate method of defining it.

The greatest technical challenge of NGS application in low level variant discovery is in differentiating between true variants and background noise. The major intrinsic sources of noise are introduced through primary PCR and sequencing (either sequencing reaction or base calling) (Kozarewa *et al.*, 2009).

Whilst it was originally suggested that the lowest detection levels of mtDNA variants was 1.5%-10% (Y. He *et al.*, 2010; Mingkun Li *et al.*, 2010; Tang and Huang, 2010; Goto *et al.*, 2011), bacterial and virological small amplicon studies were quoting 0.1-1% (Daly *et al.*, 2011; Soares *et al.*, 2012). This was the accepted limit until novel mtDNA clone and phage studies resolved at a sub 1% threshold (Brodin *et al.*, 2013; Payne *et al.*, 2013) with both quoting successful resolution of mtDNA variants at ≥0.2%. Very low level variant detection has recently been used in supporting evidence of maternal inheritance of mtDNA mutations at a 0.1-0.5% threshold (Y. Guo *et al.*, 2013).

1.6.5 Bioinformatic considerations

It has previously been discussed (Section 1.6.4) that noise can be generated biologically through the process of NGS. However there is another consideration that comes into play when determining very low level variants; noise can paradoxically be induced artificially through poorly designed bioinformatics pipelines.

Perfectly duplicate reads must initially be removed; this is because duplicate reads may contain base errors that have been introduced through PCR (amplification stage), only unique reads should be used to avoid any potential false calls from these duplicates (Pireddu *et al.*, 2011).

The sequencing depth or 'coverage' of the data is a determining factor of what the lowest variant is that can be statistically resolved; i.e. if 10 supporting reads are required and if the sequencing coverage is 1,000, 10 reads in 1000 would resolve at 1%. Therefore the lowest detection level for calling mtDNA heteroplasmy would be 1%. Driving the sequencing coverage up statistically increases the low level variant resolution power (10 in 10,000 is 0.1%). Therefore, certain sequencing coverage is required to call variants using a particular threshold or MAF (minimum allele frequency).

The readily achievable minimum sequencing coverage/depth of NGS technology is usually found in the range of >5000x. However, this figure is often hugely exceeded based on an average run producing approximately 10Gb data and 30 samples of whole mtDNA genome per run. This thesis presents data with coverage data of typically 10,000-20,000x. Coverage is not of course, the only determinant of variant calling frequency. The base-calling error rate must also be taken into consideration when defining the lower resolution limit too. If the error rate o the machine is relatively high, then no matter what the sequencing coverage is, there will be no improvement in detection level, as described in Figure 1.6.

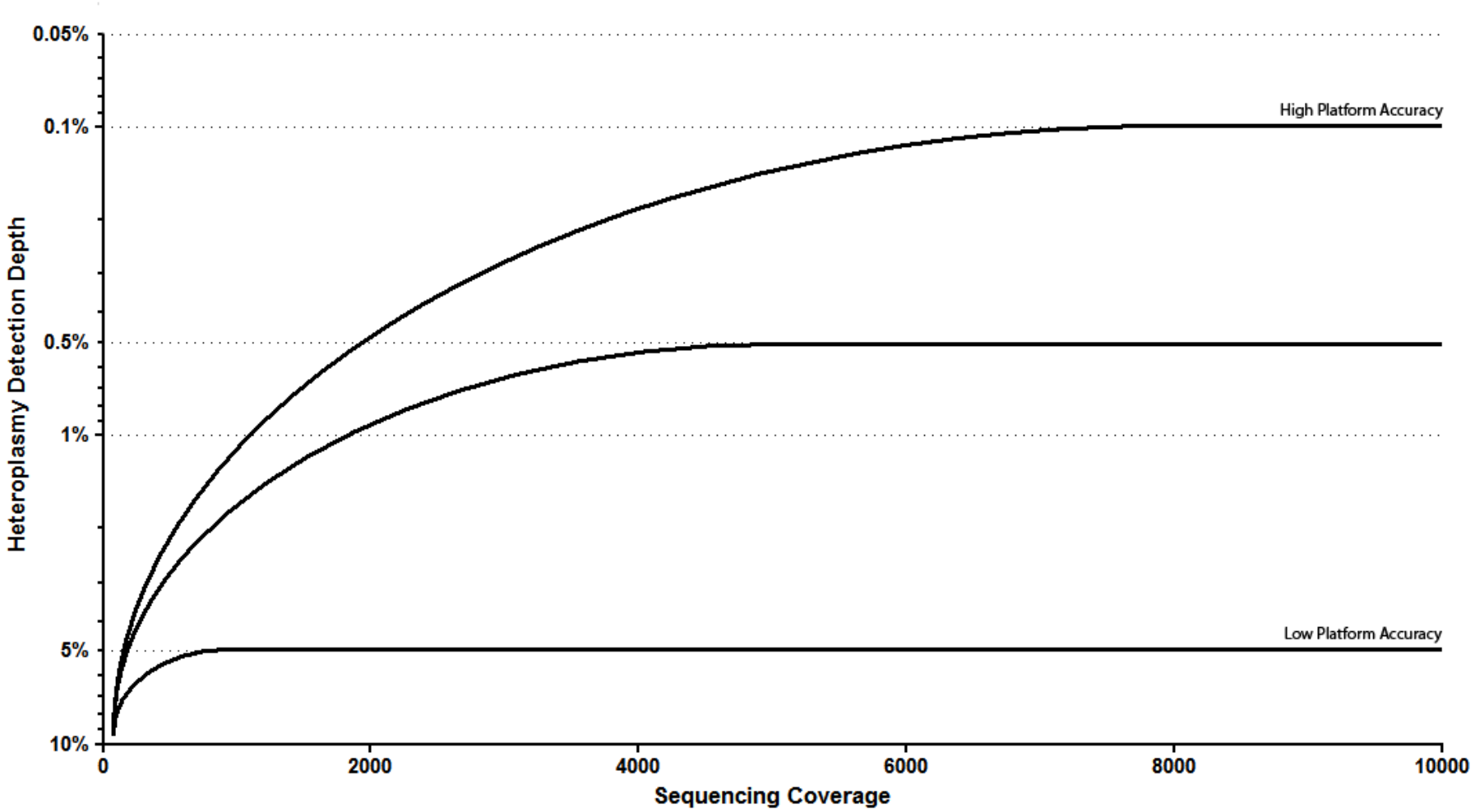

Figure 1.6 The intrinsic relationship of next generation sequencing platform base-calling error rate and sequencing coverage/depth, to the variant calling level/heteroplasmy.

Incorrect mapping of sequencing reads may also be a further source of error; although, read lengths may vary between platform chemistries. Theoretically, by taking the smallest NGS sequencing read length of 30bp, it is almost impossible to misalign the read to the human genome, i.e. perfect sequencing and perfect match to consensus sequence (Horner *et al.*, 2010); therefore, alignment of the mitochondrial genome should result in no mismatches. However, in reality, there's always a remaining cohort of reads that are termed 'unmapped' by the aligner and it's therefore likely, that another group exist termed 'incorrectly mapped'. Read mapping sometimes fails and variant calling can be affected by this. It has also been described that certain variant callers work better with certain aligners (Ulahannan *et al.*, 2013), albeit for no comprehensibly logical reason. However, the truth exists that unmapped reads may contain true sequence variants and this can be vital when assessing variants at the sub 1% heteroplasmy level.

1.6.6 Future of next generation sequencing

Since the advent of NGS, it has intertwined itself into all aspects of scientific research and continues to do so in the clinical setting. There have been a large number of improvements from speed to cost over the past decade of sequencing technologies, which only appears to be improving. The number of users of NGS continually increases (unpublished data, Qiagen, Manchester, UK) and specifically in the field of mitochondrial research, the understanding of how to uniquely utilise NGS, also increases.

NGS allows for both direct (targeted) and indirect (whole genome) sequencing of mitochondria, which is greatly improving the range of data available for mitochondrial research. The ability to detect low level variants is continually improving, with many studies confidently quoting variants, at or below, the 1% heteroplasmy level (Craven *et al.*, 2010; Calvo *et al.*, 2012; Brodin *et al.*, 2013; Dames *et al.*, 2013; Payne *et al.*, 2013; Gardner *et al.*, 2014). This is undoubtedly going to improve in the forthcoming years especially with the description of 'next' next generation sequencing technology (or third generation sequencing) to be single molecule high throughput sequencing.

1.7 Human Immunodeficiency Virus

The Human Immunodeficiency Virus (HIV) is a member of the retrovirus family and it is the causative agent for the acquired immune deficiency syndrome (AIDS). The HIV virus can be broken down into two species, HIV-1 and HIV-2. The species most prominently found in HIV-infected individuals is HIV-1. HIV-2 species is mainly isolated to West African infected persons (Peeters *et al.*, 2013).

1.7.1 HIV/AIDS pandemic

AIDS was first recognised as a disease in 1981 when an increasing number of homosexual men succumbed to opportunistic infections. The retrovirus, HIV-1, was subsequently deemed the causative agent of AIDS and since then has become one of the most devastating infections to have emerged in recent history.

Over the recent decades, HIV-1 has been quoted to have been the underlying reason for over 25 million deaths and infected more than 60 million individuals (Merson *et al.*, 2008).

1.7.2 HIV classification, structure and life cycle

HIV-1 is not one virus; it has been found to be comprised of four very distinct lineages, each of which resulted from independent cross-species transmission events from the Simian Immunodeficiency Virus (SIV). These lineages are termed M, N, O and P; group M was the first to be discovered and represents the causative form of the HIV pandemic (Sharp and Hahn, 2011). HIV-1 groups M and N are believed to have evolved from SIV chimpanzee strain (SIVcpz) and HIV-1 groups O and P from SIV gorilla strain (SIVgor).

The first insights into the ultra-structure of HIV-1 were not reported until nearly two decades after the report of the HIV index case, despite a developed understanding of the HIV lifecycle. A seminal paper in 1998 (Kwong *et al.*, 1998) highlighted the crystal structure of the virus and further elucidated the mechanisms behind viral entry into the host cell.

HIV glycoprotein 120 (gp120) found on the surface of the virus envelope, binds to the two-domain soluble CD4 construct and induces a conformational change of gp120, resulting in the exposure of chemokine (CCR5 and CXCR4) receptors to facilitate viral fusion and entry (Moore, 1997). Once the viral capsids are inside the host cell, reverse transcriptase (R.T.) begins to initiate the formation of double stranded viral DNA

through replication of the single stranded RNA molecules. The viral DNA migrates to the nucleus and integrates with the host's DNA through the enzyme integrase. Once integrated, the cell remains infected until the cell is destroyed. The provirus DNA is then replicated when normal cellular transcription occurs, creating more viral RNA. After translation of the RNA into viral proteins, the assembly of a new virus begins and particles move the cells outer membrane. To propagate, the host cell proteins cut the virus bud from the cells outer membrane and thereby create a new virus peptide (Barre-Sinoussi *et al.*, 2013) (overview of process in Figure 1.7).

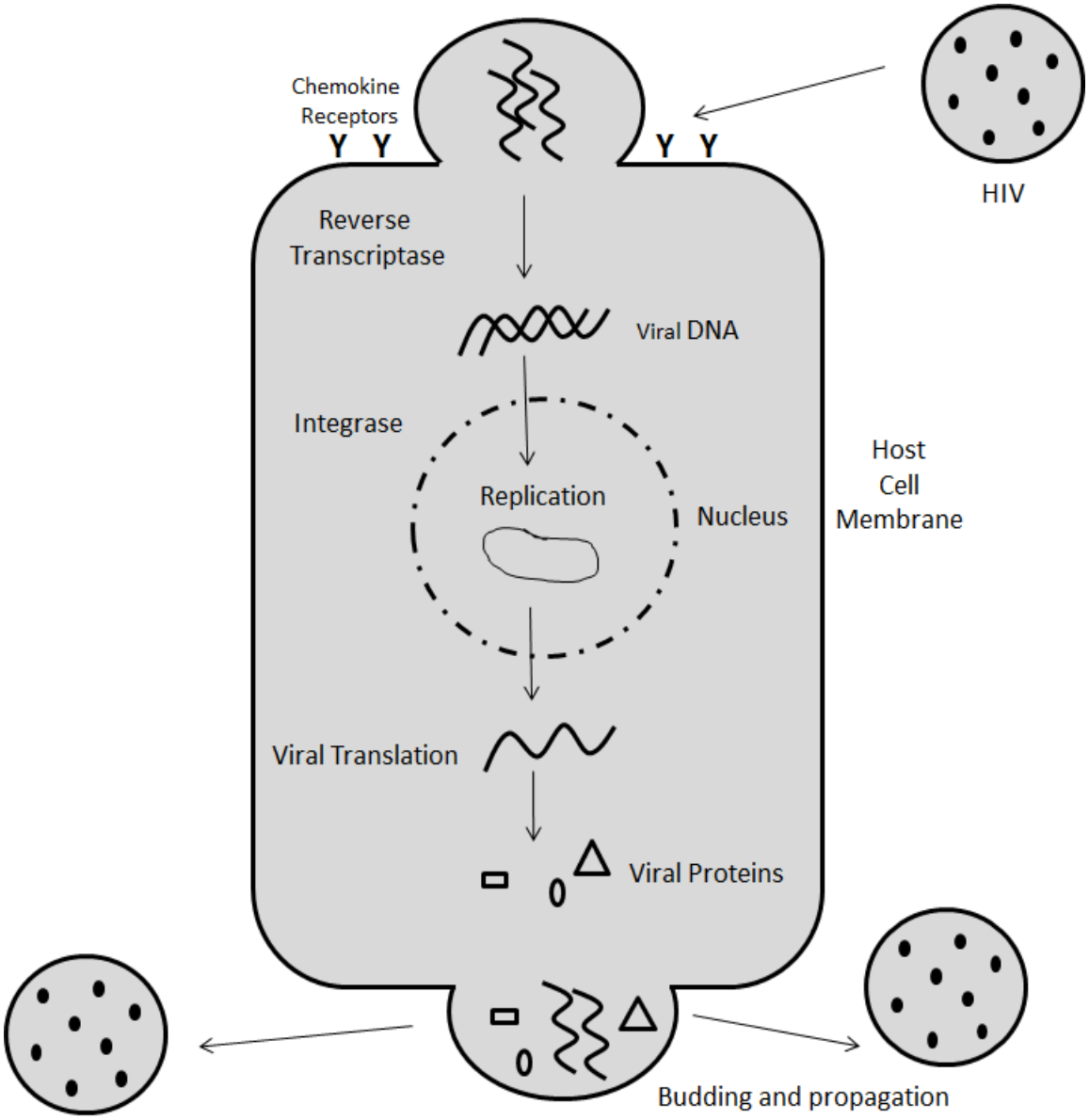

Figure 1.7 A schematic overview of the life cycle of HIV inside of a host cell.

1.8 HIV Therapy

The first drug was released in 1987 (zidovudine – AZT) and paved the way to the advent of highly active antiretroviral therapy (HAART) that is now the basis of HIV therapy (Rodriguez-Novoa *et al.*, 2006).

1.8.1 HAART and pharmacologic mechanisms

The management of HIV normally consists of the use of multiple (three or more) antiretroviral drugs that are comprised of two or more drug classes. This form of therapy is known as HAART (highly active antiretroviral therapy). Although there are multiple steps in the HIV replication process that can be targeted pharmaceutically, HAART usually consists of the use of two drugs from the nucleoside analogue reverse transcriptase inhibitor (NRTI) class and a third drug from either the protease inhibitor (PI) class, or the non-nucleoside analogue reverse transcriptase inhibitor (NNRTI) class (M. C. Dalakas *et al.*, 2001). However, there are other drug classes which may be used in HAART (as shown below in Table 1.2). The main reason for the success of HAART in treatment of HIV is the multi-step targeting system of multiple drugs to ensure that the virus is unlikely to escape through resistance and therefore allows durable suppression of viral replication and recovery of the immune system.

Currently, there are five drug classes which target different stages of the HIV cell cycle, namely these are: Entry inhibitors, NRTIs, NNRTIs, integrase inhibitors and PIs. Table 1.2 describes the mechanism of action for each drug class and the commonly used drugs within each.

Drug Class	Mechanism of action	Members
Entry/fusion inhibitors	Prevents the binding and fusion of HIV-1 to the host cell by blocking either T-cell CCR5 or CXCR4 receptors, or fusion of the virus to the host cell membrane.	Enfuvirtide Maraviroc
Nucleoside analogue reverse transcriptase inhibitor	Prevents the conversion of HIV RNA to DNA by inhibiting the action of HIV reverse transcriptase through premature chain termination.	Zalcitabine (ddC) Didanosine (ddI) Stavudine (d4T) Zidovudine (AZT) Tenofovir (TDF) Abacavir (ABC) Lamivudine (3TC) Emtricitabine (FTC)
Protease inhibitor	Blocks viral protease enzyme which is necessary to produce mature virons upon budding from the host cell.	Atazanavir Darunavir Indinavir Nelfinavir Saquinavir Loprinavir
Non-nucleoside analogue reverse transcriptase inhibitor	Prevents conversion of HIV RNA into DNA by binding to an allosteric site of HIV reverse transcriptase.	Efavirenz Nevirapine Rilpivirine Delavirdine Etravirine
Integrase inhibitors	Prevents viral DNA integrating into the DNA of the host cell by blocking the enzyme integrase.	Elvitegravir Dolutegravir Raltegravir

Table 1.2 A breakdown of each HIV drug class used in HAART and the associated mechanism of action by which they work. A non exhaustative list of drugs are also reported for each class.

1.8.2 Future directions of HIV therapy

Although current HAART therapy is highly efficient in combatting the virus, patients still present with residual viraemia and in as little as 2 weeks after HAART cessation, HIV RNA plasma levels return to a measurable level, highlighting the inability of HAART to completely remove HIV from the patient proving its purpose as a mere suppression tool (Joseph K. Wong *et al.*, 1997; Palmer *et al.*, 2008).

There are a range of renewed research efforts into finding the infamous vaccine cure for HIV which are showing some promise (Rerks-Ngarm *et al.*, 2009; Gao *et al.*, 2014). However, taking into consideration the persistent HIV latent reservoirs, an eradication cure is currently unlikely (Han *et al.*, 2007). Therefore, a drive towards a functional cure rather than an eradication cure is receiving more interest, with an aim of following the 'cancer model' where patients have an undetectable viral load with no disease progression and do not require other treatment to maintain the condition (i.e. HAART) (Autran *et al.*, 2011).

There's also continued research effort into developing new pharmaceuticals for HAART with a more efficacious and suppressive effect to treat HIV. New drugs include: new classes, new drugs within current classes, better efficacy and better tolerability. However, the main aim of such research is to inevitably improve long-term health of individuals receiving HAART, especially those displaying specific HAART side effects.

1.9 HIV Therapy and Associated Pathologies

Two decades prior to this writing, the main aim of HIV treatment was to prevent death of patients through the contraction of opportunistic infections, such as tuberculosis. Once therapy became more efficacious and the advent of HAART, the viral load suppression issue and avoiding opportunistic infections became under control. However, with this success came another concern, drug toxicities.

Antiretroviral toxicity was found to comprise of a range of organ specific toxicity, which was strongly associated with NRTIs. Although toxicity from protease inhibitors and non-NRTIs (NNRTIs) was also implicated, the major contributor to the new HIV pathology was NRTIs; evidence of which was found both *in vivo* and *in vitro* (P and Nelson A, 2006; Gardner *et al.*, 2013).

The range of presenting pathologies encompasses the entire biological system including: cardiomyopathy, nephropathy, hepatotoxicity, peripheral blood toxicity and lipodystrophy. The most commonly ascribed disorders include: lactic acidosis, skeletal myopathy and neuronal pathology, which are described in greater detail below (see Sections 1.9.1, 1.9.2 and 1.9.3).

1.9.1 Myopathy

The study that can be regarded as the seminal paper describing NRTI-associated pathology was published in 1988, one year on from the release of the first NRTI (AZT), which implicated AZT as the causative agent of myopathy in HIV patients. It was proposed that AZT, rather than HIV was the cause of the myopathy in AZT exposed patients (Helbert *et al.*, 1988). *Helbert and colleagues* described how the myopathy was found to resolve upon AZT cessation, but not on dose reduction.

Mitochondria were later reported to be the target of AZT myopathy, with findings in rodents that were exposed for 35 days with AZT, presented with selective changes in striated muscle, which were found to be ultra-structurally localised to the mitochondria upon high resolution electron microscopy analysis. The electron micrographs highlighted that there was a high abundance of abnormal mitochondrial architecture and an apparent lysis of mitochondria throughout specimens (W. Lewis *et al.*, 1992).

A short number of years later, another NRTI drug was implicated as being a causative agent for mitochondrial induced myopathy, ddC (C. H. Chen and Cheng, 1989). Although, as detailed below (Section 1.9.2), ddC was implicated in neuropathy in a

significantly higher number of cases than myopathy. It was later found in the landmark Delta trial that there was no increased prevalence of myopathy in the AZT+ ddC arm than the AZT-only arm (Darbyshire, 1996).

1.9.2 Neuropathy

Arguably the most prevalent complication amongst the HIV-infected community is neuropathy, with up to 60% of infected individuals displaying symptoms of some form of neuropathy, whether it be a severe case of distal sensory polyneuropathy, or minor axonal injury (Kamerman *et al.*, 2012).

ddC and ddI are the NRTIs most strongly implicated as being the causative agents for neuropathy in the HIV-infected community. There was strong association with axonal injury amongst subjects receiving ddC, with 55% of mitochondria morphologically abnormal and significant mitochondrial copy number depletion (up to 80% compared to normal) (Marinos C. Dalakas, 2001).

Given that ddC was discontinued in mid-1990 and ddI is rarely used clinically, it is surprising that the prevalence of neuropathy still has not disappeared. It has been recently found that d4T has been strongly associated with the incidence of neuropathy, with 21% of patients receiving d4T displaying symptoms of neuropathy (Affandi *et al.*, 2008; van Oosterhout *et al.*, 2012).

It is crucial to note that HIV itself has also been strongly implicated as being an underlying agent for causing sensory neuropathy among HIV-infected individuals through the neurotoxicity of HIV proteins (Moyle, 2000). Although often clinically indistinguishable, mitochondrial pathology is not always present in HIV-sensory neuropathy (Zhou *et al.*, 2007).

1.9.3 Lactic acidosis

The association of NRTI induced pathology driven through mitochondrial involvement is supported further by the fact, that lactic acidosis is prevalent among such patients (Gerschenson and Brinkman, 2004). The earliest claims of lactic acidosis and NRTIs implicated the ‘d-drugs’ (ddC, ddI and d4T) as the causative agent; however, AZT has also been implicated too (Baram and Cooke, 1993).

1.9.4 Lipodystrophy

The morphological signs of lipodystrophy in HIV-infected patients were first described approximately two years after the introduction of protease inhibitors (PIs) (Martínez, 1998). HIV-infected patients being treated with such drugs were found to present with a selective thinning of subcutaneous fat tissue in the trunk and extremities.

However, it was later found that the manifestations of lipodystrophy may actually be the consequence of mitochondrial damage from exposure to NRTIs and not PIs, as previously thought (U. A. Walker *et al.*, 2002). Due to the ability of NRTIs to cause a decrease in the content and quality of mtDNA content, tissue specific effect on fat tissue presented as wasting, likely through a reduction in mitochondrial encoded respiratory chain proteins and increased apoptosis (Kotler *et al.*, 2003).

It is interesting to note that lipodystrophy is not well associated with the most potent pol γ inhibitors, ddI and ddC, suggesting that there are likely to be additional adverse effects of the thymidine analogues within adipose tissue beyond mtDNA depletion.

1.9.5 Other complications

Protease inhibitors have also been implicated as a causative agent for HIV pathologies. The two most common pathologies described in the literature are diarrhoea and lipodystrophy (Martinez *et al.*, 2001; Mukhopadhyay *et al.*, 2002). There's also a suggestion that patients exposed to protease inhibitors also express an increased risk of cardiovascular disease; however, the increased risk was found to be marginal (Iloeje *et al.*, 2005).

Despite the potential implications for PIs to be the causative agents for HIV disorders, there have been studies implying that PIs can actually promote neuroprotection during HIV therapy. The mechanism ascribed to this process is through the inhibition of both Caspase-dependent and independent mitochondrial apoptosis, among many other potential direct and indirect mechanisms (Hisatomi *et al.*, 2008). However, a full discussion of PIs and other HAART drug complications is beyond the scope of this thesis.

1.9.6 Pol γ hypothesis

The proposed cellular mechanism driving mitochondrial NRTI side effects is known as the polymerase γ hypothesis (Lund *et al.*, 2007). Through structural mimicry, NRTIs are mistaken as nucleosides by pol γ during the process of replication of mtDNA. Although

there are five species of human DNA polymerase, for an unknown reason, pol γ is the only enzyme to possess a significant affinity for incorporating NRTIs, besides the drug target HIV reverse transcriptase.

Through incorporation by pol γ, the elongating mtDNA strand is truncated as a result of premature termination of mtDNA replication. Termination of mtDNA replication results in a depletion of mtDNA content (copy number) and ultimately leads to a decrease in respiratory chain protein production. Although mtDNA partly codes for sub units of some of the respiratory chain complexes, the most evident decrease of protein is found in complex IV, COX (see Section 1.1.3), as it is completely encoded by mtDNA and commonly used as the hallmark of mtDNA depletion and dysfunction.

Through the wealth of *in vitro* data, a hierarchy of NRTI inhibition of pol γ has become apparent. The NRTIs with stronger affinity for pol γ, are also more likely to be the cause of mtDNA depletion and cellular dysfunction in subjects (M. C. Dalakas *et al.*, 2001; Deveaud *et al.*, 2005; Cherry *et al.*, 2006; Bourdon *et al.*, 2007). The described hierarchy of pol γ inhibition is often deemed as follows: ddC > ddI > d4T ≥ AZT > 3TC (lamivudine) = ABC = TDF (Gardner *et al.*, 2013). There are also implications for HIV itself, affecting mitochondria and causing mitochondrial toxicities such as HIV-myopathy and HIV-neuropathy.

1.9.7 Mitochondrial DNA mutations and antiretroviral therapy

There has been *in vitro* data to support the notion of mtDNA levels recovering once the causative NRTI has been removed, in keeping with the polymerase γ hypothesis (G. A. McComsey *et al.*, 2005b). However, due to the increased number of HIV-individuals in industrialised countries now receiving NRTIs that are not inhibitors of pol γ, such as TDF and ABC (Wendelsdorf *et al.*, 2009), it begs the question, whether there are persistent effects of mitochondrial pathologies. The likely candidate here is the induction of mtDNA mutations (Bartley *et al.*, 2001).

There have been a range of studies implicating mtDNA mutations as a cause of mitochondrial NRTI-induced pathologies. In a seminal paper, an NRTI treated individual was described to have a fatal case of lactic acidosis, associated with an increased level of large scale mtDNA deletions (Bartley *et al.*, 2001), although it doesn't necessarily prove causality. The first longitudinal study which found an increase in mtDNA mutations through NRTI therapy came from studying mutations in PBMCs

of patients. The observations found development of new mtDNA mutations through NRTI therapy (Martin *et al.*, 2003).

Despite the great significance of mtDNA large scale deletions, there have been only a few studies specifically assessing the level of deletions in post-mitotic tissue in the HIV-setting (Lehmann *et al.*, 2011; Payne *et al.*, 2011). It is particularly interesting to see an increased level of mtDNA point mutations and deletion mutations in post-mitotic tissues, especially neuromuscular samples (Payne *et al.*, 2011). These data indicate the importance of quantifying mutations in patients who are currently or have been exposed to NRTIs, especially those that are known to be strong pol γ inhibitors.

The aforementioned process of polymerase γ inhibition, coupled with the accumulation of mutations through NRTI therapy, has led the current research field to pose the questions: Do mutations occur? How do they occur? What is the significance of them? Potential mechanisms are illustrated in Figure 1.8.

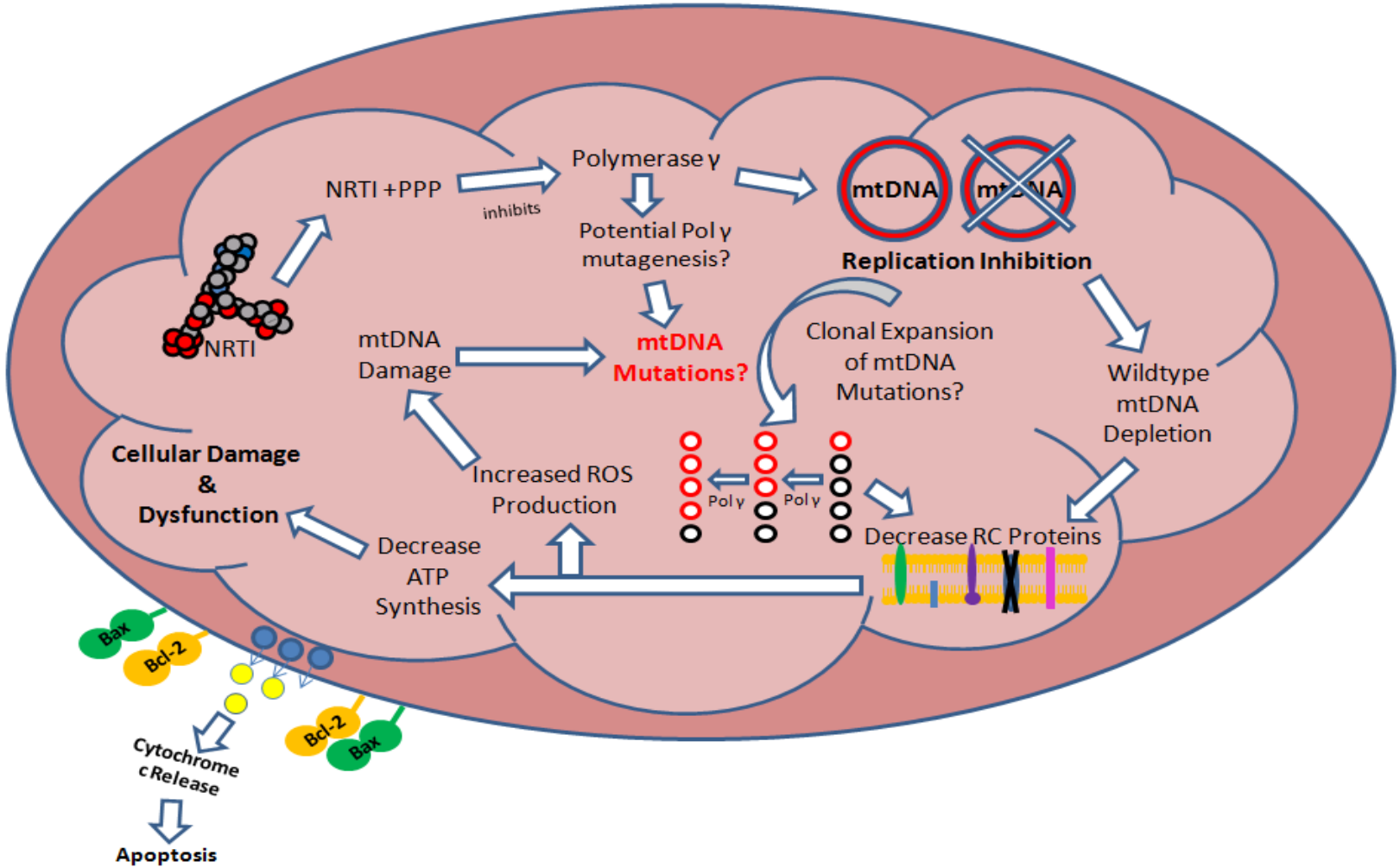

Figure 1.8 A schematic highlighting the polymerase γ hypothesis and the cellular consequences caused through NRTI inhibition of pol γ. The current paradigm that mtDNA mutations are also implicated in dysfunction is also implicated (Gardner *et al.*, 2013).

1.9.8 HIV and ageing

The positive impact that HAART has had on viral suppression and disease progression has led to a greatly improved mortality for HIV-infected individuals, with a large number of people now over the age of 50 (Psaros C *et al.*, 2014). However, co-morbidities are more common in older rather than young patients. There have been several studies that have observed cardiovascular disease, diabetes and osteoporosis are prominent among the HIV community, all of which are typically associated with ageing (Deeks and Phillips, 2009; Guaraldi *et al.*, 2011). Given the established associations between NRTIs and mtDNA, and between mtDNA and cellular ageing, there is a plausible link between the three which tells the tale of how the HIV community may present with premature aged phenotypes, known as accelerated ageing (an overview is shown below in Figure 1.9).

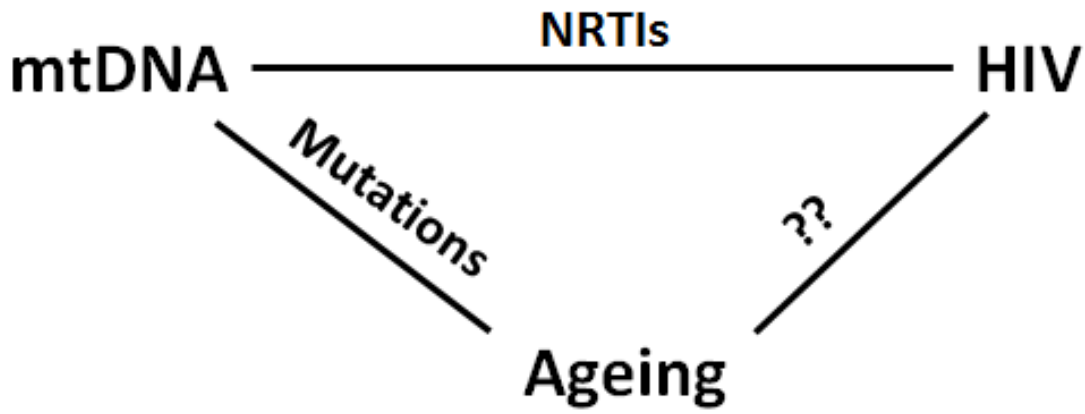

Figure 1.9 The overview of the process of HIV ageing potential driven through a relationship of mitochondrial mutations and the effect of NRTI therapy from HIV therapy on mtDNA.

The specific influence that HAART and specifically NRTIs have on mitochondria and ultimately ageing, are challenging to measure. This includes a lack of simple measures of ageing, and the lack of long-term follow up of patient receiving HAART. Such measures are critical for establishing the link between markers of intrinsic ageing the outcome for the HIV-infected individual.

With reference to Figure 1.8, measuring three key mitochondrial markers would give a clearer indication of ageing events in patients; Firstly, measurement of the level of mitochondrial content depletion; secondly, the mtDNA mutation load and finally, the amount of specific mitochondrial defects present, such as apoptosis and

COX-deficiency. All three are strongly associated with pol γ inhibition and severity of pathology, would give an indication of signs of progression or acceleration of premature ageing (see Section 1.9.6). However, ageing is multifactorial and it may be increased in HIV through mitochondrial dysfunction, driven through NRTIs by an interesting hypothesis (White, 2001; Anthony *et al.*, 2006; Payne *et al.*, 2011; Hulgan and Gerschenson, 2012).

In addition to NRTI induced accelerated ageing, there is also evidence to suggest that protease inhibitors may also contribute towards ageing. One study found an increased level of oxidative stress, senescence markers and inflammation of the endothelial cells in patients receiving PI treatment (Lefèvre *et al.*, 2010). Although this suggests that accelerated ageing could potentially be a result of more than just NRTI exposure, it also poses the fundamental question of ageing and its own causative factors outside of mtDNA involvement.

Chapter 2. Research Aims

In my Ph.D. project, I embarked upon elucidating the mechanisms behind the involvement of NRTIs (nucleoside analogue reverse transcriptase inhibitors) and the accumulation of mitochondrial DNA mutations, using the following aims:

1. To develop *in vitro* models to study the behaviour of mitochondrial DNA mutations, both deletions and point mutations, during and after the exposure of NRTIs.

2. To design and implement methods of detecting very low level (<1% heteroplasmy) mitochondrial DNA mutations through ultra-deep next generation sequencing technologies and apply them within the setting of NRTI exposure.

Chapter 3. Materials and Methods

Table of Contents

The core methods are presented within this chapter. Details of specific experimental design and chapter specific methods are described in the results chapters.

3.1 Cell culture

All cell culture was performed in a biological class II laminar airflow safety cabinet (HeraSafe; Thermo-Scientific, Hampshire, UK) to ensure a sterile environment to minimise infection of cultures. Phosphate buffered saline (PBS) and media was warmed to 37^{o}C in a temperature controlled water bath (Grant Bath JB AQUA 18 PLUS; Thermo-Scientific, Hampshire, UK) for 45 minutes prior to starting cell culture. All cultures were incubated in a cell culture humidifier at 37^{o}C and 5% CO_2 (Heracell 150i; Thermo-scientific, Hampshire, UK).

3.1.1 Trans-mitochondrial cybrid cell line maintenance

Trans-mitochondrial cybrids (as kindly provided by Professor C T Moraes, Miller school of medicine, University of Miami, U.S.A.) were used in the mtDNA deletion studies. They have been previously found to contain a single large-scale mtDNA deletion spanning the major arc from positions m.7982- 15504 (Diaz *et al.*, 2002) at a heteroplasmy level of ~70%. Two batches of these immortalised trans-mitochondrial cybrids were used in this project: ΔH2.1#4 and ΔH2.1#2 up to a passage of 30.

Optimisation experiments indicate that these cells were only cultured in T_{25} flasks (Greiner Bio-one, Stonehouse, UK) owing to the rapid growth rate observed in these cells. Smaller plates (6/9 well) plates may be used if short culturing times are used and propagation is not required.

The normal growth media of these cells was high glucose DMEM (Dulbecco's Modified Eagle's Medium; Gibco, Life-Technologies, Paisley, UK) supplemented with: 10% fetal calf serum (Sigma-Aldrich, Dorset, UK), 20µM gentamicin (Sigma-Aldrich, Dorset, UK) and 200µM uridine (Sigma-Aldrich, Dorset, UK). The cells were grown until 70-80% confluent before splitting or freezing. When culturing cells from batch ΔH2.1#2, it was determined that they require a higher volume of media so were grown in 10ml media in opposed to 5ml with cells from batch ΔH2.1#4.

3.1.2 Fibroblast maintenance

Two fibroblast cell lines containing a mitochondrial d-loop mutation (m.414T>G) were used and were a gift from Dr G Saretzki (Newcastle University, UK). These cells were derived from aged individuals. Cells were used from passage 5-11 to prevent abnormal

fluctuations in the heteroplasmy of the mtDNA mutation or abnormal cellular behaviour, which may result in skewed results and interpretation of the data.

The normal growth media of these cells was high glucose DMEM (Gibco, Life-Technologies, Paisley, UK) supplemented with: 10% fetal calf serum (Sigma-Aldrich, Dorset, UK) and 5% penicillin/streptomycin (Sigma-Aldrich, Dorset, UK). The cells were grown until 70-80% confluency in T_{25} flasks (Greiner Bio-one, Stonehouse, UK) before splitting or freezing.

3.1.3 Cell line propagation

To sub-culture and propagate the cell lines (cybrid and fibroblast), the media was aspirated and the cells washed with PBS (Oxoid, Thermo-Scientific, Hampshire, UK) and then treated with 5ml of 1x Trypsin-EDTA (Life-Technologies, Paisley, UK) for 5 minutes to detach them from the T_{25} flasks. The cells were transferred to a 15ml tube then pelleted at 1,300rpm for 5 minutes in a centrifuge (Eppendorf). The supernatant was aspirated and the pellet re-suspended in 1ml PBS. A cell count performed (see Section 3.1.6) and the correct dilution of cells re-suspended in 5ml of media in a new T_{25} flask (Greiner Bio-one, Stonehouse, UK) containing 5ml of media, and then replaced in a cell culture humidifier at 37^{o}C and 5% CO_2. The remaining cells were pelleted and stored at -20^{o}C until required for DNA extratcion.

3.1.4 Cryovial preservation of cells – freezing and thawing

Cells that were required for long term storage were frozen down and stored in liquid nitrogen. After splitting the cells (see 3.1.3) at 70%+ confluency, the pelleted cells were re-suspended in 1ml of freezing media containing: FCS (Sigma-Aldrich, Dorset, UK) and 10% Dimethyl Sulfoxide (DMSO; Sigma-Aldrich, Dorset, UK) which had been filtered through a 0.22µM filter (PALL, Portsmouth, UK). The re-suspended cells were placed into a 2ml cryovial (Greiner Bio-one, Stonehouse, UK) and into a Mr Frosty™ filled with isopropanol (Thermo-Scientific, Hampshire, UK) as per manufacturer's instructions, to allow for a slow freezing process of -1^{o}C per minute. The Mr Frosty™ was placed at -80^{o}C for 24 hours before transferring to liquid nitrogen for long term storage.

When retrieving cells from liquid nitrogen, cells were quickly thawed at 37^{o}C in a temperature controlled water bath (Grant Bath JB AQUA 18 PLUS; Thermo-Scientific, Hampshire, UK) and placed into a T_{25} culture flask containing 8ml media as previously

described (3.1.1 and 3.1.2) and incubated at 37°C and 5% CO_2 (HeraSafe; Thermo-Scientific, Hampshire, UK).

3.1.5 *In vitro* NRTI drug concentrations

NRTI concentrations were chosen to reflect that which has been previously determined to mimic expected *in vivo* concentrations (Lund *et al.*, 2007). The final concentrations used within all experiments (unless otherwise stated) are indicated in Table 3.1. All NRTI powders were dissolved in DMSO.

Drug (Abbreviation)	Concentration	Manufacturer
Zidovudine (AZT)	7.1µM	Sigma-Aldrich, Dorset, UK
Stavudine (d4T)	3.6µM	Sigma-Aldrich, Dorset, UK
Didanosine (ddI)	11.8µM	Sigma-Aldrich, Dorset, UK
Tenofovir disoproxil fumarate (TDF)	1.3µM	Santa Cruz Biotechnology, Texas, USA

Table 3.1 The final *in vitro* NRTI concentrations used in studies with their respective manufacturer details.

3.1.6 Haemocytometer cell counting

Cell counting and viability was performed using a haemocytometer (Improved Neubauer; Hawksley, Lancing, UK) and Trypan blue staining (Sigma-Aldrich, Dorset, UK). After cell pelleting and re-suspension in 1ml of PBS (Oxoid, Thermo-Scientific, Hampshire, UK), 50µL of the cells was added to a new 200µL Eppendorf microtube, mixed with 50µL of trypan blue and allowed to incubate for no longer than 5 minutes at room temperature. The trypan blue is a diazo dye and termed a vital stain. This means it will only stain cells that are dead/not viable because it is actively removed from the membrane of living cells.

After the incubation, 10µL of the cells was loaded into the haemocytometer for counting by placing the cell mixture onto the chamber and covered using a glass cover slip. The cells in the four corner squares of the counting grid were counted and an average taken (1 mm^2; volume 0.1 µL). This number was then imputed to the following equation to calculate cells per ml.

$$Cells\ per\ ml\ (n) = (Average\ cell\ count) \cdot Trypan\ dilution\ factor\ (2)\ \cdot\ 10^4$$

3.1.7 Mycoplasma detection

Each time a cell line was newly cultured or frozen down (cybrids and fibroblasts), cells were tested for mycoplasma infection using a luminescent detection kit (Lonza, UK), as per the manufacturer's guidelines. Briefly, 100µL of cleared cell supernatant was added to a fresh 1.5ml Eppendorf microtube. 100µL of re-constituted MycoAlert™ reagent was added to the sample and incubated at room temperature for 5 minutes, after which, a luminescence reading was taken (reading A, on a luminescent plate reader). 100µL of the MycoAlert™ substrate was then added to the sample and left to incubate for 10 minutes. The final luminescence reading was then taken (reading B) the ratio of reading B to A was calculated. The sample is deemed as mycoplasma negative if the ratio is below 0.9; borderline and re-test if between 1-1.2; mycoplasma positive if greater than 1.2. All cultures that were found to be mycoplasma positive were destroyed with the 2ml of 2% (w/v) Virkon® (Du Point, Hertfordshire, UK).

3.2 Single cell preparation

The molecular analyses of single cells from *in vitro* studies were performed using a novel technique. On the day of cell splitting and propagation, the remaining cells from the split were counted as previously described (see Section 3.1.6) and approximately 10,000 cells were re-suspended in 1ml of media (supplemented according to the experimental condition).

The edge of the membrane on a polyethylene naphthalate (PEN) membrane glass slide (Leica-microsystems, Milton Keynes, UK) was coated with a hydrophobic layer using an immunostaining PAP pen (VectorLabs, Peterborough, UK), see Figure 3.1.

The 1ml of re-suspended cells was then evenly distributed across the membrane slide and placed into a petri dish (VWR, Leicestershire, UK) at 37^{o}C and 5% CO_2 for 12 hours to allow the cells to fully adhere and morphologically re-shape. After 12 hour incubation the excess media was aspirated from the membrane slide, the slide was washed once with PBS and then left to air dry in a microbiology class II laminar airflow cabinet for 30 minutes.

Figure 3.1 Illustrative example of a membrane slide with the membrane highlighted in faded blue and the black exterior of the membrane representing the hydrophobic layer used to isolate cells for single cell analyses

3.2.1 Laser microdissection

Single cells were captured using Laser Microdissection (LMD) on an AS LMD platform and product-specific proprietary software (Leica-microsystems, Milton Keynes, UK). Live imaging was captured using a HV-C20 3CCD digital colour camera (Hitachi, Northamptonshire, UK).

Prior to dissection, machine calibration was performed. First, the reference point for collection tube alignment was checked by ensuring that the aperture ring was centralised on the viewing screen (Options > Settings> Move Reference Point). Next, the laser was calibrated by creating reference cutting marks on a sample slide. This was performed by loading a PEN membrane microscope slide (Leica-microsystems, Milton Keynes, UK) (membrane side down) onto the stage of the microscope and selecting specimen on the software toolbar; this moves the stage into position ready for viewing through the microscope eyepiece, or live camera feed through the software. The magnification was changed to cutting view (10x), then the laser calibration function was selected (Laser > Calibrate). The laser automatically cuts three crosses on the top and bottom corners of the viewing screen of the membrane. Next, using the mouse cursor, the centre of these cutting marks was selected. The system was now fully optimised for cutting.

All sample cutting was performed using the following settings; aperture, 10; intensity, 30; speed, 5; and bridge, medium. Single cells were cut leaving a small gap around the cell to avoid damaging the cell, and collected individually in sterile 0.5ml collection tubes (Thermo-scientific, Hampshire, UK). After capture, visualisation of the tube cap was performed to ensure successful capture. Collection tubes were subjected to 15 minutes of UV radiation using a UV hood before use, in order to reduce cross-contamination and to limit the static charge of the collection tube caps that may result in single cells being repelled, rather than collected. Once the desired number of cells was collected, all were processed using the single cell lysis method (see Section 3.2.2).

3.2.2 Lysis protocol

Once cells had been captured using LMD (see Section 3.2.1) the capture tubes were centrifuged (eppendorf) at 13,000rpm for 10 minutes to ensure that the cell was at the bottom of the eppendorf.

Whilst spinning, the lysis buffer was prepared and contained: 50mM Tris-HCL pH 8.5, 1% tween-20 (Sigma-Aldrich, Dorset, UK), 200ng/ml proteinase K (Invitrogen™,

Life-Technologies, Paisley, UK) and deionised water (sterilised by autoclave) made up to 1ml. Excess lysis buffer was disposed of and newly made every time.

After centrifugation, 12µL of lysis buffer was added to each capture tube and incubated for 16 hours at 55^{o}C in a temperature controlled digital dry water bath (Benchmark Scientific, New Jersey, U.S.A.). This step was followed by 10 minute incubation at 95^{o}C to denature the proteinase K. Subsequent molecular analyses were immediately performed after lysis at room temperature, to prevent further degradation of the sample.

3.3 DNA Extraction from cultured cells

DNA was extracted from cultured cells using the Blood and Tissue DNA extraction kit (Qiagen, Manchester, UK); all specific reagents were provided in the kit. The frozen cells, once thawed, were re-suspended in 200µL PBS (Oxoid, Thermo-Scientific, Hampshire, UK) and treated with 20µL of proteinase K solution. After thorough vortexing, 200µL of Buffer AL was then added. The sample was incubated for 10 minutes at 56^{o}C with frequent vortexing. After 10 minutes, or when the solution had become significantly less viscous, 200µL of 100% ethanol was added and mixed by vortexing, ensuring a homogenous solution was produced prior to continuing the rest of the procedure. This solution was then added to a mini-spin column and centrifuged at 8000rpm for 1 minute. The flow through was discarded, and a new collection tube used to spin (at 8000rpm for 1 minute) with 500µL AW1 solution. This was repeated with solution AW2, centrifuging at 14,000 rpm for 3 minutes. Finally, the column was placed in to a 1.5ml eppendorf microtube with 200µL AE elution buffer. The elute was quantified using a NanoDrop 2000 UV-Vis (Thermo-Scientific, Hampshire, UK) spectrophotometer and stored at -20^{o}C until required.

3.4 Polymerase chain reaction (PCR)

All standard polymerase chain reaction (PCR) reactions were performed using a hot start *taq* polymerase and amplified using a Veriti® thermocycler (Applied Biosystems, Life-Technologies, Paisley, UK). The master mix was set up in a Bio-Air UV PCR hood to ensure the samples were sterile and in a contamination free area. 50ng of each DNA sample was used and amplified in the following 25µL master mix reaction (all PCR reagents from Bioline, London, UK): 1x ImmoBuffer, 2mM dNTPs, 200nM forward Primer (IDT), 200nM reverse primer (IDT), 2.5mM $MgCl_2$ and autoclaved PCR-grade deionised water, 5U Immolase™ polymerase.

The product was then amplified under the following conditions: an initial denaturation step at 95^{o}C for 10 minutes, then 30 cycles of denaturation at 95^{o}C for 1 minute, annealing at T_m for 1 minute, finally extension at 72^{o}C for 30 seconds/Kb, with a final extension step at 72^{o}C for 10 minutes. The samples were finally cooled and stored at -20^{o}C until required.

3.5 Agarose gel electrophoresis of PCR amplicons

DNA fragments were separated using electrophoresis on an agarose gel utilising ethidium bromide as an intercalating fluorescent dye. According to the expected product size, a 0.8-2% (w/v) agarose gel was used to resolve the products; agarose (Helena Biosciences Europe, Gateshead, UK) was dissolved in 100ml 1x TAE buffer (Helena biosciences Europe, Gateshead, UK) and heated for 2 minutes in an 800W microwave oven. Once cooled, 0.4µg of ethidium bromide was added to the solution and then poured into a flat, horizontal gel casting tray and allowed to solidify at room temperature. 5µL of the PCR product was electrophoresed with 5µl orange G running buffer (50% dH_20, 50% glycerol [v/v] with few grains of orange G to colour the solution). Each gel had a designated sizing lane which was loaded with 5µL of the appropriate DNA ladder. Running time for all gels was 60 minutes under 55V in 1x TAE buffer. The gel was then imaged using a GelDocIt transilluminator gel imaging system (UVP, California, U.S.A).

3.6 Mitochondrial DNA monoplex quantitative polymerase chain reaction (qPCR)

Quantitative PCR (qPCR) provides a reliable and reproducible method to measure the amount of DNA being amplified in 'real time' during the PCR process. For all targets, an appropriate standard curve was employed to assess amplification efficiency, and a DNA negative control to monitor contamination.

Templates of genes of interest, for standard curves, were amplified using the standard PCR procedure (see Section 3.4) and the products were separated and imaged using a PCR product size specific agarose gel (see Section 3.5) using previously designed lab primers from Professor Chinnery's lab (for primers, see Table 3.2) and a DNA ladder (1Kb plus; Thermo-Scientific, Hampshire, UK).

Gene	Forward Primer 5' – 3'	Reverse Primer 5' – 3'	T_m (^{0}C)	Product Size (bp)
Template				
β2M	CGCAATCTCCAGTGACAGAA	GCAGAATAGGCTGCTGTTCC	60	1092
MT-ND1	CAGCCGCTATTAAAGGTTCG	AGAGTGCGTCATATGTTGTTC	60	1040
MT-ND4	ATCGCTCACACCTCATATCC	TAGGTCTGTTTGTCGTAGGC	60	1072
qPCR				
β2M	CACTGAAAAAGATGAGTATGCC	AACATTCCCTGACAATCCC	62.5	231
MT-ND1	ACGCCATAAAACTCTTCACCAAAG	GGGTTCATAGTAGAAGAGCGATGG	62.5	111
MT-ND4	ACCTTGGCTATCATCACCCGAT	AGTGCGATGAGTAGGGGAAGG	62.5	107

Table 3.2 Pre-designed primer details for mitochondrial (*MT-ND1* & *MT-ND4*) and nuclear (*β2M*) genes used in qPCR analyses

The fragments were extracted and purified from the gel using the QIAquick gel extraction kit (Qiagen, Manchester, UK), as per the manufacturer's recommendations. In summary, the band of interest was cut out of the gel quickly to minimise UV exposure, using a round end scalpel, carefully removing all traces of excess agarose. The gel slice was weighed and placed into a sterile 1.5ml eppendorf, and 1 volume (100µL) of buffer QG added per 100mg of gel and incubated at 50^{o}C until all the gel was dissolved; gel melting was facilitated by regular vortexing to evenly distribute heat throughout the mixture. Following this, 1 volume of isopropanol was added and then the entire mixture added to a QIAquick spin column and centrifuged for 1 minute at 13,000rpm. The flow-through was discarded and the column membrane was washed with 750µL of buffer PE by centrifugation for 1 minute at 13,000rpm. The flow-through was discarded and 50µL of elution buffer added to the spin column membrane, incubated for 10 minutes at room temperature, before final centrifugation at 13,000rpm for 1 minute, with the flow-through column replaced with a 1.5ml eppendorf microtube. The elute DNA concentration was quantified using the NanoDrop 2000 UV-Vis (Thermo-Scientific, Hampshire, UK) and stored at -20^{o}C until required.

The number of template copies per µL for each template was calculated using the following equation:

$$Copies\ per\ \mu L = \frac{Concentration}{molecular\ weight} \cdot K$$

Where; concentration = DNA in ng/µL (expressed as 10^{9}); molecular weight = template length (bp) x 2 x 330; and K = Avogadro's constant (6.022×10^{23} mol^{-1}).

The copies per µL of each template was standardised to 1×10^{10} and a serial dilution in 1Log$_{10}$ dilution steps, was amplified along with a DNA negative control on each qPCR plate. This was performed in 25µL reactions in a 96-well plate (Bio-Rad, Hertfordshire, UK), sealed using microplate 'B' plate sealers, using the fluorescent dye iQ™ SYBR® Green Supermix (containing $MgCl_2$, dNTPs, iTaq™ polymerase, SYBR® Green I, enhancers, stabilizers and fluorescein) (Bio-Rad). Each reaction contained: 1 x iQ™ SYBR® Green Supermix (Bio-Rad), 4mM forward and 4mM reverse primers (see Table 3.2) and PCR-grade autoclaved sterile deionised water (to make up to 25µL reaction). The cycling conditions consisted of initial denaturation at 95^{o}C for 3 minutes, followed by 40 cycles of denaturation at 95^{o}C for 10 seconds, with annealing and

extension at 62.5°C for 1 minute. A melt curve was also generated to analyse the specific and nonspecific PCR (i.e. primer-dimers) products generated during the reaction by 10 second incubation steps during which the temperature was increased from 65°C to 95°C in 0.5°C increments (see Figure 3.2).

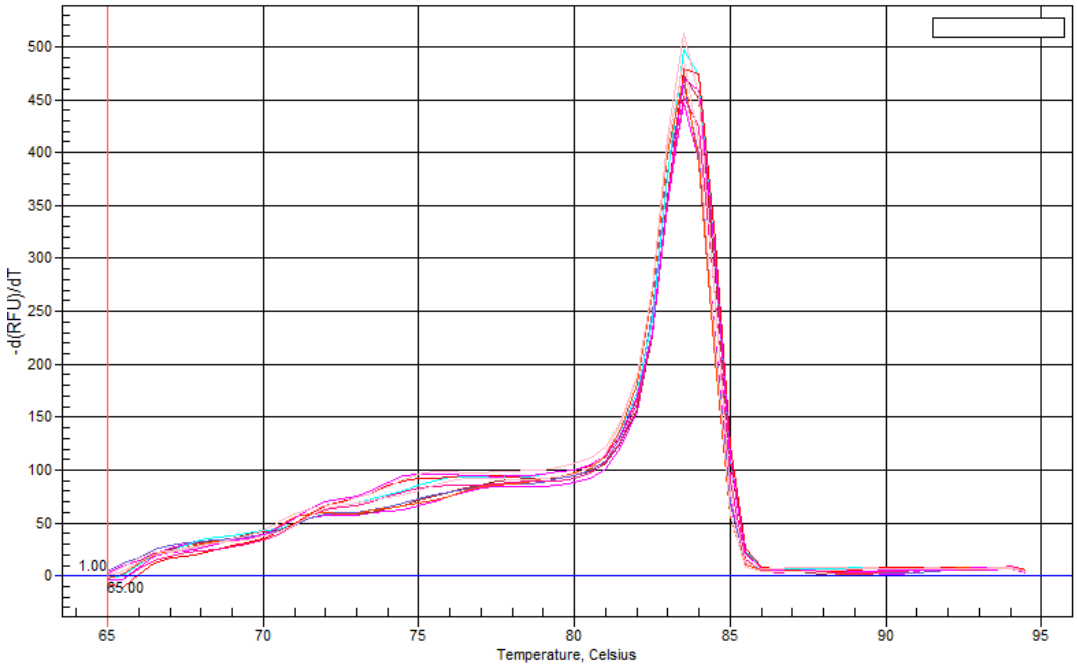

Figure 3.2 Melt curve of *MT-ND1* standard curve serial dilution generated using the iQ5™ thermocycler by iQ™ SYBR® Green method, showing the single expected product.

The serial dilution curve was plotted using the Ct (cycle threshold) and $\log_{10}$ concentration (copies per µL) to determine the amplification efficiency and coefficient of determination (R^2) to ensure the values lie between 98-102% efficiency and R^2 value of >0.98 (see Figure 3.3). All standards were loaded in triplicate.

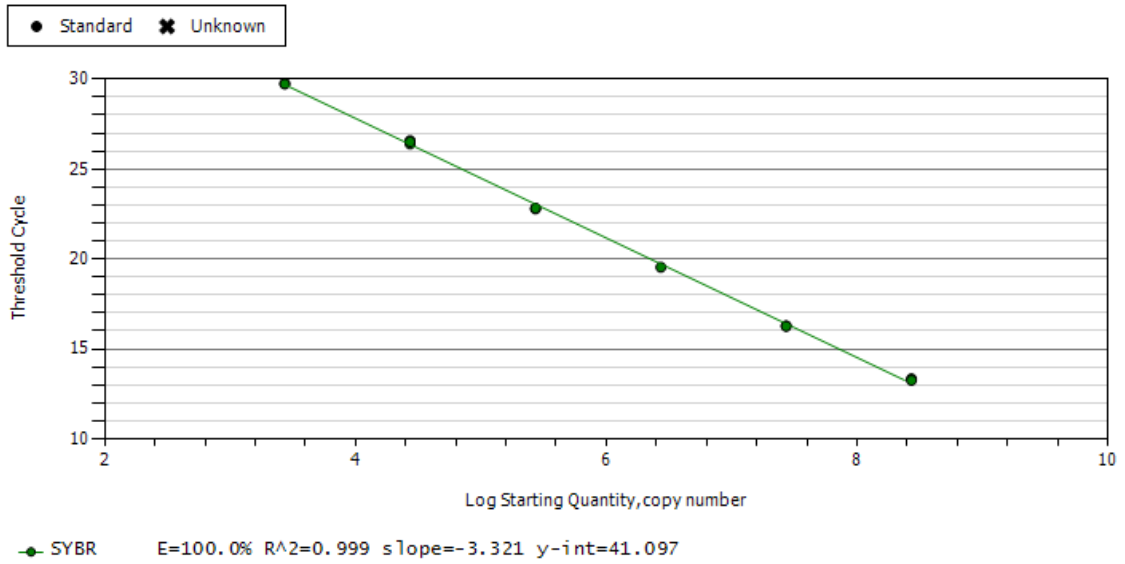

Figure 3.3 Standard curve of serial dilution for *MT-ND1* template from 10^8 – 10^3 copies per µL, highlighting optimal assay efficiency, generated using iQ5™ thermocycler by the iQ™ SYBR® Green method.

3.6.1 Mitochondrial multiplex qPCR design and optimisation

The quantification of the mitochondrial DNA deletion level and the relative mtDNA copy number was performed using a novel triplex qPCR assay. A specific probe based assay was designed due to improved accuracy, specificity and sensitivity compared to the traditional fluorescent dye incorporation method (as described previously in Section 3.6) to ensure improved accuracy when measuring low target content.

The amplicon probes were designed as nested fragments within the qPCR primer regions previously described (see Table 3.2). Given the primers' previously proven performance and specific amplification (see Figure 3.2 & Figure 3.3); it was deemed unnecessary to redesign the primers.

The probes were designed to be 20-25bp in size (see Table 3.3) and amplicon specificity was determined by the use of primer-BLAST (J. Ye *et al.*, 2012) to avoid amplification of homologous nuclear DNA regions, or multiple mitochondrial targets. Each probe was designed with a fluorophore that had emission spectra that didn't overlap with either of the other two probes to ensure background cross-talk didn't occur between targets. The probe quencher, responsible for absorbing excess excitation energy from the fluorophore, was chosen based on the greatest overlap between the fluorophore emission spectra and the absorbance spectra of the quencher (Ryazantsev *et al.*, 2014).

Gene	Fluorophore	Quencher	Amplicon sequence 5’ – 3’	Amplicon Size (bp)
β2M	FAM	BHQ-1	CCGTGTGAACCATGTGACTTTGTC	231
MT-ND1	HEX	BHQ-1	ACCCGCCACATCTACCATCACCCTC	111
MT-ND4	Cy5	BHQ-2	CAACCAGCCAGAACGCCTGAACGCA	107

Table 3.3 qPCR probes properties, used to specifically amplify each gene target within the multiplex qPCR reaction

When running the multiplex reaction it was initially found that there was a delayed and frequent 'negative amplification' (initial reduction in fluorescence signal) present in the nuclear *β2M* target (see Figure 3.4B). The inhibition 'negative amplification' was postulated to be due to PCR competition between the mitochondrial genes due to their greater abundance in relation to the *β2M*, meaning that reagents may have reached limiting dilutions for *β*2M. Alternatively, the effect may be due to the mitochondrial genes amplifying prior to the *β*2M gene, and the release of quenchers from the mitochondrial specific probes which result in an unintended quenching of the *β2M* fluorophore.

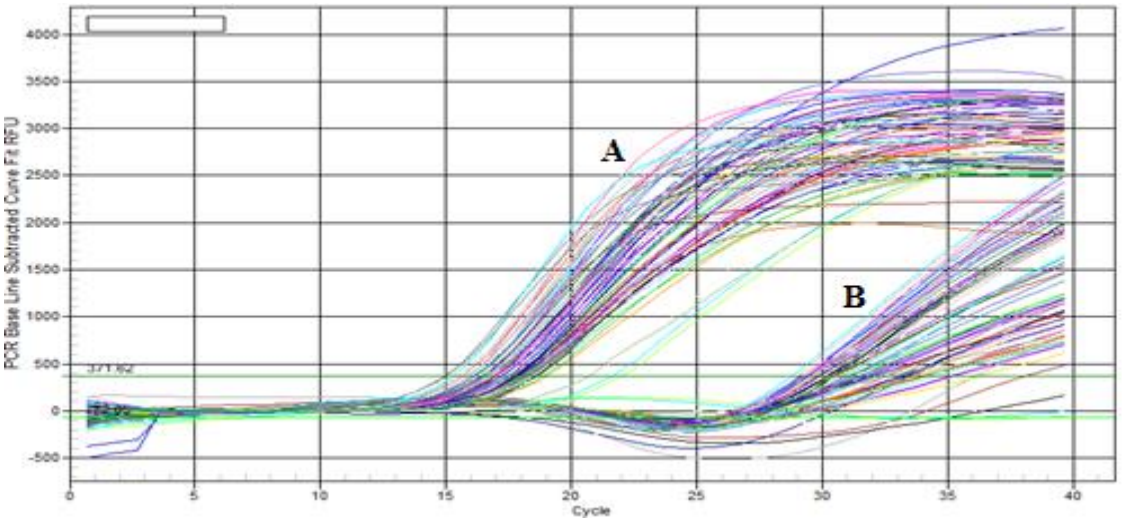

Figure 3.4 Delayed and negative amplification of nuclear target *β2M* (B) due to PCR competition between the abundant mitochondrial targets: *MT-ND1* (A) & *MT-ND4* (amplification data not shown). Amplifcation charts are of cybrid samples.

In an attempt to prevent PCR competition, a primer limiting concentration experiment was performed to determine the primer concentration at which no change in sample Ct was detected compared with monoplex and a reduction/elimination of *β2M* amplification inhibition was found in the multiplex reaction. Primer concentrations were 300nM for all targets; therefore, monoplex (single probe) reactions were run side by side with a 2-plex using a primer dilution series (mitochondrial gene dilutions of *MT-ND1* and *MT-ND4* with *β2M* concentrations unchanged – see Figure 3.5 for plate map example). It was found that the lowest concentration that left the Ct value for *MT-ND1* and *MT-ND4* unchanged was 75nM. This had a positive impact on the negative fluorescence resulting in the elimination of PCR competition (see Figure 3.6) and a linear amplification of all targets.

300nM	300nM	300nM		300nM/300nM	300nM/300nM	300nM/300nM
250nM	250nM	250nM		250nM/300nM	250nM/300nM	250nM/300nM
200nM	200nM	200nM		200nM/300nM	200nM/300nM	200nM/300nM
150nM	150nM	150nM		150nM/300nM	150nM/300nM	150nM/300nM
100nM	100nM	100nM		100nM/300nM	100nM/300nM	100nM/300nM
75nM	75nM	75nM		75nM/300nM	75nM/300nM	75nM/300nM
50nM	50nM	50nM		50nM/300nM	50nM/300nM	50nM/300nM
Negtiave	Negtiave	Negtiave		Negative	Negative	Negative
MT-ND1/4 monoplex				*MT-ND1/4 - b2M* multiplex		

Figure 3.5 An example plate map of primer limiting experiment to determine optimal primer concentrations for mitochondrial gene target primers.

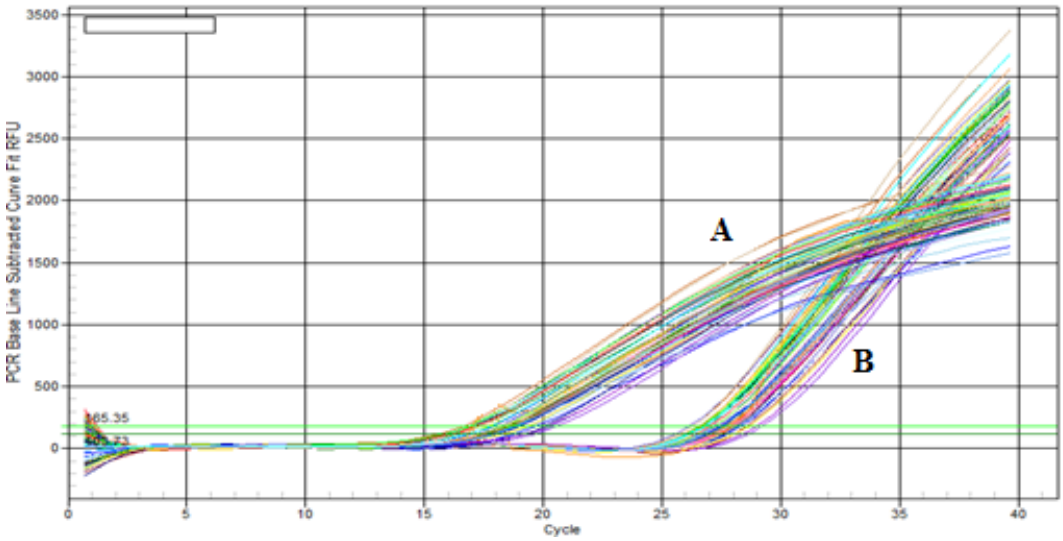

Figure 3.6 Amplification of *β2M* (B) with optimised mitochondrial target primer concentrations of 75nM for *MT-ND1* (A) and *MT-ND4* (amplification chart not shown). Amplification charts are of cybrid samples as illustrated in Figure 3.4.

The optimised multiplex assay used in the determination of mitochondrial deletion level and copy number quantification was performed using the iQ5™ thermocycler (Bio-Rad, Hertfordshire, UK) in 96-well clear plates (Bio-Rad), sealed with microplate 'B' plate sealers (Bio-Rad) in 25µL reactions containing: iQ™ Multiplex Powermix (Bio-Rad), 300nM *β2M* forward and reverse primers, 75mM *MT-ND1* primers, 75nM *MT-ND4* primers, 200nM each β2M, *MT-ND1* and *MT-ND4* probe, autoclaved PCR-grade sterile deionised water. All DNA samples were standardised approximately 50ng/µL.

The multiplex qPCR 2-step protocol consisted of an initial denaturation step at 95^{o}C for 3 minutes, followed by 40 cycles of denaturation at 95^{o}C for 10 seconds and an annealing/extension step at 60^{o}C for 1 minute.

3.6.2 Mitochondrial deletion level quantitation

MT-ND1 is a mitochondrial gene rarely deleted (L. He *et al.*, 2002), whilst *MT-ND4* is found in the centre region of the 'major arc' encompassing the majority of large scale mtDNA deletions (including that found in the trans-mitochondrial cybrid cell line). Therefore, given that in a non-deleted mtDNA molecule the Ct values of both genes should be the same, calculating the difference between the Ct values of the two will give the relative level of mtDNA deletion. To do this, the ΔΔCt method (Livak and Schmittgen, 2001) was implemented by using a calibration control mtDNA deletion-negative sample, and ensuring that the qPCR assay was performing at optimal efficiency by the use of a serial dilution of relevant standards and only accepting a run efficiency of 98-102% (see Section 3.6). The following equations were used to calculate the mitochondrial deletion level:

$$\Delta\Delta Ct = \left(Ct_{MT-ND4(x)} - Ct_{MT-ND4(C)}\right) - \left(Ct_{MT-ND1(x)} - Ct_{MT-ND1(C)}\right)$$

Where, x = sample and C = calibrator.

$$Deletion\ Level\ (\%) = 100 - (2^{\Delta\Delta Ct})$$

The level of deletion in samples was cross-checked with the result of a positive sample, which contained a previously determined deletion level in order to ensure that the calculated level was within ±3% of the 'true' deletion level.

3.6.3 Mitochondrial copy number determination

The mitochondrial copy number was calculated in a similar manner to the deletion level by making a calculation between the conserved *MT-ND1* and the nuclear housekeeping gene of known copies of *β2M* ,using the quantitative method. The SQ (starting quantity) was determined automatically for each sample based on an extrapolation of the standard curve values (assuming 100% efficiency); therefore using these figures, the mtDNA copy number per cell was calculated using the following equation:

$$mtDNA\ Total\ Copy\ Number = \frac{ND1\ SQ}{(\beta 2M\ \frac{SQ}{2})}$$

The above calculation will give total mitochondrial copy number per cell as it is presumed that copies of *MT-ND1* are sufficiently conserved to be equivalent to total mtDNA. The mtDNA wild type copy number per cell can also be calculated when using

the multiplex assay (see Section 3.6.1) due to the amplification of *MT-ND4* when a full mtDNA molecule is present, and therefore the SQ value is calculated. The equation for relative wild type copy number per cell therefore becomes:

$$mtDNA\ Wild\ Type\ Copy\ Number = \frac{ND4\ SQ}{(\beta 2M \frac{SQ}{2})}$$

Finally, the relative mutant (deleted) copy number per cell can also be calculated due to the simple law of summation. By subtracting the wild type copy number from the total, the result will leave mutant copy number:

$$mtDNA\ Mutant\ Copy\ Number = \ (Total\ CN) - (Wild\ Type\ CN)$$

3.7 Long range PCR for multiple mitochondrial DNA deletions

To investigate the presence of mtDNA deletions (single and multiple), a long range PCR method was implemented using the Takara LA polymerase (Takara, Clontech).

A 10kb mtDNA fragment was amplified. Each reaction was run on a Veriti® thermocycler (Applied Biosystems, Life-Technologies, Paisley, UK) in 50µL reactions containing: 1x LA buffer II, 2.5mM dNTPs, 200nM forward primer (see Table 3.4), 200nM reverse primer, autoclaved PCR-grade deionised water (made up to 50µL), 60ng DNA template and 2.5U *LA Taq* polymerase.

DNA was amplified under the follow conditions: an initial denaturation step at 94^{o}C for 1 minute, followed by 35 cycles of denaturation at 94^{o}C for 30 seconds, annealing step at 58^{o}C for 30 seconds and an extension at 68^{o}C for 11 minutes, with a final extension time of 12 minutes at 72^{o}C, before cooling the sample to 4^{o}C.

The 5µL of sample was electrophoresed on a 0.8% agarose gel as previously described (see Section 3.5) using 1kb plus DNA ladder (Thermo-Scientific, Hampshire, UK) to quantify the size of the bands. Each sample should produce a band at the ~10kb marker. Any other additional bands will correspond with a deletion of a specific size. 10kb minus the size of the additional band will give an approximate size of the deletion.

Name	Primer 5' – 3'	T_m (^oC)	Product size (bp)
10kb mtDNA Deletion	F - CCCTCTCTCCTACTCCTG R - CAGGTGGTCAAGTATTTATGG	58 58	9931

Table 3.4 Long range PCR primer properties used to screen for mtDNA deletions

3.8 Cloning

Cloning was performed using the pGEM-T easy vector system (Promega, Southampton, UK). Incubations at 37^{o}C were performed in a temperature controlled microbiology incubator (Thermo-Scientific, Hampshire, UK).

3.8.1 Product preparation

First, PCR amplification of the region of interest to clone was performed using a polymerase that inserted 3' A (adenine) overhang; Takara LA (Takara, Clontech) was used to do this. A 50µL PCR reaction containing: 1x LA buffer II, 2.5mM dNTPs, 200nM forward primer, 200nM reverse primer, autoclaved PCR-grade deionised water, 60ng DNA template and 2.5U *LA Taq* polymerase, was ran on a Veriti® thermocycler (Applied Biosystems, Life-Technologies, Paisley, UK) under the following conditions: initial denaturation at 94^{o}C for 1 minutes, followed by 30 cycles of denaturation at 94^{o}C for 30 seconds, annealing for 30 seconds (T_m specific), extension step at 68^{o}C for 1min/Kb, followed by a single final extension step of 10 minutes at 72^{o}C. The entire product was then run on a size specific gel as previously described (see Section 3.5). The product was gel-extracted using the QIAquick gel extraction kit (Qiagen, Manchester, UK) as previously described within Section 3.6 and DNA concentration quantified.

3.8.2 Ligation

Next, a ligation reaction was set up using the vector and ligase included in the cloning kit (Promega, Southampton, UK). Two ligation reactions were set up, a positive control and the extracted PCR product. Each ligation reaction consisted of: 2x Rapid ligation buffer, 50ng pGEM-T easy vector, 3U T4 DNA ligase, autoclaved PCR-grade deionised water and the volume of PCR product determined from the following equation:

$$\frac{ng\ of\ vector\ x\ size\ of\ insert\ (Kb)}{size\ of\ vector\ (Kb)} \cdot insert : vector\ molar\ ratio$$

For the positive control ligation, 2µL was loaded into the ligation. In an attempt to maximise the number of positive transformants from the reaction, ligation occurred at 16^{o}C for 16 hours.

3.8.3 Transformation

Ligation transformation was performed using 'one shot' JM109 high efficiency competent cells (Promega, Southampton, UK) using LB broth (Lysogeny broth) and LB agar plates as growth medium (see Table 3.5). After autoclaving and cooling to approximately 55^{o}C, the LB agar was supplemented with: 500µL Ampicillin (50mg/ml; Sigma-Aldrich, Dorset, UK), 750µL X-Gal (50ng/µL; Promega, Southampton, UK) and 250µL IPTG (1M; Promega, Southampton, UK). The mixture was then aseptically poured (~15-20ml) into 90mm plates underneath the flame of a Bunsen burner (Greiner Bio-one, Stonehouse, UK) removing any bubbles that may have formed and were left to solidify at room temperature on a level surface. Once cooled, plates were labelled and placed in an incubator at 37^{o}C or 4^{o}C for long term storage.

Reagent	LB Broth	LB Agar
Agar	N/A	7.5g
Tryptone	5.0g	5.0g
$NaCl_2$	5.0g	5.0g
Yeast extract	2.5g	2.5g
Autoclaved deionised water	500ml	500ml

Table 3.5 Reagents and corresponding measurements to make 500ml LB broth and LB Agar

Once the ligation reactions had been incubated for 16 hours, the tubes were briefly centrifuged and 1µL added to a pre-aliquoted amount of competent cells (25µL) on ice, gently agitating the tube to ensure the ligation and cells were sufficiently mixed. The tubes were incubated on ice for 20 minutes and then subjected to heat shock using 42^{o}C water using a digital temperature controlled dry heat bath with 250µL water in a well (Benchmark Scientific, New Jersey, U.S.A), for 50 seconds. After which, the tube was immediately placed back onto ice for 2 minutes. Then 475µL of LB broth medium was added to the cells transformed with ligation reactions and incubated at 37^{o}C for 1 hour. Finally, 100µL of the transformation culture was added onto the LB plates and spread around the agar to ensure an even coverage and growth of single cultures. This was performed using a glass spreader, dipped in ethanol and flamed briefly in a Bunsen burner. The plates were incubated for 16 hours at 37^{o}C.

After 16 hours incubation, a small number of white coloured colonies were selected and colony PCR performed to check for the correct insert. The growth of blue colonies indicates a successful transfection, but an unsuccessful ligation. Sterile 0.2ml 8 strip PCR tubes (STARLAB, Milton Keynes, UK) were used for the PCR. 10µL of autoclaved PCR-grade deionised water was added to one of the PCR tubes, one of the white colonies was selected using a clean eppendorf and dipped into the water before streaking onto another agar plate and then the tip placed in 5ml of LB broth (50ml falcon tube; Greiner Bio-one, Stonehouse, UK) supplemented with 5µL Ampicillin (50mg/ml; Sigma-Aldrich, Dorset, UK. The agar plate was incubated at 37^{o}C to expand the growth of single colonies and the broth/tip was placed at 37^{o}C, agitated at 150rpm for 16 hours. The PCR tube containing the dipped colony was used in colony PCR, treating the 10µL of dipped colonies as DNA. The PCR procedure described in product preparation was used (see Section 3.8.1). After checking for the presence of the correct size PCR product on an agarose gel (see Section 3.5) and the LB broth containing the colony selection tip had been incubated for 16 hours, the tube was centrifuged at 8000rpm for 10 minutes to pellet the colonies. The plasmid colony pellets were then frozen at -20^{o}C until required.

3.8.4 Plasmid purification

The DNA from the pelleted plasmid colonies (see Section 3.8.3) were processed using the using the Miniprep kit (Qiagen, Manchester, UK). The pellet was re-suspended in 250µL buffer P1 and transferred to a sterile 1.5ml eppendorf. 250µL of buffer P2 was then added and mixed thoroughly. 350µL of Buffer N3 was immediately added to the mixture and thoroughly mixed by vortexing. The tube was then centrifuged at 13,000rpm for 10 minutes to form a compact white pellet. The supernatant was then applied to a QIAprep spin column (supplied in kit) and centrifuged for an additional 1 minute at 13,000rpm. The column membrane was then washed by adding 750µL buffer PE and centrifuging at 13,000rpm for 1 minute. The flow-through was discarded and 50µL elution buffer was added to the spin column membrane and incubated at room temperature for 10 minutes prior to spinning at 13,000rpm for 1 minute into a sterile 1.5ml eppendorf microtube. The DNA was quantified using a NanoDrop and frozen at -20^{0}C until required.

3.9 Pyrosequencing

To quantify the level of mitochondrial D-loop point mutation m.414T>G a pyrosequencing assay was designed. Pyrosequencing is a method of DNA sequencing (sequencing by synthesis) short amplicons. Locus specificity is achieved by the use of specific primary PCR primers to generate ~200bp amplicon and then a third sequencing primer (which is biotinylated to allow for specific isolation of product) is used to amplify the 8-10 base-pairs of interest to quantify a mitochondrial heteroplasmic mutation within a sample. The following technique was performed using the PyroMark™ Q24 (Qiagen, Manchester, UK) using shallow 24 well sequencing plates (Qiagen, Manchester, UK) and 0.2ml sterile 8 strip PCR tubes (Greiner Bio-one, Stonehouse, UK).

3.9.1 Pyrosequencing m.414T>G assay design

Primers were designed using the PyroMark™ assay design software v2.0.6 (Qiagen, Manchester, UK) according to manufacturer's recommendations and primers with the highest quality score (100 indicating best predicted PCR performance) were chosen for the assay (see Table 3.6).

Primer	Sequence 5' – 3'	Primer position	Product size (bp)	T_m (^oC)
Forward	CACAGCCACTTTCCACACAGA	256-276	244	60
Reverse	5biosG/CGGGGGTTGTATTGATGAGATT	499-478		60
Sequencing	TTTATCTTTTGGCGG	399-413	*N/A*	60

Table 3.6 Pyrosequencing m.414T>G primer properties as designed by the PyroMark™ Assay design software.

The assay method was designed using the PyroMark™ Q24 software (Qiagen, Manchester, UK), which was also used for running the assay. A 10bp region around the m.414 position was inserted with the change (T>G) expressed as the IUPAC nucleotide code, K. The dispensing order of the nucleotides when sequencing was generated and the assay locked from accidental editing.

3.9.2 Pyrosequencing m.414T>G procedure

A primary PCR of samples was initially performed, as previously described (see Section 3.4. The PCR cycling conditions consisted of an initial denaturation of 95^{o}C for 10

minutes, followed by 30 cycles of denaturation for 30 seconds at 96^{o}C, annealing at 60^{o}C for 30 seconds and extension time of 30 seconds at 72^{o}C. The final extension step was at 72^{o}C for 10 minutes. 5µL of the PCR reaction was run on a 2% agarose gel to check for purity and presence of the correct band size (see Section 3.5). The size of the band was checked using the DNA hyperladder IV (Bioline, London, UK).

Once it was proven that the PCR reaction had generated the correct product size, 10µL of each PCR sample was then transferred to a new sterile 8 strip PCR tube (STARLAB, Milton Keynes, UK). To each sample, 70µL of binding buffer mixture was added containing: 40µL binding buffer (Qiagen, Manchester, UK), 2µL sepharose beads (GE Life science, Amersham Place, UK) and 28µL autoclaved PCR-grade deionised water. The caps were securely fastened and transferred to a plate agitator (Quantifoil Instruments GmbH, Jena, Germany) for 10 minutes at 2000rpm. On each run a mutant positive control, mutant negative control and 50:50 (mutant: wild type) sample was also included.

The PyroMark™ Q24 workstation was used for preparing the samples for sequencing. Once the samples had been agitated for 10 minutes, they were placed into the deep plate support wells, located at the front of the workstation. The handheld processing unit (‘hedgehog’) was cleaned by switching the vacuum pump on and washing thoroughly with de-ionised water. It was then placed and left in the parking station. The Q24 wash reservoirs were then filled to the visible line with 100% ethanol (Sigma-Aldrich, Dorset, UK), denaturation solution (Qiagen, Manchester, UK) 1x wash buffer (Qiagen, Manchester, UK) and autoclaved PCR-grade de-ionised water. The wash buffer was filled slightly above the line to ensure full cleaning of the hedgehog pins occurred.

The annealing buffer mixture was then made up containing: 300nM pyrosequencing primer and annealing buffer (Qiagen, Manchester, UK). A shallow 24 well sequencing plate (Qiagen, Manchester, UK) was placed on the workstation and 25µL of annealing buffer added to each well.

Using the hedgehog unit, all liquid was aspirated. The captured beads on the hedgehog pins were then processed through the ethanol for 15 seconds, denaturation buffer for 15 seconds and then the wash buffer for 30 seconds. After this step, the suction was disabled on the hedgehog and the hedgehog held vertically to allow excess liquid to drain from the pins, to avoid any interference with the sequencing reaction. After 30

seconds of drying, the pin heads were placed into the annealing buffer mixture on the sequencing reaction plate to release the beads into the mixture. The sequencing plate was then heated for 2 minutes at 80^{o}C using a temperature controlled dry water bath hot block (Benchmark Scientific, New Jersey, U.S.A) and a reaction plate heating unit (Qiagen, Manchester, UK). The sequencing plate was then placed into the pyrosequencing machine.

The PyroMark™ assay run file was generated using the PyroMark™ Q24 software by selecting File > new run. The m.414T>G assay file previously designed (see Section 3.9.1) was loaded onto the plate map and the correct sample names given to each well on the plate. The instrument cartridge was selected and a name given to the run. After saving the run file onto a removable device, the pre-run information was recorded (enzyme, substrate and nucleotide volumes).

The sequencing cartridge was then loaded with the pre-determined volumes (as calculated by the PyroMark™ software based on number of reactions) of enzyme substrate and nucleotides (PyroMark™ Gold reagents; Qiagen, Manchester, UK) and the cartridge placed into the dispensing unit in the machine. For a full sequencing plate the typical volumes loaded were; enzyme, 119µL; substrate 119µL; and 57µL of each nucleotide.

Once the assay was finished, the data was analysed using the PyroMark™ Q24 software. The pyrogram produced (see Figure 3.7) was graphed and statistical comparisons made accordingly (GraphPad™ Prism v5, CA, U.S.A). The sequencing cartridge was cleaned using deionised water and the sequencing reaction plate disposed of.

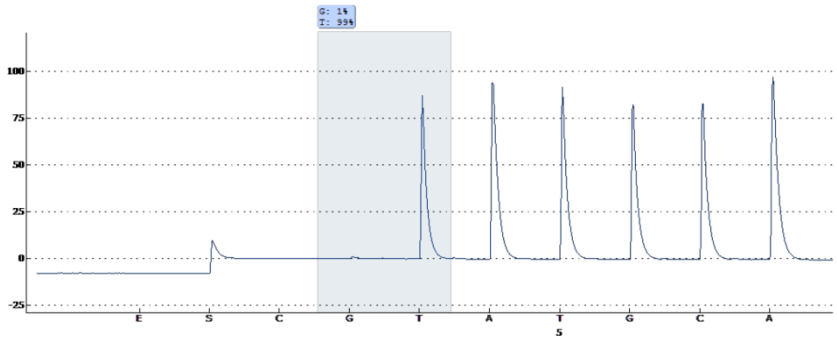

Figure 3.7 A representative pyrogram generated from a pyrosequencing reaction of a positive wild type control (100% T). The first three base codes are the following controls; E, enzyme; S, substrate and C, internal control.

3.9.3 Generation of assay standards

To ensure that each pyrosequencing run was performed optimally and efficiently, a series of standards were made with determined heteroplasmic levels of m.414T and m.414G. To achieve a pure culture of populations, cloning was performed (see Section 3.8) using an aliquot of DNA extracted from fibroblast m.414T>G cells. The wild type and mutant specific DNA generated from cloning was then used on every pyrosequencing run in the following way: 100% T, 100% G and 50:50 (T: G). The 'peaks' of each sample were visually inspected any samples flagged as 'red' with warnings by the software were excluded.

To assess the quantification accuracy of the assay, a standard curve was generated by making multiple dilutions of wild type to mutant, ranging from 0:100 (mutant: wild type) through to 100:0 (mutant: wild type). The data was plotted as observed to expected heteroplasmy level and is expressed the Figure 3.8.

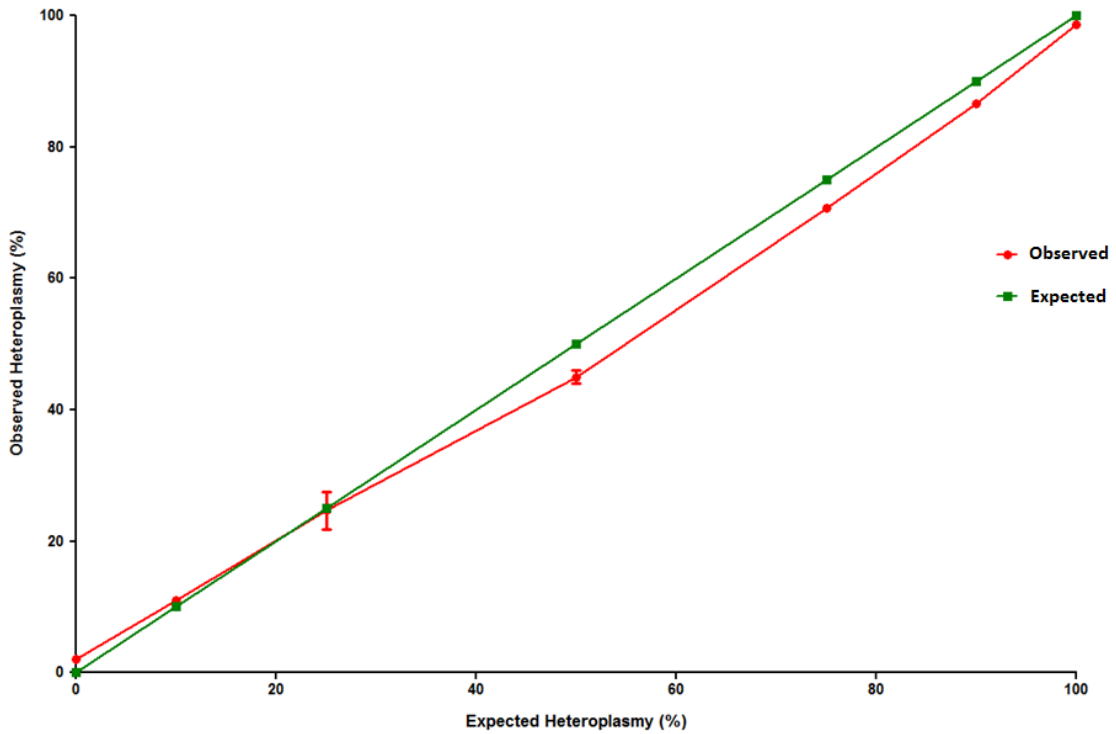

Figure 3.8 The observed heteroplasmy and the expected heteroplasmy of pyrosequencing standard curve dilutions. Green line indicates where y=x and the red line is the standard curve.

3.10 Next generation sequencing

To assess the low level mutation burden in cells treated with NRTIs (human trans-mitochondrial cybrids and fibroblasts) and aged mouse (*Mus musculus*) samples exposed to antiretrovirals; next generation sequencing (NGS) was performed using the

MiSeq™ platform (Illumina, Cambridge, UK). All pipetting was performed using sterile filter tips (STARLAB, Milton Keynes, UK).

3.10.1 Long range PCR primer design and amplicon enrichment

In order to amplify the mtDNA region of interest, long range PCR primers were designed for each experiment (see Table 3.7).

The primers used for human cells were designed based on the following rationale. The fibroblasts required two ~9kb overlapping fragments in order to encompass the entire mtDNA genome. The cybrid primers were designed to specifically only amplify either mutant mtDNA (i.e. molecules that would contain the deletion from m.7982-15504), or wild type DNA (see Figure 3.9). A size specific extension time for mutant mtDNA was performed to ensure that only the mutant mtDNA was amplified using the mutant primers as the wild-type mtDNA would require a longer extension time for successful PCR of the mutant amplicon primers. All primers were screened using primer-BLAST to ensure there was no amplification of nuclear homologous regions.

The mouse primers were designed using the reference mouse mtDNA genome (NC_005089.1). Three overlapping ~6kb fragments were designed to amplify the entire mtDNA genome, and screened for absence of Numt (nuclear mitochondrial homologues) amplification using primer-BLAST (J. Ye *et al.*, 2012).

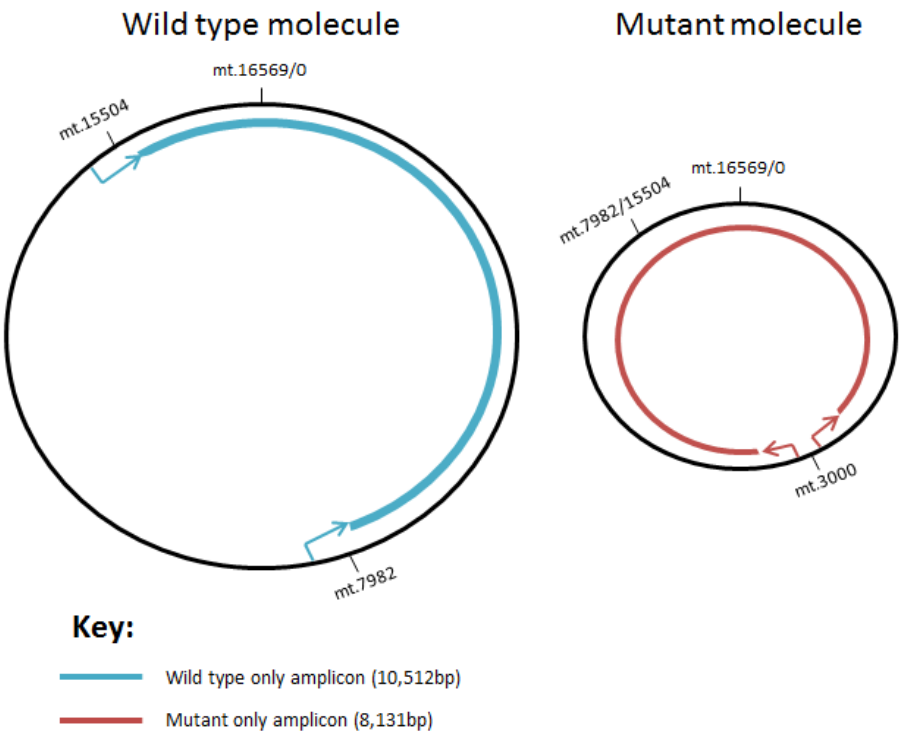

Figure 3.9 Primer design rationale in the amplification of mutant and wild type specific bands of trans-mitochondrial cybrid cells exposed to NRTIs, sequecned using the Illumina MiSeq™ platform.

Name	Forward sequence 5' – 3'	Reverse Sequence 5' – 3'	T_m (^oC)
Mouse			
Fragment A	GCCAATGAAATGGGAAGAAA	CGATGTCAAGGGATGAGTTG	57
Fragment B	GCAACCCTACACGGAGGTAA	TGATGGTTTGGGAGATTGGT	57
Fragment C	TCCTACTGGTCCGATTCCAC	TAGAAACCGACCTGGATTGC	57
Human			
Cybrid Deletion	TCGCCCTATTCTTCATAGCC	GATTACTCCGGTCTGAACTC	61
Cybrid Wild type	ATTCATCGACCTCCCCACC	GTATAAGAGATCAGGTTCGTC	61
Fibroblast A	CCCTCTCTCCTACTCCTG	CAGGTGGTCAAGTATTTATGG	68
Fibroblast B	CATCTTGCCCTTCATTATTGC	GGCAGGATAGTTCAGACGG	68

Table 3.7 Primer properties for NGS experiments on human (trans-mitochondrial cybrids and fibroblasts) and mouse mtDNA designed using the mitochondrial reference sequences for the mouse and human (NC_005089.1 and NC_012920, respectively).

The PCR was performed using the PrimeSTAR GXL polymerase (Takara, Clontech) that has been specifically designed for NGS amplicon generation (Takara). Each fragment was amplified using a Veriti® thermocycler (Applied Biosystems, Life-Technologies, Paisley, UK); the fragments were amplified in 25µL reactions using the following: 1x GXL buffer (Takara, Clontech), 2.5mM dNTPs, 300nM forward primer, 300nM reverse primer, 1M Betaine (mouse fragments only), autoclaved PCR-grade deionised water and 1.25U GXL *taq* polymerase. The PCR cycling conditions can be found in Table 3.8.

Primers	Denaturation x1	Denaturation x30	Annealing x30	Extension x30	Extension x1
Mouse					
Fragment A	98°C (30s)	98°C (30s)	57 °C (30s)	68 °C (8.5 mins)	72°C (10 mins)
Fragment B	98°C (30s)	98°C (30s)	57 °C (30s)	68 °C (10.5 mins)	72°C (10 mins)
Fragment C	98°C (30s)	98°C (30s)	57 °C (30s)	68 °C (8.5 mins)	72°C (10 mins)
Human					
Cybrid Deletion	98°C (30s)	98°C (30s)	61 °C (30s)	68 °C (8 mins)	72°C (12 mins)
Cybrid Wild type	98°C (30s)	98°C (30s)	61 °C (30s)	68 °C (8 mins)	72°C (12 mins)
Fibroblast A	95°C (5 min)	98°C (10s)	68 °C (15 mins)		72°C (10 mins)
Fibroblast B	95°C (5 mins)	98°C (10s)	68 °C (15 mins)		72°C (10 mins)

Table 3.8 The PCR cycling conditions for fragment amplification using the GXL *taq* polymerase (Takara, Clontech). The temperature and cycling time (in brackets) is shown.

All fragments were electrophoresed on a 1% agarose gel as previously described (see Section 3.5) using 5µL of product. To ensure that the reaction was contamination free and only specific product of the correct size was amplified, 5µL of the DNA ladder 1kb plus (Thermo-Scientific, Hampshire, UK) was loaded and the PCR product size cross referenced using this.

3.10.2 Sample cleaning and pooling

Once the fragments for each sample had been amplified and checked, the DNA was then purified to remove any excess dNTPs, primers, primer-dimers and salts, in excess from the PCR reaction. To do this, the samples were purified using a bead based system (Agencourt AMPure XP; Beckman Coulter, High Wycombe, UK) implementing a magnetic plate to facilitate the process.

The purification process was carried out as follows: the Agencourt magnetic bead solution were brought to room temperature and gently agitated to fully re-suspend the mixture. Next, 1.8x PCR reaction volume of beads was added to each reaction on a 96-well PCR plate. The PCR reaction and bead reagent was mixed thoroughly to generate a homogenate solution by pipette mixing 10-12 times. The mixture was incubated at room temperature for 5 mins before placing on a 96-well super magnetic plate (Agencourt SPRIPlate; Beckman Coulter, High Wycombe, UK) for 2 minutes to separate the beads from the solution. During the incubation period, the beads will bind to the DNA fragments and therefore, separation of beads from solution will extract and purify the DNA. Next, the clear solution was aspirated leaving only a brown ring of beads in each well. These were then washed twice using 200µL of 70% ethanol (Sigma-Aldrich, Dorset, UK), incubating between washes for 30 seconds. All washing steps were performed on the magnetic plate. To ensure that all ethanol had been fully removed the plate was then air dried at room temperature for 5 minutes. The magnetic beads were then re-suspended in 30µL elution buffer (ultra-pure PCR grade deionised water) away from the magnetic plate. The plate was sealed using universal 96-well plate sealers (Thermo scientific, Hampshire, UK) and gently agitated to ensure full re-suspension of the beads in the water. Finally, the beads were placed back onto the magnetic plate for 5 minutes and the elutant transferred to a new plate.

The DNA concentration of the elutant was quantified using a Qubit® 2.0 Fluorometer (Life-Technologies, Paisley, UK). Fragments were then pooled into one well in equimolar quantities (all fragments per sample) and the final concentration of the sample DNA was adjusted to 2ng/µL, ready for fragment re-sequencing preparation.

3.10.3 Fragment re-sequencing

Fragment re-sequencing preparation for the MiSeq™ (Illumina, Cambridge, UK) was performed using the newest paired-end chemistry version at time of running (v3). The sample amplicons were diluted to 0.2ng/µL and 5µL (1ng total) used in the Nextera XT (Illumina, Cambridge, UK) tagmentation process. The sequencing primers and DNA barcodes were incorporated into the samples and a bead clean up performed, as previously described (see Section 3.10.2) with the bead to sample ratio changed from 1.8:1 to 0.5:1. The library concentration was standardized, diluted and loaded onto the MiSeq™ along with a 15% positive spike-in sequencing control (PhiX; Illumina, Cambridge, UK). Targeted re-sequencing was then initiated on the MiSeq™ machine.

3.11 Bioinformatics

Next generation sequencing (NGS) data was analysed by an optimised 'in-house' pipeline which was run using the UNIX operating platform and the command line scripting language associated with 'common' bioinformatics approaches.

Data generated from a next generation sequencing platform is presented, in its most raw form, as a .fasta and/or .fastq (fasta quality file) file(s). When using a sequencing platform that implements a colour based detection system, such as the SOLiD (Applied Biosystems, Life-Technologies, Paisley, UK) then the raw data is presented in a specific fasta and fastq known as .csfasta or .csfastq (colour space fasta) which must be handled with such consideration in mind (see Section 7.4.1).

The general principle of an NGS data analysis pipeline (see Figure 3.10) consists of the initial alignment of the raw data files to a reference/consensus sequence (usually the rCRS for mitochondrial sequencing). Once alignment has been performed, the data needs to be sorted to ensure only sequencing reads that have been fully mapped to the genome are kept and then processed through a variant calling tool. Generally, only SNPs and small deletions (≤50bp) will be called with variant callers such as VarScan (Koboldt *et al.*, 2009; Payne *et al.*, 2014). The final step in analysis is variant annotation and filtering.

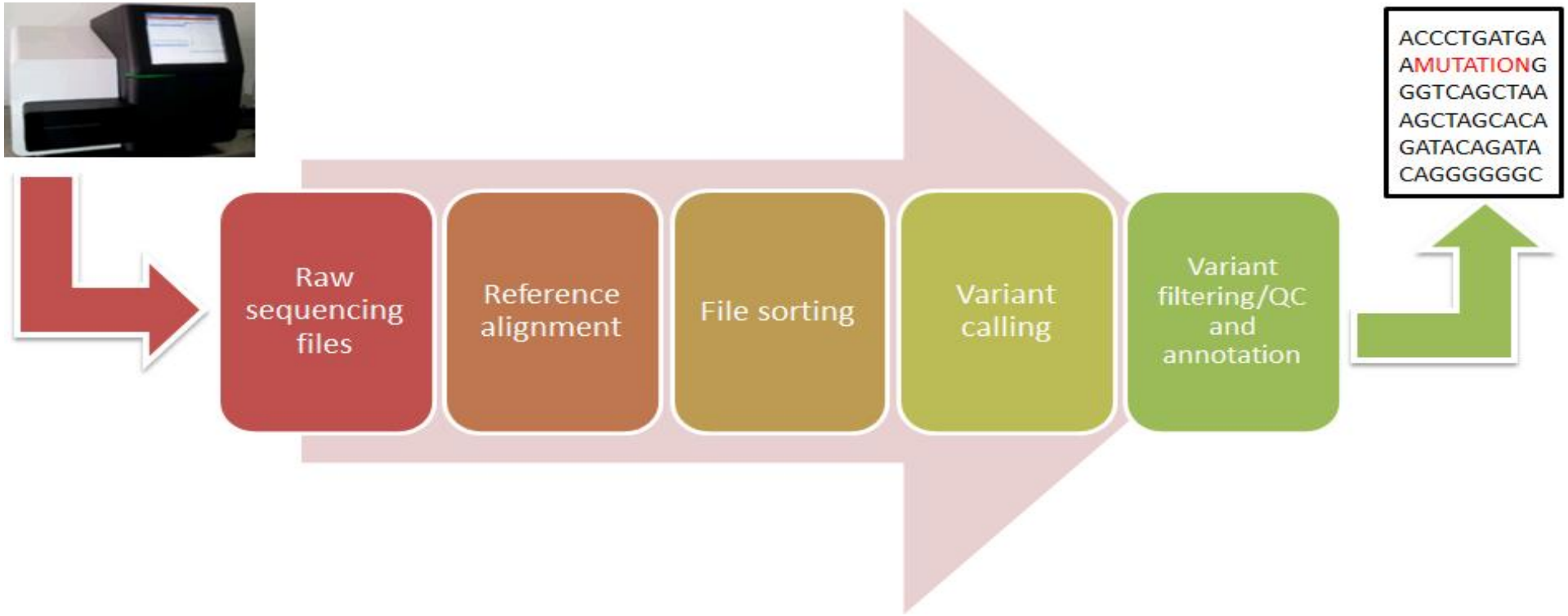

Figure 3.10 Data flow diagram illustrating the main steps in an NGS data analysis pipeline. Briefly, the raw sequencing files are aligned to the reference genome, sorted/manipulated and then passed through a variant calling tool. The variant list is then filtering through QC steps and annotated.

3.11.1 NGS analysis pipeline

All NGS data analysis was run on a UNIX based 'grid' computer cluster network system (Lampredi2; IGM, Newcastle University, UK). The analysis was broken down into three stages: primary, secondary and tertiary, which were pieced together into one analysis 'script' that processed a single sample from raw sequencing data to annotated files that were read into Excel Spreadsheet (Microsoft, Reading, UK). The following method was used for the SOLiD™ (Applied Biosystems, Life-Technologies, Paisley, UK) and the MiSeq™ (Illumina, Cambridge, UK) platforms' data analysis.

The primary analysis stage consisted of aligning the data to the reference mitochondrial only sequence file; human, rCRS (NC_012920); mouse, Mus mitochondrial sequence (NC_005089.1). The alignment tool used was BWA v0.7.4 (H. Li and Durbin, 2009) implementing the 'MEM' (maximum exact matches) algorithm for optimised alignment of short sequencing reads to produce a '.SAM' (sequence alignment map) output file. MEM alignment performs a sequence alignment based on similarity matches (as opposed to a global alignment which forces alignment based on a best match principle to cover the reference genome). Prior to alignment, the reference file was 'indexed' using BWA to prepare the file for processing with BWA.

The secondary stage of analysis consisted of sorting and preparing the data files for the variant calling tool. To do this, the SAM files were initially converted in BAM (binary alignment map file) using SAMtools v0.1.8 (H. Li *et al.*, 2009). To do this the 'import' function was used followed by the 'sort' function to organise the reads into a coherent order. The sorted bam file was then indexed to allow faster read access to the file information by software using the SAMtools 'index' function. Finally, the read duplicates were removed, or 'marked' in the data files (for background, see 1.6.5) using the Picard v1.85 tool (http://picard.sourceforge.net) and the 'remove duplicates' function, creating a log file and noting that the files were sorted.

The final or tertiary step of data analysis was to call and annotate the predicted variants. This consisted of the following steps; VarScan v2.3.1 (Koboldt *et al.*, 2009), was used to call variants; ANNOVAR (K. Wang *et al.*, 2010), was used to annotate the exonic variants. Before VarScan was used to call potential variants, the sorted bam files were converted into a 'pileup' file, which describes base-pair information to facilitate variant calling and thus making the bam files 'interpretable' to VarScan; The SAMtools 'mpileup' function was implemented to do this. Each pileup file was read into VarScan

using the 'pileup2snp' function to call the variants, creating a tab delimited VCF (variant call format) output file. The parameters used for calling variants are shown in Table 3.9.

Species	Minimum Coverage	Phred	Variant Frequency (%)	Minimum supporting reads
SOLiD™				
Human	1500	30	0.1	10
MiSeq™				
Human Cybrids	5000	30	0.3	10
Human Fibroblasts	3000	30	0.3	8
Mouse	3000	30	0.3	8

Table 3.9 Variant calling parameters used in the VarScan tool for the SOLiD™ and MiSeq™ platforms data analysis. Minimum coverage – sequecning depth; Phred – sequencing quality score (see Section 1.6.2); variant frequency – variant level within the sample; minimum supporting reads – the number of reads required to make a variant call.

The variants called were annotated using ANNOVAR (most recent version) as per the user manual; however, firstly, a fasta file and text file containing the mitochondrial coding gene sequences and the relevant transcripts was made for human and mouse mitochondria using the *Ensembl* (Flicek *et al.*, 2014) gene database (instructions for creating files are found in ANNOVAR user manual).

The files generated by ANNOVAR were converted into a tab delimited file, rather than space separated, and read into excel for viewing, filtering or graphing.

3.12 Statistical analysis

For statistical analyses, either SPSS statistics v19 (IBM) or, GraphPad™ Prism v5 (GraphPad™ Software, CA, USA) was used and the statistical test used in the analyses are specified in each results chapter.

Two-sided *p*-values expressed as ≤ 0.05, were deemed as significant.

Chapter 4. Assessment of Mitochondrial DNA Damage in HIV-infected Individuals

Table of Contents

4.1 Background

Patients receiving highly active antiretroviral therapy (HAART) to combat HIV infection, especially those who have been/are exposed to the older nucleoside analogue reverse transcriptase inhibitors such as zalcitabine (ddC), didanosine (ddI), stavudine (d4T) and zidovudine (AZT), often present with pathologies associated with mitochondrial toxicity. This apparent toxicologic effect of NRTIs has been described as resulting from the affinity of NRTIs for the mitochondrial DNA polymerase γ, resulting in mitochondrial DNA replication inhibition (Höschele, 2006). The newer NRTIs such as tenofovir (TDF) and abacavir (ABC) have not been associated with such pathologies, arguably due to their lack of affinity for polymerase γ. The sub-group of older NRTIs, known as the 'd-drugs' (dideoxynucleoside analogues, ddC, ddI, d4T), have been found to have the highest affinity for polymerase γ.

The histochemical analysis of skeletal muscle samples from patients treated with the pol γ NRTIs indicates that some patients have a mosaic mitochondrially encoded COX (cytochrome c oxidase) deficiency (Birkus *et al.*, 2002).

There is a strong link between mitochondrial DNA deletions and the process of ageing (Khrapko *et al.*, 1999; Z. Cao *et al.*, 2001; Herbst *et al.*, 2007). For example, the mitochondrial DNA 'common deletion' mutation (a 4977bp deletion spanning the 'major arc' of the genome) accumulates with age (Y. F. Zhang, 2007; Gendron *et al.*, 2012; P. Wang *et al.*, 2013).

It has also been suggested that mitochondrial DNA point mutations may contribute towards (or at the very least accumulate) with the process of ageing. There has been an identification of specific 'hot spots' within the mtDNA control region where point mutations, such as m.414T>G may accumulate in post-mitotic tissue such as skeletal muscle and brain during human ageing (Michikawa *et al.*, 1999; Del Bo *et al.*, 2003). The implications of such mutations have been debated, with some claims of them being neutral (Seibel *et al.*, 2008) whereas other authors cite an impact on mtDNA replication, and their presence within a number of 'pre-cancerous' cells (Z. Liu *et al.*, 2002; Kassem *et al.*, 2011).

4.2 Experimental Aims

The aim of this study was to determine the effects of HIV status, and NRTI treatment exposure on the mitochondrial DNA control region point mutation m.414T>G, and the mitochondrial DNA δ4977 common deletion mutation.

4.3 Experimental design and methods

4.3.1 Patient cohort

All individuals gave informed consent and was formally approved by the local research ethics committee. Lower limb skeletal muscle samples from HIV-infected subjects had been previously obtained by percutaneous biopsy. Samples from HIV-uninfected control subjects had been previously obtained at the time of elective lower limb surgery, and submitted to the local Mitochondrial Control Tissue Bank (MCTB). A total of 26 HIV-infected/NRTI exposed individuals were used along with 12 HIV-infected/NRTI naïve and nine MCTB controls. DNA was previously extracted from these samples (Nucleon DNA extraction kit, Tepnel life science, Manchester, UK) and stored at -80°C.

4.3.2 Pyrosequencing

In order to investigate the heteroplasmy level of the control region (D-loop), m.414T>G mutation, pyrosequencing was performed as previously described (see Section 3.9.2). Technical replicates were taken in duplicate and the average reading reported. Each sequencing run included three controls: a 100% mutant, 100% wild type and 50:50 mut: wild type, as described in Section 3.9.3.

4.3.3 Common deletion qPCR quantification assay

The proportional level of mitochondrial DNA m.δ4977bp common deletion (m.8470 - m.13447) was quantified using a real-time qPCR assay as previously described (Payne *et al.*, 2011). The level of common deletion (CD) per mitochondrial DNA copy was determined by comparison with the copy number of *MT-ND1*. The result was expressed on a logarithmical scale to improve resolution of the lower level of CD detection.

The qPCR CD standards were generated (see Table 4.1) using a sample of DNA from a patient with KSS (Kearns Sayre syndrome) with a known level of the specific δ4977 mitochondrial DNA common deletion mutation. The standards were processed as previously described (see Section 3.6). The standards were diluted serially from 10^8 - 10^2 copies per cell and amplified in triplicate on the qPCR run, with an assessment of the amplification efficiency made for both reactions.

The CD-*MT-ND1* qPCR assay, perform on the Bio-Rad CFX system in skirted deep-well, white bottom plates (Bio-Rad), and was performed in 10µL reactions consisting of: 1x Evagreen supermix (Bio-Rad), 625nM primers (forward and reverse; Table 4.1) and

approximately 50ng DNA per sample, with each amplified in triplicate, using a positive control with a known common deletion present. The cycling conditions comprised: 98^{o}C initial denaturation for 2 minutes, followed by 40 cycles at 98^{o}C for 5 seconds and annealing at 60^{o}C for 20 seconds. A melt curve was taken at the end of analysis.

Name	Forward 3'-5'	Reverse 3'-5'	T_m (^oC)	Product size (bp)
CD qPCR standards				
CD standard	TCCTAACACTCACAACAAAAC	GTTAGGTAGTTGAGGTCTAGG	58	1158
CD qPCR amplification primers				
CD qPCR	CCCACCATAATTACCCCCATAC	GGAGTAGAAACCTGTGAGGAAAGG	60	115

Table 4.1 The qPCR template standards and qPCR primers used in the m.δ4977 mitochondrial common deletion quantification qPCR assay.

4.3.4 Large scale deletion determination

To investigate the presence and number of large scale deletion mutations in the samples, a long range PCR was used to screen for mtDNA deletions as previously described (see Section 3.7).

The PCR products were electrophoresed on a 1% agarose gel, as previously described (see Section 3.5) using the DNA ladder 1kb plus (Thermo-Scientific, Hampshire, UK) to determine the number of bands present.

4.3.5 Statistical analyses

The following tests were performed using SPSS: non-parametric correlations (spearman's rank and Mann-Whitney), univariate linear regression and student *t*-tests.

4.4 Results

4.4.1 m.414T>G quantification

The mitochondrial D-loop point mutation, m.414T>G, was quantified using pyrosequencing and the heteroplasmy level has been graphed and expressed as mutant percentage level for all three groups of subjects (Figure 4.1). The lower detection limit of the pyrosequencing assay was deemed to be 2%, based upon my previous optimisation assay, as described in Section 3.9.3. All patient data can be found in Appendix A.

The most striking feature of the heteroplasmy graph is the large distribution of the mutation level in the HIV^+/NRTI-exposed condition compared to the other two groups. The highest heteroplasmy limit was found to be 9% in this group. Approximately half of subjects (14 out of 26) in the HIV+/NRTI-exposed group showed heteroplasmy levels clustered at the 2% level, consistent with no detectable m.414T>G mutation. The remaining subjects (12 of 26) showed increased levels which ranged from 2.5% up to the 9%.

The HIV^+/NRTI-naïve group showed a heteroplasmy level across all subjects of 2%, representing the lower level of resolution of the assay, and therefore may be considered to show no detectable m.414T>G heteroplasmy. The MCTB control samples also showed a similar pattern with eight of the nine subjects presenting with the lower limit of 2% and one sample contained 3% heteroplasmy.

The distribution of the heteroplasmy level in the HIV^+/NRTI-exposed group was assessed for correlations with previously collected, anonymised, clinical treatment data. Highly significant, positive correlations were found with the heteroplasmy level and the length of NRTI treatment ($p<0.001$) across the NRTI-exposed group. The m.414T>G level was also found to correlate with exposure to pol γ inhibiting NRTI's, ddC, ddI and d4T ($p=0.02$) with a Mann-Whitney test also indicates significance ($p=0.03$) compared to the untreated group.

In order to determine the effects of the specific NRTIs, univariate linear regression within the HIV^+/NRTI-exposed group was performed to investigate the above correlation of m.414T>G heteroplasmy level and the pol γ inhibiting NRTIs. Prior history of exposure to each of these NRTIs (AZT, ddI, ddC, d4T) was found to be

statistically significantly associated with m.414T>G heteroplasmy level, with *p*-values of 0.022, 0.006, 0.006 and 0.006, respectively.

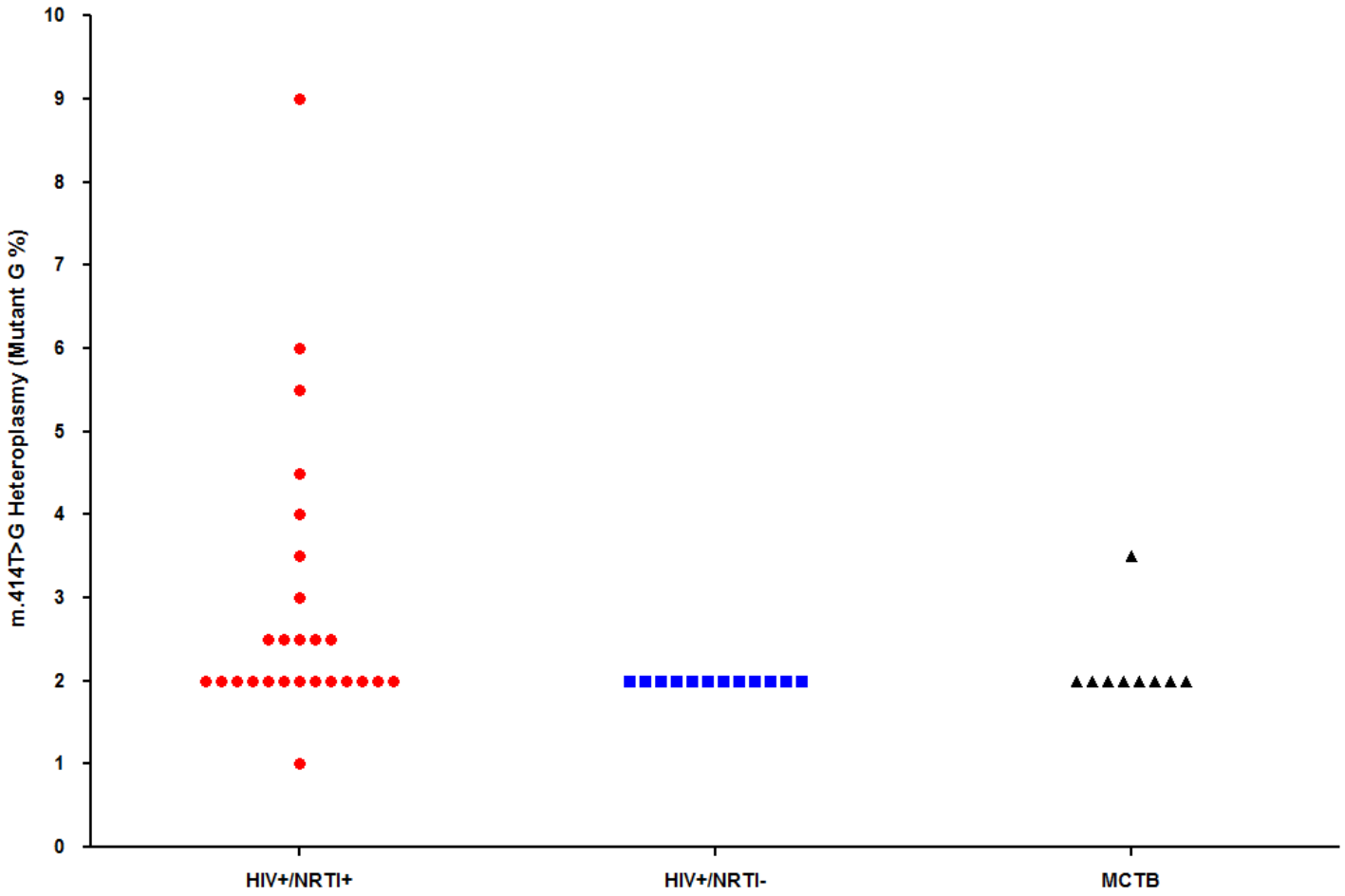

Figure 4.1 The heteroplasmy level of m.414T>G, quantified by the use of pyrosequencing in the muscle samples of lower limb in HIV$^+$/NRTI-exposed individuals, HIV$^+$/NRTI-naïve individuals and control muscle samples (MCTB). The lower detection limit of the assay was determined to be 2%. Lines indicate group average.

4.4.2 Quantitation of large scale deletions

The screening of large scale deletions in the samples using long range PCR (see Section 4.3.4) indicated that there was a range of deletions present (Figure 4.2).

In the HIV^+/NRTI-exposed individuals and 12 out of 26 were found to have deletions present. This represents 46% of samples containing a deletion, with eight of the 12 showing multiple bands/deletions from the LR-PCR, representing 31% of all HIV^+/NRTI-exposed samples.

Within the $HIV^{+/}$NRTI-naïve group, three out of the 12 samples were found to contain deletions. This represents 25% of samples containing a deletion and of these; two were found to contain multiple bands/deletions, representing 17% of the all samples. In the untreated, MCTB control samples, there were no deletions present in any of the nine samples.

The deletion data was categorised into three groups: no deletions, one deletion, and multiple deletions. This was then correlated (non-parametric) with exposure to pol γ inhibiting NRTIs and duration of treatment. The associations were found to be statistically significant in both cases, with more deletions found in those treated with pol γ inhibiting NRTIs (p=0.036), and in those with greater length of treatment exposure (p=0.028). Individuals exposed to d-drugs were found to have more deletions (p=0.009) than those never exposed. Age did not appear to be a contributor to number of deletions (p=0.624).

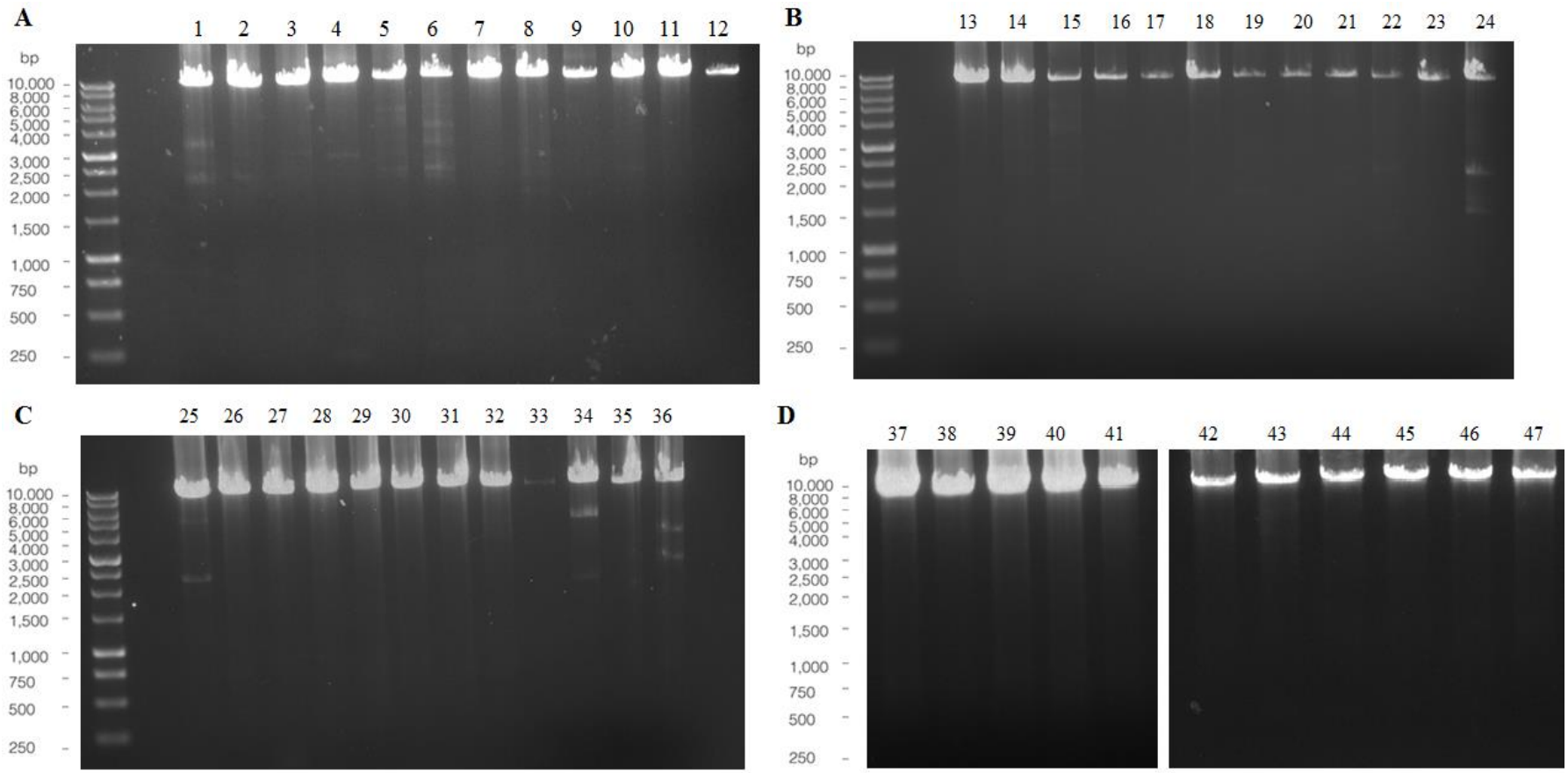

Figure 4.2 Four panel image illustrating the long range PCR gel images of large scale deletions. A) 1 – 12, HIV$^+$/NRTI-exposed samples; sample 3, HIV$^+$/NRTI-naïve B) 13 – 24, HIV$^+$/NRTI-exposed samples C) samples 25 – 27, HIV$^+$/NRTI-exposed samples; 28 – 36, HIV$^+$/NRTI-naïve samples D) comprised of two gel images – 37 – 38; HIV$^+$/NRTI-naive samples, 39 – 41 MCTB; 42 – 47, remaining MCTB samples.

4.4.3 Common deletion quantification

The level of the δ4977 mitochondrial DNA common deletion mutation (CD) was quantified using qPCR, expressed as CD per mtDNA copy and presented logarithmically (Figure 4.3). One HIV^+/NRTI-exposed sample was excluded from the analysis due to a lack of adequate DNA.

The HIV^+/NRTI-exposed subjects show a statistically higher CD level compared to the HIV^+/NRTI-naïve individuals using a two-tailed *t*-test (p=0.024). The mean CD level was found to be -4.39 $\log_{10}$ (CD)/mtDNA (±S.D 0.9), whereas the mean CD in the HIV^+/NRTI-naïve subjects was -5.16 $\log_{10}$ CD/mtDNA (±S.D 0.93).

The level of CD was found to correlate (non-parametric) with months of treatment (p=0.04) and was specifically associated with exposure to the d-drugs (ddI, d4T and ddC) compared with no d-drugs exposure (p=0.01). The cumulative months of exposure to d-drugs (one or more) was also found to be statistically significantly associated with CD level (p=0.015). The CD level was found to be statistically significant when compared to the total (mitochondrial toxic + mitochondrial safe) cumulative treatment duration (p=0.009).In addition to these correlations, there was also evidence of the level of CD correlating with age of the individual (p=0.034) and the percentage of COX-deficient cells in individuals (p=0.013).

The CD level in the MCTB control individuals was found to have range of data -3.59 to -5.60 $\log_{10}$ (CD)/mtDNA, with a mean level of -4.34 $\log_{10}$ (CD)/mtDNA (±S.D 0.69). The difference between MCTB and HIV^+/NRTI-naïve was found to be not significant (p=0.05).

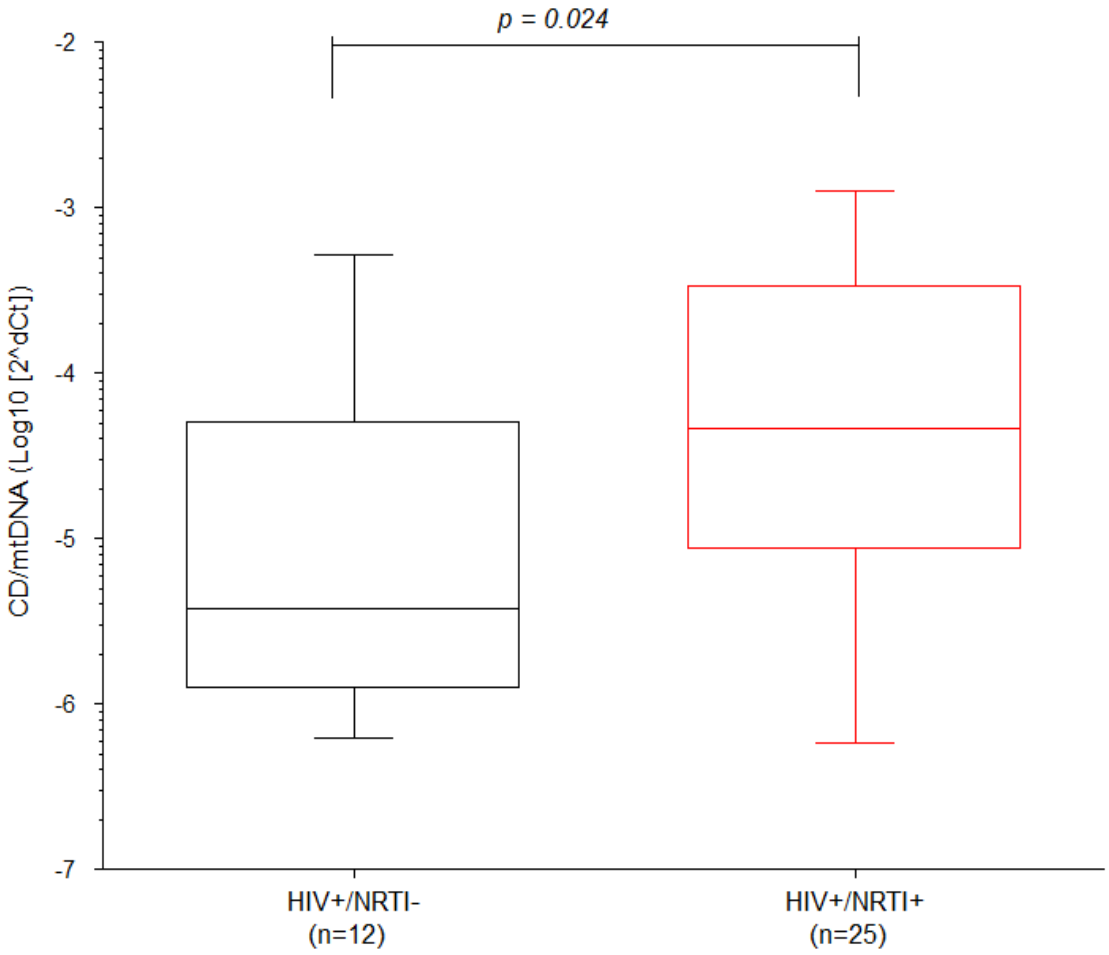

Figure 4.3 The quantification of mitochondrial DNA common deletion, expressed as $\log_{10}$ CD/mtDNA. The black coloured box and whisker plot represents the HIV/NRTI-naïve individuals and the red represents the HIV/NRTI-exposed individuals which was found to be significantly higher (p=0.024) by a two-tailed t-test.

4.5 Discussion

The aim of this study was to assess the levels of the mitochondrial DNA control region m.414T>G point mutation, the common deletion mutation, and the number of large scale deletion mutations present in skeletal muscle DNA samples from control/healthy individuals, HIV^+/NRTI-naïve individuals and HIV^+/NRTI-exposed individuals.

4.5.1 m.414 T>G data summary

The presence of the point mutation m.414T>G (see Figure 4.1) was evident in only the HIV^+/NRTI-exposed individuals. NRTI-naïve and MCTB control individuals showed no evidence of detectable mutation (within the 2% heteroplasmy limits of resolution of the assay), except in one control sample with a 3% heteroplasmy level detected.

The level of m.414T>G has been previously associated with age in skeletal muscle tissue (Del Bo *et al.*, 2003). Although there were no aged individuals included in this study, the average age amongst the HIV^+/NRTI-naïve and HIV^+/NRTI-exposed groups were similar (35.7 and 40 years old, respectively). Therefore it is not surprising that there was no signal that point mutation level was associated with age in this study. However these data do suggest that HIV^+/NRTI-exposed individuals present with an 'aged' m.414 genotype, plausibly suggestive of an accelerated intrinsic ageing effect.

The evidence that the level of m.414T>G increases through NRTI exposure was further supported by strong associations with the NRTI treatment duration ($p<0.001$), and the specific correlations with exposure to the 'old' (pol γ inhibiting) NRTIs (AZT, d4T, ddI and ddC; p-values = 0.022, 0.006, 0.006, 0.006). None of the HIV^+/NRTI-exposed individuals who had received only new NRTIs, displayed high levels of m.414T>G; all levels were comparable to the MCTB individuals with the highest of 3%, suggesting that HIV itself does not cause an increase in the m.414T>G point mutation.

4.5.2 Mitochondrial deletion data summary

The level of CD in the MCTB controls was not significantly different from the HIV^+/NRTI-naïve individuals ($p=0.05$). This suggests that the HIV infection itself does not contribute towards an increased CD level, with the mean age of the MCTB group and HIV^+/NRTI-naïve group being highly similar, 32 and 35 years old, respectively. There were no large scale deletion mutations detected by LR-PCR in the MCTB samples, in-keeping with the lower CD level.

In the HIV^+ groups, there was a significant increase in CD level in the NRTI-exposed individuals compared to the naïve individuals (p=0.024) (see Figure 4.3).

The previously described (Maagaard *et al.*, 2006) notion of an increased deletion level in NRTI-exposed individuals with COX-deficient cells, which was also found to statistically correlate in this dataset, with a higher CD level in those with a higher percentage of COX- cells (p=0.013). The measurement of deletions by either LR-PCR or CD-assay indicates that any increase in deletions may be due to either clonal expansion within the COX-deficient cells or an increase at a lower level across many cells. However, the percentage of total COX-deficient cells gives an indication as to how much clonal expansion is occurring in a sample.

The association of multiple mtDNA deletions with exposure to d-drugs, such as didanosine was also found (p=0.036) and in a time dependent manner linking d-drug exposure time with the accumulation of one or more deletions (p=0.026). As the clinical data was based on lifetime exposure and not current (i.e. none of the subjects were receiving d-drugs at time of biopsy), the effects from the d-drugs are likely to be not reversible suggesting they have caused irreversible damage.

There was no age correlation found with the number of deletions (p=0.624), although there was with the level of CD (p=0.034). The accumulation of CD over time has been correlated with age in a number of different tissues (Y. F. Zhang, 2007; Gendron *et al.*, 2012; P. Wang *et al.*, 2013). The data presented here supports this notion and suggests that due to the higher CD level present in the HIV^+/NRTI-exposed individuals, there may be an accelerated accumulation of the CD that is irrespective of age and is due to the time dependent exposure to NRTIs (p=0.01). The length of exposure time ranged from 29 months to 193 months, with the individual exposed to NRTIs for 193 months also presenting with the highest percentage of COX-deficient cells (4.5%) and the second highest CD level at -2.91 $\log_{10}$ (CD)/mtDNA; the individual exposed for the least number of months was found to have the second lowest CD level of -6.03 $\log_{10}$ (CD)/mtDNA. Overall these data therefore show a strong dependence on duration of NRTI exposure.

4.5.3 Study limitations

There are three limitations to this study that may affect the understanding of the data. These are:

- The mtDNA CD measurement is not fully representative of mtDNA deletions. The CD does not represent the total mtDNA deletion population and has been recently shown to differ in levels between tissues and individuals, especially in highly energetic tissues such as muscle and brain (Phillips *et al.*, 2013). Although LR-PCR was used for screening for other deletions, it is only a semi-quantitative technique.
- The measurement of one point mutation (m.414T>G) is not representative of the entire mtDNA genome. The behaviour of the mutation may also depend upon the underlying haplogroup. This point mutation was chosen for two reasons: the correlation of the m.414T>G mutation with age and the fact it is known to be a neutral mutation.
- Although there is an expansion of the m.414T>G present in some samples, this does not indicate whether it is an event present in only select cells, such as COX-deficient cells, or whether it is across the board in many cells as a low level.

4.6 Chapter conclusion

These data demonstrate an increase in mtDNA point mutations (m.414T>G) and large-scale deletion mutations (including the δ4977 common deletion; CD) in association with NRTI exposure. These effects are principally seen in those patients exposed to the older, pol γ inhibiting NRTIs.

These NRTI-mediated effects are comparable with the changes expected much later in life in association with normal ageing.

.

Chapter 5. The *in vitro* effects of NRTIs on the Behaviour of Mitochondrial DNA Deletion Mutations in Mitotic and Post-mitotic Models

Table of Contents

5.1 Background

Large scale mitochondrial DNA deletions are an important pathological feature of mitochondrial disease (D. C. Wallace, 1992; Schaefer *et al.*, 2008) and there is a large amount of evidence to suggest that they are involved in the process of intrinsic ageing, with large scale deletions accumulating with age seen among a variety of mammalian species (Khrapko *et al.*, 1999; Z. Cao *et al.*, 2001; Herbst *et al.*, 2007). Such studies elegantly demonstrate that deletion accumulation in ageing occurs through generation of a 'clonal' population that is dominated by a single large scale deletion within an individual cell.

The cellular mechanism by which defective mitochondria accumulate within cells is currently unknown; however, there are a number of theories with varying evidences to suggest how this occurs. The earliest suggestion comes under the umbrella term 'vicious cycle theory' (N. Arnheim and Cortopassi, 1992a). This theory postulates that typical mutations cause the host mitochondria to generate an increased amount of ROS (see Section 1.1.3), thereby promoting or inducing further mutations in an accelerate fashion (de Grey, 2005). This theory is flawed by the lack of multiple deletions present in ageing and diseased patients, which are populated with purely one single mutant form. The theory does however, give support to the idea that mutations arise at the same point in the mitochondrial genome, which has been discussed regarding the timing of mutations in ageing (Khrapko, 2011).

The two most commonly quoted theories to explain clonal expansion are random drift and a replicative advantage mechanism. The former has been neatly illustrated via computer simulations that the process of deletion accumulation in ageing humans (Chinnery and Samuels, 1999; Elson *et al.*, 2001), using previously demonstrated mitochondrial replicative dynamics (Shadel and Clayton, 1997), can occur purely through random drift alone. These studies mimic mutation accumulation observations seen in patients with late onset mitochondrial disorders; however, more recent simulations suggest that random drift would only explain clonal expansion in long lived mammals, such as humans, and not for short lived species such as rodents (Kowald and Kirkwood, 2013).

Clonal expansion by a replicative advantage mechanism has been supported by a number of studies (D. C. Wallace, 1992; Z. Cao *et al.*, 2001) which arises from the idea

that a deletion mutation creates a reduction in the size of the genome and therefore a reduction in the replication time. Whilst this theory would make logical sense, it has been criticised on the basis that the time required for replication of mtDNA is in the order of 75 minutes (Clayton, 1982), which isn't sufficient to cause a replicative difference based on the fact the half-life of a mtDNA is approximately 8-23 days (Korr *et al.*, 1998).

The case for clonal expansion of mutations isn't restricted to purely mitochondrial disease patients and aged individuals; there's increasing evidence that HIV-infected individuals receiving NRTIs also have an increased mtDNA mutation burden and the levels previously found are comparable to that of ageing individuals but it is unknown how they arise (Payne *et al.*, 2011).

5.2 Experimental aims

The aims of the study were to determine:

- The relative effects of NRTI-exposure on the replication of wild-type and deleted mtDNA.
- Whether NRTI exposure leads to new deletion formation through *de novo* mutagenesis.
- Whether there is a dose response effect of NRTIs on deleted mtDNA?
- Whether there are differences between mtDNA behaviour in mitotic and post-mitotic cells in terms of mtDNA wild-type and mutant copy number

5.3 Experimental design and methods

5.3.1 Trans-mitochondrial cybrid cell culture

Trans-mitochondrial cybrid cell lines were cultured and propagated as previously described (see Section 3.1.1 and 3.1.3); cells were grown in T_{25} flasks (Greiner Bio-one, Stonehouse, UK) and split when reaching 70-80% monolayer confluency with sufficient media available to allow for cells to remain in the exponential growth phase. All cells were checked before freezing and routinely throughout the experiment for the presence of mycoplasma infection (see Section 3.1.7).

5.3.2 Rat cerebellar neuron harvesting and propagation

The pregnant WISTAR rat was euthanized by administration of carbon dioxide. The E18 (embryonic developmental day 18) embryos were removed and decapitated, transferring the heads to a small petri dish (VWR, Leicestershire, UK) containing Hanks Balanced Salt Solution (HBSS; Sigma-Aldrich, Dorset, UK).

The cerebellum was separated from the cortices and placed into another small petri dish containing HBSS. The cerebellum meninges were carefully removed and then the cerebellum finely chopped using a round end scalpel in HBSS solution, before centrifugation for 3 minutes at 1,300rpm. The supernatant was removed and 5ml of Accutase (Gibco, Life-Technologies, Paisley, UK) added and incubated for 20 minutes at 37^{o}C, with regular agitation. Prepared DNase (Sigma-Aldrich, Dorset, UK) as per manufacturer's guidelines and 500µL was added to the suspension, and then centrifuged at 1,300rpm for 3 minutes. The supernatant was removed and 3ml FCS (Sigma-Aldrich, Dorset, UK) added. The suspension was then triturated using a 19G 1ml needle, followed by centrifugation at 1,600rpm for 3 minutes. The supernatant was removed and the cell pellet re-suspended in 3ml of the DMEM culture medium containing: 1% penicillin/streptomycin (Sigma-Aldrich, Dorset, UK) and 2% B27 (Gibco, Life-Technologies, Paisley, UK).

A cell count was performed using a haemocytometer and trypan blue as previously described (see Section 3.1.6), seeding approximately 2,000,000 cells per well in a 6 well plate (Greiner Bio-one, Stonehouse, UK) which had been treated with Poly-L-Lysine (Sigma-Aldrich, Dorset, UK). Poly-L-Lysine treatment consisted of adding 2ml solution to each well for 20 minutes, incubated at 37^{o}C. The cells were supplemented with 2ml

of the appropriate media, which was refreshed every 72hrs. After the first 72 hours, neurons begin to form neurites and networks (see Figure 5.1); however, they do not replicate and are post-mitotic cells.

At the end of the experiment, media was aspirated and the cells treated with 1x Trypsin-EDTA (Life-Technologies, Paisley, UK) for 5 minutes to detach them from the wells. The cells were then pelleted at 1,300rpm for 5 minutes. The supernatant was aspirated and the pelleted stored at -20^{o}C until required.

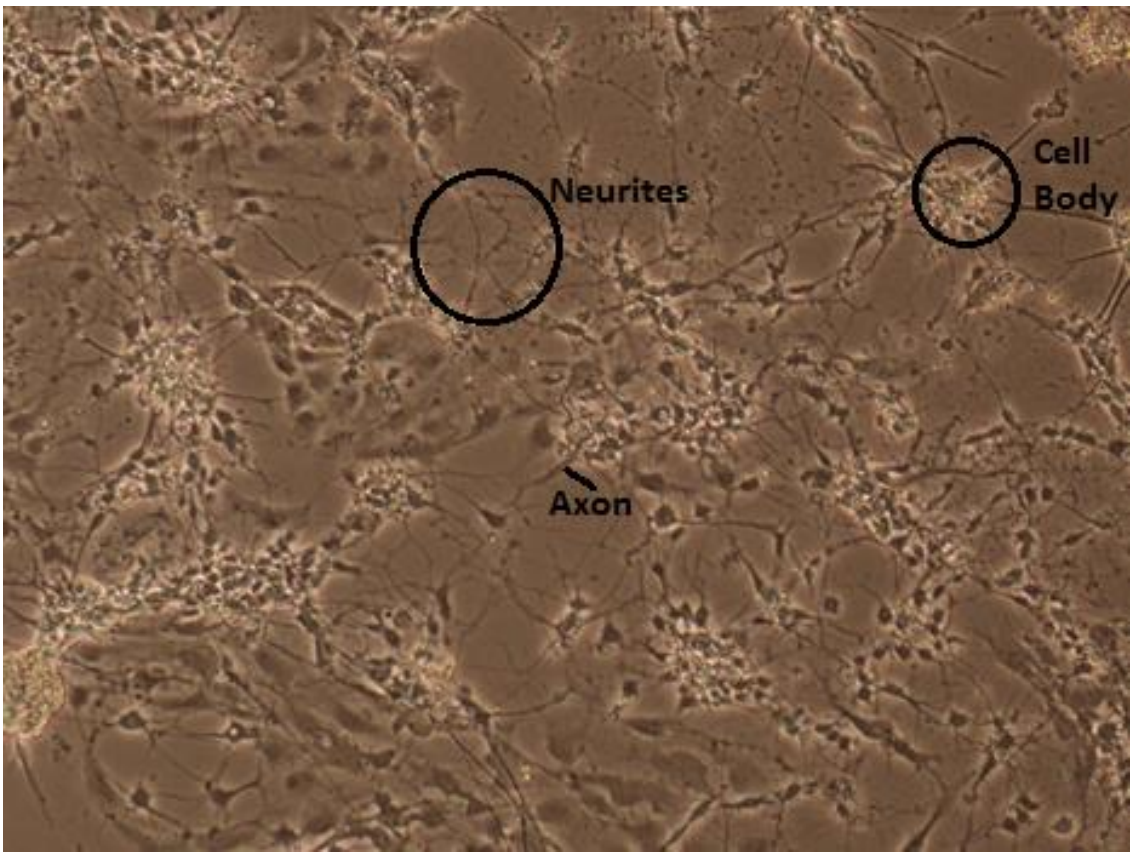

Figure 5.1 E18 cerebellar rat neurons 72 hours post culturing; the neurons have formed neurites between each neuron.

5.3.3 NRTI treatment experimental design

The NRTI exposure-repopulation assay was performed using the NRTIs and concentrations as previously described (see Section 3.1.5).

Exposure of the drug starts at experimental day 0 and exposure times were determined through previous pilot work (Dr Brendan Payne, IGM, Newcastle University, UK).

In order to determine whether or not cells treated with NRTIs were subject to a molecular bottleneck mechanism affecting the deletion mutation, cybrids were treated with ddI at physiological concentration as previously described (see Section 3.1.5) in 10 biological replicates with 10 untreated biological replicates for 14 days before remove the ddI and exposing the cells for a further 14 days to normal growth media (see Figure 5.4).

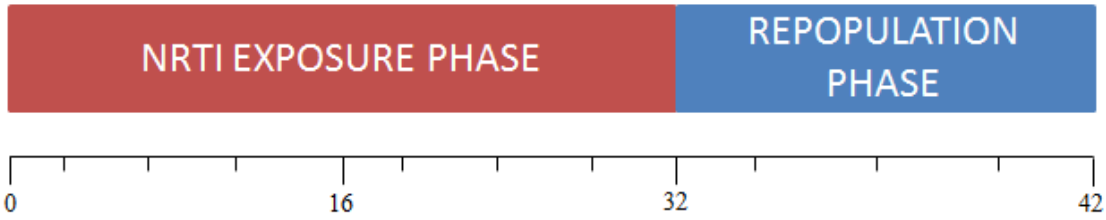

Figure 5.2 NRTI exposure-repopulation assay protocol for physiological dosing experiment of trans-mitochondrial cybrids

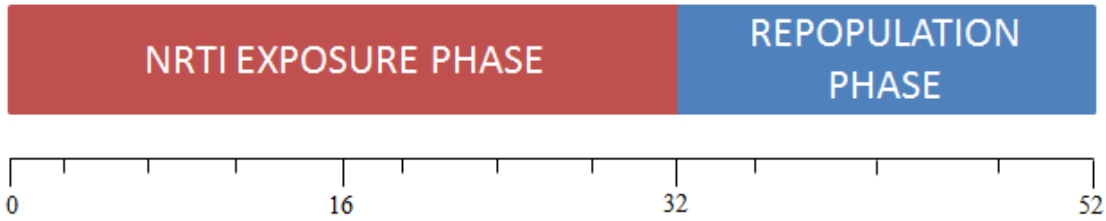

Figure 5.3 NRTI exposure-repopulation assay time course for the dose-response (10x physiological) study performed using trans-mitochondrial cybrids.

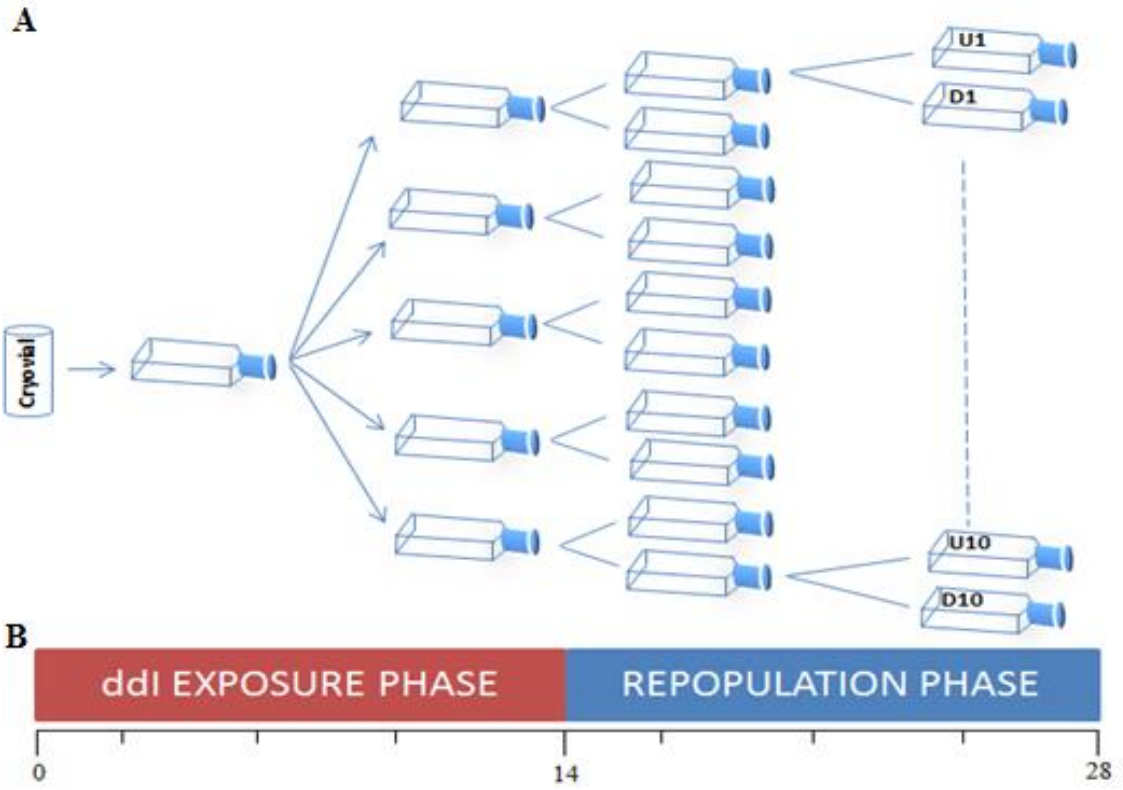

Figure 5.4 ddI exposure-repopulation assay procedure for assessment of genetic bottleneck mechanism in trans-mitochondrial cybrids grown in T_{25} flasks from a single cryovial for the untreated (U1-10) replicates and ddI treated (D1-10) replicates (A), with assay time course illustrated in days (B).

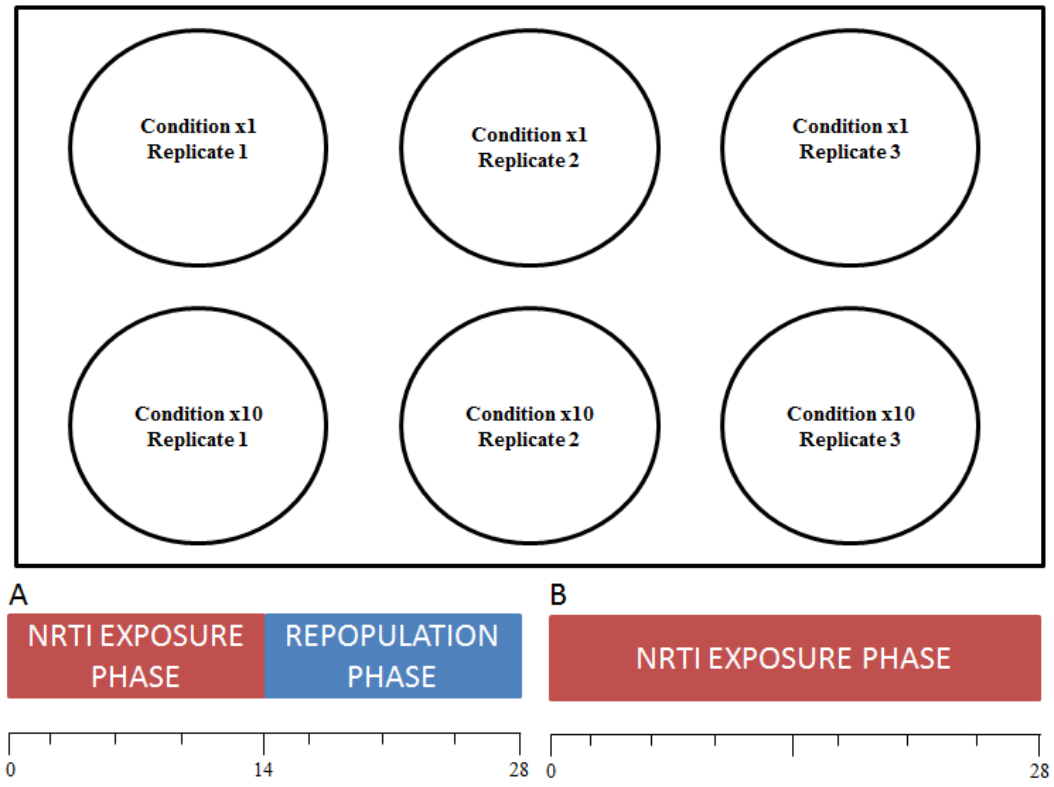

Figure 5.5 Experimental 6-well plate highlighting typical set up of NRTI condition layout for harvested cerebellar E18 neurons in NRTI exposure-repopulation (A) and NRTI exposure only (B) assays.

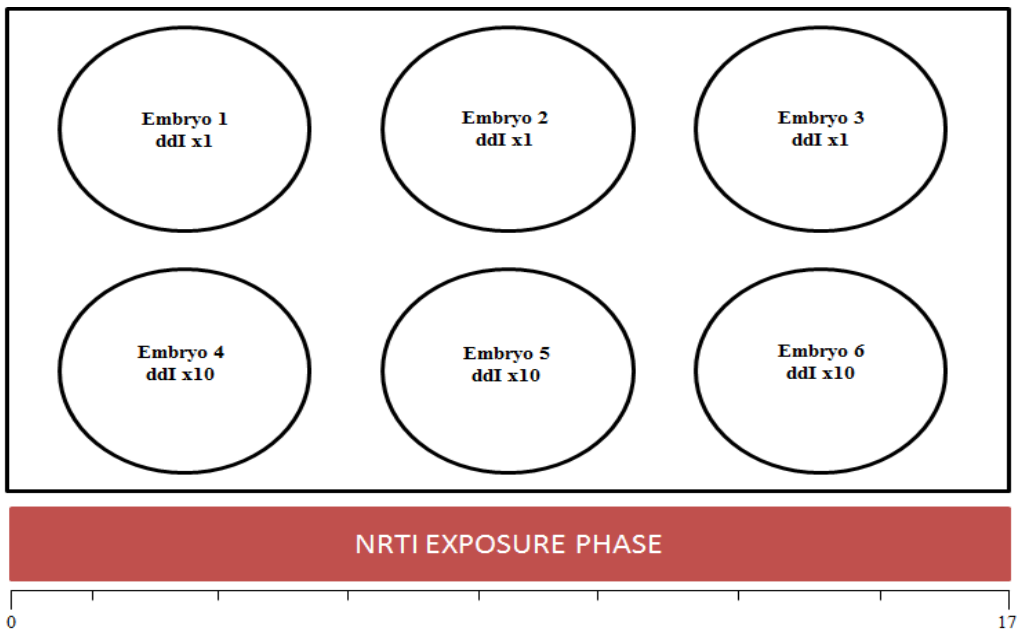

Figure 5.6 Experimental 6-well plate highlighting set up of ddI treatment only for individual E18 rat cerebellar neurons in treatment only assay. The embryo cerebellum was halved with one half used in the ddI condition and the other used in the untreated condition.

5.3.4 DNA extraction protocol

For the cybrid cells, each time they were split, the residual cells that were not re-seeded in propagation were span at 1300rpm for 5 minutes, supernatant removed and the pellet frozen at -80^{o}C, until required. The DNA was then extracted using the DNeasy® blood and tissue extraction kit (Qiagen, Manchester, UK) as previously described (see Section 3.3). The same protocol was applied to the extraction of DNA from the neuronal cells at the end of the study.

5.3.5 Mitochondrial network fragmentation analysis

In order to make an assessment of mitochondrial network morphology changes before, during and after NRTI exposure at the physiological concentration of each NRTI in the cybrids, two time points throughout the exposure phase of the experiment were chosen for assessment; day 13 and 27.

Cells were seeded at a density of 25,000 cells per well in a 12-well plate (Greiner bio-one, Stonehouse, UK) on 16mm round cover slips (VWR) and allowed to adhered for 48 hours at 37^{o}C and 5% CO_2, in normal cybrid growth media with corresponding treatment NRTI (see Section 5.3.1). The media was then aspirated out of the wells and the cells on the cover slip washed twice in PBS (Oxoid, Thermo-Scientific, Hampshire, UK) before the addition of 0.5ml/well of staining solution consisting: DMEM cell media (see Section 5.3.1), 1µg/µL Hoechst 33342 (Molecular Probes, Life-Technologies, Paisley, UK) and 100nM MitoTracker® Green (Molecular probes, Life-Technologies, Paisley, UK). To allow for incorporation of the dyes, the wells were incubated for 30 minutes at 37^{o}C and 5% CO_2. Following this, the staining solution was aspirated and the cover slips washed three times with PBS and stored in culture medium prior to visualisation on a microscope.

The cover slips were removed from the wells using fine point tweezers and the excess media removed using absorbent tissue before being inverted (cell side facing down) onto a microscope slide with a small drop (5µL) of PBS between them to prevent the cells drying. The mitochondrial network morphology and nucleus structure were then captured using an upright widefield Axio Imager.Z1 epifluorescence (Carl Zeiss, Cambridge, UK) with an oil-immersion objective (plan apochromat 63x / N.A. = 1.4) on an AxioCam HRm high resolution digital camera using AxioVision v4.6 software (all Carl Zeiss, Cambridge, UK). The MitoTracker® Green probe stained the mitochondrial

network and was imaged using the GFP (green) channel. The Hoechst 33342 probe stained the nucleus and was detected using the DAPI (blue) channel. Each cell was captured in a series of z-stack 'slices' of 20 images at 0.2µM separation before being saved as a compressed single file including metadata. Approximately 10 single cells were imaged, depending upon cell separation on the cover slip, image clarity and the hydration status of the cells.

Prior to analysing the morphology the images went through a process of deconvolution using Huygens Essential Software (Scientific Volume Imaging, Netherlands). This process recovers the image that has been degraded through the physical processes involved in capturing the image by using a point spread function to remove the 'spreading' or 'blurring'. The deconvolved images were then used to analyse the mitochondrial network morphology using the following parameters; garbage, 50; seed, 0%; threshold, 5%. The mitochondrial length and volume for each cell analysed was copied into GraphPad™ Prism v5 (GraphPad™ Software, CA, USA) for visual graphing and statistical analyses.

5.3.6 Quantification of human mitochondrial deletion and copy number

The deletion level and mitochondrial copy number was calculated using the multiplex qPCR assay as previously described (see Section 3.6.1) using previously described analysis methods for deletion level and copy number calculation (see Sections 3.6.2 and 3.6.3). The concentration of each DNA sample was standardised to 25-50ng loaded into each 25µL reaction. Each sample was run in duplicate on the reaction plate and allowing for a 0.5Ct difference between sample replicates.

Mitochondrial copy number was expressed relative to the untreated group for each time point. Therefore, when sample copy number is identical to untreated copy number, relative copy number ratio would equal one.

5.3.7 Measurement of *de novo* deletion formation

To investigate whether or not NRTI exposure induces new deletion mutations, a long range PCR was used to screen for mtDNA deletions as previously described (see Section 3.7). The products were run on a 1% agarose gel, as previously described (see Section 3.5) using the DNA ladder 1kb plus (Thermo-Scientific, Hampshire, UK) to determine the size of the bands.

5.3.8 Quantification of embryonic rat neuron mitochondrial copy number

To assess the level of E18 neuronal mitochondrial copy number, a novel qPCR assay was designed. The assay was designed in the same manner and implemented on the same principles as the human qPCR SYBR™ Green method.

Primers (standards and qPCR) were designed to amplify a mitochondrial region and a nuclear housekeeping gene. Unlike human mtDNA, the D-loop region in rats has a large conserved region (Reyes *et al.*, 2003), primers were designed in this region for the mitochondrial target; β-actin gene was chosen as the nuclear target. The qPCR primers were designed to be nested within the template standards (see Table 5.1). The standards curves were determined as previously described (see Section 3.6); all samples ran in triplicate (see Figure 5.7) on the iQ5™ thermocycler (Bio-Rad), with the mean and standard deviation calculated.

Gene	**Forward Sequence 5' – 3'**	**Reverse sequence 5' – 3'**	**T_m**
Template			
D-loop	CCATTCATTATCGCCGCCCT	CCTTCATGCCTTGACGGCTA	60
β-Actin	CAAAGCTTAACTTTCCCGGCCC	AGTCCTTCTGACCCATACCCA	60
qPCR			
D-loop	TCCCCAAGCATATAAGCATGTAA	TGGTGCATGTCTAATAACACAGA	60
β-Actin	GGAACTCTTCCTCTCCCCCT	CGCCCTCGCCCAACC	60

Table 5.1 *Rattus norvegicus* template and qPCR primer properties. Primers were designed using genbank accession numbers X14848 and V01217 for the D-Loop and *β-Actin*, respectively.

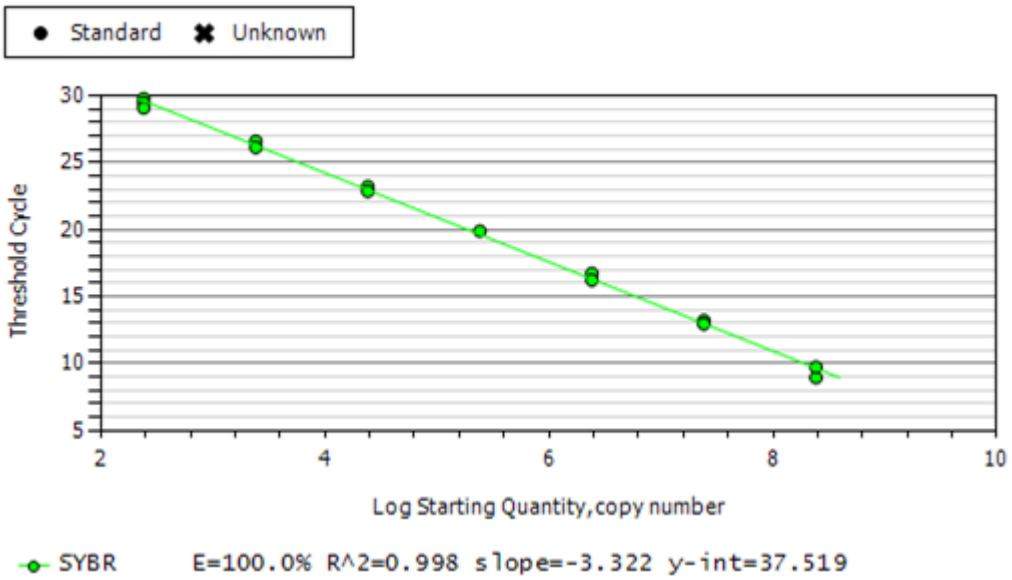

Figure 5.7 Standard curve dilution series (10^8 – 10^2 copies/µL) of nuclear reference gene, *β-Actin*, indicating good eficiency and R^2, run on the iQ5™ thermocycler using the SYBR™ Green method.

The total mtDNA copy number for each sample was then calculated using the same method as previously described (see Section 3.6.3); the SQ (starting quantity), which was determined automatically for each sample based on the standard curve values and efficiency. Therefore, assuming 100% efficiency, the mtDNA copy number was calculated using the following equation:

$$mtDNA\ Total\ Copy\ Number = \frac{D-Loop\ SQ}{(\beta-Actin\frac{SQ}{2})}$$

5.3.9 Detection of embryonic rat neuron mtDNA common deletion

To quantify the level of rat mtDNA common deletion, a novel technique was developed. The mtDNA common deletion in rats is slightly smaller than that found in humans, but still occurs between homologous repeat regions of 16bp and is referred to as m.δ4834. The presence of m.δ4834 was detected using small amplicon PCR by designing specific primers (see Table 5.2) around the breakpoint region (i.e. m.8103 and 12952) which would generate a small ~300bp amplicon (~150bp either side of the break point) when the m.δ4834bp was present in a sample.

Name	Forward primer 5' – 3'	T_m (^oC)	Amplicon Size (bp)
δ4834bp	F - GGTCTACCAATTGTTGTGACCAT R - ATGCTAGGCGTTTGATTGGA	59	298

Table 5.2 Primer properties of *Rattus norvegicus* m.δ4834 small amplicon PCR

The amplification of the product was performed using 50ng of each DNA sample in the following 25µL reactions containing: 10x ImmoBuffer, 2mM dNTPs, forward Primer and reverse primer (10µM), 50mM $MgCl_2$ and autoclaved PCR-grade deionised water, 5U Immolase *taq* (Bioline, London, UK).

The product was then amplified under the following conditions: an initial denaturation step at 95^{o}C for 10 minutes, then 40 cycles of denaturation at 95^{o}C for 1 minute, annealing at 59^{o}C for 1 minute, finally extension at 72^{o}C for 30 seconds, with a final extension step at 72^{o}C for 10 minutes. The product was then run on a 2% agarose gel as previously described (see Section 3.5), using the DNA ladder Hyperladder IV (Bioline, London, UK) to determine the size of the band.

5.3.10 Single cell analysis

To assess the deletion level within single cybrid cells during the NRTI exposure protocols, single cell analysis was performed as previously described (see Section 3.1.7). Approximately 20 cells were captured and lysed (see Sections 3.2.2) for each condition studied at the end of the experiment and 5µL of the lysate loaded into the downstream molecular analyses (deletion quantification, see Section 5.3.6).

5.4 Results

5.4.1 Effects of NRTI-exposure on mtDNA

The total, mutant and wild type mtDNA copy number (relative ratio to the untreated) of each NRTI treatment group for the cybrid cells is displayed below (see Figure 5.8, Figure 5.9 and Figure 5.10). Each data point is expressed as the mean values derived from the biological and qPCR technical replicates with error bars plotted as standard deviations (S.D). All data is expressed on logarithmic scales relative to untreated.

The ddI condition throughout the experiment has a reduced total copy number compared to the untreated especially throughout the exposure phase with a depletion maximum present at experimental day 22 with a ratio of 0.59 (±S.D 0.04) with minor fluctuations until the end of the exposure phase where at experimental day 29 depletion, it was at 0.7 (±S.D 0.00). During the post-exposure phase the copy number appears to repopulate by experimental day 36 at 1.18 (±S.D 0.23); however, throughout the first eight days of the repopulation phase, the spread between replicates was large with approximately 25% difference between them.

The d4T and AZT conditions present with a small amount of depletion of total copy number during the exposure phase with depletion maximum present at 0.59 (±S.D 0.10) for the d4T and 0.64 (±S.D 0.04) for the AZT by experimental day 22. The general trend appears to fluctuate in line with the untreated condition.

The total copy number of the TDF condition shows hardly any depletion throughout the exposure phase. The post-exposure phase shows the condition plateau to a comparable level as the untreated, reaching a relative to untreated copy number of 0.95 (±S.D 0.32) by the end of the experimental at experimental day 42.

The mutant copy number graph (Figure 5.9) illustrates the copy number behaviour to be highly comparable with total copy number. Across all conditions there's minimal if any difference to the total copy number and fluctuates through the experiment.

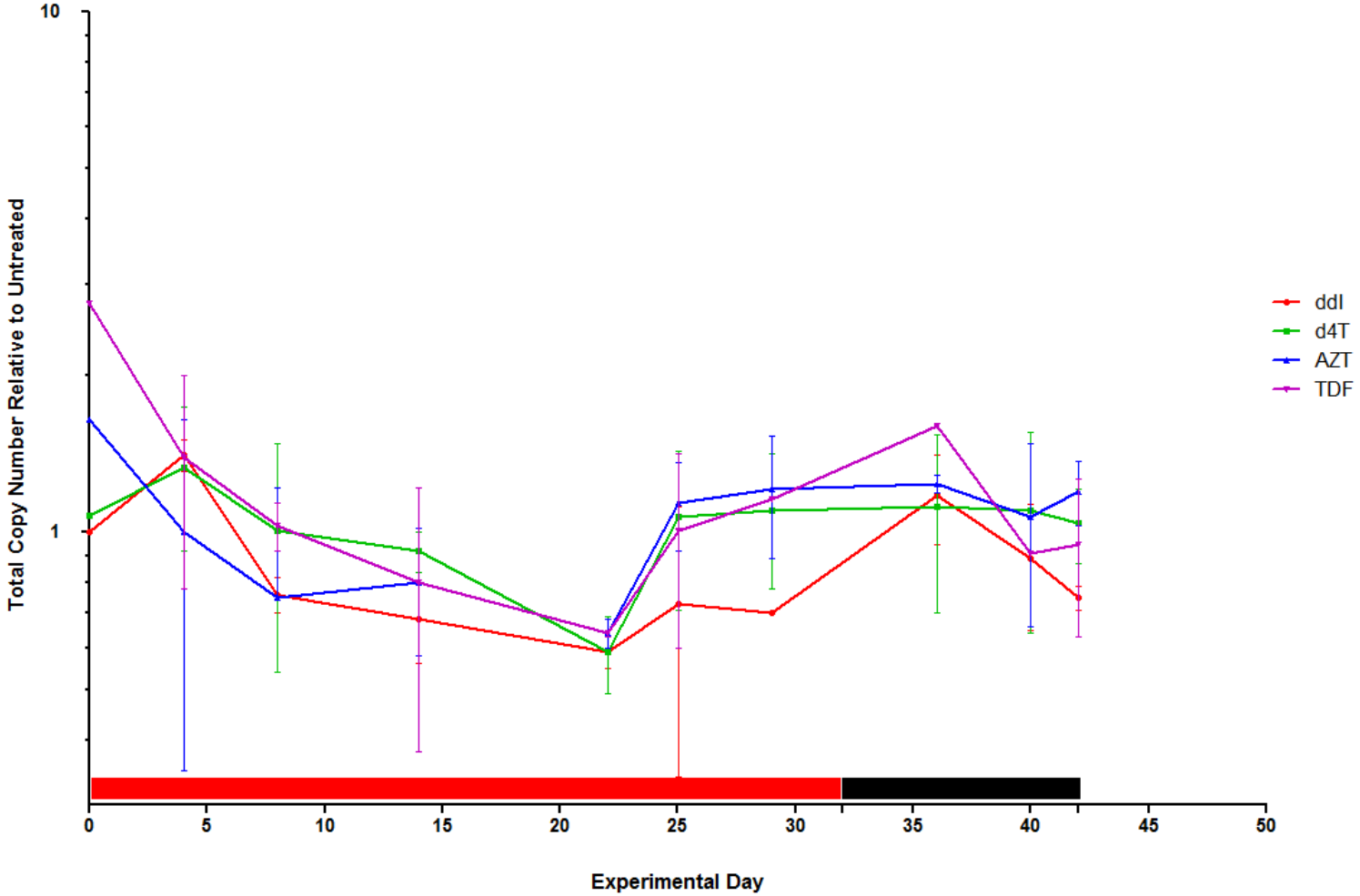

Figure 5.8 The total mitochondrial copy number of trans-mitochondrial cybrids during the NRTI exposure-repopulation assay. Red bar indicates exposure phase and black indicates untreated phase. Error bars are standard deviations of the mean.

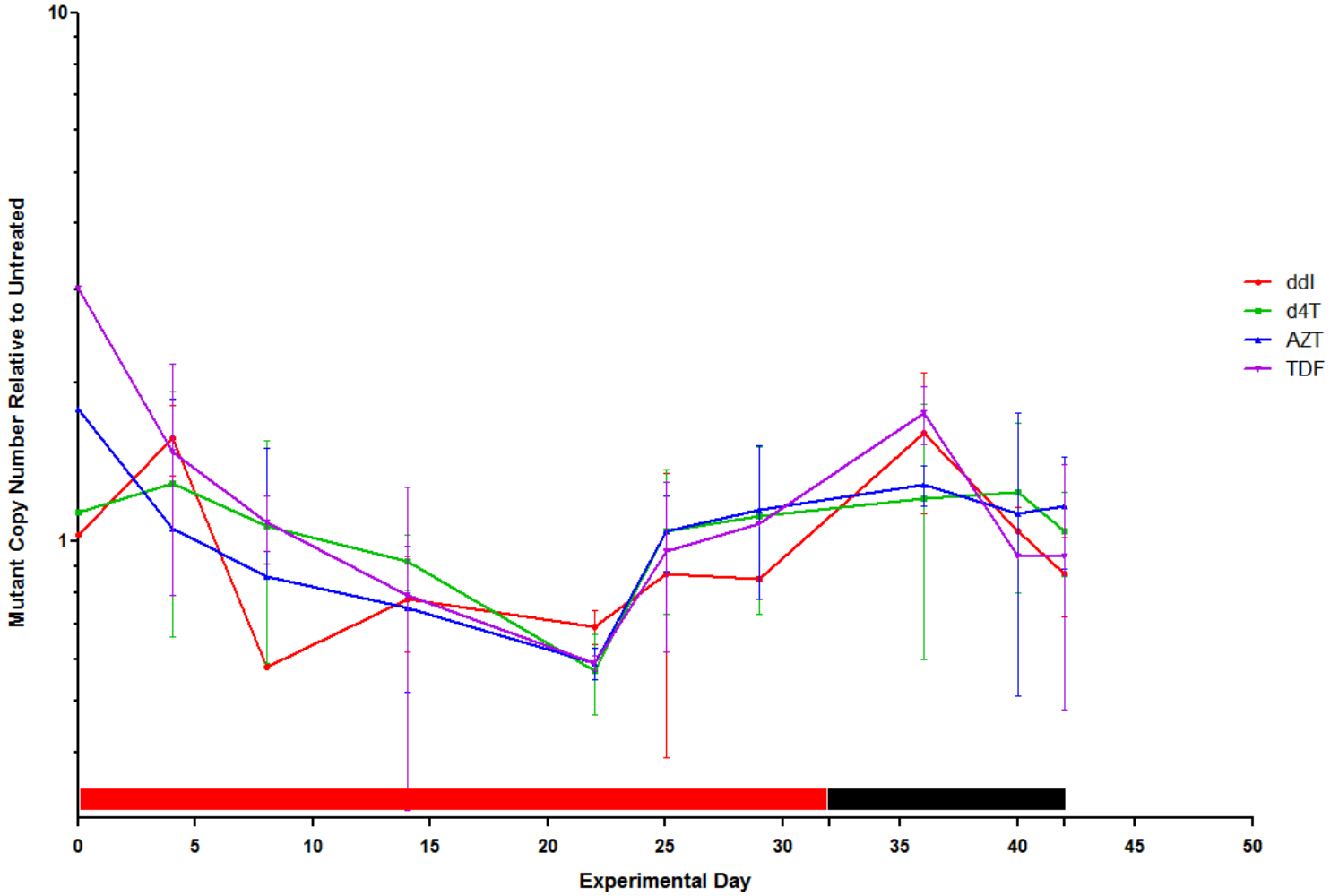

Figure 5.9 The mutant mitochondrial copy number of trans-mitochondrial cybrids during the NRTI exposure-repopulation assay. Red bar indicates exposure phase and black indicates untreated phase. Error bars are standard deviations of the mean.

The picture for the wild-type tDNA is very different. The most clear observation is the depletion in the ddI throughout the exposure phase, reaching a depletion maximum of 0.14 (±S.D 0.03) at experimental day 29 ($p<0.001$; Table 5.3*)*. The repopulation sees the copy number begin to increase and reaches a repopulation maximum of 0.57 (±S.D 0.49) at experimental day 40. It is evident however, that there's a large spread between the two biological replicates, contributing towards the large standard deviation values present throughout the repopulation phase.

The d4T, AZT and TDF conditions show little depletion throughout the exposure phase and are comparable to the total and mutant copy number graphs (Figure 5.8 and Figure 5.9). All three conditions reach an experimental depletion maximum at experimental day 22 of 0.67 (±S.D 0.11), 0.85 (±S.D 0.09) and 0.87 (±S.D 0.06), respectively. All three conditions then increase at experimental day 29 to greater than the untreated condition with a d4T increasing to 1.03 (±S.D 0.09), AZT 1.42 (±S.D 0.1) and TDF 1.49 (±S.D 0.01).

During the repopulation phase, the d4T, AZT and TDF conditions can be seen to fluctuate around a similar level as the untreated with a repopulation maximum found to be 1.03 (±S.D 0.11), 1.29 (±S.D 0.15) and 1.40 (±S.D 0.31). None of the exposure and post-exposure data for d4T, AZT and TDF were found to be significantly different to the untreated with *p*- values displayed in Table 5.3, as determined by a student *t*-test at experimental day 29.

Condition	**Copy number (relative to untreated)**	**S.D**	***p*-value**
ddI	0.14	0.03	<0.001
d4T	1.03	0.09	0.865
AZT	1.42	0.1	0.471
TDF	1.49	0.31	0.159

Table 5.3The *p*-values for the exposure-repopulation data compared to the untreated values at experimental day 29.

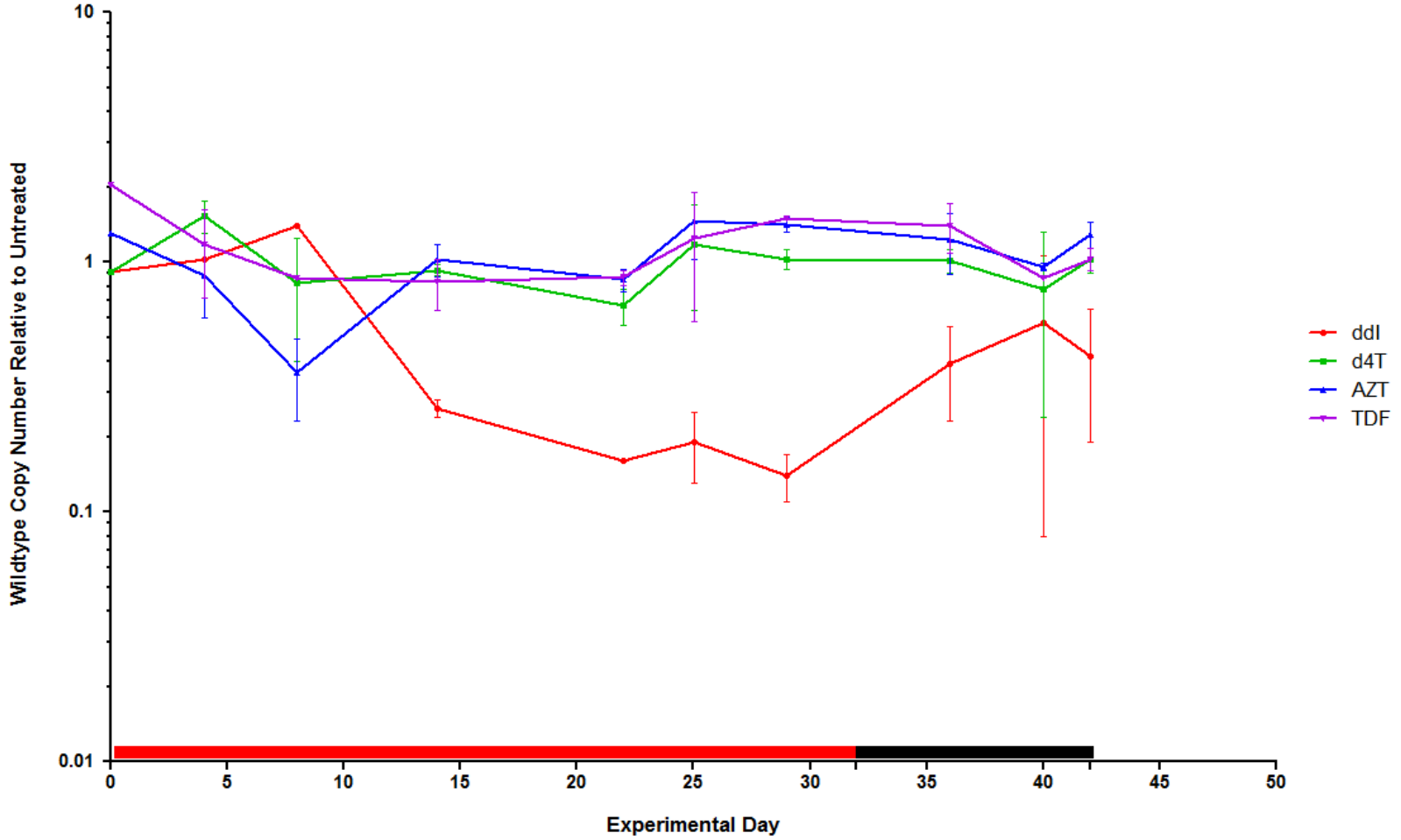

Figure 5.10 The wild type mitochondrial copy number of trans-mitochondrial cybrids during the NRTI exposure-repopulation assay. Red bar indicates exposure phase and black indicates untreated phase. Error bars are standard deviations of the mean.

The mitochondrial deletion heteroplasmy has been graphed and expressed as deletion level percentage for the cybrid cells during the NRTI exposure and post-exposure assay (Figure 5.11). Each data point is expressed as the mean values from the biological and qPCR technical replicates, with error bars plotted as standard deviations.

The striking change in deletion heteroplasmy is seen exclusively in the ddI condition. During the exposure phase, the heteroplasmy level increases from 75.3% to 95.6% (±S.D 1.6%) by experimental day 29 and it was found to be statistically significantly different to the untreated ($p<0.01$). This shift represents an experimental heteroplasmy shift maximum of 20.4%. The heteroplasmy level reduces during the post-exposure phase to a heteroplasmy of 85.6% (±S.D 7.17%) at experimental day 42. The overall shifts during both the exposure and post-exposure phase was found to be highly significant ($p<0.001$) at experimental day 42 (*t*-test) compared to the untreated.

The d4T, AZT and TDF conditions appeared to fluctuate slightly throughout the experiment and remain comparable to the untreated condition.

The overall final shift in heteroplasmy of the d4T, AZT or TDF condition was found not to be significant; d4T, $p=0.847$; AZT, $p=0.589$; TDF, $p=0.810$. Figures derived using the experimental day 42 time point, implementing a two tailed student *t*-test.

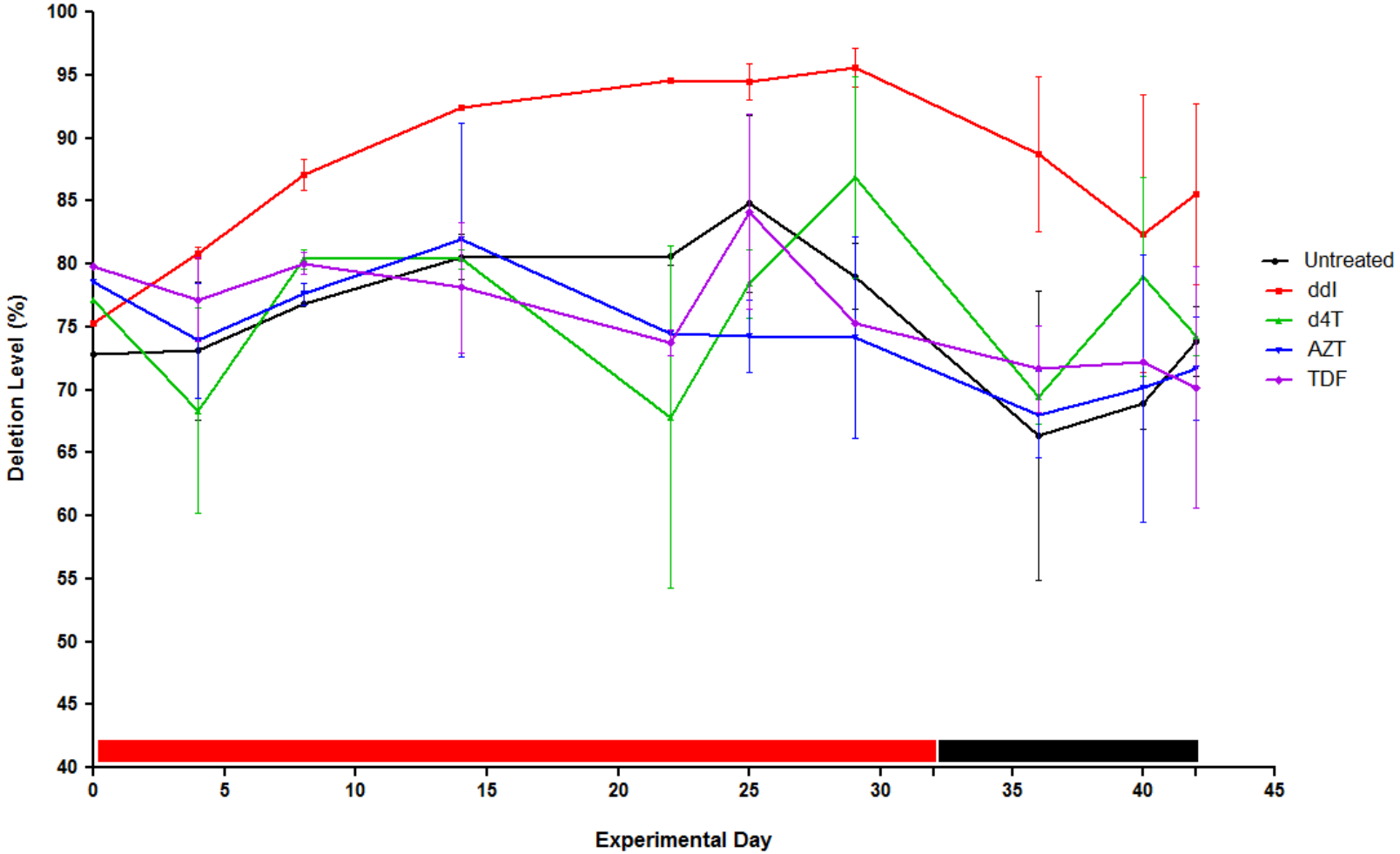

Figure 5.11 The heteroplasmy distribution of large scale deletion within cybrid cells during the NRTI exposure-repopulation assay. Red bar indicates exposure phase, black bar indicates repopulation phase. Error bars are standard deviations of the mean.

5.4.2 Effect of NRTI exposure on mitochondrial network structure

The average length and average number of mitochondrial networks for each condition at experimental day 13 and 27, during the NRTI exposure-repopulation assay, were estimated for approximately 10 single cells and the data graphed as column graphs with error bars expressed as standard deviations of the mean (see Figure 5.12 and Figure 5.13).

The average mitochondrial length (see Figure 5.12) at experimental day 13 appears to be almost constant across the conditions; However, conditions ddI and TDF were found to have a statistically significant smaller average network length, compared to the untreated condition, with *p*-values of 0.01 and <0.001, respectively. Whereas at experimental day 27: ddI, d4T and TDF, were all statistically significant in mean length compared to the untreated condition.

The average mitochondrial length in the untreated condition reduces by 25.92% ($p<0.001$) by experimental day 27, and a 15.26% ($p=0.135$) reduction in condition AZT; whereas, ddI, d4T and TDF conditions all increase accordingly: 8.64% ($p=0.540$), 33.58% ($p=0.018$) and 29.42% ($p=0.018$) (see Table 5.4).

Experimental Condition	**Day 13 Mean Length (µM)**	**Day 27 Mean Length (µM)**	**Mean percentage change**	***p*-value**
Untreated	3.24 (0.47)	2.40 (0.66)	- 25.92%	<0.001
ddI	3.01 (0.42)	3.27 (0.79)	+ 8.64%	0.540
d4T	2.68 (0.44)	3.58 (0.89)	+ 33.58%	0.0180
AZT	3.08 (0.59)	2.61 (0.63)	- 15.26%	0.135
TDF	2.38 (0.32)	3.08 (0.79)	+ 29.42%	0.0180

Table 5.4 Average mitochondrial length changes in all conditions from experimental day 13 to experimental day 27 during the NRTI exposure-repopulation assay. Expressed as mean and S.D (in brackets) of 10 replicates, as measure in µM. Percentage change, from day 13 to 27, measures gain (+) and losses (-). *P*-values calculated by student *t*-test.

The average number of mitochondrial networks (see Figure 5.13) at experimental day 27 decreases from experimental day 13 in all NRTI conditions; however the number of networks actually increases in the untreated condition by 74.73%. The condition with

the greatest decrease in number of mitochondrial networks was d4T with a 78.10% decrease. AZT condition decreased the least from experimental day 13 to 27 with a 24.31% decrease. The ddI and TDF conditions were found to be decreased by 33.90% and 75.82%, respectively, from experimental day 13 to 27 (see Table 5.5).

Experimental Condition	**Day 13 Mean Number (µM)**	**Day 27 Mean Number (µM)**	**Mean Percentage Change**
Untreated	56.6 (26.35)	131.33 (60.02)	74.73%
ddI	94.5 (26.70)	60.6 (30.18)	-33.90%
d4T	122.5 (64.13)	44.4 (16.59)	-78.10%
AZT	114.56 (75.23)	90.25 (67.55)	-24.31%
TDF	126.45 (82.37)	50.63 (42.34)	-75.82%

Table 5.5 Average number of mitochondrial networks across all conditions with mean changes from experimental day 13 to experimental day 27 during the NRTI exposure-repopulation assay. Expressed as mean and S.D (in brackets) of 10 replicates. Percentage change, from day 13 and 27, measures gains (+) and losses (-).

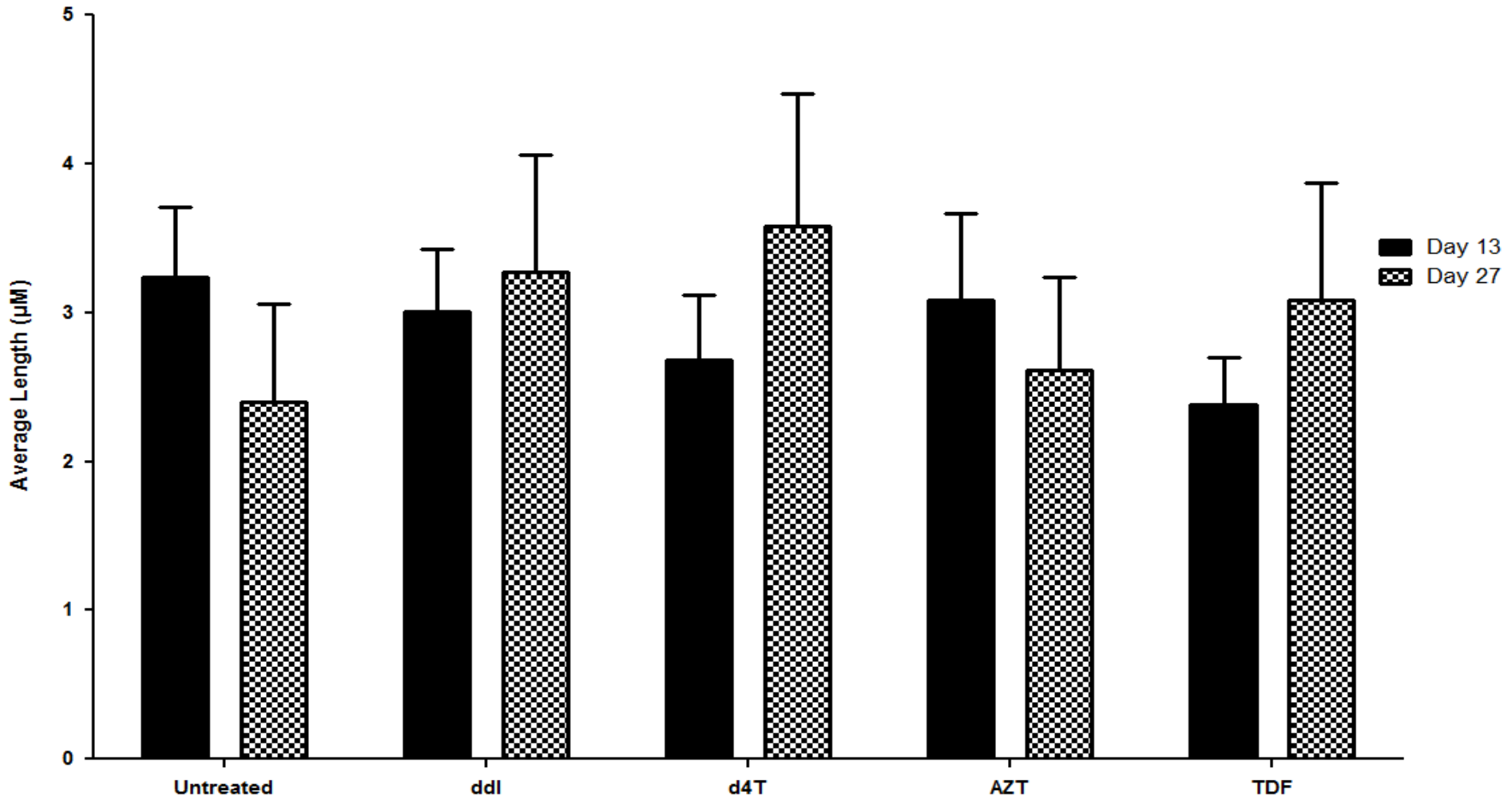

Figure 5.12 The average mitochondria network length, calculated from 10 individual cells for each treatment condition, at experimental day 13 and 17, during the NRTI exposure-repopulation assay. Error bars are standard deviations of the mean.

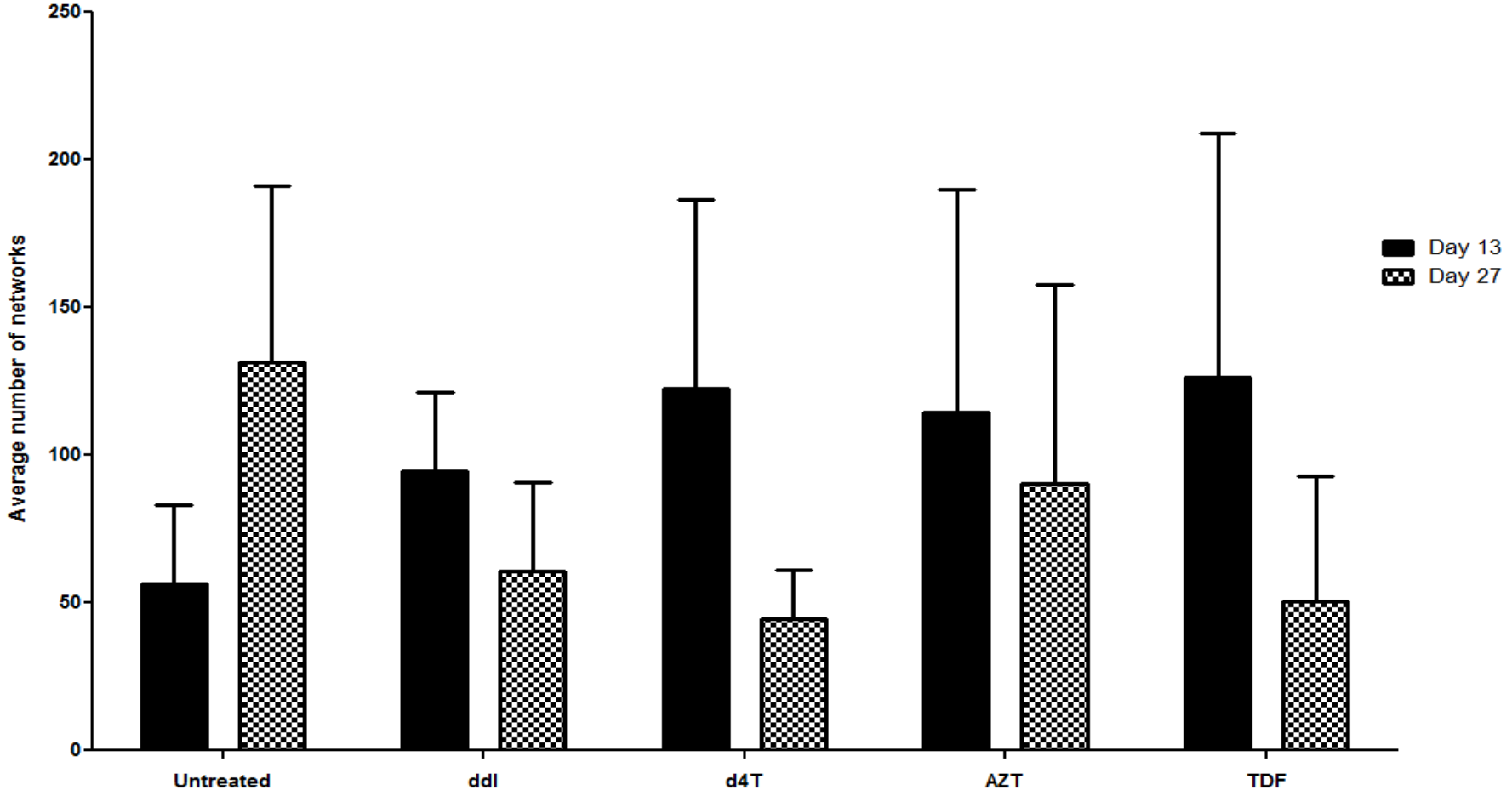

Figure 5.13 The average number of mitochondrial networks calculated from 10 single cells for each treatment condition, at experiemntal days 13 and 27, during the NRTI expsoure-repopualtion assay. Error bars are standard deviations of the mean.

5.4.3 Modelling NRTI-induced molecular bottleneck

The total, mutant and wild type mitochondrial copy number of each of the 10 replicates of ddI exposure for the cybrid cells is displayed below (see Figure 5.14, Figure 5.15 and Figure 5.16). No results are presented for experimental day 28 due to DNA extraction failure.

A molecular bottleneck effect on mutation heteroplasmy would be predicted to manifest as an increase in the spread of mutant heteroplasmy levels following the period of mtDNA depletion ('the bottleneck').

Overall, the data highlights three key findings:

- The total copy number in both conditions shows fluctuation with minor depletion present in the ddI-exposed condition.
- The mutant copy number behaves almost identically to the total copy number with the untreated found to fluctuate throughout the experiment and the ddI-exposed condition showing minor depletion.
- The wild-type copy number in the ddI-exposed condition behaves in the opposite manner to the total and mutant cop number, with a large depletion present by the end of ddI-exposure, at experimental day 14 in all replicates. The untreated condition wild-type copy number fluctuates throughout the experiment, in-line with the total mutant copy number.

The mitochondrial large scale deletion heteroplasmy of each replicate (10 biological replicates) in the ddI exposure-repopulation study to assess for a genetic bottleneck effect in the mitochondrial trans-mitochondrial cybrid cell line is displayed below (see Figure 5.17). Each data point is expressed as the mean values from the qPCR technical replicates with error bars plotted as standard deviations.

The key finding from this graph is all of the ddI-exposed replicates increase in mtDNA heteroplasmy deletion level during the exposure phase for the first 14 experimental days. The replicates were found to congregate at almost the same level at a mean of 97% heteroplasmy across the replicates. After the exposure phase, the deletion level slowly decreases until experimental day 24. The untreated conditions were found to have a fluctuating heteroplasmy level throughout the experiment.

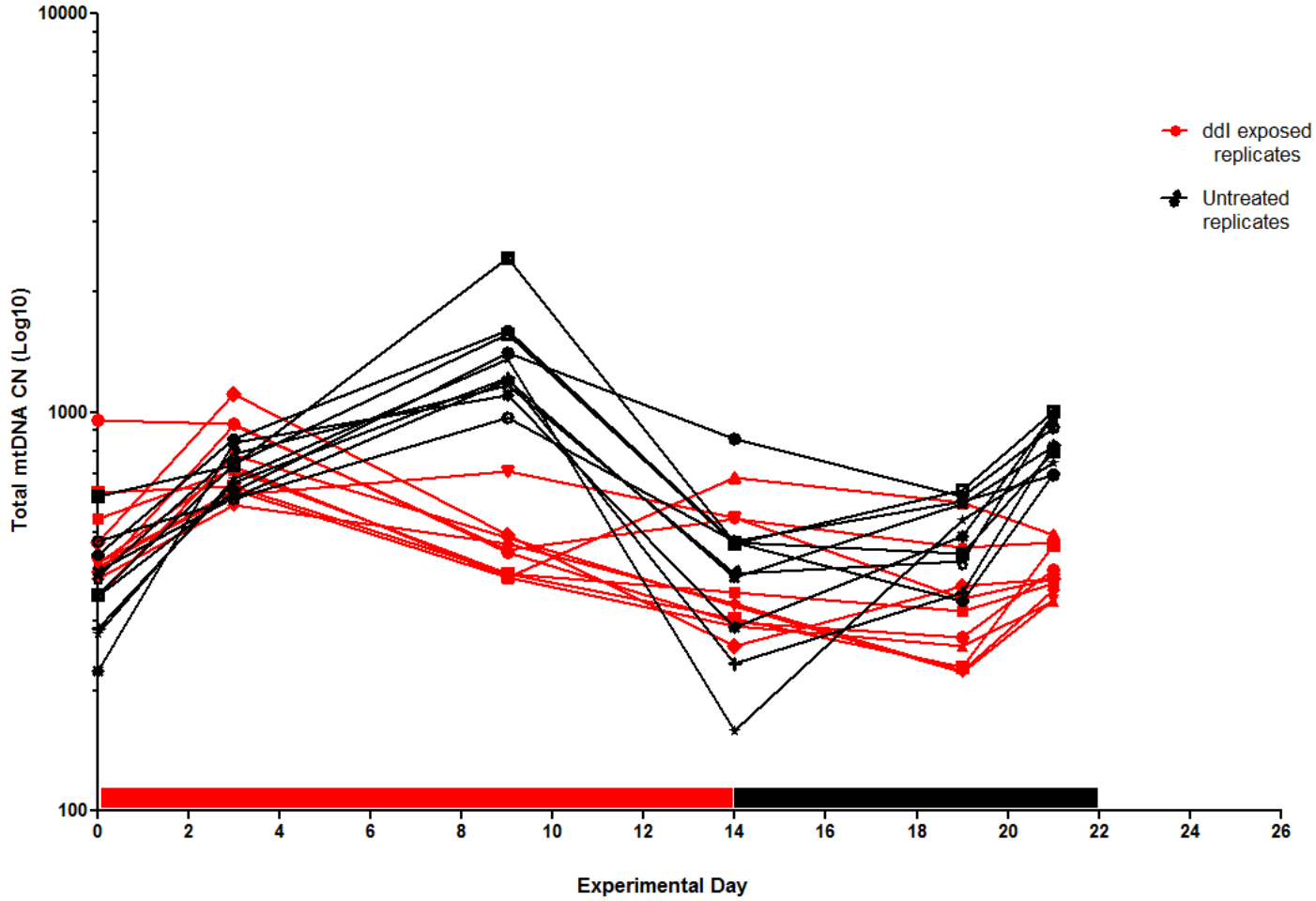

Figure 5.14 Total mitochondrial copy number through ddI exposure-repopulation assay to assess for a mitochondrial genetic bottleneck. Red bar indicates ddI exposure phase and black bar indicates repopulation phase. The red lines indicate ddI exposed replicates and the black indicate untreated replicates.

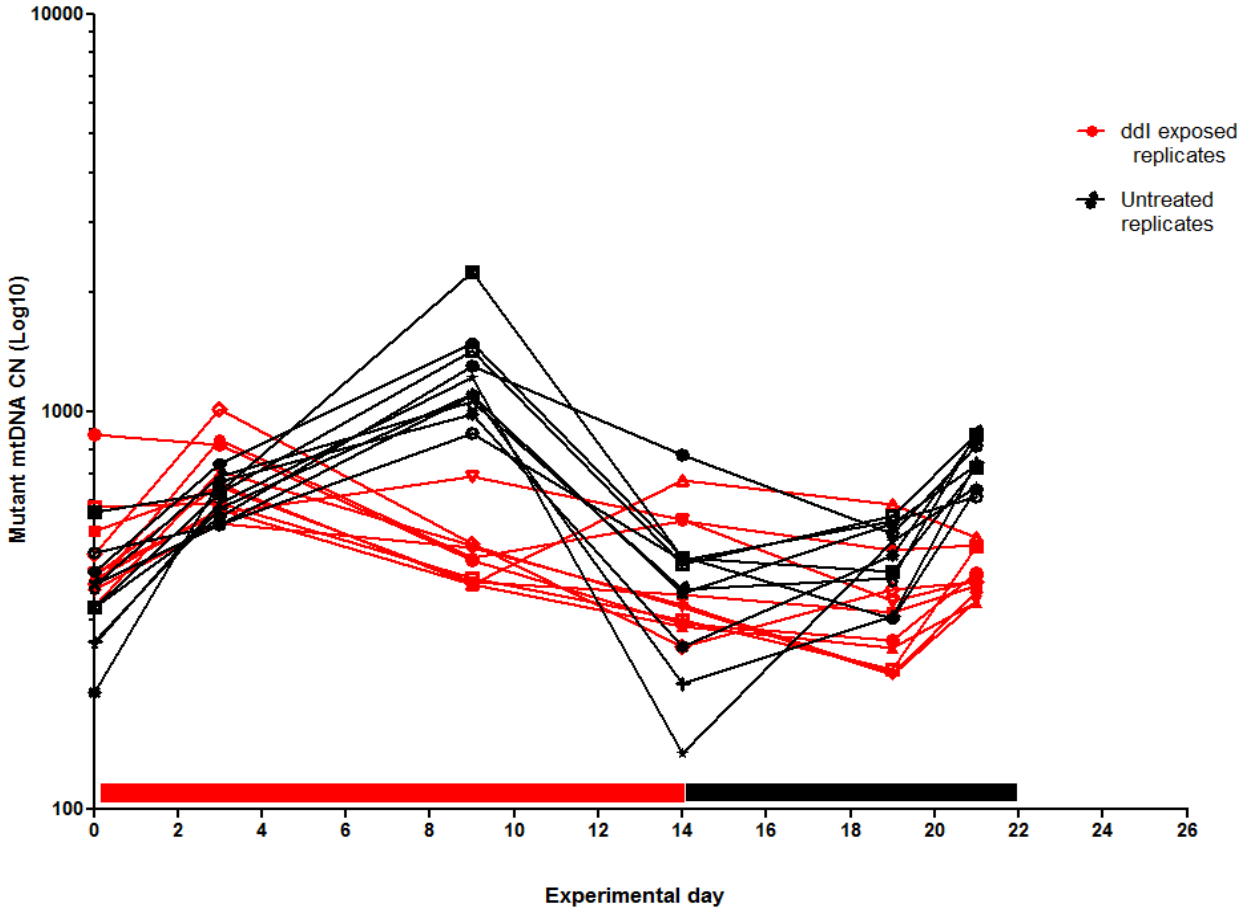

Figure 5.15 The mutant mitochondrial copy number through ddI exposure-repopulation assay to assess for a mitochondrial genetic bottleneck. Red bar indicates ddI exposure phase and black bar indicates repopulation phase. The red lines indicate ddI exposed replicates and the black indicate untreated replicates.

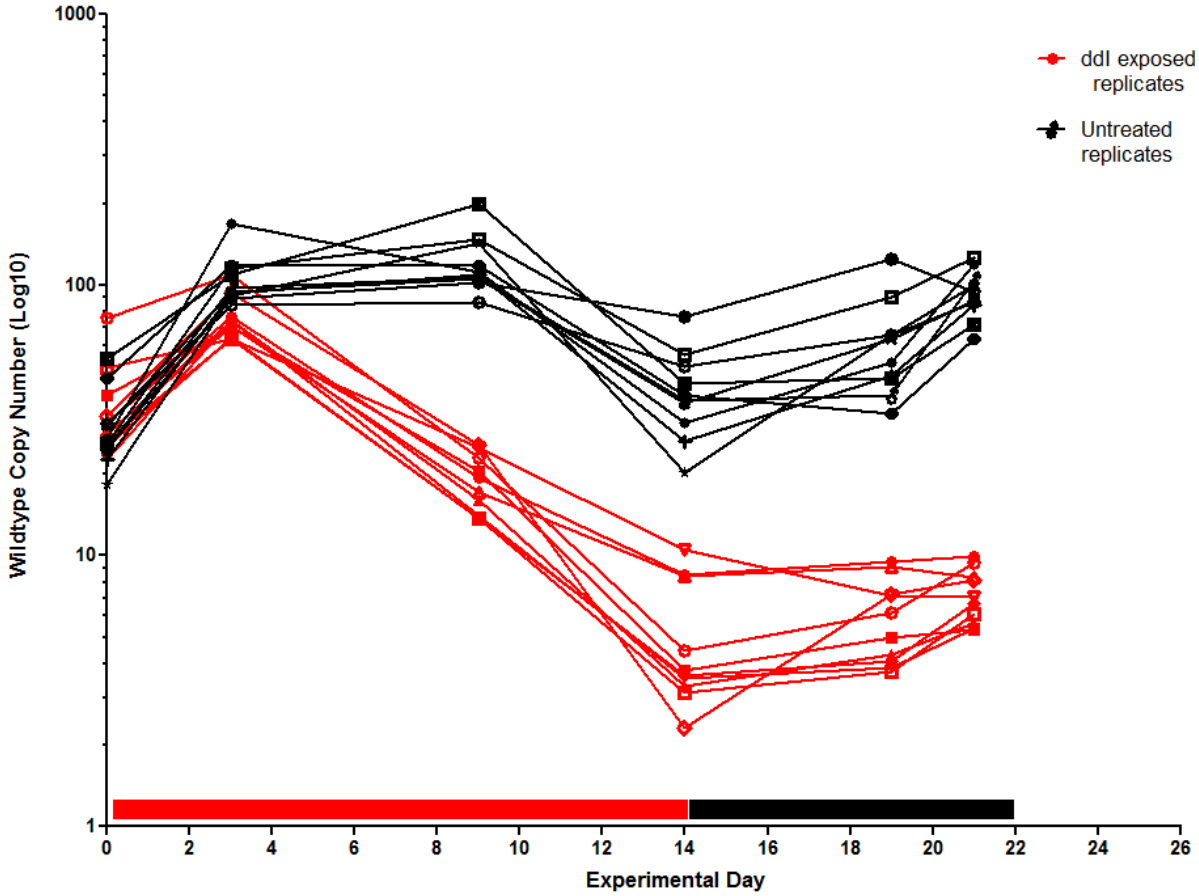

Figure 5.16 The wild type mitochondrial copy number through ddI exposure-repopulation assay to assess for a mitochondrial genetic bottleneck. Red bar indicates ddI exposure phase and black bar indicates repopulation phase. The red lines indicate ddI exposed replicates and the black indicate untreated replicates.

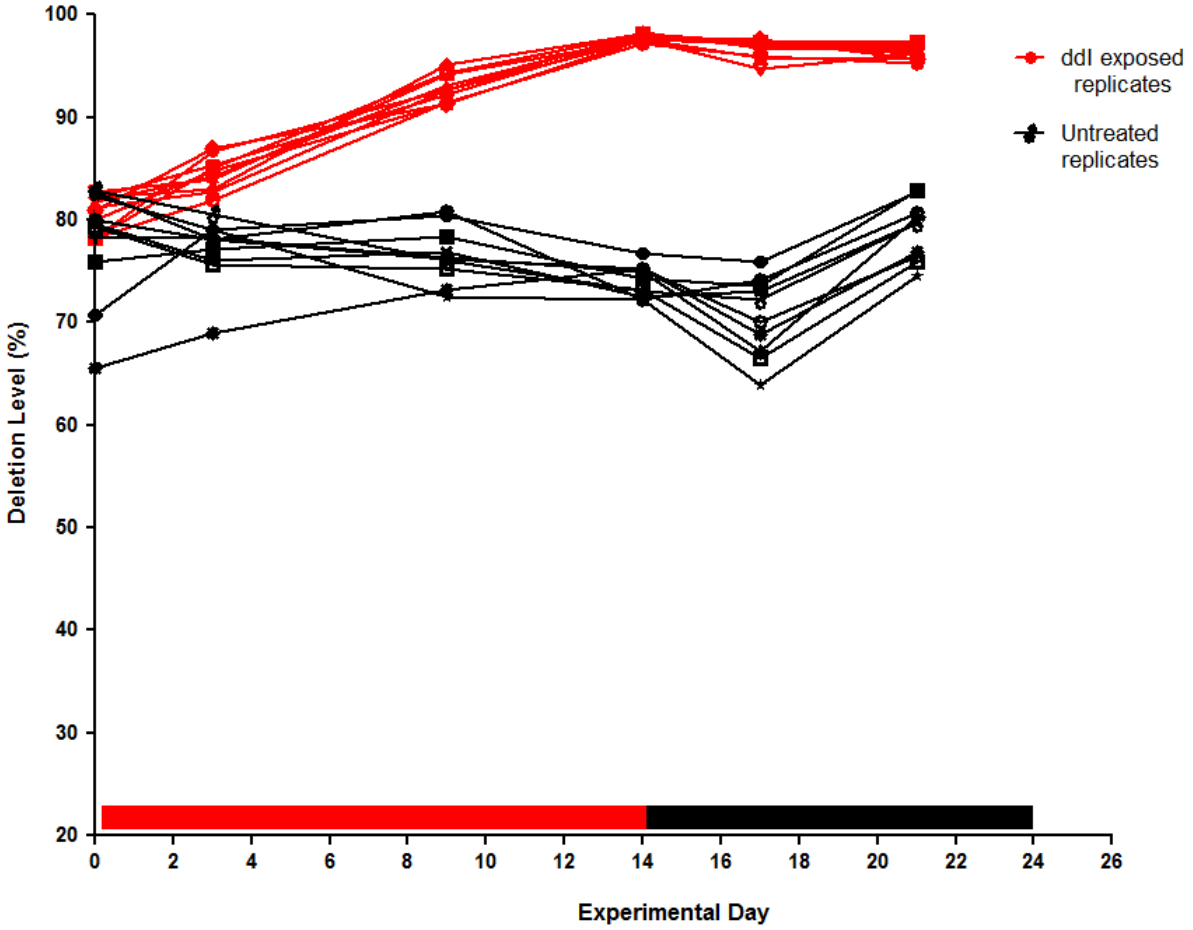

Figure 5.17 The distribution of a large scale deletion number through ddI exposure-repopulation assay to assess for a mitochondrial genetic bottleneck. Red bar indicates ddI exposure phase and black bar indicates repopulation phase. The red lines indicate ddI exposed replicates and the black indicate untreated replicates.

5.4.4 Effect of NRTI exposure on the distribution of mtDNA deletion mutation in single cells

The heteroplasmy distribution of single cells isolated from three replicates in the trans-mitochondrial cybrid cell line experiment at experimental day 24, is graphed below (see Figure 5.18). Each point on the graph indicates a single biological replicate from one cell. The mean heteroplasmy of the pooled replicates for the untreated and ddI exposed group are shown in Figure 5.18.

The most striking feature of the single cell heteroplasmy levels in Figure 5.18 is the distribution difference between the untreated and the treated (ddI exposed) groups. All of the exposed to ddI have a very tight spread of data at the top end of the graph around the 90% heteroplasmy level.

The untreated conditions show a large spread of heteroplasmy levels across each single cell. Untreated replicates had a range of heteroplasmy levels across the single cells of 9.25% - 100%.

Statistical comparison of the two groups indicates that they're highly different (p<0.0001). The average heteroplasmy level of all the pooled single cells is 94.46%. The average heteroplasmy level of the untreated pooled single cells is 65.83%. These levels are comparable to the homogenate levels described in Section 5.4.3, previously.

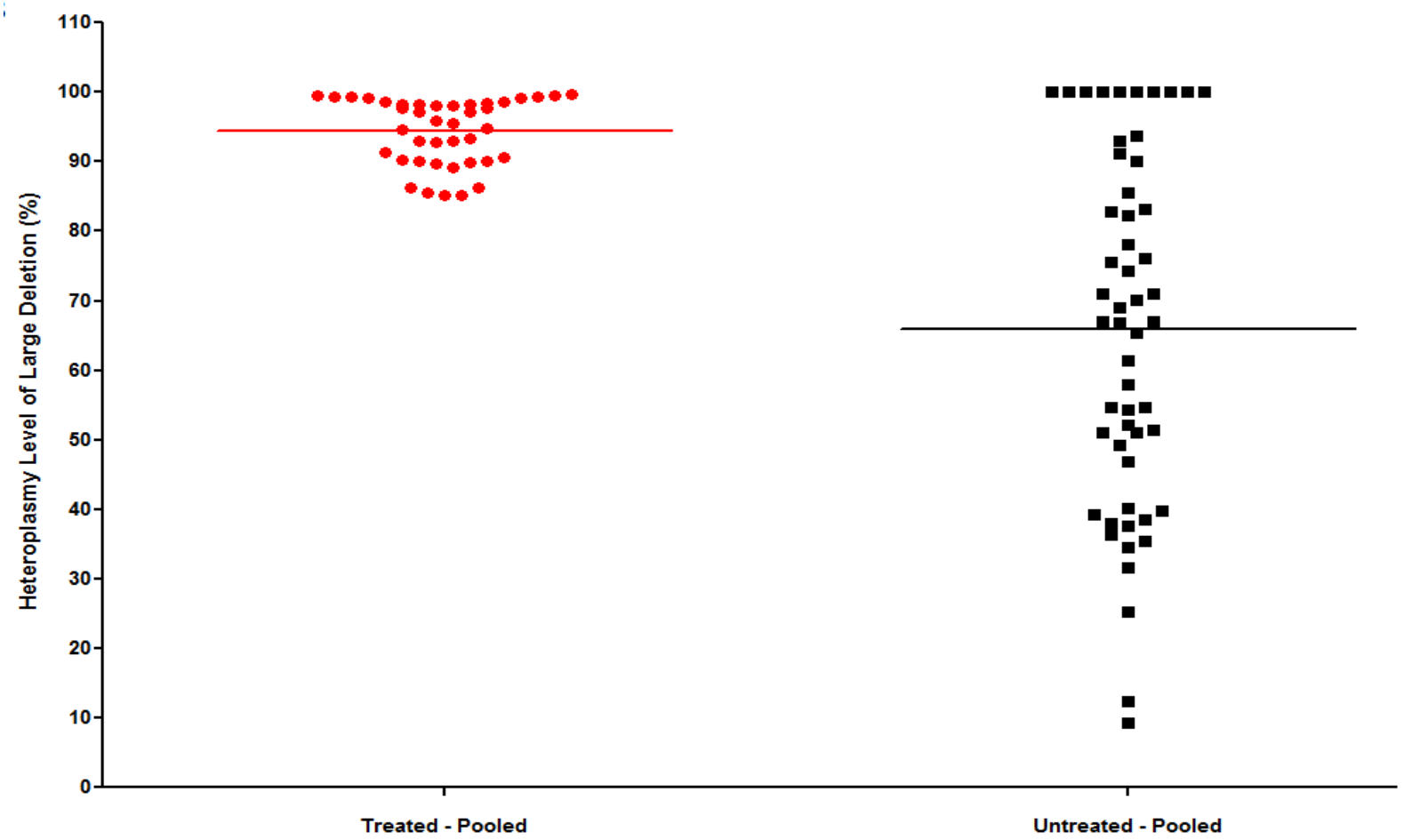

Figure 5.18 The heteroplasmy level distribution of the large scale deletion mutation within single cells extracted from replicates at experimental day 24. The pooled heteroplasmy distribution of single cells across all replicates for the untreated and ddI exposed group. Red indicate cell that have been exposed to ddI and black indicates untreated cells.

5.4.5 Dose-response effect on mtDNA content and large scale deletion

In order to examine the dose-dependent effects of NRTI exposure, 10x physiological dosing was studied. The total, mutant and wild type mitochondrial copy number (relative to the untreated) of each NRTI treatment group for the cybrid cells is displayed below (see Figure 5.19, Figure 5.20 and Figure 5.21). Each data point is expressed as the mean values derived from the biological and qPCR technical replicates, with error bars plotted as standard deviations (S.D).

The total copy number graph presented below (Figure 5.19), indicates severe depletion in the 10x ddI condition, evident from the first time point at experimental day 5, with 82% depletion present (0.18 ±S.D 0.00045) when compared to the untreated. Depletion continues throughout the exposure phase up to experimental day 32 with an experimental depletion maximum of 0.04 (±S.D 0.005). The depletion present in the exposure phase was found to be statistically significant when compared to the untreated with *p*-value <0.001, at experimental day 32. The mtDNA content was found to recover completely by the end of the post-exposure phase.

The total copy number during the d4T exposure phase slightly fluctuates around the comparable level to the untreated for the entire exposure and post-exposure phases with a level of 1.52 (±S.D 0.18) at experimental day 32, before ending at 1.07 (±S.D 0.155) by experimental day 52.

The AZT condition shows little depletion of total mitochondria during the exposure phase, but was found to be slightly higher than the untreated at experimental day 32 at 1.26 (±S.D 0.03) b. The total copy number is then seen to fluctuate a total relative copy number of 1.16 (±S.D 0.007) by the end of the experiment at experimental day 52.

The TDF condition was found to be comparable to the untreated condition throughout the experiment.

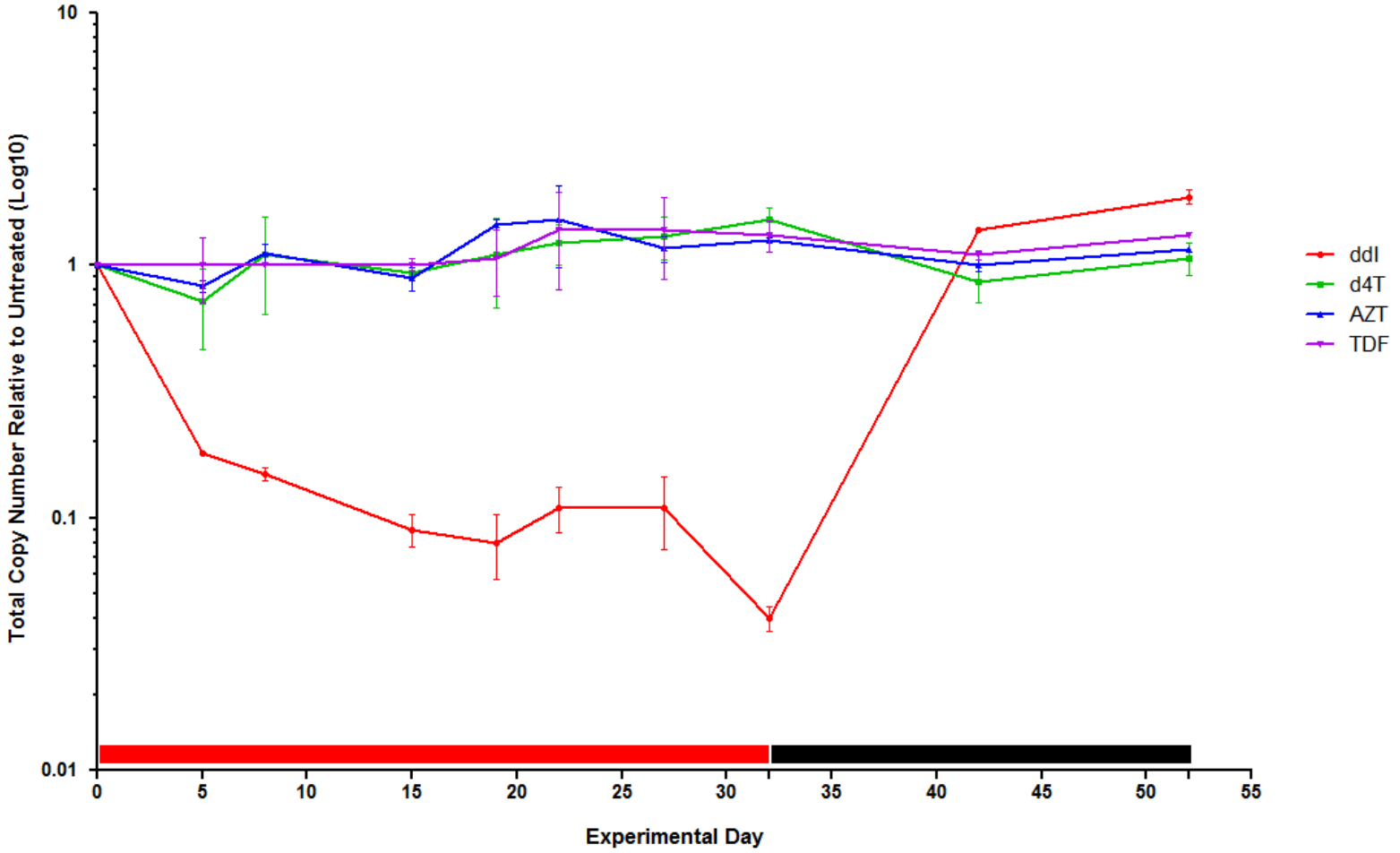

Figure 5.19 The total mitochondrial copy number of trans-mitochondrial cybrids during the NRTI exposure-repopulation, dose response assay. Red bar indicates exposure phase and black indicates untreated phase. Error bars are standard deviations of the mean.

The mutant copy number graph is presented below (Figure 5.20) and appears highly comparable to the total copy number graph (Figure 5.19). The most striking feature is the mtDNA depletion present in ddI-exposed cells.

The ddI condition was found to have the biggest depletion during the exposure phase with the first time point at experimental day 5, indicating 79% depletion compared to the untreated (0.21 ±S.D 0.005). The depletion continues to decrease throughout the exposure phase and reaches a depletion maximum value of 0.038 (±S.D 0.001). The exposure phase was found to be statistically significant compared to the untreated (p<0.001), at experimental day 32. The post-exposure phase highlights the mutant copy number increasing to almost 2-fold greater than the untreated by experimental day 42 (1.98 ±S.D 0.029) and almost 3-fold greater by the end of the experiment at experimental day 52 (2.83 ±S.D 0.176).

The d4T, AZT and TDF conditions were highly comparable to the total copy number at experimental day 32 and the end of the experiment at experimental day 52.

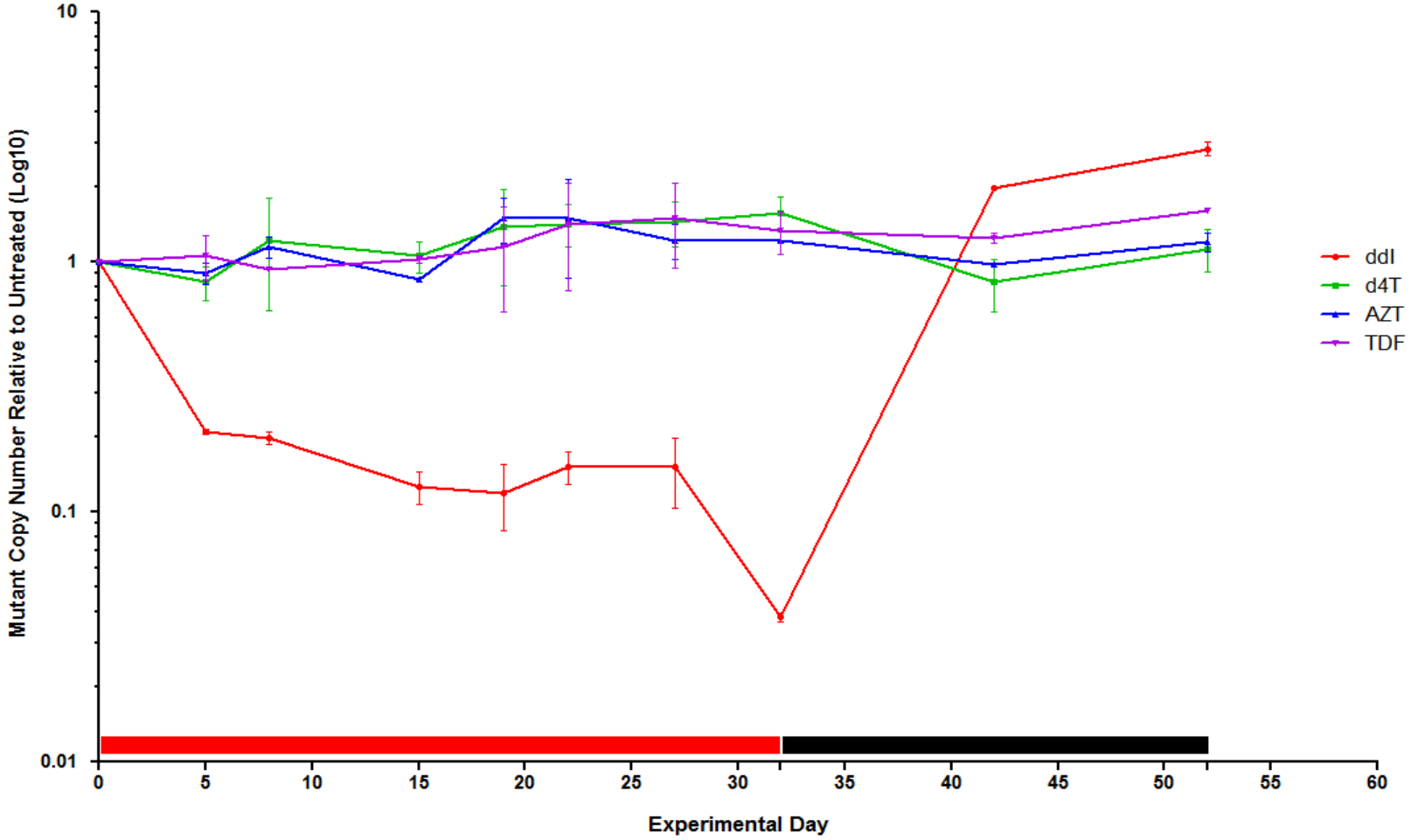

Figure 5.20 The mutant mitochondrial copy number of trans-mitochondrial cybrids during the NRTI exposure-repopulation, dose response assay. Red bar indicates exposure phase and black indicates untreated phase. Error bars are standard deviations of the mean.

The wild type copy number graph (Figure 5.21) is presented below. The most striking feature of the graph is the apparent permanent depletion in the ddI condition. The other conditions appear similar to the untreated throughout the entire experiment.

The ddI condition was found to severely and quickly deplete for both the total and mutant copy number; however, the wild type was affected the most. The experimental depletion maximum was found to be present at experimental day 19 with relative to untreated copy number ratio of 0.009 (±S.D 0.0009). At experimental day 32, end of exposure phase, the copy number was found to be 0.035 (±S.D 0.012). The ddI exposure wild type copy number data was found to be highly statistically significant when compared to the untreated at experimental day 32 ($p<0.001$). The wild type copy number barely repopulated during the repopulation phase increasing from 0.035 (±S.D 0.012) at experimental day 32 to 0.057 (±S.D 0.012) by experimental day 42. This appeared to plateau by experimental day 52 with only a minor increase to 0.058 (±S.D 0.012).

The wild type copy number in the d4T condition shows mild depletion during the majority of the exposure phase with experimental depletion maximum present at experimental day 19, at 0.63 (±S.D 0.09). The depletion severity reduces from this point throughout the rest of the exposure phase and reaches an experimental high of 1.28 (±S.D 0.04) at experimental day 32. This depletion was not found to be significant when compared to the untreated ($p=0.103$).

The wild type copy number in the AZT and TDF conditions were found to be barely affected during the exposure phase with an exposure phase. The wild type copy number is highly similar to the untreated during the whole of the experiment displaying minor fluctuations.

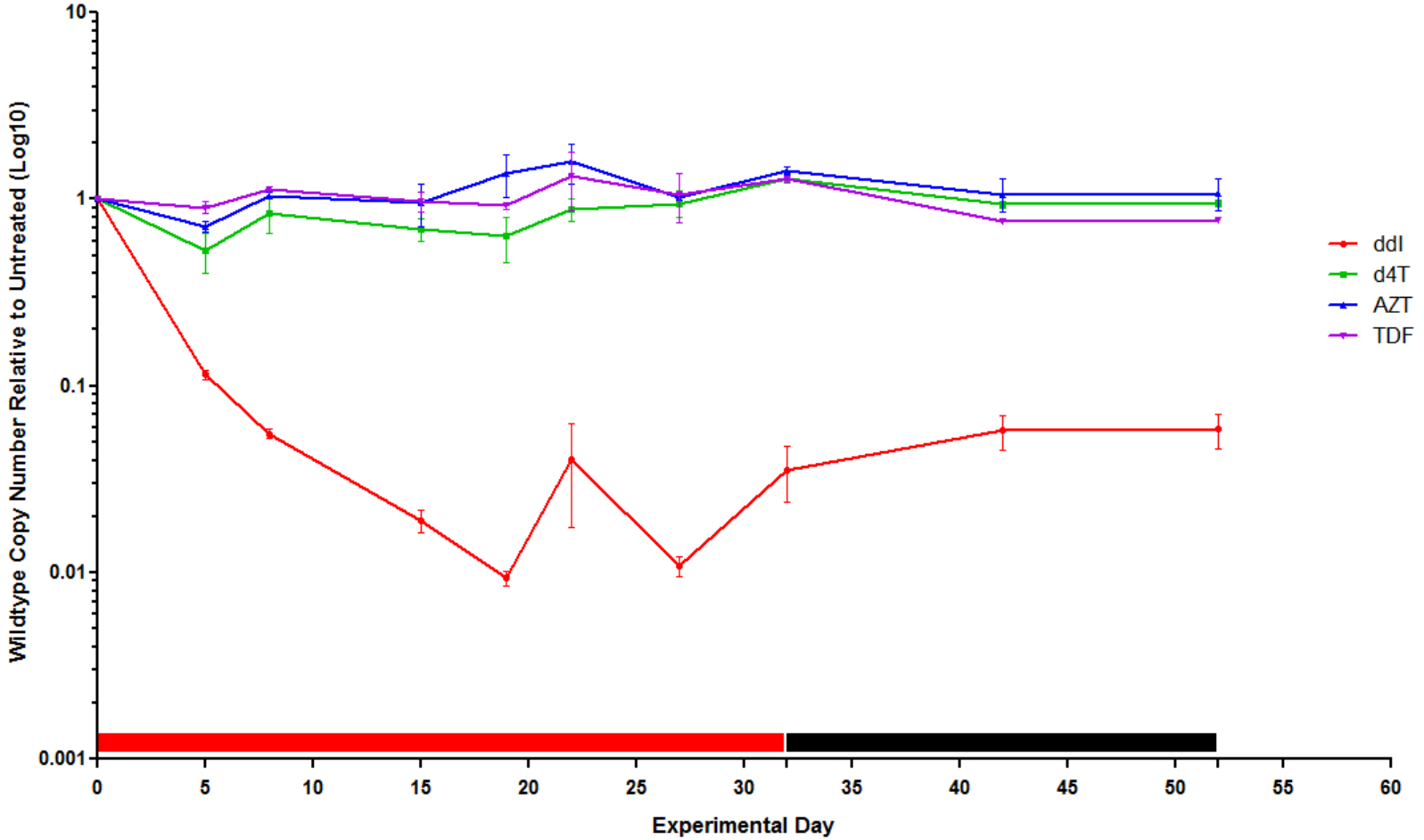

Figure 5.21 The wild type mitochondrial copy number of trans-mitochondrial cybrids during the NRTI exposure-repopulation, dose response assay. Red bar indicates exposure phase and black indicates untreated phase. Error bars are standard deviations of the mean.

The mitochondrial deletion heteroplasmy level has been graphed and expressed as deletion level percentage, for the cybrid cells during the NRTI exposure-repopulation dose response assay (Figure 5.22). Each data point is expressed as the mean value, derived from the biological replicates and qPCR technical replicates, with error bars plotted as standard deviations of the mean.

The untreated condition shows very little change in deletion level throughout the whole experiment ranging from 63.9% (±S.D 7.5%) to 82.3% (±S.D 1.2%).

The ddI condition markedly affects heteroplasmy levels during the exposure phase, with a rapid increase in heteroplasmy from 65.37% at experimental day 0, to 81.2% (±S.D 2.2%) at experimental day 5. The heteroplasmy level continues to increase throughout the exposure phase and reaching the experimental exposure phase maximum level of 97.4% (±S.D 0.5%) at experimental day 27. The heteroplasmy change was found to be highly statistically significant during the exposure phase when compared to the untreated at experimental day 32 ($p<0.001$) at experimental day 32. The post-exposure phase indicates very little change in heteroplasmy level; at experimental day 42, the heteroplasmy level was found to be 99.1% (±S.D 0.3%), which decreases to 98.9% (±S.D 0.2%) by experimental day 52.

The d4T condition was found to have heteroplasmy shifts present during the exposure phase reaching an experimental increase maximum of 89.1% (±S.D 4.3%) by experimental day 32. During the repopulation phase, the heteroplasmy level quickly reduces to 69.4% (±S.D 4.6%) at experimental day 42 68.7% (±S.D 3.7%) by the end of the repopulation phase at experimental day 52.

The AZT and TDF conditions remained comparable with the untreated condition.

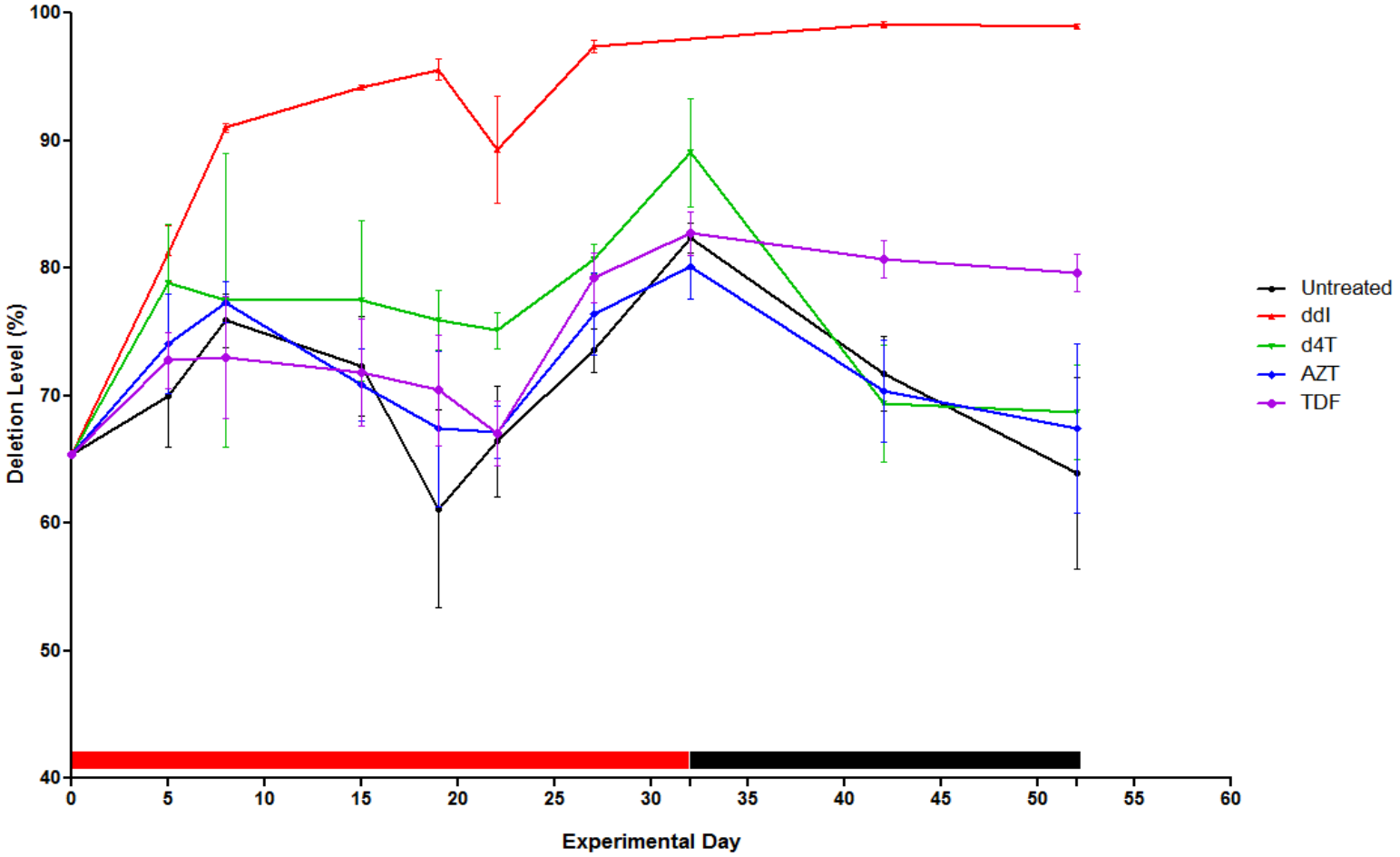

Figure 5.22 The heteroplasmy distribution of large scale deletion within cybrid cells during the NRTI exposure-repopulation dose response assay. Red bar indicates exposure phase, black bar indicates repopulation phase. Error bars are standard deviations of the mean.

5.4.6 Effect of NRTI-exposure on mtDNA content in embryonic rat neurons

The total mitochondrial copy number for E18 harvested rat neurons is graphed below (Figure 5.23) expressed as relative to the untreated condition. The data point is expressed as the mean of the technical and biological replicates (qPCR) and error bars plotted as standard deviations of the mean.

The physiological dosing of the NRTIs (Figure 5.23A) illustrates total copy number depletion is present (compared to the untreated) after 2 weeks exposure and 2 weeks repopulation of ddI, d4T and AZT to 0.43 (±S.D 0.07), 0.28 (±S.D 0.05) and 0.67 (±S.D 0.006), respectively. The TDF condition was found to be increased to 1.25 (±S.D 0.18).

It was found that in the ddI, d4T and AZT conditions, with 4 weeks physiological NRTI concentration exposure, depletion is almost half as severe than the 2 weeks exposure 2 weeks repopulation experiment; ddI was 0.8 (±S.D 0.003); d4T was found to be at 0.63 (±S.D 0.04) and AZT was found to be depleted to 0.92 (±S.D 0.36) compared to the untreated. The TDF condition was comparable to the untreated condition at 1.03 (±S.D 0.12).

All of the 10x physiological NRTI concentration conditions (see Figure 5.23B), at both 4 weeks exposure and after a 2 weeks exposure-repopulation, were found to be depleted when compared to the untreated. The TDF condition was found to be depleted the most after 2 weeks exposure-repopulation to a value of 0.41 (±S.D 0.007). The ddI condition was depleted to 0.5 (±S.D 0.23) compared to the untreated. D4T condition was found to be depleted to 0.53 (±S.D 0.15) and the AZT condition showed the least depletion at 0.75 (±S.D 0.01).

After 4 weeks exposure of the 10x physiological NRTI concentration exposure, ddI was found to be severely depleted to 0.09 (±S.D 0.002) relative to the untreated condition. The d4T condition was found to be depleted to half of the untreated, 0.5 (±S.D 0.022). The TDF condition was found to be depleted to 0.69 (±S.D 0.22) of the untreated and the AZT condition was depleted the least at 0.76 (±S.D 0.04), compared to the untreated condition.

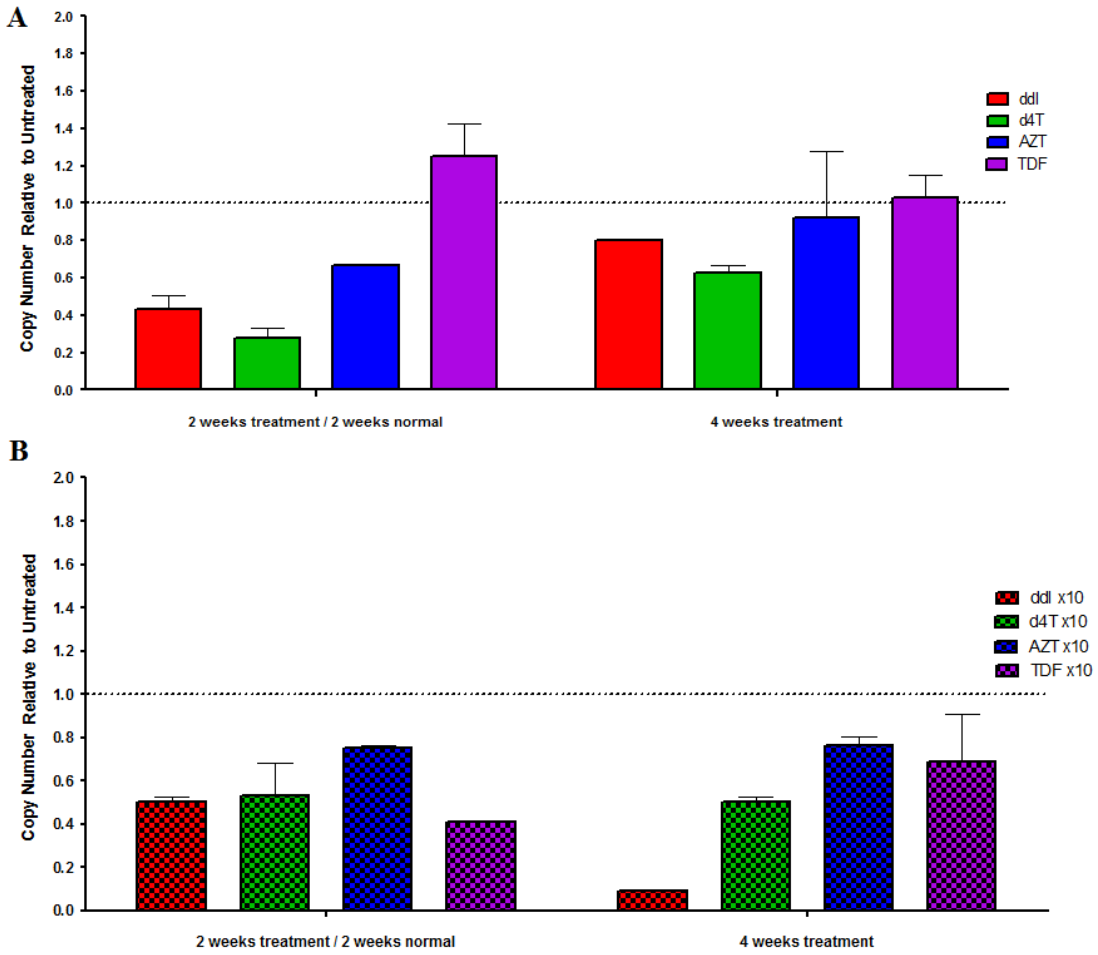

Figure 5.23 The total mitochondrial copy number, relative to untreated condition, of harvested E18 rat neurons at physiologicial (A) and 10x physiological dose (B) in NRTI exposure-repopualtion assay, and NRTI exposure only assay. The dotted line represents the relative untreated ratio level. Error bars are standard deviations of the mean.

5.4.7 Effect of NRTI exposure to individual E18 rat neurons

The total mitochondrial copy number for individual E18 rat neurons is graphed below (Figure 5.24 and Figure 5.23) expressed as absolute mitochondrial copy number. The data points are expressed as the mean of the technical replicates (qPCR) and error bars plotted as standard deviations of the mean. Only data from three embryos is presented as the neurons of the other conditions died before the end of the experiment.

Embryo 1 was found to display severe depletion in the ddI condition compared to the untreated condition with a depletion of 60.3%. The copy number of the ddI condition was found to be 36.84 copies per cell (±S.D 2.78) at day 17. The untreated condition for embryo 1 was found to be 92.78 copies per cell (±S.D 18.00).

Embryo 2 was found to display very little depletion when compared to the untreated condition with 94.82 copies per cell (±S.D 0.9) compared to 113.14 copies per cell (±S.D 6.80; illustrating 16.12% depletion between the treated condition and the untreated.

The ddI condition in Embryo 3 was found to be subject to 45% depletion with the ddI condition found to be 40.11 copies per cell (±S.D 1.02) compared to the untreated condition with 72.86 copies per cell (±S.D 7.89).

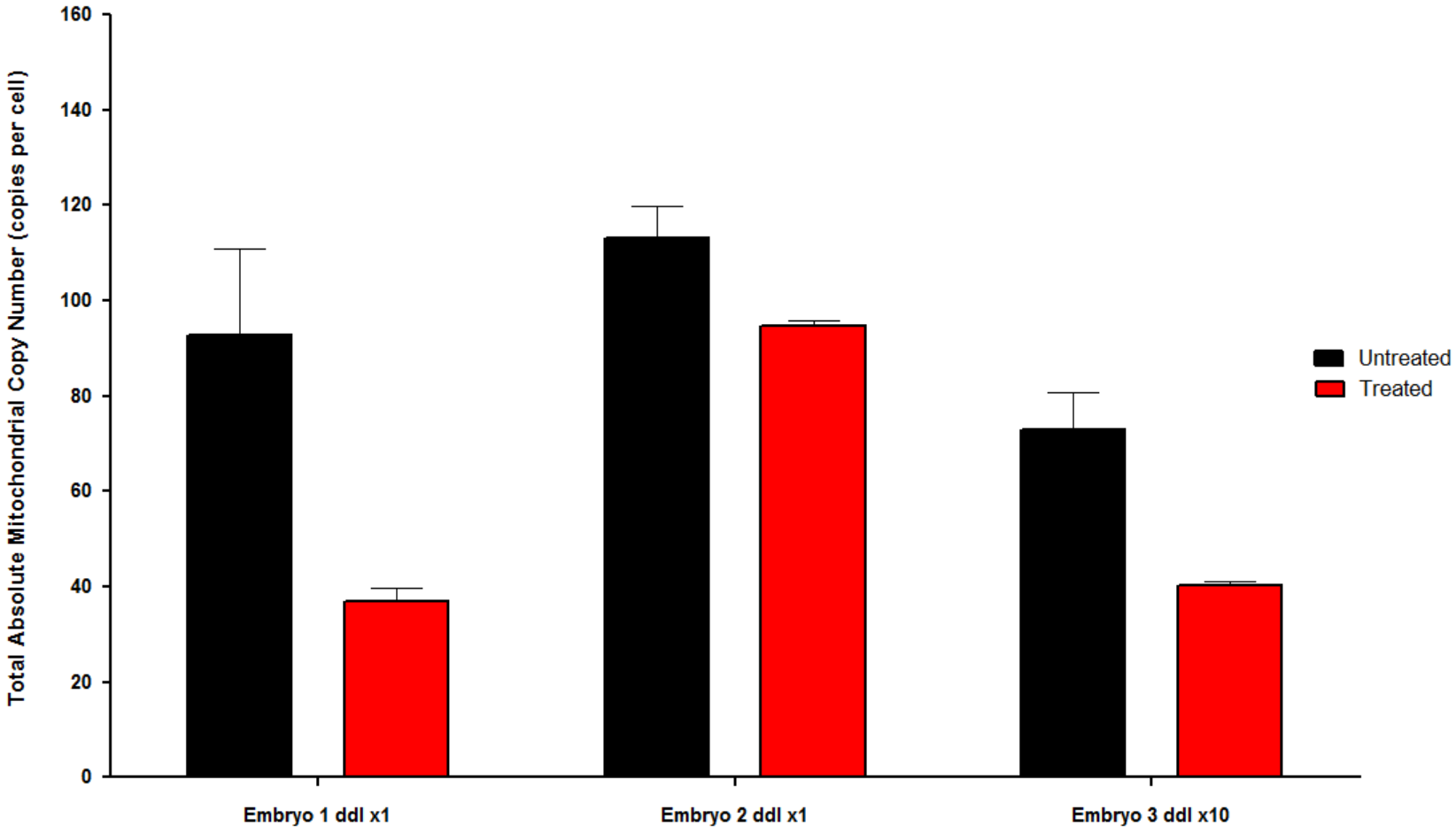

Figure 5.24 The Absolute mitochondrial copy number determined for individual E18 rat embryo neurons for the untreated and treated Section of the brain. Red bar indicates 17 days ddI exposure and black indicates 17 days untreated condition. Error bars are standard deviations of the mean.

5.4.8 Assessment of *de novo* mutagenesis

All of the aforementioned studies included a specific screening assessment for *de novo* mutagenesis of large scale deletions through NRTI exposure.

In the human cybrid samples, there was no presence of new deletions after any of the NRTI exposure conditions at a physiological or a 10x physiological dose (see Figure 5.25). The top 10Kb band is the undeleted amplification of 10Kb and the lower band is the amplification of the mtDNA in the deleted mtDNA molecules, missing approximately 4Kb of the region amplified.

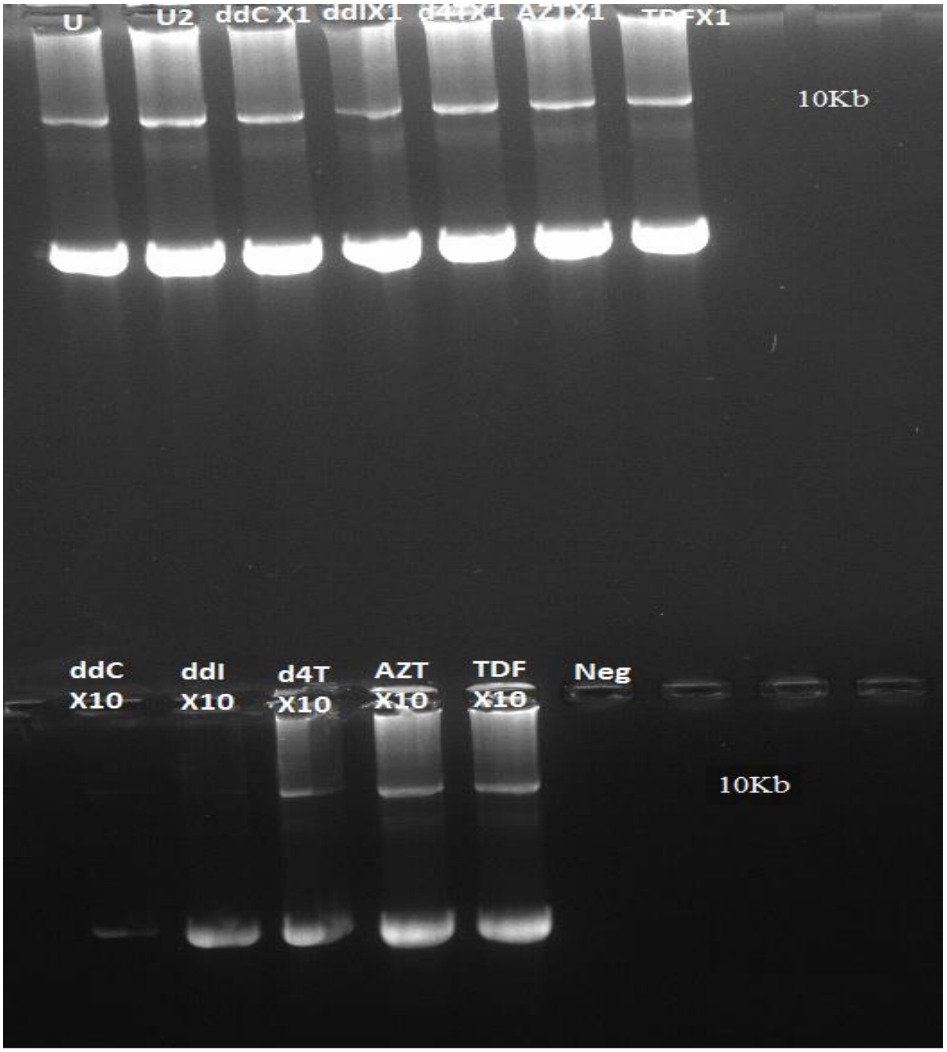

Figure 5.25 The agarose gel image from long range PCR of physiological and 10x physiological NRTIs doses at the end of the exposure phase showing no change from the untreated condition and no apparent new deletion formation.

The assessment of the mtDNA common deletion within the rat samples was performed using small amplicon qPCR. There was no change after any of the NRTI conditions at either of the concentrations used (see Figure 5.26).

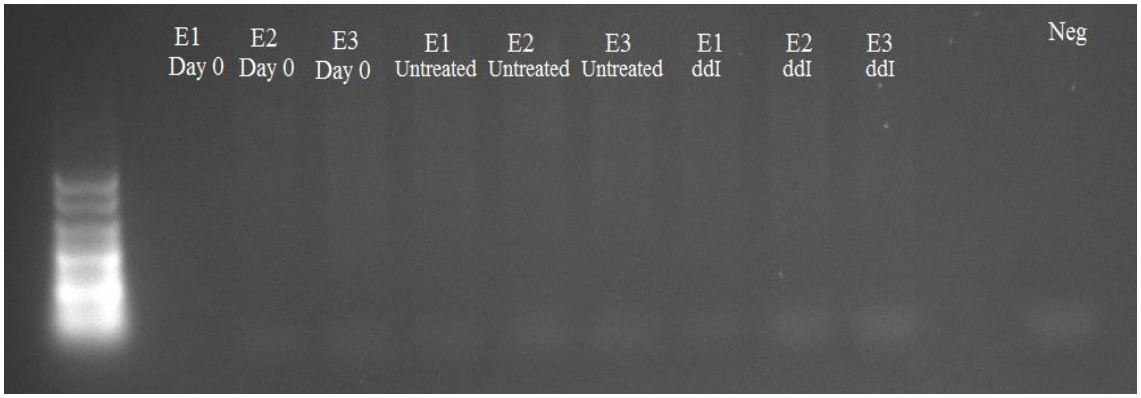

Figure 5.26 The small amplicon qPCR of rat common deletion the dosing of embryonic rat neurons harvested from individual embryonic brains showing no band; therefore, no common deletion before or after treatment with ddI.

5.5 Discussion

The aim of this experiment was to model the behaviour of a mitochondrial DNA deletion mutation in the presence of NRTIs (in multiple concentrations) to try and establish the mechanisms by which NRTI-exposure might affect mtDNA deletions, in order to increase the understanding of previous *in vivo* observations.

This was done by assessing three different scenarios:

- Is there a replicative advantage of the mutant mtDNA?
- Is there a molecular bottleneck mechanism?
- Is there presence of *de novo* mutagenesis?

5.5.1 Effect of NRTI exposure on mtDNA deletion mutation: data summary

The physiological concentration of NRTI exposure to the trans-mitochondrial cybrid cell line indicates that ddI was the only NRTI able to cause any statistically significant change to the deletion level and copy number. The clinically safe NRTI, TDF, showed no significant change in copy number or deletion heteroplasmy.

In Figure 5.8 there was observed to be a large distribution of copy number values at experimental day 0 across the conditions (see Figure 5.8). This is likely due to the nature of how the experiment was set up. One flask was split 1:15 to make all the biological conditions three days prior to experimental day 0 (five conditions in biological duplicate). A harsh split will lead temporarily to greater spread in observations; also, cell division may slow due to the density of cells being too low, but mtDNA replication continues, leading to temporarily elevated mtDNA CN per cell.

As this was the first experiment of the series of studies in this chapter, all other experiments were designed and experimentally set up with this artefact in mind and the required number of flasks was made more than three days in advance of experimental day 0, to allow copy number to stabilise.

The copy number graphs indicate that for both total and mutant, the ddI condition is slightly lower than the other conditions; however, the most apparent effect is seen in the wild type, where there is significant depletion present. This is the first indicator that there's a preferential inhibition of the wild-type molecule replication. This decrease in wild type compared to the untreated condition was found to be statistically significant ($p<0.001$).

The mutant molecule is approximately 46% smaller than the wild type and the replication time of a full sized mtDNA molecule has been estimated to be 1-2 hours (Davis and Clayton, 1996). Based on these data, almost twice as many deleted mtDNA molecules could potentially replicate than wild type molecules in a given time period. In normal treatment conditions, this time advantage doesn't appear to make any difference, as seen with the stable turnover in the untreated condition and the constant heteroplasmy level. However, in the presence of polymerase γ inhibition, where there's an apparent 'brake' applied to replication, this size advantage is unmasked.

Despite known polymerase γ inhibition by d4T exposure and at lower concentrations than used here *in vitro* (Stankov *et al.*, 2010), there was no severe depletion present in this study with a depletion maximum of 31%. This wasn't sufficient to cause an evident shift in heteroplasmy level either indicating that *in vitro* large depletion is required to cause a shift that would have physiological consequence. This is perhaps a consequence of the cybrids being more resistant to depletion than other cell lines, such as primary cells.

There was no *de novo* mutagenesis detected with no change in the number of presence of large scale mtDNA deletions after any of the NRTI exposures.

5.5.2 Mitochondrial network fragmentation analysis data summary

The mitochondrial network fragmentation analysis revealed that all NRTI conditions appeared to induce a decrease in the number of networks present after 27 days of exposure, although arguably within the variability of the untreated cells. Indicating all NRTIs despite a varying affinity for polymerase γ, are capable of causing mitochondrial network fragmentation. This phenomenon has been described with the exposure of protease inhibitors but not NRTIs (Roumier *et al.*, 2005). However, there's unpublished evidence from *Hans van der Speks'* laboratory in the Netherlands, suggesting thymidine analogues, such as d4T and AZT, can increase mitochondrial network size. The assessment of mitochondrial network length, revealed there was a statistically significant increase in the length of mitochondrial networks in the d4T condition after 27 days exposure (Table 5.4), supporting the data found in the Netherlands.

5.5.3 Assessment molecular bottleneck: data summary

Given that ddI exposure induces significant mtDNA depletion, we speculated that the mtDNA deletion mutations might be subject to a molecular bottleneck effect. This

possibility was assessed by implementing a large number of biological replicates (n=10) to assess how the deletion behaves after a strong bottleneck has been applied through depletion of copy number. A 14 day ddI exposure was previously found to cause sufficient depletion; in the case of a bottleneck mechanism driving the increased heteroplasmy of the deletion in all replicates, it would be expected that once ddI was removed, each replicate would repopulate and have a different heteroplasmy level depending upon the make-up of the genetic population created through the bottleneck/depletion. The repopulation phase illustrated that all replicates behave in the exact same manner as they come out of the bottleneck. This strongly excludes a bottleneck mechanism and adds support to the notion of replicative advantage.

In case the study of multiple experimental replicates was insufficient to capture a change in spread of deletion levels indicative of a molecular bottleneck effect, therefore a single cell analysis was also performed. Analysis of the untreated condition revealed a wide spread of heteroplasmy levels (7% - 100%) between individual cells. In contrast the ddI-exposed cells showed an overwhelming population of cells with high level heteroplasmy level of >95%, comparable with data found in the homogenate analysis of the flasks, which was found to be highly statistically significant (p0.0001). These data highlight that:

- The untreated cybrid cell line is a very heterogeneous population of cells (despite the very steady heteroplasmy level found in the homogenates)
- The overall increase in heteroplasmy level in the ddI treatment may be accounted for by very large shifts within cells that previously had very low heteroplasmy levels.
- The combination of the strength of this (selective) effect of ddI treatment, and the initial wide spread in untreated, would completely preclude detecting an increased spread in the ddI-treated.
- These data suggest there is not a bottleneck effect for deletions - or at least that its completely overwhelmed by the selective effect.

5.5.4 NRTI exposure-repopulation dose response: data summary

The supra-physiological concentration of 10x physiological concentration was used in the trans-mitochondrial cybrids to assess for a dose dependent response across the NRTIs.

The most severely affected was the ddI condition with rapid depletion present, 82% within the first 5 experimental days. This depletion was evident in both the mutant and wild type copy number data suggesting that at 10x concentration, the replicative advantage mechanism present at the physiological concentration, is not sufficient to protect even the smaller molecule from severe depletion. The copy number does recover during the repopulation phase of the experiment, albeit only really the mutant molecule. The fact that mutant copy number quickly recovers suggests that there's a replicative advantage. This is also supported by the fact that the heteroplasmy doesn't really decrease during repopulation.

The repopulation phase was increased to allow for a greater assessment after the severe depletion induced by the higher concentration of the drugs. Despite doubling the repopulation phase time to 20 experimental days, it was not sufficient for wild type copy number to repopulate enough to cause a 'downward' shift (decrease) in the heteroplasmy level of the deletion. *In vivo*, this may translate as an irreversible cellular defect. This data is supported by the findings of a previous study using similar trans-mitochondrial cybrids where they were exposed to ethidium bromide (a highly potent compound in inducing mtDNA depletion) and then allowed to recover. The findings revealed that the smaller deleted molecule repopulated faster than the wild type and to a level, comparable to the experimental starting mtDNA level (Diaz *et al.*, 2002).

The d4T condition illustrates a classical dose response effect. There was increased depletion of the mutant copy number and an increased heteroplasmy level during the exposure phase. Although it was required to increase the physiological concentration by 10-fold, a longer exposure than the relatively short 32 days of a physiological concentration may be sufficient to elicit a comparable response to that seen with ddI. Therefore d4T may have less relevant effect on a deletion mutation *in vivo.* The decrease of copy number was more evident in the wild type than the mutant (37% depletion maximum vs. 17.1% maximal depletion, respectively), further supporting clonal expansion through a size dependant advantage.

A supra-physiological dose would perhaps be expected to induce a certain level of depletion, given the depletion present at the physiological concentration of 41% at experimental day 22 (Figure 5.8). The fact there's an opposite effect, a potentially inverse dose response, is strikingly similar to the previous observations of another mitochondrial safe NRTI, abacavir (ABC), also increasing mtDNA copy number (Bulst

et al., 2012). However, there's evidence to suggest that TDF actually induces subclinical mitochondrial toxicity in patients receiving the NRTI, especially when compared to patients receiving ABC rather than TDF (Maggi *et al.*, 2012). However, there's no *in vivo* data to support these findings of NRTI up regulating mitochondrial copy number and given the fact that the heteroplasmy level wasn't affected, there's no real clinical importance (in regards to mitochondrial phenotypes) in this finding either. It could be proposed that a patient, who is still receiving an NRTI which is known to cause depletion, could be co-supplement with ABC or TDF, in an attempt to reverse or reduce any depletion.

It is also noted that the previously described hierarchy of inhibition by NRTI's *in vitro* was almost perfectly depict in each of the experiments with ddI causing the greatest depletion, followed by d4T, AZT and finally TDF. The effects were not confined to one cell type with the effects seen in both mitotic and post-mitotic cell types.

Finally, there was no *de novo* mutagenesis detected with no change in the number of presence of large scale mtDNA deletions after any of the NRTI exposures (even at 10x dose).

5.5.5 Embryonic rat neuron exposure to NRTIs: data summary

Two experiments using E18 rat neurons were used to assess the behaviour of mitochondria in post-mitotic cells through NRTI exposure.

The hierarchy of depletion was visible in the harvested cells using the physiological concentration and 10x physiological concentration (Figure 5.23). The ddI exposure caused the greatest copy number depletion; however, interestingly, the TDF condition increased in copy number (as previously discussed in 5.5.4, above). The results were not as dramatic as expected using a cell type containing very low mtDNA copy number (< 100 copies per cell), although it is unclear the exact turnover of mtDNA within these cells.

When the rat embryos were segregated and neurons extracted from individual brains, rather than pooling all of them as above, ddI exposure with only the physiological concentration was found to cause greater depletion within 17 days, compared to the 10x physiological dose.

There were no *de novo* deletions detected in any of the conditions, despite previous reports of detectable levels of rat mtDNA 'common deletion' in normal embryonic brain and young (3 months) liver samples (Petruzzella *et al.*, 1992; Nicklas *et al.*, 2004). Although it doesn't serve as a good model for studying deletion behaviour, it does reinforce the idea of clonal expansion of pre-existing mutations rather than inducing new mutations.

5.5.6 Limitations

The experiments presented within this chapter also carried recognised limitations. The use of a trans-mitochondrial cybrid does not strictly simulate 'normal' cells and does leave a gap in the understanding of how the effects seen *in vitro* would translate *in vivo*. However, the cell line implemented was done so, on the grounds that it is the only plausible method of successfully studying deletions *in vitro*.

The E18 rat neurons gave a unique *in vitro* insight into post-mitotic behaviour in the presence of NRTIs; however, given the low copy number and nature of the cells, it may be seen as a poor model to modelling the behaviour of deletion mutation(s) in post-mitotic cells as results between replicates appeared very variable biologically.

5.6 Chapter conclusion

The analyses of this chapter have strongly indicated:

- That the expansion of a pre-existing deletion mutation during the presence of pol γ inhibiting NRTIs is driven by a size dependent advantage mechanism.
- The effect is very dependent upon the strength of the polymerase γ inhibition and the dose of the NRTI.
- There is no evidence for *de novo* deletion mutagenesis through NRTI exposure.

The overall chapter conclusion is that in the presence of partial polymerase γ inhibition *in vivo*, an mtDNA deletion may clonally expand through a greater chance of replication.

Chapter 6. Modelling the Behaviour of Mitochondrial DNA Point Mutations in the Presence of NRTI's

Table of Contents

6.1 Background

Point mutations in mtDNA accumulate in a variety of tissues during the process of ageing in an uneven manner across cells (Pallotti *et al.*, 1996). This principle has been best described in the polymerase γ defector mouse model where mtDNA mutations were found to accumulate at a rapid rate due to a defective polymerase γ, and the mouse presented with premature ageing phenotypes (Trifunovic *et al.*, 2004). Nonetheless, it neatly illustrates the importance of point mutation and ageing.

There have been reports of specific point mutations accumulating with age in certain 'hot spot' regions; the mitochondrial non-coding region, the D-loop, has been found to be susceptible to point mutations with age (Kennedy *et al.*, 2013). Specifically, m.189A>G and m.414T>G, with the latter associated with age in a multiple of tissues (Michikawa *et al.*, 1999; Del Bo *et al.*, 2003).

An accumulation of mtDNA point mutations within individual cells in HIV-infected individuals receiving NRTI therapy has been previously described (Payne *et al.*, 2011). Given the association of D-loop mutations with normal ageing and the nature of it being a 'hot spot' for accumulating somatic mutations, it is therefore important to determine whether or not NRTI therapy is mutagenic, or whether it imposes an accelerated clonal expansion of point mtDNA mutations.

Clonal expansion has been best studied in deletion mutations with a range of postulated theories (see Section 1.2.9); however, there are studies that have tried to unify one theory of clonal expansion for both types of mutation (Coller *et al.*, 2002). The theory of replicative advantage is potentially plausible on the basis of a point mutation changing the regulatory region of mtDNA, as opposed to a size dependant driving mechanism. However, neutral drift is generally accepted as a more plausible mechanism for point mutations.

Due to the reduction of mitochondrial copy number through NRTI exposure, another potential mechanism behind the shifts in point mutation clonal expansion maybe explained by a genetic bottleneck creating segregation of mutant molecules, as described in embryogenesis through the mtDNA copy number reduction in oocytes (Cree *et al.*, 2008). This theory has been modelled for a period of depletion/repopulation due to NRTI exposure over many years in humans (Payne *et al.*, 2011). In contrast, it may in fact be that NRTIs are mutagenic for mtDNA point mutations, as previous

studies have suggested using human, mouse and *in vitro* models (Martin *et al.*, 2003; DaleM Walker *et al.*, 2004; Y. Zhang *et al.*, 2014).

6.2 Experimental Aims

The aims of the experiments were to assess:

- Whether there is evidence of mutagenesis for mtDNA point mutations through NRTI exposure.
- Whether point mutation heteroplasmy shift through NRTI exposure
- If any shift is present, whether this is mediated through a molecular bottleneck effect.
- Whether point mutations behave in a manner suggestive of a replicative advantage (as seen with deletions) during NRTI exposure.

6.3 Experimental design and methods

6.3.1 Fibroblast culture

Two human fibroblast cell lines, derived from skin biopsies of aged patients (80 & 91 years old; cell lines nr68 and nr100, respectively) harbouring a point mutation within the control, D-loop, region of the mitochondrial DNA (m.414T>G; Dr G Saretzki, Newcastle University, UK) were used, with a previously visualised genotype of approximately 50% heteroplasmy using sanger sequencing (Dr Angela Pyle, IGM, Newcastle University, UK). Cells were grown in cell culture using high glucose DMEM (Gibco, Life-Technologies, Paisley, UK) supplemented with: 10% fetal calf serum (Sigma-Aldrich, Dorset, UK) and 5% penicillin/streptomycin (Sigma-Aldrich, Dorset, UK) as previously described (see Section 3.1.2). The cells were cultured in T_{25} flasks (Greiner Bio-one, Stonehouse, UK) at 37^{o}C and 5% CO_2 to a confluency of 70-90% before splitting and propagating (see Section 3.1.3).

At the point of splitting and propagating the cells during each study, the residual cells were pelleted by centrifugation at 8,000rpm for 5 minutes, the supernatant removed and DNA extracted, as previously described (3.3). Each time point was then subject to mitochondrial copy number determination (see Section 6.3.3), *de novo* deletion assessment (see Section 6.3.4) and m.414T>G heteroplasmy quantification (see Section 6.3.5).

6.3.2 NRTI procedure

The NRTI exposure-repopulation assays were performed using the NRTIs and concentrations as previously described (see Section 3.1.5) with the additional inclusion of zalcitabine (ddC; Sigma-Aldrich, Dorset, UK) at 1µM concentration.

Fibroblasts were grown from one inital vial, thawed from liquid nitrogen storage (see Section 3.1.4) to ensure all flasks and conditions throughout each study were from the same origin.

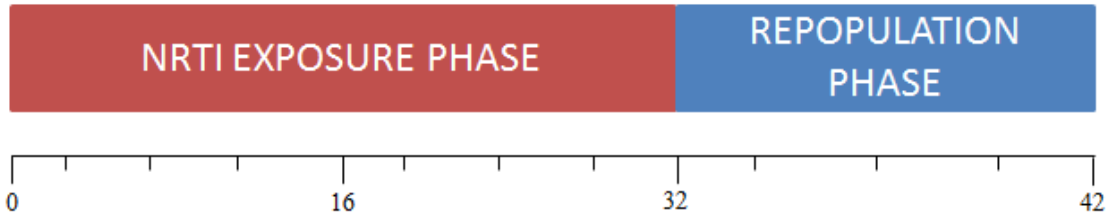

Figure 6.1 NRTI exposure-repopulation procedure for assessment of m.414T>G behaviour during and after NRTI physiological treatment

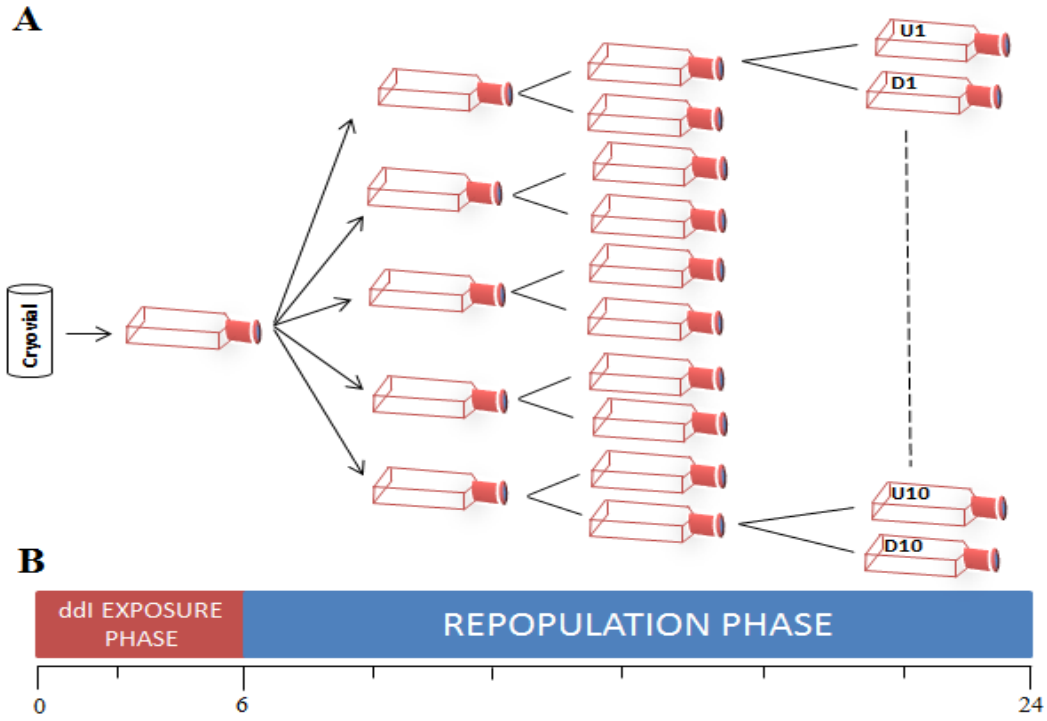

Figure 6.2 The ddI exposure-repopulation assay procedure for investigation of a bottleneck mechanism of a point mutaion through ddI exposure. Derviation procedure of untreated replicates (U1-10) and ddI exposed replicates (D1-10) is shown in A, with the time assay time course illustrated in B.

6.3.3 Mitochondrial copy number assessment

In order to investigate the mtDNA total copy number levels, multiplex qPCR was performed, as previously described (see Section 3.6.1); however, the assay was run as a duplex assay to determine mitochondrial copy number (*MT-ND1* and *β2M* – see Table 3.3). Running conditions and reaction make up remained coherent with that previously described (see Section 3.6.1). The concentration of each DNA sample was standardised and approximately 50ng was loaded into each qPCR 25µL reaction. Each sample was run in duplicate on the reaction plate.

Total mitochondrial copy number was calculated using the SQ values of each sample as determined using the standard curves (see Section 3.6.3 for equation). The mitochondrial copy number was then expressed as a ratio relative to the untreated condition at that time point. Therefore, a relative copy number ratio of one indicates that the condition shows the same mtDNA copy number level per cell as the untreated conditions.

6.3.4 Long range PCR determination of *de novo* deletions

To investigate whether or not NRTI exposure induces deletion mutations in the fibroblast cell lines exposed to NRTIs, in comparison to the untreated conditions, a long range PCR was used to screen for mtDNA deletions as previously described (see Section 3.7).

The PCR products were run on a 1% agarose gel, as previously described (see Section 3.5) using the DNA ladder 1kb plus (Thermo-Scientific, Hampshire, UK) to determine the size of the bands and predict an approximate size of any deletion(s) present in each sample. The technique is used as a qualitative method rather than a quantitative.

6.3.5 Point mutation heteroplasmy assessment

In order to investigate the effects of NRTI exposure on the heteroplasmy level of the d-loop, m.414T>G mutation, pyrosequencing was performed as previously described (see Section 3.9.2).

6.3.6 Single cell analysis

In order to investigate the distribution of the m.414T>G mutation of the samples used in the ddI-exposure repopulation study, single cell analysis was performed as previously described (see Section 3.1.7) and single cells isolated by the use of laser microdissection (see Section 3.2.1) from untreated and treated replicates one to five.

Approximately 20 cells were captured and lysed (see Sections 3.2.2) for each replicate assessed at experimental day 24 in the study. 5µL of the lysate was loaded into the downstream molecular analyses (m.414T>G quantification by pyrosequencing, see Section 6.3.5, above, increasing cycles from 30-40 for the primary PCR). The data was compared on a single replicate basis and then pooled to express the data as an overall perspective of mutation level behaviour.

6.4 Results

6.4.1 Effect of NRTI exposure on mtDNA copy number

The mitochondrial DNA copy number (ratio relative to the untreated) of each NRTI treatment group for cell lines nr100 and nr68 are displayed below (see Figure 6.3 and Figure 6.4). Each data point is expressed as the mean values derived from the biological and qPCR technical replicates with error bars plotted as standard deviations (S.D). Relative copy number is expressed on a logarithmic scale to increase resolution at the lower end. Due to the slow growth of the cells, there was no time point taken between experimental day 22 and 42.

During NRTI exposure of ddC and ddI in both cell lines, a significant depletion in total mitochondrial copy number was present throughout the exposure phase which didn't repopulate in either of the cell lines once the drugs were removed. These findings were calculated to be statistically significant by means of a 2-tailed student *t*-test (see Table 6.1).

ddC caused the greatest copy number depletion during exposure with a exposure phase maximum depletion level of 0.03 (±S.D 0.003) in cell line nr68 by day 14 and 0.039 (±S.D 0.03) by experimental day 22 in cell nr100. The repopulation phase didn't show any increase in copy number in either cell line by experimental day 42, with both reaching experimental maximum depletion of 0.02 (±S.D 0.001) in nr68 and 0.007 (±S.D 0.00001) in nr100.

ddI had the second greatest effect on copy number in both cell lines. The exposure phase maximum depletion level was found to be 0.12 (±S.D 0.049) by experimental day 14 in cell line nr68; maximum depletion level was found to be 0.05 (±S.D 0.06) by experimental day 22 in cell line nr100. The repopulation phase of both cell lines after ddI exposure behaved in the same manner as that seen in ddC, with no apparent repopulation present and both cell lines reaching the experimental maximum depletion level of 0.01 (±S.D 0.003) by experimental day 42 in cell line nr68 and 0.03 (±S.D 0.0005) in cell line nr100.

None of the other NRTI treatment groups: d4T, AZT and TDF, produced any significant depletion in mitochondrial copy number during exposure (see Table 6.1). It is noted however, in both cells lines, d4T, AZT and TDF treatment groups, there was an increase in relative copy number at experimental day 42 and the effect was found to be more

pronounced in cell line nr100. TDF caused the greatest increase with a 3.3-fold increase when compared to the untreated at experimental day 42 (see Figure 6.3) in nr100 and a 1.4-fold increase compared to the untreated in cell line nr68.

Condition	**Nr100 ratio to untreated**	***p*-value nr100 vs. Untreated**	**Nr68 ratio to untreated**	***p*-value nr68 vs. Untreated**
ddC	0.0067	<0.001	0.02	<0.001
ddI	0.03	<0.001	0.01	<0.001
d4T	3.219	0.6105	1.77	0.7622
AZT	2.94	0.3922	1.34	0.8472
TDF	3.25	0.3100	1.374	0.7887

Table 6.1 *p*-values from student t-test comparison of total mitochondrial copy number of all NRTI conditions compared to the untreated in cell lines nr68 and nr100 at experiemntal day 42 along with the corresponding ratio relative to untreated copy number level; 95% confidence interval wa used. Significant values a highlighted in red.

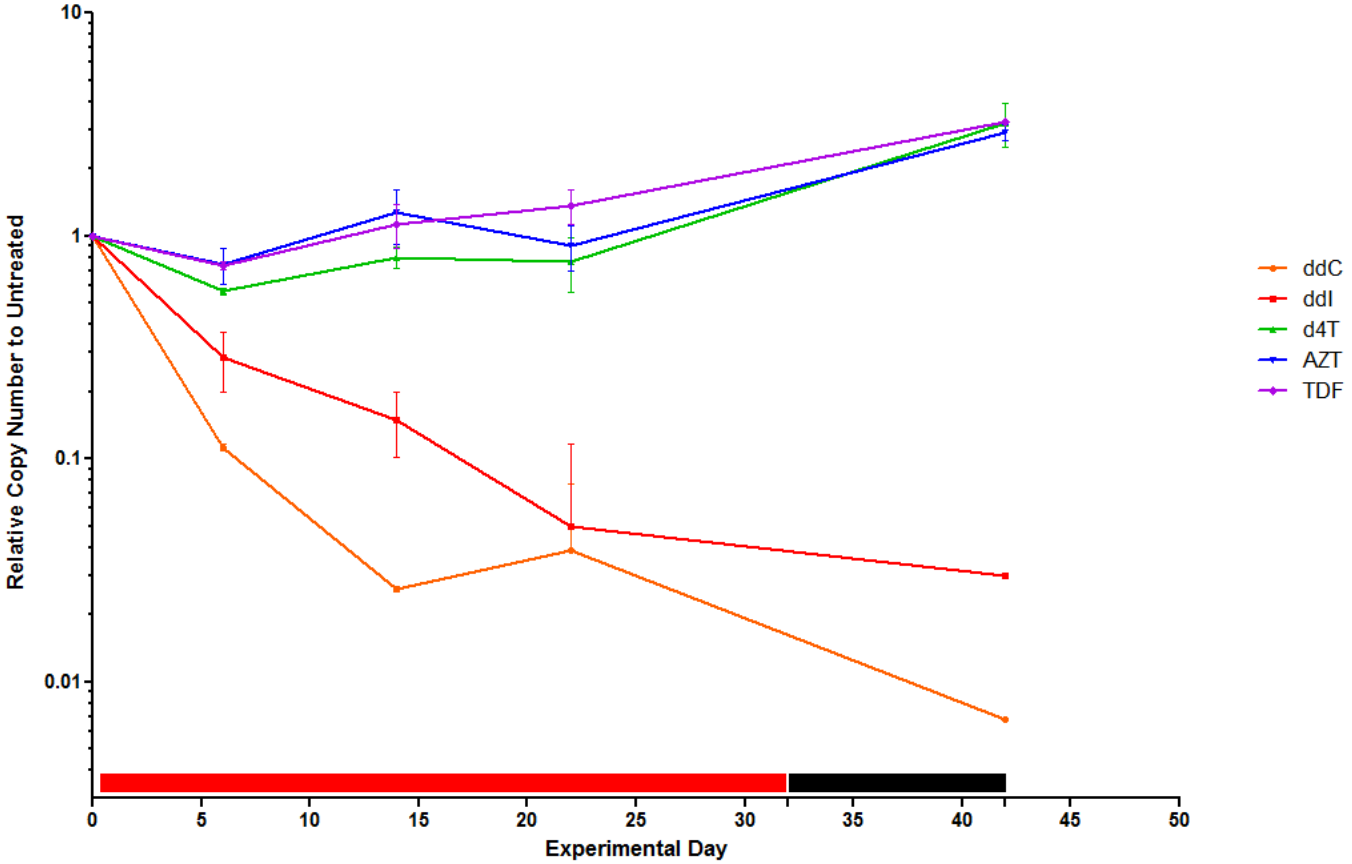

Figure 6.3 Mitochondrial copy number graph expressed as relative to untreated condition for the point mutation fibroblast cell line, nr100. Red bar indicates exposure time and black indicates repopulation. Error bars are standard deviations of the mean.

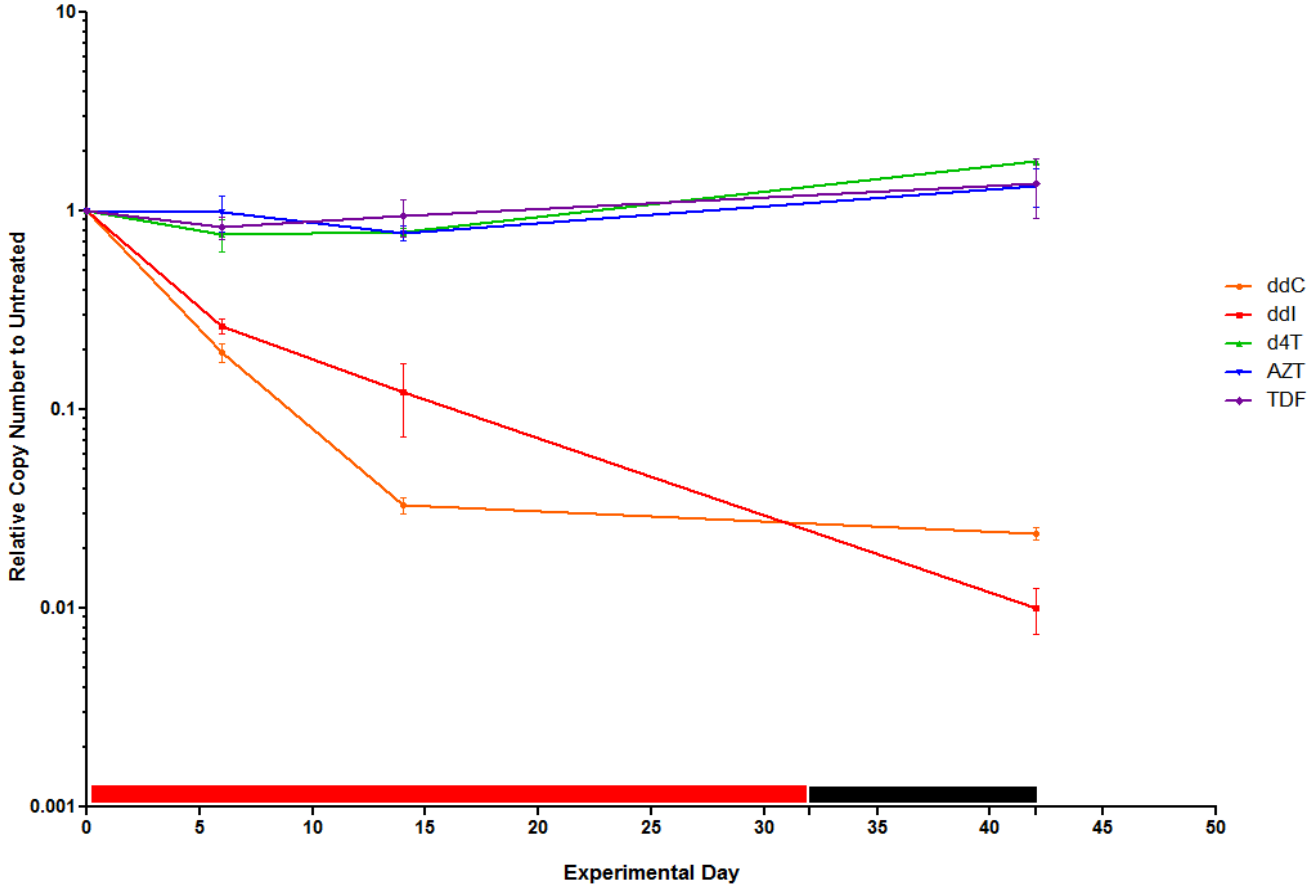

Figure 6.4 The mitochondrial copy number expressed as relative to the untreated group for point mutation fibroblast cell line, nr68. Red bar indicates exposure time and black indicates repopulation. Error bars are standard deviations of the mean.

6.4.2 Effects of NRTI exposure on point mutation level

The mitochondrial non-coding region point mutation m.414T>G heteroplasmy level has been graphed and expressed as mutant percentage level for both cell lines: nr100 and nr68 (Figure 6.5 and Figure 6.6, respectively). Each data point is expressed as the mean values from the biological and pyrosequencing technical replicates with error bars plotted as standard deviations. Due to the slow growth of the cells, there was no time point taken between experimental day 22 and 42.

During NRTI exposure phase, there was no apparent change in overall heteroplasmy level of the point mutation in any of the conditions, in either of the cell lines at experimental day 42 (see Table 6.2).

Condition	**Nr100 Heteroplasmy**	***p*-value nr100 vs. Untreated**	**Nr68 Heteroplasmy**	***p*-value nr68 vs. Untreated**
Untreated	53%		35.5%	
ddC	35.5%	0.7352	42%	0.4226
ddI	46.5%	0.2965	27%	0.4024
d4T	43.5%	0.7818	29%	0.5918
AZT	39%	0.9235	35%	0.9175
TDF	26.5%	0.1299	31.5%	0.3333

Table 6.2 The *p*-values from 2 way student t-test of m.414T>G mutant heteroplasmy level during NRTI exposure-repopulation experiment at experimental day 42, compared to the untreated condition.

The untreated cell line shows minor fluctuation during the NRTI exposure phase and highlights that the variation present in treatment groups is in line with the untreated condition. At experimental day 42, the heteroplasmy of the untreated was found to be slightly increased in cell line nr100 to 53% (±S.D 8.48%). Conversely, in cell line nr68 at experimental day 42, the heteroplasmy was found to have decreased with a heteroplasmy of 35.5% (±S.D 2.12%).

Despite the very significant mtDNA depletion (see Figure 6.3 and Figure 6.4), neither of the ddC or ddI conditions shows any effect on heteroplasmy level in either of the cell lines during the NRTI exposure phase. However, some effect is seen at the end of the repopulation (experimental day 42). The ddC condition can be seen to shift down in heteroplasmy in cell line nr100, with a heteroplasmy level of 35.5% (±S.D 3.53%); whilst in cell line nr68, the ddC heteroplasmy level at experimental day 42 is found to be almost exactly the same as the starting level (44%) at 42% (±S.D 9.89).

The d4T condition shows little change during the exposure phase in either cell line. The biggest change was present at the end of the repopulation phase with cell line nr100 decreasing to 43.5% (±S.D 4.95%) and nr68 decreasing to 29% (±S.D 2.82).

The AZT condition behaves in a similar manner to that seen in the d4T condition with minor fluctuations during the exposure phase in both cell lines. The biggest change is an overall downward trend in the repopulation phase with both nr100 and n68 decreasing at experimental day 42 to a heteroplasmy level of 39% (±S.D 0%) and 35% (±S.D 2.82%), respectively.

The TDF condition shows little change during the exposure phase compared to the untreated in either cell line. The condition was however, found to decrease the most of all treatment conditions in the nr100 cell line to a heteroplasmy level of 26.5% (±S.D 4.95%) at experimental day 42. Cell line nr68 was found to also decrease in heteroplasmy level to 31.5% (±S.D 0.7%) at experimental day 42.

The key information of this data is when the trend of the heteroplasmy data for both cell lines is assessed (Figure 6.5 and Figure 6.6). It is evident that there's a large spread across the conditions in which the greatest effect can be seen at experimental day 42, where all NRTI conditions can also be seen to have a widespread distribution between replicates i.e. large S.D error bars between biological replicates.

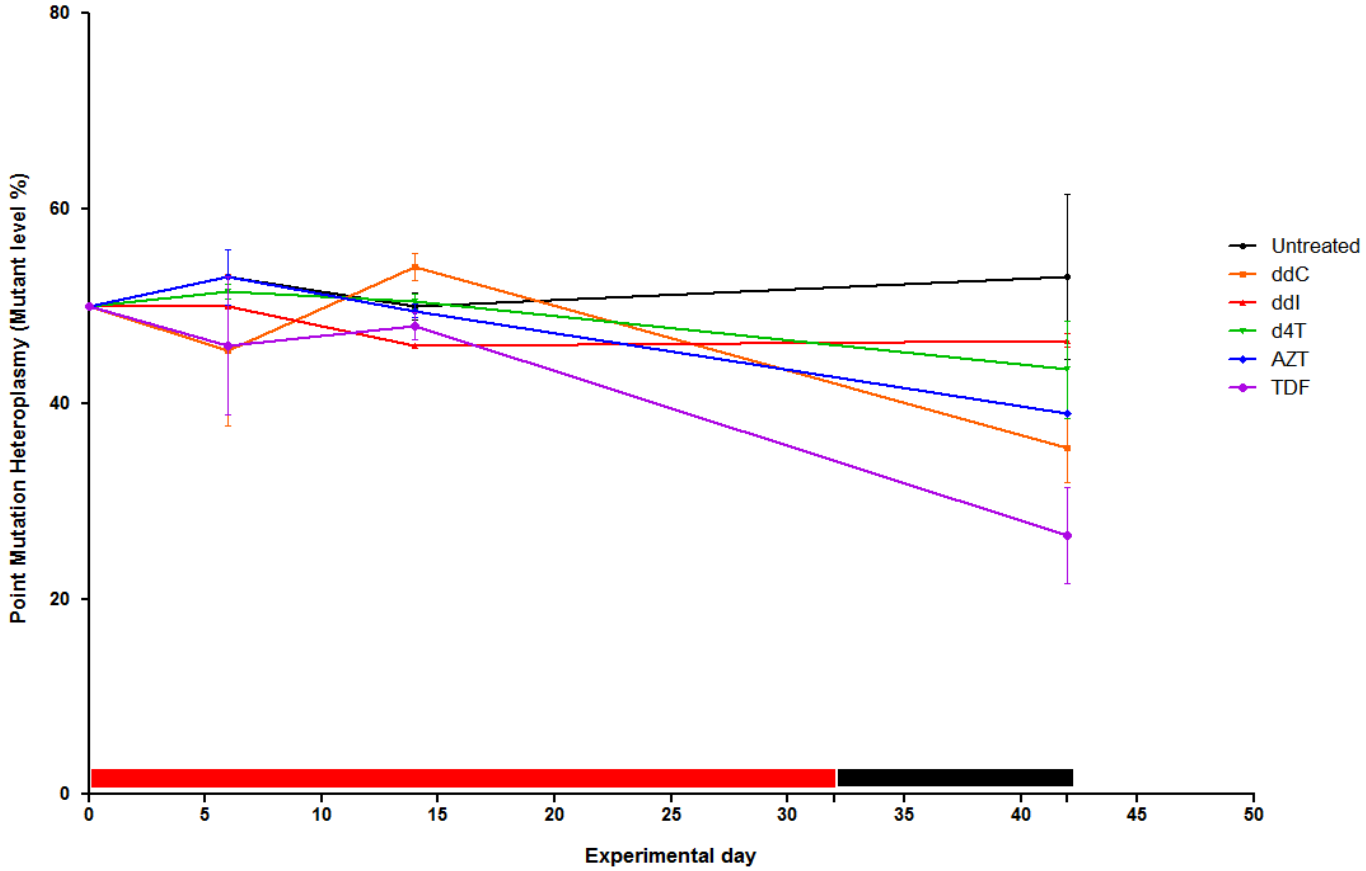

Figure 6.5 m.414T>G mutation level expressed as percentage as mutant percentage (G) determined by pyrosequencing in cell line nr100, during physiological exposure of NRTIs. Red bar indicates exposure time and black indicates repopulation. Error bars are standard deviations of the mean.

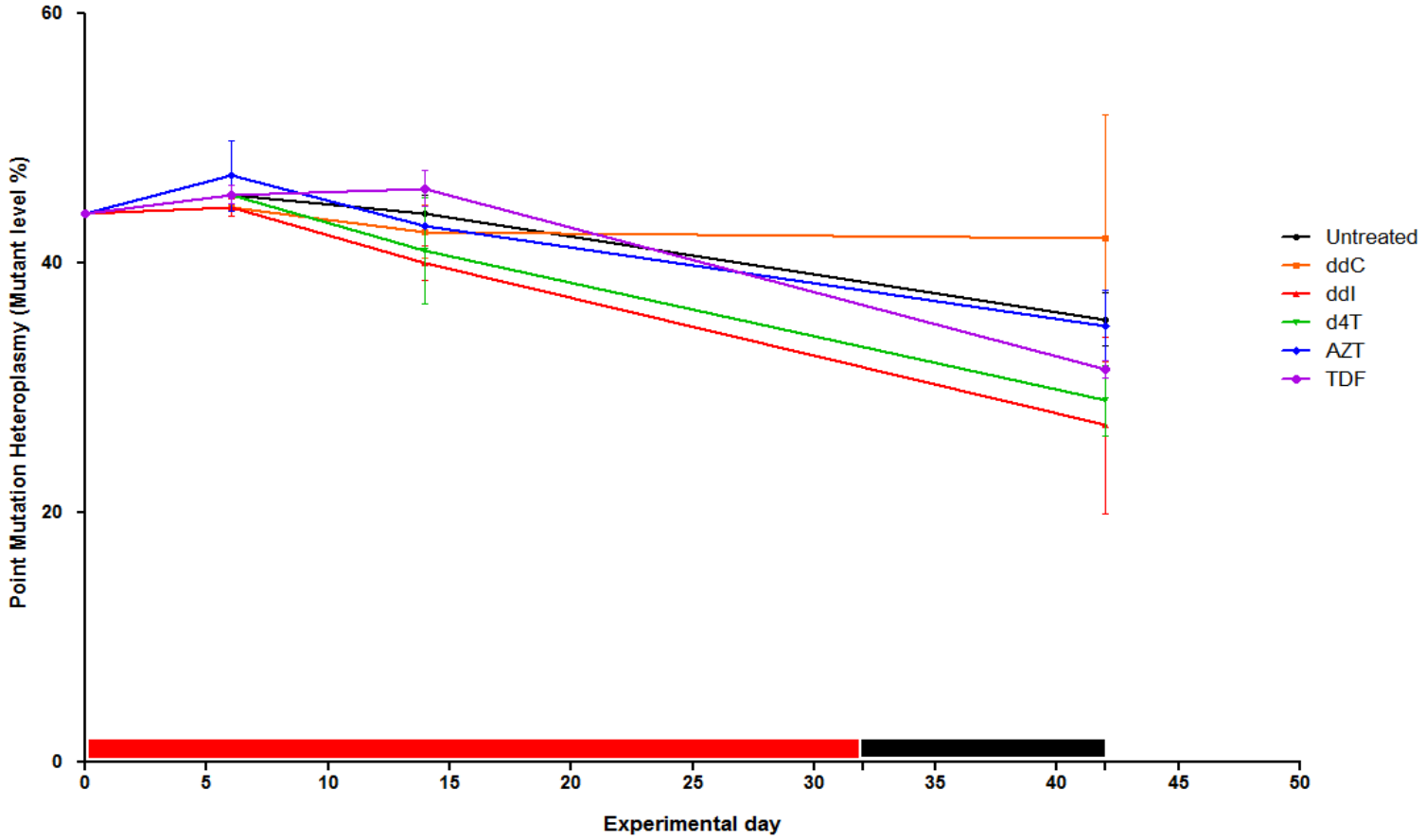

Figure 6.6 m.414T>G mutation level expressed as mutant percentage (G) determined by pyrosequencing in cell line nr68, during physiological exposure of NRTIs. Red bar indicates exposure time and black indicates repopulation. Error bars are standard deviations of the mean.

6.4.3 Assessment of the presence of a molecular bottleneck

The mitochondrial copy number of each replicate (10 biological replicates) in the ddI exposure-repopulation study to assess for a genetic bottleneck effect in the mitochondrial point mutation m.414T>G nr100 cell line is displayed below (see Figure 6.7). Each data point is expressed as the mean values derived from the qPCR technical replicates with error bars plotted as standard deviations (S.D). Six days of ddI was chosen due to the lack of repopulation found in the data presented in Section 6.4.1, after 32 days exposure.

The untreated replicates all have a mean starting value of 31.06 mtDNA copies per cell (±S.D 6.95) and with a range of 22-47 copies per cell across the 10 replicates. All replicates share the same fluctuating trend throughout the entire experiment. The overall distribution of the variation between untreated replicates throughout the experiment was found not to be significant by means of ANOVA analysis (*p*-value=0.962).

The 10 ddI exposed samples share a similar spread of copy number at experimental day 0 as the untreated with a mean copy number per cell as 31 (±S.D 5.94) with replicate spread found to be 22-42 copies per cell. All replicates exposed to ddI show severe depletion during the experiment with the copy number in each replicate decreasing by the end of ddI exposure at day 6 to 9 copies per cell (± S.D 1.07), with a range of 8-11 copies per cell. The overall decreasing trend slowly continues into the recovery phase where mean copy number was found to be 6 copies per cell (±S.D 1.67), with a range of 4-8 copies per cell. All replicates appear to repopulate quickly over the following 10 days to reach an mean copy number of 21 copies per cell (±S.D 3.33) by experimental day 24.

The difference in copy number per cell at both experimental day 6 and experimental day 14 in the ddI treated cells was found to be highly statistically significant when compared to the untreated (*t*-test; p-value<0.001).

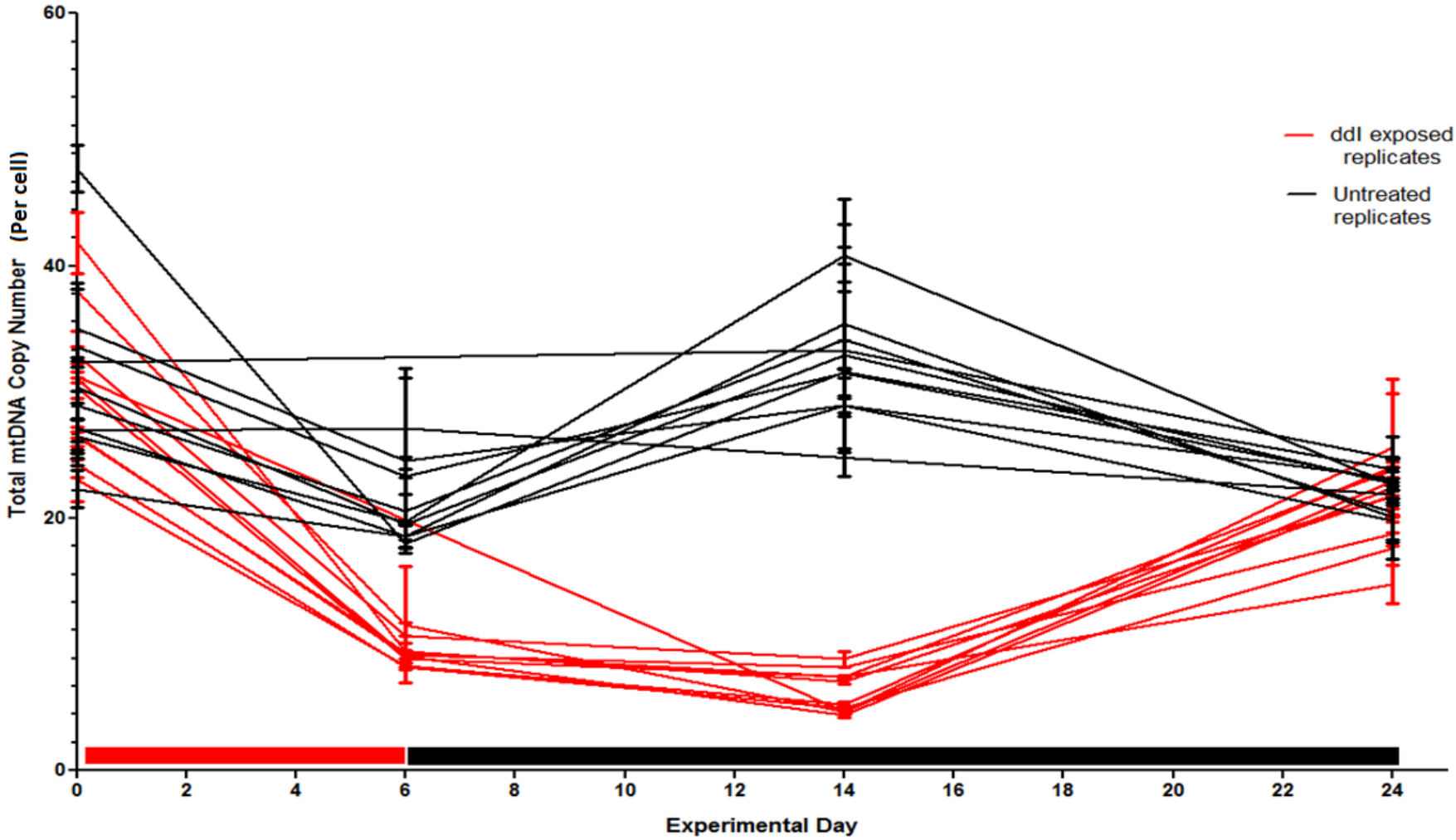

Figure 6.7 Total mitochondrial copy number for assessment of genetic bottleneck mechanism through ddI exposure. Red lines bar indicates ddI exposure phase and black bar indicates repopulation phase. Error bars are standard deviations of the mean.

The mitochondrial non-coding region point mutation m.414T>G heteroplasmy of each replicate (10 biological replicates) in the ddI exposure-repopulation study to assess for a genetic bottleneck effect in the mitochondrial point mutation m.414T>G nr100 cell line is displayed below (see Figure 6.8). Each data point is expressed as the mean values from the pyrosequencing technical replicates with error bars plotted as standard deviations.

The untreated replicates have a very tight spread of heteroplasmy levels across replicates with average heteroplasmy 24.7% (±S.D 1.7%) at experimental day 0. Overall, the untreated replicates show a slight upward trend throughout the experiment with a tight data spread amongst the replicates at each time point. The average heteroplasmy increased to 27.3% (±S.D 1.8%) by the end of the exposure phase at experimental day 6 and further increased to 33.7% (±S.D 1.0%) by experimental day 14. The maximum average heteroplasmy seen in replicates was found at the end of the experiment, experimental day 24, with heteroplasmy level of 37.3% (±S.D 1.). Due to the consistent increasing heteroplasmy trend of the untreated, the distribution of the untreated replicates across the time course was found to be borderline significant when the trend analysed using 2-sample Kolmogorov-Smirnov (K-S) test, to assess for normality.

The ddI treated replicates also share a tight spread of heteroplasmy distribution at experimental day 0, with a mean heteroplasmy of 24.7% (±S.D 1.3%). The spread of replicates ranged from 23% - 27%. The heteroplasmy remained largely unchanged by the end of ddI exposure at experimental day 6 (24.9% ±S.D 1.7%). The most noticeable change in heteroplasmy distribution was found at experimental day 14, where average heteroplasmy across replicates remained similar to experimental day 6 at 22.2% (±S.D 5.6%); however, the spread increased to 15.5% across the replicates with a range of 12.5% - 28% across the 10 replicates. The heteroplasmy spread then reduced with an average heteroplasmy level was found to be at a similar level to the untreated at day 24. The data distribution was found to be not normally distributed with a highly statistically significant K-S p-value < 0.001, when comparing experimental day 14 to experimental day 6 amongst all of the ddI treated conditions.

The heteroplasmy levels of all ddI replicates were found to be statistically significant at experimental day 14, when compared to the level of all untreated replicates (p-value < 0.001; Mann-Whitney test chosen due to data distribution not deemed as normal from K-S test).

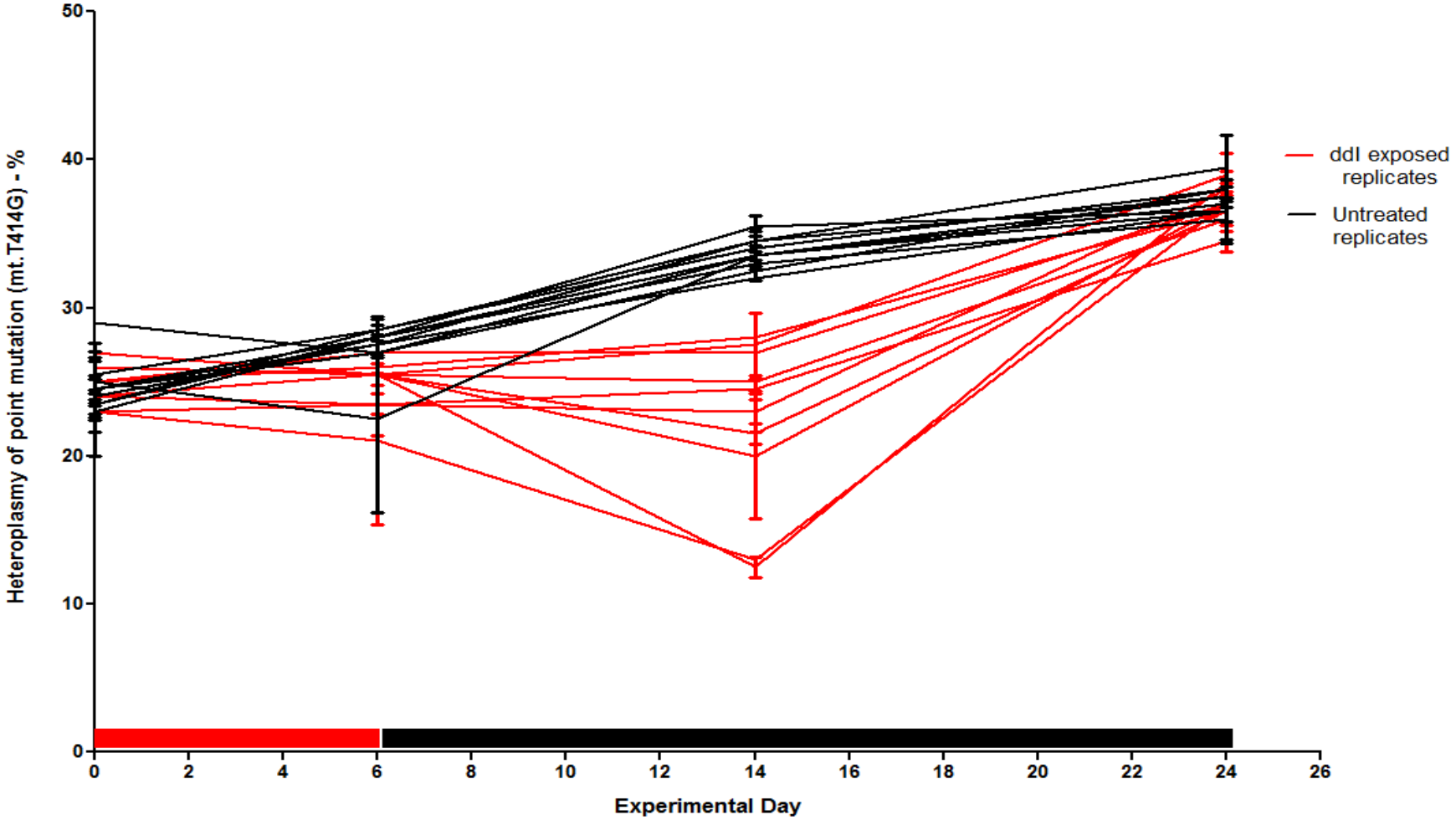

Figure 6.8 The heteroplasmy level distribution of the 414T>G mutant level during the assessment of a genetic bottleneck mechanism through ddI exposure. Red bar indicates ddI exposure phase and the black bar indicates repopulation phase. Error bars are standard deviations of the mean.

6.4.4 Single cell analyses

The heteroplasmy distribution of single cells isolated from replicates one to five at experimental day 24, in the assessment of mitochondrial genetic bottleneck through ddI exposure in nr100 point mutation fibroblast cell line experiment is graphed below (see Figure 6.9). Each point on the graph indicates a single replicate from one cell of the corresponding experimental biological replicate. The mean heteroplasmy of the single cells for all replicate (pooled) are highlighted in Figure 6.9; where n=100 cells from five biological replicates in the NRTI-exposed and untreated conditions.

The most striking feature of the single cell heteroplasmy levels is the distribution among the single cells (Figure 6.9). In total (untreated and ddI treated), there was a range of 5%-86% amongst all of the cells. The average of the single cells for the condition and replicate was found to be comparable to that found in the homogenate replicate values (Table 6.3).

Comparing the overall distribution of heteroplasmy levels of pooled single cells within replicates from ddI exposed to untreated (Figure 6.9B) indicates there's no difference in distribution with a *p*-value of 0.28 (*t*-test, two-tailed). Single cell averages were found to be 26.6% and 29.9% for the untreated and ddI exposed groups, respectively, with a range of 4% - 82% for the untreated replicates and 5% - 87% for the treated replicates.

Replicate	Mean Untreated (%)	SD Untreated (%)	Mean ddI (%)	SD ddI (%)
1	39.5	2.1	37.0	1.4
2	37.0	0.0	36.5	0.7
3	36.5	0.7	39.0	1.4
4	37.5	0.7	38.0	0.0
5	38.0	0.0	36.0	0.0
Pooled	**37.7**	**0.7**	**37.3**	**0.7**

Table 6.3 The heteroplasmy of m.414T>G of each replicate used in the single cell analysis from experimental day 24 of the assessment of mitochondrial genetic bottleneck through ddI exposure experiment.

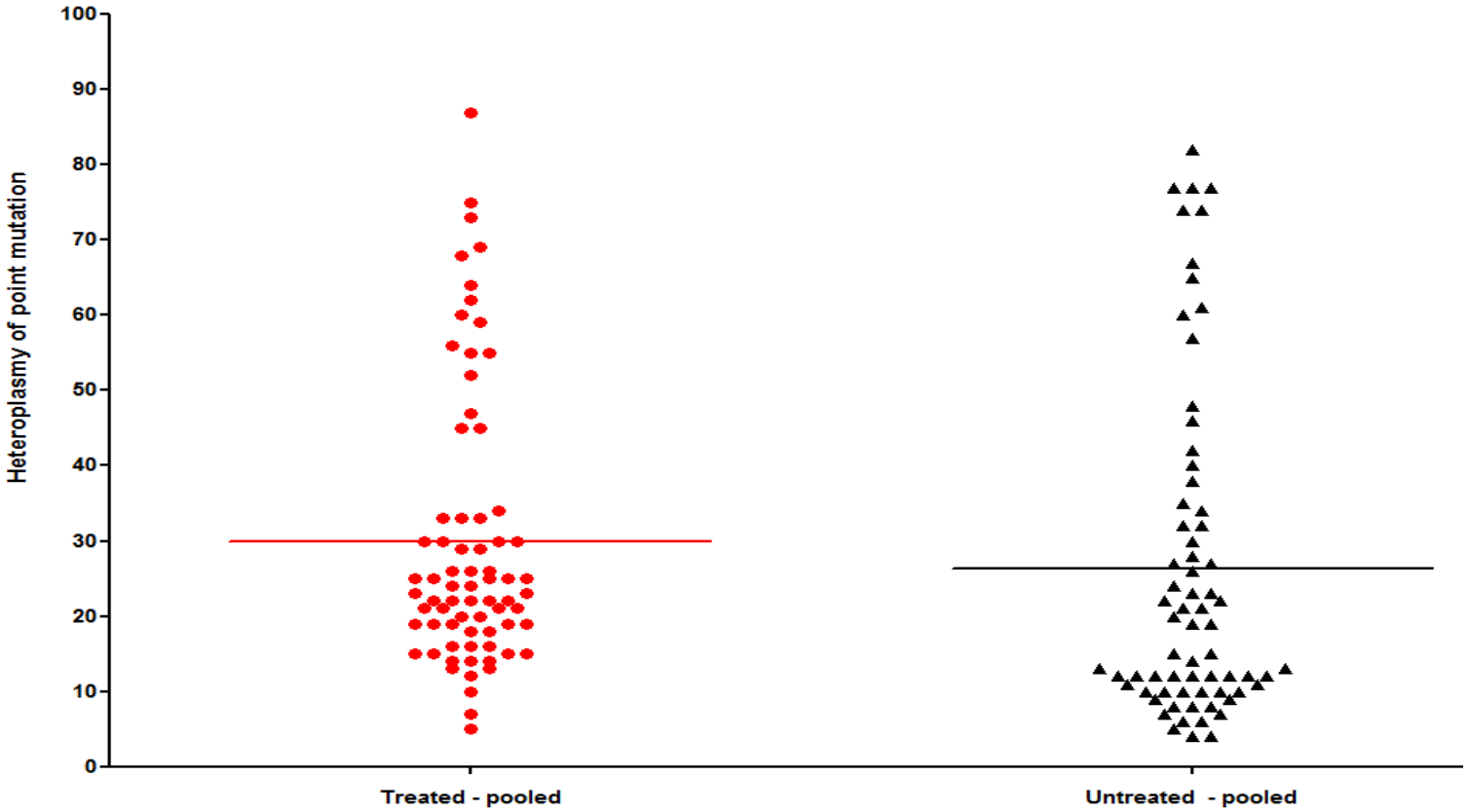

Figure 6.9 The heteroplasmy level distribution of mutation m.414T>G in single cells extracted from five replicates at experimental day 24, in the assessment of mitochondrial genetic bottleneck mechanism through ddI exposure study (results Section 6.4.3). Red indicate cell that have been exposed to ddI and black indicates untreated cells. Mean values are indicated for each replicate.

6.4.5 *De novo* mutagenesis formation

No *de novo* deletions were found in any of the samples, in any of the NRTI treatments before or after exposure, see Figure 6.10, in the same vain as described in Chapter 5.

An assessment for *de novo* mutagenesis of point mutations was performed using ultra-deep next generation sequencing technology and is discussed in Chapter 7.

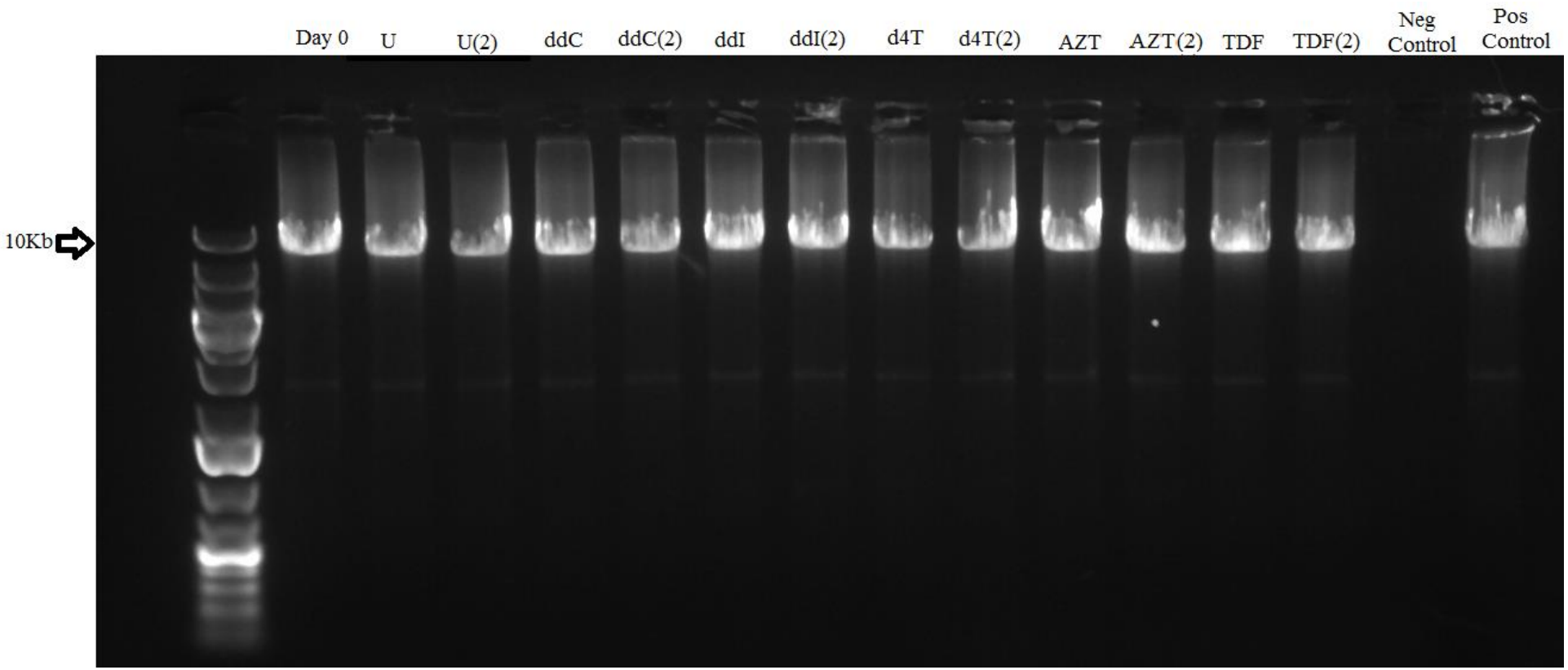

Figure 6.10 The agarose gel electrophoresis images of m.414T>G fibroblasts exposed to NRTIs at experimental day 42, end of exposure for the nr100 cell line, the same image was produced for the nr68 cell line. Despite the significant depletion in the ddC, a large 10Kb band was still produced, highlighting the presence of mtDNA content through repopulation after exposure.

6.5 Discussion

The aim of this experiment was to model the behaviour of a mitochondrial DNA point mutation that has been associated with ageing, through NRTI exposure using *in vitro* approaches. There were two main questions to be answered in this study:

- Is there any shift of point mutation level through NRTI exposure?
- If so, does this shift appear to be a replicative advantage mechanism or through a molecular bottleneck?

In addition to these questions, to further upon the assessment of *de* novo mutagenesis discussed in Chapter 6, as screening for *de novo* large scale deletions across multiple NRTIs was also performed, albeit not an aim of the study, the technique was used due to the cell line being of different nature (primary cell line, with a point mutation rather than deletion).

6.5.1 NRTI exposure-repopulation study data summary

The mitochondrial DNA copy number was found to decrease as expected in conditions ddC and ddI (Kornblum *et al.*, 2013); however, at a faster rate than expected, compared to the cybrid data from Chapter 5. This extreme depletion by day 6 of the experiment highlights the clearly differences and cell line specific sensitivity when compared to the depletion data the in trans-mitochondrial cybrid cell lines (see Chapter 5), which is more comparable to the 10x physiological dose, rather than the physiological dose, as used here. The cells may also have been affected more severely due to the age of the cells (derived from aged individuals). This gives insights into the effects of NRTIs on older cells too, and suggests an aged individual (>80 years old) receiving NRTI treatment, may be subject to faster depletion and NRT-associated effects.

Given the sensitivity of the fibroblasts to ddC and ddI, it is not surprising that there was depletion present in the mild polymerase γ inhibitor, d4T with 44% depletion compared to the untreated in nr100 after 6 days exposure, which is line with previously published data (Stankov *et al.*, 2010). A more mild 23% depletion was seen in cell line nr68 after 6 days exposure to d4T. In both cases, the copy number recovers to similar level as the untreated and once in the repopulation phase, increasing to greater than the untreated with 3.2-fold and 1.7-fold increases in cell lines nr100 and 68, respectively. This is likely through a over-compensatory mechanism driven by the mild depletion present and following NRTI removal.

The copy number in AZT and TDF conditions for both cell lines were found to suffer minimal depletion during exposure. However, the repopulation phase was the most interesting with 3-fold increasing in cell line nr100 for both conditions, and almost 2-fold increases in nr68, when compared to the untreated. In the case of AZT, this could be explained by a compensatory mechanism of the cell to produce more mitochondrial DNA following mild inhibition of replication; AZT caused ~23% depletion in both cell lines. The lack of any real depletion in the TDF conditions is in keeping with previous data.

The severe depletion in the ddC and ddI conditions during the exposure phase appears to create such a large reduction in copy number, to <1% of untreated, that during the recovery phase, there was no evident repopulation. This is likely due to the copy number being so low, that there were insufficient mtDNA copies remaining from which to repopulate. That is, the cells were approaching Rho-0 status. However, a larger recovery period would have been required to control for an almost complete loss of mtDNA copies.

The behaviour of the point mutation heteroplasmy during the NRTI exposure phase was similar across all conditions in both cell lines. Although there was no time point taken at experimental day 32 (end of exposure), due to the slow growing nature of the cells *in vitro*, at day 14 where severe depletion was present in the ddC and ddI conditions, there was no significant change to point mutation heteroplasmy. This indicates that there is no overall increase in the mutation and given that depletion is present, there is no replicative advantage mechanism by depleting copy number, in contrast to the large-scale deletion model in chapter 5.

6.5.2 Assessment for a molecular bottleneck mechanism: data summary

A molecular bottleneck effect can be created by a period of depletion followed by repopulation of mtDNA content. During recovery, mutations may be subject to accelerated drift, which will be predicted to manifest as increased spread in mutation heteroplasmy levels in this experimental design.

The ddI exposure, created a bottleneck, as seen by the depletion of total copy number at experimental day 6 (Figure 6.7); however, the heteroplasmy level shows no change across any of the replicates until the ddI had been removed and repopulation was allowed to occur. The tight spread of point mutation heteroplasmy levels present at

experimental day 6 rapidly increase to a statistically significant spread compared to the untreated (p-value<0.001) exactly as would be predicted with a molecular bottleneck (Khrapko, 2008).

The single cell data indicates that there is already a wide spread of m.414T>G heteroplasmy levels present between individual cells within the population. There are two key points from these analyses:

- The very large spread at baseline precludes showing an increase in spread after the bottleneck (it would be an increased, not decreased, spread expected).
- The single cell measurements may have been taken too late, as the replicates show the large spread at the intermediate time point during replication, at this point (end of recovery phase, experimental day 24), everything appears to have comes back together at the end of the repopulation phase.

6.5.3 Study limitations

The greatest limitation of this study is the use of fibroblast cell lines from aged individuals with a very slow replication time *in vitro*. This reduced the number of time points that were taken throughout experiments and may have meant that transient phenomena during repopulation/recovery were missed.

Another consideration is that the point mutation used in this experiment was a neutral 'non-disease' causing mutation. This mutation was chosen due to its presence in fibroblasts and the known associations with ageing. However, there is a possibility that the kinetics of a neutral mutation may be different to a disease causing point mutation. Nevertheless, a bottleneck mechanism would be expected to apply, regardless of mutation type *in vitro*.

6.6 Chapter conclusion and further comments

The data presented in this chapter strongly support the following conclusions:

- No systematic increase in m.414 point mutation heteroplasmy, due to NRTI exposure, despite it being deemed a likely hot spot for mutagenesis.
- There is evidence for a bottleneck mechanism during ddI exposure; this would potentially explain the *in vivo* data showing clonal expansion of point mutations.

It is finally worth noting the significant evidence to suggest that fibroblast cell lines are highly sensitivity to NRTI exposure with a clear hierarchy matching that of polymerase γ inhibition (Gardner *et al.*, 2013).

Chapter 7. The use of Next Generation Sequencing in the Detection of Very Low Level Mitochondrial DNA Mutations

Table of Contents

7.1 Background

The use of next generation sequencing (NGS) as a tool for analysing mtDNA has resulted in a breadth of different analyses performed from mitochondrial mutation load quantification to rare variant detection across multiple species (Y. He *et al.*, 2010; Ameur *et al.*, 2011; Y. Guo *et al.*, 2013; L. C. Greaves *et al.*, 2014). There have been a range of research efforts in developing methodologies on how to perform sequencing in the first instance and than techniques to analyse the data produced, as reviewed in Section 1.9.8.

The ability to successfully interpret NGS data is intrinsically linked to the quality of data itself, but also extrinsically in how the data is analysed. The main contributors to intrinsically linked factors affecting how data is interpreted are sequencing quality and quantity (sequencing depth). Although a large quantity of high quality sequencing data is often sought and regarded as being the matter of an optimal sequencing run, the ability to make interpretations of the data can be defined as behaving in a sigmoidal fashion, due to the summation of a variety of error rates (see Figure 1.6).

The extrinsic factors affecting the resolution of very low level variants is generally comprised of defining the minimum thresholds to call a variant. Other factors include assessing the level of sequencing strand bias, reference mapping quality and base call accuracy, to ensure variants lie within the binomial distribution.

There have been a number of recent studies that have concluded that the lowest level which mtDNA variants can be successfully resolved is in the region of ≥0.2% variant frequency. Such methodologies implemented the use of sequencing of cloned mtDNA in order to define the lowest level at which variants should be called in biological samples (Payne *et al.*, 2013).

There have been a number of studies suggesting that NRTI exposure can induce mtDNA mutagenesis through a number of different mechanisms; however, there were methodological limitations associated with the studies (Susan-Resigna *et al.*, 2007). The data presented in Chapter 6, suggests that this may not be the case and an increase in point mutation is mediated through a molecular bottleneck effect. The aforementioned studies also fail to illustrate how an induced point mutation would reach a high heteroplasmy level, and previous modelling data suggests it would require a substantial mutation rate to lead to a biochemical defect (Elson *et al.*, 2001).

7.2 Experimental aims

The experimental aims of this study were to use NGS to:

- Design and optimise a pipeline to define the detection limit of very low level mitochondrial DNA variants.
- Assess fibroblast and cybrid cells exposed *in vitro* to NRTIs for *de novo* mtDNA point mutagenesis.
- Develop a model for analysing low level mtDNA point mutations through a putative molecular bottleneck.
- Assess mutation burden in aged mouse samples exposed to NRTIs compared to aged control samples.

7.3 Experimental design and methods

7.3.1 Sample sets and DNA

Three sets of samples were used in this study: human samples, *in vitro* cells and aged mouse tissue samples.

DNA from the blood and skeletal muscle of a patient with a confirmed MNGIE disorder (mitochondrial neurogastrointestinal encephalomyopathy; *TYMP* mutation 22q12.32-qter), along with age matched control skeletal muscle, were studied. MNGIE DNA was provided by Professor Rita Horvath (Institute of Genetic Medicine, Newcastle University, UK). These samples were used to explore the lower limits of resolution of an NGS assay for mtDNA mutations, owing to the previously described presence of specific mutational motifs within MNGIE mtDNA.

DNA was extracted as previously described for the *in vitro* samples (fibroblasts and trans-mitochondrial cybrids; see Section 3.3). The m.414 T>G fibroblasts (cell line nr100), were exposed to a physiological dose of each NRTI (ddC, ddI, d4T, AZT and TDF) for 32 days and then placed into normal growth media for a further 10 days. Experimental day 0 (baseline) and experimental day 22 for all NRTI conditions was used in this study.

The trans-mitochondrial cybrid cell line was exposed to a physiological and 10x physiological dose of ddI, d4T, AZT and TDF for 32 days and then allowed to grow in normal media conditions (repopulation phase) for 10 days and 20 days, respectively. The time point used in this NGS study consisted of: experimental day 0, experimental day 32 and the end of repopulation phase, for each of the NRTI exposure conditions, specifically amplifying the mutant and wild-type molecules (see Section 3.10.1).

DNA from three mouse tissues: heart, liver and gastrocnemius, was kindly provided by Professor Joanna Poulton (University of Oxford, UK) from two aged mice (2 years old), one exposed to zidovudine *in utero* (0.15mg/ml; GlaxoSmithKline, UK) and throughout life, and the other a normal age matched control. The mice were classified as first cousins, sharing a maternal lineage. Original protocol is previously described (Morten *et al.*, 2005).

7.3.2 Next generation sequencing and bioinformatics

The DNA from the MNGIE and control samples were prepared using three overlapping fragments and pooled in equimolar quantities by Dr Brendan Payne and kindly provided for this study. The samples were sequenced using the SOLiD™ platform (Applied Biosystems, Life Technologies, Paisley, UK) by Genomic Research, University of Liverpool, UK.

The DNA from mouse and cells grown *in vitro* was prepared for sequencing as previously described in Sections 3.10.1, 3.10.2 and 3.10.3. They were sequenced using the Illumina MiSeq™ (Illumina, Cambridge, UK) platform by Dr Jonathan Coxhead (IGM, Newcastle University, UK).

All bioinformatics was performed as previously described (see Section 3.11.1). Minimum coverage and supporting reads data was generated based on run performance the optimal coverage to encapsulate at least >95% of the mtDNA genome.

7.3.3 Pipeline optimisation: strand bias

In order to determine the optimum bioinformatics parameters for analysing mtDNA low level variants, the MNGIE samples were used as positive controls for a stereotypic mtDNA mutation motif (nAT>C) (Nishigaki *et al*., 2004) and the variant calling parameters (as described in Table 3.9) were refined based on this pattern (as described in Section 7.4.1). The two-fold strand bias filter was implemented after optimising by comparing the number of 'lost' or excluded variants dropped reduced through a series of thresholds (ranging from five-fold to two-fold). The remaining variants were deemed likely to be signal rather than sequencing noise and included in the analysis. Filtering for strand bias is generally considered a stringent quality control step. This is especially important for sequencing platforms using short reads and strand bias is more prevalent in such systems, compared to long read sequencing systems (Yan Guo *et al.*, 2012). The exact cut off threshold for very low level variants is further discussed within this Chapter, under Section 7.4.1.

7.3.4 Pipeline optimisation: NGS amplicon generation

To further assess the optimisation of the polymerase used in the amplification of mtDNA products, an mtDNA sequencing clone from control *Mus musculus* was generated as previously described (see Section 3.8) using a 1188bp amplicon to generate the clone insert (see Table 7.1).

The clone was amplified using PrimeSTAR GXL (Takara, Clontech) as previously described (see Section 3.10.1) as well as Takara LA (Takara, Clontech), Phusion High fidelity polymerase (Thermo-Scientific, Hampshire, UK), Q5 (NEB, Herts, UK) using the same master mix as previously described for the PrimeSTAR GXL and amplified as previously described (see Section 3.10.1). Sequencing was performed on the MiSeq™ (Illumina, Cambridge, UK).

It was found that the number of variants predicted in the clone across the variety of polymerases (see Table 7.2) was lowest in the Takara PrimeSTAR GXL preparation and therefore, this was the chosen polymerase for all subsequent NGS experiments on the MiSeq™ platform. The raw data indicated a mean sequencing depth for each sample >30,000 fold, with the exception of the Phusion HF, which was found to be 17,316 fold.

Primer	Sequence 5' – 3'	mtDNA Position	T_m (^oC)	Size (bp)
Clone	F - TCCTACTGGTCCGATTCCAC	5' m.12533	63	1188
	R - TGATGGTTTGGGAGATTGGT	3' m.13721		

Table 7.1 The primer properties used to generate amplicon of mtDNA sequencing clone.

Polymerase	Sequencing Depth	Variants called (Post-filtering)
Takara LA	38,554	15
Takara PrimeSTAR GXL	30,051	2
Phusion HF	17,316	4
Q5	30,319	5

Table 7.2 The number of variants predicted post filtering, in the mouse (*Mus musculus*) clone using the diffferent primary LR-PCR enrichment polymerases.

7.4 Results

7.4.1 Defining the detection limit of very low level mitochondrial DNA variants

Human mtDNA amplicons (MNGIE and control) were sequenced on the AB SOLiD platform. The mean sequencing depth across the mtDNA genome was >16,000 for all three samples, with >95% of the whole mtDNA genome in all samples covered with a read depth of >5,000.

The expected stereotypic pattern (termed nAT>C, where n represent the length of the preceding poly 'A' tract) was seen extensively within the MNGIE skeletal muscle and blood samples, throughout the whole genome at all heteroplasmy variant threshold levels (initially all putative variants at ≥0.1% were considered). There was however no apparent increase in total mtDNA point mutation load in either MNGIE tissue compared with the control. Variant data can be found in Appendix B.

Variant threshold cumulative frequency 'bins' were then set at 0.1% intervals, from a variant heteroplasmy frequency of 0.1% to 1% in the MNGIE skeletal muscle and the control skeletal muscle datasets. The number of variants detected at these thresholds is plotted in Figure 7.1, stratified by matches to the nAT>C mutation motif and non-matches (all other variants), with the nAT>C mutational burden represented as a percentage of the total mutational burden at each threshold.

The most striking feature is the black bars representing stereotypical mutations throughout the MNGIE sample at all heteroplasmy thresholds. At ≥0.1% the stereotypical mutation burden was at the lowest of all the thresholds studied, when expressed as a proportion of the total mutational burden (27.8%). The proportion of stereotypical mutations was found to consistently increase in a progressive manner with an increased detection threshold, with 72.7% of total variants at ≥1% matching the nAT>C mutation pattern.

The control sample has a consistently low level of stereotypical mutations of approximately 10% of total mutations from ≥0.1% and ≥0.7%, with no stereotypical mutations found above the ≥0.8% heteroplasmy detection threshold. The highest overall stereotyped mutation burden was found at ≥0.6% threshold (at 14.3%), which was almost one fifth of that found at the same threshold level in the MNGIE sample (68% stereotyped mutation burden).

The variant detection threshold bin intervals were then narrowed to 0.02% between the 0.2% and 0.3% detection level (as a putative range for eventual threshold cut-off). This increased the resolution at the very low level and revealed the consistent trend of an increased proportion of stereotypical nAT>C mutations with increasing detection threshold in the MNGIE sample, and a consistent decreasing trend in the control.

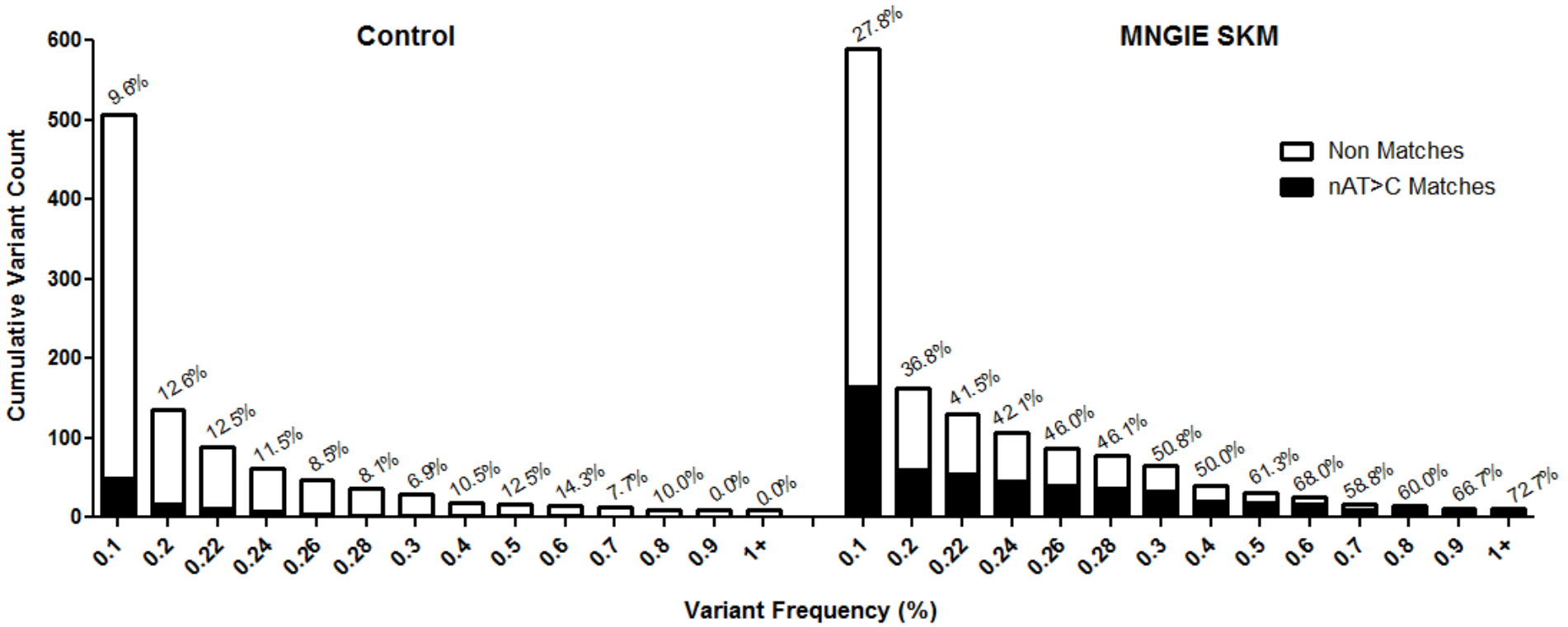

Figure 7.1 The cumulative variant count of the nAT>C matches and non-matches (all other variants) found in the MNGIE skeletal muscle (SKM) and control, categorised into variant frequency thresholds of 0.1%, with an increased resolution at 0.2-0.3% with increased 'bins' of 0.02%.

An ANOVA was performed to compare the distributions of nAT>C mutations in both samples at very low level variant detection thresholds, the *p*-values are tabulated below in Table 7.3.

As highlighted in Table 7.3, it was found that 0.3% was the statistically significant level at which the stereotypic mutations were different in the MNGIE compared to the control. To further investigate a more precise cut off, the threshold bins were decreased to 0.02%, between ≥0.2 - 0.3%. The *p*-values are tabulated below in Table 7.4. The ANOVA revealed that ≥0.22% was the lowest statistically significant detection level (Figure 7.2).

The number of variants matching the nAT>C pattern at was broken down into their constituent poly A tract lengths, where n= 1 - 4, at the candidate detection threshold of ≥0.22% (see Figure 7.3). There was only one stereotypical mutation detected in the control sample where cumulative tract lengths were n = 2, 3 or 4. This compared with 18, eight and seven respectively, in the MNGIE sample. If we consider a mutational motif of n = 4, there are 22 sites in the mtDNA genome at which a mutation could occur (i.e. AAAAT). There were seven mutations found in the MNGIE sample that matched this mutation motif, representing 31% of the total number of sites in the mtDNA genome where this mutation could theoretically occur.

Detection Threshold (%)	ANOVA *p*-value
≥0.1	0.255
≥0.2	0.087
≥0.3	<0.001
≥0.4	0.001
≥0.5	0.001

Table 7.3 The *p*-values calculated from an ANOVA of the nAT>C mutation distribution in the control and MNGIE skeletal muscle from very low level variant detection thresholds of ≥0.1% - 0.5%, in 0.1% grouping bins with statistical significant values highlighted in red.

Detection Threshold (%)	ANOVA *p*-value
≥0.2	0.087
≥0.22	0.034
≥0.24	<0.001
≥0.26	<0.001
≥0.28	<0.001
≥0.3	<0.001

Table 7.4 The ANOVA *p*-values of very low level nAT>C variants between ≥0.2-0.3% to further increase the resolution of the statistically significant threshold.

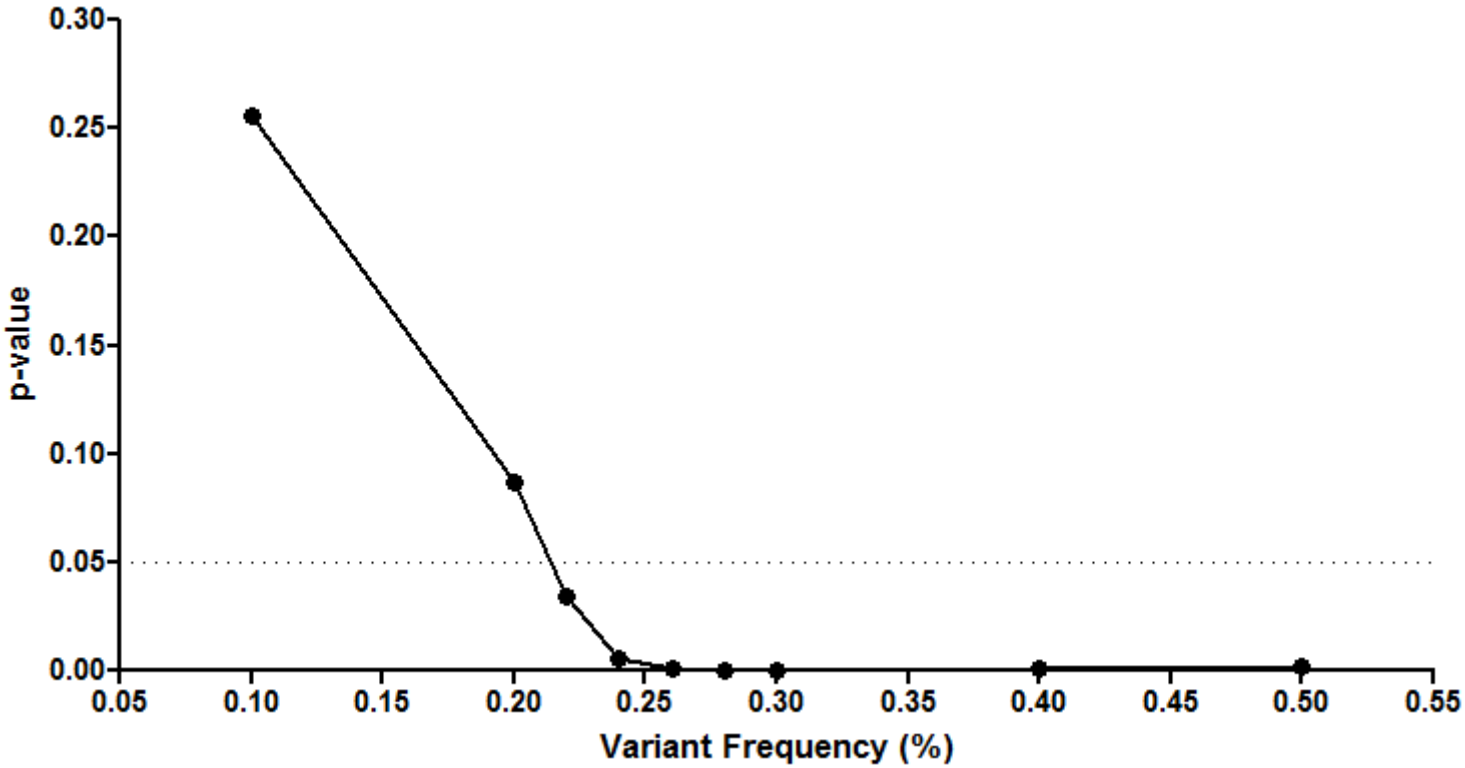

Figure 7.2 The p-value calculated from an ANOVA of nAT>C distribution between MNGIE skeletal muscle (see Figure 7.2) and control for different very low level cut-off frequency thresholds.

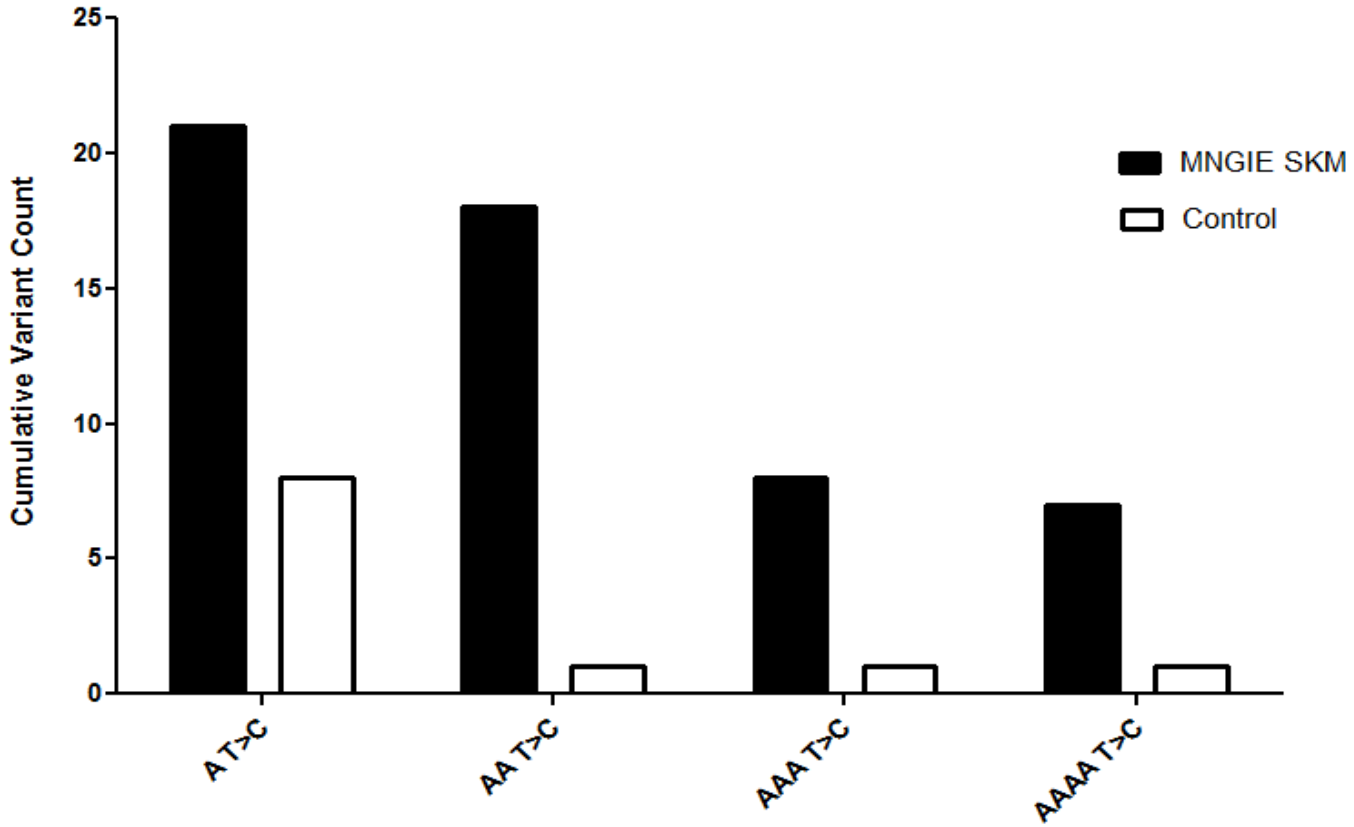

Figure 7.3 The distribution of nAT>C matches, where n=1-4, in the MNGIE skeletal muscle (black bar) and control (white bar). MNGIE SKM indicates MNGIE skeletal muscle.

7.4.2 Assessment of NRTI *de novo* mutagenesis

>98% of the mtDNA genome was covered at >3,000 fold sequencing depth in the fibroblasts cell line (nr100) in the following treatment groups: baseline (day 0), untreated, ddI, d4t, AZT and TDF (day 22). The ddC had slightly lower sequencing depth with >95% of the mtDNA genome covered at >3,000 fold.

The frequency of the m.414T>G mutation was sought and found to be in line with the results (illustrated in Figure 6.5, previously generated by pyrosequencing in Chapter 6). For this anlysis only, the bioinformatics variant calling criterion was reduced to compensate for poor sequencing depth between mtDNA regions m.262 – 514, to 1,000 fold coverage and 5 supporting variant reads (see Table 7.5). All conditions appear to be in line with the experimental day 0 and the untreated condition heteroplasmy. The ddC condition was very poorly covered and therefore was excluded from the analyses.

The count of heteroplasmic variants found in the coding region of the untreated and experimental day 0 samples were highly similar at the ≥0.3% detection threshold. The untreated condition was found to have 26 variants compared to 21 at experimental day 0. This was found to be a consistent level among the NRTI conditions. The count of very low level variants in the D-Loop (see Table 7.6) was also fairly consistent amongst all samples (complete list of D-loop variants available in Appendix C).

	Day 0	Untreated	ddI	d4T	AZT	TDF
414 Variant Frequency - NGS	55.0%	50.6%	46.6%	51.9%	49.3%	45.7%
414 NGS coverage	3121	4489	3753	3125	2567	2405

Table 7.5 The frequency distribution of m.414T>G in all samples, calculated using a reduced calling criterion of 1,000 minimum coverage and 5 supporting reads, along with the 414 site coverage. The variant frequency was compared to that found in Chapter 6, with no change visually apparent.

	Day 0	Untreated	ddC	ddI	d4T	AZT	TDF
HVS1 (m.16024-16365)	3	2	4	2	0	2	1
Central region (m.16366-56)	0	0	0	0	0	1	0
HVS2 (m.57-372)	0	2	0	2	0	0	0
m.373-575	0	2	0	3	0	4	0
Total	**3**	**6**	**4**	**7**	**0**	**7**	**1**

Table 7.6 The number of variants found in each condition at experimental day 0 and experimental day 22 from the fibroblast m.414T>G nr100 NRTI exposure-repopualtion experiment (see Section 6.4.2).

7.4.3 Assessment of a bottleneck mechanism through NRTI exposure

The cybrid samples were broken down into three categories in this analysis; point mutation behaviour on the wild-type molecule through exposure of NRTIs, at a physiological concentration; point mutation behaviour on the wild-type molecule through exposure of NRTIs, at a 10x physiological concentration; and the point mutation behaviour on the deleted mutant molecule through the exposure of ddI, at 10x physiological concentration. In addition to these analyses, a further assessment of *de novo* mutagenesis was also assessed in the samples exposed to NRTIs during the exposure phase.

All wild-type samples (amplicons derived from undeleted mtDNA, see Figure 3.9) were found to have >98% of the amplified mtDNA genome region, at a sequencing coverage of >5,000 depth, with an average read depth across all samples of 10,681. There was no evidence of increased mutagenesis in coding or D-Loop point mutation counts, with both regions consistently producing a similar number of point mutations at both physiological and 10x physiological doses.

The variants found at experimental day 0, in both groups (physiological concentrations and 10x physiological concentrations) were defined as the starting variant heteroplasmy level. These samples were compared to all other conditions in the group to identify shared variants present across all samples. The level of point mutation shift was calculated from day 0 to end of NRTI exposure and end of repopulation for each NRTI condition, at ≥0.3% variant frequency, in both biological replicates. The average level of shift from baseline, day 0 (between biological replicates) for the physiological concentration and 10 x physiological concentrations of NRTI exposure are shown in Figure 7.4 and Figure 7.5, respectively.

Figure 7.4A illustrates the level of heteroplasmy shift of the shared variants at the end of NRTI exposure at a physiological concentration. There is little shift present in any of the samples, with all conditions demonstrating small positive heteroplasmy shifts (increased level) from experiment day 0 of approximately 1.2%. Figure 7.4B illustrates that the level of heteroplasmy shift from experimental day 0 to the end of recovery shared small shifts with most NRTIs, which ranged from 2.24% (±S.D 1.05%) in the d4T condition, to 2.68% (±S.D 1.23%) in the untreated condition. However, the ddI exposure was found to have increased heteroplasmy shifts with an average across all the

six variants of 6.33% (±S.D 2.76%). A t-test (equal variance) found the difference in heteroplasmy shift in the ddI condition was statistically significant (p=0.008) compared with the untreated. All other conditions were statistically comparable to the untreated.

Figure 7.5A shows the level of heteroplasmy shift of seven point mutations across the NRTI 10x physiological concentration conditions. There is a small increased level in all point mutations in each condition with ddI found to have the smallest average increase of 2.1% in four variants (three omitted due to lack of sequencing depth) compared to an average of 3.92% increase in the untreated condition. Figure 7.5B shows the heteroplasmy shifts from day 0 to the end of the repopulation phase in the NRTI x10 physiological concentration conditions. All conditions show a heteroplasmy increase in each variant, with an average increased in the untreated of 5.87% (±S.D 1.76%). The d4T and AZT conditions show a similar average increase with 6.64% (±S.D 1.34%) and 5.36% (±S.D 1.77%), respectively. The TDF condition shows an average heteroplasmy increase of 7.82% (±S.D 2.37%) and the ddI condition found to display an average heteroplasmy shift of 12.26% (±S.D 0.72%). The increased heteroplasmy shifts in the ddI was only condition to be statistically significant when compared to the untreated (p<0.001). All shared variant data is expressed in Appendix D and Appendix E for the deleted amplicon.

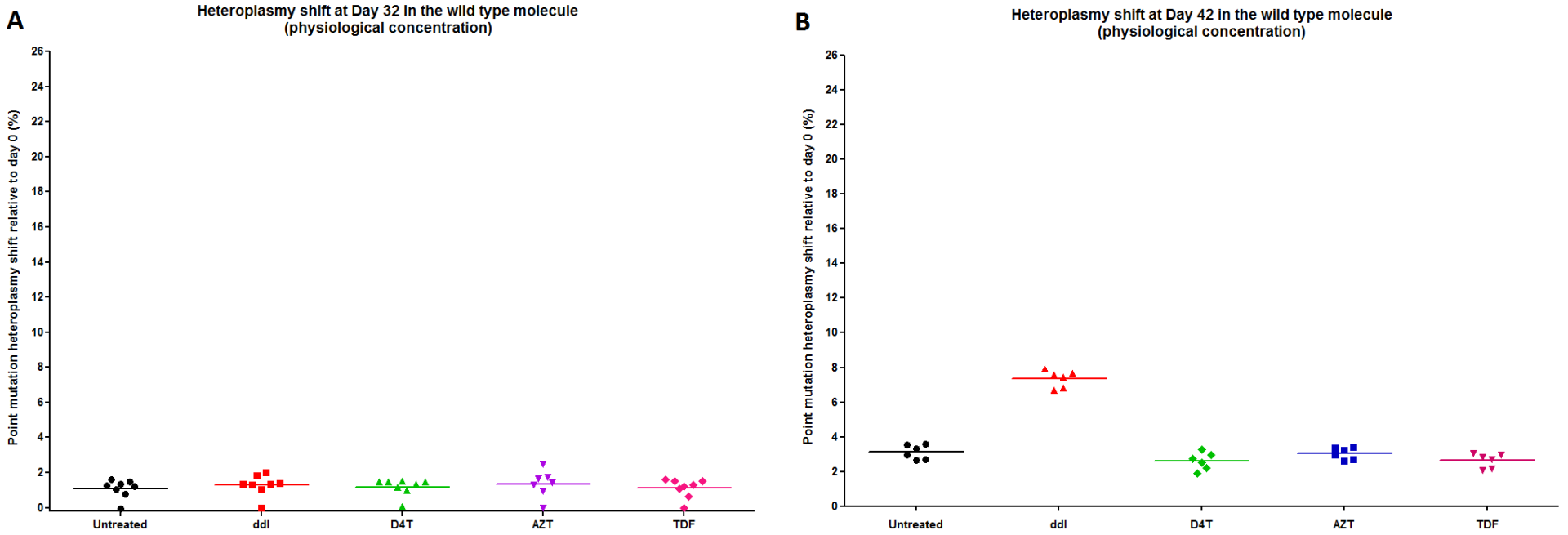

Figure 7.4 The heteroplasmy shifts of 6 shared variants across NRTI conditions at experimental day 0, 32 and 42 from the NRTI-exposure repopulation cybrid experiment. A) Heteroplasmy shifts present at the end of NRTI-exposure. B) Heteroplasmy shifts present in variants at end of repopulation.

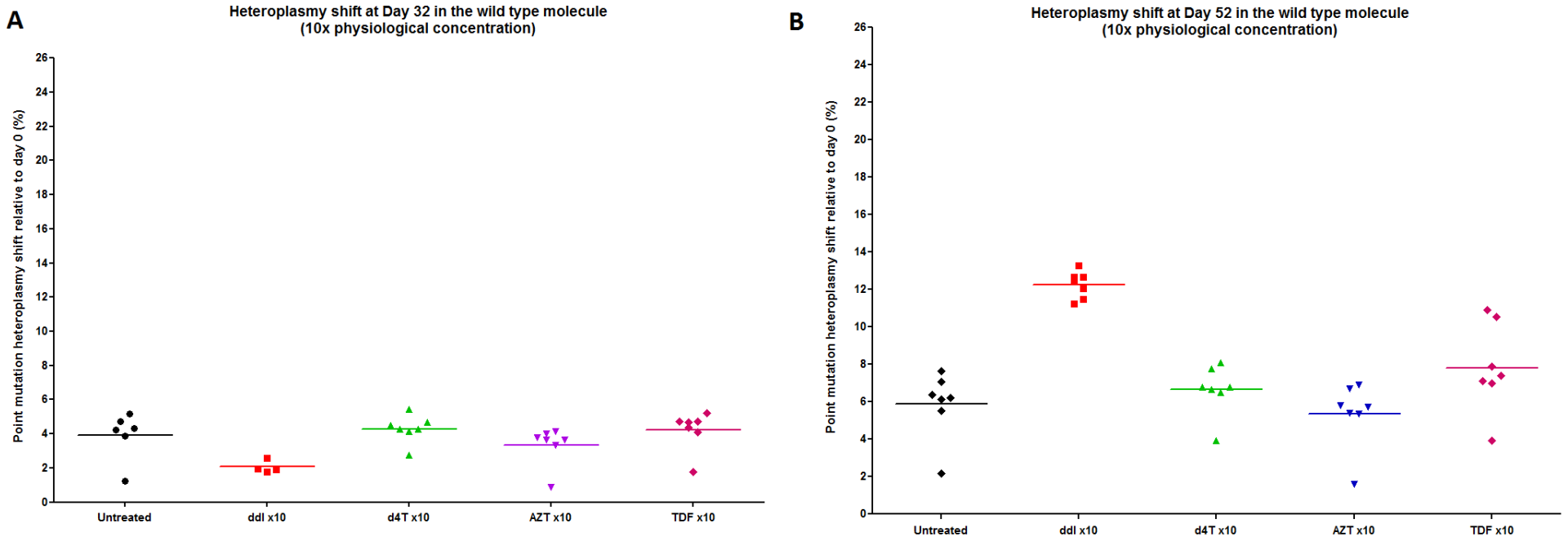

Figure 7.5 The heteroplasmy shifts of 7 shared variants across NRTI conditions at experimental day 0, 32 and 52 from the NRTI dose-response cybrid experiment. A) Heteroplasmy shifts present at the end of NRTI-exposure. Shifts in ddI of two variants were covered at <5,000 depth. B) Heteroplasmy shifts present in variants at end of repopulation.

All mutant only samples (derived from deleted mtDNA molecules) had a coverage of >98% of the amplified mtDNA genome region at a sequencing depth of >5,000. The average sequencing coverage across the samples was 16,066 fold.

The same analysis was performed as with the wild-type molecule only samples, with shared variants between experimental day 0, untreated and ddI x10 conditions (chosen due to a depletion of deleted mtDNA present at 10x ddI exposure compared with minimal depletion at physiological concentration, Chapter 5; Figure 5.21), at the end of NRTI exposure (day 32) and the end of the repopulation phase (day 52). The heteroplasmy shifts of 16 variants in both conditions at day 32 and day 42, is plotted in Figure 7.6A & B, respectively.

The most striking feature of both graphs is the large spread of heteroplasmy shifts in the ddI x10 condition at both time points. In comparison to the wild-type molecule only samples, where shifts were found to be increased, the mutant molecule only samples show both increases and decreases.

Figure 7.6A highlights an average shift in the ddI x10 condition as a (-)1.23% (±S.D 0.86%) decrease, compared to 2.62% (±S.D 0.78) increase in the average heteroplasmy in the untreated condition at the end of ddI x10 exposure phase. The ddI heteroplasmy shift of each variant was found to be significantly different to the untreated shift (p=0.0024).

Figure 7.6B shows the heteroplasmy shifts at the end of repopulation, which is almost the identical scenario to the end of exposure. The average shift in the ddIx10 condition was a (-)1.18% (±S.D 1.01%) decrease, compared to a 2.23% (±S.D 0.78%) increase in the untreated. The mean difference in shifts between the conditions was found to be statistically significant (p=0.01).

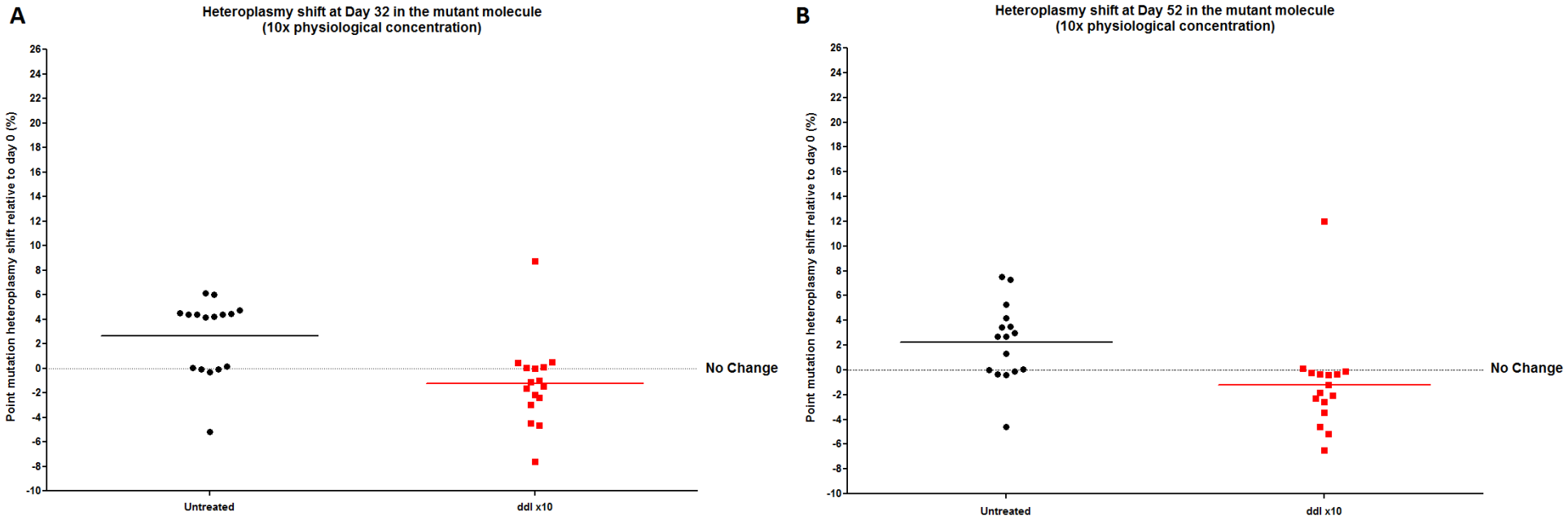

Figure 7.6 The heteroplasmy shift from day 0 of 16 shared variants in the mutant only amplicon for the control (black) and ddI x10 concentration (red). A) Point mutation heteroplasmy shift from experimental day 0 to end of NRTI-exposure phase. B) Point mutation heteroplasmy shifts from experimental day 0 to end of repopulation phase.

7.4.4 Intra-mitochondrial molecule comparison – mutant vs wild-type

As a unique opportunity to analyse the two molecules (deleted mutant and wild-type) coexisting in the same mtDNA sample, the shared variants within the mutant and wild-type molecules of the untreated samples were studied along with the heteroplasmy levels of those variants (see Table 7.7) and any differences between the molecules (intra-molecule shift) calculated (actual base position heteroplasmies, not shifts from base line). This was performed in both biological replicates for the experimental day 52 time point (10x physiological dose response experiment).

Across the shared variants, there is an interesting range of increases, decreases and positions that appear highly similar, in the mutant amplicon compared to the undeleted amplicon. The biggest shift in replicate 1 was found at mtDNA position 515 with an increase in the mutant of 16.66%. The largest decrease in heteroplasmy in replicate 1 was found to be at mtDNA position 16126 with a 5.71% change. The largest increase in replicate 2 was found at mtDNA position 2707, with a change of 15.60% and the largest decrease in heteroplasmy in the mutant replicate 2 molecules found at mtDNA position 16126, with a change of 7.28%.

Untreated Replicate 1			
MtDNA Position	**Mutant Molecule**	**Wild-type Molecule**	**Intra-molecule Shift**
150	10.18%	10.44%	-0.26%
185	9.96%	9.36%	0.60%
295	17.87%	7.95%	9.92%
515	17.05%	0.39%	16.66%
1811	12.18%	0.32%	11.86%
3010	15.21%	0.41%	14.80%
4188	6.65%	0.39%	6.26%
4216	5.95%	0.45%	5.50%
4640	4.18%	0.58%	3.60%
16069	5.53%	8.79%	-3.26%
16093	4.75%	9.77%	-5.02%
16126	4.65%	10.36%	-5.71%
16343	6.22%	8.57%	-2.35%
16519	8.24%	8.24%	0.00%

Untreated Replicate 2			
MtDNA Position	**Mutant Molecule**	**Wild-type Molecule**	**Intra-molecule Shift**
150	11.30%	12.32%	-1.02%
185	10.86%	10.93%	-0.07%
1811	13.30%	0.35%	12.95%
2707	15.95%	0.35%	15.60%
3010	15.28%	0.43%	14.85%
4188	5.67%	0.33%	5.34%
4216	5.12%	0.36%	4.76%
4640	4.46%	0.51%	3.95%
16069	5.88%	9.57%	-3.69%
16093	4.87%	11.23%	-6.36%
16126	4.73%	12.01%	-7.28%
16343	6.64%	10.45%	-3.81%
16519	8.42%	9.83%	-1.41%

Table 7.7 The heteroplasmic shared variants of the mutant and wild-type molecule in the untreated cybrid sample replicate 1 (left), and replicate 2 (right) at experimental day 52. A heteroplasmy shift was calculated between the mutant and wild-type molecules for each variant position and displayed in intra-molecule shift column for each replicate. Green indicates increase in mutant, yellow indicates no change and red indicates decrease in mutant. The intra-molecule shift indicates the amount of difference in heteroplasmy for a given position between the mutant and wild-molecule.

7.4.5 Assessment of mouse tissues after zidovudine exposure

The mean sequencing depth of the tissues from the control and AZT-exposed mouse samples is shown in Table 7.8. The mouse mtDNA clone was also included on this specific sequencing run and found to have a mean coverage depth of 56,666 fold.

Sample	Mean Sequencing Coverage/Depth
Control gastrocnemius	9,950
Control liver	8,970
Control heart	10,801
AZT gastrocnemius	13,945
AZT liver	16,477
AZT heart	13,746
GXL clone	56,666

Table 7.8 The mean sequencing coverage/depth of four tissues from a 2 year old control mouse and 2 year old AZT exposed mouse.

There were no variants found in the clone across the mtDNA genome, above a ≥0.3% variant threshold.

The number of variants across the coding region of the mtDNA indicated that there were no distinct differences between AZT-exposed and control in type of mutation (synonymous, non-synonymous, missense and mtRNA variants; see Table 7.9).

The mtDNA coding region for mitochondrial subunit COX I was found to be a common site for synonymous, non-synonymous and missense variants which was exclusive to the AZT tissues.

An assessment of the number of variants across the mitochondrial D-Loop in all samples indicated distinct differences between the AZT-exposed and the control samples (see Table 7.10; full variant data in Appendix F). The AZT gastrocnemius tissue was found to have the highest number of D-Loop low level variants with 13, compared to two variants in the control gastrocnemius tissue. AZT heart was found to have ten D-Loop mutations and six variants in liver. The respective tissues in the control sample revealed no mutations in liver and two variants in heart.

Samples	Synonymous	Non synonymous	Missense	RNA's
Control gastrocnemius	2	5	2	4
Control liver	3	3	0	2
Control heart	4	23	5	7
AZT gastrocnemius	**2**	**11**	**5**	**4**
AZT liver	**1**	**5**	**0**	**2**
AZT heart	**4**	**16**	**7**	**15**

Table 7.9 The mouse variants across coding regions of the AZT-exposed and control samples represented as their constituent variant type (non-synonymous, synonymous etc.). The bolded samples are the AZT exposed tissues and non-bold samples are from the control mouse.

A Fisher's exact test was performed comparing the D-Loop variant count in the same tissue between control and AZT-exposed mouse samples; with a final overall assessment of pooled, total D-Loop variants across all tissues in each mouse (see Table 7.10). The count of variants was found to be statistically significant for each tissue in the AZT-exposed mouse, with the overall burden in the AZT-exposed mouse found to be strongly statistically significant (*p*-value <0.0001) compared to the control.

Tissue	D-Loop Mutations	Fisher's exact test (*p*-value)
Control gastrocnemius	2	0.007
AZT gastrocnemius	**13**	
Control liver	0	0.031
AZT liver	**6**	
Control heart	2	0.038
AZT heart	**10**	
Total in control	4	<0.0001
Total in AZT	**29**	

Table 7.10 The number of mitochondrial D-Loop mutations found in the liver, heart and gastrocnemius tissues of AZT-exposed and control mice. A Fisher's exact test was performed comparing number of variants between tissues along with a pooled comparison of total variants in the AZT mouse compared to total in the control. Red indicates statistically significant difference.

The total number of D-Loop mutations pooled from the three tissues was categorised in to cumulative counts found in variant heteroplasmy level thresholds levels 0.1% bins, ranging from detection level of ≥0.3% to ≥1% in the AZT-exposed and control samples (see Figure 7.7). The number of mutations in the AZT-exposed mouse sample was found to be higher than the control at each threshold. The number of variants in the control sample was found to be zero by a detection threshold of 0.5%.

A three tissue shared variant comparison was performed in the D-loop of the AZT tissues. Only one variant was found to be shared in all three tissues, at similar levels in each tissue, consisting of: 0.48%, 0.44% and 0.62%, for gastrocnemius, liver and heart respectively at mitochondrial position m.16103. The remaining variants in the d-loop of the samples were unshared; AZT gastrocnemius, 12; AZT liver, 5; AZT heart, 9.

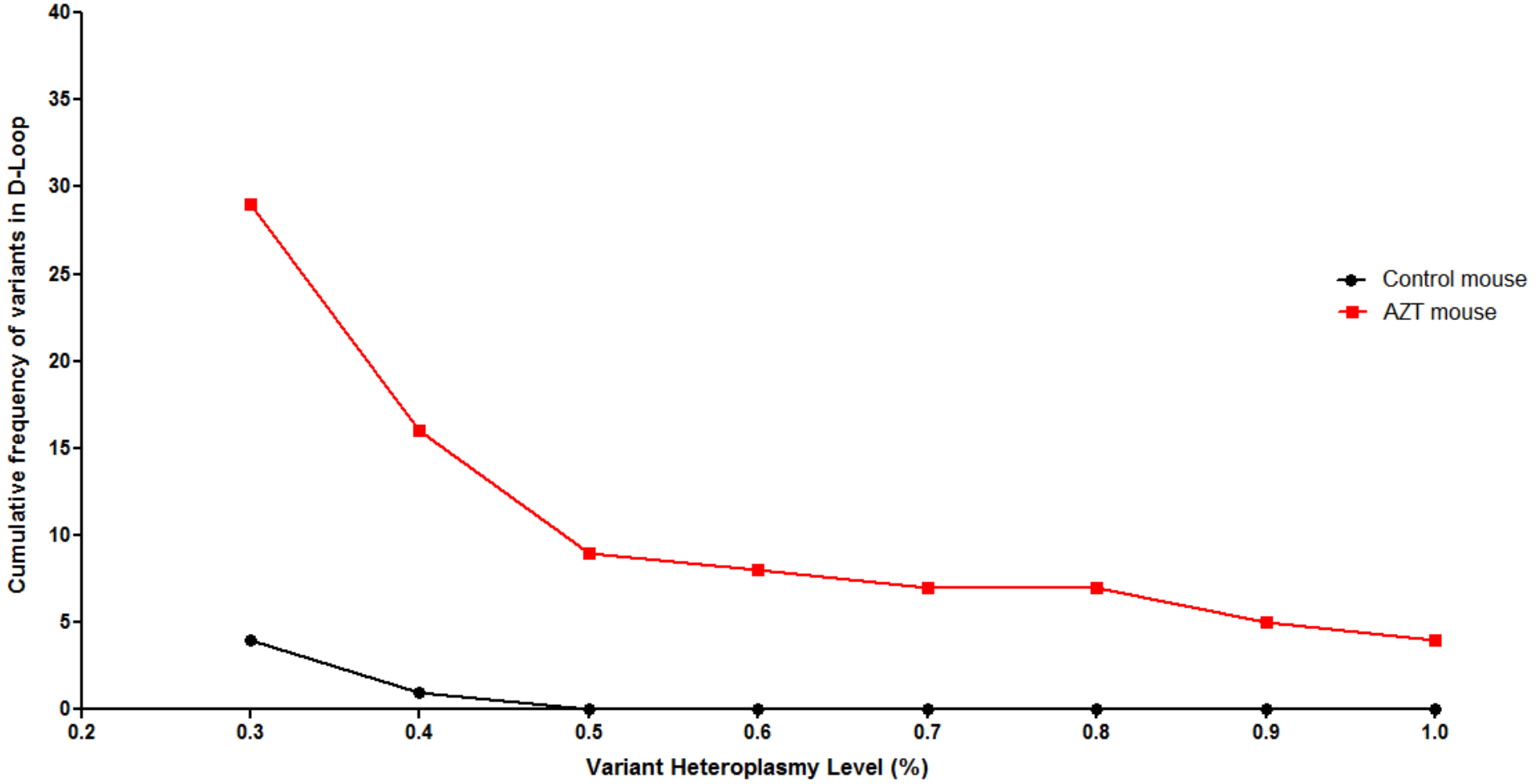

Figure 7.7 The cumulative frequency distribution graph of pooled D-Loop variants across the four tissues representative of the variant load present in the control aged mouse (black) and AZT-treated aged mouse (red).

7.5 Discussion

7.5.1 Defining lower limits of resolution for very low level mtDNA variants using NGS: data summary

In the course of this experiment, I implemented a novel methodology to derive empirical lower limits of resolution for characterising mtDNA low level variants. Ultimately, this revealed strikingly similar results to those thresholds previously claimed (He *et al.*, 2010; (Payne *et al.*, 2013), with an optimised bioinformatics approach suggested for mtDNA variant detection.

This study was performed by taking advantage of the metabolic disorder, MNGIE, which is known to induce a characteristic mtDNA mutation motif (nAT>C; Nishigaki *et al*., 2004) and comparing the motif frequency with a healthy control. It is inevitable with deep re-sequencing experiments on NGS that at very low variant heteroplasmy levels there will be a mixture of signal and noise detected. In reality, the noise will not stop abruptly above a certain threshold, but rather will taper off. Thus the higher the heteroplasmy detection threshold employed the greater the confidence will be that any detected variant is biological. However for the analysis of mtDNA somatic mutations, it is precisely the lowest level variants that are likely to be of the greatest interest, and therefore it is desirable to have an analytical method which allows the lowest possible heteroplasmy threshold to be employed. Furthermore, bioinformatics tools, such as stringent strand bias filtering will remove large amounts of noise, but little signal, so that lower heteroplasmy thresholds can be interrogated. Previous methods for defining lower limits of resolution have either used kit phage DNA (which may not capture mtDNA-specific sequencing problems), or specific mtDNA clones (which cannot cover the whole mtDNA genome and therefore may not be stringent enough). Given that nAT>C variants are very rare in the control sample (whether due to biological variants, or noise), we can be very confident that the vast majority of nAT>C variants detected in the MNGIE samples are therefore real biological signal. It was therefore empirically derived that the statistically lowest level where stereotypical mutational signal in the MNGIE sample outweighed the control sample was at ≥0.22% variant frequency/heteroplasmy level (see Figure 7.2). A more conservative approach was also taken by assessing the level at which the proportion of the nAT>C motifs was >50% of total mutational burden in the MNGIE sample. Therefore at a very low level of ≥0.3%

heteroplasmy we can be confident that the vast majority of variants reported using our pipeline will be true biological mutations.

The methodology also implemented a stringent quality control feature of using a maximum 2-fold bi-directional strand bias filtering process. This technique is especially important for short read length deep sequencing, as with the SOLiD™ and Illumina platforms. Most variant calling tools implement a level of strand filtering, however, I argue that the level of filtering built into the software was designed for whole genome sequencing studies, with innately lower read depth, and not designed for the large amount of sequencing depth and increased level of noise found in with deep sequencing of smaller molecules, such as the mtDNA genome. A stringent strand bias filter accounts for under-represented variants (low supporting reads within the high sequencing depth), especially when analysing data at very low detection levels (<1%), which improves the overall reliability of variants being called lying within a binomial distribution.

An often overlooked step of NGS preparation is the polymerase for initial amplicon enrichment and amplification before sequencing. Here I present a novel small study of polymerase that is marketed by the manufacturers as optimised for NGS purposes. It is evident that false calls and an increased error rate is especially prevalent when using the *Takara* LA *taq*. The optimal performing polymerase for lowest error rates was found to be the *Takara* GXL PrimeSTAR. Using this polymerase, variants were negligible, indicating a 'clean' pipeline.

This study carries the possible limitation that the pipeline was optimised on data from one platform (SOLiD) and it was assumed to translate across other platforms (with potentially different base calling error rates). It was however deemed acceptable in this case, as previous studies have performed quad-platform comparison and found little to no difference in calling of low level variants (Archer *et al.*, 2012).

7.5.2 Assessment of NRTI *de novo* mutagenesis data summary

It has been previously suggested that the mtDNA point mutation load increase is driven through a mutagenic effect of the NRTIs. A documented study suggests this increased point mutation load occurs during the exposure to AZT in skeletal muscle of HIV-infected patients presenting with myopathy (Casademont *et al.*, 1996). This effect is potentially comparable to inherited defects of *POLG* that result in an accumulation of

mtDNA mutations, as previously shown (Payne *et al.*, 2013). *Casademont et al.*, suggest the underlying mechanism is due to a lack of exonuclease function of pol γ, caused through inhibition by NRTIs. Alternatively, inhibition of mitochondrial thymidine kinase (TK2) may result in a nucleotide pool imbalance and induce mutations (analogous to that seen in MNGIE, presented in Section 7.4.1). This would be expected to present in NGS data as either a significantly increased number of total base positions across the genome that are mutated, or a large increase in very low level point mutations.

The NGS of samples exposed to a range of NRTIs *in vitro* in two cell lines (fibroblasts and trans-mitochondrial cybrids) revealed that there was no mutation burden increase across the coding region or in the D-Loop, where an increased level of specific point mutations presenting as clonal expansions within cells has been described in NRTI-exposed individuals (Martin *et al.*, 2003; Payne *et al.*, 2013).

7.5.3 Assessment of a molecular bottleneck through NRTI exposure data summary

The data presented in Chapter 5 & 6, indicated that there are the conditions for a possible molecular bottleneck effect, especially in opposed to strong inhibitors of pol γ, such as ddI. *In vivo*, this would drive an increase of point mutation heteroplasmy through NRTI exposure, due to accelerated neutral drift. Therefore, further confirmation of this hypothesis was sought through NGS deep sequencing of trans-mitochondrial cybrids that had been exposed to a range of NRTIs. It was found that the level of heteroplasmy shift during exposure of the NRTIs was minimal and entirely comparable to the drift seen in the untreated sample (see Figure 7.4A & Figure 7.5A). However, at the end of the repopulation phase (after NRTI removal), the heteroplasmy shifts were significant in the wild-type molecules (undeleted) of the ddI-exposed cells at both physiological and 10x physiological concentration. The expansions correlate with the increase of mtDNA copy number after depletion found in Chapter 5 (see Figure 5.10& Figure 5.21) during the NRTI-exposure repopulation assays. Further work could be implemented to advance the understanding of this by using a dose-range with ddI to display the different levels of shift from a greater/tighter bottleneck as depletion increases through from 1x physiological to 10x. This effect also appears to be subject to a dose-dependent response with the 10x physiological dose creating a greater mtDNA depletion and also a greater heteroplasmy shift (see Figure 7.5B), the process is described in Figure 7.8.

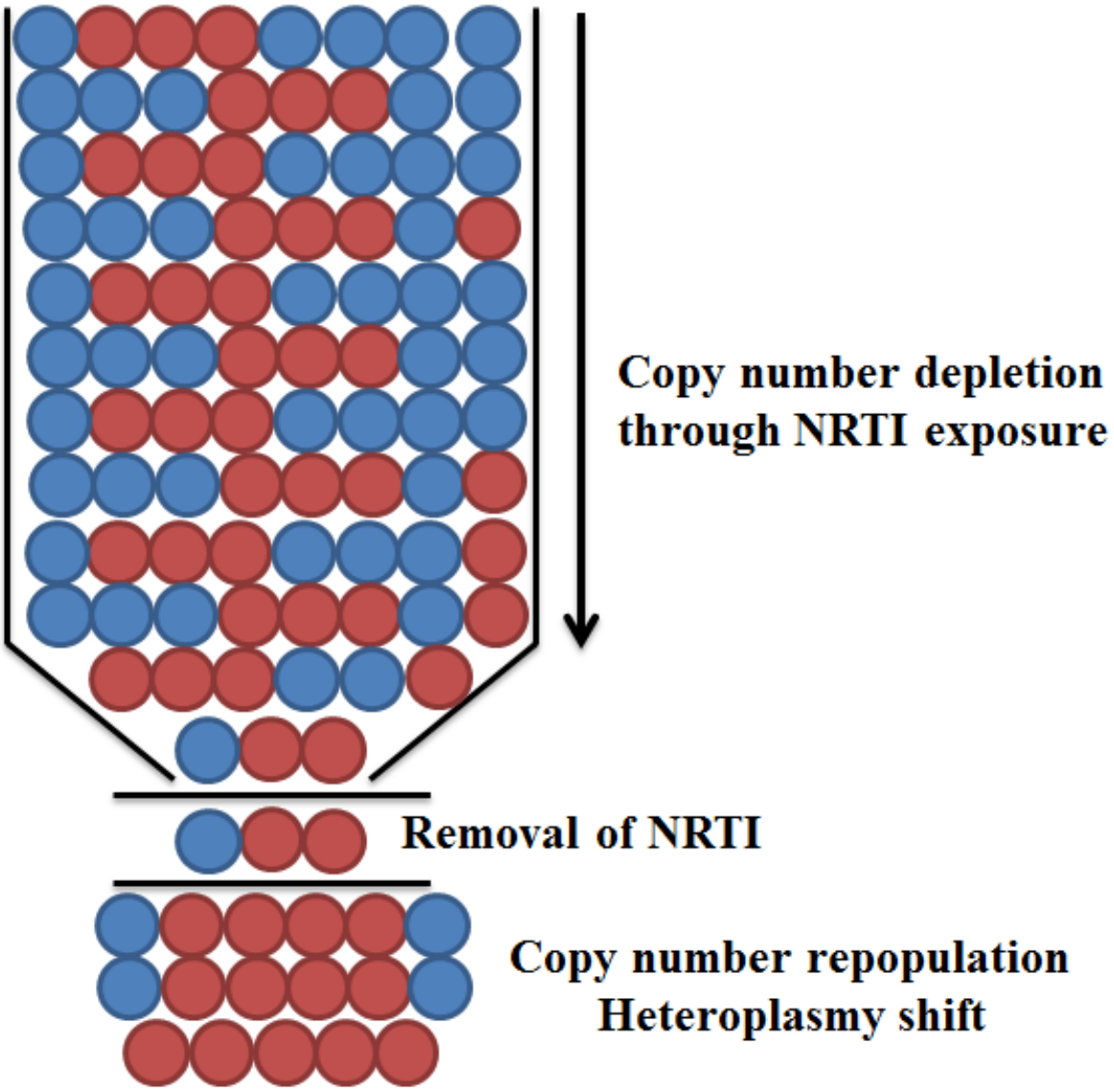

Figure 7.8 The representation of the process of mitochondrial copy number depletion through a bottleneck caused by NRTI exposure, resulting in shift of mutation level after removal of NRTI and repopulation of the mtDNA population/copy number.

The distribution of point mutations in mtDNA comparison between the mutant (deleted) and wild-type (undeleted) molecules highlighted an interesting feature of an apparent spread of heteroplasmy levels of shared variants in the mutant compared to the wild-type. Although this data is preliminary, I interpret and propose a phenomenon of point mutation clonal expansion on deleted mtDNA molecules within a heteroplasmic population. The mechanism underlying this is based upon the fact the deleted mtDNA is potentially already subject to a replicative size dependent mechanism for survival and increased turnover. Through this the molecule is subject to increased rounds of replication and therefore an associated point mutation could accumulate as described with point mutations such as the MELAS m.3243 (Picard *et al.,* 2014). This data is summarised in three main points:

- The baseline mtDNA point mutations appear to be different in some cases on the deleted and wild-type mtDNA molecules. In patients, these may have existed as distinct populations from birth and drifted through time, potentially due to the replicative difference between deleted and non-deleted molecules.
- In principle, these data suggests a 'piggy-back' mechanism in that each mtDNA point mutation behaves differently, according to the mtDNA species it is associated with.
- This could be further assessed through patient biopsies to look at deleted and undeleted mtDNA to perform the same analysis and ensure this phenomenon extends to *in vivo* observations.

7.5.4 Deep sequencing of multiple mice tissues exposed to AZT data summary

The data presented in Chapter 4 indicated that there was a shift in a point mutation associated with NRTI exposure in tissue from HIV-infected individuals. This finding supported previously described data in physiological samples from humans (Payne *et al.*, 2013). There have been studies assessing the mtDNA point mutation status through NRTI exposure in other mammals such as rodents, however, unfortunately they didn't implement the best methodology. Therefore, the opportunity was taken to implement the optimised NGS methodology in an attempt to elaborate upon this issue.

There is evidence to suggest that *in utero* exposure to NRTIs can result in mitochondrial dysfunction (Brogly *et al.,* 2007). Three tissues from mice exposed to NRTIs *in utero* and throughout life were sequenced using the NGS platform, MiSeq™ (Illumina,

Cambridge, UK). The results indicated that there was no increase in mutation load or heteroplasmy levels in the coding region; however, there was an increase in mutation load in the D-Loop. This is somewhat surprising given the previous findings in this chapter (see Section 7.5.3). However, it is interesting to note the increased mutation burden in the non-coding region, which was especially seen in the conserved region 2 of the mouse mtDNA. This may indicate a specific hotspot for point mutation increase during NRTI exposure. This region may be comparative to the D-Loop hotspot for accumulating point mutations during age, found in humans. The statistically significant increase across all three tissues suggests that this is not a chance finding. The data indicates that this is likely to be a mutagenic effect, as we see an increase in low level mutations at different positions across the different tissues, rather than increased heteroplasmy of common mutations.

The major drawback of this study is the lack of biological replicates used in the mouse study for further supporting (or refuting) the mutation load increase in the mouse mtDNA D-Loop with AZT exposure. An insightful addition to this study would have been an age range of mice representing different exposure times of AZT, which would have given a greater insight into any time dependency associated with the effect seen. However, the time point used in this study represents the extreme level of time exposure, especially within a mouse model.

7.6 Chapter conclusions

There are three major conclusions that can be drawn from the data presented in this chapter:

- The discovery and quantitation of very low level mtDNA variants can be successfully resolved at the range of ≥0.22-0.3%, by implementing the specific methods described.
- NRTIs do not appear to cause *de novo* point mutation mutagenesis in cells exposed *in vitro,* to a range of NRTIs and concentrations. However, tissue samples may be subject to an increased mutation load in specific mtDNA hotspots over a very prolonged timescale, *in vivo.*
- Low level mtDNA point mutations are subject to a molecular bottleneck effect when exposed to NRTIs, which are strong pol γ inhibitors (such as ddI). This will lead to enhanced clonal expansion *in vivo*, due to accelerated drift.

Chapter 8. Increasing Mitochondrial Copy Number through the up-regulation of Mitochondrial Biogenesis

Table of Contents

8.1 Background

The abundance of mitochondria within a cell is determined by mitochondrial biogenesis and division of the organelle through tightly regulated transcription factors and signalling pathways (Attardi and Schatz, 1988; Moyes and Hood, 2003). It therefore follows that the greater the numbers of mitochondria that are present within a cell, the greater the number of mtDNA copies there will be.

Peroxisome proliferator activated receptor γ co-activator-1α (PGC-1α) is a transcriptional co-factor that induces mitochondrial biogenesis (see Section 1.2.5) through the interaction with different transcription factors (Jornayvaz and Shulman, 2010). Activation of PGC-1α is known to be mediated by a number of co-factors. Of these co-factors, AMPK (AMP-activated protein kinase) is the major regulator that has been described to directly affect PGC-1α activity through a phosphorylation pathway (Canto and Auwerx, 2009).

SIRT1 (sirtuin 1) is another co-factor which up-regulates PGC-1α expression through de-acetylation of PGC-1α/ERR complex (oestrogen related receptor). The PGC-1α/ERR complex is implicated to have a major role in regulating biogenesis (Wu *et al.*, 1999). The interaction of PGC-1α with nuclear respiratory factors (NRF-1 and 2) and the PPAR's (Peroxisome proliferator activated receptors α, β and γ) result in an up-regulation of mitochondrial biogenesis through promoting mtDNA transcription and translation (Canto and Auwerx, 2009).

There are candidate compounds (drugs and nutritional supplements) said to have a beneficial impact upon mitochondrial biogenesis by promoting an up-regulation of the regulatory pathway of PGC-1α (for greater detail, see Section 1.2.5).

Those implicated to have such an involvement in mitochondrial biogenesis up-regulation, are pioglitazone, bezafibrate, AICAR (AICA ribonucleotide) and resveratrol. Briefly (as depicted in Figure 8.1), pioglitazone and bezafibrate are compounds used in the clinical setting to treat diabetes (pioglitazone) and manage hypercholesterolaemia (bezafibrate). Both pharmaceuticals cause an up-regulation of the PPAR co-factors, which in turn have a feedback mechanism that also up-regulates PGC-1α, and vice versa (L. Li *et al.*, 2011). AICAR is an analogue of AMP (adenosine monophosphate) and increases the up-regulation of AMPK. Resveratrol is a type of

phenol known as phytoalexin, which has been found to increase PGC-1α expression through SIRT1 (Davinelli *et al.*, 2013).

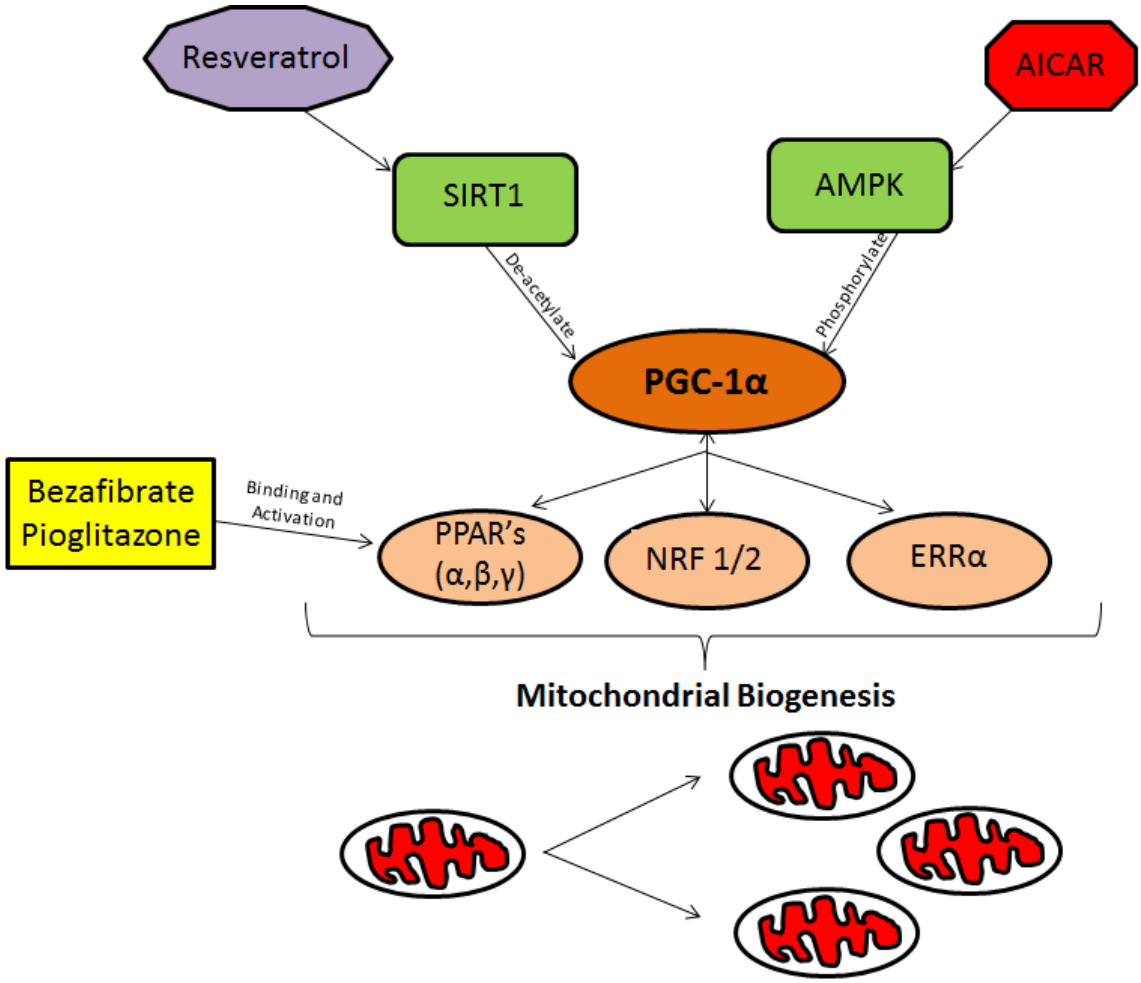

Figure 8.1 An overview of the induction of PGC-1α through bezafibrate, AICAR, resveratrol and pioglitazone supplementation.

The mitochondrial copy number depletion present during NRTI therapy is well established and has been discussed extensively throughout the previous chapters of this thesis. There have been a small number of studies assessing co-treatment with compounds that may reduce mitochondrial damage; namely, Acetyl-L-carnitine, co-enzyme Q, uridine and most recently, vitamin D (Hart *et al.*, 2004; U. A. Walker *et al.*, 2006; Cherry *et al.*, 2010; Campbell *et al.*, 2013; Sinha *et al.*, 2013).

8.2 Experimental aim

The experimental aims of this study are:

- To assess whether compounds associated with the up-regulation of mitochondrial biogenesis, can reduce the extent of mitochondrial copy number depletion during NRTI exposure;
- To assess the extent to which such compounds can prevent the effects of NRTI exposure on mtDNA deletion mutation heteroplasmy.

8.3 Experimental design and methods

8.3.1 Trans-mitochondrial cybrid cell line culture

Trans-mitochondrial cybrid cell lines were cultured and propagated as previously described (see Section 3.1.1 and 3.1.3) using T_{25} culturing flasks (Greiner Bio-one, Stonehouse, UK).

8.3.2 Exposure procedure

The concentration of compounds matched those that have been used in previously described *in vitro* studies, which were found to have a positive effect on PGC-1α expression (Miglio *et al.*, 2009; Davinelli *et al.*, 2013; Noe *et al.*, 2013). The co-supplementation set up is described in Table 8.1 with 14 days exposure procedure used for each of the conditions.

Condition	Concentration
Untreated	N/A
ddI	11.8µM
Bezafibrate	100µM
Bezafibrate + ddI	100µM + 11.8µM
AICAR	100µM
AICAR + ddI	100µM+ 11.8µM
Resveratrol	10µM
Resveratrol + ddI	10µM + 11.8µM
Pioglitazone	1µM
Pioglitazone + ddI	1µM+11.8µM

Table 8.1 The co-supplementation procedures of the supplements and the NRTI, ddI. All conditions were run in biological triplicate.

8.3.3 Mitochondrial deletion and copy number quantification

The deletion level and mitochondrial copy number was calculated using the multiplex qPCR assay as previously described (see Section 3.6.1) and using the previously described analysis methods for deletion level and copy number calculation (see Sections 3.6.2 and 3.6.3). The concentration of each DNA sample was standardised within the range of 5-10ng/µL concentration, with 5µL loaded into each 25µL reaction. Each sample was run in duplicate on the reaction plate and allowing for a 0.5Ct difference between sample replicates.

Mitochondrial copy number was expressed as a relative ratio to the untreated group for each time point.

8.3.4 DNA extraction

The residual cybrid cells that were not re-seeded in propagation were span at 1300rpm for 5 minutes, supernatant removed and the pellet frozen at -80^{o}C, until required. The DNA was then extracted using the DNeasy® Blood and Tissue extraction kit (Qiagen, Manchester, UK) as previously described (see Section 3.3).

8.4 Results

8.4.1 Effects of supplement conditions on mitochondrial DNA copy number

The relative total, mutant and wild type mitochondrial copy number (relative ratio to the untreated) of each co-supplementation group for the cybrid cells is displayed below (see Figure 8.2, Figure 8.3 and Figure 8.4). Each data point is expressed as the mean values derived from the biological and qPCR technical replicates with error bars plotted as standard deviations (S.D). All graphs are plotted on a logarithmic scale to improve resolution of relative to untreated copy number at the lower end. Bezafibrate exposure was found to be toxic to the cells.

The relative total copy number in the ddI only condition displays evident depletion throughout the 14 days of exposure. Relative copy number by experimental day 7 was 0.27 (±S.D 0.077), which was found to be the maximal level of depletion throughout the exposure. The depletion appears to improve throughout the experiment ending on experimental day 14 at relative copy number of 0.47 (±S.D 0.43). Depletion throughout the experiment was found to be significant (p=0.05) compared to untreated condition.

The relative total copy number in the AICAR condition (without ddI) was found to be raised by experimental day 3 with a 1.39-fold higher total copy number than the untreated (1.39 ±S.D 0.35). This represented the largest change throughout the experiment, with fluctuations between 0.94 at experimental day 7 (±S.D 0.28) to 0.72 (±S.D 0.48) at experimental day 14. Overall the data was not found to be significantly different from the untreated (p=0.33).

The relative total copy number in the AICAR + ddI condition was opposite of that found in the AICAR only (without ddI) condition. Depletion was evident from experimental day 3 at 0.89 (±S.D 0.25) which continued to decrease throughout the exposure. The relative copy number was at 0.08 (±S.D 0.01) by experimental day 14, representing maximal depletion level. The depletion levels were found to be statistically significant compared to the untreated condition (p=0.02).

The relative total copy number in the resveratrol treated condition (no NRTI) showed a progressive increase throughout the exposure, reaching a maximum by the end of the exposure at experimental day 14 at 4.26 (±S.D 0.59). The upward increase increasing in copy number was found to be statistically significant (p=0.02).

The relative total copy number in the resveratrol + ddI condition was found to be very different to that of the resveratrol only condition with a progressive depletion throughout the entire exposure phase with the maximal level of depletion was found at experimental day 14 at 0.0018 (±S.D 0.001). The relative copy number data was found to be statistically significant when compared to the untreated ($p<0.01$).

The relative total copy number of the pioglitazone (no NRTI) condition was found to be 2.73-fold (±S.D 0.12) greater than the untreated at experimental day 3 than the untreated. This level decreased to 0.85 (±S.D 0.23) by experimental day 7. The relative copy number level fluctuated throughout the remainder of the experiment. The data overall was not found to be significantly different from the untreated ($p=0.23$).

The total copy number in the pioglitazone + ddI condition was found to initally increase in line with the pioglitazone only condition to 2.42 (±S.D 0.17) by experimental day 3. The relative copy number was then found to decrease throughout the rest of the exposure phase, reaching an experimental depletion maximum at 0.22 (±S.D 0.03) by experimental day 14.

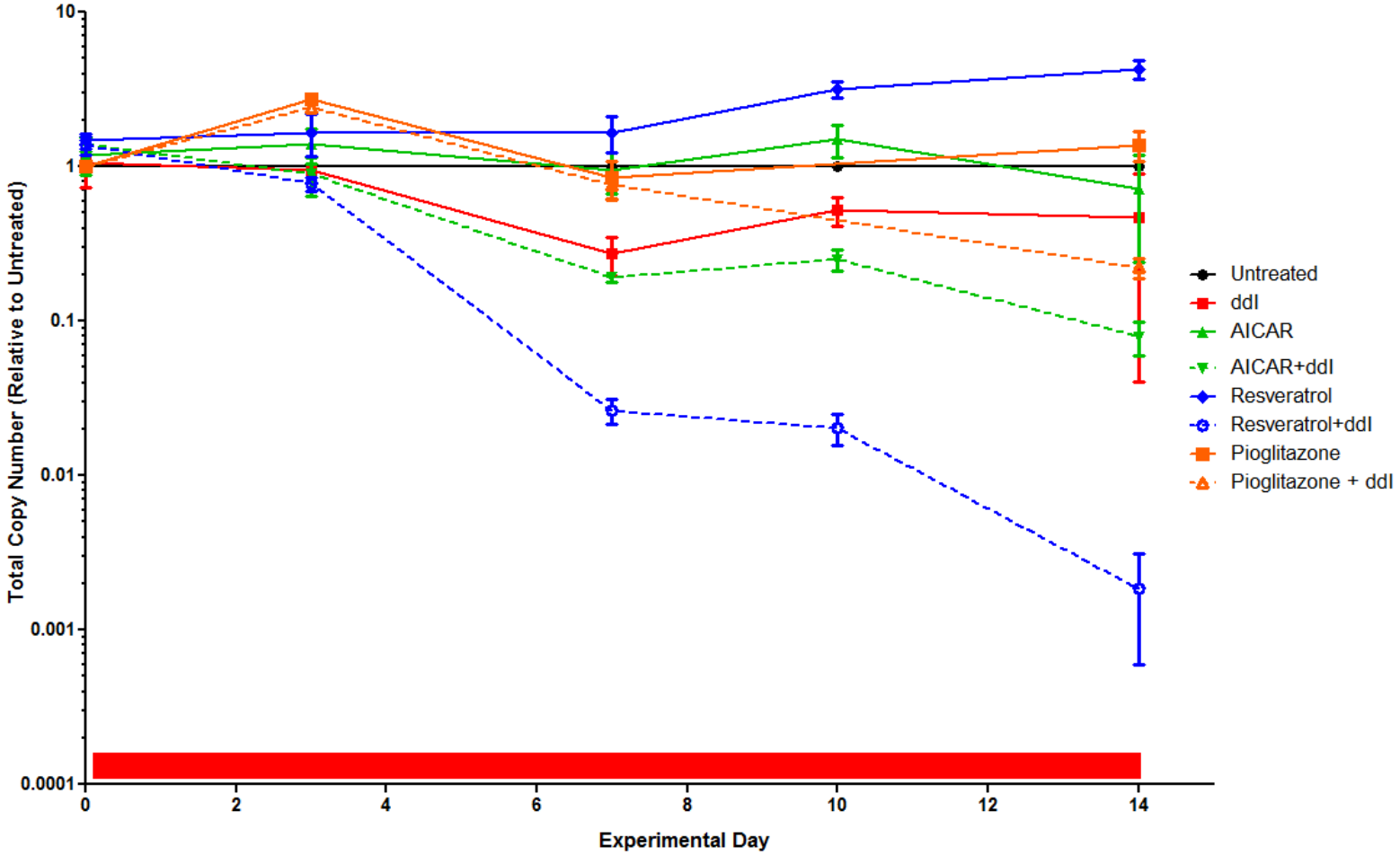

Figure 8.2 The total copy number of trans-mitochondrial cybrids during the ddI co-supplementation assay. Red bar indicates exposure phase and dotted lines represent the NRTI co-supplementation condition. Error bars are standard deviations of the mean.

The mutant copy number (Figure 8.3) was found to behave in a very similar manner to the total copy number across all conditions.

The mutant copy number in the ddI condition was found to reach exposure depletion maximum of 0.29 (±S.D 0.08), at experimental day 7. The level increased throughout the remainder of the experiment to 0.49 (±S.D 0.19). The level of depletion throughout exposure was found to be significantly different from the untreated (p=0.02).

The mutant copy number in the AICAR (no NRTI) condition was found to fluctuate throughout the exposure phase in a highly similar manner to that seen with the total copy number. There was found to be a reduction in the mutant copy number by the end of the exposure phase of 0.71 (±S.D 0.47). The fluctuation levels were not found to be significantly different than the untreated (p=0.48).

The mutant copy number in the AICAR+ ddI condition decreased throughout the entire exposure phase. The experimental depletion maximum value was reached at the end of the exposure phase of 0.08 (±S.D 0.02). The decreasing levels in mutant copy number were found to be statistically significant when compared to the untreated (p=0.01).

The mutant copy number in the resveratrol condition (no NRTI) was found increase throughout the experiment before reaching the exposure maximum level of 4.37 (±S.D 0.63) at experimental day 14. The increased copy number throughout the exposure was found to be significant compared to the untreated (p=0.02).

The mutant copy number in the resveratrol+ ddI condition was found to deplete to 0.79 (±S.D 0.11) by experimental day 3. The decrease continues throughout the exposure phase as seen in the total copy number before reaching the depletion maximum at experimental day 14 at 0.0016 (±S.D 0.001). The decrease was found to be significant (p=0.01) when compared to the untreated.

The mutant copy number in the pioglitazone (no NRTI) exposure condition was found to increase by almost 3-fold at experimental day 3 (2.94 ±S.D 0.26). The relative mutant copy number then fluctuates throughout the exposure phase before reaching 1.43 (±S.D 0.35) at experimental day 14. The changes were not found to be statistically significant, despite the high copy number at experimental day 3.

The pioglitazone+ ddI behaves in a highly similar manner to the pioglitazone only exposure for the first 7 experimental days with large fluctuations in relative mutant copy

number. The end of the exposure phase, experimental day 14, was found to be very different to the pioglitazone only condition with a large decrease in mutant copy number found at 0.24 (±S.D 0.04). The data overall was not significantly different from the untreated condition (p=0.71).

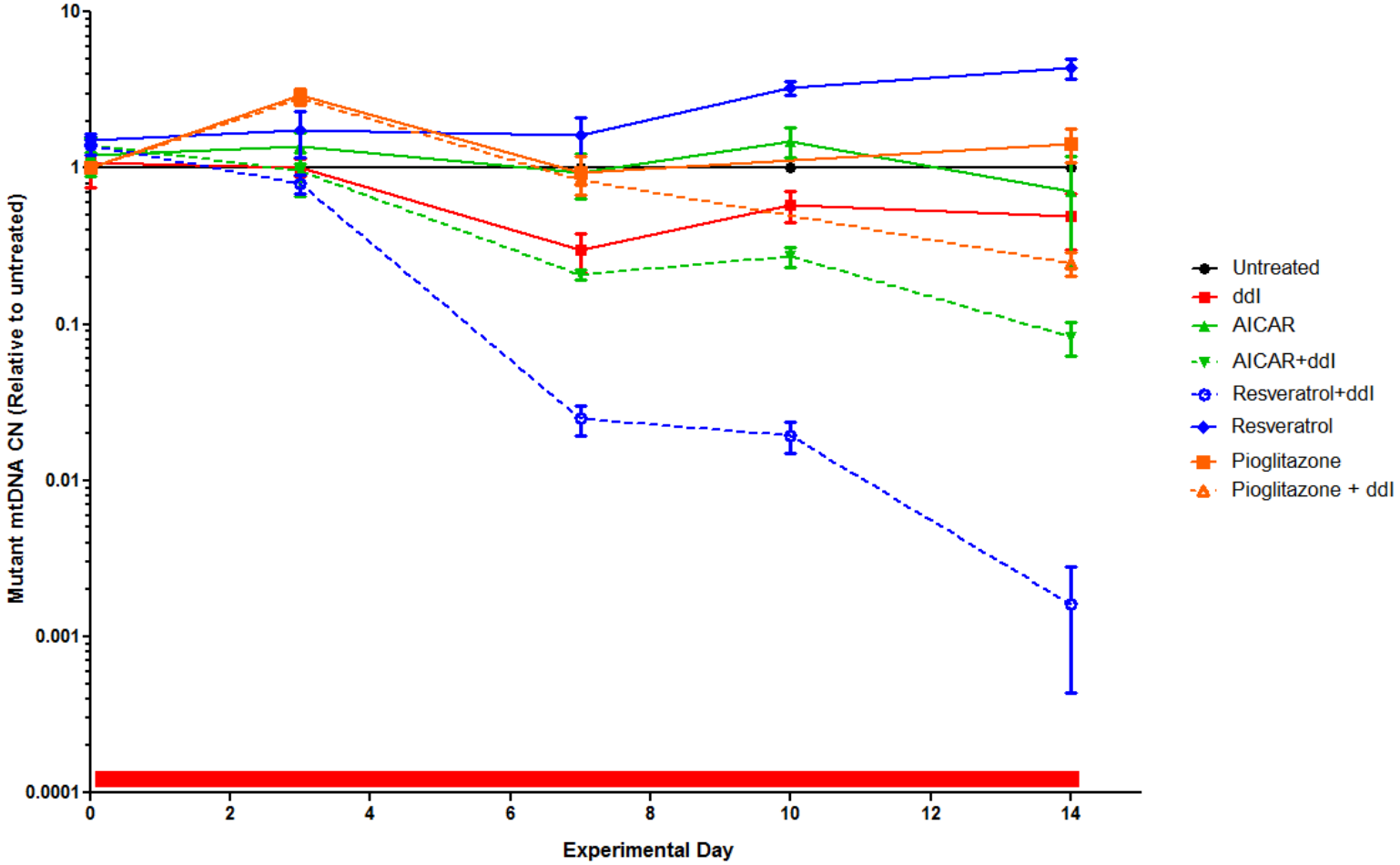

Figure 8.3 The mutant copy number of trans-mitochondrial cybrids during the ddI co-supplementation assay. Red bar indicates exposure phase and the dotted lines represent the ddI co-supplementation conditions. Error bars are standard deviations of the mean.

The relative wild type copy number data (Figure 8.4) indicates that the ddI condition suffers severe depletion from the onset. At experimental day 3, the wild type copy number was found to be 0.55 (±S.D 0.04). The decrease reached a depletion maximum at experimental day 14 of 0.03 (±S.D 0.015). The depletion was found to be highly significant compared to the untreated condition (p<0.0001).

The relative wild type copy number for the AICAR condition (no NRTI) was found to fluctuate throughout the entire exposure phase, with an initial increase to 1.51 (±S.D 0.66) by experimental day 3. The exposure depletion maximum was found to be 0.86 (±S.D 0.49) at experimental day 14. The data was not found to be significantly different to the untreated (p=0.19).

The AICAR+ ddI condition displayed severe depletion during the exposure phase. The wild type copy number was found to severely deplete by experimental day 3 to 0.48 (±S.D 0.07), which continued throughout the exposure phase, before reaching depletion maximum of 0.01 (±S.D 0.0009) at experimental day 14. The depletion was found to be highly significant when compared to the untreated condition (p<0.0001).

The relative wild type copy number in the resveratrol only condition (no NRTI) was found to have a sustained increased level throughout the exposure phase compared to the untreated condition. The level ranged an experimental maximal value of 2.33 (±S.D 0.41) at experimental day 10 and slightly decreased to 2.25 (±S.D 0.13) at experimental day 14. The data was found to be statistically different from the untreated (p=0.03).

The relative wild type copy number during the resveratrol+ ddI condition was found to deplete throughout the entire exposure phase. By experimental day 7, the value was found to significantly plummet to 0.02 (±S.D 0.009). The level remained the same through experimental day 10 and found to reach an experimental depletion maximum of 0.006 (±S.D 0.002). The decreased levels were found to be highly significant compared to the untreated (p<0.001).

The pioglitazone (no NRTI) exposure was found to induce a fluctuating response in the relative wild type copy number, which was highly similar to the untreated condition throughout the entire experiment. The difference was not found to be statistical significant (p=0.62).

The relative wild type copy number during the pioglitazone+ ddI exposure was found to initially increase to 1.50 (±S.D 0.04) by experimental day 3before fluctuating throughout the exposure phase and reaching an experimental depletion maximum at experimental day 14 of 0.04 (±S.D 0.01). The data was not found to be statistically significant (p=0.45), despite the depletion present at the end of the exposure phase.

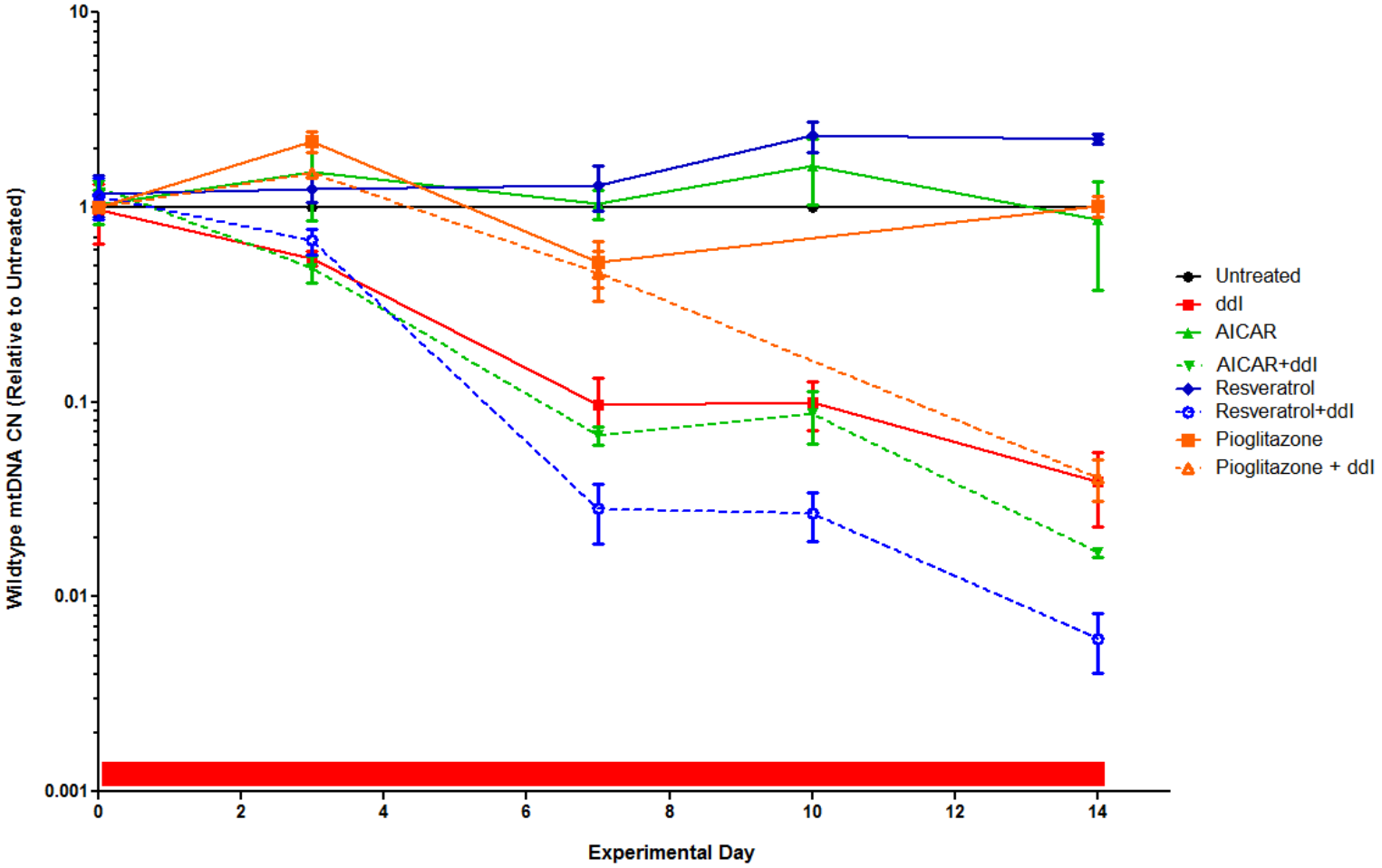

Figure 8.4 The wild type copy number of trans-mitochondrial cybrids during the ddI co-supplementation assay. Red bar indicates exposure phase, dotted lines indicate ddI co-supplementation condition. Error bars are standard deviations of the mean.

8.4.2 Effects of supplement conditions on a mitochondrial DNA deletion heteroplasmy distribution

The mitochondrial deletion heteroplasmy level distribution in each of the co-supplementation groups has been graphed and expressed as deletion percentage for the cybrid cells (Figure 8.5). Each data point is expressed as the mean values from the biological and qPCR technical replicates, with error bars plotted as standard deviations.

The most striking feature of the heteroplasmy levels across the condition is the difference in shifts between the conditions. Four conditions, ddI, AICAR+ ddI, resveratrol and Pioglitazone+ ddI, showed an increased heteroplasmy during the experiment compared to the untreated. The AICAR, resveratrol + ddI and pioglitazone only conditions remain at a comparable level to the untreated.

Compared to previous experiments, the untreated condition was found to have a slight increasing trend in heteroplasmy level by the end of the experiment. The distribution ranged from 66.44% (±S.D 1.91%) to 72.86% (±S.D 0.60%).

The ddI condition was found to progressively increase in deletion heteroplasmy from the onset of the exposure phase. At experimental day 3, the heteroplasmy was increased to 79.42% (±S.D 0.36%). The level continued to increase throughout the exposure phase. The experimental maximal value was found at experimental day 14, with a level of 98.51% (±S.D 0.2%). The increase was found to be statistically significant compared to the untreated (p=0.03).

The heteroplasmy level during AICAR exposure (no NRTI) was found to be highly similar to the untreated condition (p=0.30). Very little change was found during the exposure phase with the heteroplasmy level increasing in the latter half of the exposure phase with the maximal value for the AICAR condition of 72.81% (±S.D 1.20), at experimental day 14.

The heteroplasmy level was found to increase progressively in the AICAR+ ddI exposure condition, in a similar manner to the ddI only condition. The starting heteroplasmy level of 70.56% (±S.D 0.91%) then increased to an experimental maximum value at experimental day 14 at 95.45% (±S.D 0.18%). The increase was found to be statistically significant compared to the untreated condition (p=0.02), although highly similar to the heteroplasmy level of the ddI only condition.

The heteroplasmy level during resveratrol (no NRTI) exposure was also found to have an increasing trend throughout the experiment; although, it wasn't found to be significant compared to the untreated condition (p=0.15). The experimental maximum value was found at experimental day 14 at a heteroplasmy level of 86.76% (±S.D 0.66%).

The heteroplasmy level in the resveratrol+ ddI condition was found to show little variation during exposure. The heteroplasmy level was found to increase to 75.86% (±S.D 2.18%) at experimental day 7, which was found to be the experimental maximum. The final heteroplasmy level was calculated at experimental day 10 to be 74.08% (±S.D 1.97%). The changes were not significantly different between day 0-10 (p=0.64). There was no experimental day 14 heteroplasmy level calculated due to such little difference between the Ct values, corresponding with the low copy number levels.

The heteroplasmy of the pioglitazone condition (no ddI) was found to fluctuate throughout the experiment from 59.39% (±S.D 0.4%) to the final heteroplasmy level at experimental day 14 of 70.74% (±S.D 1.21%). The data was not found to be significant compared to the untreated (p=0.64).

The heteroplasmy level of the pioglitazone+ ddI condition was found to increase throughout the exposure phase and reached an experimental maximum level of 92.75% (±S.D 0.37%) at experimental day 14.

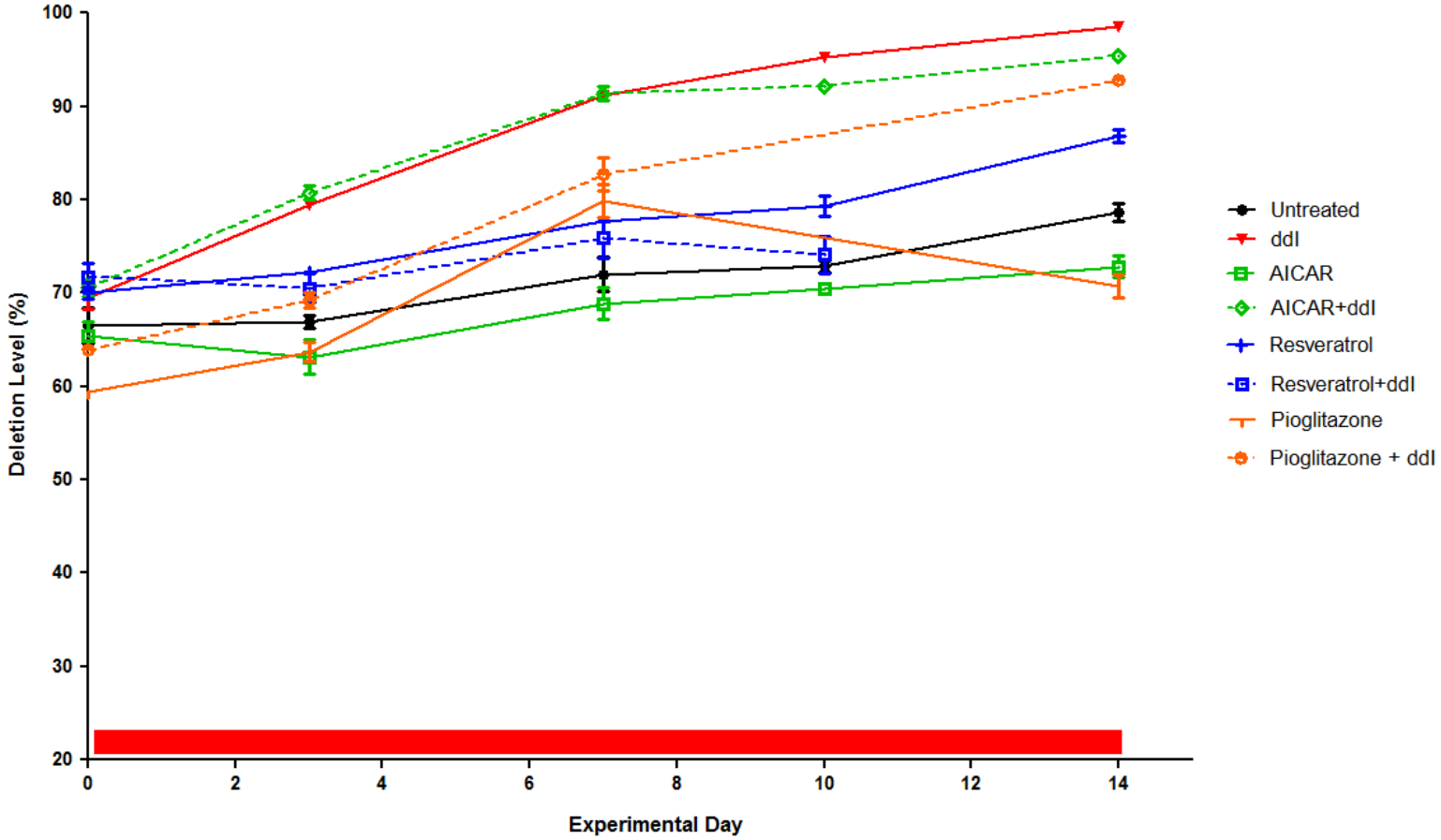

Figure 8.5 The distribution of the large scale deletion heteroplasmy level in the trans-mitochondrial cybrids during the ddI co-supplementation assay. Red bar indicates exposure phase. Error bars are standard deviations of the mean.

8.5 Discussion

The experimental aim of this study was to assess whether compounds associated with up-regulating mitochondrial biogenesis, could help reduce the amount of mitochondrial copy number depletion and prevent heteroplasmy shift of a large scale deletion (as previously discussed in Chapter 5). This was carried out using trans-mitochondrial cybrid cells which contained a large scale deletion and exposing them to the NRTI, ddI. Single-therapy exposure of each compound (AICAR, resveratrol, bezafibrate and pioglitazone) was performed as well as co-supplementation (ddI + compound).

8.5.1 Bezafibrate data summary

Bezafibrate acts as a pan-agonist for all three PPAR isoforms, PPAR α, β and γ. Although it is a licenced pharmaceutical for the clinical treatment of dyslipidaemia, the up-regulation of PGC-1α from PPAR activation during exposure, has been indicated to have beneficial effects on mitochondrial function. These include improved OXPHOS function and respiratory capacity *in vitro* (Bastin *et al.*, 2008; Srivastava *et al.*, 2009).

The cells within this experiment were found to live no longer than 48 hours in the presence of bezafibrate at 100µM. All three biological replicates were found to have a markedly reduced replicative capacity before eventually dying off. This is surprising based on the results from previous *in vitro* studies which used a higher concentration of bezafibrate (250µM and 400µM) for a longer exposure time (five days) than managed in this study, with successful up-regulation of PGC-1α and a positive impact on the mitochondria (Noe *et al.*, 2013).

8.5.2 AICAR data summary

AICAR, an analogue of AMP, has been found to be a strong activator of AMPK and an up-regulator of mitochondrial biogenesis. AICAR has been described as a potentially beneficial compound for patients with a mitochondrial complex I deficiency (Golubitzky *et al.*, 2011). However, there are no reports of copy number behaviour in the setting of antiretroviral therapy.

The copy number (total, mutant and wild type) in the AICAR only condition was found to fluctuate throughout the exposure phase. Although, there was an increase present after 3 days of exposure, in all DNA species, of approximately 40%, this increased level appeared to be temporary and transient. With a longer exposure time illustrating a copy number became comparable with the untreated condition. The heteroplasmy level did

not change much throughout the exposure, despite an increase in copy number at experimental day 3 and 10. This suggests that there is no preferential increase of either of the mtDNA species (wild type or mutant) when copy number is increased.

However, the AICAR+ ddI condition indicated a more severe depletion and accelerated onset of depletion when compared to the ddI only condition. The decrease was found to be significant worse than the ddI condition (p=0.02). The depletion of both mtDNA species was evident, but more so in the wild type DNA, suggesting the smaller mutant molecule maintained a replicative advantage. Wild type copy number was depleted to 1% of the untreated, whereas the mutant was depleted to only 7% of the untreated by the end of the exposure phase of 14 days. The increase in heteroplasmy was eventually identical to the ddI only condition.

8.5.3 Resveratrol data summary

Resveratrol is a naturally occurring compound that has been used in recent studies due to its ability to activate SIRT1 of the PGC-1α pathway. It has been found to preserve mitochondrial function and reduce ROS production *in vitro* (B. Wang *et al.*, 2014a) and that there's a strong link between increasing SIRT1 expression and mitochondrial copy number (Sato *et al.*, 2014).

The mitochondrial copy number in the resveratrol only condition was found to steadily increase throughout the exposure phase to a significantly higher level than the untreated (p=0.02). After 14 days exposure, the total copy number was found to be more than 4-times greater than the untreated condition. However, the heteroplasmy level of the deletion was also found to increase during the exposure and by the experimental day 14, it was found to be raised to 86.76% (±S.D 0.66%). This is due to an increased level of mutant copy number, despite both species increasing, the wild type only increased 2.25-fold greater than the untreated; whereas the mutant increased 4.37-fold greater than the untreated.

The resveratrol+ ddI condition showed profound depletion. At experimental day 14, the copy number of all mtDNA species was found to be <1% of the untreated condition. The depletion was also found to deplete to a lower level than the ddI alone condition in all mtDNA species. Despite the decrease in copy number, the heteroplasmy remains constant, unlike the increase in heteroplasmy in the ddI alone condition. In the AICAR condition (Section 8.5.2, above), there was an increased heteroplasmy level due to

mutant copy number being maintained at a higher level than the wild type (fitting with a size dependent mechanism); however, the resveratrol+ ddI condition appears to decrease both the mutant and wild type and the same rate. This may be due to the fact that the depletion occurs so rapidly, the differential difference in size makes no difference. Alternatively, of course, there could be complete cessation of mtDNA turnover.

8.5.4 Pioglitazone data summary

Pioglitazone is a member of the thiazoline family of pharmaceuticals used in the management of diabetes, clinically. There's growing evidence to suggest that PPAR agonists such as pioglitazone, are able to induce mitochondrial biogenesis (Guilherme *et al.*, 2008). This induction has been best demonstrated in adipose tissue where mitochondrial function in cells has been found to improve (Rosen and Spiegelman, 2006).

The total copy number in the pioglitazone only condition was found to increase after a short exposure time of 3 days where it was raised to 2.73-fold greater than the untreated. All mtDNA species were subject to a temporary, initial increase in copy number, which was not sustained and decreased back to a similar level to the untreated by the end of exposure. The fact that pioglitazone increases PGC-1α through the same pathway as bezafibrate but resulted in an increased copy number initially, with no apparent negative impact on the cells, is certainly worth noting.

The pioglitazone+ ddI condition was found to suffer the same fate as the other co-supplementation conditions with little protective effect present. A significant improvement was present after 3 days, suggesting a short term protective effect. This was found to quickly disappear by experimental day 7 with depletion present of 24% (total). The heteroplasmy of the sample increased during the exposure as the depletion occurred, with more depletion present in the wild type copy number. The heteroplasmy at experimental day 3, when the copy number was found to increase, remained at a similar level to experimental day 0. This indicates that the induction of biogenesis wasn't mtDNA species specific.

8.5.5 Study limitations

This study carried some limitations. There was no cell line specific dose toxicity assessment performed and concentrations were based on previous studies using a range

of different cell types, from primary to immortal. This approach may not give a good indicator as to the best concentrations used in these cells and could be a contributing factor as to why the bezafibrate appeared to be so toxic. A future study could select one or two compounds from the compounds used in this study for a more comprehensive analysis implementing a dose-ranging design.

The cell line used was an osteosarcoma cybrid cell line and some studies have found that cancerous cell lines are severely affected (slow in cell replication with defects in mitochondrial OXPHOS and induction of apoptosis) in the presence of compounds that have been found to be beneficial in 'healthy/normal cells', i.e. AICAR (Jose *et al.*, 2011). This may explain why there was a limited beneficial effect seen in the cells over a long exposure phase.

8.6 Chapter conclusions

This is the first study to the author's knowledge that has assessed the behaviour of an mtDNA large scale deletion and the mtDNA copy number in the presence of NRTI exposure, during dual-therapy or 'co-supplementation' of an NRTI with a compound known to up-regulate mitochondrial biogenesis.

The most striking feature of this study is the general negative impact that the 'beneficial' compounds have on cells when co-supplemented with ddI. The compounds appeared to exacerbate or potentiate the depletion effect of ddI within the cells, with severe copy number depletion and increased heteroplasmy throughout exposure present; despite in some instances a transient increase in copy number.

The longer exposure time (compared to previous studies of these compounds) used in this study highlighted that the compounds which have previously been found to be beneficial to mitochondrial function after a short exposure phase i.e. bezafibrate and AICAR, was found to be only transient and not beneficial with longer exposure time.

This data gives insights into what may happen to individuals taking these compounds in the long term and suggest the only real beneficial compound is resveratrol, with significantly increased mitochondrial copy number. However, the large scale deletion was found to increase throughout the exposure and suggests that an amplification of pre-existing mutations may indeed occur in otherwise healthy individuals and these compounds may prove to further accelerate the HIV-associated ageing phenomena described previously in this thesis in individuals whom are receiving NRTI therapy.

Chapter 9. Genetic Susceptibility to Severe Mitochondrial-mediated Side Effects in HIV-infected Malawian's treated with Stavudine

Table of Contents

9.1 Background

Patients receiving the NRTI, stavudine (d4T), present with high rates of peripheral neuropathy (PN) and lipodystrophy (LD). Although d4T administration has been virtually eliminated in the developed countries, there are a large number of HIV-infected individuals still receiving d4T in the developing countries of sub-Saharan Africa, specifically through 'antiretroviral rollout' programmes due to the low cost and availability of the drug.

While the link between d4T exposure and PN/LD is suggestive of a causative role (Menezes *et al.*, 2011; Pujades-Rodriguez *et al.*, 2011; Singh *et al.*, 2014), the underlying mechanisms between d4T toxicity and mitochondria have not been fully elucidated.

Pol γ is encoded by two nuclear genes: *POLG* (chromosome 15q25) and *POLG2* (chromosome 17q) (Longley *et al.*, 1998); *POLG* encodes for the catalytic subunit of pol γ and *POLG2*, encodes for the accessory or processivity subunit of pol γ. Monogenic mutations of the catalytic subunit have been associated with mitochondrial related disorders.

A large array of genetic abnormalities and secondary mitochondrial defects have been associated with inherited mutations in *POLG* (Hudson and Chinnery, 2006). However, the impact of such mutations within the setting of anti-retroviral treatment is largely unknown.

There have been a very small number of studies assessing the extent to which genetic predisposition plays a role in NRTI induced pathology, through *POLG* polymorphisms. Two polymorphisms that have been associated with lipodystrophy and neuropathy are R964C and E1143G, respectively (Yamanaka *et al.*, 2007; Lee-Jun C. Wong *et al.*, 2008). However, African SNPs in *POLG* are not very well characterised, and given that *POLG* in the Caucasian population is known to be only mildly polymorphic, it could be that it is even less polymorphic in Africans.

9.2 Experimental Aim

The aim of this study is to assess whether severe d4T toxicity (PN and LD), in African individuals is mediated by monogenic mutations in *POLG*.

9.3 Experimental design and methods

9.3.1 Patient Cohort

DNA (extracted using Nucleon DNA extraction kit, Tepnel life science, Manchester, UK) from a cohort of Malawian HIV-infected patients receiving d4T (n=253) were sent on dry ice from Blantyre, Malawi (van Oosterhout *et al.*, 2012) to Newcastle upon Tyne, UK. Based upon their clinical assessment 12 months after initiation of d4T therapy, the patients were categorized into three groups; no side effects, currently receiving d4T; mild side effects, currently receiving d4T; severe side effects, stopped d4T therapy. There were fourteen severely affected individuals, of which, ten samples yielded good quality DNA and were included in this study. All patient had given informed consent and the study was approved by the institutional review board.

9.3.2 Whole genome amplification

Due to the low volume of DNA that was provided for each sample, whole genome amplification, REPLI-g (Qiagen, Manchester, UK) was used. REPLI-g works on an isothermal genome amplification system which works down to DNA concentration of 10ng/µL. to each 1µL of template (10ng/µL), 1µL of buffer D1 was added and mixed by vortexing. The samples were incubated at room temperature for 3 minutes before adding 2µL of buffer N1. Then a master mix of 16µL was added to each well (15µL reaction buffer and 1µL Ultrafast polymerase). This was incubated at 30^{o}C for 16hrs before a 3 minute polymerase denaturation step at 65^{o}C. The samples were stored at -20^{o}C until required.

9.3.3 PCR

POLG is comprised of twenty-three exons. Of these, only twenty-two are actually coding while 'exon one', which was initially thought to be a coding region, was later found to be non-coding and therefore, it was excluded from the sequencing of these patients. A series of primers were designed for each of the twenty-two coding exons. PCR of each sample was performed as previously described (materials and methods Section 3.4) using the sequences as detailed below in Table 9.1. The PCR reaction was mplified on a Veriti® thermocycler (Applied Biosystems, Life Technologies, Paisley, UK) with cycling conditions consisting: an initial denaturation step of 95^{o}C for 10 minutes, followed by 30 cycles of denaturation at 95^{o}C for 1 minutes, an annealing step

at T_m (see Table 9.1) for 1 minute and an extension step at 72^{o}C for 1 minute. A final extension step at 72^{o}C for 10 minutes was performed before cooling the samples to 4^{o}C.

All products were imaged on a 2% agarose gel (w/v) as per materials and methods (Section 3.5) and an approximate quantification of each product was made using an arbitrary comparison of band strength compared to the DNA ladder Hyperladder IV (Bioline, London, UK) to ensure that the correct size was amplified with no contamination present in the negative control.

Exon	Forward Sequence	Reverse Sequence	T_m
2a	GCCAGTAAAAGAAGCCAAGC	CTCCTTGCCCGAAGATTTG	63
2b	AGCCGCAAGTGCTATCCTC	GGCCTGCAACAGCAAGTT	63
2c	GACAACCTGGACCAGCACTT	AACACATCAGCGCTCCCTAC	63
3	TAGGTGTGCAGTGGTTGTTG	AACCACTGAGATTAGGGCTC	63
4	TCCACACCACCAAGCAGTGGT	AGAGGGGGTCCCAAGCACTAT	65
5+6	ATAGTGCTTGGGACCCCCTCT	TACCAGGAACACACTGACC	65
7	ATGGGATGATATTGTTCCCATTT	AGTCCACTAGGGCAGGGCTA	65
8	GCTCTCAGGAGAGAGGTAGCC	GGGAAGACAATCAGGAGCAG	55
9	AGGTAGGGTAGGGTAGGGGT	CTGAGAATGGAGCAAGGGTA	63
10	GGGACATTGTGAGAGAGAGA	CACTCTTTCCACTAGCCTGAG	61
11+12	CAGAGTGGGCATCTGGTAAT	AAGAGGAAGCCCTTTCCACC	61
13	ACAGTTTCAGGCCCTTTTCC	TGTGCCTGAAATCACACTCTG	61
14	AGGTTCTGGGCTCAGTGTTG	GGCACCAGGACCAAAAGTAG	61
15+16	AGTGAGGCTGGGTAATGGAG	CAGGGTCCTTTTCATGATCC	61
17	TCTCTAAAGCCATCCCCTCAG	AGCTCAGGAACATTCTGC	63
18	GCATGCATGGTGAGCAGGAG	GTAATGGGCAGGAGATAGAACAG	65
19+20	TGAACATTCCTTGCCAAGGC	TCTGCCCATGCTCCAAAGGTA	61
21	GCTTCTACCCTGGAGTTAATTG	CAAGGAACGCTCACCCAAAG	65
22	GTCATTGCTCCAGGAGTGAT	CTGAGTCAAGAGTGGATTCTC	65
23	CCTTTTGACCTTAGCATTAAGC	CTACTGAAAAATGGCTGGCC	65

Table 9.1 POLG primers all primers were designed using primer-BLAST and the POLG reference sequence NC_000015.9. A gradient for each primer was performed ranging from 55 – 65^{o}C to determine the optimum cycling condition.

9.3.4 ExoFAP-IT Protocol

The excess of PCR reagents were removed from the samples using the ExoFAP-IT (Thermo Fisher Scientific Inc., Loughborough, UK) clean up system to prevent any interaction with the subsequent sequencing reaction. The reaction works by implementing two hydrolytic enzymes to breakdown the excess primers and dNTPs, Exonuclease I (Exo I) and shrimp alkaline phosphatase (FastAP), respectively. The clean ups were performed in the 96 well full skirted sequencing plate (Applied Biosystems, Life Technologies, Paisley, UK) using 20ng of PCR product with 0.5µL Exo I and 1µL FastAP, on ice. The reaction was incubated at 37^{o}C for 15 minutes before inactivation of the enzymes by 15 minute incubation at 85^{o}C.

9.3.5 BigDye® Terminator v3.1 sequencing

The BigDye® Terminator v3.1 sequencing chemistry (Applied Biosystems, Life Technologies, Paisley, UK) was used to prepare the products for sequencing immediately after clean up. For each reaction, 1µL of BigDye® was added to 2µL BigDye® sequencing buffer, 1µL primer of sequencing directionality (10µM) , 5µL of the cleaned ExoFAP-IT product and made up to 20µL using PCR grade water. The reaction was heated to 96^{o}C for 1 minute, followed by 25 cycles of 96^{o}C for 10 seconds, 50^{o}C for 5 seconds and 60^{o}C for 4 minutes, in a PCR Thermocycler (Applied Biosystems, Life Technologies, Paisley, UK). The reaction was performed in the same 96 well sequencing plate as the ExoFAP-IT reaction.

9.3.6 Ethanol precipitation

Ethanol precipitation was performed to purify the sequencing reaction by removing any excess salt and to concentrate the DNA. The reaction was performed in the 96 well sequencing plate. To each well 2µL of 125mM EDTA (Ethylenediaminetetraacetic acid) with 2µL of 3M sodium acetate solution was added. 70µL of 100% ethanol was then added and the plate sealed, mixed by inversion and incubated at room temperature for 15 minutes. The plate was span at 2000g for 30 minutes, then the supernatant removed by gentle spinning at 100g for 10 seconds. 70µL of 70% ethanol was then added to each well and then a further spinning process at 1650g for 15 minutes. The supernatant was removed in the same manner again before air drying in the dark for 10 minutes to allow full evaporation of excess ethanol to prevent inhibition of the sequencing reaction.

9.3.7 Hi-Di™ re-suspension and sequencing analysis

Each well with precipitated product was re-suspended in 10µL of Hi-Di™ (Applied Biosystems, Life Technologies, Paisley, UK) which is a highly deionized formamide that facilitates the product sequencing in a capillary electrophoresis system. The plate was denatured at 95^{o}C for 2 minutes prior to being loaded into the ABI 3130*xl* genetic analyser (Applied Biosystems, Life Technologies, Paisley, UK). The data was analysed using SeqScape® v2.1.1 (Applied Biosystems, Life technologies, Paisley, UK).

9.4 Results

No pathogenic mutations were found in exons 3-23 in any of the ten patient samples; the sequencing results of exon 2 (nine patients, one sample produced poor results) are reported in Table 9.2, no significant CAG repeat variation was detected; One sample presented with a SNP in exon 2, p.R42Q (see Table 9.2), which has been previously reported (rs74382477) and predicted not to be disease causing.

Patient Number	Sex	Age	Months Treatment	Sequencing Observation
55	F	30	12	CAG 10/10 Homozygous
73	F	44	18	CAG 10/10 Homozygous
94	F	33	24	CAG 10/11 Heterozygous
137	F	48	15	Poor Sequencing across exon
143	M	58	24	CAG 10/10 Homozygous, p.R42Q
166	M	61	15	CAG 10/11 heterozygous
193	M	54	18	CAG 10/10 Homozygous
218	F	30	18	CAG 10/11 Heterozygous
238	F	30	12	CAG 11/11 Homozygous
249	F	31	15	CAG 10/11 Heterozygous

Table 9.2 Exon 2 CAG *POLG* sequencing observations indicate no abnormal repeat variation from that found in the general population.

9.5 Discussion

There were no mutations found in the samples sequenced; despite previous literature quoting two monogenic mutations in *POLG* associated with lipodystrophy (R964C) and neuropathy (E1143G) (Yamanaka *et al.*, 2007; Lee-Jun C. Wong *et al.*, 2008). Although this study didn't find an association with these mutations, it may be due to the rarity in the population and a larger cohort would have revealed an association.

Another consideration is the geographical origin of the sample, which raises the question as to whether or not the quoted mutations were not found in this study due to the distinct differences in the mitochondrial background haplotype (African vs European). A subsequent study on this patient cohort revealed that in those that did not present with lipodystrophy, there was a distinct prevalence of sub-haplogroup L3e, suggesting a protective effect (Kampira *et al.*, 2013). The presence of R964C and E1143G may be rare in European individuals, but potentially even more so in African's.

An alternative theory to monogenic *POLG* mutations being pathogenic is that R964C and E1143G are merely very rare polymorphisms that mediate an increased susceptibility for toxicity. This can also be scientifically supported in an evolutionary sense, as natural selection would not have selected against NRTI toxicity-susceptible determinants.

The $CAG_{(n)}$ repeat found in exon 2, encoding a polyglutamine tract, has been associated with disease, depending upon the number of repeats (Rovio *et al.*, 2001; Anvret *et al.*, 2010). 10 CAG repeats has been found to be present in 88-96% of the Eurasian population (Malyarchuk *et al.*, 2005), with up to three CAG additional repeats on an allele deemed as normal variation; deleting the CAG repeat was found to have no effect on respiratory chain function *in vitro* (Spelbrink *et al.*, 2001). Although this data is not available for African populations, it was assumed that 10-13 CAG repeats would also be normal in Malawians. On this basis, none of the patients presented with an abnormal CAG repeat region, suggesting a further lack of association with *POLG* mutations and LD/PN.

In addition to *POLG* mutations, there have been a number of other genetic associations with PN. In a multinational study of 294 patients, there was a strong positive correlation with age and height with developing d4T PN (Cherry *et al.*, 2009). A study from the same group also showed that Australian HIV-infected individuals receiving d4T were

more susceptible to PN when expressing TNF alpha-1031*2 cytokine profile (Cherry *et al.*, 2008).

Despite the fact that this study was relatively small with only ten patient samples expressing severe side effects, it was the first to the author's knowledge where genetic predisposition to NRTI toxicities has been assessed in sub-Saharan African subjects.

9.5.1 Study Limitations

The two major drawbacks of this study was the small cohort of patients and the lack of baseline neuropathy data for each patient limited the conclusions drawn from the data.

9.6 Chapter Conclusions

In conclusion, the data suggests that *POLG* monogenic mutations are not associated with d4T associated mitochondrial toxicity in Malawians. A larger population study would improve the resolution of these claims and increase the chance of detecting any rare mutations in *POLG*. Therefore, it cannot be fully ruled out that *POLG* mutations may contribute to severe toxicities in d4T treatment in a small number of individuals.

Chapter 10. General Discussion and Concluding Statement

Table of Contents

10.1 Research overview

The research described within this thesis was undertaken to further understand the underlying mechanism(s) driving the apparent mtDNA defects present in HIV-infected individuals receiving NRTI therapy.

The aims of the thesis were to address the following questions:

1. To develop *in vitro* models to study the behaviour of mitochondrial DNA mutations, both deletions and point mutations, during and after the exposure of NRTIs.

2. To design and implement methods of detecting very low level (<1% heteroplasmy) mitochondrial DNA mutations through ultra-deep next generation sequencing technologies and apply them within the setting of NRTI exposure.

These questions have been addressed by implementing a range of molecular techniques and developing new models and methodologies to build upon the current understanding in the literature regarding the effect of NRTI therapy on HIV-infected individuals. The following sections consist of a discussion of the results found and how it relates to the wider issue of ageing in HIV.

10.2 Anti-HIV therapy and implications for mitochondria

The level of mitochondrial DNA mutations, both point mutations and deletions, is generally accepted to increase with age in healthy individuals and has been characterised in a number of post-mitotic tissues, and has been associated with functional cellular mitochondrial defects; specifically, an increased level of mitochondrial COX-deficiency (N. Arnheim and Cortopassi, 1992a; Kopsidas *et al.*, 1998; Khrapko *et al.*, 1999; Del Bo *et al.*, 2003; Bua *et al.*, 2006; Yu-Wai-Man *et al.*, 2010; P. Wang *et al.*, 2013; L. C. Greaves *et al.*, 2014).

The exposure to NRTIs in HIV-infected individuals has been shown to result in a range of pathophysiological conditions such as neuropathy, myopathy and lipodystrophy, driven through a mitochondrial mechanism (Lim and Copeland, 2001; Kohler and Lewis, 2007). A known hierarchy of pol γ inhibition of NRTIs is well accepted along with the notion of mtDNA depletion and the polymerase γ hypothesis (Höschele, 2006).

The data presented in Chapter 4, highlighted the results from an assessment of the level of mtDNA large scale deletions, the mtDNA common deletion (CD) and the level of an age associated point mutation (Del Bo *et al.*, 2003) m.414T>G, in skeletal muscle of three groups of individuals: healthy controls, HIV^+/NRTI-naïve individuals and HIV^+/NRTI-exposed individuals.

The data from this study indicates that there is an association of increased m.414T>G heteroplasmy level, in muscle homogenate DNA from patients receiving NRTIs. This association also correlates with the hierarchy of pol γ inhibiting NRTIs, with the strongest associations made with ddC and ddI, which are documented as being very strong inhibitors of pol γ activity (Lim and Copeland, 2001).

Amongst the HIV-infected individuals who were NRTI-naïve, as well as the healthy controls, there was only one sample (healthy control) that displayed a minor increased level (1% increase) in the point mutation compared to the background level. The increased level of m.414T>G suggests an accelerated intrinsic ageing effect of NRTI exposure.

These observations were extended to large scale deletion mutations. There was a significant increase of both total large-scale deletions, and the specific ‘common deletion’ (CD) in NRTI-exposed individuals compared to NRTI-naïve and control individuals, irrespective of age. The exposure to strong pol γ inhibiting drugs also

correlated with the increased CD level and large scale deletions. The number of large scale deletions in NRTI-exposed individuals was also found to correlate with multiple large scale deletions in individuals who have been exposed to ddI, d4T and ddC.

These findings were correlated with previous clinical data, and the increased deletion levels (both CD and large scale deletions) positively correlated with an increased level of COX-deficient fibres, potentially leading to increased clonal expansions. This further supports previous claims of an increased level of mtDNA deletions in COX-negative fibres of patients who had received NRTIs (Maagaard *et al.*, 2006; Payne *et al.*, 2011).

The conclusion drawn from this study is that NRTI-mediated effects are comparable with the changes expected much later in life in association with normal ageing, especially when exposed to strong pol γ inhibitors. However, as the exposure to strong pol γ inhibitors in these patients is a historical occurrence (as with the majority of HIV-infected individuals in industrialised countries), the fact that the mtDNA defects are still present indicates a non-reversible phenomenon with implications for future health of such patients. The observations define the basis for studying mechanisms of clonal expansion in the rest of the thesis.

10.3 Clonal expansion and HIV-therapy

The literature contains a wealth of data for NRTI exposure and mtDNA copy number depletion, with an established NRTI pol γ inhibition hierarchy defined (Höschele, 2006). There is limited data however, that highlights the impact of NRTIs on mtDNA mutations during exposure in HIV-infected individuals, and no previous exploration of mechanisms.

There are three plausible mechanisms for clonal expansion in HIV-infected individuals: a molecular bottleneck effect from copy number depletion, a replicative advantage through a size dependent mechanism and thirdly *de novo* mutagenesis.

10.3.1 Mitochondrial DNA deletion mutation behaviour

An assessment for a molecular bottleneck mechanism was performed firstly with the use of multiple biological replicates for both normal conditions and NRTI-exposure.

The cells exposed to a strong pol γ inhibiting NRTI, ddI, showed an increase in heteroplasmy throughout the exposure, corresponding with a depletion of wild type mitochondrial DNA copy number. After exposure, where a bottleneck was created through copy number depletion, the heteroplasmy across all conditions slowly reduced, as opposed to an increased spread of heteroplasmies across the different flasks. These data were then further supported by assessment of single cells across multiple replicates. Following ddI exposure, all single cell heteroplasmy levels of the deletion were found to be shifted towards the same heteroplasmy level, comparable with the level found in the homogenate analyses. These data suggests that there is no molecular bottleneck effect on deletions.

To further explore this issue, two assessments for a replicative advantage mechanism were performed. The use of a physiological and supra-physiological dose of cells containing a large scale deletion was employed.

The first suggestion in the literature of a mutant size dependent replicative advantage in the presence of NRTIs was described in fibroblasts of patients with Kearns-Sayre syndrome exposed to the NRTIs, ddC and AZT (strong and mild pol γ inhibiting activity, respectively). The wild type mtDNA level was found to deplete whereas the mutant level appeared no show no change after exposure (H. Wang *et al.*, 1996). *Wang and colleagues* used a primary cell line to study the KSS deletion, which is usually rapidly lost during cell culture and not an ideal model to use in such a study; therefore,

the model used in this thesis is much preferable for long cell culture experiments with repeatability.

It was found that the deletion level in cells exposed to the strong pol γ inhibiting NRTI, ddI, increased throughout exposure due to depletion of the wild type mtDNA, but little depletion in the mutant mtDNA copy number. This is suggestive of a replicative advantage of the mutant (deleted) molecule, resulting in an increased heteroplasmy level of the deletion and supports the long lived idea that mtDNA deletions accumulate in a size dependent manner (D. C. Wallace, 1992). This effect was also seen in the cells exposed to the supra-physiological dose but at a substantially faster rate, as well as in the d4T treatment (a mild pol γ inhibitor), illustrating a linear relationship between size dependent accumulation and a 'brake' on mtDNA replication. Furthermore, in all NRTI conditions and at supra-physiological dose, there were no new deletions found during or after exposure, in any of the cell lines studied.

These data therefore describes a novel method for studying mtDNA deletions for long periods of time *in vitro*. The findings were extended to post-mitotic embryonic neuronal cells exposed to physiological and supra-physiological dosing of NRTIs, along with the expected depletion pattern across NRTI conditions.

10.3.2 Mitochondrial DNA point mutation behaviour

In a similar vein to the assessment of mtDNA deletion mutation behaviour through the exposure of NRTIs, an assessment was made for mtDNA point mutations (Chapter 6).

The fundamental difference between deciphering between these possible mechanisms of mutation behaviour lies within the findings of the NRTI exposure phase and the repopulation phase. An increase of heteroplasmy during exposure would indicate a replicative advantage as with the increase seen with an mtDNA deletion in both the HIV-setting (Chapter 5) and the non-HIV setting (Diaz *et al.*, 2002); Whereas a bottleneck mechanism would manifest most obviously, after the nadir of depletion.

The exposure of NRTIs to cells containing a point mutation proved to have little overall effect on the heteroplasmy level. The phase of interest was found to be during the repopulation where there was a dispersion of heteroplasmies across multiple replicates after exposure to pol γ inhibiting NRTIs. These data neatly fits with a bottleneck mechanism given the extent to which mtDNA content depletion occurs, creating a limited pool of molecules during repopulation. The shifts seen are then produced

through the removal of pol γ replication inhibition and an expansion of the restricted pool of point mutation heteroplasmy. This is also in keeping with the data presented by *Cree and colleagues* with the segregation of mtDNA genotypes during embryogenesis through a bottleneck mechanism of mtDNA content reduction (Cree *et al.*, 2008).

These data do not completely rule out the potential of a mutageneic effect of NRTIs inducing *de novo* mutations during exposure. This has been supported by numerous studies highlighting an increase in mtDNA mutations in a range of animal models and *in vitro* work. However, another longitudinal study failed to detect any new mtDNA mutations through NRTI exposure (G. McComsey *et al.*, 2005a; Lehmann *et al.*, 2011; K. Liu *et al.*, 2013; Y. Zhang *et al.*, 2014).

The issue was therefore addressed through the next generation sequencing (NGS) of cells exposed to NRTIs during and after exposure (Chapter 7). The major drawback of the aforementioned studies was their inability to differentiate between mutagenesis and clonal expansion; the former was therefore assumed. By implementing NGS to assess for *de novo* mutagenesis and the qPCR data discussed here, a clearer picture was formed.

10.3.3 Mitochondrial DNA *de novo* mutagenesis

Firstly, before analysing cells from the *in vitro* experiments (Chapters 5 & 6), the methodology of analysing very low level variants using NGS was sought.

The optimal analysis of very low level mtDNA mutations is debated and there have been various studies implementing ultra-deep next generation sequencing approaches to address the question of *how low can we go?* Despite an array of different methodologies implemented and designed to answer such question, there have been similar results among a handful of papers defining the low level detection threshold of mtDNA mutations, is in the 0.2% heteroplasmy range (Y. He *et al.*, 2010; Mingkun Li *et al.*, 2010; Payne *et al.*, 2013).

The fundamental issue with the aforementioned studies is that they never addressed biological samples in the optimisation of the analyses. Therefore, in Chapter 7, I present samples from a rare mitochondrial disorder known to display a stereotypical mutation motif (MNGIE; nAT>C) (Nishigaki *et al.*, 2003), that were utilised in a novel way to define the low level mutation pattern, through NGS technologies. It was concluded that the statistically lowest threshold that signal exceeds noise in a biological sample is at

≥0.22%, which is strikingly similar to previous studies. To ensure a greater stringency and confidence in calling very low level mutations, the threshold was set to ≥0.3%.

The newly defined threshold level and analysis pipeline was used to investigate the issue of *de novo* mutagenesis of NRTIs. A novel approach of sequencing *in vitro* cells during and after NRTI exposure was performed in two different cell lines. The results proved the same in both studies, there was no increase in total number of mutations, during or after NRTI exposure, including strong pol γ inhibitors, such as ddI, and thymidine analogues such as d4T, with have previously been suggested to be possibly mutagenic (Carter *et al.*, 2007).

By comparing the shared mutations across all conditions of the system, it was possible to measure the heteroplasmy shifts of a number of point mutations simultaneously. The findings revealed little change during NRTI exposure; however, after the NRTIs were removed, in particular in the strong pol γ inhibitors, there was a statistically significant shift of heteroplasmy levels that coincided with the recovery of mtDNA content. These data strongly supports the findings described in Chapter 4 and Chapter 6, of a molecular bottleneck mechanism underlying point mutation behaviour.

The analysis of the NRTI exposed, trans-mitochondrial cybrids was extended by performing a unique analysis of sequencing specifically the deleted/mutant mtDNA and the wild-type/undeleted mtDNA, within NRTI-exposure conditions. By modelling the individual mtDNA species, it was possible to analyse the behaviour of pre-existing mutations through ddI exposure on both mtDNA species, as well as the behaviour of the wild-type molecule mutations in the presence and absence of all NRTIs used in this thesis. These preliminary data reveal for the first time, a distinct difference between mutation behaviour on the wild-type and the mutant molecules, with a preferential bias for mutation heteroplasmy accumulation on the deleted molecule in an apparent ‘piggy-back’ mechanism in the untreated molecules.

10.3.4 Mitochondrial DNA mutagenesis hot spots

Despite the analyses discussed in the previous section, it cannot be completely ruled out that *de novo* mutagenesis does not occur, especially in particular hot spots for mutation accumulation such as the d-loop.

Mice exposed to the NRTI, AZT, were found at the end of 2 years of exposure to have an increased number of mutations in the d-loop compared to an age matched control.

There was no difference found in the coding region between the mice. It is possible however, that AZT is maybe unique in this situation, with previous claims of AZT specific mutagenesis. This is plausible through an inhibition of pol γ exonuclease proof-reading activity (Lim and Copeland, 2001). Alternatively, there is evidences to suggest that inhibition of mitochondrial thymidine kinase by AZT (a thymidine analogue) would be sufficient to create a nucleotide pool imbalance (purine/pyrimidine levels) and thus induce mtDNA mutations; this mechanism has recently been proposed as a mechanism present in other NRTIs too, such as ddI (Rylova *et al.*, 2005; Sun *et al.*, 2014). However, these data is very preliminary and limited conclusions can be drawn from it.

The main issue a mutagenesis hypothesis is that low level induced mutations would require a long time to clonally expand to the high levels seen in the middle-age patients presented within this thesis (Chapter 4), this idea is also substantiated by previous modelling data of point mutations (Elson *et al.*, 2001). Clonal expansions are more likely to have more functional affect anyway, as they compromise the function of individual cells, whereas mutagenesis of mutations will be low level and therefore, not functional relevant immediately (unless they subsequently go through clonal expansions).The other problem is an imbalance in nucleotide pool in metabolism disorders, such as MNGIE, result in specific mutation motifs or an overwhelming type of mutation created due to the imbalance of nucleotides; this was not the case in any of the AZT exposed conditions sequencing in this study.

10.4 Genetic predisposition

Finally, an under researched area that could contribute towards mitochondrial toxicity in HIV-infected individuals, is a genetic predisposition to certain pathologies. There have been only a small number of studies assessing the extent to which genetic predisposition contributes towards the described NRTI-induced pathology, specifically through *POLG* (pol γ encoding gene) polymorphisms. There have namely been two polymorphisms that have been associated with common pathologies, lipodystrophy and neuropathy, which were R964C and E1143G, respectively (Yamanaka *et al.*, 2007; Lee-Jun C. Wong *et al.*, 2008).

As the majority of developed countries no longer use strong pol γ inhibiting NRTIs, it is difficult to gain samples for a large genetic predisposition study; however, the use of such drugs in sub-Saharan Africa is still highly prevalent, making HIV-infected individuals from this region a highly useful for studying a potential genetic predisposition to pathologies.

A cohort of 10 Malawian HIV-infected subjects exposed to stavudine for approximately 12 months, displaying severe side effects such as neuropathy and lipodystrophy, were assessed for monogenic *POLG* mutations for associations with NRTI induced pathology (Chapter 9). The selection of a severe phenotype was intended to amplify any genetic signal. There were no pathogenic mutations found in the samples sequenced. Although this study didn't find *POLG* mutations, this may be due to their likely rarity in the population. A larger cohort may therefore have revealed a patient with an underlying *POLG* mutation.

The overall conclusion of this dataset is limited due to the small number; however, this pilot study does indicate that monogenic *POLG* mutations are unlikely to make a major contribution towards a susceptibility to severe stavudine-induced side effects.

10.5 Reversing Mitochondrial DNA defects from NRTI exposure

The mtDNA copy number depletion present during NRTI therapy is well established and has been described in all of the *in vitro* studies presented in this thesis. However, there have been only a small number of recent studies assessing co-treatment with compounds (indicated as mitochondrial biogenesis up-regulators) that may reduce mitochondrial damage during NRTI exposure; namely, Acetyl-L-carnitine, co-enzyme Q, uridine and most recently, vitamin D (Hart *et al.*, 2004; U. A. Walker *et al.*, 2006; Cherry *et al.*, 2010; Campbell *et al.*, 2013; Sinha *et al.*, 2013).

The consensus on the effects of these treatments on mitochondria in general is not clear, and the specific effect on mtDNA mutations has not been investigated. Therefore in Chapter 8, I set out to assess whether co-supplementation could help reduce the amount of mitochondrial copy number depletion and prevent a heteroplasmy shift of a large scale deletion *in vitro,* when exposed to an NRTI.

Despite previous claims in the literature that mitochondrial protection was conferred through the exposure of bezafibrate (Noe *et al.*, 2013), the cells in this study that were exposed to bezafibrate were unable to survive, despite the use of a lower concentration than the *Noe and colleagues* study. This may be explained by previous claims that bezafibrate increases the lowered respiratory function of cancerous cells, and therefore slows the growth (X. Wang and Moraes, 2011). The entire system in this study was comprised of cancerous cells (cybrids from osteosarcoma nuclear background); therefore, when the slowing growth effect of bezafibrate affected all cells in the system, it appeared that all of the cells had died due to a toxicity of the exposure.

AICAR and pioglitazone were found to both have a protective function in the short term (first three days of exposure) by themselves, which is in line with previously found beneficial effects of AICAR and pioglitazone on mitochondria (Jose *et al.*, 2011; Sauerbeck *et al.*, 2011). However, after this timeframe, there was a negative effect seen on mtDNA copy number when exposure exceeded five days.

In the presence of an NRTI, the mtDNA copy number was depleted quicker and more severely than the NRTI condition only. This is contradictory to what has been recently published, with AICAR being heralded as a ‘mitochondria builder’ and provides protection against other mitochondrial disorders such as Alzheimer’s (Du *et al.*, 2014). Myotubes exposed to *in vitro* to AICAR have been found to increase mitochondrial

biogenesis, however decrease oxygen consumption over time (Spangenburg *et al.*, 2013). Both conditions increased the level of mitochondrial deletion heteroplasmy during exposure with the NRTI, although there was no difference in the AICAR and pioglitazone only conditions, when compared to the untreated condition.

The only condition within this study that proved to be beneficial to increasing mtDNA copy number in the long term was resveratrol which was found to increase mitochondrial copy number throughout the exposure (no ddI). Resveratrol has been found in a number of short exposure studies to have beneficial effects on mitochondria and implicated to have protective properties against metabolic disease through an affinity for DNA repair and modulation of cellular metabolism (Davinelli *et al.*, 2013; B. Wang *et al.*, 2014a; R. Wang *et al.*, 2014b). However, in the presence of an NRTI, the mitochondrial copy number depleted severely and to the greatest extent across all conditions (including NRTI).

The heteroplasmy was found to increase in the resveratrol only condition. The expansion of the deleted mtDNA was accelerated through a non-specific increase in mitochondrial biogenesis. This defies the logic of increasing mitochondrial copy number after depletion and indicates that in individuals receiving such treatment, as with those after ischaemia (R. Wang *et al.*, 2014b), exposure to resveratrol may improve mitochondrial function but it may also accelerate expansion of pre-existing mtDNA deletions. This could prove to have detrimental effects in the long term and may be relevant to normal ageing as well as HIV/NRTIs.

Overall, this study provides valid insights into relatively long term *in vitro* exposure of compounds used in the up-regulation of mitochondrial biogenesis and that they may not always exhibited a positive impact on mtDNA. This indicates that the up-regulation of biogenesis through PGC-1α is only beneficial when mtDNA is allowed to freely replicate (unlike in the case of pol γ inhibition). In the presence of ddI, the incorporation of the drug maybe increased through a faster depletion driven by the faster turnover when biogenesis is greatly up-regulated.

The main conclusion drawn from these data for the HIV setting is that improving mitochondrial copy number through up-regulating biogenesis during NRTI-exposure is an inappropriate method which may exacerbate the effect of NRTIs on mtDNA defects.

10.6 The role of mitochondrial ageing in HIV infection

Overall, this thesis has highlighted four key aspects that are vital for the understanding of issues surrounding HIV accelerated ageing and mitochondrial genetics:

- There is a lack of evidence to suggest that NRTIs induce *de novo* mutagenesis in mitochondrial DNA
- Mitochondrial DNA deletions appear to accumulate in a size dependent manner in the presence of pol γ inhibitors
- Mitochondrial DNA point mutations are subject to a molecular bottleneck driven through the depletion of mtDNA content
- There is insufficient evidence to suggest that NRTI-induced mtDNA copy depletion and expansion of mutations can be beneficially modulated through pharmaceutical co-supplementation

The expansion of mtDNA deletion mutations is supportive of the current paradigm that mtDNA deletions accumulate through clonal expansion with age and contribute towards aged diseases, such as Alzheimer's and Parkinson's (Swerdlow and Khan; Bender *et al.*, 2006). Such pathologies have also been found to present in young HIV-infected individuals (Anthony *et al.*, 2006). The pathologies generally present in tissues of HIV-infected individuals that with high energy demand and are known to be related to mitochondrial dysfunction (Schapira, 2012).

Rather than the previously described mutagenesis theory of point mutation accumulation through NRTI exposure, this thesis has neatly refuted this idea in a multitude of conditions and cell types; however, there still remains the possibility that through time, the D-loop could be a hot spot for mutagenesis driven through NRTIs. This maybe a consequence of the reduced proofreading activity of pol γ, exerted through NRTI incorporation (W Lewis *et al.*, 2003), thereby mimic the effects seen in the polymerase γ mouse model (Trifunovic *et al.*, 2004) and accelerating ageing.

The mtDNA defects presented here are expected to be irreversible and therefore, will still pose issues for the individuals, despite them no longer receiving the pol γ inhibiting NRTIs. The issue also remains as to the long term effects of mild or weak pol γ inhibitors on HIV-infected individuals. Although, little to no defects was seen during the *in vitro* studies (Chapters 5 & 6), the defects may only appear after prolonged use (that is not representable *in vitro*).

However, how relevant are these molecular changes on ageing, and do they have any functional relevance for HIV-infected patients? It is possible that patients do have an increased level of mtDNA damage, but due to the incomplete understating of ageing, the defects go unnoticed. I would suggest that increased mutations would ultimately drive an accelerated physiological decline in multiple tissues and organs within the body (as seen in normal human ageing). The challenge is to measure this decline in HIV patients. This is an issue in ageing research in general, as the definitive longitudinal studies are very difficult to do. The use of simian model would likely address this issue, although the lengthy nature of such a study may not be feasible. Alternatively, it would be worth correlating the mtDNA data of patients in a large-scale epidemiological study with age-associated diseases in HIV patients; the mtDNA findings of a longitudinal study could also be correlated with physiological measurements in HIV patients.

Although these data strengthen the case in general for mitochondrial involvement in HIV-associated ageing, given the complexity of the normal human ageing process, we should also consider other potential aspects of ageing in HIV infection, including: the effects of NRTIs on other ageing-associated molecular pathways, the effects of other anti-retroviral drug classes, and the effects of the virus itself. For example, there is clinical data to suggest that certain NRTIs (including AZT and d4T) can also cause telomere shortening and contribute toward accelerated ageing (Strahl and Blackburn, 1996; Bollmann, 2013). Consideration must also be given to oxidative stress exerted by protease inhibitors (Xinwen Wang *et al.*, 2007), which might also contribute to the molecular damage seen in ageing (including to mitochondria), and raises the question that it is not only NRTIs that are the causative agent of HIV-associated ageing. Overall however, the effects of NRTIs provide the most compelling case for promoting and accelerating ageing, in HIV-infected individuals.

10.7 Concluding remarks

This thesis has started to explore the complex mechanisms underlying mtDNA mutation load increase through NRTI exposure. The data are suggestive that there are two separate mechanisms at play when considering the clonal expansion of mtDNA mutations: a molecular bottleneck for point mutations, and a replicative advantage for deletions. There is little if any evidence for mutagenesis. This thesis is the first research that attempts to pull apart these mechanisms, with the aim of an increased understanding that would ultimately aid the management of HIV-infected individuals.

Overall I conclude that the answer to the question posed in 2012 by *Fisher and colleagues* of '*HIV and ageing: premature ageing or premature conclusions?'* (Fisher and Cooper, 2012) is that the data within this thesis support the notion of an acceleration of normal cellular 'intrinsic' ageing. This appears to be propagated through mitochondrial mechanisms in HIV-infected individuals that present as irreversible changes in mtDNA. The implications of these findings are most pertinent to individuals who are (or have previously) been exposed to strong pol γ inhibiting NRTIs, raising the issue of a unique iatrogenic effect that not only has implications on the HIV-infected community, which should be followed longitudinally to assess the ultimate health consequence(s). Due to recent data suggesting that clonal expansion may be the most important factor affecting mtDNA in normal ageing, here I present a unique model for manipulating it.

Chapter 11. Appendices

Appendix A

Appendix A consists of the relevant clinical and research data associated with the patient presented in Chapter 4.

ID	Group	CD - Log10	Number of Deletions	m.414 (%)	COX-fibres (%)	Age	Months Treated	ddI Duration	d4T Duration	ddC Duration	AZT duration	Months D-drug
DH340N	NAIVE	-5.76	0	2	0	24.8	0	0	0	0	0	0
MG400N	NAIVE	-5.25	2	2	0.28	48.8	0	0	0	0	0	0
MT460N	NAIVE	-5.95	0	2	0	23.3	0	0	0	0	0	0
MJ250N	NAIVE	-3.28	0	2	0.5	46.3	0	0	0	0	0	0
KW623N	NAIVE	-5.59	0	2	0.18	45.3	0	0	0	0	0	0
PH535N	NAIVE	-4.22	0	2	0	50.2	0	0	0	0	0	0
SR326N	NAIVE	-4.07	0	2	0.11	31.8	0	0	0	0	0	0
MM456N	NAIVE	-6.2	0	2	0.11	26.9	0	0	0	0	0	0
SH630N	NAIVE	-5.78	1	2	0.07	34.2	0	0	0	0	0	0
PM420N	NAIVE	-6.15	2	2	0	32.2	0	0	0	0	0	0
IB525N	NAIVE	-4.53	0	2	0	27.4	0	0	0	0	0	0
DD120N	NAIVE	-5.18	0	2	0	36.9	0	0	0	0	0	0
BJ520A	NRTI	-4.49	0	1	0.13	47.1	103.2	0	0	0	103.2	0
CA200A	NRTI	-4.33	2	2.5	0	30.9	42.2	0	0	0	42.2	0
CC200A	NRTI	-4.49	0	2	0.19	46.3	37.2	0	0	0	37.2	0
DE152A	NRTI	-6.23	0	2	0	38.3	100.1	0	0	0	100.1	0
DM456A	NRTI	-5.26	2	2	0.14	30.9	36.3	0	0	0	36.3	0
IM200A	NRTI	-3.4	2	2.5	0.28	32.4	93.6	0	0	0	93.6	0
JO400A	NRTI	-4.33	2	2	0	29.6	49.3	0	0	0	49.3	0
KG420A	NRTI	-3.36	1	2	0.39	44.3	51	0	0	0	51	0
MW320A	NRTI	-4.04	0	9	0	42.3	125.7	0	0	0	125.7	0
PW322A	NRTI	-4.69	0	2	0	32.8	97.6	0	0	0	97.6	0
RF500A	NRTI	-3.25	1	2	1.59	35.5	139.6	0	0	0	139.6	0

SA255A	NRTI	-6.03	0	5.5	0.68	47.9	29.8	0	0	0	29.8	0
SB616A	NRTI	-5.44	0	2	0	34.5	82.3	0	0	0	82.3	0
SM000A	NRTI	-4	0	4	0.04	33	54.3	0	0	0	54.3	0
SS162A	NRTI		0	2	1.94	49.5	96.7	0	0	0	96.7	0
BM254S	NRTI	-4.85	0	2.5	0.1	45.3	70.4	0	0	0	40	0
CM250S	NRTI	-3.39	2	2.5	0.84	39.8	54	0	0	0	51	0
IJ525S	NRTI	-2.91	2	4.5	1.3	49.4	193.3	1	1	1	10	89
JC462S	NRTI	-4.05	1	6	4.9	48.2	151	1	1	1	9	57
MH340S	NRTI	-5.69	0	2	1.4	43	118.2	0	0	0	101	0
PD145S	NRTI	-2.89	2	3	2.8	50.2	138.5	1	1	0	4	72
CR255T	NRTI	-4.49	1	3		39	12.4	0	0	0	0	0
DK200T	NRTI	-4.16	2	2		39.5		0	0	0	0	0
JS163T	NRTI	-5.38	0	2.5		25.9	33.6	0	0	0	0	0
PM640T	NRTI	-3.53	0	2		50.4	17	0	0	0	0	0
TR200T	NRTI	-4.31	0	2		34.2	17.8	0	0	0	0	0
M9	MCTB	-5.6	0	2	0.2	42	0	0	0	0	0	0
M12	MCTB	-4.08	0	2	0	24	0	0	0	0	0	0
M13	MCTB	-4.92	0	2	0	21	0	0	0	0	0	0
M17	MCTB	-4.14	0	3.5	0.2	52	0	0	0	0	0	0
M7	MCTB	-3.74	0	2		36	0	0	0	0	0	0
M6	MCTB	-3.79	0	2		47	0	0	0	0	0	0
M4	MCTB	-3.84	0	2		18	0	0	0	0	0	0
M2	MCTB	-4.78	0	2		31	0	0	0	0	0	0

Table 11.1 The patient clinical and research data from HIV-infected individuals used in the study presented in Chapter 4.

Appendix B

Appendix B consists of the MNGIE and control variants presented in the next generation sequencing pipeline optimisation section in Chapter 7 (Section 7.4.1).

Position	5' Sequence	VarFreq	Ref	Var
6767	TGTG	0.20%	A	G
7385	AGTG	0.20%	A	G
8196	AAACC	0.20%	A	G
9399	AA	0.20%	A	G
13269	AGG	0.20%	A	G
9145	TGTC	0.20%	G	A
4569	A	0.20%	G	T
4703	AACAA	0.20%	T	A
3183	ATA	0.20%	T	C
4708	AC	0.20%	T	C
5873	CG	0.20%	T	C
9088	CCC	0.20%	T	C
11661	CCA	0.20%	T	C
12563	CCCAGC	0.20%	T	C
13254	CCAC	0.20%	T	C
13271	GAC	0.20%	T	C
1590	C	0.21%	A	G
9639	AATC	0.21%	A	G
1243	CC	0.21%	T	C
2017	C	0.21%	T	C
2279	A	0.21%	T	C
3344	CCCA	0.21%	T	C
5162	TATC	0.21%	T	C
6225	CCC	0.21%	T	C
6367	GGTG	0.21%	T	C
8548	CGCT	0.21%	T	C
8705	AACCA	0.21%	T	C
9850	TTTCC	0.21%	T	C
11770	CAC	0.21%	T	C
12477	TCAG	0.21%	T	C
13260	AAG	0.21%	T	C
14502	AAA	0.21%	T	C
15976	AAC	0.21%	T	C
4662	A	0.22%	A	G
7542	TTAC	0.22%	A	G
7138	AAAA	0.22%	T	A
504	A	0.22%	T	C
6235	AC	0.22%	T	C
7668	CCC	0.22%	T	C
9219	AA	0.22%	T	C
9592	AAG	0.22%	T	C
10512	ATC	0.22%	T	C

11517	CCCCC	0.22%	T	C
12255	AACA	0.22%	T	C
14979	AA	0.22%	T	C
15744	CC	0.22%	T	C
4727	AT	0.23%	A	G
6720	GT	0.23%	A	G
13836	AAC	0.23%	A	G
2636	G	0.23%	G	A
6620	GG	0.23%	T	C
6844	A	0.23%	T	C
8746	A	0.23%	T	C
9367	A	0.23%	T	C
12005	AA	0.23%	T	C
14952	A	0.23%	T	C
3003	G	0.24%	A	G
3711	GC	0.24%	A	G
12900	CATG	0.24%	A	G
8250	AAATAG	0.24%	G	A
2510	AACA	0.24%	T	C
4028	AA	0.24%	T	C
5311	A	0.24%	T	C
6076	CG	0.24%	T	C
8786	GGAC	0.24%	T	C
9321	CCAC	0.24%	T	C
10264	AAAT	0.24%	T	C
14166	AA	0.24%	T	C
2148	AGAG	0.25%	A	G
6629	A	0.25%	A	G
9964	CTGT	0.25%	A	G
13224	AAAA	0.25%	T	A
11299	TCAC	0.25%	T	C
12797	TTGC	0.25%	T	C
13433	A	0.25%	T	C
13488	ACC	0.25%	T	C
15831	AA	0.26%	T	A
1119	AAA	0.26%	T	C
3055	AAG	0.26%	T	C
3631	CC	0.26%	T	C
11895	AAC	0.26%	T	C
15511	AA	0.26%	T	C
2392	AA	0.27%	T	C
7749	AACA	0.27%	T	C
11324	AAAC	0.27%	T	C

1995	AA	0.28%	A	G
10263	AA	0.28%	A	G
2256	CC	0.28%	T	C
3964	CCCC	0.28%	T	C
10983	CCCC	0.28%	T	C
15741	CC	0.28%	T	C
15852	A	0.28%	T	C
385	A	0.29%	A	G
1358	AA	0.29%	A	G
16043	GCAG	0.29%	A	G
1974	AAG	0.29%	A	T
4396	CCCCA	0.29%	T	C
15873	AAAA	0.29%	T	C
1694	AC	0.30%	T	C
5277	AA	0.30%	T	C
9581	AAA	0.30%	T	C
12851	AA	0.30%	T	C
13369	GGG	0.30%	T	C
13520	ACA	0.30%	T	C
8108	A	0.31%	A	G
9609	A	0.31%	T	C
10432	AATG	0.32%	A	G
4703	AACAA	0.32%	T	C
8547	TCGC	0.32%	T	C
10989	AA	0.34%	T	C
14048	AAA	0.34%	T	C
9380	GATG	0.35%	G	A
7880	AAA	0.35%	T	C
9405	A	0.35%	T	C
4370	AAAA	0.36%	T	C
7275	GGC	0.36%	T	C
6121	A	0.37%	T	C
12919	CCAAC	0.37%	T	C
14540	AAAA	0.37%	T	C
2253	AAC	0.38%	T	C
2104	AA	0.39%	A	G
13135	CCCCCTA	0.39%	G	A
1632	AAC	0.39%	T	C
5136	AAC	0.40%	T	C
10403	A	0.41%	A	G
11070	AAA	0.42%	T	A
13444	ACT	0.43%	T	C
15639	CCA	0.43%	T	C

10158	AAAAA	0.44%	T	A
5726	AA	0.44%	T	C
9110	AA	0.44%	T	C
12913	ACAC	0.45%	T	C
7317	CATG	0.51%	A	G
7258	AAACA	0.51%	T	C
10805	A	0.51%	T	C
8469	CCACC	0.52%	T	A
5857	AG	0.54%	A	G
12631	ATGG	0.59%	T	C
3521	A	0.61%	T	C
6406	AA	0.62%	T	C
2371	AA	0.64%	T	C
11087	AACA	0.64%	T	C
408	TTT	0.65%	T	A
10893	AAA	0.68%	T	C
11070	AAA	0.69%	T	C
16172	AA	0.69%	T	C
1974	AAG	0.70%	A	G
10221	AAAA	0.74%	T	C
3861	TG	0.80%	A	G
7091	ACTG	0.83%	A	G
3653	AA	0.89%	T	C
16311	TAG	0.98%	T	C
3915	AAGG	1.02%	G	A
2075	AAA	1.28%	T	C
7440	AAAA	1.61%	T	C
15831	AA	1.75%	T	C
5628	AAA	1.89%	T	C
13446	TTC	2.39%	A	T
7138	AAAA	2.79%	T	C
2233	AAAAAA	2.90%	T	C
10158	AAAAA	2.97%	T	C
13879	AAAA	4.60%	T	C
189	A	4.93%	A	G

Table 11.2 The variants detected in the MNGIE skeletal muscle patient sample after bioinformatics processing. The 5' sequence is the sequence extracted before the 'Var' base, which differed from the rCRS 'Ref' base. The heteroplasmy or 'VarFreq' indicates the variant frequency found in the sample. All homoplasmic variants were removed and inverted at ≥ 90%.

Position	5' Sequence	VarFreq	Ref	Var
8129	AA	0.20%	A	G
10840	AC	0.20%	A	G
7755	AG	0.20%	A	G
2866	TCC	0.20%	A	G
4721	TG	0.20%	A	G
6146	TG	0.20%	A	G
11167	TG	0.20%	A	G
3838	A	0.20%	T	C
10510	A	0.20%	T	C
8109	AA	0.20%	T	C
5158	AC	0.20%	T	C
7943	AC	0.20%	T	C
11814	AC	0.20%	T	C
1847	CC	0.20%	T	C
4933	CC	0.20%	T	C
5892	CC	0.20%	T	C
9850	CC	0.20%	T	C
7869	CCC	0.20%	T	C
4800	CCCC	0.20%	T	C
2010	GG	0.20%	T	C
13500	GG	0.20%	T	C
6345	TC	0.20%	T	C
6640	TC	0.20%	T	C
4939	TCC	0.20%	T	C
2656	TG	0.20%	T	C
2875	AA	0.21%	A	G
14658	AA	0.21%	A	G
4529	AC	0.21%	A	G
5944	AG	0.21%	A	G
10055	AG	0.21%	A	G
14564	CC	0.21%	A	G
4128	CG	0.21%	A	G
7987	CG	0.21%	A	G
7961	A	0.21%	T	C
8634	A	0.21%	T	C
15852	A	0.21%	T	C
1694	AC	0.21%	T	C
11295	AC	0.21%	T	C
11978	AC	0.21%	T	C
13988	AC	0.21%	T	C
4480	CCC	0.21%	T	C
4492	CG	0.21%	T	C

6763	CG	0.21%	T	C
3801	CT	0.21%	T	C
11276	GC	0.21%	T	C
6620	GG	0.21%	T	C
4905	TCC	0.21%	T	C
10841	A	0.22%	A	G
6581	GG	0.22%	A	G
13269	GG	0.22%	A	G
15766	GG	0.22%	A	G
5311	A	0.22%	T	C
14060	A	0.22%	T	C
629	C	0.22%	T	C
6558	CC	0.22%	T	C
7695	CC	0.22%	T	C
8477	CCC	0.22%	T	C
1950	CG	0.22%	T	C
8079	TG	0.22%	T	C
4662	A	0.23%	A	G
4986	AA	0.23%	A	G
4021	CC	0.23%	A	G
7851	CG	0.23%	A	G
9629	GG	0.23%	A	G
3000	TCC	0.23%	A	G
3841	A	0.23%	T	C
7800	A	0.23%	T	C
4055	AC	0.23%	T	C
4248	AT	0.23%	T	C
5016	CC	0.23%	T	C
5338	CC	0.23%	T	C
11790	CC	0.23%	T	C
2069	CCC	0.23%	T	C
14050	TC	0.23%	T	C
4722	GA	0.24%	A	G
7749	A	0.24%	T	C
4903	AAA	0.24%	T	C
3847	CC	0.24%	T	C
7501	CC	0.24%	T	C
5873	GC	0.24%	T	C
11579	GC	0.24%	T	C
7854	GG	0.24%	T	C
1887	AG	0.25%	A	G
8923	GC	0.25%	A	G
2651	TCC	0.25%	A	G

15139	A	0.25%	T	C
11792	TC	0.25%	T	C
9480	TT	0.25%	T	C
12654	TG	0.26%	A	G
15383	CC	0.26%	T	C
3976	TCC	0.26%	T	C
4220	TG	0.26%	T	C
10988	A	0.27%	A	C
13395	GA	0.27%	A	G
4690	A	0.27%	T	C
5116	AT	0.27%	T	C
12011	GC	0.27%	T	C
13574	TC	0.27%	T	C
4728	A	0.28%	A	G
12177	TG	0.28%	A	G
7282	A	0.28%	T	C
6365	GG	0.28%	T	C
13836	AC	0.29%	A	G
5186	TG	0.29%	A	G
2778	GG	0.29%	T	C
3793	TCC	0.29%	T	C
12919	AC	0.30%	T	C
5162	TC	0.31%	T	C
15141	TG	0.31%	T	C
10676	TG	0.32%	C	T
7220	CG	0.32%	T	C
8548	CT	0.32%	T	C
11770	AC	0.33%	T	C
15976	AC	0.35%	T	C
482	C	0.37%	T	C
2256	TCC	0.39%	T	C
8506	A	0.43%	T	A
6076	CG	0.43%	T	C
13878	AAA	0.47%	A	G
189	A	0.59%	A	G
15301	TT	0.59%	G	A
14783	AAAA	0.67%	T	C
9540	AA	0.71%	T	C
14766	AAAA	0.75%	C	T
3394	GC	0.78%	T	C
16126	TG	0.92%	T	C
16223	CC	1.10%	C	T
11719	GG	1.16%	G	A

11914	AC	1.17%	G	A
12705	AT	1.20%	C	T
8701	A	1.43%	A	G
16519	GGG	1.44%	T	C
16342	AT	1.57%	T	C
16390	GG	1.96%	G	A
16274	AG	3.12%	G	A

Table 11.3 The variants detected in the control skeletal muscle patient sample after bioinformatics processing. The 5' sequence is the sequence extracted before the 'Var' base, which differed from the rCRS 'Ref' base. The heteroplasmy or 'VarFreq' indicates the variant frequency in the sample. All homoplasmic variants were removed and inverted at ≥ 90%.

Appendix C

Appendix C consists of all variants detected in the fibroblast cells exposed to NRTIs from Chapter 6.

Gene	Position	Ref	Var	Average depth	Var Freq	Mutation type	Nucleotide Change	Protein Change
D-Loop_A	304	C	A	3950	0.38%			
D-Loop_A	309	C	T	3823	0.85%			
D-Loop_A	414	T	G	4009	50.64%			
D-Loop_A	515	A	G	5077	0.45%			
RNR1	709	G	A	18687	0.57%			
RNR1	1024	G	A	25388	2.77%			
RNR2	1709	G	A	25785	4.18%			
RNR2	1999	A	G	25217	2.90%			
RNR2	2623	A	G	23785	1.15%			
RNR2	2844	G	A	24008	0.74%			
RNR2	3003	A	C	25927	3.01%			
ND1	4111	C	T	26296	0.56%	synonymous	c.C805T	p.L269L,
ND2	5435	C	A	6876	0.31%	synonymous	c.C966A	p.P322P,
ND2	5439	C	A	7125	0.34%	nonsynonymous	c.C970A	p.P324T,
ND2	5450	C	A	7167	0.43%	synonymous	c.C981A	p.P327P,
ND2	5894	A	G	20258	0.99%			
COX1	5992	G	A	20455	1.12%	nonsynonymous	c.G89A	p.G30D,
COX1	7055	A	G	34525	37.03%	synonymous	c.A1152G	p.G384G,
ATP6	8545	G	A	15886	9.66%	nonsynonymous	c.G19A	p.A7T,
ATP6	9057	C	T	22964	0.35%	synonymous	c.C531T	p.A177A,
COX3	9921	G	A	14433	45.17%	nonsynonymous	c.G715A	p.A239T,

ND4	12128	T	C	20978	16.65%	nonsynonymous	c.T1369C	p.F457L,
ND5	12425	A	C	18824	0.49%	nonsynonymous	c.A89C	p.N30T,
ND5	13028	C	A	8586	0.38%	nonsynonymous	c.C692A	p.P231H,
ND5	13037	C	A	8330	0.40%	nonsynonymous	c.C701A	p.P234H,
ND5	13434	A	G	15387	44.38%	nonsynonymous	c.A1098G	p.I366M,
ND5	13701	A	G	7570	2.68%	synonymous	c.A1365G	p.K455K,
ND5	13708	G	A	7195	0.38%	nonsynonymous	c.G1372A	p.A458T,
CYTB	15293	T	G	21940	1.11%	nonsynonymous	c.T547G	p.F183V,
CYTB	15543	C	T	29021	2.17%	nonsynonymous	c.C797T	p.P266L,
D-Loop_B	16093	T	C	27667	99.58%			
D-Loop_B	16129	G	A	26634	99.52%			

Table 11.4 The Variants detected in the untreated fibroblast cell line sequenced on the MiSeq using at experimental day 22.

Gene	Position	Ref	Var	Average Depth	Var Freq	Mutation Type	Nucleotide change	Protein change
D-Loop_A	174	C	A	7485	0.31%			
D-Loop_A	309	C	T	3218	0.92%			
D-Loop_A	414	T	G	3330	46.61%			
D-Loop_A	459	C	A	3135	0.32%			
D-Loop_A	515	A	G	4776	0.50%			
RNR1	709	G	A	17663	0.53%			
RNR1	1024	G	A	24191	6.85%			
RNR2	1709	G	A	24553	4.05%			
RNR2	1999	A	G	24049	2.27%			
RNR2	2623	A	G	22757	1.49%			
RNR2	2629	A	G	23980	0.54%			
RNR2	2844	G	A	22592	0.94%			
RNR2	3003	A	C	24726	2.08%			
ND1	4111	C	T	25096	0.79%	synonymous	c.C805T	p.L269L,
COX1	5894	A	G	18581	0.52%			
COX1	5992	G	A	18991	0.93%	nonsynonymous	c.G89A	p.G30D,
COX1	7055	A	G	33177	37.26%	synonymous	c.A1152G	p.G384G,
COX2	7794	C	T	16789	0.41%	nonsynonymous	c.C209T	p.A70V,
ATP6,ATP8	8545	G	A	14441	11.21%	nonsynonymous	c.G19A	p.A7T,
COX3	9921	G	A	12807	47.72%	nonsynonymous	c.G715A	p.A239T,
ND4	10938	C	A	8998	0.37%	nonsynonymous	c.C179A	p.P60H,
ND4	12128	T	C	19933	17.67%	nonsynonymous	c.T1369C	p.F457L,
tRNA-Ser2	12259	T	C	17984	0.32%			

ND5	13434	A	G	14759	46.56%	nonsynonymous	c.A1098G	p.I366M,
ND5	13701	A	G	7133	5.64%	synonymous	c.A1365G	p.K455K,
ND5	13708	G	A	6790	0.52%	nonsynonymous	c.G1372A	p.A458T,
ND5	14051	C	A	18305	0.36%	nonsynonymous	c.C1715A	p.S572Y,
CYTB	15293	T	G	21213	1.11%	nonsynonymous	c.T547G	p.F183V,
CYTB	15543	C	T	27879	1.41%	nonsynonymous	c.C797T	p.P266L,
D-Loop_B	16093	T	C	26410	99.67%			
D-Loop_B	16129	G	A	25407	99.45%			

Table 11.5 The variants detected in the ddI-exposed fibroblast sample detected using the MiSeq platform from experimental day 22.

Gene	Position	Ref	Var	Average depth	Var Freq	Mutation type	Coding Gene	Nucleotide change	Protein change
RNR1	709	G	A	14429	0.55%				
RNR1	1024	G	A	20314	4.63%				
RNR2	1709	G	A	21068	3.99%				
RNR2	1999	A	G	20933	1.57%				
RNR2	2523	C	A	19676	0.39%				
RNR2	2623	A	G	19153	1.60%				
RNR2	2629	A	G	20152	0.56%				
RNR2	2844	G	A	19180	1.14%				
RNR2	3003	A	C	20962	1.52%				
ND1	4111	C	T	20825	0.73%	synonymous	ND1	c.C805T	p.L269L,
ND2	5487	T	C	4878	0.31%	nonsynonymous	ND2	c.T1018C	p.S340P,
COX1	5992	G	A	16662	0.66%	nonsynonymous	COX1	c.G89A	p.G30D,
COX1	7055	A	G	28524	31.68%	synonymous	COX1	c.A1152G	p.G384G,
COX2	7794	C	T	13309	0.46%	nonsynonymous	COX2	c.C209T	p.A70V,
COX2	8138	A	G	12826	0.82%	nonsynonymous	COX2	c.A553G	p.T185A,
ATP6	8545	G	A	11850	9.27%	nonsynonymous	ATP6	c.G19A	p.A7T,
ATP6	9057	C	T	18157	0.30%	synonymous	ATP6	c.C531T	p.A177A,
COX3	9921	G	A	10523	41.64%	nonsynonymous	COX3	c.G715A	p.A239T,
ND4	12128	T	C	16010	15.48%	nonsynonymous	ND4	c.T1369C	p.F457L,
ND5	12418	A	C	15066	0.47%	nonsynonymous	ND5	c.A82C	p.K28Q,
ND5	12425	A	C	14512	0.50%	nonsynonymous	ND5	c.A89C	p.N30T,

ND5	13309	G	A	12815	0.31%	nonsynonymous	ND5	c.G973A	p.A325T,
ND5	13434	A	G	11970	39.77%	nonsynonymous	ND5	c.A1098G	p.I366M,
ND5	13701	A	G	5597	4.18%	synonymous	ND5	c.A1365G	p.K455K,
ND5	13708	G	A	5250	0.44%	nonsynonymous	ND5	c.G1372A	p.A458T,
CYTB	15543	C	T	23488	2.35%	nonsynonymous	CYTB	c.C797T	p.P266L,

Table 11.6 The variants detected in the d4T-exposed fibroblast sample detected using the MiSeq platform from experimental day 22.

Gene	Position	Ref	Var	Average depth	Var Freq	Mutation type	Nucleotide change	Protein change
D-Loop_A	53	G	T	5662	0.39%			
D-Loop_A	498	C	T	3084	0.42%			
D-Loop_A	515	A	G	3752	0.46%			
D-Loop_A	536	C	A	5240	0.31%			
D-Loop_A	543	C	A	5420	0.33%			
RNR1	709	G	A	14830	0.59%			
RNR1	750	A	G	15265	99.69%			
RNR1	880	C	A	18516	0.39%			
RNR1	904	C	A	17856	0.30%			
RNR1	1024	G	A	20290	4.96%			
RNR1	1045	G	T	20234	0.36%			
RNR1	1414	C	A	22103	0.31%			
RNR2	1709	G	A	20681	4.17%			
RNR2	1999	A	G	19709	2.21%			
RNR2	2043	C	A	19640	0.30%			
RNR2	2405	C	A	18486	0.36%			
RNR2	2425	A	C	19430	0.41%			
RNR2	2443	C	A	18357	0.32%			
RNR2	2553	G	T	19317	0.35%			
RNR2	2623	A	G	19145	1.68%			
RNR2	2629	A	G	20222	0.59%			
RNR2	2819	G	T	18007	0.31%			
RNR2	2844	G	A	18843	0.74%			

RNR2	2878	G	T	18557	0.31%			
RNR2	2909	G	T	18125	0.37%			
RNR2	3003	A	C	20726	2.04%			
ND1	3553	C	A	4362	0.41%	nonsynonymous	c.C247A	p.L83I,
ND1	3586	C	A	4189	0.48%	nonsynonymous	c.C280A	p.P94T,
ND1	3588	C	A	4205	0.33%	synonymous	c.C282A	p.P94P,
ND1	3594	C	A	4243	0.33%	synonymous	c.C288A	p.V96V,
ND1	4086	C	A	17376	0.32%	synonymous	c.C780A	p.V260V,
ND1	4111	C	T	17762	0.60%	synonymous	c.C805T	p.L269L,
ND1	4153	G	T	16816	0.34%	nonsynonymous	c.G847T	p.D283Y,
ND2	4814	C	A	14499	0.31%	synonymous	c.C345A	p.V115V,
ND2	5147	G	T	9678	0.33%	synonymous	c.G678T	p.T226T,
ND2	5435	C	A	4298	0.33%	synonymous	c.C966A	p.P322P,
ND2	5438	C	A	4402	0.39%	synonymous	c.C969A	p.T323T,
ND2	5439	C	A	4409	0.39%	nonsynonymous	c.C970A	p.P324T,
ND2	5444	C	A	4438	0.34%	nonsynonymous	c.C975A	p.F325L,
ND2	5449	C	A	4446	0.34%	nonsynonymous	c.C980A	p.P327H,
ND2	5450	C	A	4397	0.52%	synonymous	c.C981A	p.P327P,
ND2	5456	C	A	3904	0.36%	synonymous	c.C987A	p.L329L,
COX1	5992	G	A	15094	0.83%	nonsynonymous	c.G89A	p.G30D,
COX1	6054	G	T	16417	0.41%	nonsynonymous	c.G151T	p.D51Y,
COX1	6211	G	T	15977	0.31%	stoploss	c.G308T	p.X103L,
COX1	6280	G	T	16926	0.32%	stoploss	c.G377T	p.X126L,
COX1	6621	C	A	28905	0.39%	nonsynonymous	c.C718A	p.H240N,
COX1	6808	G	T	30373	0.33%	nonsynonymous	c.G905T	p.R302L,
COX1	6871	G	T	32131	0.30%	stoploss	c.G968T	p.X323L,
COX1	6962	G	T	31619	0.31%	synonymous	c.G1059T	p.L353L,

COX1	7055	A	G	28457	34.04%	synonymous	c.A1152G	p.G384G,
COX1	7418	C	A	15384	0.31%	nonsynonymous	c.C1515A	p.F505L,
COX2	7979	G	T	14970	0.31%	nonsynonymous	c.G394T	p.D132Y,
COX2	8138	A	G	14636	0.57%	nonsynonymous	c.A553G	p.T185A,
COX2	8168	C	A	13843	0.32%	nonsynonymous	c.C583A	p.Q195K,
COX2	8263	C	A	14604	0.32%	synonymous	c.C678A	p.T226T,
ATP6	8545	G	A	12525	8.70%	nonsynonymous	c.G19A	p.A7T,
ATP6	8852	G	T	17638	0.32%	stoploss	c.G326T	p.X109L,
COX3	9209	G	T	19226	0.34%	nonsynonymous	c.G3T	p.M1I,
COX3	9528	C	A	11485	0.34%	nonsynonymous	c.C322A	p.P108T,
COX3	9921	G	A	10799	43.42%	nonsynonymous	c.G715A	p.A239T,
ND3	10164	C	A	5560	0.32%	nonsynonymous	c.C106A	p.P36T,
ND3	10182	G	T	5107	0.31%	nonsynonymous	c.G124T	p.D42Y,
ND3	10194	C	A	5374	0.39%	nonsynonymous	c.C136A	p.P46T,
ND3	10327	C	A	6258	0.34%	stopgain	c.C269A	p.S90X,
ND3	10371	G	T	6434	0.33%	stopgain	c.G313T	p.E105X,
ND3	10375	G	T	6188	0.31%	stoploss	c.G317T	p.X106L,
ND4	10806	G	T	9966	0.36%	stoploss	c.G47T	p.X16L,
ND4	10917	C	A	8264	0.36%	nonsynonymous	c.C158A	p.S53Y,
ND4	10938	C	A	7510	0.39%	nonsynonymous	c.C179A	p.P60H,
ND4	10939	C	A	7778	0.35%	synonymous	c.C180A	p.P60P,
ND4	10949	C	A	7203	0.49%	nonsynonymous	c.C190A	p.P64T,
ND4	10950	C	A	7164	0.32%	nonsynonymous	c.C191A	p.P64H,
ND4	11018	G	T	9685	0.33%	stopgain	c.G259T	p.E87X,
ND4	11835	G	T	19299	0.33%	stoploss	c.G1076T	p.X359L,
ND4	12128	T	C	19272	15.78%	nonsynonymous	c.T1369C	p.F457L,
ND5	12418	A	C	17814	0.42%	nonsynonymous	c.A82C	p.K28Q,

ND5	12425	A	C	17189	0.48%	nonsynonymous	c.A89C	p.N30T,
ND5	12971	C	A	9211	0.31%	nonsynonymous	c.C635A	p.P212Q,
ND5	13028	C	A	7569	0.50%	nonsynonymous	c.C692A	p.P231H,
ND5	13035	C	A	7219	0.37%	synonymous	c.C699A	p.L233L,
ND5	13037	C	A	7236	0.32%	nonsynonymous	c.C701A	p.P234H,
ND5	13038	C	A	7228	0.33%	synonymous	c.C702A	p.P234P,
ND5	13197	C	A	12665	0.31%	nonsynonymous	c.C861A	p.F287L,
ND5	13434	A	G	12703	42.11%	nonsynonymous	c.A1098G	p.I366M,
ND5	13649	C	A	6968	0.55%	nonsynonymous	c.C1313A	p.P438H,
ND5	13683	C	A	6444	0.31%	synonymous	c.C1347A	p.T449T,
ND5	13701	A	G	6836	4.36%	synonymous	c.A1365G	p.K455K,
ND5	13707	G	T	5641	0.32%	synonymous	c.G1371T	p.L457L,
ND5	13708	G	A	6484	0.52%	nonsynonymous	c.G1372A	p.A458T,
ND5	14006	G	T	15630	0.35%	stoploss	c.G1670T	p.X557L,
CYTB	15230	C	A	20290	0.32%	nonsynonymous	c.C484A	p.Q162K,
CYTB	15466	G	T	23460	0.34%	nonsynonymous	c.G720T	p.M240I,
CYTB	15500	G	T	21625	0.32%	nonsynonymous	c.G754T	p.D252Y,
CYTB	15543	C	T	24499	2.09%	nonsynonymous	c.C797T	p.P266L,
D-Loop_B	16093	T	C	22236	99.66%			
D-Loop_B	16145	G	T	17535	0.35%			

Table 11.7 The variants detected in the AZT-exposed fibroblast sample detected using the MiSeq platform from experimental day 22.

Gene	Position	Ref	Var	Average depth	Var Freq	Mutation type	Nucleotide change	Protein change
COX1	5992	G	A	13126	0.86%	nonsynonymous	c.G89A	p.G30D,
COX2	8138	A	G	13815	0.48%	nonsynonymous	c.A553G	p.T185A,
ATP6,ATP8	8545	G	A	12929	9.48%	nonsynonymous	c.G19A	p.A7T,
COX3	9921	G	A	11597	44.17%	nonsynonymous	c.G715A	p.A239T,
ND4	12128	T	C	17835	17.11%	nonsynonymous	c.T1369C	p.F457L,
ND5	13434	A	G	12682	43.44%	nonsynonymous	c.A1098G	p.I366M,
ND5	13708	G	A	6057	0.33%	nonsynonymous	c.G1372A	p.A458T,
ND5	13787	T	C	5951	0.37%	nonsynonymous	c.T1451C	p.L484P,
CYTB	14766	C	T	14950	0.53%	nonsynonymous	c.C20T	p.T7I,
CYTB	14798	T	C	14417	0.57%	nonsynonymous	c.T52C	p.F18L,
CYTB	15293	T	G	19057	0.84%	nonsynonymous	c.T547G	p.F183V,
CYTB	15452	C	A	23028	0.57%	nonsynonymous	c.C706A	p.L236I,
CYTB	15543	C	T	22603	1.50%	nonsynonymous	c.C797T	p.P266L,
COX2	7948	C	A	16250	0.63%	stopgain	c.C363A	p.Y121X,
ND1	4111	C	T	17150	0.72%	synonymous	c.C805T	p.L269L,
COX1	6572	C	A	27548	0.35%	synonymous	c.C669A	p.A223A,
COX1	7028	C	T	27123	0.50%	synonymous	c.C1125T	p.A375A,
COX1	7055	A	G	25961	35.29%	synonymous	c.A1152G	p.G384G,
COX2	7792	C	G	14878	0.94%	synonymous	c.C207G	p.P69P,
COX3	9251	A	T	16925	0.52%	synonymous	c.A45T	p.P15P,
ND5	13701	A	G	6375	6.21%	synonymous	c.A1365G	p.K455K,
RNR1	709	G	A	12008	0.37%			
RNR1	1024	G	A	16714	7.37%			

RNR2	1709	G	A	17562	4.07%			
RNR2	1999	A	G	17232	1.53%			
RNR2	2182	G	A	16072	0.43%			
RNR2	2623	A	G	16235	1.62%			
RNR2	2629	A	G	17140	0.71%			
RNR2	2844	G	A	16105	1%			
RNR2	3003	A	C	17475	1.49%			
tRNA-Thr	15928	G	A	24306	0.31%			
D-Loop_B	16093	T	C	19342	99.43%			

Table 11.8 The variants detected in the TDF-exposed fibroblast sample detected using the MiSeq platform from experimental day 22.

Appendix D

Appendix D consists of the shared variant data in the wild-type molecules of the cybrids exposed to NRTIs at end of depletion and end of repopulation compared to the experimental day 0 values.

	Position	Average Shift				
		Untreated	ddI	D4T	AZT	TDF
	150	1.61%	2.00%	1.48%	2.47%	1.62%
	185	1.47%	1.83%	1.53%		1.50%
	4216	-0.08%	0.00%	0.06%	-0.02%	-0.02%
	15466	1.20%	1.36%	1.35%	1.44%	1.23%
	16069	1.24%	1.30%	1.46%	1.72%	1.30%
	16093	1.35%	1.36%	1.47%	1.63%	1.53%
	16343	0.76%	1.05%	0.99%	0.94%	0.64%
	16519	1.05%	1.37%	1.17%	1.30%	1.10%
Mean shift of all bp (n=8)		**1.00%**	**1.18%**	**1.15%**	**1.17%**	**1.04%**
SD of mean shift (n=8)		**0.53%**	**0.60%**	**0.49%**	**0.77%**	**0.55%**

Table 11.9 The mean shift from experimental day 0 for all of the shared variants across the NRTI conditions in the cybrids exposed to x1 concentration at experimental day 32.

	Pos.	Average Shift				
		Untreated	**ddI**	**d4T**	**AZT**	**TDF**
	150	3.60%	7.65%	1.88%	3.36%	2.97%
	185	0.04%	0.15%	0.11%	0.02%	0.07%
	15466	2.94%	7.46%	2.72%	2.95%	2.68%
	16069	3.30%	7.59%	2.98%	3.21%	2.84%
	16093	3.54%	7.93%	3.26%	3.40%	3.07%
	16343	2.67%	6.82%	2.19%	2.60%	2.09%
	16519	2.70%	6.71%	2.52%	2.69%	2.18%
Mean shift of all bp (n=7)		**2.68%**	**6.33%**	**2.24%**	**2.60%**	**2.27%**
SD of mean shift (n=7)		**1.23%**	**2.76%**	**1.05%**	**1.18%**	**1.04%**

Table 11.10 The mean shift from experimental day 0 for all of the shared variants across the NRTI conditions in the cybrids exposed to x1 concentration at experimental day 52.

	Position	Average Shift				
		Untreated	ddI	D4T	AZT	TDF
	150	1.73%		1.53%	0.00%	1.50%
	185			0.06%	0.00%	-0.02%
	15466	1.12%	1.69%	1.35%	1.44%	1.23%
	16069	1.15%	1.51%	1.46%	1.72%	1.30%
	16093	1.24%	1.67%	1.47%	1.63%	1.53%
	16343	0.95%	1.20%	0.99%	0.94%	0.64%
	16519	1.34%		1.17%	1.30%	1.10%
Mean shift of all bp (n=7)		**1.16%**	**1.52%**	**1.08%**	**1.17%**	**0.96%**
SD of mean shift (n=7)		**0.27%**	**0.23%**	**0.52%**	**0.73%**	**0.55%**

Table 11.11 The mean shift from experimental day 0 for all of the shared variants across the NRTI conditions in the cybrids exposed to x10 concentration at experimental day 32 .

	Pos.	Untreated	ddI	d4T	AZT	TDF
		Average Shift	Average Shift	Average Shift	Average Shift	Average Shift
	150	7.64%	13.28%	8.08%	6.92%	10.92%
	185	7.05%	12.44%	7.78%	6.70%	10.53%
	15466	2.19%	12.66%	3.93%	1.60%	3.91%
	16069	5.51%	11.25%	6.49%	5.39%	6.97%
	16093	6.13%	11.46%	6.80%	5.81%	7.88%
	16343	6.21%	12.04%	6.64%	5.35%	7.39%
	16519	6.37%	12.68%	6.80%	5.73%	7.11%
Mean shift of all bp (n=7)		**5.87%**	**12.26%**	**6.64%**	**5.36%**	**7.82%**
SD of mean shift (n=7)		**1.76%**	**0.72%**	**1.34%**	**1.77%**	**2.37%**

Table 11.12 The mean shift from experimental day 0 for all of the shared variants across the NRTI conditions in the cybrids exposed to x10 concentration at experimental day 52.

Appendix E

Appendix E consists of the shared variant data in the deleted molecules of the cybrids exposed to NRTIs at end of depletion and end of repopulation compared to the experimental day 0 values.

Mitochondrial Position	Average Shift in untreated to Day 32	Average Shift in untreated to Day 52
150	4.22%	5.27%
1389	-0.08%	-0.38%
1811	6.14%	7.54%
2221	-0.07%	0.06%
2707	4.49%	7.28%
4188	6.01%	3.42%
4248	0.05%	0.01%
4640	4.71%	1.34%
5893	0.14%	-0.32%
6209	-0.32%	-0.14%
15940	-5.18%	-4.60%
16069	4.40%	3.00%
16093	4.39%	2.72%
16126	4.41%	2.72%
16343	4.42%	3.49%
16519	4.17%	4.20%

Table 11.13 The average shift from experimental day 0 to experimental days 32 and 52 in the untreated deleted amplicon. (-) indicates a loss of heteroplasmy and no sign indicates a gain in heteroplasmy.

Pos	Average Shift in ddI to Day 32	Average Shift in ddI to Day 52
150	-4.67%	-5.15%
1389	0.47%	-0.24%
1811	-4.48%	-4.62%
2221	0.00%	-0.33%
2707	-7.59%	-6.48%
4188	-2.40%	-0.33%
4248	0.09%	0.09%
4640	-2.17%	-1.23%
5893	0.53%	-0.11%
6209	0.05%	-0.42%
15940	8.73%	12.02%
16069	-1.67%	-2.57%
16093	-1.13%	-2.09%
16126	-1.00%	-1.84%
16343	-1.49%	-2.29%
16519	-2.95%	-3.42%

Table 11.14 The average shift from experimental day 0 to experimental days 32 and 52 in the ddI x10 deleted amplicon. (-) indicates a loss of heteroplasmy and no sign indicates a gain in heteroplasmy.

Appendix F

Appendix F consists of the variant data of the three tissues sequenced on the MiSeq for the control and AZT-exposed aged mice.

Control Mouse							
Gene	**Ref**	**Var**	**Average depth**	**Var Freq**	**Mutation type**	**Nucleotide change**	**Protein change**
16S_rRNA	G	T	9161	0.31%			
16S_rRNA	G	T	8566	0.30%			
cytochrome_c_oxidase_I	G	T	3888	0.36%	stoploss	c.G968T	p.X323L,
cytochrome_c_oxidase_I	G	T	3371	0.39%	nonsynonymous	c.G1090T	p.D364Y,
cytochrome_c_oxidase_I	C	A	7226	0.33%	nonsynonymous	c.C1183A	p.H395N,
cytochrome_c_oxidase_I	G	T	7969	0.30%	nonsynonymous	c.G1219T	p.D407Y,
cytochrome_c_oxidase_II	G	T	9200	0.37%	stoploss	c.G194T	p.X65L,
tRNA_arginine	A	T	6392	0.41%			
tRNA_arginine	C	A	6624	0.33%			
NADH_dehydrogenase_6	A	C	7314	0.92%	nonsynonymous	c.T376G	p.Y126D,
NADH_dehydrogenase_6	G	C	6220	0.66%	nonsynonymous	c.C349G	p.L117V,
cytochrome_b	C	A	8265	0.34%	synonymous	c.C369A	p.V123V,
cytochrome_b	C	G	7488	0.32%	synonymous	c.C1020G	p.G340G,
D-Loop	C	A	4566	0.33%			
D-Loop	T	C	6155	0.31%			
AZT Mouse							
Gene	**Ref**	**Var**	**Average depth**	**Var Freq**	**Mutation type**	**Nucleotide change**	**Protein change**
12S_rRNA	G	T	7405	0.31%			
12S_rRNA	C	A	7590	0.33%			
12S_rRNA	C	A	7190	0.32%			
12S_rRNA	C	A	7640	0.30%			

NADH_dehydrogenase_2	C	A	18718	12.27%	nonsynonymous	c.C49A	p.P17T,
cytochrome_c_oxidase_I	C	A	16845	0.31%	nonsynonymous	c.C6A	p.F2L,
cytochrome_c_oxidase_I	G	T	15782	0.31%	nonsynonymous	c.G151T	p.D51Y,
cytochrome_c_oxidase_I	G	T	17040	0.31%	nonsynonymous	c.G634T	p.D212Y,
cytochrome_c_oxidase_I	G	T	15902	0.33%	stoploss	c.G824T	p.X275L,
cytochrome_c_oxidase_I	C	A	8029	0.35%	synonymous	c.C954A	p.V318V,
cytochrome_c_oxidase_I	G	T	4369	0.34%	nonsynonymous	c.G1090T	p.D364Y,
cytochrome_c_oxidase_I	G	T	10039	0.33%	nonsynonymous	c.G1219T	p.D407Y,
cytochrome_c_oxidase_I	G	T	9708	0.34%	stoploss	c.G1226T	p.X409L,
cytochrome_c_oxidase_II	C	A	11372	0.33%	synonymous	c.C426A	p.V142V,
cytochrome_c_oxidase_III	G	T	11162	0.46%	stoploss	c.G47T	p.X16L,
NADH_dehydrogenase_3	G	T	7217	0.32%	nonsynonymous	c.G124T	p.D42Y,
NADH_dehydrogenase_5	C	T	12190	0.46%	nonsynonymous	c.C793T	p.P265S,
NADH_dehydrogenase_6	A	C	10561	0.58%	nonsynonymous	c.T376G	p.Y126D,
NADH_dehydrogenase_6	G	C	9349	0.49%	nonsynonymous	c.C349G	p.L117V,
NADH_dehydrogenase_6	G	T	9353	0.31%	stopgain	c.C29A	p.S10X,
tRNA_glutamic_acid	G	T	10175	0.30%			
cytochrome_b	G	T	10653	0.37%	nonsynonymous	c.G58T	p.D20Y,
cytochrome_b	G	T	12928	0.32%	stoploss	c.G230T	p.X77L,
D-Loop	C	A	4557	0.31%			
D-Loop	C	A	4341	0.51%			
D-Loop	C	A	4540	0.40%			
D-Loop	A	C	3852	0.80%			
D-Loop	C	A	4315	0.42%			
D-Loop	C	A	4401	0.48%			

D-Loop	T	C	3312	4.86%			
D-Loop	T	C	3295	1.21%			
D-Loop	G	T	3134	0.45%			
D-Loop	C	A	3184	0.38%			
D-Loop	C	A	3396	0.44%			
D-Loop	C	A	3520	0.34%			
D-Loop	G	A	3103	0.81%			

Table 11.15 The variant data for gastrocnemius mouse tissue for the control and AZT-exposed mouse.

Control Mouse							
Gene	**Ref**	**Var**	**Average depth**	**Var Freq**	**Mutation type**	**Nucleotide change**	**Protein change**
12S_rRNA	G	T	5536	0.31%			
16S_rRNA	C	A	9187	0.33%			
16S_rRNA	C	A	12481	0.34%			
cytochrome_c_oxidase_I	C	A	14769	0.30%	nonsynonymous	c.C683A	p.P228Q,
cytochrome_c_oxidase_I	A	G	15618	0.42%	synonymous	c.A795G	p.K265K,
cytochrome_c_oxidase_I	T	A	6786	0.71%	synonymous	c.T972A	p.L324L,
cytochrome_c_oxidase_I	T	G	6223	0.31%	nonsynonymous	c.T993G	p.N331K,
cytochrome_c_oxidase_I	G	T	5932	0.39%	nonsynonymous	c.G1149T	p.M383I,
cytochrome_c_oxidase_I	G	T	7283	0.33%	stoploss	c.G1481T	p.X494L,
cytochrome_c_oxidase_I	C	A	7201	0.35%	nonsynonymous	c.C1496A	p.P499H,
tRNA_serine_1	C	A	6109	0.31%			
tRNA_serine_1	C	A	5953	0.32%			
cytochrome_c_oxidase_II	C	A	7743	0.35%	nonsynonymous	c.C232A	p.L78I,
cytochrome_c_oxidase_II	C	A	7938	0.31%	synonymous	c.C426A	p.V142V,
cytochrome_c_oxidase_II	C	A	7379	0.38%	nonsynonymous	c.C497A	p.P166H,
ATP_synthase_6	C	A	6782	0.37%	nonsynonymous	c.C223A	p.L75I,
ATP_synthase_6	G	T	6766	0.33%	stoploss	c.G326T	p.X109L,
ATP_synthase_6	G	T	6810	0.37%	nonsynonymous	c.G462T	p.M154I,
cytochrome_c_oxidase_III	G	T	7506	0.36%	nonsynonymous	c.G3T	p.M1I,
cytochrome_c_oxidase_III	G	T	8296	0.31%	nonsynonymous	c.G662T	p.R221L,
NADH_dehydrogenase_3	G	T	4830	0.37%	stoploss	c.G230T	p.X77L,
tRNA_arginine	A	T	4826	0.42%			
tRNA_arginine	T	A	4720	0.38%			
NADH_dehydrogenase_4L	G	T	5660	0.34%	nonsynonymous	c.G52T	p.G18W,

NADH_dehydrogenase_4	C	A	5466	0.33%	nonsynonymous	c.C460A	p.L154I,
NADH_dehydrogenase_4	G	T	6768	0.33%	stoploss	c.G1076T	p.X359L,
NADH_dehydrogenase_5	G	T	5526	0.31%	stoploss	c.G293T	p.X98L,
NADH_dehydrogenase_5	T	C	7922	0.63%	nonsynonymous	c.T772C	p.F258L,
NADH_dehydrogenase_5	C	T	8864	0.51%	nonsynonymous	c.C793T	p.P265S,
NADH_dehydrogenase_5	C	A	8976	0.30%	synonymous	c.C795A	p.P265P,
NADH_dehydrogenase_5	T	A	15534	0.40%	nonsynonymous	c.T1430A	p.I477N,
NADH_dehydrogenase_5	G	T	12027	0.32%	nonsynonymous	c.G1579T	p.G527W,
NADH_dehydrogenase_6	A	C	7211	1.66%	nonsynonymous	c.T376G	p.Y126D,
NADH_dehydrogenase_6	G	C	6259	1.66%	nonsynonymous	c.C349G	p.L117V,
NADH_dehydrogenase_6	C	A	6541	0.37%	nonsynonymous	c.G283T	p.G95W,
NADH_dehydrogenase_6	C	A	6722	0.31%	nonsynonymous	c.G181T	p.G61W,
NADH_dehydrogenase_6	C	A	6483	0.35%	nonsynonymous	c.G106T	p.G36W,
cytochrome_b	G	T	7565	0.36%	nonsynonymous	c.G3T	p.M1I,
cytochrome_b	G	T	8930	0.36%	nonsynonymous	c.G496T	p.G166W,
D-Loop	T	C	4841	0.35%			
D-Loop	G	T	7421	0.42%			
AZT-exposed Mouse							
Gene	**Ref**	**Var**	**Average depth**	**Var Freq**	**Mutation type**	**Nucleotide change**	**Protein change**
12S_rRNA	C	A	5620	0.34%			
12S_rRNA	G	T	5859	0.32%			
12S_rRNA	C	A	5986	0.30%			
12S_rRNA	C	A	8698	0.31%			
12S_rRNA	G	T	9939	0.43%			
tRNA_valine	G	T	10964	0.31%			
16S_rRNA	C	A	11908	0.31%			

16S_rRNA	G	T	14317	0.30%			
16S_rRNA	C	A	13139	0.32%			
16S_rRNA	C	A	13069	0.31%			
16S_rRNA	G	T	13683	0.31%			
16S_rRNA	C	A	10183	0.30%			
NADH_dehydrogenase_1	C	A	15625	0.37%	nonsynonymous	c.C124A	p.P42T,
NADH_dehydrogenase_1	C	A	15450	0.30%	synonymous	c.C264A	p.P88P,
NADH_dehydrogenase_1	C	A	18103	0.33%	nonsynonymous	c.C850A	p.Q284K,
NADH_dehydrogenase_2	C	A	20257	16.32%	nonsynonymous	c.C49A	p.P17T,
cytochrome_c_oxidase_I	G	T	15925	0.35%	nonsynonymous	c.G113T	p.R38L,
cytochrome_c_oxidase_I	C	A	17247	0.36%	stopgain	c.C467A	p.S156X,
cytochrome_c_oxidase_I	C	A	16323	0.33%	nonsynonymous	c.C683A	p.P228Q,
cytochrome_c_oxidase_I	C	A	8499	0.31%	nonsynonymous	c.C943A	p.P315T,
cytochrome_c_oxidase_I	G	T	4372	0.32%	nonsynonymous	c.G1090T	p.D364Y,
cytochrome_c_oxidase_I	G	T	10370	0.52%	nonsynonymous	c.G1219T	p.D407Y,
cytochrome_c_oxidase_I	G	T	9987	0.33%	stoploss	c.G1226T	p.X409L,
cytochrome_c_oxidase_I	C	A	11174	0.36%	nonsynonymous	c.C1396A	p.L466I,
cytochrome_c_oxidase_II	C	A	11370	0.36%	synonymous	c.C426A	p.V142V,
ATP_synthase_6	G	T	9918	0.31%	nonsynonymous	c.G395T	p.G132V,
cytochrome_c_oxidase_III	G	T	11995	0.43%	nonsynonymous	c.G3T	p.M1I,
cytochrome_c_oxidase_III	G	T	11941	0.32%	stoploss	c.G47T	p.X16L,
cytochrome_c_oxidase_III	G	T	12100	0.33%	stoploss	c.G296T	p.X99L,
cytochrome_c_oxidase_III	G	T	10751	0.37%	stopgain	c.G706T	p.E236X,
NADH_dehydrogenase_3	G	T	7550	0.40%	stoploss	c.G338T	p.X113L,
tRNA_arginine	A	T	7849	0.36%			
tRNA_arginine	G	T	7588	0.41%			
NADH_dehydrogenase_4	G	T	8182	0.31%	stopgain	c.G340T	p.E114X,

NADH_dehydrogenase_5	C	A	9099	0.34%	synonymous	c.C288A	p.V96V,
NADH_dehydrogenase_5	C	A	11916	0.30%	nonsynonymous	c.C692A	p.P231Q,
NADH_dehydrogenase_6	A	C	9220	0.85%	nonsynonymous	c.T376G	p.Y126D,
NADH_dehydrogenase_6	G	C	7898	0.67%	nonsynonymous	c.C349G	p.L117V,
NADH_dehydrogenase_6	C	A	8273	0.40%	synonymous	c.G210T	p.T70T,
tRNA_glutamic_acid	G	T	8889	0.36%			
cytochrome_b	G	T	9077	0.37%	nonsynonymous	c.G58T	p.D20Y,
cytochrome_b	C	A	10863	0.36%	nonsynonymous	c.C363A	p.F121L,
D-Loop	C	A	6039	0.35%			
D-Loop	G	T	8866	0.33%			
D-Loop	C	A	8282	0.34%			
D-Loop	C	A	4799	0.31%			
D-Loop	C	A	4098	0.32%			
D-Loop	C	A	3644	0.36%			
D-Loop	C	A	3627	0.33%			
D-Loop	C	A	3386	0.47%			
D-Loop	C	A	3523	0.62%			
D-Loop	C	A	3601	0.36%			

Table 11.16 The variant data for heart mouse tissue for the control and AZT exposed mouse.

Control Mouse							
Gene	**Ref**	**Mut**	**average depth**	**Var Freq**	**Mutation type**	**Nucleotide change**	**Protein change**
12S_rRNA	C	A	4176	0.31%			
cytochrome_c_oxidase_I	T	A	4225	0.31%	synonymous	c.T972A	p.L324L,
tRNA_arginine	A	T	5243	0.36%			
NADH_dehydrogenase_5	G	T	5203	0.33%	nonsynonymous	c.G247T	p.D83Y,
NADH_dehydrogenase_6	A	C	7954	0.78%	nonsynonymous	c.T376G	p.Y126D,
NADH_dehydrogenase_6	G	C	7098	0.61%	nonsynonymous	c.C349G	p.L117V,
cytochrome_b	C	A	6662	0.30%	synonymous	c.C375A	p.A125A,
cytochrome_b	C	G	7219	0.32%	synonymous	c.C1020G	p.G340G,
AZT-exposed Mouse							
Gene	**Ref**	**Mut**	**average depth**	**Var Freq**	**Mutation type**	**Nucleotide change**	**Protein change**
NADH_dehydrogenase_2	C	A	23133	20.88%	nonsynonymous	c.C49A	p.P17T,
cytochrome_c_oxidase_I	T	A	9165	0.41%	synonymous	c.T972A	p.L324L,
cytochrome_c_oxidase_I	G	A	13053	0.90%	nonsynonymous	c.G1261A	p.V421I,
tRNA_arginine	A	T	8820	0.39%			
tRNA_arginine	T	A	8636	0.34%			
NADH_dehydrogenase_5	C	T	17234	0.36%	nonsynonymous	c.C793T	p.P265S,
NADH_dehydrogenase_6	A	C	10907	0.99%	nonsynonymous	c.T376G	p.Y126D,
NADH_dehydrogenase_6	G	C	9388	0.96%	nonsynonymous	c.C349G	p.L117V,
D-Loop	A	C	4141	1.66%			
D-Loop	C	A	4754	0.44%			
D-Loop	T	C	3513	4.13%			

D-Loop	T	C	3535	0.93%			
D-Loop	C	A	3486	0.34%			
D-Loop	C	A	3105	0.32%			

Table 11.17 The variant data for liver mouse tissue for the control and AZT exposed mouse.

Chapter 12. References

Affandi, J. S., Price, P., Imran, D., Yunihastuti, E., Djauzi, S. and Cherry, C. L. (2008) 'Can we predict neuropathy risk before stavudine prescription in a resource-limited setting?', *AIDS Res Hum Retroviruses*, 24(10), pp. 1281-4.

Ameur, A., Stewart, J. B., Freyer, C., Hagström, E., Ingman, M., Larsson, N.-G. and Gyllensten, U. (2011) 'Ultra-Deep Sequencing of Mouse Mitochondrial DNA: Mutational Patterns and Their Origins', *PLoS Genet*, 7(3), p. e1002028.

Andrews, R. M., Kubacka, I., Chinnery, P. F., Lightowlers, R. N., Turnbull, D. M. and Howell, N. (1999) 'Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA', *Nat Genet*, 23(2), p. 147.

Anthony, I. C., Ramage, S. N., Carnie, F. W., Simmonds, P. and Bell, J. E. (2006) 'Accelerated Tau deposition in the brains of individuals infected with human immunodeficiency virus-1 before and after the advent of highly active anti-retroviral therapy', *Acta Neuropathol*, 111(6), pp. 529-38.

Anvret, A., Westerlund, M., Sydow, O., Willows, T., Lind, C., Galter, D. and Belin, A. C. (2010) 'Variations of the CAG trinucleotide repeat in DNA polymerase gamma (POLG1) is associated with Parkinson's disease in Sweden', *Neurosci Lett*, 485(2), pp. 117-20.

Archer, J., Weber, J., Henry, K., Winner, D., Gibson, R., Lee, L., Paxinos, E., Arts, E. J., Robertson, D. L., Mimms, L. and Quiñones-Mateu, M. E. (2012) 'Use of Four Next-Generation Sequencing Platforms to Determine HIV-1 Coreceptor Tropism', *PLOS ONE*, 7(11), p. e49602.

Arnheim, N. and Cortopassi, G. (1992a) 'Deleterious mitochondrial DNA mutations accumulate in aging human tissues', *Mutat Res*, 275(3-6), pp. 157-67.

Arnheim, N. and Cortopassi, G. (1992b) 'Deleterious mitochondrial DNA mutations accumulate in aging human tissues', *Mutation Research/DNAging*, 275(3–6), pp. 157-167.

Arnoult, D., Grodet, A., Lee, Y.-J., Estaquier, J. and Blackstone, C. (2005) 'Release of OPA1 during Apoptosis Participates in the Rapid and Complete Release of Cytochrome c and Subsequent Mitochondrial Fragmentation', *Journal of Biological Chemistry*, 280(42), pp. 35742-35750.

Ashrafi, G. and Schwarz, T. L. (2013) 'The pathways of mitophagy for quality control and clearance of mitochondria', *Cell Death Differ*, 20(1), pp. 31-42.

Attardi, G. and Schatz, G. (1988) 'Biogenesis of mitochondria', *Annu Rev Cell Biol*, 4, pp. 289-333.

Autran, B., Descours, B., Avettand-Fenoel, V. and Rouzioux, C. (2011) 'Elite controllers as a model of functional cure', *Current Opinion in HIV and AIDS*, 6(3), pp. 181-187 10.1097/COH.0b013e328345a328.

Avula, S., Parikh, S., Demarest, S., Kurz, J. and Gropman, A. (2014) 'Treatment of mitochondrial disorders', *Curr Treat Options Neurol*, 16(6), p. 292.

Baar, K., Wende, A. R., Jones, T. E., Marison, M., Nolte, L. A., Chen, M., Kelly, D. P. and Holloszy, J. O. (2002) 'Adaptations of skeletal muscle to exercise: rapid increase in the transcriptional coactivator PGC-1', *FASEB J*, 16(14), pp. 1879-86.

Bach, D., Pich, S., Soriano, F. X., Vega, N., Baumgartner, B., Oriola, J., Daugaard, J. R., Lloberas, J., Camps, M., Zierath, J. R., Rabasa-Lhoret, R., Wallberg-Henriksson, H., Laville, M., Palacín, M., Vidal, H., Rivera, F., Brand, M. and Zorzano, A. (2003) 'Mitofusin-2 Determines Mitochondrial Network Architecture and Mitochondrial Metabolism: A NOVEL REGULATORY MECHANISM ALTERED IN OBESITY', *Journal of Biological Chemistry*, 278(19), pp. 17190-17197.

Baram, D. and Cooke, J. (1993) *Lactic acidosis and AIDS*. Ann Intern Med. 1993 Aug 15;119(4):345.

Barre-Sinoussi, F., Ross, A. L. and Delfraissy, J. F. (2013) 'Past, present and future: 30 years of HIV research', *Nat Rev Microbiol*, 11(12), pp. 877-83.

Bartley, P. B., Westacott, L., Boots, R. J., Lawson, M., Potter, J. M., Hyland, V. J. and Woods, M. L. I. (2001) 'Large hepatic mitochondrial DNA deletions associated with l-lactic acidosis and highly active antiretroviral therapy', *AIDS*, 15(3), pp. 419-420.

Bastin, J., Aubey, F., Rotig, A., Munnich, A. and Djouadi, F. (2008) 'Activation of peroxisome proliferator-activated receptor pathway stimulates the mitochondrial respiratory chain and can correct deficiencies in patients' cells lacking its components', *J Clin Endocrinol Metab*, 93(4), pp. 1433-41.

Behar, D. M., van Oven, M., Rosset, S., Metspalu, M., Loogvali, E. L., Silva, N. M., Kivisild, T., Torroni, A. and Villems, R. (2012) 'A "Copernican" reassessment of the human mitochondrial DNA tree from its root', *Am J Hum Genet*, 90(4), pp. 675-84.

Bender, A., Krishnan, K. J., Morris, C. M., Taylor, G. A., Reeve, A. K., Perry, R. H., Jaros, E., Hersheson, J. S., Betts, J., Klopstock, T., Taylor, R. W. and Turnbull, D. M. (2006) 'High levels of mitochondrial DNA deletions in substantia nigra neurons in aging and Parkinson disease', *Nat Genet*, 38(5), pp. 515-7.

Birkus, G., Hitchcock, M. J. and Cihlar, T. (2002) 'Assessment of mitochondrial toxicity in human cells treated with tenofovir: comparison with other nucleoside reverse transcriptase inhibitors', *Antimicrob Agents Chemother*, 46(3), pp. 716-23.

Bogenhagen, D. and Clayton, D. A. (1977) 'Mouse L cell mitochondrial DNA molecules are selected randomly for replication throughout the cell cycle', *Cell*, 11(4), pp. 719-27.

Bollmann, F. M. (2013) 'Telomerase inhibition may contribute to accelerated mitochondrial aging induced by anti-retroviral HIV treatment', *Medical Hypotheses*, 81(2), pp. 285-287.

Bonilla, E., Tanji, K., Hirano, M., Vu, T. H., DiMauro, S. and Schon, E. A. (1999) 'Mitochondrial involvement in Alzheimer's disease', *Biochim Biophys Acta*, 1410(2), pp. 171-82.

Bourdon, A., Minai, L., Serre, V., Jais, J. P., Sarzi, E., Aubert, S., Chretien, D., de Lonlay, P., Paquis-Flucklinger, V., Arakawa, H., Nakamura, Y., Munnich, A. and Rotig, A. (2007) 'Mutation of RRM2B, encoding p53-controlled ribonucleotide reductase (p53R2), causes severe mitochondrial DNA depletion', *Nat Genet*, 39(6), pp. 776-80.

Brierley, E. J., Johnson, M. A., Lightowlers, R. N., James, O. F. and Turnbull, D. M. (1998) 'Role of mitochondrial DNA mutations in human aging: implications for the central nervous system and muscle', *Ann Neurol*, 43(2), pp. 217-23.

Brodin, J., Mild, M., Hedskog, C., Sherwood, E., Leitner, T., Andersson, B. and Albert, J. (2013) 'PCR-induced transitions are the major source of error in cleaned ultra-deep pyrosequencing data', *PLoS One*, 8(7), p. e70388.

Bua, E., Johnson, J., Herbst, A., Delong, B., McKenzie, D., Salamat, S. and Aiken, J. M. (2006) 'Mitochondrial DNA-deletion mutations accumulate intracellularly to detrimental levels in aged human skeletal muscle fibers', *Am J Hum Genet*, 79(3), pp. 469-80.

Bulst, S., Holinski-Feder, E., Payne, B., Abicht, A., Krause, S., Lochmuller, H., Chinnery, P. F., Walter, M. C. and Horvath, R. (2012) 'In vitro supplementation with deoxynucleoside monophosphates rescues mitochondrial DNA depletion', *Mol Genet Metab*, 107(1-2), pp. 95-103.

C, P., J, B., K, R. G., C, B., E, P. and S, S. (2014) 'Reflections on living with HIV over time: exploring the perspective of HIV-infected women over 50.', *Aging and Mental Health*.

C, R. (1995) 'Oxidative damage to mitochondrial DNA and its relationship to

ageing.', *Int J Biochem Cell Biol*, 27, pp. 647-653.

Calvo, S. E., Compton, A. G., Hershman, S. G., Lim, S. C., Lieber, D. S., Tucker, E. J., Laskowski, A., Garone, C., Liu, S., Jaffe, D. B., Christodoulou, J., Fletcher, J. M., Bruno, D. L., Goldblatt, J., Dimauro, S., Thorburn, D. R. and Mootha, V. K. (2012) 'Molecular diagnosis of infantile mitochondrial disease with targeted next-generation sequencing', *Sci Transl Med*, 4(118), p. 118ra10.

Campbell, G. R., Pallack, Z. T. and Spector, S. A. (2013) 'Vitamin D attenuates nucleoside reverse transcriptase inhibitor induced human skeletal muscle mitochondria DNA depletion', *AIDS*, 27(9), pp. 1397-401.
Canto, C. and Auwerx, J. (2009) 'PGC-1alpha, SIRT1 and AMPK, an energy sensing network that controls energy expenditure', *Curr Opin Lipidol*, 20(2), pp. 98-105.
Cao, L., Shitara, H., Sugimoto, M., Hayashi, J., Abe, K. and 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 Genet*, 5(12), p. e1000756.
Cao, Z., Wanagat, J., McKiernan, S. H. and Aiken, J. M. (2001) 'Mitochondrial DNA deletion mutations are concomitant with ragged red regions of individual, aged muscle fibers: analysis by laser-capture microdissection', *Nucleic Acids Res*, 29(21), pp. 4502-8.
Carter, M. M., Torres, S. M., Cook, D. L., McCash, C. L., Yu, M., Walker, V. E. and Walker, D. M. (2007) 'Relative mutagenic potencies of several nucleoside analogs, alone or in drug pairs, at the HPRT and TK loci of human TK6 lymphoblastoid cells', *Environmental and Molecular Mutagenesis*, 48(3-4), pp. 239-247.
Casademont, J., Barrientos, A., Grau, J. M., Pedrol, E., Estivill, X., Urbano-Márquez, A. and Nunes, V. (1996) 'The effect of zidovudine on skeletal muscle mtDNA in HIV-1 infected patients with mild or no muscle dysfunction', *Brain*, 119(4), pp. 1357-1364.
Chakrabarty, S., D'Souza, R. R., Kabekkodu, S. P., Gopinath, P. M., Rossignol, R. and Satyamoorthy, K. (2014) 'Upregulation of TFAM and mitochondria copy number in human lymphoblastoid cells', *Mitochondrion*, 15, pp. 52-8.
Chen, C. H. and Cheng, Y. C. (1989) 'Delayed cytotoxicity and selective loss of mitochondrial DNA in cells treated with the anti-human immunodeficiency virus compound 2',3'-dideoxycytidine', *J Biol Chem*, 264(20), pp. 11934-7.
Chen, H., Detmer, S. A., Ewald, A. J., Griffin, E. E., Fraser, S. E. and Chan, D. C. (2003) 'Mitofusins Mfn1 and Mfn2 coordinately regulate mitochondrial fusion and are essential for embryonic development', *The Journal of Cell Biology*, 160(2), pp. 189-200.
Cherry, C. L., Affandi, J. S., Imran, D., Yunihastuti, E., Smyth, K., Vanar, S., Kamarulzaman, A. and Price, P. (2009) 'Age and height predict neuropathy risk in patients with HIV prescribed stavudine', *Neurology*, 73(4), pp. 315-20.
Cherry, C. L., Mobarok, M., Wesselingh, S. L., Fain, R., Weinstock, S., Tachedjian, G., Srivastava, S., Tyssen, D. P., Glass, J. D. and Hooker, D. J. (2010) 'Ubisol-Aqua: coenzyme Q10 prevents antiretroviral toxic neuropathy in an in vitro model', *Curr HIV Res*, 8(3), pp. 232-9.
Cherry, C. L., Nolan, D., James, I. R., McKinnon, E. J., Mallal, S. A., Gahan, M. E., Lal, L., McArthur, J. C. and Wesselingh, S. L. (2006) 'Tissue-specific associations between mitochondrial DNA levels and current treatment status in HIV-infected individuals', *J Acquir Immune Defic Syndr*, 42(4), pp. 435-40.
Cherry, C. L., Rosenow, A., Affandi, J. S., McArthur, J. C., Wesselingh, S. L. and Price, P. (2008) 'Cytokine Genotype Suggests a Role for Inflammation in Nucleoside Analog-Associated Sensory Neuropathy (NRTI-SN) and Predicts an Individual's NRTI-SN Risk', *AIDS Res Hum Retroviruses*, 24(2), pp. 117-23.
Chinnery, P. F. and Samuels, D. C. (1999) 'Relaxed replication of mtDNA: A model with implications for the expression of disease', *Am J Hum Genet*, 64(4), pp. 1158-65.
Christian, B. E. and Spremulli, L. L. (2009) 'Evidence for an active role of IF3mt in the initiation of translation in mammalian mitochondria', *Biochemistry*, 48(15), pp. 3269-78.
Clay Montier, L. L., Deng, J. J. and Bai, Y. (2009) 'Number matters: control of mammalian mitochondrial DNA copy number', *J Genet Genomics*, 36(3), pp. 125-31.
Clayton, D. A. (1982) 'Replication of animal mitochondrial DNA', *Cell*, 28(4), pp. 693-705.
Clayton, D. A. (2003) 'Mitochondrial DNA replication: what we know', *IUBMB Life*, 55(4-5), pp. 213-7.
Coller, H. A., Bodyak, N. D. and Khrapko, K. (2002) 'Frequent intracellular clonal expansions of somatic mtDNA mutations: significance and mechanisms', *Ann N Y Acad Sci*, 959, pp. 434-47.

Cooke, M. S., Evans, M. D., Dizdaroglu, M. L. and Lunec, J. (2003) 'Oxidative DNA damage: mechanisms, mutation, and disease', *The FASEB Journal*, 17(10), pp. 1195-1214.
Corral-Debrinski, M., Horton, T., Lott, M. T., Shoffner, J. M., Beal, M. F. and Wallace, D. C. (1992) 'Mitochondrial DNA deletions in human brain: regional variability and increase with advanced age', *Nat Genet*, 2(4), pp. 324-9.
Corral-Debrinski, M., Horton, T., Lott, M. T., Shoffner, J. M., McKee, A. C., Beal, M. F., Graham, B. H. and Wallace, D. C. (1994) 'Marked changes in mitochondrial DNA deletion levels in Alzheimer brains', *Genomics*, 23(2), pp. 471-6.
Coskun, P. E., Beal, M. F. and Wallace, D. C. (2004) 'Alzheimer's brains harbor somatic mtDNA control-region mutations that suppress mitochondrial transcription and replication', *Proc Natl Acad Sci U S A*, 101(29), pp. 10726-31.
Cox, R., Platt, J., Chen, L. C., Tang, S., Wong, L. J. and Enns, G. M. (2012) 'Leigh syndrome caused by a novel m.4296G>A mutation in mitochondrial tRNA isoleucine', *Mitochondrion*, 12(2), pp. 258-61.
Craven, L., Tuppen, H. A., Greggains, G. D., Harbottle, S. J., Murphy, J. L., Cree, L. M., Murdoch, A. P., Chinnery, P. F., Taylor, R. W., Lightowlers, R. N., Herbert, M. and Turnbull, D. M. (2010) 'Pronuclear transfer in human embryos to prevent transmission of mitochondrial DNA disease', *Nature*, 465(7294), pp. 82-5.
Cree, L. M., Samuels, D. C., de Sousa Lopes, S. C., Rajasimha, H. K., Wonnapinij, P., Mann, J. R., Dahl, H. H. and Chinnery, P. F. (2008) 'A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes', *Nat Genet*, 40(2), pp. 249-54.
D'Antona, G., Nabavi, S. M., Micheletti, P., Di Lorenzo, A., Aquilani, R., Nisoli, E., Rondanelli, M. and Daglia, M. (2014) 'Creatine, L-Carnitine, and 3 Polyunsaturated Fatty Acid Supplementation from Healthy to Diseased Skeletal Muscle', *Biomed Res Int*, 2014, p. 613890.
Dalakas, M. C. (2001) 'Peripheral neuropathy and antiretroviral drugs', *Journal of the Peripheral Nervous System*, 6(1), pp. 14-20.
Dalakas, M. C., Semino-Mora, C. and Leon-Monzon, M. (2001) 'Mitochondrial alterations with mitochondrial DNA depletion in the nerves of AIDS patients with peripheral neuropathy induced by 2'3'-dideoxycytidine (ddC)', *Lab Invest*, 81(11), pp. 1537-44.
Daly, G. M., Bexfield, N., Heaney, J., Stubbs, S., Mayer, A. P., Palser, A., Kellam, P., Drou, N., Caccamo, M., Tiley, L., Alexander, G. J., Bernal, W. and Heeney, J. L. (2011) 'A viral discovery methodology for clinical biopsy samples utilising massively parallel next generation sequencing', *PLoS One*, 6(12), p. e28879.
Dames, S., Chou, L.-S., Xiao, Y., Wayman, T., Stocks, J., Singleton, M., Eilbeck, K. and Mao, R. (2013) 'The Development of Next-Generation Sequencing Assays for the Mitochondrial Genome and 108 Nuclear Genes Associated with Mitochondrial Disorders', *The Journal of Molecular Diagnostics*, 15(4), pp. 526-534.
Darbyshire, J. H. (1996) 'Delta: a randomised double-blind controlled trial comparing combinations of zidovudine plus didanosine or zalcitabine with zidovudine alone in HIV-infected individuals', *The Lancet*, 348(9023), pp. 283-291.
Davinelli, S., Sapere, N., Visentin, M., Zella, D. and Scapagnini, G. (2013) 'Enhancement of mitochondrial biogenesis with polyphenols: combined effects of resveratrol and equol in human endothelial cells', *Immun Ageing*, 10(1), p. 28.
Davis, A. F. and Clayton, D. A. (1996) 'In situ localization of mitochondrial DNA replication in intact mammalian cells', *J Cell Biol*, 135(4), pp. 883-93.
De Alwis, N., Hudson, G., Burt, A. D., Day, C. P. and Chinnery, P. F. (2009) 'Human liver stem cells originate from the canals of hering', *Hepatology*, 50(3), pp. 992-993.
de Grey, A. D. (2005) 'Reactive oxygen species production in the mitochondrial matrix: implications for the mechanism of mitochondrial mutation accumulation', *Rejuvenation Res*, 8(1), pp. 13-7.
De la Mata, M., Garrido-Maraver, J., Cotan, D., Cordero, M. D., Oropesa-Avila, M., Izquierdo, L. G., De Miguel, M., Lorite, J. B., Infante, E. R., Ybot, P., Jackson, S. and Sanchez-Alcazar, J. A.

(2012) 'Recovery of MERRF fibroblasts and cybrids pathophysiology by coenzyme Q10', *Neurotherapeutics*, 9(2), pp. 446-63.
de la Monte, S. M., Luong, T., Neely, T. R., Robinson, D. and Wands, J. R. (2000) 'Mitochondrial DNA damage as a mechanism of cell loss in Alzheimer's disease', *Lab Invest*, 80(8), pp. 1323-35.
Deeks, S. G. and Phillips, A. (2009) 'HIV infection, antiretroviral treatment, ageing, and non-AIDS related morbidity.', *BMJ*, 338(a3172).
Del Bo, R., Crimi, M., Sciacco, M., Malferrari, G., Bordoni, A., Napoli, L., Prelle, A., Biunno, I., Moggio, M., Bresolin, N., Scarlato, G. and Pietro Comi, G. (2003) 'High mutational burden in the mtDNA control region from aged muscles: a single-fiber study', *Neurobiol Aging*, 24(6), pp. 829-38.
Deveaud, C., Beauvoit, B., Hagry, S., Galinier, A., Carriere, A., Salin, B., Schaeffer, J., Caspar-Bauguil, S., Fernandez, Y., Gordien, J. B., Breilh, D., Penicaud, L., Casteilla, L. and Rigoulet, M. (2005) 'Site specific alterations of adipose tissue mitochondria in 3'-azido-3'-deoxythymidine (AZT)-treated rats: an early stage in lipodystrophy?', *Biochem Pharmacol*, 70(1), pp. 90-101.
Diaz, F., Bayona-Bafaluy, M. P., Rana, M., Mora, M., Hao, H. and Moraes, C. T. (2002) 'Human mitochondrial DNA with large deletions repopulates organelles faster than full-length genomes under relaxed copy number control', *Nucleic Acids Res*, 30(21), pp. 4626-33.
Dimauro, S. and Davidzon, G. (2005) 'Mitochondrial DNA and disease', *Ann Med*, 37(3), pp. 222-32.
Du, L. L., Chai, D. M., Zhao, L. N., Li, X. H., Zhang, F. C., Zhang, H. B., Liu, L. B., Wu, K., Liu, R., Wang, J. Z. and Zhou, X. W. (2014) 'AMPK Activation Ameliorates Alzheimer's Disease-Like Pathology and Spatial Memory Impairment in a Streptozotocin-Induced Alzheimer's Disease Model in Rats', *J Alzheimers Dis*.
Dufour, C. R., Wilson, B. J., Huss, J. M., Kelly, D. P., Alaynick, W. A., Downes, M., Evans, R. M., Blanchette, M. and Giguere, V. (2007) 'Genome-wide orchestration of cardiac functions by the orphan nuclear receptors ERRalpha and gamma', *Cell Metab*, 5(5), pp. 345-56.
Durham, S. E., Bonilla, E., Samuels, D. C., DiMauro, S. and Chinnery, P. F. (2005) 'Mitochondrial DNA copy number threshold in mtDNA depletion myopathy', *Neurology*, 65(3), pp. 453-5.
Edgar, D., Shabalina, Y. Camara, A. Wredenberg, M.A. Calvaruso, L. Nijtmans, J. Nedergaard, B. Cannon, N.G. Larsson and Trifunovic, A. (2009) 'Random point mutations with major effects on protein-coding genes are the driving force behind premature aging in mtDNA mutator mice. ', *Cell Metab*, 10, pp. 131-138.
Ekstrand, M. I., Falkenberg, M., Rantanen, A., Park, C. B., Gaspari, M., Hultenby, K., Rustin, P., Gustafsson, C. M. and Larsson, N. G. (2004) 'Mitochondrial transcription factor A regulates mtDNA copy number in mammals', *Hum Mol Genet*, 13(9), pp. 935-44.
Elo, J. M., Yadavalli, S. S., Euro, L., Isohanni, P., Gotz, A., Carroll, C. J., Valanne, L., Alkuraya, F. S., Uusimaa, J., Paetau, A., Caruso, E. M., Pihko, H., Ibba, M., Tyynismaa, H. and Suomalainen, A. (2012) 'Mitochondrial phenylalanyl-tRNA synthetase mutations underlie fatal infantile Alpers encephalopathy', *Hum Mol Genet*, 21(20), pp. 4521-9.
Elson, J. L., Samuels, D. C., Turnbull, D. M. and Chinnery, P. F. (2001) 'Random intracellular drift explains the clonal expansion of mitochondrial DNA mutations with age', *Am J Hum Genet*, 68(3), pp. 802-6.
Ewing, B. and Green, P. (1998) 'Base-calling of automated sequencer traces using phred. II. Error probabilities', *Genome Res*, 8(3), pp. 186-94.
Ewing, B., Hillier, L., Wendl, M. C. and Green, P. (1998) 'Base-calling of automated sequencer traces using phred. I. Accuracy assessment', *Genome Res*, 8(3), pp. 175-85.
Fisher, M. and Cooper, V. (2012) 'HIV and ageing: premature ageing or premature conclusions?', *Current Opinion in Infectious Diseases*, 25(1), pp. 1-3 10.1097/QCO.0b013e32834f14fa.
Flicek, P., Amode, M. R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., Gil, L., Giron, C. G., Gordon, L., Hourlier, T., Hunt, S., Johnson, N., Juettemann, T., Kahari, A. K., Keenan, S., Kulesha, E., Martin, F. J., Maurel, T., McLaren, W. M.,

Murphy, D. N., Nag, R., Overduin, B., Pignatelli, M., Pritchard, B., Pritchard, E., Riat, H. S., Ruffier, M., Sheppard, D., Taylor, K., Thormann, A., Trevanion, S. J., Vullo, A., Wilder, S. P., Wilson, M., Zadissa, A., Aken, B. L., Birney, E., Cunningham, F., Harrow, J., Herrero, J., Hubbard, T. J., Kinsella, R., Muffato, M., Parker, A., Spudich, G., Yates, A., Zerbino, D. R. and Searle, S. M. (2014) 'Ensembl 2014', *Nucleic Acids Res*, 42(Database issue), pp. D749-55.

Frazier, A. E., Chacinska, A., Truscott, K. N., Guiard, B., Pfanner, N. and Rehling, P. (2003) 'Mitochondria use different mechanisms for transport of multispanning membrane proteins through the intermembrane space', *Mol Cell Biol*, 23(21), pp. 7818-28.

Gabaldon, T. and Huynen, M. A. (2004) 'Prediction of protein function and pathways in the genome era', *Cell Mol Life Sci*, 61(7-8), pp. 930-44.

Gao, F., Bonsignori, M., Liao, H. X., Kumar, A., Xia, S. M., Lu, X., Cai, F., Hwang, K. K., Song, H., Zhou, T., Lynch, R. M., Alam, S. M., Moody, M. A., Ferrari, G., Berrong, M., Kelsoe, G., Shaw, G. M., Hahn, B. H., Montefiori, D. C., Kamanga, G., Cohen, M. S., Hraber, P., Kwong, P. D., Korber, B. T., Mascola, J. R., Kepler, T. B. and Haynes, B. F. (2014) 'Cooperation of B Cell Lineages in Induction of HIV-1-Broadly Neutralizing Antibodies', *Cell*, 158(3), pp. 481-91.

Gardner, K., Hall, P. A., Chinnery, P. F. and Payne, B. A. (2013) 'HIV Treatment and Associated Mitochondrial Pathology: Review of 25 Years of in Vitro, Animal, and Human Studies', *Toxicol Pathol*.

Gardner, K., Payne, B. A., Horvath, R. and Chinnery, P. F. (2014) 'Use of stereotypical mutational motifs to define resolution limits for the ultra-deep resequencing of mitochondrial DNA', *Eur J Hum Genet*.

Garrido-Maraver, J., Cordero, M. D., Monino, I. D., Pereira-Arenas, S., Lechuga-Vieco, A. V., Cotan, D., De la Mata, M., Oropesa-Avila, M., De Miguel, M., Bautista Lorite, J., Rivas Infante, E., Alvarez-Dolado, M., Navas, P., Jackson, S., Francisci, S. and Sanchez-Alcazar, J. A. (2012) 'Screening of effective pharmacological treatments for MELAS syndrome using yeasts, fibroblasts and cybrid models of the disease', *Br J Pharmacol*, 167(6), pp. 1311-28.

Garrido-Maraver, J., Cordero, M. D., Oropesa-Avila, M., Fernandez Vega, A., de la Mata, M., Delgado Pavon, A., de Miguel, M., Perez Calero, C., Villanueva Paz, M., Cotan, D. and Sanchez-Alcazar, J. A. (2014) 'Coenzyme q10 therapy', *Mol Syndromol*, 5(3-4), pp. 187-97.

Gendron, S. P., Mallet, J. D., Bastien, N. and Rochette, P. J. (2012) 'Mitochondrial DNA common deletion in the human eye: a relation with corneal aging', *Mech Ageing Dev*, 133(2-3), pp. 68-74.

Gerschenson, M. and Brinkman, K. (2004) 'Mitochondrial dysfunction in AIDS and its treatment', *Mitochondrion*, 4(5-6), pp. 763-77.

Ghivizzani, S. C., Madsen, C. S., Nelen, M. R., Ammini, C. V. and Hauswirth, W. W. (1994) 'In organello footprint analysis of human mitochondrial DNA: human mitochondrial transcription factor A interactions at the origin of replication', *Mol Cell Biol*, 14(12), pp. 7717-30.

Giguere, V. (2008) 'Transcriptional control of energy homeostasis by the estrogen-related receptors', *Endocr Rev*, 29(6), pp. 677-96.

Giles, R. E., Blanc, H., Cann, H. M. and Wallace, D. C. (1980) 'Maternal inheritance of human mitochondrial DNA', *Proc Natl Acad Sci U S A*, 77(11), pp. 6715-9.

Giuliani, C., Barbieri, C., Li, M., Bucci, L., Monti, D., Passarino, G., Luiselli, D., Franceschi, C., Stoneking, M. and Garagnani, P. (2014) 'Transmission from centenarians to their offspring of mtDNA heteroplasmy revealed by ultra-deep sequencing', *Aging (Albany NY)*, 6(6), pp. 454-67.

Golubitzky, A., Dan, P., Weissman, S., Link, G., Wikstrom, J. D. and Saada, A. (2011) 'Screening for active small molecules in mitochondrial complex I deficient patient's fibroblasts, reveals AICAR as the most beneficial compound', *PLoS One*, 6(10), p. e26883.

Goto, H., Dickins, B., Afgan, E., Paul, I. M., Taylor, J., Makova, K. D. and Nekrutenko, A. (2011) 'Dynamics of mitochondrial heteroplasmy in three families investigated via a repeatable re-sequencing study', *Genome Biol*, 12(6), p. R59.

Graham, B. H. (2012) 'Diagnostic challenges of mitochondrial disorders: complexities of two genomes', *Methods Mol Biol*, 837, pp. 35-46.

Grazina, M. M., Diogo, L. M., Garcia, P. C., Silva, E. D., Garcia, T. D., Robalo, C. B. and Oliveira, C. R. (2007) 'Atypical presentation of Leber's hereditary optic neuropathy associated to mtDNA 11778G>A point mutation--A case report', *Eur J Paediatr Neurol*, 11(2), pp. 115-8.
Greaves, L. C., Barron, M. J., Plusa, S., Kirkwood, T. B., Mathers, J. C., Taylor, R. W. and Turnbull, D. M. (2010) 'Defects in multiple complexes of the respiratory chain are present in ageing human colonic crypts', *Experimental Gerontology*, 45(7–8), pp. 573-579.
Greaves, L. C., Nooteboom, M., Elson, J. L., Tuppen, H. A., Taylor, G. A., Commane, D. M., Arasaradnam, R. P., Khrapko, K., Taylor, R. W., Kirkwood, T. B., Mathers, J. C. and Turnbull, D. M. (2014) 'Clonal Expansion of Early to Mid-Life Mitochondrial DNA Point Mutations Drives Mitochondrial Dysfunction during Human Ageing', *PLoS Genet*, 10(9), p. e1004620.
Guaraldi, G., Orlando, G., Zona, S., Menozzi, M., Carli, F. and Garlassi, E. (2011) 'Premature age-related comorbidities among HIV-infected persons compared with the general population', *Clin Infect Dis*, 53, pp. 1120-1126.
Guilherme, A., Virbasius, J. V., Puri, V. and Czech, M. P. (2008) 'Adipocyte dysfunctions linking obesity to insulin resistance and type 2 diabetes', *Nat Rev Mol Cell Biol*, 9(5), pp. 367-77.
Guo, Y., Li, C. I., Sheng, Q., Winther, J. F., Cai, Q., Boice, J. D. and Shyr, Y. (2013) 'Very low-level heteroplasmy mtDNA variations are inherited in humans', *J Genet Genomics*, 40(12), pp. 607-15.
Guo, Y., Li, J., Li, C.-I., Long, J., Samuels, D. and Shyr, Y. (2012) 'The effect of strand bias in Illumina short-read sequencing data', *BMC Genomics*, 13(1), p. 666.
Gyllensten, U., Wharton, D., Josefsson, A. and Wilson, A. C. (1991) 'Paternal inheritance of mitochondrial DNA in mice', *Nature*, 352(6332), pp. 255-7.
Haack, T. B., Haberberger, B., Frisch, E. M., Wieland, T., Iuso, A., Gorza, M., Strecker, V., Graf, E., Mayr, J. A., Herberg, U., Hennermann, J. B., Klopstock, T., Kuhn, K. A., Ahting, U., Sperl, W., Wilichowski, E., Hoffmann, G. F., Tesarova, M., Hansikova, H., Zeman, J., Plecko, B., Zeviani, M., Wittig, I., Strom, T. M., Schuelke, M., Freisinger, P., Meitinger, T. and Prokisch, H. (2012) 'Molecular diagnosis in mitochondrial complex I deficiency using exome sequencing', *J Med Genet*, 49(4), pp. 277-83.
Han, Y., Wind-Rotolo, M., Yang, H.-C., Siliciano, J. D. and Siliciano, R. F. (2007) 'Experimental approaches to the study of HIV-1 latency', *Nat Rev Micro*, 5(2), pp. 95-106.
Harman, D. (1956) 'Aging: a theory based on free radical and radiation chemistry', *J Gerontol*, 11(3), pp. 298-300.
Hart, A. M., Wilson, A. D. H., Montovani, C., Smith, C., Johnson, M., Terenghi, G. and Youle, M. (2004) 'Acetyl-l-carnitine: a pathogenesis based treatment for HIV-associated antiretroviral toxic neuropathy', *AIDS*, 18(11), pp. 1549-1560.
Hazkani-Covo, E., Zeller, R. M. and Martin, W. (2010) 'Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes', *PLoS Genet*, 6(2), p. e1000834.
He, L., Chinnery, P. F., Durham, S. E., Blakely, E. L., Wardell, T. M., Borthwick, G. M., Taylor, R. W. and Turnbull, D. M. (2002) 'Detection and quantification of mitochondrial DNA deletions in individual cells by real-time PCR', *Nucleic Acids Res*, 30(14), p. e68.
He, Y., Wu, J., Dressman, D. C., Iacobuzio-Donahue, C., Markowitz, S. D., Velculescu, V. E., Diaz, L. A., Jr., Kinzler, K. W., Vogelstein, B. and Papadopoulos, N. (2010) 'Heteroplasmic mitochondrial DNA mutations in normal and tumour cells', *Nature*, 464(7288), pp. 610-4.
Helbert, M., Fletcher, T., Peddle, B., Harris, J. R. and Pinching, A. J. (1988) 'Zidovudine-associated myopathy', *Lancet*, 2(8612), pp. 689-90.
Herbst, A., Pak, J. W., McKenzie, D., Bua, E., Bassiouni, M. and Aiken, J. M. (2007) 'Accumulation of mitochondrial DNA deletion mutations in aged muscle fibers: evidence for a causal role in muscle fiber loss', *J Gerontol A Biol Sci Med Sci*, 62(3), pp. 235-45.
Hirano, M., Shtilbans, A., Mayeux, R., Davidson, M. M., DiMauro, S., Knowles, J. A. and Schon, E. A. (1997) 'Apparent mtDNA heteroplasmy in Alzheimer's disease patients and in normals due to PCR amplification of nucleus-embedded mtDNA pseudogenes', *Proc Natl Acad Sci U S A*, 94(26), pp. 14894-9.

Hirst, J., King, M. S. and Pryde, K. R. (2008) 'The production of reactive oxygen species by complex I', *Biochem Soc Trans*, 36(Pt 5), pp. 976-80.
Hisatomi, T., Nakazawa, T., Noda, K., Almulki, L., Miyahara, S., Nakao, S., Ito, Y., She, H., Kohno, R., Michaud, N., Ishibashi, T., Hafezi-Moghadam, A., Badley, A. D., Kroemer, G. and Miller, J. W. (2008) 'HIV protease inhibitors provide neuroprotection through inhibition of mitochondrial apoptosis in mice', *J Clin Invest*, 118(6), pp. 2025-38.
Holt, I. J., Harding, A. E. and Morgan-Hughes, J. A. (1988) 'Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies', *Nature*, 331(6158), pp. 717-719.
Holt, I. J., Lorimer, H. E. and Jacobs, H. T. (2000) 'Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA', *Cell*, 100(5), pp. 515-24.
Horner, D. S., Pavesi, G., Castrignano, T., De Meo, P. D., Liuni, S., Sammeth, M., Picardi, E. and Pesole, G. (2010) 'Bioinformatics approaches for genomics and post genomics applications of next-generation sequencing', *Brief Bioinform*, 11(2), pp. 181-97.
Höschele, D. (2006) 'Cell culture models for the investigation of NRTI-induced mitochondrial toxicity.: Relevance for the prediction of clinical toxicity', *Toxicology in Vitro*, 20(5), pp. 535-546.
Hudson, G. and Chinnery, P. F. (2006) 'Mitochondrial DNA polymerase-gamma and human disease', *Hum Mol Genet*, 15 Spec No 2, pp. R244-52.
Hulgan, T. and Gerschenson, M. (2012) 'HIV and Mitochondria: More Than Just Drug Toxicity', *Journal of Infectious Diseases*, 205(12), pp. 1769-1771.
Iborra, F. J., Kimura, H. and Cook, P. R. (2004) 'The functional organization of mitochondrial genomes in human cells', *BMC Biol*, 2, p. 9.
Iloeje, U. H., Yuan, Y., L'Italien, G., Mauskopf, J., Holmberg, S. D., Moorman, A. C., Wood, K. C. and Moore, R. D. (2005) 'Protease inhibitor exposure and increased risk of cardiovascular disease in HIV-infected patients', *HIV Med*, 6(1), pp. 37-44.
Jornayvaz, F. R. and Shulman, G. I. (2010) 'Regulation of mitochondrial biogenesis', *Essays Biochem*, 47, pp. 69-84.
Jose, C., Hebert-Chatelain, E., Bellance, N., Larendra, A., Su, M., Nouette-Gaulain, K. and Rossignol, R. (2011) 'AICAR inhibits cancer cell growth and triggers cell-type distinct effects on OXPHOS biogenesis, oxidative stress and Akt activation', *Biochim Biophys Acta*, 1807(6), pp. 707-18.
Kaguni, L. S. (2004) 'DNA polymerase gamma, the mitochondrial replicase', *Annu Rev Biochem*, 73, pp. 293-320.
Kamerman, P. R., Moss, P. J., Weber, J., Wallace, V. C. J., Rice, A. S. C. and Huang, W. (2012) 'Pathogenesis of HIV-associated sensory neuropathy: evidence from in vivo and in vitro experimental models', *Journal of the Peripheral Nervous System*, 17(1), pp. 19-31.
Kampira, E., Kumwenda, J., Van Oosterhout, J. J. and Dandara, C. (2013) 'Mitochondrial subhaplogroups and differential risk of stavudine-induced lipodystrophy in Malawian HIV/AIDS patients', *Pharmacogenomics*, 14(16), pp. 1999-2004.
Kassem, A. M., El-Guendy, N., Tantawy, M., Abdelhady, H., El-Ghor, A. and Abdel Wahab, A. H. (2011) 'Mutational hotspots in the mitochondrial D-loop region of cancerous and precancerous colorectal lesions in Egyptian patients', *DNA Cell Biol*, 30(11), pp. 899-906.
Kennedy, S. R., Salk, J. J., Schmitt, M. W. and Loeb, L. A. (2013) 'Ultra-sensitive sequencing reveals an age-related increase in somatic mitochondrial mutations that are inconsistent with oxidative damage', *PLoS Genet*, 9(9), p. e1003794.
Khrapko, K. (2008) 'Two ways to make an mtDNA bottleneck', *Nat Genet*, 40(2), pp. 134-5.
Khrapko, K. (2011) 'The timing of mitochondrial DNA mutations in aging', *Nat Genet*, 43(8), pp. 726-7.
Khrapko, K., Bodyak, N., Thilly, W. G., van Orsouw, N. J., Zhang, X., Coller, H. A., Perls, T. T., Upton, M., Vijg, J. and Wei, J. Y. (1999) 'Cell-by-cell scanning of whole mitochondrial genomes in aged human heart reveals a significant fraction of myocytes with clonally expanded deletions', *Nucleic Acids Res*, 27(11), pp. 2434-41.

Khrapko, K. and Vijg, J. (2007) 'Mitochondrial DNA mutations and aging: a case closed?', *Nat Genet*, 39(4), pp. 445-6.
Koboldt, D. C., Chen, K., Wylie, T., Larson, D. E., McLellan, M. D., Mardis, E. R., Weinstock, G. M., Wilson, R. K. and Ding, L. (2009) 'VarScan: variant detection in massively parallel sequencing of individual and pooled samples', *Bioinformatics*, 25(17), pp. 2283-5.
Kohler, J. J. and Lewis, W. (2007) 'A brief overview of mechanisms of mitochondrial toxicity from NRTIs', *Environmental and Molecular Mutagenesis*, 48(3-4), pp. 166-172.
Kopsidas, G., Kovalenko, S. A., Kelso, J. M. and Linnane, A. W. (1998) 'An age-associated correlation between cellular bioenergy decline and mtDNA rearrangements in human skeletal muscle', *Mutat Res*, 421(1), pp. 27-36.
Kornblum, C., Nicholls, T. J., Haack, T. B., Scholer, 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. and Prokisch, H. (2013) 'Loss-of-function mutations in MGME1 impair mtDNA replication and cause multisystemic mitochondrial disease', *Nat Genet*, 45(2), pp. 214-9.
Korr, H., Kurz, C., Seidler, T. O., Sommer, D. and Schmitz, C. (1998) 'Mitochondrial DNA synthesis studied autoradiographically in various cell types in vivo', *Braz J Med Biol Res*, 31(2), pp. 289-98.
Kota, B. P., Huang, T. H. and Roufogalis, B. D. (2005) 'An overview on biological mechanisms of PPARs', *Pharmacol Res*, 51(2), pp. 85-94.
Kotler, D. P., Ionescu, G., Johnson, J. A., Inada, Y., He, Q., Engelson, E. S. and Albu, J. B. (2003) 'Studies of adipose tissue metabolism in human immunodeficiency virus-associated lipodystrophy', *Clin Infect Dis*, 37(2), pp. S47-51.
Kowald, A. and Kirkwood, T. B. (2013) 'Mitochondrial mutations and aging: random drift is insufficient to explain the accumulation of mitochondrial deletion mutants in short-lived animals', *Aging Cell*, 12(4), pp. 728-31.
Kozarewa, I., Ning, Z., Quail, M. A., Sanders, M. J., Berriman, M. and Turner, D. J. (2009) 'Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes', *Nat Meth*, 6(4), pp. 291-295.
Krishnan, K. J., Greaves, L. C., Reeve, A. K. and Turnbull, D. M. (2007) 'Mitochondrial DNA mutations and aging', *Ann N Y Acad Sci*, 1100, pp. 227-40.
Kukat, C., Wurm, C. A., Spahr, H., Falkenberg, M., Larsson, N. G. and Jakobs, S. (2011) 'Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA', *Proc Natl Acad Sci U S A*, 108(33), pp. 13534-9.
Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J. and Hendrickson, W. A. (1998) 'Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody', *Nature*, 393(6686), pp. 648-59.
Kyriacou, K., Mikellidou, C., Hadjianastasiou, A., Middleton, L., Panousopoulos, A. and Kyriakides, T. (1999) 'Ultrastructural diagnosis of mitochondrial encephalomyopathies revisited', *Ultrastruct Pathol*, 23(3), pp. 163-70.
Lamont, P. J., Surtees, R., Woodward, C. E., Leonard, J. V., Wood, N. W. and Harding, A. E. (1998) 'Clinical and laboratory findings in referrals for mitochondrial DNA analysis', *Arch Dis Child*, 79(1), pp. 22-7.
Lane, N. and Martin, W. (2010) 'The energetics of genome complexity', *Nature*, 467(7318), pp. 929-34.
Larsson, N. G. (2010) 'Somatic mitochondrial DNA mutations in mammalian aging. ', *Annu Rev Biochem*, 79, pp. 683-706.
Lefèvre, C., Auclair, M., Boccara, F., Bastard, J.-P., Capeau, J., Vigouroux, C. and Caron-Debarle, M. (2010) 'Premature Senescence of Vascular Cells Is Induced by HIV Protease Inhibitors: Implication of Prelamin A and Reversion by Statin', *Arteriosclerosis, Thrombosis, and Vascular Biology*, 30(12), pp. 2611-2620.

Lehmann, H. C., Chen, W., Borzan, J., Mankowski, J. L. and Hoke, A. (2011) 'Mitochondrial dysfunction in distal axons contributes to human immunodeficiency virus sensory neuropathy', *Ann Neurol*, 69(1), pp. 100-10.

Lewis, W., Day, B. J. and Copeland, W. C. (2003) 'Mitochondrial toxicity of nrti antiviral drugs: an integrated cellular perspective', *Nat Rev Drug Discov*, 2(10), pp. 812-822.

Lewis, W., Gonzalez, B., Chomyn, A. and Papoian, T. (1992) 'Zidovudine induces molecular, biochemical, and ultrastructural changes in rat skeletal muscle mitochondria', *J Clin Invest*, 89(4), pp. 1354-60.

Lezza, A. M. S., Boffoli, D., Scacco, S., Cantatore, P. and Gadaleta, M. N. (1994) 'Correlation Between Mitochondrial DNA 4977-bp Deletion and Respiratory Chain Enzyme Activities in Aging Human Skeletal Muscles', *Biochemical and Biophysical Research Communications*, 205(1), pp. 772-779.

Li, H. and Durbin, R. (2009) 'Fast and accurate short read alignment with Burrows-Wheeler transform', *Bioinformatics*, 25(14), pp. 1754-60.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009) 'The Sequence Alignment/Map format and SAMtools', *Bioinformatics*, 25(16), pp. 2078-9.

Li, L., Pan, R., Li, R., Niemann, B., Aurich, A. C., Chen, Y. and Rohrbach, S. (2011) 'Mitochondrial biogenesis and peroxisome proliferator-activated receptor-gamma coactivator-1alpha (PGC-1alpha) deacetylation by physical activity: intact adipocytokine signaling is required', *Diabetes*, 60(1), pp. 157-67.

Li, M., Schönberg, A., Schaefer, M., Schroeder, R., Nasidze, I. and Stoneking, M. (2010) 'Detecting Heteroplasmy from High-Throughput Sequencing of Complete Human Mitochondrial DNA Genomes', *The American Journal of Human Genetics*, 87(2), pp. 237-249.

Lim, S. E. and Copeland, W. C. (2001) 'Differential incorporation and removal of antiviral deoxynucleotides by human DNA polymerase gamma', *J Biol Chem*, 276(26), pp. 23616-23.

Liu, K., Sun, Y., Liu, D., Yin, J., Qiao, L., Shi, Y., Dong, Y., Li, N., Zhang, F. and Chen, D. (2013) 'Mitochondrial toxicity studied with the PBMC of children from the Chinese national pediatric highly active antiretroviral therapy cohort', *PLoS One*, 8(2), p. e57223.

Liu, Z., Liu, R., Zhang, G., Hao, F. and Yan, H. (2002) '[Study on point mutations in mitochondrial DNA control region for replication DLP(6) of cultured dermal fibroblast with 8-MOP/UVA treatment]', *Zhonghua Yi Xue Yi Chuan Xue Za Zhi*, 19(5), pp. 386-8.

Livak, K. J. and Schmittgen, T. D. (2001) 'Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method', *Methods*, 25(4), pp. 402-408.

Longley, M. J., Ropp, P. A., Lim, S. E. and Copeland, W. C. (1998) 'Characterization of the Native and Recombinant Catalytic Subunit of Human DNA Polymerase γ: Identification of Residues Critical for Exonuclease Activity and Dideoxynucleotide Sensitivity', *Biochemistry*, 37(29), pp. 10529-10539.

Lund, K. C., Peterson, L. L. and Wallace, K. B. (2007) 'Absence of a universal mechanism of mitochondrial toxicity by nucleoside analogs', *Antimicrob Agents Chemother*, 51(7), pp. 2531-9.

Maagaard, A., Holberg-Petersen, M., Kollberg, G., Oldfors, A., Sandvik, L. and Bruun, J. N. (2006) 'Mitochondrial (mt)DNA changes in tissue may not be reflected by depletion of mtDNA in peripheral blood mononuclear cells in HIV-infected patients', *Antivir Ther*, 11(5), pp. 601-8.

Maggi, P., Montinaro, V., Bellacosa, C., Pietanza, S., Volpe, A., Graziano, G., Strippoli, G. F. and Angarano, G. (2012) 'Early markers of tubular dysfunction in antiretroviral-experienced HIV-infected patients treated with tenofovir versus abacavir', *AIDS Patient Care STDS*, 26(1), pp. 5-11.

Malyarchuk, B. A., Papuga, M., Grzybowski, T., Rogozin, I. B., Wozniak, M., Derenko, M. V., Rychkov, S. Y., Czarny, J., Zakharov, I. A. and Miscicka-Sliwka, D. (2005) 'Low variability of the POLG (CAG)n repeat in north Eurasian populations', *Hum Biol*, 77(3), pp. 355-65.

Man, P. Y., Griffiths, P. G., Brown, D. T., Howell, N., Turnbull, D. M. and Chinnery, P. F. (2003) 'The epidemiology of Leber hereditary optic neuropathy in the North East of England', *Am J Hum Genet*, 72(2), pp. 333-9.
Margulis, L. (1975) 'Symbiotic theory of the origin of eukaryotic organelles; criteria for proof', *Symp Soc Exp Biol*, (29), pp. 21-38.
Markley, H. G. (2012) 'CoEnzyme Q10 and riboflavin: the mitochondrial connection', *Headache*, 52 Suppl 2, pp. 81-7.
Martin, A. M., Hammond, E., Nolan, D., Pace, C., Den Boer, M., Taylor, L., Moore, H., Martinez, O. P., Christiansen, F. T. and Mallal, S. (2003) 'Accumulation of mitochondrial DNA mutations in human immunodeficiency virus-infected patients treated with nucleoside-analogue reverse-transcriptase inhibitors', *Am J Hum Genet*, 72(3), pp. 549-60.
Martínez, E. (1998) 'Fat accumulation and HIV-1 protease inhibitors', *The Lancet*, 352(9137), p. 1392.
Martinez, E., Mocroft, A., Garcia-Viejo, M. A., Perez-Cuevas, J. B., Blanco, J. L., Mallolas, J., Bianchi, L., Conget, I., Blanch, J., Phillips, A. and Gatell, J. M. (2001) 'Risk of lipodystrophy in HIV-1-infected patients treated with protease inhibitors: a prospective cohort study', *Lancet*, 357(9256), pp. 592-8.
Matthews, P. M., Ford, B., Dandurand, R. J., Eidelman, D. H., O'Connor, D., Sherwin, A., Karpati, G., Andermann, F. and Arnold, D. L. (1993) 'Coenzyme Q10 with multiple vitamins is generally ineffective in treatment of mitochondrial disease', *Neurology*, 43(5), pp. 884-90.
McComsey, G., Bai, R. K., Maa, J. F., Seekins, D. and Wong, L. J. (2005a) 'Extensive investigations of mitochondrial DNA genome in treated HIV-infected subjects: beyond mitochondrial DNA depletion', *J Acquir Immune Defic Syndr*, 39(2), pp. 181-8.
McComsey, G. A., Paulsen, D. M., Lonergan, J. T., Hessenthaler, S. M., Hoppel, C. L., Williams, V. C., Fisher, R. L., Cherry, C. L., White-Owen, C., Thompson, K. A., Ross, S. T., Hernandez, J. E. and Ross, L. L. (2005b) 'Improvements in lipoatrophy, mitochondrial DNA levels and fat apoptosis after replacing stavudine with abacavir or zidovudine', *Aids*, 19(1), pp. 15-23.
Meissner, C., Bruse, P., Mohamed, S. A., Schulz, A., Warnk, H., Storm, T. and Oehmichen, M. (2008) 'The 4977bp deletion of mitochondrial DNA in human skeletal muscle, heart and different areas of the brain: A useful biomarker or more?', *Experimental Gerontology*, 43(7), pp. 645-652.
Menezes, C. N., Maskew, M., Sanne, I., Crowther, N. J. and Raal, F. J. (2011) 'A longitudinal study of stavudine-associated toxicities in a large cohort of South African HIV infected subjects', *BMC Infect Dis*, 11, p. 244.
Merson, M. H., O'Malley, J., Serwadda, D. and Apisuk, C. (2008) 'The history and challenge of HIV prevention', *Lancet*, 372(9637), pp. 475-88.
Michikawa, Y., Mazzucchelli, F., Bresolin, N., Scarlato, G. and Attardi, G. (1999) 'Aging-dependent large accumulation of point mutations in the human mtDNA control region for replication', *Science*, 286(5440), pp. 774-9.
Miglio, G., Rosa, A. C., Rattazzi, L., Collino, M., Lombardi, G. and Fantozzi, R. (2009) 'PPARgamma stimulation promotes mitochondrial biogenesis and prevents glucose deprivation-induced neuronal cell loss', *Neurochem Int*, 55(7), pp. 496-504.
Mishra, P. and Chan, D. C. (2014) 'Mitochondrial dynamics and inheritance during cell division, development and disease', *Nat Rev Mol Cell Biol*, 15(10), pp. 634-646.
Moore, J. P. (1997) 'Coreceptors: implications for HIV pathogenesis and therapy', *Science*, 276(5309), pp. 51-2.
Morten, K., Field, P., Ashley, N., Williams, K. A., Harris, D., Hartley, M., Clark, A. and Poulton, J. (2005) *Fetal and neonatal exposure to AZT and low-protein diet affects glucose homeostasis: a model with implications for AIDS prevention*.
Moyes, C. D. and Hood, D. A. (2003) 'Origins and consequences of mitochondrial variation in vertebrate muscle', *Annu Rev Physiol*, 65, pp. 177-201.

Moyle, G. (2000) 'Clinical manifestations and management of antiretroviral nucleoside analog-related mitochondrial toxicity', *Clinical Therapeutics*, 22(8), pp. 911-936.
Mukhopadhyay, A., Wei, B., Zullo, S. J., Wood, L. V. and Weiner, H. (2002) 'In vitro evidence of inhibition of mitochondrial protease processing by HIV-1 protease inhibitors in yeast: a possible contribution to lipodystrophy syndrome', *Mitochondrion*, 1(6), pp. 511-8.
Nass, M. M. (1969) 'Mitochondrial DNA. I. Intramitochondrial distribution and structural relations of single- and double-length circular DNA', *J Mol Biol*, 42(3), pp. 521-8.
Nicklas, J. A., Brooks, E. M., Hunter, T. C., Single, R. and Branda, R. F. (2004) 'Development of a quantitative PCR (TaqMan) assay for relative mitochondrial DNA copy number and the common mitochondrial DNA deletion in the rat', *Environ Mol Mutagen*, 44(4), pp. 313-20.
Nishigaki, Y., Martí, R., Copeland, W. C. and Hirano, M. (2003) 'Site-specific somatic mitochondrial DNA point mutations in patients with thymidine phosphorylase deficiency', *The Journal of Clinical Investigation*, 111(12), pp. 1913-1921.
Noe, N., Dillon, L., Lellek, V., Diaz, F., Hida, A., Moraes, C. T. and Wenz, T. (2013) 'Bezafibrate improves mitochondrial function in the CNS of a mouse model of mitochondrial encephalopathy', *Mitochondrion*, 13(5), pp. 417-26.
Nooteboom, M., Johnson, R., Taylor, R. W., Wright, N. A., Lightowlers, R. N., Kirkwood, T. B. L., Mathers, J. C., Turnbull, D. M. and Greaves, L. C. (2010) 'Age-associated mitochondrial DNA mutations lead to small but significant changes in cell proliferation and apoptosis in human colonic crypts', *Aging Cell*, 9(1), pp. 96-99.
O'Brien, T. W. (2003) 'Properties of human mitochondrial ribosomes', *IUBMB Life*, 55(9), pp. 505-13.
Oldfors, A., Fyhr, I. M., Holme, E., Larsson, N. G. and Tulinius, M. (1990) 'Neuropathology in Kearns-Sayre syndrome', *Acta Neuropathologica*, 80(5), pp. 541-546.
Olichon, A., Baricault, L., Gas, N., Guillou, E., Valette, A., Belenguer, P. and Lenaers, G. (2003) 'Loss of OPA1 Perturbates the Mitochondrial Inner Membrane Structure and Integrity, Leading to Cytochrome c Release and Apoptosis', *Journal of Biological Chemistry*, 278(10), pp. 7743-7746.
P, H. and Nelson A, M. (2006) 'The Pathology Induced by Highly Active Antiretroviral Therapy Against Human Immunodeficiency Virus: an Update', *Current Medicinal Chemistry*, 13(26), pp. 3121-3132.
Pallotti, F., Chen, X., Bonilla, E. and Schon, E. A. (1996) 'Evidence that specific mtDNA point mutations may not accumulate in skeletal muscle during normal human aging', *Am J Hum Genet*, 59(3), pp. 591-602.
Palmer, S., Maldarelli, F., Wiegand, A., Bernstein, B., Hanna, G. J., Brun, S. C., Kempf, D. J., Mellors, J. W., Coffin, J. M. and King, M. S. (2008) 'Low-level viremia persists for at least 7 years in patients on suppressive antiretroviral therapy', *Proceedings of the National Academy of Sciences*, 105(10), pp. 3879-3884.
Parikh, S., Saneto, R., Falk, M. J., Anselm, I., Cohen, B. H., Haas, R. and Medicine Society, T. M. (2009) 'A modern approach to the treatment of mitochondrial disease', *Curr Treat Options Neurol*, 11(6), pp. 414-30.
Patil, V. A. and Greenberg, M. L. (2013) 'Cardiolipin-mediated cellular signaling', *Adv Exp Med Biol*, 991, pp. 195-213.
Pavlakis, S. G., Phillips, P. C., DiMauro, S., De Vivo, D. C. and Rowland, L. P. (1984) 'Mitochondrial myopathy, encephalopathy, lactic acidosis, and strokelike episodes: a distinctive clinical syndrome', *Ann Neurol*, 16(4), pp. 481-8.
Payne, B. A., Gardner, K. and Chinnery, P. F. (2014) 'Mitochondrial DNA mutations in ageing and disease: implications for HIV?', *Antivir Ther*.
Payne, B. A., Wilson, I. J., Hateley, C. A., Horvath, R., Santibanez-Koref, M., Samuels, D. C., Price, D. A. and Chinnery, P. F. (2011) 'Mitochondrial aging is accelerated by anti-retroviral therapy through the clonal expansion of mtDNA mutations', *Nat Genet*, 43(8), pp. 806-10.

Payne, B. A., Wilson, I. J., Yu-Wai-Man, P., Coxhead, J., Deehan, D., Horvath, R., Taylor, R. W., Samuels, D. C., Santibanez-Koref, M. and Chinnery, P. F. (2013) 'Universal heteroplasmy of human mitochondrial DNA', *Hum Mol Genet*, 22(2), pp. 384-90.

Peeters, M., Jung, M. and Ayouba, A. (2013) 'The origin and molecular epidemiology of HIV', *Expert Rev Anti Infect Ther*, 11(9), pp. 885-96.

Petruzzella, V., Fracasso, F., Gadaleta, M. N. and Cantatore, P. (1992) 'Age-dependent structural variations in rat brain mitochondrial DNA', *Ann N Y Acad Sci*, 673, pp. 194-9.

Phillips, N. R., Simpkins, J. W. and Roby, R. K. (2013) 'Mitochondrial DNA deletions in Alzheimer's brains: A review.', *Alzheimer's Dementia*.

Phillips, N. R., Sprouse, M. L. and Roby, R. K. (2014) 'Simultaneous quantification of mitochondrial DNA copy number and deletion ratio: a multiplex real-time PCR assay', *Sci Rep*, 4, p. 3887.

Pireddu, L., Leo, S. and Zanetti, G. (2011) 'SEAL: a distributed short read mapping and duplicate removal tool', *Bioinformatics*, 27(15), pp. 2159-60.

Pujades-Rodriguez, M., Dantony, E., Pinoges, L., Ecochard, R., Etard, J. F., Carrillo-Casas, E. and Szumilin, E. (2011) 'Toxicity associated with stavudine dose reduction from 40 to 30 mg in first-line antiretroviral therapy', *PLoS One*, 6(11), p. e28112.

Rerks-Ngarm, S., Pitisuttithum, P., Nitayaphan, S., Kaewkungwal, J., Chiu, J., Paris, R., Premsri, N., Namwat, C., de Souza, M., Adams, E., Benenson, M., Gurunathan, S., Tartaglia, J., McNeil, J. G., Francis, D. P., Stablein, D., Birx, D. L., Chunsuttiwat, S., Khamboonruang, C., Thongcharoen, P., Robb, M. L., Michael, N. L., Kunasol, P. and Kim, J. H. (2009) 'Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand', *N Engl J Med*, 361(23), pp. 2209-20.

Reyes, A., Nevo, E. and Saccone, C. (2003) 'DNA sequence variation in the mitochondrial control region of subterranean mole rats, Spalax ehrenbergi superspecies, in Israel', *Mol Biol Evol*, 20(4), pp. 622-32.

Rodriguez-Novoa, S., Barreiro, P., Jimenez-Nacher, I. and Soriano, V. (2006) 'Overview of the pharmacogenetics of HIV therapy', *Pharmacogenomics J*, 6(4), pp. 234-245.

Rorbach, J., Richter, R., Wessels, H. J., Wydro, M., Pekalski, M., Farhoud, M., Kuhl, I., Gaisne, M., Bonnefoy, N., Smeitink, J. A., Lightowlers, R. N. and Chrzanowska-Lightowlers, Z. M. (2008) 'The human mitochondrial ribosome recycling factor is essential for cell viability', *Nucleic Acids Res*, 36(18), pp. 5787-99.

Rosen, E. D. and Spiegelman, B. M. (2006) 'Adipocytes as regulators of energy balance and glucose homeostasis', *Nature*, 444(7121), pp. 847-53.

Ross, J. M., Coppotelli, G., Hoffer, B. J. and Olson, L. (2014) 'Maternally transmitted mitochondrial DNA mutations can reduce lifespan', *Sci. Rep.*, 4.

Rotig, A., Lebon, S., Zinovieva, E., Mollet, J., Sarzi, E., Bonnefont, J. P. and Munnich, A. (2004) 'Molecular diagnostics of mitochondrial disorders', *Biochim Biophys Acta*, 1659(2-3), pp. 129-35.

Roumier, T., Szabadkai, G., Simoni, A. M., Perfettini, J. L., Paulau, A. L., Castedo, M., Metivier, D., Badley, A., Rizzuto, R. and Kroemer, G. (2005) 'HIV-1 protease inhibitors and cytomegalovirus vMIA induce mitochondrial fragmentation without triggering apoptosis', *Cell Death Differ*, 13(2), pp. 348-351.

Rovio, A. T., Marchington, D. R., Donat, S., Schuppe, H. C., Abel, J., Fritsche, E., Elliott, D. J., Laippala, P., Ahola, A. L., McNay, D., Harrison, R. F., Hughes, B., Barrett, T., Bailey, D. M., Mehmet, D., Jequier, A. M., Hargreave, T. B., Kao, S. H., Cummins, J. M., Barton, D. E., Cooke, H. J., Wei, Y. H., Wichmann, L., Poulton, J. and Jacobs, H. T. (2001) 'Mutations at the mitochondrial DNA polymerase (POLG) locus associated with male infertility', *Nat Genet*, 29(3), pp. 261-2.

Ryazantsev, D. Y., Kvach, M. V., Tsybulsky, D. A., Prokhorenko, I. A., Stepanova, I. A., Martynenko, Y. V., Gontarev, S. V., Shmanai, V. V., Zavriev, S. K. and Korshun, V. A. (2014) 'Design of molecular beacons: 3' couple quenchers improve fluorogenic properties of a probe in real-time PCR assay', *Analyst*, 139(11), pp. 2867-72.

Rylova, S. N., Albertioni, F., Flygh, G. and Eriksson, S. (2005) 'Activity profiles of deoxynucleoside kinases and 5'-nucleotidases in cultured adipocytes and myoblastic cells: insights into mitochondrial toxicity of nucleoside analogs', *Biochem Pharmacol*, 69(6), pp. 951-60.

Santel, A. and Fuller, M. T. (2001) 'Control of mitochondrial morphology by a human mitofusin', *Journal of Cell Science*, 114(5), pp. 867-874.

Saretzki, G. (2009) 'Telomerase, mitochondria and oxidative stress', *Exp Gerontol*, 44(8), pp. 485-492.

Sato, D., Itami, N., Tasaki, H., Takeo, S., Kuwayama, T. and Iwata, H. (2014) 'Relationship between mitochondrial DNA copy number and SIRT1 expression in porcine oocytes', *PLoS One*, 9(4), p. e94488.

Sauerbeck, A., Gao, J., Readnower, R., Liu, M., Pauly, J. R., Bing, G. and Sullivan, P. G. (2011) 'Pioglitazone attenuates mitochondrial dysfunction, cognitive impairment, cortical tissue loss, and inflammation following traumatic brain injury', *Experimental Neurology*, 227(1), pp. 128-135.

Schaefer, A. M., McFarland, R., Blakely, E. L., He, L., Whittaker, R. G., Taylor, R. W., Chinnery, P. F. and Turnbull, D. M. (2008) 'Prevalence of mitochondrial DNA disease in adults', *Ann Neurol*, 63(1), pp. 35-9.

Schapira, A. H. V. (2012) 'Mitochondrial diseases', *The Lancet*, 379(9828), pp. 1825-1834.

Schreiber, S. N., Emter, R., Hock, M. B., Knutti, D., Cardenas, J., Podvinec, M., Oakeley, E. J. and Kralli, A. (2004) 'The estrogen-related receptor alpha (ERRalpha) functions in PPARgamma coactivator 1alpha (PGC-1alpha)-induced mitochondrial biogenesis', *Proc Natl Acad Sci U S A*, 101(17), pp. 6472-7.

Schwartz, M. and Vissing, J. (2002) 'Paternal inheritance of mitochondrial DNA', *N Engl J Med*, 347(8), pp. 576-80.

Seibel, P., Di Nunno, C., Kukat, C., Schafer, I., Del Bo, R., Bordoni, A., Comi, G. P., Schon, A., Capuano, F., Latorre, D. and Villani, G. (2008) 'Cosegregation of novel mitochondrial 16S rRNA gene mutations with the age-associated T414G variant in human cybrids', *Nucleic Acids Res*, 36(18), pp. 5872-81.

Shadel, G. S. and Clayton, D. A. (1997) 'Mitochondrial DNA maintenance in vertebrates', *Annu Rev Biochem*, 66, pp. 409-35.

Sharp, P. M. and Hahn, B. H. (2011) 'Origins of HIV and the AIDS pandemic', *Cold Spring Harb Perspect Med*, 1(1), p. a006841.

Shin, M. G., Kajigaya, S., McCoy, J. P., Levin, B. C. and Young, N. S. (2004) *Marked mitochondrial DNA sequence heterogeneity in single CD34+ cell clones from normal adult bone marrow*.

Singh, A., Hemal, A., Agarwal, S., Dubey, N. and Buxi, G. (2014) 'A prospective study of haematological changes after switching from stavudine to zidovudine-based antiretroviral treatment in HIV-infected children', *Int J STD AIDS*.

Sinha, A., Hollingsworth, K. G., Ball, S. and Cheetham, T. (2013) 'Improving the vitamin D status of vitamin D deficient adults is associated with improved mitochondrial oxidative function in skeletal muscle', *J Clin Endocrinol Metab*, 98(3), pp. E509-13.

Soares, A. R., Pereira, P. M. and Santos, M. A. (2012) 'Next-generation sequencing of miRNAs with Roche 454 GS-FLX technology: steps for a successful application', *Methods Mol Biol*, 822, pp. 189-204.

Spangenburg, E. E., Jackson, K. C. and Schuh, R. A. (2013) 'AICAR inhibits oxygen consumption by intact skeletal muscle cells in culture', *J Physiol Biochem*, 69(4), pp. 909-17.

Spelbrink, J. N., Li, F. Y., Tiranti, V., Nikali, K., Yuan, Q. P., Tariq, M., Wanrooij, S., Garrido, N., Comi, G., Morandi, L., Santoro, L., Toscano, A., Fabrizi, G. M., Somer, H., Croxen, R., Beeson, D., Poulton, J., Suomalainen, A., Jacobs, H. T., Zeviani, M. and Larsson, C. (2001) 'Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria', *Nat Genet*, 28(3), pp. 223-31.

Srivastava, S., Diaz, F., Iommarini, L., Aure, K., Lombes, A. and Moraes, C. T. (2009) 'PGC-1alpha/beta induced expression partially compensates for respiratory chain defects in cells from patients with mitochondrial disorders', *Hum Mol Genet*, 18(10), pp. 1805-12.
Stankov, M. V., Lucke, T., Das, A. M., Schmidt, R. E. and Behrens, G. M. (2010) 'Mitochondrial DNA depletion and respiratory chain activity in primary human subcutaneous adipocytes treated with nucleoside analogue reverse transcriptase inhibitors', *Antimicrob Agents Chemother*, 54(1), pp. 280-7.
Stojanovski, D., Koutsopoulos, O. S., Okamoto, K. and Ryan, M. T. (2004) 'Levels of human Fis1 at the mitochondrial outer membrane regulate mitochondrial morphology', *Journal of Cell Science*, 117(7), pp. 1201-1210.
Strahl, C. and Blackburn, E. H. (1996) 'Effects of reverse transcriptase inhibitors on telomere length and telomerase activity in two immortalized human cell lines', *Molecular and Cellular Biology*, 16(1), pp. 53-65.
Sun, R., Eriksson, S. and Wang, L. (2014) 'Down-regulation of mitochondrial thymidine kinase 2 and deoxyguanosine kinase by didanosine: implication for mitochondrial toxicities of anti-HIV nucleoside analogs', *Biochem Biophys Res Commun*, Jul 25(450(2), pp. 1021-6.
Swerdlow, R. H. and Khan, S. M. 'A "mitochondrial cascade hypothesis" for sporadic Alzheimer's disease', *Medical Hypotheses*, 63(1), pp. 8-20.
Swerdlow, R. H. and Khan, S. M. (2004) 'A "mitochondrial cascade hypothesis" for sporadic Alzheimer's disease', *Medical Hypotheses*, 63(1), pp. 8-20.
Tang, S. and Huang, T. (2010) 'Characterization of mitochondrial DNA heteroplasmy using a parallel sequencing system', *Biotechniques*, 48(4), pp. 287-96.
Taylor, R. W., Barron, M. J., Borthwick, G. M., Gospel, A., Chinnery, P. F., Samuels, D. C., Taylor, G. A., Plusa, S. M., Needham, S. J., Greaves, L. C., Kirkwood, T. B. and Turnbull, D. M. (2003) 'Mitochondrial DNA mutations in human colonic crypt stem cells', *J Clin Invest*, 112(9), pp. 1351-60.
Taylor, R. W. and Turnbull, D. M. (2005) 'Mitochondrial DNA mutations in human disease', *Nat Rev Genet*, 6(5), pp. 389-402.
Torroni, A. and Wallace, D. C. (1994) 'Mitochondrial DNA variation in human populations and implications for detection of mitochondrial DNA mutations of pathological significance', *J Bioenerg Biomembr*, 26(3), pp. 261-71.
Trifunovic, A., Wredenberg, A., Falkenberg, M., Spelbrink, J. N., Rovio, A. T., Bruder, C. E., Bohlooly, Y. M., Gidlof, S., Oldfors, A., Wibom, R., Tornell, J., Jacobs, H. T. and Larsson, N. G. (2004) 'Premature ageing in mice expressing defective mitochondrial DNA polymerase', *Nature*, 429(6990), pp. 417-23.
Tsuboi, M., Morita, H., Nozaki, Y., Akama, K., Ueda, T., Ito, K., Nierhaus, K. H. and Takeuchi, N. (2009) 'EF-G2mt is an exclusive recycling factor in mammalian mitochondrial protein synthesis', *Mol Cell*, 35(4), pp. 502-10.
Ulahannan, D., Kovac, M. B., Mulholland, P. J., Cazier, J. B. and Tomlinson, I. (2013) 'Technical and implementation issues in using next-generation sequencing of cancers in clinical practice', *Br J Cancer*, 109(4), pp. 827-35.
van der Bliek, A., Redelmeier, T., Damke, H., Tisdale, E., Meyerowitz, E. and Schmid, S. (1993) 'Mutations in human dynamin block an intermediate stage in coated vesicle formation', *The Journal of Cell Biology*, 122(3), pp. 553-563.
van der Bliek, A. M., Shen, Q. and Kawajiri, S. (2013) 'Mechanisms of Mitochondrial Fission and Fusion', *Cold Spring Harbor Perspectives in Biology*, 5(6).
van Oosterhout, J. J., Mallewa, J., Kaunda, S., Chagoma, N., Njalale, Y., Kampira, E., Mukaka, M. and Heyderman, R. S. (2012) 'Stavudine Toxicity in Adult Longer-Term ART Patients in Blantyre, Malawi', *PLoS ONE*, 7(7), p. e42029.
Vermulst, M., Bielas, J. H., Kujoth, G. C., Ladiges, W. C., Rabinovitch, P. S., Prolla, T. A. and Loeb, L. A. (2007) 'Mitochondrial point mutations do not limit the natural lifespan of mice', *Nat Genet*, 39(4), pp. 540-3.

Virbasius, J. V. and Scarpulla, R. C. (1994) 'Activation of the human mitochondrial transcription factor A gene by nuclear respiratory factors: a potential regulatory link between nuclear and mitochondrial gene expression in organelle biogenesis', *Proc Natl Acad Sci U S A*, 91(4), pp. 1309-13.

Wai, T., Teoli, D. and Shoubridge, E. A. (2008) 'The mitochondrial DNA genetic bottleneck results from replication of a subpopulation of genomes', *Nat Genet*, 40(12), pp. 1484-8.

Walker, D., Poirier, M., Campen, M., Cook, D., Jr., Divi, R., Nagashima, K., Lund, A., Cossey, P., Hahn, F. and Walker, V. (2004) 'Persistence of mitochondrial toxicity in hearts of female B6C3F1 mice exposed in utero to 3′ -azido-3′ -deoxythymidine', *Cardiovascular Toxicology*, 4(2), pp. 133-153.

Walker, U. A., Auclair, M., Lebrecht, D., Kornprobst, M., Capeau, J. and Caron, M. (2006) 'Uridine abrogates the adverse effects of antiretroviral pyrimidine analogues on adipose cell functions', *Antivir Ther*, 11(1), pp. 25-34.

Walker, U. A., Bickel, M., Lutke Volksbeck, S. I., Ketelsen, U. P., Schofer, H., Setzer, B., Venhoff, N., Rickerts, V. and Staszewski, S. (2002) 'Evidence of nucleoside analogue reverse transcriptase inhibitor--associated genetic and structural defects of mitochondria in adipose tissue of HIV-infected patients', *J Acquir Immune Defic Syndr*, 29(2), pp. 117-21.

Wallace, D., Singh, G., Lott, M., Hodge, J., Schurr, T., Lezza, A., Elsas, L. and Nikoskelainen, E. (1988) 'Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy', *Science*, 242(4884), pp. 1427-1430.

Wallace, D. C. (1992) 'Mitochondrial genetics: a paradigm for aging and degenerative diseases?', *Science*, 256(5057), pp. 628-32.

Wallace, D. C. (2001) 'A mitochondrial paradigm for degenerative diseases and ageing.', *Novartis Found. Symp*, 235, pp. 247-263.

Wallace, D. C. and Chalkia, D. (2013) 'Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease', *Cold Spring Harb Perspect Med*, 3(10), p. a021220.

Wang, B., Sun, J., Ma, Y., Wu, G., Tian, Y., Shi, Y. and Le, G. (2014a) 'Resveratrol Preserves Mitochondrial Function, Stimulates Mitochondrial Biogenesis, and Attenuates Oxidative Stress in Regulatory T Cells of Mice Fed a High-Fat Diet', *J Food Sci*.

Wang, H., Lemire, B. D., Cass, C. E., Weiner, J. H., Michalak, M., Penn, A. M. and Fliegel, L. (1996) 'Zidovudine and dideoxynucleosides deplete wild-type mitochondrial DNA levels and increase deleted mitochondrial DNA levels in cultured Kearns-Sayre syndrome fibroblasts', *Biochim Biophys Acta*, 1316(1), pp. 51-9.

Wang, J., Xiong, S., Xie, C., Markesbery, W. R. and Lovell, M. A. (2005) 'Increased oxidative damage in nuclear and mitochondrial DNA in Alzheimer's disease', *J Neurochem*, 93(4), pp. 953-62.

Wang, K., Li, M. and Hakonarson, H. (2010) 'ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data', *Nucleic Acids Res*, 38(16), p. e164.

Wang, P., Liu, Y. L., Han, L., Zhao, F. L., Guo, F., Wang, X. A. and Lv, Y. M. (2013) 'Mitochondria DNA 4977 bp common deletion in peripheral whole blood from healthy donors', *Biomed Environ Sci*, 26(12), pp. 990-3.

Wang, R., Liu, Y. Y., Liu, X. Y., Jia, S. W., Zhao, J., Cui, D. and Wang, L. (2014b) 'Resveratrol protects neurons and the myocardium by reducing oxidative stress and ameliorating mitochondria damage in a cerebral ischemia rat model', *Cell Physiol Biochem*, 34(3), pp. 854-64.

Wang, X., Chai, H., Yao, Q. and Chen, C. (2007) 'Molecular Mechanisms of HIV Protease Inhibitor-Induced Endothelial Dysfunction', *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 44(5), pp. 493-499 10.1097/QAI.0b013e3180322542.

Wang, X. and Moraes, C. T. (2011) 'Increases in mitochondrial biogenesis impair carcinogenesis at multiple levels', *Mol Oncol*, 5(5), pp. 399-409.

Wang, Y. and Bogenhagen, D. F. (2006) 'Human mitochondrial DNA nucleoids are linked to protein folding machinery and metabolic enzymes at the mitochondrial inner membrane', *J Biol Chem*, 281(35), pp. 25791-802.

Wendelsdorf, K. V., Song, Z., Cao, Y. and Samuels, D. C. (2009) 'An analysis of enzyme kinetics data for mitochondrial DNA strand termination by nucleoside reverse transcription inhibitors', *PLoS Comput Biol*, 5(1), p. e1000261.

White, A. J. (2001) 'Mitochondrial toxicity and HIV therapy', *Sexually Transmitted Infections*, 77(3), pp. 158-173.

Wong, J. K., Hezareh, M., Günthard, H. F., Havlir, D. V., Ignacio, C. C., Spina, C. A. and Richman, D. D. (1997) 'Recovery of Replication-Competent HIV Despite Prolonged Suppression of Plasma Viremia', *Science*, 278(5341), pp. 1291-1295.

Wong, L.-J. C., Naviaux, R. K., Brunetti-Pierri, N., Zhang, Q., Schmitt, E. S., Truong, C., Milone, M., Cohen, B. H., Wical, B., Ganesh, J., Basinger, A. A., Burton, B. K., Swoboda, K., Gilbert, D. L., Vanderver, A., Saneto, R. P., Maranda, B., Arnold, G., Abdenur, J. E., Waters, P. J. and Copeland, W. C. (2008) 'Molecular and clinical genetics of mitochondrial diseases due to POLG mutations', *Human Mutation*, 29(9), pp. E150-E172.

Wu, Z., Puigserver, P., Andersson, U., Zhang, C., Adelmant, G., Mootha, V., Troy, A., Cinti, S., Lowell, B., Scarpulla, R. C. and Spiegelman, B. M. (1999) 'Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1', *Cell*, 98(1), pp. 115-24.

Yamanaka, H., Gatanaga, H., Kosalaraksa, P., Matsuoka-Aizawa, S., Takahashi, T., Kimura, S. and Oka, S. (2007) 'Novel mutation of human DNA polymerase gamma associated with mitochondrial toxicity induced by anti-HIV treatment', *J Infect Dis*, 195(10), pp. 1419-25.

Yang, M. Y., Bowmaker, M., Reyes, A., Vergani, L., Angeli, P., Gringeri, E., Jacobs, H. T. and Holt, I. J. (2002) 'Biased incorporation of ribonucleotides on the mitochondrial L-strand accounts for apparent strand-asymmetric DNA replication', *Cell*, 111(4), pp. 495-505.

Ye, F., Samuels, D. C., Clark, T. and Guo, Y. (2014) 'High-throughput sequencing in mitochondrial DNA research', *Mitochondrion*.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S. and Madden, T. L. (2012) 'Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction', *BMC Bioinformatics*, 13, p. 134.

Yoon, Y., Krueger, E. W., Oswald, B. J. and McNiven, M. A. (2003) 'The Mitochondrial Protein hFis1 Regulates Mitochondrial Fission in Mammalian Cells through an Interaction with the Dynamin-Like Protein DLP1', *Molecular and Cellular Biology*, 23(15), pp. 5409-5420.

Yu-Wai-Man, P., Griffiths, P. G., Hudson, G. and Chinnery, P. F. (2009) 'Inherited mitochondrial optic neuropathies', *J Med Genet*, 46(3), pp. 145-58.

Yu-Wai-Man, P., Lai-Cheong, J., Borthwick, G. M., He, L., Taylor, G. A., Greaves, L. C., Taylor, R. W., Griffiths, P. G. and Turnbull, D. M. (2010) 'Somatic mitochondrial DNA deletions accumulate to high levels in aging human extraocular muscles', *Invest Ophthalmol Vis Sci*, 51(7), pp. 3347-53.

Zhang, Y., Song, F., Gao, Z., Ding, W., Qiao, L., Yang, S., Chen, X., Jin, R. and Chen, D. (2014) 'Long-term exposure of mice to nucleoside analogues disrupts mitochondrial DNA maintenance in cortical neurons', *PLoS One*, 9(1), p. e85637.

Zhang, Y. F. (2007) '[Age-dependent mitochondrial DNA 4977bp depletion in human skeletal muscle]', *Fa Yi Xue Za Zhi*, 23(6), pp. 438-40.

Zhou, L., Kitch, D. W., Evans, S. R., Hauer, P., Raman, S., Ebenezer, G. J., Gerschenson, M., Marra, C. M., Valcour, V., Diaz-Arrastia, R., Goodkin, K., Millar, L., Shriver, S., Asmuth, D. M., Clifford, D. B., Simpson, D. M., McArthur, J. C., Narc and Group, A. A. S. (2007) 'Correlates of epidermal nerve fiber densities in HIV-associated distal sensory polyneuropathy', *Neurology*, 68(24), pp. 2113-2119.