UNIVERSITY OF CAMBRIDGE

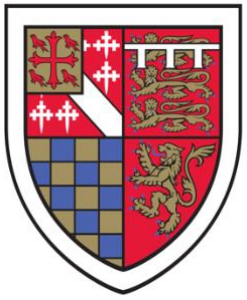
Mechanisms Controlling the Segregation of Mitochondrial DNA Heteroplasmy

Angelos Glynos

This dissertation is submitted for the degree of Doctor of Philosophy
September 2022

MRC Mitochondrial Biology Unit

Mitochondrial Genomics Laboratory



St Edmund's College

Declaration

This dissertation contains results that are the product of my own work, work that was performed collaboratively has not been included except when otherwise specified within the text and figure legends. The entirety of this body of work was carried out at the Medical Research Council (MRC) Mitochondrial Biology Unit (MBU) under the supervision of Prof. Patrick Chinnery between October 2019 and September 2022. The work presented in this thesis has not been submitted, in whole or in part, for a degree at this or any other institution and the length of it does not exceed 60.000 words, excluding figures, tables, appendices and bibliography.

Signed: [Signature redacted]

Date: 2/2/2023

Angelos Glynos, BSc, Mres,

PhD Student, MRC Mitochondrial Biology Unit, St Edmund's College, University of Cambridge

Abstract

Mechanisms controlling the segregation of mitochondrial DNA heteroplasmy
Angelos Glynos

Mutations of the mitochondrial DNA (mtDNA) are often the cause behind primary mitochondrial disorders affecting ~1:5000 individuals. However, the full extent of the impact that mtDNA mutations have is yet to be comprehensively understood. One of the main reasons behind our slow progress in the field is the multi-copied nature of mtDNA, which suggests that even healthy individuals will carry a small percentage of mutated mtDNA molecules alongside healthy ones, in a state termed heteroplasmy. In cases where the proportion of mutant to healthy mtDNA molecules reaches a critical threshold, diverse and multisystem pathological phenotypes begin to appear. While an individual's mtDNA heteroplasmy level is largely dependent on that of his maternal germline, studies have shown that there are diverse forces, both intra and extracellular in nature that drive segregation. Further complicating this phenomenon, the observed driving forces appear to be mutation- and cell type-specific in their effect.

In this dissertation I first describe my work on optimising and validating a protocol that allows us to measure single cell heteroplasmy. Developing this in-house technique, enabled us to perform high-throughput analyses of cell populations of interest while revealing for the first time the intricacies governing single mtDNA heteroplasmy variability at the single cell level. With this protocol in place, I set out to study the heteroplasmy of mouse brain- and spleen-derived populations. In this endeavour, I made use of two novel mouse models that carry a mutation on mitochondrial-tRNA Alanine (*mt-Ta*), m.5019A>G and m.5024C>T. Recording single cell heteroplasmy values at different timepoints throughout development, we observed that both mutations followed the principles of random genetic drift. The rate of drift exhibited mutation-specific patterns.

Moreover, I present a collaborative project geared towards uncovering the impact the two *mt-Ta* mutations have at the level of the transcriptome on difference cell lineages belonging to E8.5 mouse embryos. I describe the identification of 17 distinct cell lineages and their inherent variability in mtDNA transcript abundance. While no developmental disparities were observed in mutant embryos compared to controls, we did detect an upregulation of mtDNA transcripts in response to the mutation. At the same time, genes that were previously defined as epistatic suppressors/buffers were found to be downregulated. Pseudobulk analysis revealed differential expression of genes both at the level of the organism and that of the cell-lineage. Overall, mice carrying the m.5024C>T mutation seem to mount a greater compensatory transcriptional response compared to their m.5019A>G counterparts.

Finally, I explore the relationship between mtDNA heteroplasmy, copy number and the cell cycle. More specifically, making use of a fluorescent cell cycle reporter, I examine mtDNA changes along the cell cycle. Having established a consistent pattern, I assess the impact of genetic manipulation of mtDNA copy number and restriction of glycolysis on cell cycle progression. Finally, I delve into the consequences of large scale mtDNA deletions on the cell's respiratory capacity and examine whether that defect impacts their ability to complete the cell cycle.

Research Impact Statement

Student name	Angelos Glynos
USN	304336908
Department	MRC Mitochondrial Biology Unit
Supervisor	Prof. Patrick Chinnery
Thesis title	Mechanisms controlling the segregation of mitochondrial DNA heteroplasmy

Research Impact Statement:

During my first year as a PhD student at the MRC Mitochondrial Biology Unit I was a member of the lab of Prof. Massimo Zeviani. Following his departure, I joined the lab of Prof. Patrick Chinnery in October 2019, starting to work on a new project. The entirety of this thesis has been the product of my work as a member of Prof. Chinnery's lab. Consequently, the Covid-19 pandemic had a strong impact on my progress, as I was only four and a half months into my new project when the lockdowns begun.

Being away from the UK for around four months and working part-time for half a year as a result of enforced restrictions in personnel capacity, had a significant impact on the progress of projects presented in **Results Section, Chapters 3**, **4** and **6**. While we were lucky enough to have submitted the first batch of E8.5 mouse embryos for single cell RNA sequencing just before the pandemic begun, all other experiments involving mouse work and tissue culture had to be delayed.

Even though preliminary data that had been collected in the first few months were encouraging, any further experiments had to be postponed. Moreover, the timing of the lockdowns so early in my new at the time project, meant that there was not enough data for a long-term analysis nor any plans for a paper.

With the entirety of my work at the time having been halted, I decided to take advantage of the time to reflect on the progress we had made so far and, encouraged by Prof. Chinnery, planned ahead on our approach to the projects once our department re-opened. In retrospect, our discussions over the lockdown period as well as the time I spent reading on relevant topics helped at better directing our post-pandemic work.

Ultimately, the impact of the Covid-19 lockdowns brough about a significant delay in my work, which was in its entirety based on wet-lab based experiments. These delays came at a pivotal point in my PhD journey as they compounded on my transition into a new lab and project.

Student's signature *[by typing your name, you are providing your electronic signature]*	Angelos Glynos
Supervisor signature *[by typing your name, you are providing your electronic signature]*	Patrick Chinnery [Signature redacted]

Acknowledgements

First and foremost, I would like to thank my PI, Prof Patrick Chinnery who took me in as part of his lab during a period of turmoil and uncertainty in my journey as a PhD student. His continuous support and encouragement made every meeting something to look forward to. The years I spent as part of his lab are an experience I will cherish for the rest of my life. I would also like to give special thanks to Stephen Burr who I've had the privilege to work with from day one. He has been a great friend and mentor who always showed patience and kindness when introducing me to new concepts and techniques. The same goes for Michele Frison, a friend and colleague who has been instrumental in my PhD journey, as he was the first person I approached when thinking about joining Patrick's lab. Since then, his advice has always been helpful, as has his unending positivity and love for science which is a driving force keeping me and the rest of the lab going even through the toughest of times.

I would like to extend my gratitude all members of the Mitochondrial Genomics group, Zoe, Neil(the snail) and Haixin have been great colleagues, collaborators but most importantly, friends. Of course, our lab would not be complete without our lovely bioinformatics team! I would like to especially thank Malwina and Florian, for their amazing guidance and support as collaborators, providing their valuable insight from a data-analysis standpoint. My collaboration with our one and only (genius) biostatistician, Lyuba, has been an amazing experience as she keeps trying to introduce me to new concepts both in and out of science. To Camila, Mario and Brandon, I wish the best of luck with their work and I hope they enjoy their time in our lab as much as I did!

Finally, I would like to thank Prof. Massimo Zeviani for giving me the opportunity to spend four years as a PhD student at the University of Cambridge. I am extremely grateful to all the friends I made as part of the MBU family, it was their encouragement and support that helped me persist during the roughest times. I would like to extend my appreciation to my friends in Greece and abroad, past and present, for the role they've played in my journey so far. I am certain that as we go through life, we will cherish our memories together more than any results or publications.

I dedicate this thesis to my parents, Evi and Dimitris, and siblings, Laskarina and Pavlo, to whom I am forever grateful for shaping me into the man I am today. I know we will always be there for each other, through thick and thin and I am certain that all our future goals and dreams will one day become reality. Just dream big and work hard!

Contents

Abbreviations

2.3-DCPE	2[[3-(2,3-dichlorophenoxy) propyl] amino] ethanol
A	Adenine
ADP	Adenine diphosphate
AIF	Apoptosis-inducing factor
AMPARs	α-amino-3-hydroxy-5- methyl-4-isox-azoleproprionic acid receptors
ARRIVE	Animal research: reporting of in vivo experiments
ARS	Aminoacyl t-RNA synthetase
ASAP-seq	Select antigen profiling by sequencing
ATAC-seq	Assay for transposase-accessible chromatin by sequencing
ATFS-1	Activated transcription factor 1
ATM	Ataxia telangiectasia mutated protein
ATP	Adenine triphosphate
ATR	Ataxia telangiectasia and Rad3-related protein
AWERB	Cambridge animal welfare ethical review body
BrdU	Bromodeoxyuridine
BSA	Bovine serum albumin
BVSC	Blimp1-mVenus and stella-ECFP
CACT	Carnitine-acylcarnitine translocase
CCD	Charge coupled device
CDKs	Cyclin-dependent kinases
Cenp-F	Cytoskeletal-associated protein F
CHK1/2	Checkpoint kinases 1/2
CITE-seq	Cellular indexing of transcriptomes and epitopes by sequencing
CL	Cardiolipin
CNS	Central nervous system
CoA	Coenzyme A
COUP-TFII	Chicken ovalbumin upstream promoter transcription factor II
CPEO	Chronic progressive external ophthalmoplegia
CRISPR	Clustered regularly interspaced short palindromic repeats
CSBs	Conserved sequence blocks
CVS	Chorionic villus sampling
Cytc	Cytochrome c
D-foci	Degradation foci
dATPaS	Deooxyadenosine alpha thio triphosphate
ddPCR	Digital droplet PCR
DEGs	Differentially expressed genes
DNTs	Deoxyribonucleotides
Drp1	Dynamin-related protein 1
DSBs	Double strand breaks
dsDNA	Double-stranded DNA

dUTP	Deoxyuridine triphosphate
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
ER	Endoplasmic reticulum
ESCs	Embryonic stem cells
ETC	Electron transport chain
FACS	Fluoresence-associated cell sorting
$FADH_2$	Flavin adenine dinucleotide H_2
FBS	Foetal bovine serum
FELASA	Federation of European laboratory animal science associations
FSC-A	Forward scatter-area
FUCCI	Fuorescent ubiquitination-based cell cycle indicator
G0/1/2-phase	Growth phase 0/1/2
GO	Gene ontology
GRFS1	G-rich sequence factor 1
GTP	Guanosine triphosphate
GTPases	Hydrolases of guanosine triphosphate
GTPBP3	GTP Binding Protein 3
H-strand	Heavy strand
HBSS	Hank's balanced salt solution
HMG	High motility group
HSP	Heavy-strand promoter
HU	Hydroxyurea
IBM	Inner boundary membrane
IMM	Inner mitochondrial membrane
IMS	Inter-membrane space
IMTs	Inhibitors of mitochondrial transcription
indels	Insertions deletions
INFs	Interferons
INK4	Inhibitors of CDK4
IQR	Interquartile range
IRGM	Immunity related GTPase M and
ISRmt	Integrated mitochondrial stress response
KO	Knock-out
KS-test	Kolmogorov–Smirnov test
KSS	Kearns–Sayre syndrome
L-strand	Light strand
LHON	Leber's hereditary optic neuropathy
LSP	Light-strand promoter
m7G	N7 methyl guanosine
MAVS	Mitochondrial antiviral-signalling proteins
MEFs	Mouse embryonic fibroblasts

MELAS	Mitochondrial myopathy, encephalopathy, lactic acidosis, stroke-like episodes
MFN1/2	Mitofusin 1/2
MIA	Mitochondria intermembrane space assembly
MICOS	Mitochondrial contact site and cristae organizing system
MOMP	Mitochondrial outer membrane permeabilization
MPP	Mitochondrial processing peptidase
MRGs	Mitochondrial RNA granules
MRPP	Mitochondrial ribonuclease P protein
mt-DNA	Mitochondrial DNA
mt-LSU	Mitochondrial large ribosomal subunit
mt-SSU	Mitochondrial small ribosomal subunit
MTERF1	Mitochondrial transcription termination factor 1
MTFMT	mitochondrial methionine tRNA
MTO1	mt-tRNA Translation Optimization 1
mtPAP	Mitochondrial poly-A polymerase
mtRNAs	Mitochondrial RNAs
MTS	Mitochondrial targeting sequence
mtSSBs	Mitochondrial single-stranded DNA-binding proteins
NADH	Nicotinamide adenine dinucleotide
NCR	Non-coding region
nDNA	Nuclear DNA
NF-κB	Nuclear factor κB
NGS	Next generation sequencing
NLRP3	NLR family pyrin domain containing 3
NLRX1	Leucine rich repeat containing X1
NMDARs	N-methyl-D-aspartic acid receptors
Nnt	Nicotinamide nucleotide transhydrogenase
NRTIs	Nucleoside reverse transcriptase inhibitors
NSUN3	NOP2/Sun RNA Methyltransferase 3
NTB	Nitrotetrazoleum Blue
NTP	Nucleoside triphosphate
NUMTs	Nuclear mitochondrial DNA
OCR	Oxygen consumption rate
OMM	Outer mitochondrial membrane
OriH	Origin of replication for the Heavy-strand
OriL	Origin of replication for the Heavy-strand
OXPHOS	Oxidative phosphorylation
p53	Tumour protein 53
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PD	Parkinson's disease

PGCs	Primordial germ cells
PI3KKs	Phosphatidylinositol 3-kinase (PI3K)-like protein kinases
PIP	PCNA-Interacting Protein
PCNA	Proliferrating cell nuclear antigen
PKlm	Peter Kyriakopoulos loves mitochondria
PMDs	Primary mitochondrial disorders
PMF	Proton motive force
POLRMT	Mitochondrial RNA polymerase
POLγ	Polymerase gamma
PPL	Project Licence
PSD95	Postsynaptic density protein 95
PTCD3	Pentatricopeptide repeat domain 3
PUS1	Pseudouridylate synthase 1
R	Restriction point
RB	Retinoblastoma
RBCs	Red blood cells
RBCs	Red blood cells
RIs	Replication intermediates
RITOLS	Ribonucleotide incorporation through the lagging strand
RLR	RIG-I-like receptor
RNaseP	Ribonuclease P
RNR	Ribonucleotide reductase
ROS	Reactive oxygen species
RRFs	Ragged Red Fibres
rRNA	Ribosomal RNA
RT	Room temperature
S-phase	Synthesis phase
SAM	Sorting and assembly machinery
SCENIC	Single-cell regulatory network inference and clustering
scRNA-seq	Single cell RNA sequencing
SMDs	Secondary mitochondrial disorders
SSC-A	Side scatter-area
ssDNA	Single-stranded DNA
STED	Stimulated emission depletion
TAS	Termination-associated sequence
TBE	Tris-borate-EDTA
TC	Tissue culture
TCA	Tricarboxylic acid
TEFM	Transcription elongation factor, mitochondrial
TFAM	Transcription factor A
TFB1/2M	Transcription factor B1/2, mitochondrial
Tims	Translocases of the inner mitochondrial membrane
TMRT	tRNA methyltransferase 5

TMT	Tandem mass tagging
TOM	Translocase of the outer-mitochondrial membrane
TRAF6	TNF receptor associated factor 6
TRAIL	TNF-related apoptosis inducing ligand
TRIT1	tRNA isopentenyltransferase
TRMU	tRNA 5-methylaminomethyl-2-thiouridylate methyltransferase
tRNA	Transfer RNA
TRT1	tRNA nucleotidyl transferase
UbQ	Ubiquitin
UMAP	Uniform manifold approximation and projection
UMI	Unique molecular identifier
UPR^{mt}	Mitochondrial unfolded protein response
UV	Ultraviolet
VDAC	Voltage-dependent anion channel
WT	Wild-type
$\Delta\psi_m$	Mitochondrial membrane potential

Chapter 1. Introduction

1.1. The Discovery of Mitochondria

As it is often the case, not one person can be credited with the discovery of mitochondria. The organelle was first observed in 1857 by the Swiss anatomist and physiologist, Albert von Kolliker, who described the presence of subcellular vesicles in the sarcoplasm of striated muscle (van der Giezen, 2011). Physicians at the time were able to corroborate his findings and even extend them, to include many more cell types within which these fascinating structures could be identified. In 1894, these structures were termed bioblasts, by the German histologist and cytologist Richard Altman, who was able to stain them using the unsaturated lipid dye, osium tetroxide (Altmann, 1984). He made a note of their abundance within the cytoplasm, describing "the protoplasm as a colony of bioblasts" and articulated their importance as an active unit within the cell: "Microorganisms and granula are at an equivalent level and represent elementary organisms, which are found wherever living forces are acting, thus we want to describe them by the common term bioblasts. In the bioblast, that morphological unit of living matter appears to be found" (Altmann, 1894).

However, it was not until 1898 when the German microbiologist Carl Benda, a pioneer in the application of light microscopy in the study of subcellular structures, coined the term "mitochondria" (Benda, 1898). The term originates from the Greek words "mitos" (thread) and "chondros" (granule), to describe a structure that resembles a series of beads on a string. Half a century passed from the time of Benda's observation until the role of mitochondria in energy production was fully appreciated (Benda, 1898). Today, we know that mitochondrial activities span far beyond generating energy for the cell, as they are implicated in multiple cellular processes, including cell death, calcium buffering, heat production, immune cell responses just to name a few. As a result, their indispensable role within the cell, combined with the intricacy of mitochondrial genetics, make dysfunctions plaguing this organelle complex, multisystem and especially hard to diagnose.

1.2. The Role of Mitochondria

Mitochondria are present in almost all eukaryotic cells and are primarily responsible for supplying them with the energy required to carry out all basic cellular activities (Vafai and Mootha 2012). This is achieved primarily through a metabolic pathway known as Oxidative Phosphorylation (OXPHOS) that involves the oxidation of nutrients in order to eventually release chemical energy in the form of Adenine Triphosphate (ATP) (Fernandez-Vizarra and Zeviani 2021). At the same time, mitochondria are highly dynamic organelles, that participate in the propagation of intracellular signalling pathways as well as the synthesis of a plethora of cellular components.

A characteristic that sets mitochondria apart from all other organelles is that they contain their own genome, distinct from that of the nucleus, known as mitochondrial DNA (mtDNA). The human mtDNA is circular in structure and 16,569 base pairs long, encoding just 37 genes, including 13 polypeptides, 22 transfer RNA (tRNA) and 2 ribosomal RNA (rRNA) molecules (Vafai and Mootha 2012). All polypeptides encoded in the mtDNA are components of electron transport chain (ETC) complexes. In turn, the ETC is comprised of 4 macromolecular complexes central to mitochondrial activity, the role of which is to act as catalysts for the transfer of electrons from reducing equivalents to molecular oxygen (Vafai and Mootha 2012). Mitochondrial tRNA and rRNA molecules make it possible for polypeptides encoded in the mtDNA to be translated within the confines of the organelle (Pearce et a., 2017). However, out of the ~1500 proteins found within the mitochondrion, only 13 are encoded by its DNA, leaving 99% of the mitoproteome to be encoded in the nucleus and subsequently transported onto the mitochondrial surface where import takes place (Pearce et a., 2017).

1.2.1. The Evolutionary Origins of Mitochondria

The Russian biologist and botanist Konstantin Sergeevich Mereschkowski was the first person to propose the role of bacteria as the progenitors of eukaryotic organelles in his 1910 work "The Theory of Two Plasms as the Basis of Symbiogenesis, a New Study of the Origins of Organisms" (Kowallik and Martin, 2021). For many years, the endosymbiotic theory was widely accepted as our sole answer to questions surrounding the origins of mitochondria. It postulates that the evolutionary origin of mitochondria can traced back to one defining event, which involved the phagocytotic uptake of oxygen-consuming alpha-Proteobacteria within the cytosol of amitochondriate eukaryotes (Grey et al., 1999). However, the bacteria were not digested, but instead, were incorporated within the eukaryote's cytosolic structure, establishing a mutually beneficial relationship in an environment where atmospheric oxygen levels were gradually increasing (Lane and Martin, 2010).

Phylogenetic studies postulate that prior to their uptake, the bacterial symbiont's genome was comprised of more than 630 unique genes (Gabaldon and Huynen, 2003). The metabolic aspect of this early symbiotic relationship between the two primitive unicellular organisms is still hard to assess (Embley and Martin, 2006). However, it is clear that through the millennia, much of the genome that once belonged to the bacterial symbiont now resides, as part of the chromosomes, within the nucleus of eukaryotic cells (Dolezal et al., 2006).

With that in mind, it is worth mentioning that another theory around the evolutionary origins of mitochondria was developed by William F. Martin and Miklós Müller in 1998. This new 'hydrogen theory' bases itself on the common origins between mitochondria and hydrogenosomes, anaerobic, ATP-producing mitochondria (Martin and Muller, 1998). It suggests that the original host cell resembled current methanogenic, hydrogen-dependent archaea (Embley and Martin, 2006), while claiming that amitochondriate eukaryotes never existed (Embley and Martin, 2006). According to Martin and Müller, the host cell was supplied with the required hydrogen by the endosymbiont and, in turn, provided a safe environment for the mitochondrial ancestor to reside (Martin and Muller, 1998).

Despite their evolutionary origins still being a hotly debated topic, it is unquestionable that the incorporation of the mitochondrial component within the cytosol was a huge evolutionary leap, allowing for compartmentalised and thus, more efficient and energy production.

1.2.2. Mitochondrial Structure

As we have touched upon already, mitochondria are widely believed to trace their origins back to aerobic alpha-Proteobacteria, and as a result, share multiple structural characteristics with their ancestors. Our understanding of their nature was only begun in the early 1950s, when mitochondria were one of the first organelles to be examined using the emergent technology of electron microscopy (EM) (Frey et al., 2002). At the time, both Fritiof S.Sjöstrand and George E. Palade noted that mitochondria are surrounded by more than one membrane, progressing our understanding of the organelle away from the two-dimensional "beads on a string" interpretation, while providing us with a three-dimensional structure for the very first time (Frey et al., 2002). However, their early observations were slightly contradicting. While Sjöstrand's model identified both an outer and an inner boundary membrane (IBM), it also made the case for the presence of further membranes that compartmentalised the matrix of the mitochondrion (Sjostrand, 1956). Palade's observations on the other hand, made note of two distinct membranes, with the inner membrane creating folding protrusions towards the centre of the organelle, that he named "cristae mitochondriales" (Palade, 1952).

With hindsight, we can judge that Palade's model was the more accurate of the two, as it is now evident that the mitochondria are surrounded but two lipid membranes, termed the Outer Mitochondrial Membrane (OMM) and the Inner Mitochondrial Membrane (IMM) (Protasoni and Zeviani, 2021). These two phospholipid layers separate the organelle into two segments, an inner one called the matrix and an outer one called the Intermembrane Space (IMS) (Protasoni and Zeviani, 2021). As Palade correctly postulated, the IMM forms long invaginations that protrude into the mitochondrial matrix that have since been coined as cristae (Palade, 1952) **(Figure 1.1.)**.

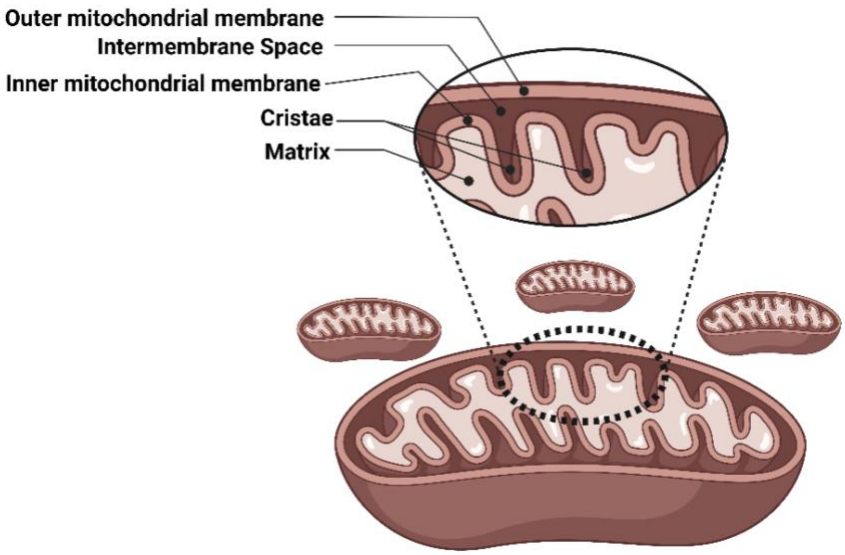


Figure 1.1. Structure of animal mitochondria.
Schematic representation of mitochondria and its most fundamental structural components: Outer mitochondrial membrane, Intermembrane space, Inner mitochondrial membrane, Cristae and Matrix. Created with BioRender.com.

In line with the endosymbiotic theory making the case for the bacterial origin of mitochondria, the IMM and the OMM differ greatly in terms of their shape, lipid composition and permeability as does the function of the transmembrane proteins that populate them (Protasoni and Zeviani, 2021). While the OMM resembles closely the architecture of membranes typically found in eukaryotes, the IMM bears all the hallmarks of bacterial membranes, exhibiting a high cardiolipin (CL) content (Cavalier-Smith, 2006). CL represents approximately 15-20% of the phospholipids composing the IMM and is widely considered a signature structural element of energy-transducing membranes due to its unique structure and conical shape which is thought of as the key behind its preference for negative curvature (Elias-Wolff, 2019).

In addition to their differences in phospholipid content, the two mitochondrial membranes can be differentiated on the basis of the transmembrane proteins they contain and, as a result, on their permeability (Shoshan-Barmatz et al., 2010). The OMM is heavily populated with voltage-dependent anion channels (VDAC) that facilitate the entry of small ions into the IMS (Shoshan-Barmatz et al., 2010). On the other hand, the IMM makes for a barrier permeable only to oxygen, carbon dioxide and water (Shoshan-Barmatz et al., 2010). At the same time, the lipid/protein ratio of the IMM is significantly higher than that of the OMM. Embedded within the cristae structures of the IMM, that serve to increase the surface area of the organelle's energy-transducing membrane, exist the protein complexes of the ETC (Schlame,

2021). The transition from the IBM portion, where the two membranes run parallel to each other, to the cristae is denoted by the presence of the mitochondrial contact site and cristae organizing system (MICOS) and is termed the cristae junctions (Palade, 1953) **(Figure 1.1.)**. With 99% of the mitoproteome having to be imported from the cytosol, it is important for the organelle to be equipped with a comprehensive machinery that will facilitate protein import and subsequent distribution to the appropriate compartments within the organelle. Proteins destined for import into the mitochondria can be categorized into two classes based on their targeting sequence that will eventually determine their final destination (Leuenberger et al., 1999). The first class, includes precursor proteins marked by a positively charged, amphipathic N-terminal sequence that is often cleavable, called a mitochondrial targeting sequence (MTS) (Vogtle et al., 2009). The average MTS ranges between 15 and 55 amino acids in length (Sickmann et al., 2003). The class of MTS-carrying preproteins includes members that are destined from the IMM, the IMS and the matrix (Dudek et al., 2013). The second class of precursor proteins targeted to the mitochondria do not include a cleavable N-terminal domain but, instead, possess an internal targeting sequence which results in their primary sequence being identical to that of the mature protein (Vogtle et al., 2009). Members of this second class of proteins include many IMS and IMM proteins, such as some Mitochondrial Carrier proteins, and all proteins embedded into the OMM (Dudek et al., 2013).

Due to the sheer number and complexity of the proteins that have to be imported and incorporated into the mitochondria, there is a series of tightly regulated steps that have to be followed. First, the hydrophobic segments of mitochondrial precursor proteins are shielded by dedicated cytoplasmic chaperones to prevent misfolding (Young et al., 2003). While proteins carrying an α-helical domain can be embedded in the OMM without any assistance, the translocases of the outer mitochondrial membrane (TOM) complex provide an initial docking site that also serves as a quality control checkpoint, at the site of entry (Dudek et al., 2013). Upon passing through the checkpoint provided by the TOM complex, proteins go through a downstream process of segregation in order to arrive at the appropriate sub-compartment within the organelle. Evidence stemming from studies in yeast, suggests β-barrel proteins are bound to small translocases of the inner mitochondrial membrane (TIMs) chaperones in order to be incorporated in the OMM (Sokol et al., 2014). In humans, the sorting and assembly machinery (SAM) is responsible for embedding proteins into the OMM (Sokol et al., 2014). Small TIMs are also responsible for guiding proteins of the mitochondrial metabolite carrier class through the IMS and onto the TIM22 complex, which is responsible for their

incorporation them into the IMM (Becker et al., 2012). Presequence containing proteins that are targeted for the IMS, pass through the TOM complex and bind onto the mitochondrial intermembrane space import assembly (MIA) machinery which catalyses folding into their mature, functional state (Dudek et al., 2013). Finally, proteins that are destined for the IMM or the matrix and contain an N-terminal presequence pass through TOM directly into TIM23 where, depending on the presence of additional import signals, are either embedded into the IMM or are released into the mitochondrial matrix (Dudek et al., 2013) **(Figure 1.2)**.

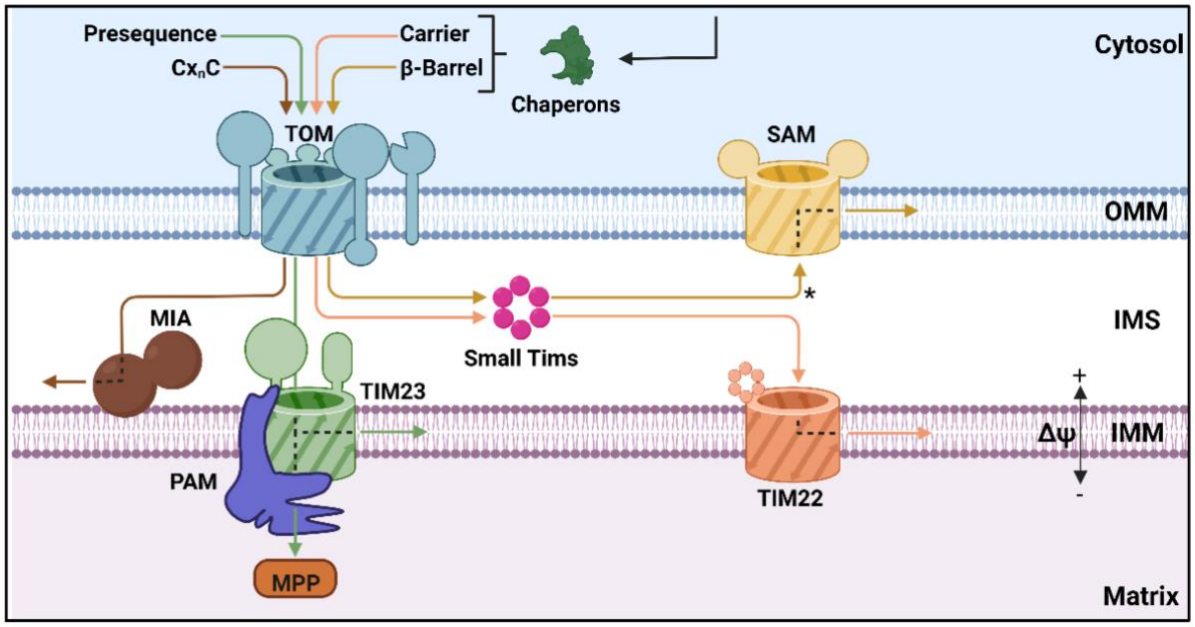


Figure 1.2. The mitochondrial protein import machinery.
Schematic representation of the pathways followed during mitochondrial protein import. Nuclear-encoded precursors of mitochondrial proteins undergo sorting and transportation to their final destination within the organelle based on the proteins targeting signals they possess. The first import step for all mitochondrial protein precursors is entry through the translocase of the outer mitochondrial membrane (TOM) complex. After passing through the TOM complex, they enter the intermembrane space (IMS) where downstream selection takes place depending on protein structure. In order for β-barrel proteins to be incorporated into the outer mitochondrial membrane (OMM), small translocases of the inner mitochondrial membrane (Tims) have to chaperone them to the sorting and assembly machinery (SAM). In the case of cysteine-rich IMS proteins that contain Cx_nC signals, their import is completed through the mitochondrial intermembrane space import and assembly (MIA) pathway. Proteins belonging to the mitochondrial carrier family are chaperoned to the translocase of the inner mitochondrial membrane 22 (TIM22) complex where they undergo incorporation into the inner mitochondrial membrane (IMM). Preproteins that need to be inserted within the IMM or translocated into the mitochondrial matrix, do so by interacting with the translocase of the inner mitochondrial membrane 23 (TIM23) complex. The presequence translocase-associated import motor (PAM) complex is responsible for the translocation of preproteins that interact with TIM23, into the mitochondrial matrix. Upon import of the preproteins into the matrix, mitochondrial processing peptidase (MPP) is responsible for the proteolytic cleavage of the presequence. Δψ, mitochondrial membrane potential; *, pathway with evidence in yeast but not human mitochondria. Created with BioRender.com.

Early observations of mitochondria could not paint a comprehensive picture of their dynamic nature. Modern live cell imaging techniques however, have revealed an interconnected tubular network that consists of highly motile organelles which constantly undergo regulated fission and fusion events in response to both internal and external stimuli (Tilokani et al., 2018). Mitochondrial fission events are characterized by the separation of an original organelle into two daughter mitochondria, while mitochondrial fusion is described as the merging of two mitochondrial into one daughter organelle. This continuous movement of organelles within the cytoplasm and balancing of the two opposing forces leads to constant changes in mitochondrial number as well as position within the cytoplasm (Liesa et al., 2009). The term "mitochondrial dynamics" is used to describe the totality of alternating states that mitochondria find themselves in (Liesa et al., 2009). The ability of mitochondria to exist in this very dynamic state comes down to a family of specialised dynamin-like guanine triphosphate hydrolases (GTPases) that drive the network towards either a fragmented or elongated state, depending amongst other factors, on the cell's energetic demands (Hoppins et al., 2007). A fused and elongated mitochondrial network serves to evenly distribute resources in the form of metabolites and mtDNA throughout the cell and is thought to primarily be employed as a defensive mechanism in times of stress (Chen et al., 2003). On the other hand, a fragmented mitochondrial network is correlated with mitochondrial dysfunction and cell death but it is vital when it comes to the selective clearance of damaged mitochondria through a specialised autophagic pathway known as mitophagy (Youle and Narendra, 2011). Despite years of work towards elucidating the role of components within the fission and fusion machinery, along with the genes that regulate them, their physiological role in human tissues such as liver and skeletal muscle has yet to be completely understood (Liesa et al., 2009).

1.2.3. Energy Production in Mitochondria

Adenosine triphosphate (ATP), discovered by Karl Lohmann in 1929, is an organic compound comprised of the nitrogenous base adenine, the sugar ribose and three phosphate molecules that serve as the foremost energy currency, driving biochemical reactions (Langen and Hucho, 2008). ATP serves this purpose when a water molecule is added, breaking off the bonds between the phosphate groups and producing adenosine diphosphate (ADP) and inorganic phosphate (Pi), a process known as hydrolysis (Weber and Senior, 2000). In order to put the importance of this molecule into perspective, one has to consider that each cell will generate

and consume 10,000,000 ATP molecules per second, while the body of an adult human requires 100-150 moles of ATP to be hydrolysed per day in order to maintain its function (Dunn and Grider, 2022). Naturally, providing a constant supply of ATP to our cells is of outmost importance and while there are other biochemical reactions that result in the generation of ATP, OXPHOS is recognised as the most efficient amongst them.

Consequently, it does not come as a surprise that mitochondrial became widely known as "the powerhouse of the cell," a term first coined by American cell biologist Philip Siekevitz in 1957 to describe the role of mitochondria as the epicentre of energy production within the cell (Siekevitz, 1957). Indeed, pioneering work done in the 1950s, 1960s and 1970s established mitochondria as a hub, serving the cell's bioenergetics needs (Pagliarini and Rutter 2013).

Cellular respiration involves a series of catabolic reactions, during the course of which, nutrients in the form of sugars, carbohydrates, proteins and fats are broken down by the cell to produce ATP (Gnaiger et al., 1995). This process can be separated into three distinct steps: glycolysis, the citric acid or tricarboxylic acid (TCA) cycle and OXPHOS (Gnaiger et al., 1995). Glycolysis takes place in the cytoplasm and involves the oxidation of one glucose molecule into two pyruvates, accompanied by the release of water, nicotinamide adenine dinucleotide (NADH) and ATP. Glycolysis itself can be further subdivided into the investment step where ATP is consumed in order to provide the energy required to initiate the breakdown of glucose, and the yield step wherein double the amount of ATP is produced than it was originally consumed (Fothergill-Gilmore and Michels, 1993). While glycolysis takes place in both fermenting and glycolytic organisms, the TCA cycle, can only be performed by respiring cells and takes place in the mitochondrial matrix of eukaryotic cells, markedly increasing the ATP output of the cell (Martinez-Reyes and Chandel, 2020). Pyruvate generated by the breakdown of glucose during glycolysis is transported into the mitochondria, where NADH and acetyl-coenzyme A (acetyl-CoA) are produced, as a result of the addition of acetyl groups onto CoA (Shi and Tu, 2015). When acetyl-CoA enters the TCA cycle, it undergoes a series of 9 reactions catalysed by 8 enzymes to produce three NADH molecules, one flavin adenine dinucleotide ($FADH_2$) molecule and one guanine triphosphate (GTP) molecule (Shi and Tu, 2015). During this process two carbon atoms are lost in the form of carbon dioxide (CO_2) as citrate is eventually converted into oxaloacetate (Martinez-Reyes and Chandel, 2020). Overall, each NADH and $FADH_2$ molecule generated in the TCA cycle corresponds to 2.5 and 1.5 molecules of ATP respectively, produced during OXPHOS (Martinez-Reyes and Chandel,

2020). The cycle can restart as acetyl synthase attaches an acetyl group onto oxaloacetate, regenerating citrate (Martinez-Reyes and Chandel, 2020) **(Figure 1.3.)**.

The final step of cellular respiration, where OXPHOS takes place, involves the ETC, the destination where the electrons being carried by the intermediates NADH and $FADH_2$ get deposited (Duchen, 2004). OXPHOS takes place within the mitochondrial matrix, through protein complexes that are embedded into the IMM and is is the cornerstone of cellular aerobic respiration (Duchen, 2004). OXPHOS is the reason why aerobes uptake oxygen, and the pathway through which the calories that we consume get oxidized (Duchen, 2004). As a result, it is responsible for ~90% of energy production within the average cell (Bergman and Ben-Shachar, 2016). During the course of OXPHOS, electrons in the form of reducing equivalents, derived from metabolised food, get carried through the complexes of the ETC which shift from reduced to oxidized states as the electrons get passed along (Martinez-Reyes and Chandel, 2020). The final electron acceptor is oxygen which gets reduced to water (Martinez-Reyes and Chandel, 2020). This process results in the generation of a proton gradient as protons (H^+) residing in the mitochondrial matrix get pumped into the intermembrane space by ETC complexes I, III and IV (Walker, 2013). The resulting electrochemical gradient drives the activity of ATP synthase (Complex V) following the principles of the chemiosmotic theory, first postulated by Peter Mitchell in 1961 (Mitchell, 1961). Under physiological conditions, the ATP synthase enzyme catalyses the formation of 1 ATP molecule from ADP + Pi for every 3 H^+ that move from the IMS, back into the mitochondrial matrix (Walker, 2013). The overall reaction is as follows:

$$ADP + Pi + {H^+}_{out} \rightleftharpoons ATP + H_2O + 3{H^+}_{in}$$

The entry point of NADH into the ETC is NADH : ubiquinone oxidoreductase, also known as respiratory Complex I. Its role is to oxidize incoming NADH and transfer two electrons, in order to reduce ubiquinone (UbQ), while at the same time pumping four protons into the IMS (Hirst, 2013). Respiratory Complex II or succinate dehydrogenase also plays a role in catalysing the conversion of succinate into fumarate in the TCA cycles and as a result, can be considered as the link between the two later steps of cellular respiration (Bezawork-Geleta et al., 2017). While being part of the ETC, Complex II, is entirely encoded within the nuclear DNA (nDNA), and is responsible for the transfer of electrons from succinate to UbQ (Cecchini, 2003). At the same time, it is the only complex of the ETC that does not pump protons into or out of the IMS (Cecchini, 2003). Respiratory Complex III, or coenzyme Q : cytochrome *c* –

oxidoreductase, is responsible for the oxidation of ubiquinol (the reduced form of UbQ), previously reduced from either Complexes I or II, and the transfer of electrons that bring about the reduction of cytochrome c (Cytc) (Zhang et al., 1998). During this process, two protons get pumped from the mitochondrial matrix into the IMS per molecule of ubiquinol being oxidised (Hinkle et al., 1991). The electrons carried by the reduced form of Cytc get accepted by the Respiratory Complex IV, also known as cytochrome c oxidase. Complex IV collects electrons from four reduced Cytc molecules and delivers them to one molecule of oxygen (O_2) and four protons (H^+), resulting in the creation of two water molecules (H_2O) and the release of great amounts of energy without necessitating the breaking of any organic compound bonds (Kaila et al., 2011). Simultaneously, Complex IV is able to utilize the energy generated in order to pump four protons from the matrix into the IMS, further increasing the proton electrochemical potential that exists across the IMM (Kaila et al., 2011). The last component of the respiratory chain is the ATP Synthase, a multiprotein complex responsible for utilising the protonmotive force built up across the IMM, as a result of the pumping of H^+ from the mitochondrial matrix into the IMS (Walker, 2013). The ATP Synthase employs a rotary mechanism of action that is powered by the backflow of protons from the IMS, where they have been actively pumped, into the mitochondrial matrix, leading to the synthesis of ATP from ADP and Pi (Walker, 2013) **(Figure 1.3.)**.

As we have seen so far, the overwhelming majority of ATP produced by the process of cellular respiration resides within the mitochondrial matrix. However, in order for the cell's energy currency to be evenly distributed throughout the cytoplasm, it is vital that it can be transported through the impermeable IMM and eventually out of the mitochondria. One of the most abundant proteins studding the IMM is the ADP/ATP carrier, whose role is to carry spent fuel in the form of ADP from all over the cell into mitochondria while, at the same time, exchanging it for newly synthesized ATP (Kunji et al., 2016).

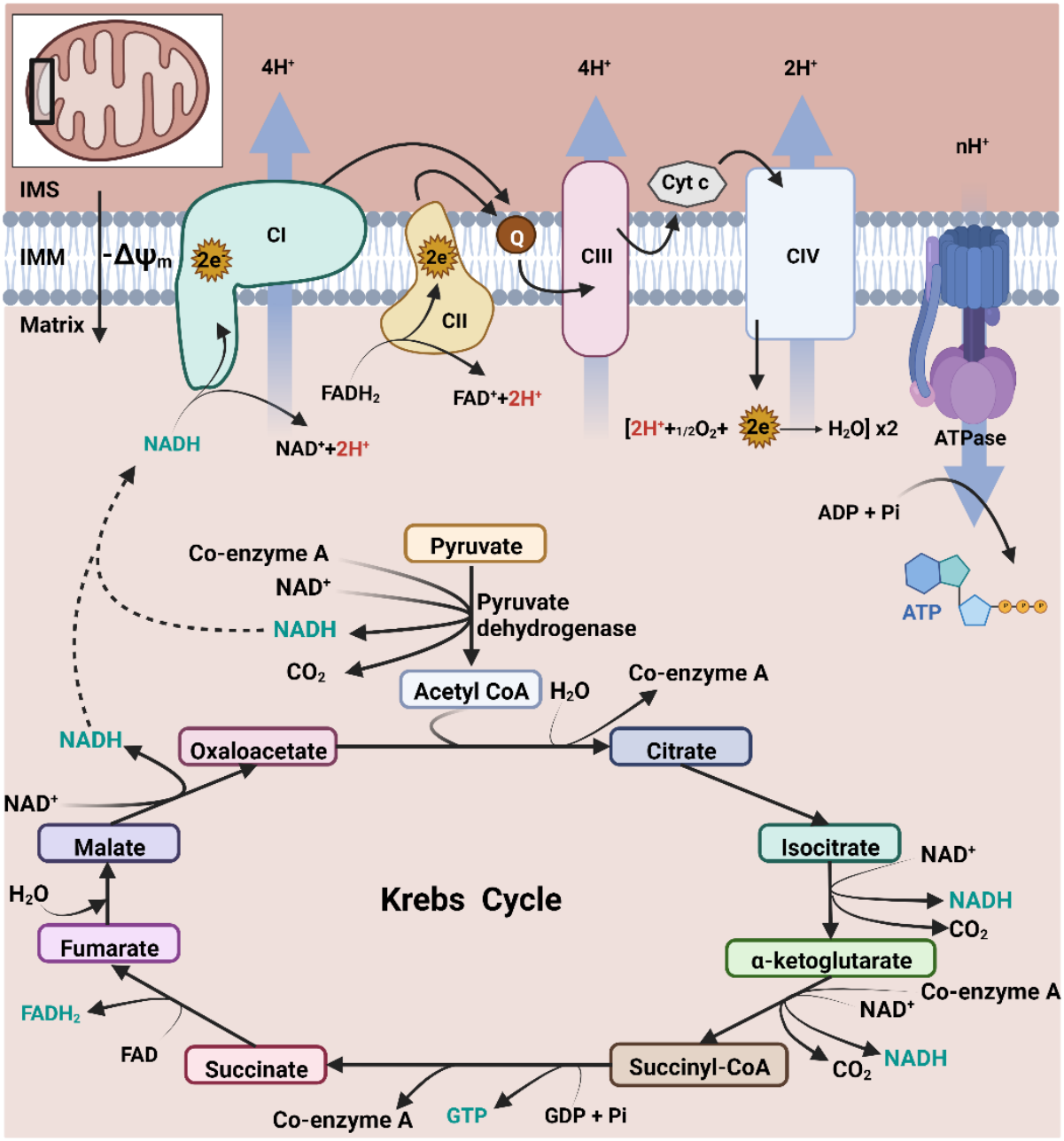


Figure 1.3. Krebs cycle and the electron transport chain.
Schematic representation of the bioenergetic reactions that are part of the Krebs cycle and the electron transport chain and result in the generation of energy in the form of adenine triphosphate (ATP) through the catalytic activity of ATP synthase. The generation of coenzyme electron carriers, nicotinamide adenine dinucleotide (NADH) and flavin adenine dinucleotide ($FADH_2$) by the Krebs cycle fuels the electron transport chain (ETC) with the required electrons. While NADH fuels the ETC through its interaction with CI which catalyses its oxidation into NAD^+, $FADH_2$ gets oxidized by CII into $FADH^+$. The transfer of electrons down the ETC drives the active pumping of protons (in the form of H^+) from the mitochondrial matrix into the intermembrane space (IMS). Finally, the proton motive force (PMF) is harvested by ATP synthase in order to catalyse the conversion of ADP and Pi into ATP, the energy currency of the cell. IMM, Inner Mitochondrial Membrane; OMM, Outer Mitochondrial Membrane; IMS, Intermembrane Space; Δψ, mitochondrial membrane potential. Created with BioRender.com.

Fatty Acid β-Oxidation (FAO) is another form of energy production taking place in mitochondria, this time, involving the breakdown of fats present on our bodies. In order for fatty acids to be imported into the mitochondrial matrix, they have to be sequentially converted into fatty acyl-CoA in the cytosol and acylcarnitine in the IMS (Rinaldo and Matern, 2002). Acylcarnitine is eventually imported into the mitochondrial matrix by carnitine-acylcanrnitine translocase (CACT) (Pande et al., 1993). Once inside the matrix, acylcarnitine reverts back into acyl-CoA and carnitine (Rinaldo and Matern, 2002). At this stage, acyl-CoA is able to undergo four steps of β-oxidation, producing NADH, $FADH_2$ and acetyl-CoA (Rinaldo and Matern, 2002). The later, enters the TCA cycle to further feed electrons into the ETC or goes on to be involved in ketogenesis or steroidogenesis (Rinaldo and Matern, 2002).

1.2.4. Mitochondria, Beyond Energy Production

While energy production was discovered first and is deservedly thought of as being central to the function of mitochondria, in recent years the organelle has been implicated in a plethora of diverse cellular processes, including redox signalling, programmed cell death, innate immunity, autophagy and calcium homeostasis and stem cell reprogramming, to name only but a few (Kamer and Mootha, 2015; Nikoletopoulou et al., 2013; Rambold and Pearce, 2018) **(Figure 1.4.)**.

Reactive oxygen species (ROS) is a broad term encompassing all highly reactive chemicals stemming from molecular oxygen (O_2) in aerobic organisms (Sies and Jones, 2020). When discussing physiologically relevant ROS, two redox signalling agents immediately come to mind, the superoxide anion radical ($O_2^{\cdot -}$) and hydrogen peroxide (H_2O_2) (Sies and Jones, 2020). Of the two, H_2O_2 is orders of magnitude more abundant in the cell (Sies and Jones, 2020) but, it is $O_2^{\cdot -}$ that can be characterized as ROS central to mitochondria, playing a role in retrograde redox signalling, while also being the underlying cause of oxidative damage in many pathophysiological conditions (Murphy, 2009). $O_2^{\cdot -}$**,** recognised as the proximal mitochondrial ROS, is generated at the matrix of the organelle as a result of the reduction of O_2 by one electron (Murphy, 2009). Sites on respiratory complexes I, II and III are responsible for the vast majority of ROS production within the mitochondrial matrix (Martinez-Cayuela, 1995). The contributions of mitochondrial ROS to disease have been studied in great detail and encompass damage to biomolecules spanning from proteins to lipid membranes and DNA (Halliwell and

Cross, 1994). At the same time, the contributions of mitochondrial ROS as a signalling agent should not be underestimated. ROS has a role in the regulation of physiological cellular senescence (Starkov, 2008) and mitochondrial biogenesis (Yoboue and Devin, 2012). Taking that into account, the levels of ROS in mitochondria should be tightly controlled, as excessive production can transform their role from merely signalling, to one harmful towards most biological molecules (Starkov, 2008).

Another role of mitochondria, outside energy production, is to manage intracellular calcium (Ca^{2+}) levels. Ca^{2+} signalling is essential for a multitude of cellular processes, including the regulation of secretion and cell migration but also, allows for muscle cell contraction and neuronal cell excitability (Yang and Huang, 2005; Tsai et al., 2014). Intracellular Ca^{2+} levels are diligently maintained through a combination of the activity of cation channels in the cell membrane and the release of Ca^{2+} from intracellular stores, mainly within the *endoplasmic reticulum* (ER) but also the mitochondrial network (Romero-Garcia and Prado-Garcia, 2019). Mitochondrial Ca^{2+} levels are of particular interest because of the close link between Ca^{2+} and mitochondrial function, metabolism and dynamics (Romero-Garcia and Prado-Garcia, 2019). Consequently, mitochondria are responsible for modulating both the timing and amplitude of Ca^{2+} signals (Romero-Garcia and Prado-Garcia, 2019). This is achievable through either their ability to buffer intracellular Ca^{2+} or through their functional interactions with other organelles, such as the ER (Deak et al., 2014). Close contact and interaction between the ER and mitochondria permit the creation of microdomains, allowing for an increase in mitochondrial Ca^{2+} buffering capacity, in line with increases in the levels of cytosolic Ca^{2+} (Jouaville et al., 1999). Physiologically-functioning mitochondria are able to buffer Ca^{2+} between 50 and 500nM in many different cell types (Imbert et al., 1995; Budd and Nicholls, 1996; Hartmann and Verkhratsky, 1998). If unchecked, elevated Ca^{2+} levels can be the trigger of cellular responses to disease states, as seen during T-cell activation (Hoth et al., 2000) and apoptosis (Pinton et al., 2008).

Programmed or regulated cell death, also known as apoptosis, is brought about not solely as a response to cellular stress or damage, but, is an integral part of normal cell development and the process of morphogenesis (Nikoletopoulou et al., 2013). Apoptosis can be set in motion extrinsically, through the activation of cell death receptors situated on the plasma membrane, such as tumor necrosis factor-α (TNFα), TNF related apoptosis inducing ligand (TRAIL) and Fas (CD95/APO1) (Nikoletopoulou et al., 2013). However, intrinsic triggering of apoptosis is

also achievable, through mitochondrial signalling (Adams, 1999; Kroemer et al., 2007). In both cases, early-stage apoptosis involves the caspase-induced cleavage of targeted substrates leads to mitochondrial outer membrane permeabilization (MOMP), widely considered the 'point of no return' as it brings about the release of cell death-promoters such as cytochrome c (Liu et al., 1996), apoptosis-inducing factor (AIF) (Susin et al., 1999), Smac/DIABLO, Htra2/Omi, Endo G and others (Green and Evan, 2002). These mitochondrial proteins have roles within the organelle that are unrelated to cell death (Newmeyer and Ferguson-Miller, 2003). However, once released from the IMS and into the cytoplasm, they set in motion a cascade of downstream events that result in apoptosome formation, chromatin condensation (Green, 2005), nucleic and DNA fragmentation as well as the shedding of apoptotic bodies (Nikoletopoulou et al., 2013).

In addition to their role in cell death, there is increasing evidence implicating mitochondria in the facilitation of antiviral signalling and antibacterial immunity following the release of ROS (West et al., 2011). Mitochondrial contribution to the mounting of innate immunity responses is an emerging field of great potential (West et al., 2011). It was the discovery of mitochondrial antiviral-signalling proteins (MAVS) that first linked mitochondria to innate immunity signalling (Seth et al., 2005). RIG-I-like receptor (RLR) signalling, involved in the recognition of viruses by the innate immune system, requires MAVS, an adaptor localized in the mitochondria that includes the nuclear factor κB (NF-κB) and the generation of pro-inflammatory cytokines alongside type I interferons (INFs) (Kawai et al., 2005; Meylan et al., 2005; Seth et al., 2005; Potter et al., 2008). Since the discovery of MAVS, other signalling molecules, including nucleotide-binding oligomerisation domain, leucine rich repeat containing X1 (NLRX1), NLR family pyrin domain containing 3 (NLRP3), immunity related GTPase M (IRGM) and TNF receptor associated factor 6 (TRAF6) have been linked to mitochondria (Arnoult et al., 2011). Consequently, our understanding of the role of mitochondria in immunity has shifted towards appreciating the ways in which innate immunity is integrated within and controlled by basic mitochondrial functions (Arnoult et al., 2011).

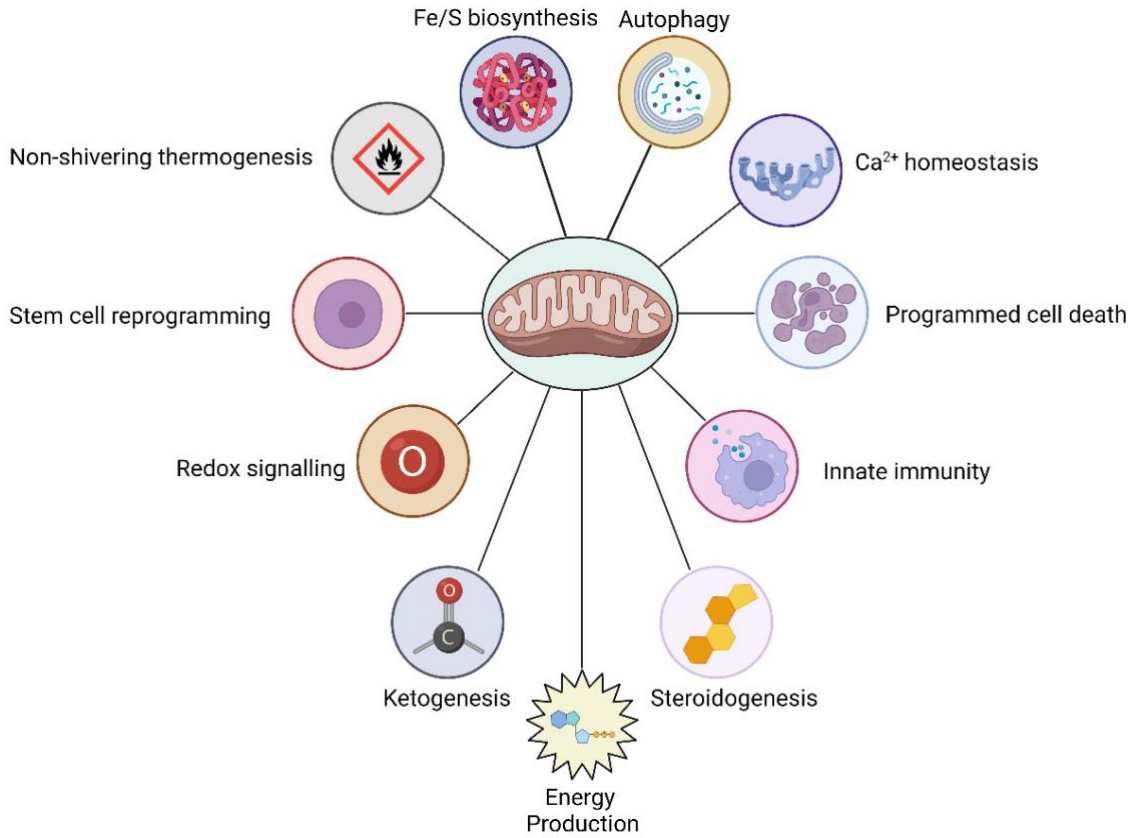


Figure 1.4. Cellular processes involving mitochondria
Schematic representation outlining the role of the mitochondrial network in a multitude of cellular functions, including: energy production, ketogenesis, redox signalling, stem cell reporgramming, non-shivering thermogenesis, iron/sulfuer cluter (Fe/S) biosynthesis, autophagy, Ca^{2+} homeostasis, programmed cell death, innate immunity and steroidogenesis. Created with BioRender.com.

1.3. The Mitochondrial DNA

Through the course of evolution, much of the DNA that originally belonged to the aerobic α-proteobacterium that first established the endosymbiotic relationship with an early, amitochondriate eukaryote, has been incorporated within the cell's nDNA (Dolezal et al., 2006). This development has stripped the mitochondrion of any last vestiges of autonomy, and cemented a co-dependent relationship between the cell as a whole and the mitochondrion (Berg and Kurland, 2000). It has been shown that the transposition of mitochondrial genes into the nDNA is a continuous and dynamic process (Leister, 2005), with many of the genes involved being essential for the maintenance of mitochondrial homeostasis (Berg and Kurland, 2000). Conversely, others are just pseudogenes known as nuclear mitochondrial DNAs (NUMTs), a term coined by the evolutionary biologist Jose V. Lopez (Lopez et al., 1994). Remnants of the ancient α-proteobacterium's DNA are still found within the mitochondrial matrix and the genes present in the mtDNA encode for essential components of the ETC (Wolstenholme, 1992).

The reason behind the incorporation of most mitochondrial genes to the nucleus of mammalian cells, with only a reduced genome being left behind, still residing within the organelle, has yet to be comprehensively explained. However, it has been argued that the selection process deciding which genes remain in the organelle was based on the encoded protein's hydrophobicity (Claros et al., 1995; Daley et al., 2002). The higher the hydrophobicity of the protein, the harder its transport through the cytosol and its import to the organelle (von Heijne, 1986). As a result, this hypothesis states that it is beneficial for the cell to have the most hydrophobic components of the ETC be translated at a close proximity to their final destination (von Heijne, 1986). An alternative hypothesis postulates that the existence on mtDNA encoded proteins serves to maintain the optimal flow of electrons through the ETC, through the *in-situ* modulation of components from Complexes I, III, IV and V (Allen, 2017). The semi-independent existence of the mtDNA allows for an individual mitochondrion to adapt the production of proteins encoded in its DNA based on its own, organelle-specific energetic needs, whilst bypassing any delays that might arise from complete dependence on nuclear-mitochondrial crosstalk (Allen, 2017). Finally, a third hypothesis makes the case for differences in genetic code between the nuclear and mitochondrial DNA being the underlying reason preventing the transfer of the remaining genes (Osawa et al., 1992). These differences exist in the coding of methionine and tryptophan in the two molecules (Osawa et al., 1992). Also, while

the mtDNA presents only two stop codons to the nDNA's three, possibly making coding regions in the mtDNA incompatible with those in the nucleus (Osawa et al., 1992).

1.3.1. Organisation of Mammalian Mitochondrial Genome

As opposed to the bi-parental inheritance that governs nDNA origins, in mammals, the mtDNA is maternally inherited in its entirety (Sato and Sato, 2012). True to its prokaryotic ancestry, the mtDNA is, multi-copied, double-stranded and circular in structure. The revised Cambridge reference sequence (rCRS) described the human mtDNA as spanning 16,569bp (Andrews et al., 1999). The mouse mtDNA on the other hand, spans 16,299bp in length (Bayona-Bafaluy et al., 2003). Sequencing the human mtDNA for the first time in 1981 (Anderson et al., 1982) and its subsequent revision in 1999 (Andrews et al., 1999) marked the beginning of the Human Genome Project, making huge inroads in our understanding of mitochondrial disease, population genetics and human evolution.

The two strands of the mtDNA are easily distinguishable based on their nitrogenous base composition. One strand, is rich in guanine and thus branded the heavy strand (H-strand), while its complementary strand, rich in cytosine is termed the light strand (L-strand) (Vinograd et al., 1963). This difference in base composition allows for the physical separation of the two strands using density centrifugation in alkaline CsCl gradients (Wells and Larson, 1972). There are 37 genes encoded in the mtDNA, unevenly distributed between the two strands, with 28 of them being on the H-strand and 9 on the L-strand. Thirteen proteins are encoded by the mtDNA, all essential subunits of ETC complexes, namely, seven Complex I subunits (ND1, ND2, ND3, ND4/4L, ND5, ND6), one Complex III subunit (CYTB), three Complex IV subunits (COI/II/III) and two Complex V subunits (ATP6/8) (Anderson et al., 1981). Twenty-two tRNAs are also encoded in the mtDNA to allow for translation of take place in the mitochondrial matrix. One tRNA corresponds to each amino acid, with the notable exception of both leucine and serine which are being represented by two distinct molecules (Anderson et al., 1981). Two rRNAs forming the 12S and 16S components of the mitoribosomes are also encoded in the mtDNA (Anderson et al., 1981) **(Figure 1.5.)**.

As opposed to the nDNA, there are no introns present within the mtDNA, however, it possesses a large noncoding region (NCR). The NCR is often home to a third DNA strand that spans over and separates that particular region of double stranded DNA (dsDNA) (Nicholls and Minczuk, 2014). This segment is termed the displacement loop (D-loop) and its common occurrence has led to the terms NCR and D-loop to be used interchangeably in the literature (Nicholls and Minczuk, 2014). However, one should keep in mind that the D-loop is not always present and when it is, it covers only a fraction of the NCR (Nicholls and Minczuk, 2014). Promoters for the replication of both strands are housed in the NCR, termed heavy strand promoter (HSP) and light strand promoter (LSP) respectively (Montoya et al., 1982; Montoya et al., 1983). Additionally, a regulatory sequence responsible for controlling the process of mtDNA replication, termed origin of heavy-strand replication (OriH) is also present within the NCR. OriH is mirrored by the origin of light-strand replication (OriL), positioned 11kb downstream, on the opposite strand (Montoya et al., 1982). The two origins of replication separate the mtDNA into two sections, one that starts with OriH and end with OriL, termed the major arch, and the remaining section, termed the minor arc (Eimon et al., 1996). With the NCR portion spanning between the LSP and OriH, three conserved sequence blocks (CSBs) can be found, known as CSB1, CSB2 and CSB3 (Walberg and Clayton, 1981). Finally, on the light strand, downstream of OriH, the termination associated sequences (TAS) is situated (Madsen et al., 1993).

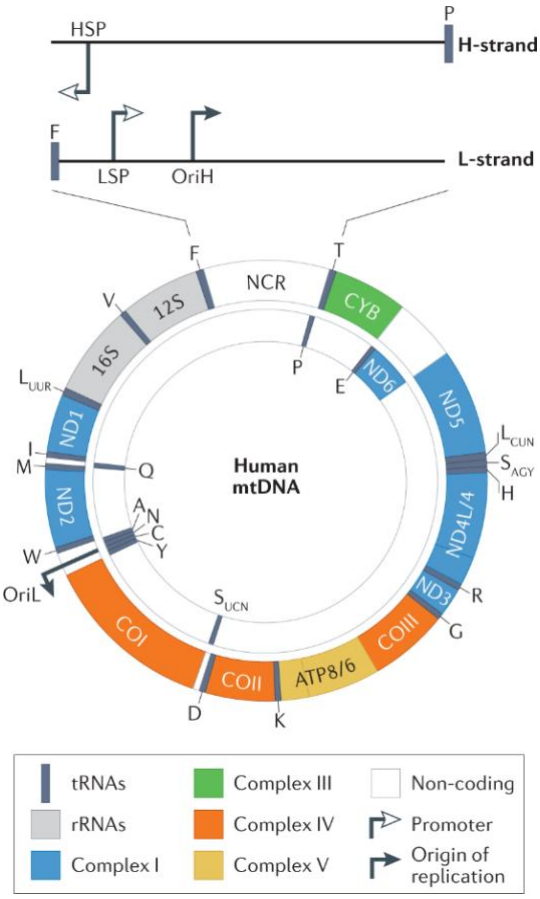


Figure 1.5. Structure of the mammalian mitochondrial DNA.
Schematic representation of core elements of the mtDNA. Spanning 16,569bps in length, the human mitochondrial DNA (mtDNA) encodes for 13 polypeptides in addition to 22tRNAs and 2rRNAs that are essential to their translation. All of these genes are densely packed and unevenly distributed between the two strands, the heavy strand (H-strand) and the light strand (L-strand). The mtDNA also contains one major non-coding region (NCR). Within there, the majority of cis-elements required for mtDNA transcription and translation are situated: heavy strand promoter (HSP), light strand promoter (LSP) and origin of replication for the heavy strand (OriH). The origin of replication for the light strand (OriL) is situated 11kb downstream of the OriH, defining a region known as the major arch. The remaining section of the mtDNA is known as the minor arch. Figure adapted from: Silva-Pinheiro and Minczuk, 2022.

1.3.2. Mitochondrial Nucleoid Structure

Early observations of mitochondrial structure established that the organelle is quite narrow with width of about 0.5μm and contour length of 5μm (Nass, 1966). This means that the DNA existing within the matrix of mitochondria has to be packaged in a way that will allow for optimal control of the genetic material. Indeed, similar to bacterial chromosomes, the mtDNA is tightly packaged into DNA-protein structures called nucleoids (Kucej and Butow, 2007). It is known that organisation of the mtDNA into nucleoids assists in anchoring the genetic material to the IMM (Wang and Bogenhagen, 2006), however the dynamics governing mtDNA nucleoid movement are not yet fully understood (Jazek and Dlaskova, 2019). Although mitochondrial fission events as of late have been subdivided into peripheral and midzone evens (Kleele et al., 2021), they are largely nucleoid-centric in nature, resulting in daughter organelles containing at least one nucleoid (Tauber et al., 2013). Early experimentation on nucleoids seemed to reveal that there were around 2-8 molecules of mtDNA per nucleoid structure (Legros et al., 2004). However, more recent work that employed the use of two-dimensional stimulated emission depletion (2D STED) super-resolution microscopy, revealed that, on average there is 1.4 molecules of mtDNA per nucleoid in human fibroblasts (Kukat et al., 2011; Kukat et al., 2015).

The most prolific protein component in the nucleoid is mitochondrial transcription factor A (TFAM) (Kanki et al., 2004). TFAM is essential for mtDNA maintenance as the first identified mitochondrial transcription factor, it also belongs to a high motility group (HMG) protein family that binds to DNA non-selectively and introduces a 180^{o} U-turn (Kukat et al., 2015). In addition to bending the DNA strands, TFAM can introduce cross-strand binding by interacting with adjacent TFAM molecules (Kukat et al., 2015). Combining the bending and strand-binding properties of TFAM with its high abundance (approximately 1000 molecules per mtDNA) (Takamatsu et al., 2002), it is clear why its function is often compared to nDNA histones. *In vitro* experiments on TFAM have also revealed a regulatory aspect to the protein, as compaction of mtDNA nucleoids correlates with decreased levels of mtDNA replication and transcription (Farge et al., 2014). At this point, it is important to consider that more tightly condensed nucleoids can coexist with more relaxed molecules (Farge et al., 2014) and while

there are hundreds of copies of mtDNA within most cells, only a portion of them undergo replication or transcription at a given point in time (Kukat et al., 2011) **(Figure 1.6.)**.

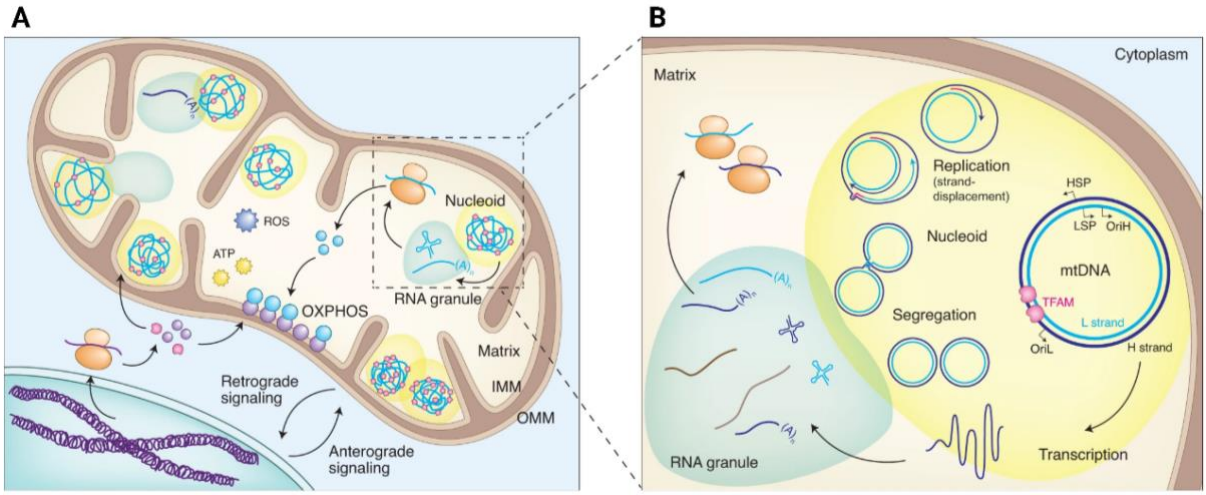


Figure 1.6. Mitochondrial DNA organisation into nucleoids.
Schematic representation of mtDNA nucleoid organisation and distribution within the organelle. **(A)** While the mitochondrial DNA (mtDNA) encodes for 13 polypeptides central to oxidative phosphorylation (OXPHOS) activity, the vast majority of the mitoproteome is encoded in the nucleus and has to be imported from the cytosol. This includes around 1,500 proteins that are involved in mtDNA maintenance and expression. The maintenance of the nucleo-mitochondrial crosstalk is dependent on a continuous stream of retrograde and antiretrograde signalling. **(B)** mtDNA lacks histones and instead is packaged in compact structures called nucleoids that are a hub for mtDNA replication and transcription. Replication of the mtDNA follows principally the tenants of the strand displacement model. First, the light strand promoter (LSP) is used for the creation of an RNA primer (depicted in pink). Then, synthesis of the H-strand (heavy strand) takes place starting at origin of replication for the heavy strand (OriH). Eventually, the nascent H-strand (depicted in dark blue) is produced though the unidirectional movement of the replisome. Once the nascent H-strand reaches origin of replication for the light strand (OriL), a stem loop structure is formed that allows for the synthesis of an RNA primers used to initiate the replication of the light strand (L-strand) (depicted in light blue). The separation of the two newly formed mtDNAs is achieved through decatenation. MtDNA transcription initiation takes place at the heavy strand promoter (HSP) and LSP accordingly and results in the formation of a polycistronic RNA molecule that will eventually undergo the steps required for processing and maturation within the RNA granule, situated in close proximity to the nucleoid. mtDNA replication, transcription and translation will be discussed in greater detail in **Introduction Section, Chapter 1.3.4**. and **5**. IMM, Inner Mitochondrial Membrane; OMM, Outer Mitochondrial Membrane. Figure adapted from: Fu et a., 2020.

1.3.3. Mitochondrial DNA Replication and Copy Number Control

While nDNA of a cell replicates only once within each cell cycle, during the synthesis phase (S-phase), in preparation for cell division, the mtDNA has a high turnover rate (Chinnery and Samuels, 1999). Individual mtDNA molecules undergo replication in a semi-autonomous manner, independent of the nDNA, following a pattern known as relaxed replication (Birky, 1994). The central role mitochondrial play in many of the cell's metabolic processes makes the faithful replication of mtDNA a necessity. DNA polymerase γ (POLγ) is the only mtDNA polymerase found in mammals, and is responsible for the replication of both the heavy and light strand (Falkenberg and Gustafsson, 2020). POLγ is a heterodimer, made up of one POLγA subunit with catalytic activity and two accessory POLγB subunits (Young et al., 2015). Human POLγA subunit possesses 5´-3´polymerase, 5´-deoxyribose phosphate lyase and 3´-5´exonuclease activities (Longley et al., 1998), the latter of the three bestowing it proofreading abilities and allowing the correction of base misincorporations. This characteristic enables POLγ to have an error frequency of less than $1x10^{-6}$ per nucleotide, making it a particularly reliable polymerase (Longley et al., 2001). The accessory POLγB subunits provide stability to the molecule, increasing its processivity and thus, allowing it to replicate long spans of DNA (Carrodeguas et al., 1999). Both subunits of POLγ are essential to its activity and ablation of either, in conjunction with the lack of built-in redundancy, leads to a drastic reduction in the total mtDNA content of the cell and, eventually, developmental arrest at an early embryonic stage (Hance et al., 2005; Humble et al., 2013).

While POLγ is the sole polymerase present within mitochondria, there are other components whose presence is essential for the faithful replication of the mtDNA. POLγ alongside the DNA helicase TWINKLE and the mitochondrial single-stranded DNA-binding proteins (mtSSBs) comprise the minimal mtDNA replication machinery *in vitro* (Korhonen et al., 2004). TWINKLE is the only helicase required for mtDNA replication (Peter and Falkenberg, 2020). Its activity and structure are similar to that of homohexameric ring DNA helicases, possessing a helicase, a linker and a 5-primase domain (Korhonen et al., 2003). Being a helicase, TWINKLE is able to unwind the double stranded mtDNA in a 5´- 3´ direction, fuelling its activity through the hydrolysis of nucleotide triphosphate (NTP) (Peter and Falkenberg, 2020). Upon unwinding of the mtDNA, evolutionarily conserved mtSSBs attach to the single stranded DNA (ssDNA) in a manner that lacks sequence-specificity and serves to protect the strands

against degradation, while also maintaining the function of the now single stranded portions of mtDNA (Meyer and Laine, 1990; Shereda et al., 2008; Pestryakov and Lavrik, 2008). Studies in fruit flies (*Drosophila melanogaster*) revealed that mtSSBs helped to increase fly POLγ processivity through the initiation of mtDNA synthesis (Williams and Kaguni, 1995; Farr et al., 1999). Following the pattern reported in the absence of POLγ, fly models that lacked mtSSBs displayed rapid depletion of their mtDNA, accompanied by stagnation of cell proliferation during development (Maier et al., 2001). Prior to the initiation of mtDNA replication by POLγ, mitochondrial RNA polymerase (POLRMT) has to prime the sites where replication originates (OriH and OriL) (Wanrooij et al., 2008). When presented with ssDNA, POLRMT exhibits low processivity, resulting in the formation of short RNA primers (25–75nt in length) that are then used by POLγ to kick off the process of mtDNA replication (Wanrooij et al., 2008).

To date, there is no consensus on the model followed for the replication of mtDNA. The three primary hypotheses include the strand displacement model, the strand coupled replication model and a model that endorses ribunocleotide incorporation through the lagging strand (RITOLS) **(Figure 1.7.)**.

The strand displacement model of mtDNA replication first emerged in the early 1970s, as the result of EM imaging (Kasamatsu et al., 1971). It described how synthesis of the leading strand takes place at a specific site on the mtDNA (now known as OriH) and proceeds for about two thirds of the molecule until it reaches a second site (now known as OriL) where the synthesis of the lagging strand begins (Kasamatsu et al., 1971). It is the formation of a hairpin structure at OriL that allows for the reversal of the unidirectional DNA synthesis, this time back towards OriH (Clayton, 1982). POLRMT makes use of this hairpin structure to generate the RNA primer that is later going to be used by POLγ to carry out the L-strand synthesis (Wanrooij et al., 2012). Following the initiation of L-strand synthesis, the two mtDNA strands continue to be synthesized in parallel, until both are complete and the entire molecule is successfully replicated. While notably different from the processes surrounding nDNA replication, the strand replacement model is not without precedent as a difference in replication initiation between the leading and lagging strands has been observed in viral DNA and plasmids (Khan, 1997).

The strand coupled replication model was given birth to following the first two-dimensional agarose gel electrophoresis of mtDNA originating from rat livers (Holt et al., 2000). This study claimed that a proportion of mitochondrial replication intermediates (RIs) were made up of duplex DNA, leading the authors to conclude that they were the product of simultaneous replication of both the leading and lagging strands (Holt and Reyes, 2012). Although at the time quite preliminary, this observation made a strong case for the ability of the lagging strand to initiate its synthesis at different positions on the mtDNA (Holt and Reyes, 2012). Studies would later find that replication could be initiated within a broad span of the mtDNA that covers several kilobases (kb), thus making a case for the lagging strand's discontinued synthesis as well as the presence of Okazaki-like fragments (Reyes et al., 2005; Holt et al., 2000). The maturation of these fragments is required prior to the ligation, and even though the mitochondrial proteome includes components of the Okazaki maturation pathways (Ruhanen et al., 2011), there is yet no evidence that Okazaki fragments exist within human mitochondria.

The RITOLS model of mtDNA replication makes the case against the asynchronous model of replication described by the strand displacement theory. It bases this claim on observations of unusually large DNA molecules forming RI arcs that were particularly slow, originating from mitochondria that had undergone high levels of purification (Yang et al., 2002). These RIs were resistant to any treatment with either duplex or single-stranded DNA nucleases and, having ruled out the possibility that proteins, lipids or carbohydrates were involved, RNA was singled out as the primary suspect (Holt and Reyes, 2012). Indeed, the susceptibility of these intermediates to RNase H, provided the evidence required to cement the presence of RNA/DNA hybrids (Yang et al., 2002). Consequently, this model proposes that during replication, the single strand of mtDNA is not subject to single stranded DNA binding protein 1 (SSBP1) coating, but instead, this role is performed by RNA originating from mitochondrial transcripts, which is then used by POLγ for the synthesis of DNA (Yang et al., 2002). Further support for this theory came when RNA was observed to be coating unwound, single stranded mtDNA at the point of the replication fork (Reyes et al., 2013).

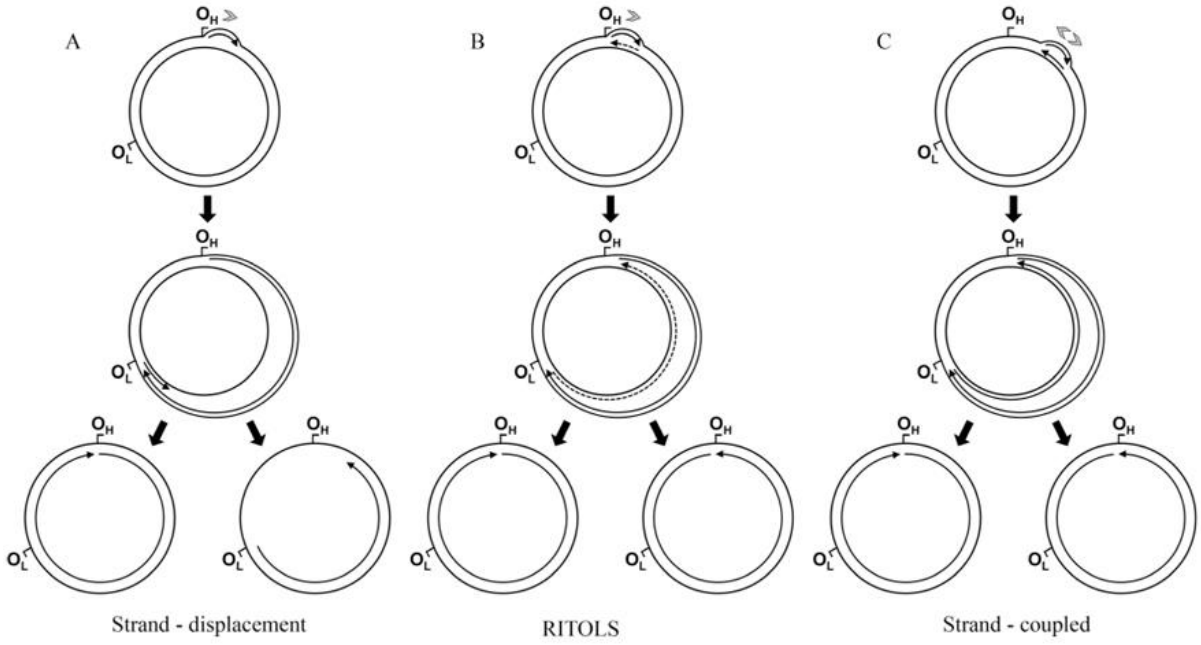


Figure 1.7. Possible models describing mammalian mitochondrial DNA replication.
Schematic representation of the three most prominent models explaining the steps undertaken during mammalian mtDNA replication. **(A)** The strand displacement model, **(B)** The RITOLS (RNA incorporated through the lagging strand) model. **(C)** The leading and lagging strand-coupled model. Arrows on the replicating mtDNA point towards the 5'-3' direction followed by DNA synthesis; continuous lines represent DNA while dashed ones represent RNA (the possible presence of short RNA primers is not represented, only long stretches of RNA described by the RITOLS model); grey arrowheads at and downstream of O_H indicate the model-dependent directionality and number of replication forks; O_H (origin of replication for the heavy strand); O_L (origin of replication for the light strand). Figure adapted from: McKinney and Oliveira 2013.

While for the most part there are two copies of nuclear DNA within the nucleus of each mammalian cell, one derived from each parent, the same is not true for mtDNA. mtDNA is maternally inherited (Sato and Sato, 2012) and since the point of oocyte maturation, it is evident that multiple copies of it are essential for the proper function of the cell (Wai et al., 2010). In order to address this phenomenon, the term copy number has been coined to refer to the number of mtDNA molecules that exist within a cell at a given time (Longchamps et al., 2020). With the notable exception of red blood cells (RBCs) in the human body carry many hundreds and sometimes up to many thousands of copies of mtDNA (Shuster et al., 1988). The copy number of a cell is tightly correlated to its energy requirements, with neurons, muscle cells and hepatocytes being known for their particularly high copy number (Herbst et al., 2021; Wisniewski et al., 2016). Conversely, cell types such as epithelial or immune cells can be sustained by a much more limited network (Ding et al., 2015; Cohen et al., 2016). The responsibility of regulating mtDNA copy number falls to the components of the cell's

replisome machinery (Montier et al., 2009). POLγ, TFAM, TWINKLE and SSBP1 protein levels have all been shown to have an effect in mtDNA copy number regulation (Montier et al., 2009). A relatively recent study carried out in *Drosophila melanogaster* has also linked nDNA-encoded subunits of ATP synthase to the regulation of mtDNA copy number (Fukuoh et al., 2014). Indeed, the regulation of mtDNA copy number through the years has been shown to be influenced by multiple factors, including but not limited to cytoskeletal proteins (Reyes et al., 2011), mitochondrial dynamics (Gaziev et al., 2014) and biogenesis (Montier et al., 2013), abundance of chaperone proteins (Santos et al., 2011), availability of metabolic and carrier proteins (Thompson et al., 2016) as well as exonucleases and proteases (Filograna et al., 2021). Overall, copy number control is central to cell and tissue development from the point of oocyte fertilization as in the absence of sufficient mtDNA content, mitochondria-dependent apoptosis can be triggered (Benkhalifa et al., 2014).

1.3.4. Mitochondrial DNA Transcription

As previously described, mitochondria play a central role in a number of important cellular processes and in, order to optimize their activity, have retained a greatly reduced genome, comprised of 13 polypeptides, 22 tRNA and 2 rRNA molecules (Gorman et al., 2016). The expression of mitochondrially encoded genes is of vital importance to the cell, and as a result, the process of mtDNA transcription is coordinated very closely with that of nDNA-encoded genes (Boczonadi et al., 2018).

Initiation of mtDNA transcription takes place within the NCR, and more specifically, at the HSP and LSP respectively, resulting in the formation of two polycistronic RNAs that span the length of almost the entire mitochondrial genome (Gustafsson et al., 2016). POLRMT binds specifically to promoter elements within the mtDNA, but is unable to initiate RNA synthesis on its own (Rusecka et al., 2018). Instead, it needs to form a transcription initiation complex alongside mitochondrial transcription factor B2 (TFB_2M) and TFAM (Rusecka et al., 2018). Indeed, it is TFAM that first binds to the promoters upstream to the transcription initiation site, bringing about a 180^o U-shaped bend in the strand (Falkenberg et al., 2002). It is also involved in tethering the N-terminal region of POLRMT, thus recruiting it onto the promoter site while TFB_2M introduces structural changes to the RNA polymerase, breaking down duplex DNA

and bringing about promoter opening, while trapping the non-template single stranded mtDNA strand (Hillen et al., 2017).

Once the transcription initiation step is complete, TFB_2M is removed from the complex and replaced by the mitochondrial transcription elongation factor (TEFM) (Minczuk et al., 2011; Posse et al., 2015). TEFM is involved in enhancing the processivity of POLRMT, allowing it to generate longer RNA transcripts (Jiang et al., 2019). It has also been suggested that TEFM allows the process of transcription elongation to overcome regions of mtDNA that might cause stalling or even premature termination, such as oxidative lesions or secondary RNA structures formed by G-quadruplexes in CSB2 (Hillen et al., 2017).

The exact mechanism surrounding transcription termination is still a hotly debated topic, with questions being asked on whether the function of mitochondrial transcription termination factor 1 (MTERF1), is required for the termination of all transcripts originating from the three different promoters found in the control region (Barchiesi and Vascotto, 2019). However, it is generally agreed upon that termination of the complete HSP transcript takes place at the D-loop's 3′-end where the TAS is located, while the LSP transcript terminates downstream of the 16S rRNA (Gustafsson et al., 2016). The imbalance in mitochondrial gene distribution between the two strands is of course reflected in the resulting transcripts, with the LSP mRNA encoding only Nd6 and 8tRNAs, while the HSP mRNA encodes for the remaining 12 proteins, 14 tRNAs and 2 rRNAs (Gustafsson et al., 2016). Following their completion, the two polycistronic RNA transcripts are separated from their complementary mtDNA strands and segmented into individual mRNAs in order to be translated by the mitochondrial ribosome (D'Souza and Minczuk, 2018) **(Figure 1.8.)**.

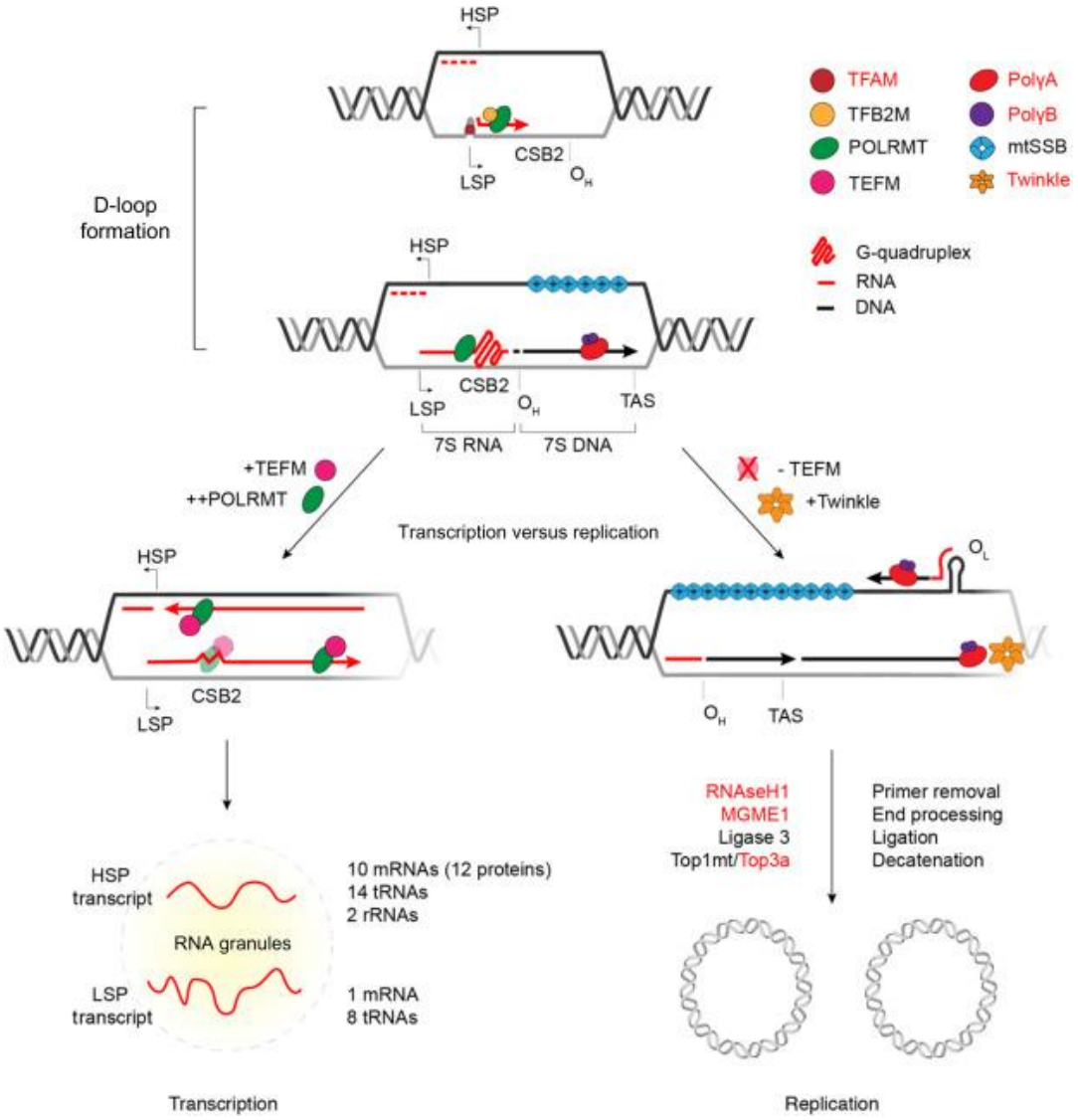


Figure 1.8. The role mitochondrial DNA D-loop in replication vs transcription.
Schematic representation of the switch between mtDNA replication and transcription. **(Top Section)** D-loop formation is a product of mitochondrial transcription factor A (TFAM)-promoted transcriptional activity that originates at the light-strand promoter (LSP). Both mitochondrial RNA polymerase (POLRMT) and mitochondrial transcription factor B2 (TFB2M) are then recruited at the LSP. Their recruitment brings about the transcription of just a 200nt-long RNA molecule as the G-quadruplex sequence prevents any further progress. DNA polymerase γ (Polγ) then uses this short RNA construct (termed 7S RNA) as a primer to generate the 7S DNA, a span of DNA that leads to the disruption of the heavy-light strand hybridisation. Mitochondrial single stranded DNA binding proteins (mtSSBs) are then recruited in order to stabilise the newly single-stranded portion of DNA. **(Lower Left Section)** In order for transcription to be carried out, higher levels of POLRMT and mitochondrial transcription elongation factor (TEFM) are recruited. This allows for the bypassing of the G-quadruplex complex and the initiation of a genome-wide transcription event that ends with the generation of two polycistronic RNA transcripts, one originating at the HSP and one at the LSP. Further processing of these molecules takes place within the RNA granules. **(Lower Right Section)** In the case of mtDNA replication as described by the strand displacement model, TFEM is absent while Twinkle is recruited at the termination-associated sequence (TAS), bringing about the formation of the replisome. mtDNA replication will be completed as RNA primers are removed and two new molecules undergo end processing, ligation and are eventually separated through decatenation. Mitochondrial pathologies arise in humans when proteins indicated in *red* are mutated. Figure adapted from: Fontanesi et al., 2020.

1.3.5. Mitochondrial DNA Translation

From the moment of their synthesis to that of their degradation, mitochondrial RNAs (mtRNAs) have to undergo multiple processing steps of modification and maturation in order for the proteins encoded by the mtDNA to be successfully translated. The processes of transcription and translation within the mitochondria have to be coordinated both in terms of their timing but also with regard to their spatial organization (Barchiesi and Vascotto, 2019). To satisfy this need, mtRNAs undergo maturation and processing within mitochondrial RNA granules (MRGs), dynamic structures that are responsible for the regulation of post transcriptional processing, making sure that mtRNAs have fully matured prior to the initiation of protein synthesis (Mai et al., 2017; Pearce et al., 2017).

In order for the exact function of the MRGs to be deciphered, a known component of the complex, endogenous G-rich RNA sequence-binding factor 1 (GRSF1) was used as a bait for immunoprecipitation (Antonicka and Shoubridge, 2015). The result was compelling, as the presence of known components such as ribonuclease P (RNaseP) was substantiated (Jourdain et al., 2013), while the discovery of hSUV3 and Polynucleotide Phosphorylase (PNPase), both components of the mitochondrial degradosome, gave credence to the theory that MRGs are not just hubs of mtRNA processing and maturation but are also involved in its degradation and turnover (Borowski et al., 2014). Immunoprecipitation followed by mass spectrometry analysis of proteins associated with the primary polycistronic transcript, revealed an abundance of components responsible for its post-translational processing, including mitochondrial poly-A polymerase (mtPAP), mitochondrial ribonuclease P protein (MRPP), pentatricopeptide repeat domain 3 (PTCD3) and dimethyladenosine transferase 1 (TFB1M) (Barchiesi and Vascotto, 2019). Furthermore, structural components of the large and small mitochondrial ribosomal subunits (mt-LSU and mt-SSU), proteins involved in mitochondrial translation, factors involved in the process of ribosomal composition and breakdown as well as aminoacyl tRNA synthetases were identified as being part of MRGs (Barchiesi and Vascotto, 2019). Overall, this data suggests that the function of MRGs is not very dissimilar to that of the nucleolus, as they are implicated in translation regulation and mitoribosome biogenesis (Tu and Barrientos, 2015; Barrientos, 2015).

The two polycistronic transcripts that are generated as a result of mtDNA transcription have to undergo processing in order to release individual mRNAs. Within these polycistrons, tRNAs separate most of the rRNA and mRNA-coding regions in a pattern described as the mt-tRNA punctuation model (Anderson et al., 1981; Ojala et al., 1981). This model describes how the majority of individual mRNA coding sequences encoded within the H-strand polycistron, are strictly preceded and followed by tRNA sequences and that endonucleolytic cleavage takes place precisely before and after the tRNA sequence (Ojala et al., 1981), resulting in the release of the messenger and ribosomal mtRNAs. However, it should be noted that there are exceptions to this rule, notably ND5, Cytb and ATP6/8 **(Figure 1.8.)**.

Once released from the polycistron, individual mRNAs have to undergo steps of post-translational modification that exist to enhance their stability, in a way distinct and more simplistic to the one seen in nDNA-encoded mRNAs (Barchiesi and Vascotto, 2019). The first distinction between the maturation of nuclear-encoded mRNAs versus those encoded within the mtDNA is the lack of introns as well as that of a canonical 5′CAP modification (Barchiesi and Vascotto, 2019). At the same time, while both cytosolic and mitochondrial mRNAs have a poly-A tail attached to their 3′-end, that of mt-mRNAs is a lot shorter (Temperley et al., 2010).

Indeed, the poly-A tail of nuclear-encoded mRNAs measures around 250 nucleotides in length, while that of mt-mRNAs ranges between 45 and 55 nucleotides with ND6 and sometimes ND5 reportedly not undergoing any polyadenylation (Temperley et al., 2010). Moreover, the mt-mRNA 3′-end poly-A tail length varies depending on the cell type (Temperley et al., 2010). Synthesis of the mt-mRNA poly-A tails comes down to the non-canonical mtPAP localising within MRGs, providing evidence in support of the theory that the first steps in mt-mRNA maturation take place co-transcriptionally (Tomecki et al., 2004; Nagaike et al., 2005; Bai et al., 2011; Wilson et al., 2014). While the exact contribution of the 3′-end poly-A tail's addition to mt-tRNA has not yet been fully understood (Nagaike et al., 2005), it has been demonstrated that in the absence of mtPAP there is a notable decrease in mitochondrial gene translation leading to downstream complications, impairing the cell's respiratory capacity (Rorbach et al., 2014). One possible function of the poly-A tails is to complete the stop codon at the 3′-end of the mt-mRNA (Anderson et al., 1981; Van Haute et al., 2015). Following the tRNA's nucleolytic cleavage from the primary transcript, seven out of thirteen encoded mRNAs do not terminate in a stop codon, but are left with a 3′ ending in either "U" or "UA" (Barchiesi and

Vascotto, 2019). Consequently, the addition of a 3′-end poly-A tail can result in the completion of the missing stop codon. Another striking difference between nuclear and mitochondrial-encoded mRNA is the 5′CAP modification. While a N7 methyl guanosine (m7G) cap at the 5′-end of the molecule is used to promote stability and translation efficiency in nuclear-encoded mRNA, the same is not true for their mitochondrial counterparts (Barchiesi and Vascotto, 2019). Instead, mt-mRNA harbor nicotinamide adenine dinucleotide (NAD^+) as their 5′CAP, which instead of promoting stability, stimulates the mRNA's decay (Kiledjian, 2018).

Similar to the structure of cytosolic ribosomes, mitoribosomes are also comprised of two subunits, the a larger (16S) and a smaller one (12S), both encoded within the mtDNA, as well as protein components that are encoded in the nucleus and have to be imported from the cytosol (Brown et al., 2014; Greber et al., 2015; Kaushal et al., 2014; Amunts et al., 2015). Mitochondrial rRNAs undergo modifications upon release from the polycistrons, similar to but less extensive than the one undergone by cytosolic ribosomal components. However, in stark contrast to their cytosolic counterparts, they do not undergo any nucleolytic processing (Decatur and Fournier, 2002; Piekna-Przybylska et al., 2008; Van Haute et al., 2015). The last stages in the lifespan of mtRNA are believed to take place in structures called degradation foci (D-foci) which are primarily hosting the mitochondrial degradosome but at the same time include many enzymes the role of which has yet to be determined (Borowski et al., 2013; Szczesny et al., 2010). There is evidence that suggests D-foci co-localise with MRGs but it whether they are a completely distinct entity from them or just a sub-population has not yet been addressed (Barchiesi and Vascotto, 2019).

1.3.6. Mitochondrial Transfer RNAs

The overwhelming majority of RNA molecules are first synthesized as immature precursors and their conversion to their final, functional form is a central step in the process of gene expression. All RNA components required for translation to take place within mitochondria, are encoded by the mtDNA. Upon release from the primary transcript, mt-tRNAs need to reach their mature form and final confirmation that is necessary in order for translation of the mt-mRNA to be carried out (Barchiesi and Vascotto, 2019).

Nucleoside modifications are a key structural feature of all tRNAs (Duechler et al., 2016; Hori et al., 2014). These modifications are performed post transcriptionally, and play a vital role in ensuring the tRNA's proper function and maintained stability (Suzuki et al., 2011). However, mt-tRNAs have a much more limited number of modified nucleosides compared to their cytoplasmic counterparts (Suzuki et al., 2011). The chemical modifications undergone by mt-tRNAs can be classified into two categories: modifications that are in place to grant a tRNA the folding and structural stability required, and those that are responsible for implementing adjustments in order to optimize its interaction with other translation factors (Bai et al., 2011). These chemical modifications and the consequences that are brought about by their disruption are discussed in more detail in **Introduction Section, Chapter 1.4.2**.

It is generally accepted that mt-tRNAs, share the same ‘cloverleaf’ structure, characterized by three loops and four stems, as their cytosolic counterparts (Florentz et al., 2003; Kogelnik et al., 1997; Martin, 1994). Also common between the two is the presence of a three-base anti-codon, making sure that the codon present on the mRNA is appropriately recognized, as well as that of the acceptor stem which, when maturation is complete, is charged with the amino acid of choice (Florentz et al., 2003; Helm et al., 2000). However, the secondary structure of mt-tRNAs is often thought of as atypical when compared to that of cytosolic ones (Florentz et al., 2003). The two most structurally distinct mt-tRNAs are mt-tRNA$^{Ser\ (AGY)}$, which lacks the entirety of the D-arm, and mt-tRNA $^{Ser\ (UCN)}$, where there the anticodon and D stem are linked by only one nucleotide (Helm et al., 2000). Mitochondrial tRNAs generally exhibit greater variability in the size of the T and D arms while at the same time having a shorter variable region (Yarham et al., 2010). While structural flexibility is central to mt-tRNAs, conserved Watson-Crick base pairs that provide the molecule its distinct tertiary structure, are still present

(Florentz et al., 2003). It is not uncommon for mismatches to be present within mt-tRNAs, non-classical and weaker G-U pairs can be found within heavy tRNAs that are transcribed from the L-strand, while more easily degradable UpA and CpA and dinucleotides can be found in light mt-tRNAs (Helm et al., 2000). Overall, tRNAs can be classified into four distinct categories **(Figure 1.9.)**.

First, Type 0 includes all canonical tRNAs found in all walks of life and generally maintain their characteristic cloverleaf secondary structure that in order to be converted into the final tertiary L-shaped structure, has to undergo multiple interactions between looks and helices (Suzuki et al., 2011). The mt-tRNAs that are member of Type I (mt-tRNA $^{Ser\ (UCN)}$), possess an unconventional cloverleaf structure that maintains canonical interactions between the D-loop and T-loop while at the same time having unique structural features, such as and extended anticodon stem, a short D-loop and one base only between the D and acceptor stems (Watanabe et al., 1994). Type II includes a lot of mt-tRNAs that are defined by a lack of canonical interactions between the D and T-loops (Suzuki et al., 2011). At the same time, these two loops are characterized by variable sequence and size (Suzuki et al., 2011). Finally, as previously mentioned, mt-tRNA$^{Ser\ (AGY)}$ lacks the entirety of the D-loop and is classified as Tpe III, being the only known case of a tRNA deviating entirely from the cloverleaf structure.

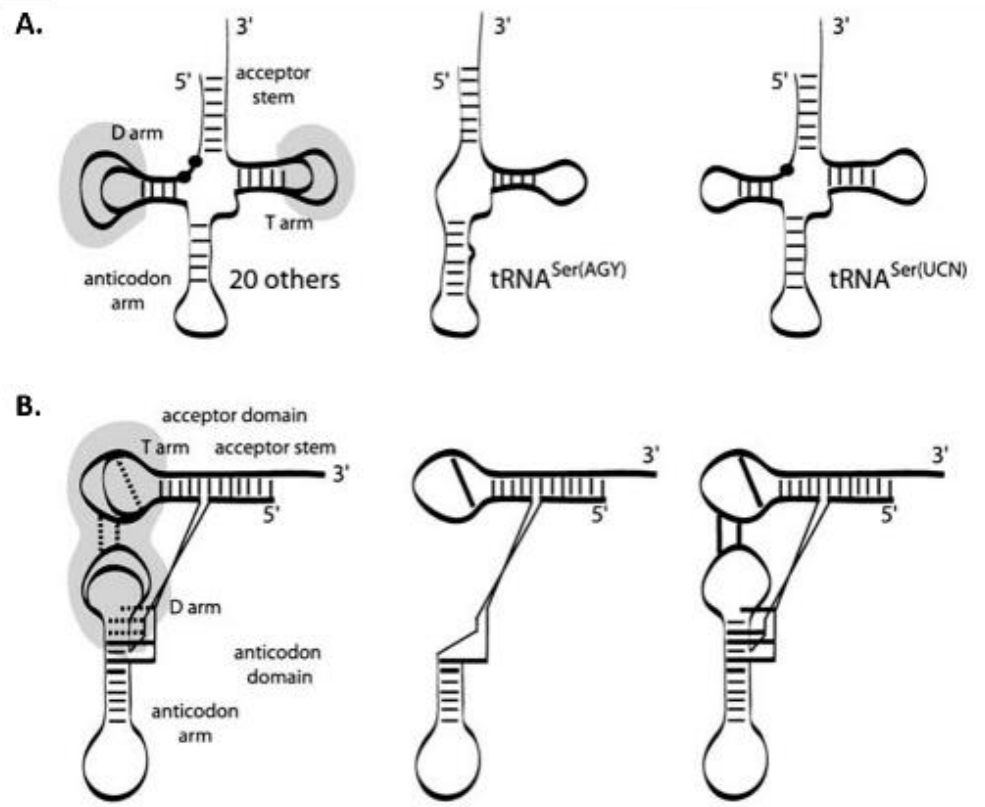


Figure 1.9. Structural characteristics of human mitochondrial tRNAs.
Schematic representation of the two- and three-dimentional structures of human mt-tRNAs (mitochondrial tRNAs). **(A)** Typical two-dimentional structure of human mt-tRNAs. On the left, the general secondary structure shared by 20 mt-tRNAs, possessing the classic cloverleaf structure with D, T and anticodon arms. In the middle, the structure of mt-tRNA$^{Ser\ (AGY)}$ that completely lacks the D-arm. On the right, the structure of mt-tRNA$^{Ser\ (UCN)}$ is characterised by a shorter connector 1 region and a longer anticodon stem. **(B)** Typical tertiary structure of human mt-tRNAs. Bold lines represent long-distance interaction, while dashed lines signify triple interactions that are not strictly conserved. Domains that are highlighted in grey represent regions where the type and number of interactions present vary greatly between tRNA molecules. Figure adapted from: Florentz et al., 2003.

1.4. Mitochondria in Disease

The very first case where mitochondrial dysfunction was identified as the origin of disease was in 1958 by the Swedish endocrinologist Rolf Luft, now widely recognised as the father of mitochondrial medicine. Dr. Luft at the time described a young patient suffering from euthyroid hypermetabolism, who despite having a high-caloric food uptake, exhibited weight loss accompanied by excessive sweating. Follow-up work stemming from these initial observations resulted in the discovery of uncoupled mitochondria present in the patient's muscles (Ernster et al., 1959). Dr. Luft's ground-breaking report was the spark that gave birth to the field of mitochondrial medicine and since his time, mutations in more than 300 mitochondrial genes have been implicated as the underlying cause in mitochondrial disorders, with more than 150 unique syndromes of mitochondrial disease having been identified (Stenton and Prokisch, 2018).

Given the central role that mitochondria play as the cell's primary energy production machinery, diseases of the mitochondrion are genetic disorders characterized by OXPHOS deficiency. At this point is it imperative to draw the distinction between primary mitochondrial disorders (PMDs) that are genetic diseases brought about as a result of mutations directly hampering the cell's OXPHOS activity or the mitochondrial machinery required in order for it to be carried out, and secondary mitochondrial disorders (SMDs) (Niyazov et al., 2016). SMDs can be brought about by genes whose function has nothing to do with OXPHOS or mitochondrial quality control, such dysfunctions often go hand in hand with a wide range of non-mitochondrial hereditary conditions and in some cases, can even be given rise to by environmental factors (Niyazov et al., 2016). For purposes of clarity and to prevent the conflation of the two terms, when used in this dissertation, the term 'mitochondrial disorders' will be referring exclusively to PMDs.

Primarily impacting the cell's ability to produce ATP, mitochondrial disorders comprise the most prevalent type of inherited metabolic disorders and the second most prevalent of inherited neurological conditions (Gorman et al., 2016). Mitochondrial disorders are brought about as a result of mutations affecting either genes encoded within the mtDNA, or, nDNA-encoded genes that give rise to mitochondrial proteins. Being governed by such complicated genetics, means that mitochondrial disorders can follow different inheritance patters, including

autosomal, X-linked and maternal inheritance in the case of mtDNA mutations (Gorman et al., 2016). In certain rare cases, the appearance of *de novo* mutations can also be the cause for disease (Sallevelt et a., 2017). One of the most challenging aspects when it comes to dealing with mitochondrial diseases in the clinic, is the striking pathological heterogeneity exhibited between patients that results in diagnostic delays (Gorman et al., 2016). In patients who develop mitochondrial disorders as a result of mutations in the mtDNA, shedding light into the inheritance and the factors affecting clinical presentation becomes even more challenging as within each individual cell in the patient's body, there is a ratio of mutant to wild-type (WT) mtDNA molecules that reside in a state of dynamic coexistence called heteroplasmy (Melton, 2004). Determining the heteroplasmy level of a given mutation is crucial as the higher the percentage of mutant mtDNA that gets accumulated within the cell is, the more likely it is for the cell's mitochondrial activity to be disrupted (Melton, 2004).

For a long time, mitochondrial disorders were considered to be rare, however, advances in clinical diagnosis coupled with genetic testing have set the prevalence at about 1:5000 (Melton, 2004). The most frequent presentation of mitochondrial disorders is that of neurological syndromes (Smeitink et al., 2006; Schapira, 2006; Lin and Beal, 2006). The genetic heterogeneity in combination with the biochemical complexity of mitochondrial disorders, bring about highly heterogeneous clinical manifestations (Chinnery, 1993). These can range from single-tissue lesions as seen in Leber's hereditary optic neuropathy (LHON), to more widespread lesions seen in myopathies, cardiopathies and encephalopathies, to even multisystem disorders whose onset can vary from neonate to adult life (Chinnery, 1993).

Mitochondrial diseases can surface at any age and their hallmark clinical heterogeneity means that the patient can develop a range of symptoms that at first glance might appear unrelated (Gorman et al., 2016). Usually, the ubiquitous nature of the organelle and the dependency of all tissues on constant energy production, means that such disorders are multisystem **(Figure 1.10.)**, more often than not affecting tissues and organs that depend highly upon aerobic metabolism (McFarland et al., 2010). Patients suffering from mitochondrial disorders usually deteriorate with age, while the syndromes are defined by high morbidity and mortality (McFarland et al., 2010).

Most common clinical features seen in paediatric patients include lactic acidosis, psychomotor delay, generalized hypotonia as well as signs of cardiorespiratory failure (Zeviani and Carelli, 2003). Leigh syndrome is the most common and best characterized early onset mitochondrial encephalopathy (Chinnery, 1993). Adult patients of mitochondrial disorders often show signs of myopathy that are associated with variable central nervous system (CNS) involvement that can entail hearing loss, seizures, ataxia and polyneuropathy among others (Zeviani and Carelli, 2003). There is a host of biochemical and morphological feature that characterize most of adult-onset mitochondrial disorders, with the most prominent being the formation of ragged red fibres (RRFs) from scattered muscle fibres (Chinnery, 1993).

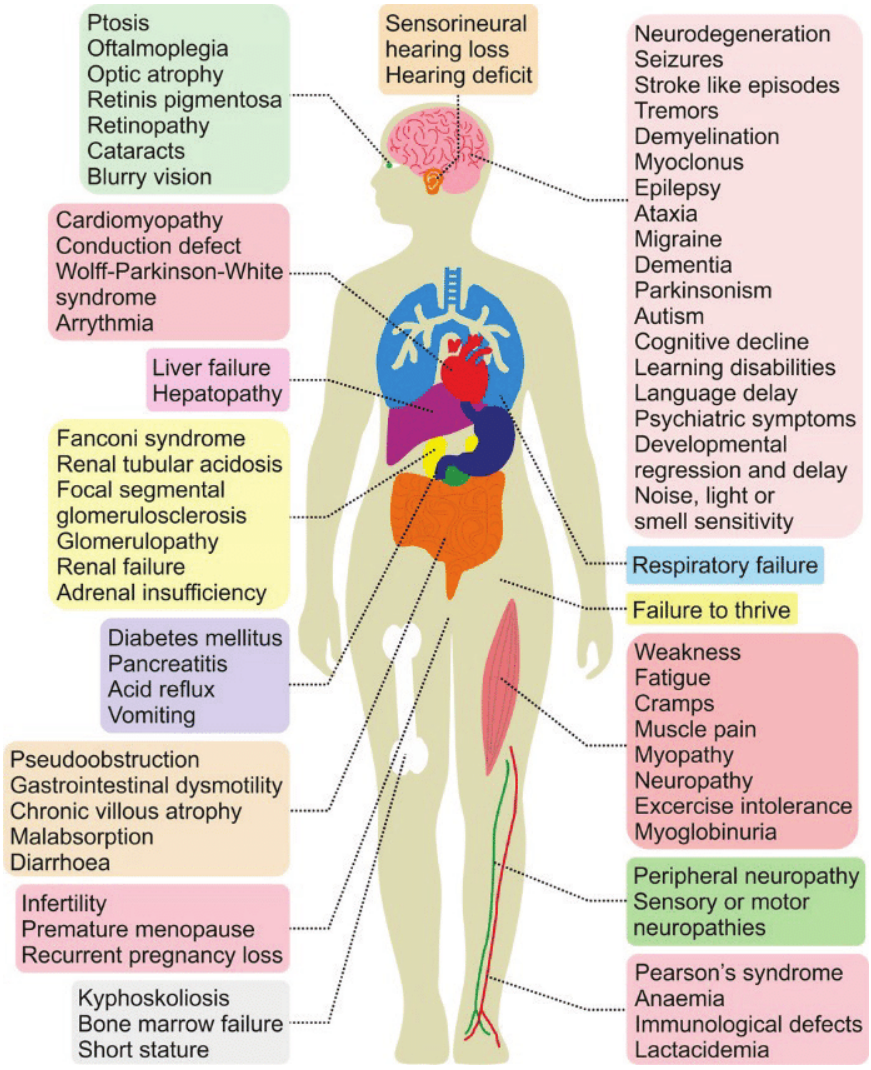


Figure 1.10. Mitochondrial disease heterogeneity.
Clinical presentations of mitochondrial disorders tend to be extremely inconsistent as they exhibit a wide range of symptoms. These symptoms can show further variation in terms of their severity, ranging from mild to quite severe disorders, impacting primarily the nervous system. Figure adapted from: Konarikova et al., 2020.

1.4.1. Mutations in Nuclear versus Mitochondrial Genes

With evolution driving the majority of genes encoding for mitochondrial proteins out of the organelle, and incorporating them within the nDNA, it does not come as a surprise that while the majority of adult patients exhibit mtDNA mutations (Gorman et al., 2015), there are at least 300 known nDNA-encoded genes harbouring mutations that lead to mitochondrial disorders (Abicht et al., 2018). Expanding our understanding of the molecular basis behind mitochondrial disorders has been hampered by the fact that to date, we have yet to indentify the entirety of the proteins comprising the mitoproteome (Andreoli et al., 2004; Cotter et al., 2004), which is

estimated to consist of 1,500 proteins (Lopez et al., 2000). Consequently, despite their prevalence, the identification of nuclear genes involved in mitochondrial disorders has moved at a much slower pace compared to efforts made towards identifying and characterizing mitochondrial DNA mutations (Zeviani et al., 2003). A combination of genetic and clinical criteria has been used in order to classify nuclear genes that are involved in mitochondrial disorders into four categories (Zeviani, 2001). The first two groups of genes include structural components of the ETC and their assembly factors respectively, while group three includes genes involved in the regulation of mtDNA stability (Zeviani, 2001). Finally, the last group of nuclear genes under this classification includes factors involved in mitochondrial biogenesis (Zeviani, 2001).

In contrast to the diploid nature of the nDNA, the mtDNA is multi-copied and as a result, most mutations coexist alongside WT copies in heteroplasmic state. In principle, the higher the level of heteroplasmy of a given mutation, the more likely it is for a disease phenotype to be developed. While the mitochondrial network attempts to buffer the presence of mutations through metabolic compensation, there is a certain threshold of heteroplasmy that exceeds the network's buffering capabilities. This heteroplasmy threshold is mutation and cell-type dependent (Stewart and Chinnery, 2015). The first type of mtDNA mutation leading to disease are point mutations that include base substitutions as well as base insertions or deletions (indels). Mitochondrial DNA point mutations have an estimated population prevalence of about 1:200 and constitute a major cause of disease in humans (Chinnery et al., 2012). Point mutations have been recorded in every single one of the 37 genes encoded within the human mtDNA and have been associated with a wide spectrum of clinical presentations, that spans from non-syndromic sensorial deafness, to mitochondrial myopathy, encephalopathy, lactic acidosis, stroke-like episodes (MELAS) (Alston et al., 2017). Symptoms are not age-dependent and can present during childhood as well as adulthood, with about 75% of the cases being inherited and the remaining 25% arising through spontaneous (*de novo*) mutations (Sallevelt et al., 2017). The presence of a mtDNA bottleneck during the process of oocyte maturation (discussed in more detail in **Introduction Section, Chapter 1.4.4.**), means that a mother that carries a mutation at a heteroplasmy level that is below the threshold required to present clinical symptoms, can give rise to offspring that exceed that threshold and as a result, suffer from that pathogenic condition (Wilson et al., 2016). Consequently, it is incredibly hard to predict the risk carried in successive pregnancies and amniocentesis or chorionic villus sampling (CVS) are the only ways of informing reproductive choices by prenatally screening foetal tissue

heteroplasmy (Nesbitt et al., 2014). Single, large-scale deletions of the mtDNA are not as common as point mutations, with a population prevalence of about 1.5 : 100,000 (Gorman et al., 2015). There are three main disease phenotypes associated with single, large-scale deletions, namely chronic progressive external ophthalmoplegia (CPEO), accounting for around 65% of cases, Kearns–Sayre syndrome (KSS), accounting for around 30% of cases and Pearson syndrome, accounting for less than 5% of cases (Mancuso et al., 2015). Of the three, Pearson syndrome is the more severe, with patients presenting sideroblastic anaemia and pancreatic dysfunction early in life, it is often fatal in infancy (Rotig et al., 1990). In contrast to gene rearrangements occurring in the nDNA, single, large-scale deletions are given rise to during embryonic development in a sporadic fashion and have a low risk of recurrence (Chinnery et al., 2004).

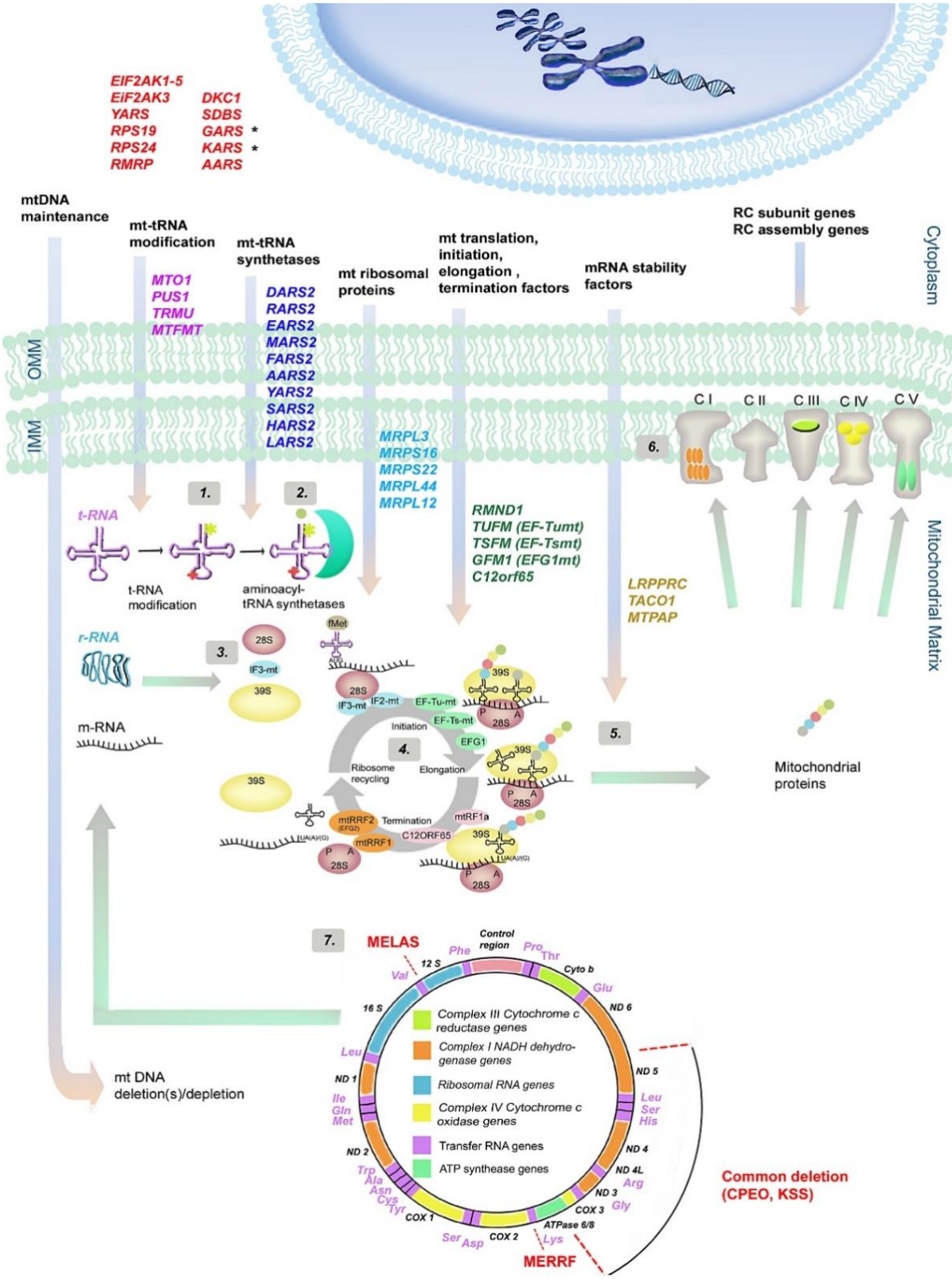

EIF2AK1-5
EIF2AK3
YARS
RPS19
RPS24
RMRP
DKC1
SDBS
GARS *
KARS *
AARS
mtDNA maintenance
mt-tRNA modification
mt-tRNA synthetases
mt ribosomal proteins
mt translation, initiation, elongation , termination factors
mRNA stability factors
RC subunit genes
RC assembly genes
Cytoplasm
OMM
IMM
MTO1
PUS1
TRMU
MTFMT
DARS2
RARS2
EARS2
MARS2
FARS2
AARS2
YARS2
SARS2
HARS2
LARS2
MRPL3
MRPS16
MRPS22
MRPL44
MRPL12
RMND1
TUFM (EF-Tumt)
TSFM (EF-Tsmt)
GFM1 (EFG1mt)
C12orf65
LRPPRC
TACO1
MTPAP
C I
C II
C III
C IV
C V
6.
1.
2.
t-RNA
t-RNA modification
aminoacyl-tRNA synthetases
r-RNA
3.
28S
IF3-mt
39S
fMet
IF2-mt
EF-Tu-mt
EF-Ts-mt
EFG1
Initiation
Elongation
Ribosome recycling
Termination
mtRRF2
(EFG2)
mtRRF1
mtRF1a
C12ORF65
m-RNA
4.
5.
Mitochondrial proteins
Mitochondrial Matrix
mt DNA deletion(s)/depletion
7.
MELAS
MERRF
Common deletion (CPEO, KSS)
Control region
12 S
16 S
Cyto b
ND 1
ND 2
ND 3
ND 4
ND 4L
ND 5
ND 6
COX 1
COX 2
COX 3
ATPase 6/8
Complex III Cytochrome c reductase genes
Complex I NADH dehydro-genase genes
Ribosomal RNA genes
Complex IV Cytochrome c oxidase genes
Transfer RNA genes
ATP synthease genes

Figure 1.11. Human mitochondrial protein synthesis defects.
Schematic representation of mutations in both nuclear and mitochondrial DNA that lead to defective mitochondrial proteins synthesis. Preceding mitochondrial proteins synthesis are the mechanisms ensuring mtDNA maintenance, and faithful replication and transcription. Mutations in nuclear genes responsible for carrying out these processes often result in mtDNA depletion and/or the formation of deletions. A different host of nuclear-encoded genes have to be imported from the cytosol in order to guarantee reliable mtDNA translation. These genes can be divided into different functional groups. The first group of genes includes members that are involved in the translation of proteins within the cytosol (highlighted *red*). A few of these genes (denoted with a *star*) are imported into the mitochondria where they function as aminoacyl t-RNA synthetases (ARSs). **1.,** Highlighted *purple* are nuclear-encoded genes that perform mitochondrial tRNA (mt-tRNA) modifications. **2.,** Highlighted *dark blue* are ten mitochondrial ARSs that have been implicated in human mtDNA translational defects when mutated. **3.,** Highlighted in *light blue* are nuclear-encoded genes coding for mitoribosomal proteins which when mutated are known to cause translational defects. **4.,** Highlighted in *green* are genes that play a key role in carrying out the three steps of translation: initiation, elongation and termination. **5.,** Nuclear-encoded genes involved in mRNA stability and translation activation that also, when mutated, impact the synthesis of mitochondrial proteins. All respiratory complexes (with the notable exception of CII), require a combination of nuclear and mitochondrial-encoded polypeptides. The nuclear-encoded portion of respiratory complex subunits have to be synthesised and transported to the mitochondrial matrix, alongside their assembly factors. There, combined with the 13 mtDNA-encoded polypeptides, the assembly of subunits into functional enzymatic complexes and supercomplexes takes place. **6.,** Fully assembled enzyme complexes embedded in the inner mitochondrial membrane (IMM) making up the electron transport chain (ETC). **7.,** Schematic representation of the circular mtDNA structure, depicting all encoded genes: 13 proteins, 2 r-RNAs and 22 t-RNAs, in addition to the control region (also known as non-coding region). The diagram denotes mutations leading to MERF and MELAS syndromes as well as common deletions. Figure adapted from: Mitochondria: Boczonadi and Horvath, 2014.

1.4.2. Mitochondrial Transfer RNAs in Disease

A well-regulated process of mtDNA transcription and translation is essential for the maintenance of a functional mitochondrial network and everything that entails for the cell as a whole. As a result, mutations within genes involved in the regulation of these processes are the underlying cause for approximately a third of all mitochondrial disorders (Boczonadi and Horvath, 2014). Advances in the field of next generation sequencing have allowed us to identify several new genes associated with these disorders, and helped shed light into new pathogenic mechanisms, providing us with new insight into previously unknown molecular pathways.

More specifically, mt-tRNA modifications play a crucial role in the regulation of cellular energy output in response to local requirements, therefore, dysfunction in the application of said modifications, can bring forth disorders directly linked to mt-tRNAs (Boczonadi and Horvath, 2014). The first category of mitochondrial disorders brought about as a result of defects in mt-tRNAs involve alterations in the wobble base pairing, a form of interaction

between nucleotides found in RNA that does not adhere to the Watson-Crick rules of complementarity. The tRNA 5-methylaminomethyl-2-thiouridylate methyltransferase (*TRMU*) gene encodes for an enzyme involved in the thiouridylation of mt-tRNA^{Lys}, mt-tRNA^{Glu} and mt-tRNA^{Gln}. This process is essential for translational accuracy and requires the presence of cysteine, an amino acid that is considered essential during the first months of life (Sturman et al., 1970). Autosomal recessive mutations in TRMU are the cause of reversible infantile liver failure (Zeharia et al., 2009; Schara et al., 2011; Uusimaa et al., 2011; Gaignard et al., 2013). The majority of patients who suffer from this condition and manage to survive the first year of life, exhibit spontaneous and complete recovery (Boczonadi et al., 2015). In the same fashion as TRMU, mt-tRNA Translation Optimization 1 (*MTO1*) is involved in the catalysis of 5-taurinomethylation of uridine found in wobble base pairs within mt-tRNA^{Lys}, mt-tRNA^{Glu} and mt-tRNA^{Gln} (Boczonadi and Horvath, 2014). 5-taurinomethylation leads to an increase in the efficiency and accuracy of mt-mRNA translation (Li et al., 2003) and the first patients who carried recessive mutations on that gene, exhibited severe infantile hypertrophic cardiomyopathy (Ghezzi et al., 2012). GTP Binding Protein 3 (*GTPBP3*) is responsible for the timely generation of 5-taurinomethyluridine (τm(5)U) in the wobble position within the anticodon of mt-tRNAs (Kopajtich et al., 2014). Patients carrying recessive mutations in *GTPBP3* exhibited lactic acidosis and neurological symptoms, with some of them showing evidence of hypertrophic cardiomyopathy (Kopajtich et al., 2014). Finally, NOP2/Sun RNA Methyltransferase 3 (*NSUN3*) codes for a 5-methylcytosine (m(5)C) methyltransferase responsible for the methylation and formylation of mt-tRNA^{Met} (Van Haute et al., 2016). A patient with loss of NSUN3 function was described as having skeletal muscle OXPHOS deficiency, in addition to microcephaly, developmental decay, nystagmus, failure to thrive, external ophlalmoplegia, muscle weakness and recurrent lactic acidosis (Van Haute et al., 2016).

Mutations affecting genes coding for proteins responsible for the modification of position 37 on mt-tRNAs, a position known for the presence of hypermodifications, have also been found to lead to mitochondrial dysfunction. Autosomal recessive mutations affecting the tRNA methyltransferase 5 (*TMRT*) gene, have been correlated to hypomodifications of position 37 guanosine residues, especially within skeletal muscle (Powel et al., 2015). Individuals carrying mutations in TMRT exhibit dysfunction of the respiratory chain in skeletal muscle and lactic acidosis, while the dysfunction has also been associated with complex hereditary spastic paraparesis (Tarnopolsky et al., 2017). tRNA isopentenyltransferase (*TRIT1*) is responsible for

introducing i6A37 modifications in both cytosolic and mitochondrial tRNAs (Yarham et al., 2014). Mutations of the *TRIT1* gene have been associated with severe combined respiratory chain defects, encephalopathy and myoclonic epilepsy (Yarham et al., 2014). Additional processes that when disrupted directly affect the activity of mt-tRNAs include formylation of the mitochondrial methionine tRNA (*MTFMT*), pesudouridylation (carried out by Pseudouridylate synthase 1 (*PUS1*)) and 3′-CCA addition (performed by tRNA nucleotidyl transferase (*TRT1*)) (Wedatilake et al., 2016; Chen and Patton, 1999; Bykhovskaya et al., 2004). The severity of these conditions and their direct link to the maintenance and optimisation of mt-tRNA activity underlines the importance of mtDNA translation in maintaining a functional respiratory chain, especially within tissues characterised by high energetic demands.

1.4.3. Inheritance of Mitochondrial DNA Mutations

Since the mid 1970's it has been established that in mammals, the inheritance of mtDNA takes place strictly down the maternal line (Hutchison et al., 1974; Pyle et al., 2015). As a result, human mtDNA does not undergo the process of recombination, a mechanism that serves to prevent the transmission of mutations that have accumulated within the nDNA to the next generation, while serving to maintain genetic variability (Hagstrom et al., 2014). Moreover, despite the reliable nature of POLG, there is evidence that errors occur as a result of successive rounds of mtDNA replication (Ameur et al., 2011). Additionally, as the mtDNA resides within the mitochondrial matrix in close proximity to the IMM, it is adjacent to the ETC, which constitutes a major source of ROS. Consequently, the mtDNA is subject to more oxidative stress than its nuclear counterpart (Shokolenko et al., 2009). As a result, it is predicted that the mtDNA should be subject to the Muller's ratchet and eventually succumb under the forces of a mutational meltdown (Muller, 1964).

Indeed, there is evidence supporting the accumulation of mutations within the mtDNA and their use in the study of population genetics has resulted in the formation of phylogenetic trees, diagrams indicating the generational link between human populations. While the human mitochondrial DNA can be traced back to a single ancestral mtDNA foremother termed 'Mitochondrial Eve' (Cann et al., 1987), classification of variability in sequences between populations results in the formation of clusters that can be distinguished from one another on the basis of polymorphisms, these clusters are called haplogroups (Schurr et al., 1999; Torroni

et al., 1997). Although we have not yet been able to actively record the fixation of a polymoprphism at the population level, their origin should, in theory, coincide with the emergence of a heteroplasmic variant. Following the appearance of this variant, intergenerational selection in favour of the polymorphism would result in the establishment of a now homoplasmic variant that acts as a signifier of a new branch of the phylogenetic tree. At this point, it should be noted that the process of haplogroup emergence must be more complicated than initially suggested, as studies have found that very low level heteroplasmic variants are present within all humans (Payne et al., 2013), and the prevention of accumulation at higher heteroplasmies suggests the presence of regulatory forces that make sure the emergence of a disease state is largely avoided. Further reinforcing this point, more recent work has shown that the vast majority of polymorphisms, when present at a high enough heteroplasmy percentage within an organism, are the cause for mitochondrial dysfunction (Stewart and Chinnery, 2015).

The presence of rapid shifts in heteroplasmy within an organism is a phenomenon that has been recorded throughout many different animal models both vertebrate and invertebrate. Observed in Holstein cows for the very first time, it was made clear that the coexistence of two mtDNA genotypes, differing from one another only as a result of an adenine to guanine base transition, resulted in the generation of an additional Hae III recognition site within the same organism, resulted in the fixation of one or the other over the span of a few generations (Hauswirth and Laipis, 1982). Following the same principles, in 1983, flies that carried two types of mtDNA molecules that differed in their length (termed Short “S” and Long “L”) displayed a rapid drift towards either the S or the L genotype without showing any signs of preference towards one of the two (Solignac et al., 1983). The same principle was also demonstrated using conplastic mice, animals where the nuclear genome of one strain is backcrossed into the cytoplasm, and as a result the mtDNA of another. In this experiment, mice carried at the same time both the NZB and BALB mtDNA genotypes and exhibited striking shifts in heteroplasmy within just one generation (Jenuth et al., 1996). While there appeared to be a selection taking place intergenerationally, when comparing maternal heteroplasmy to that of the pups, the heteroplasmy range observed between individual pups was symmetrically distributed around the maternal heteroplasmy value (Jenuth et al., 1996). Further studies on conplastic mice, this time combining the C57BL/6N and the NZB/BINJ and haplotypes, did not exhibit clear evidence in favour of selection (Meirelles and Smith, 1997). Instead, the authors described that while in some cases there appears to be a rapid segregation towards homoplasmy that does not

require more than two generations to be established, in other lineages there seems to be a stable transmission of both haplotypes (Meirelles and Smith, 1997). Studies performed on human embryos of mothers who carried the m.3243A>G mutation, revealed evidence that the mutation follows stochastic (non-chromosome-dependent) segregation patterns (Monnot et al., 2011). This observation was in line with prior work on both oocytes and embryonic data (Brown et al., 2001; Steffann et al., 2006). Further, more extensive work was carried out using 577 mother-child pairs that transmitted a range of pathogenic mutations: m.3243A>G, m.8344A>G, m.3460G>A, m.8993T>G/C, m.8344A>G and m.11778G>A (Wilson et al., 2016). This study did not present any evidence in support of selection during transmission but demonstrated that the rate of segregation in a human pedigree is mutation-dependent (Wilson et al., 2016). When taken together, all of the above studies suggest that a variety of heteroplasmic models across a range of different species display a pattern of genetic drift that is akin to the stochastic segregation of heteroplasmic mtDNA variants across generations.

Going against the evidence provided above, there have been cases where segregation of heteroplasmy from mother to offspring appears to be subject to active selection rather than random drift. In 2012, a study was published looking three pathogenic mutations (m.8993T>C/G, m.3243A>G and m.8344A>G) that were found to be overrepresented in 327 unrelated patients or pedigrees (Hellebrekers et al., 2012). This study concluded that most of the oocytes derived from carriers of pathogenic mutations will contain heteroplasmies that are below the pathogenic threshold, with the m.3243A>G cohort showing selection against the mutation (Hellebrekers et al., 2012). Furthermore, work on conplastic mice that contained heteroplasmic mtDNA made up of both NZB and 129S6 genotypes in a C57BL/6J nuclear background, showed that across generations there was a preferential reduction in the percentage of NZB mtDNA (Sharpley et al., 2012). Following the trend of active selection against heteroplasmic variants down within the germline, work done in human germ cells in addition to early-stage embryos revealed that nonsynonymous mutation present in regions on the mtDNA coding for proteins, tRNA or even within the NCR were selected against (Floros et al., 2018).

While most of the work on mtDNA heteroplasmy segregation has been focused on the germline, through the study on mother-offspring pairs, more recently there have been attempts to elucidate heteroplasmy dynamics within somatic tissues. Work carried out in isolated mitochondrial from specific tissues systems of *Caenorhabditis elegans* demonstrated that somatic lineages are less prone to propagate pathogenic mtDNA mutations compared to the germline, a fact that was attributed to the increased rate of mtDNA replication found in this tissue (Ahier et al., 2018). Experiments centered around heteroplasmic mice have concluded that the artificial disruption of the homoplasmic state, sets in motion active haplotype segregation that, in most tissues, follows a consistent, non-random pattern (Jenuth et al., 1997; Sharpley et al., 2012; Burgstaller et al., 2014). In addition, conplastic mice have been shown to possess functionally different OXPHOS systems and as a result, exhibit variation in their metabolic profiles (Latorre-Pellicer et al., 2016). Based on those premises, Ana Victoria Lechuga-Vieco and colleagues attempted to define somatic heteroplasmy dynamics using heteroplasmic mice that carried both NZB and C57Bl/6J mtDNA (Lechuga-Vieco et al., 2020). Examining the segregation bias of the two haplotypes, revealed that highly proliferative tissues displayed a preference for C57Bl/6J mtDNA, while more quiescent ones appeared to select for NZB (Lechuga-Vieco et al., 2020), suggesting that on the C57BL/6JOlaHsd nuclear background, the C57Bl/6J hapltype promotes glycolysis, while NZB promotes lipid metabolism (Latorre-Pellicer et al., 2019). Their work confirmed that the heteroplasmic state has a significant and deleterious effect on cell metabolism, at the same time the observed differences in OXPHOS performance are strongly associated with both cell type and mtDNA genotype (Lechuga-Vieco et al., 2020). Furthermore, non-random mtDNA segregation was recorded, with most tissues selecting for one of the two haplotypes in an intracellular and cell-type dependent manner that could be influenced by tampering with genetic, nutritional and pharmacological factors (Lechuga-Vieco et al., 2020). It has long been theorised that the liver as a tissue welcomes the age-dependent accumulation of nonsynonymous mtDNA mutations that will lead to reduction in its metabolic capacity (Li et al, 2015; deGrey, 1997). The reasoning underpinning this theory, termed 'survival of the slowest,' is that a reduction in the organs metabolic processes will eventually result in in the production of fewer damaging metabolic by-products thus, increasing the healthspan of the organ (Li et al, 2015; deGrey, 1997). However, in direct contradiction of this theory, recent work done on mice carrying both NZB and C57BL/6 mtDNA on a C57BL/6 nuclear background demonstrated that the liver's tendency to accumulate the NZB haplotype with age was not only enhancing the tissue's

respiratory capacity per DNA molecule, but also was the product of selective autophagic removal of the BL6 mtDNA (Tostes et al., 2022).

In conclusion, the examples provided above paint a very complex picture with germline and somatic tissues being characterised by both random drift and preferential selection of mtDNA variants existing within a heteroplasmic state. The mechanism governing mtDNA heteroplasmy segregation is still largely unknown and while most of the work to date has been geared towards understanding dynamics within the germline, segregation within somatic tissues might hold the key to uncovering the complex pathophysiology of mitochondrial disorders.

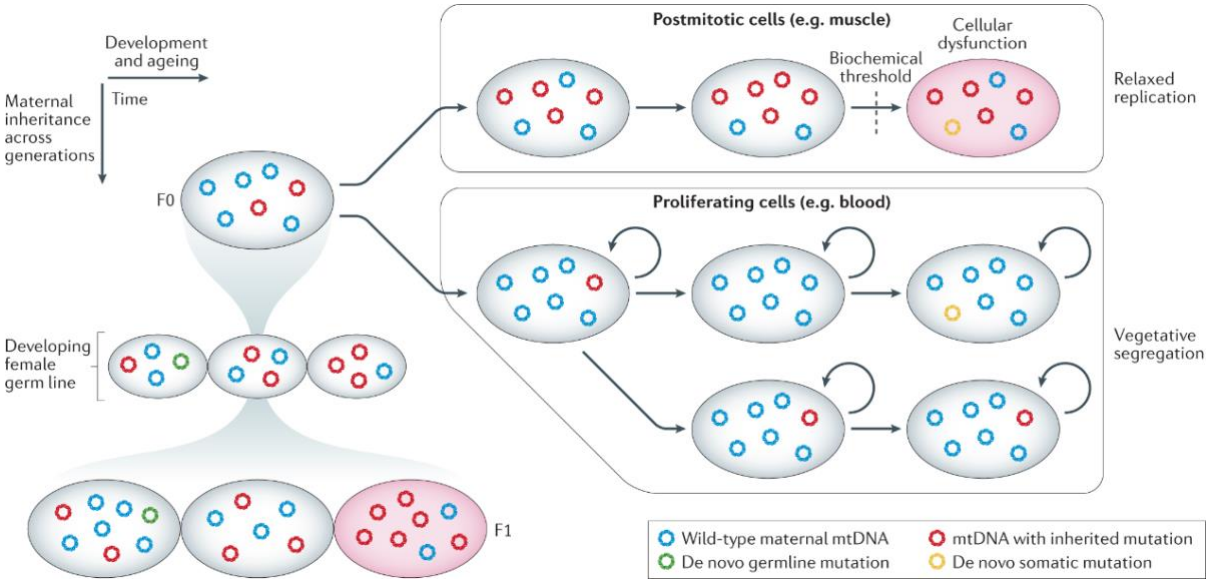


Figure 1.12. Extreme heteroplasmy shifts during inheritance and through development.
Schematic representation of the shifts in mitochondrial DNA (mtDNA) heteroplasmy that take place over the generations and through ageing. Levels of mtDNA heteroplasmy can undergo changes through lifespan of an organism, as a result of different forces. Vegetative segregation, the arbitrary partitioning and replication of organelles within the cytoplasm can result in shifts within proliferating cells. In the case of non-dividing postmitotic cells, relaxed replication governing the mtDNA can lead to increased numbers of mutant copies, possibly pushing the cell over the bioenergetics threshold and causing cellular dysfunction. When it comes to generational shifts in mtDNA heteroplasmy, major changes are brought about as a result of the germline bottleneck which drives radical heteroplasmy shifts. Consequently, oocytes that possesses high levels of mutant mtDNA will, when fertilised, generate organisms that carry high levels of heteroplasmy in all of their tissues. In addition to inherited heteroplasmic mutations, *de novo* mutagenesis can result in the introduction of novel mutation in a cell and by extension a tissue. Finally, while the bioenergetics impact of high levels of pathogenic mtDNA mutations has been evident for years, there is increasing evidence that suggests low level heteroplasmy could also have physiological impact on a cell. Figure adapted from: Stewart and Chinnery, 2021.

1.4.4. Random Drift and the Mitochondrial DNA Bottleneck

In an attempt to explain the rapid shift in mtDNA heteroplasmy first observed in Holstein cows, it was proposed that a rapid reduction in the overall mtDNA copy number takes place within the germline (Jenuth et al., 1996). Termed the 'bottleneck theory,' such a drastic reduction in copy number would result in sampling a subpopulation of the mtDNA present at a given point within the cell and thus bring about radical changes in the offspring's heteroplasmy (Jenuth et al., 1996). Subsequent experiments studying mouse oogenesis verified that hypothesis and demonstrated that at the point of fertilization, the oocyte contains around 150,000 copies of mtDNA, while primordial germ cells (PGCs), the stem cells responsible for giving birth to the gametes, only have about 200 copies of mtDNA (Cree et al., 2008; Wai et al., 2008). This result was further verified by mathematical modelling that showcased shifts in mtDNA copy number through germ cell development (Cree et al., 2008). Studies in other organisms have corroborated these findings. Work on zebrafish (*Danio rerio*) described a PGC mtDNA copy number of 170 that would eventually result in mature oozytes that hosted 10^7 copies of mtDNA each (Otten et al., 2016). A similar pattern was observed in sheep (*Ovis aries*), where the mtDNA copy number within primordial follicles (a developmental stage between PGCs and the mature oocyte) was found to be 605 compared to the almost 750,000 copies that were present within mature oocytes (Cotterill et al., 2013), thus further reinforcing the concept of a developmental bottleneck within the germline being the driving force behind abrupt shifts in heteroplasmy across generations. Fruit fly oocyte development despite following a different path when it comes during germ cell specification also showed a rapid increase in mtDNA copy number while transitioning from PGCs, having about 1000 copies per cell, to mature oocytes, with around 10^5 copies (Hurd et al., 2015). Overall, the body of research outlined above, gives credence to the bottleneck theory of germ cell development as the explanation behind the drastic shifts in mtDNA heteroplasmy across generations. This rapid decrease in mtDNA copy number observed at the same stage of oocyte development across so many different species was justified as being a consequence of a reduction in mtDNA replication (Piko and Taylor, 1987; Aiken et al., 2008).

It is important to note that the evidence of a decrease in copy number as being behind the driving force behind the bottleneck theory has not been left unchallenged. An alternative suggestion was proposed by Liqin Cao and his colleagues, who postulated that the lower mtDNA copy number seen at the PGC stage is not a unique phenomenon and indeed, the copy numbers seen at that stage of oocyte development even exceed those observed in many mature somatic cells (Cao et al., 2007; Cao et al., 2009). Instead, they proposed that the driving force behind the germline bottleneck is the separation of the totality of mtDNA within the cell into segregation units and this process of compartmentalisation is what leads to the drastic shift in heteroplasmy from mother to offspring (Cao et al., 2007; Cao et al., 2009). Further scrutiny of the initial hypothesis that the mtDNA bottleneck takes place as a result of a rapid reduction in copy number, came from the Shoubridge lab. They were also able to verify the striking differences in copy number between PGCs and mature oocytes but when measuring heteroplasmy, they found that it remained unchanged (Wai et al., 2008). Interestingly, the point during oocyte development when they were able to detect changes in heteroplasmy was folliculogenesis, the process responsible for the maturation of the ovarian follicle (Wai et al., 2008). As a result, a third theory was given birth to. This time proposing that the drastic shift in mtDNA heteroplasmy observed from mother to offspring is a result of the selective replication of only a subset of the mtDNA molecules present within a PGC (Wai et al., 2008). Despite being a very closely studied field, at the time there is no clear consensus behind the mechanism driving mtDNA bottleneck in the germline. However, the three theories discussed here are not mutually exclusive, and indeed, it is likely that the rapid and stochastic mtDNA heteroplasmy shift from mother to offspring, comes as a result of a combination of these mechanisms **(Figure 1.13.)**.

The principles behind the germline mtDNA bottleneck can potentially play a central role when it comes to somatic cell development. The presence of a somatic bottleneck could result in major differences between different cell types and if this takes place early on in development, could result in tissues and organs within an organism exhibiting strikingly different heteroplasmy percentages (Zhang et al., 2018). Such a behaviour in mtDNA inheritance could be the reason behind the vastly different clinical presentations that are a characteristic of mitochondrial diseases, even between members of the same family (Stewart and Chinnery, 2015). Consistent with the theory of a somatic developmental bottleneck, a rapid reduction in mtDNA copy number is not only present within the germline in mice (Cao et al., 2007) and there is mounting evidence that the same can be true in humans (Floros et al., 2018). Patients

treated with nucleoside reverse transcriptase inhibitors (NRTIs) could also exhibit a similar pattern of accelerated selection of a subset of mtDNA, as NRTIs lead to a transient reduction depletion of the mtDNA (Arnaudo et al., 1991). This observation opens the door to the possibility that environmental factors that can reduce mtDNA copy number might play a role in promoting the proliferation of mtDNA molecules that carry pathogenic polymorphisms thus, disproportionately affecting the mutation burden of certain tissues or cell types during life (Zhang et al., 2018).

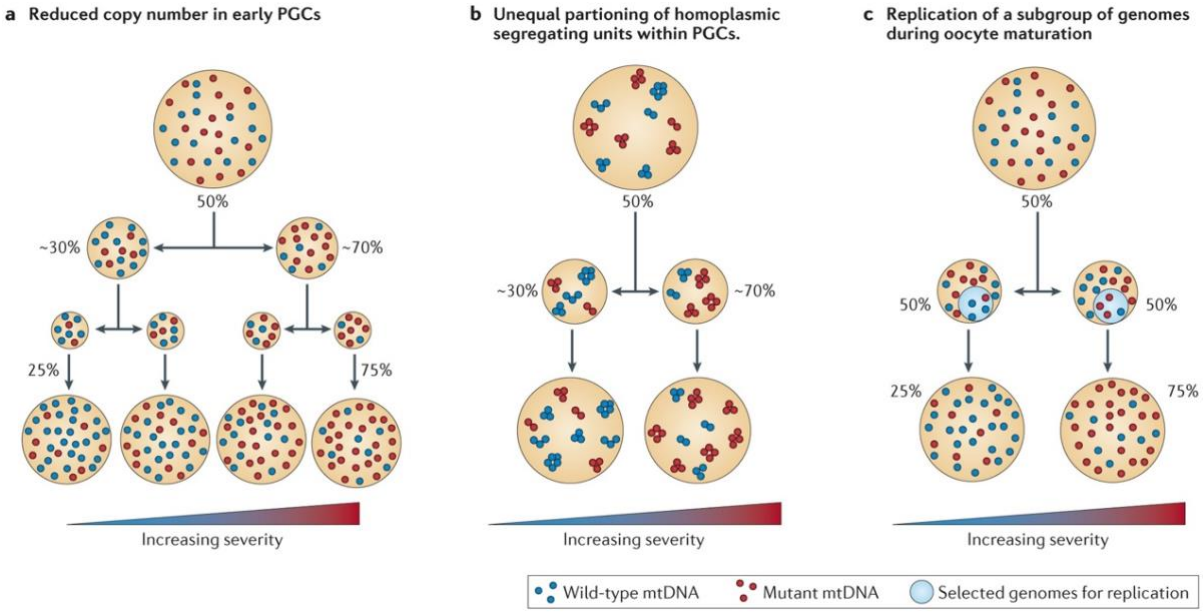


Figure 1.13. Generation of the mitochondrial DNA genetic bottleneck.
Schematic representation showing the different theories behind the generation of a mitochondrial DNA (mtDNA) genetic bottleneck within the germline. **(a)** The smaller the mtDNA content of a cell, the faster the rate of vegetative segregation is. Any resulting inequalities in the copy number of mutant (red) vs wild-type (WT) (blue) mtDNA that take place at the stage of the primordial germ cell (PGC) will inevitably result in heteroplasmy discrepancies within mature oocytes. **(b)** The segregation of mitochondria within the mitochondrial network into mutant and WT clusters, will reduce the number of units being segregated and thus, accelerate the rate of heteroplasmic drift. **(c)** Selective replication of a subpopulation of mtDNA molecules during the post-natal phase of oocyte maturation. Figure adapted from: Stewart and Chinnery, 2015.

1.4.5. Mechanisms Behind Mitochondrial DNA Selection

As discussed so far, there is a plethora of evidence and hypotheses in the literature surounding the principles and mechanisms governing the random genetic drift of mtDNA both within at outside the germline. Similar studies have been trying for years to establish what might be the underlying forces driving for selection against deleterious mutations that would lead to their extinction from the germline.

In order to investigate the behaviour of mtDNA mutations of differing severity during germline transmission, Weiwei Fan and his colleagues generated a mouse model that carried two pathogenic mutant haplotypes through the transplantation of embryonic stem cells (ESCs) into a blastocyst (Fan et al., 2008). The first variant (m.13885insC) was severe and affected mt-Nd6 while the second, was a milder variant (m.6589T>C) affecting *mt-COX1* (Fan et al., 2008). The authors demonstrated that the more deleterious, mt-ND6 mutation, was actively and rapidly selected against within the germline and eliminated by the fourth generation (Fan et al., 2008). On the other hand, the milder mt-Cytc mutant remained within the mouse mitochondrial genome for many more generations, even though there was evidence suggesting it was the cause of mitochondrial myopathy and cardiomyopathy to the offspring it was being inherited to (Fan et al., 2008). Even though the direct mechanism of selection remained unclear, the authors concluded that selection against mtDNA mutations in the germline is largely based upon the severity of the mutation at hand, as less severe mutations seem to be governed by milder selective pressures. In this case, the mutation leading to the higher levels of ROS production (m.13885insC) could have led to increased apoptosis of proto-oocytes that at this stage of development undergo quality control that results with only 30% of them completing the process of meiotic maturation (Hussein, 2005; Tilly and Tilly, 1995). In a similar endeavour, the maternal transmission of nonsynonymous mutations generated in mtDNA mutator mouse were examined (Stewart et al., 2008). This mouse model possessed a proofreading-deficient POLγ that resulted in the increased rate of incorporation of mtDNA mutations (Trifunovic et al., 2004; Kujoth et al., 2005). It was observed that purifying selection was taking place within the germline as nonsynonymous mutations harboured within protein-coding genes of the mtDNA were promptly selected-against (Stewart et al., 2008). The take-home message presented in these two studies is very similar as mutations that are having the greater impact on the organelle's capacity to perform it core function, that of energy

production, appear to be removed from the germline. The use of mitochondrial-targeted restriction enzymes in *Drosophila melanogaster* aimed at introducing nonsynonymous thermos-sensitive mutations in mt-ND2 and mt-Cytc gave further credence to this hypothesis as an increase in temperature led to the introduction of the mutations and a decrease in OXPHOS (Ma et al., 2014). As a result, these now pathogenic haplotypes, were actively selected against and prevented from passing down the germline (Ma et al., 2014).

Alternative hypotheses make the case for heteroplasmic mtDNA variants leading to transcriptional responses that in turn drive variant selection. Backcrossing of the mtDNA mutator mouse led to the generation of strains that carry mtDNA mutations. Over several generations of female mtDNA-mutator mice backcrossing, a novel mouse model was generated, carrying two heteroplasmic mutations on mt-tRNA^{met} (m.3875delC) and mt-tRNA^{cys} (m.5245T>C) (Freyer et al., 2012). The two mutations behaved quite differently, with the mt-tRNA^{cys} mutant being driven towards homoplasmy while exhibiting no signs of mitochondrial disorder, while the mt-tRNA^{met} mutant was conserved in the germline and selected against later on, during somatic tissue development (Freyer et al., 2012). At this point it should be mentioned that the contribution of the D257A PolgA exo− in generating this model meant that other mutations linked to the two pathogenic ones underlined by the authors were present within the mtDNA. The authors hypothesised that a transcriptional compensatory response was mounted by the cell in reaction to the decrease in OXPHOS capacity (Freyer et al., 2012). In that way, the cells, and by extension, the mutation, was able to escape negative selection (Freyer et al., 2012). Further credence to the concept of transcriptional responses to mtDNA heteroplasmy was provided by a study carried out on *C. elegans* carrying multiple mtDNA deletions, where evidence was provided of the mitochondrial unfolded protein response (UPR^{mt}) being triggered by the stress activated transcription factor 1 (ATFS-1), leading to the preservation of the mtDNA deletion levels (Lin et al., 2016).

Backcrossing of the mutator mouse has given birth to another mouse model, this one carrying a mutation of the mt-tRNA^{Ala} (mt-Ta) (m.5024C>T) (Kauppila et al., 2016). Male mice carrying this mutation exhibited reduced fat content, lean mass and total body mass when compared to their control counterparts (Kauppila et al., 2016). Both males and females were characterised by increased heart mass, especially at higher heteroplasmy levels (Kauppila et al., 2016). Furthermore, the use of COX/SDH staining revealed the presence of COX-deficient cells within the colonic crypt epithelium of younger mice and extended to the colonic smooth

muscle and cardiomyocytes of 40-week-old mice that exceeded 60% heteroplasmy (Kauppila et al., 2016).

Overall, this model does not exhibit strong selection against the mutation, but neither it is neutrally transmitted as a strong upper threshold at 80% means that no mice exceeding this threshold being born (Kauppila et al., 2016). Paradoxically, further work on this model, revealed that heteroplasmy segregation down the germline is more complicated than it first appears, as the mutation seems to be positively selected in mother-pup pairs when the mother heteroplasmy is below 55%, while at the same time, selected against when maternal heteroplasmy exceeds 60% **(Figure 1.14.)** (Zhang et al., 2021). Moreover, reinforcing the metabolic compensation theory surrounding mitochondrial activity and the concept of a disease threshold that is in play in mitochondrial dysfunctions, only when the m.5024C>T mutation exceeded 45% heteroplasmy noticeable changes in mt-Ta steady state levels begun to appear, a result that was in line with the COX deficiency observed in the colonic crypts of the same animal model (Kauppila et al., 2016).

Taken together, these models present the first attempts made to understand the mechanisms behind mtDNA selection. It is clear that as a field, mtDNA heteroplasmy dynamics still harbours a lot of unanswered questions as publications paint a very complex picture. At the moment, one question seems to be the most pressing, since the transmission of mtDNA is mutation-dependent, what forces dictate which mutations are selected against and which are tolerated, thus either giving directionality to the heteroplasmy shift or leaving it to follow the principles of random genetic drift?

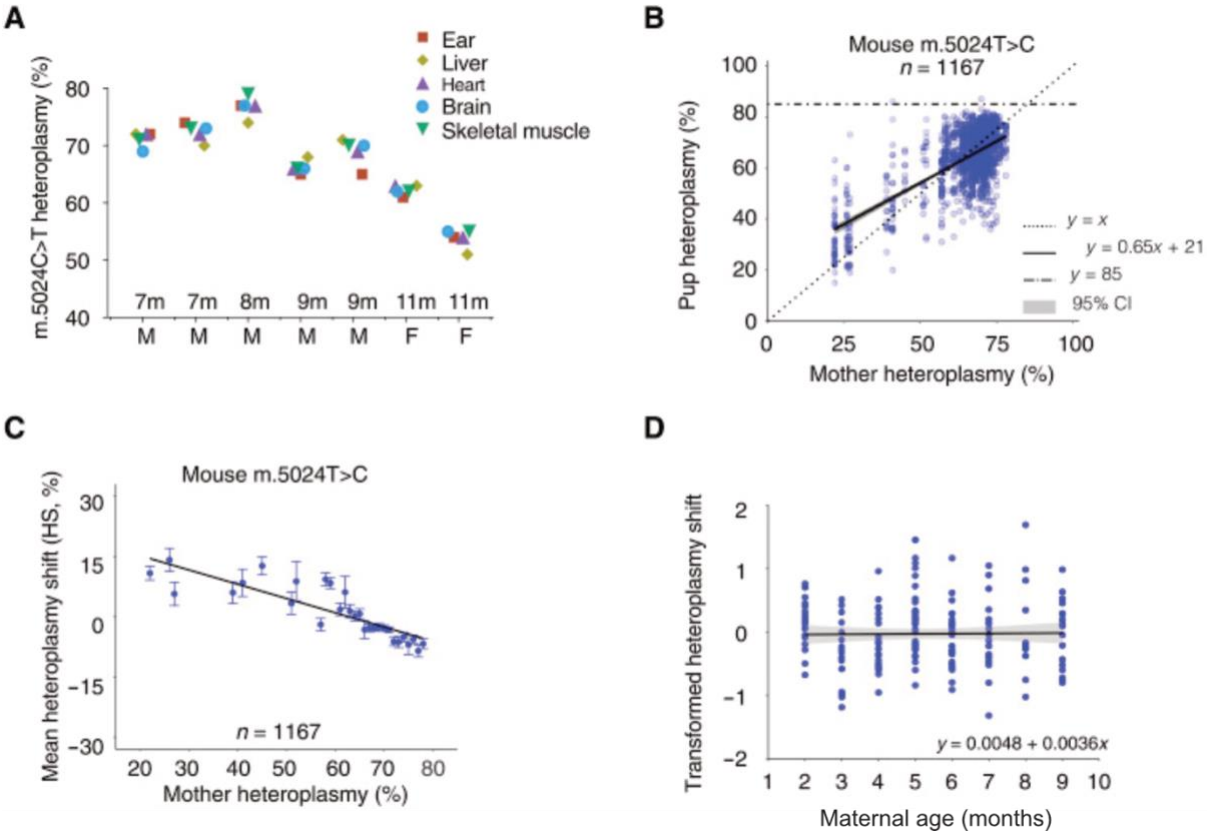


Figure 1.14. Evidence in support of both positive and negative selection of the m.5024C>T mutation during transmission.

Mitochondrial DNA (mtDNA) heteroplasmy measurements in mice carrying the Blimp1-mVenus and stella-ECFP (BVSC) reporters and the m.5024C>T $tRNA^{Ala}$ mutation. **(A)** Heteroplasmy levels of m.5024C>T $tRNA^{Ala}$ were measured in the ear, liver, heart, brain and skeletal muscle of male (M) and female (F) 7,8,9 and 11 months (m)-old mice. **(B)** Heteroplasmy transmission from 49 mothers to 1167 offspring. y = x signifies transmission of heteroplasmy from mother to pup in a 1:1 ratio. The line of best fit is represented by: y = mx + c which corresponds to the data within the 95% confidence interval for the regression line (highlighted in grey). The horizontal line at y=85% represents the upper heteroplasmy threshold, with no animals being born exceeding that value. Error bars represent SEM. **(C)** The heteroplasmy shift (difference in m.5024C>T $tRNA^{Ala}$ heteroplasmy levels) between the mothers and pups introduced in **(B)** While mothers with lower heteroplasmy levels tend to produce offspring carry higher heteroplasmy, pups that come from high heteroplasmy mothers have themselves lower heteroplasmy levels compared to their mothers. **(D)** Correlation between maternal age at the time of birth and transformed heteroplasmy shift of the offspring. The line of best fit (y = mx + c) for the data within the 95% confidence interval for the regression line (highlighted in grey, P = 0.5787) shows that maternal age has no effect on the average transformed offspring heteroplasmy shift. Mothers (n = 15); offspring (n = 17); litters (n = 33); available maternal ages are shown in months (m). Figure adapted from: Zhang et al., 2021.

1.5. The Cell Cycle

It was in 1953, the same year the structure of the DNA was discovered, that Alma Howard and Stephen Pelc were the first scientists to provide a timeframe for cellular lifespan (Plec and Howard, 1955). Studying cell proliferation in the roots of *Vicia faba*, they described a phase when cell division takes place, the pre-S-phase (also termed growth 1 (G_1-phase)), the S-phase during which DNA synthesis takes place and the pre-mitotic phase, or G_2 (Howard and Pelc, 1986). Their observations marked the birth of the cell cycle concept as we know it today.

In its entirety, the cell cycle is a tightly regulated process which ensures that genetic information stored within the cell's DNA, is successfully duplicated and passed down to the next generation of cells. It is crucial that during the cell cycle, nDNA is faithfully replicated, in order for daughter cells to acquire an exact copy of the parent cell's genetic material. As described by Howard and Pelc, the cell cycle is divided into three discrete intervals. First, while within the G_1 phase, the cell responds to the sensing of extracellular growth factors and mitogens (Israels and Israels, 2001). Having been stimulated, the cell goes into S-phase, characterised mainly by the process of DNA synthesis (Israels and Israels, 2001). Finally, growth 2 (G_2-phase) is the phase that precedes cell division of mitosis (M-phase) (Israels and Israels, 2001) where an increase in cell size takes place in anticipation of cell division. M-phase is characterised by the formation of the mitotic spindle in a bipolar manner, leading to the separation of sister chromatids to the two centrosomes immediately prior to cell division (Karsenti and Vernos, 2001). Involving a series of steps that are all necessary for timely and regulated cell proliferation, the cell cycle is subject to a series of monitoring junctions, also known as checkpoints. These checkpoints are in place in order to make sure that the cell has successfully completed each cell cycle phase before moving to the next. As a result, checkpoints that guarantee cell integrity are positioned towards the end of G_1-phase and in the transition between G_2/M-phase (Murray et al., 1994). This prevents growth and propagation of damaged or mutated cells. Quiescent cells that have temporarily or permanently had their cell cycles suspended are referred to as existing in Growth 0-phase (G_0-phase). The transition in and out of G_0 in the average cell takes place in early G_1 and is dependent upon extracellular stimuli in the form of growth factors and mitogens (Israels and Israels, 2001).

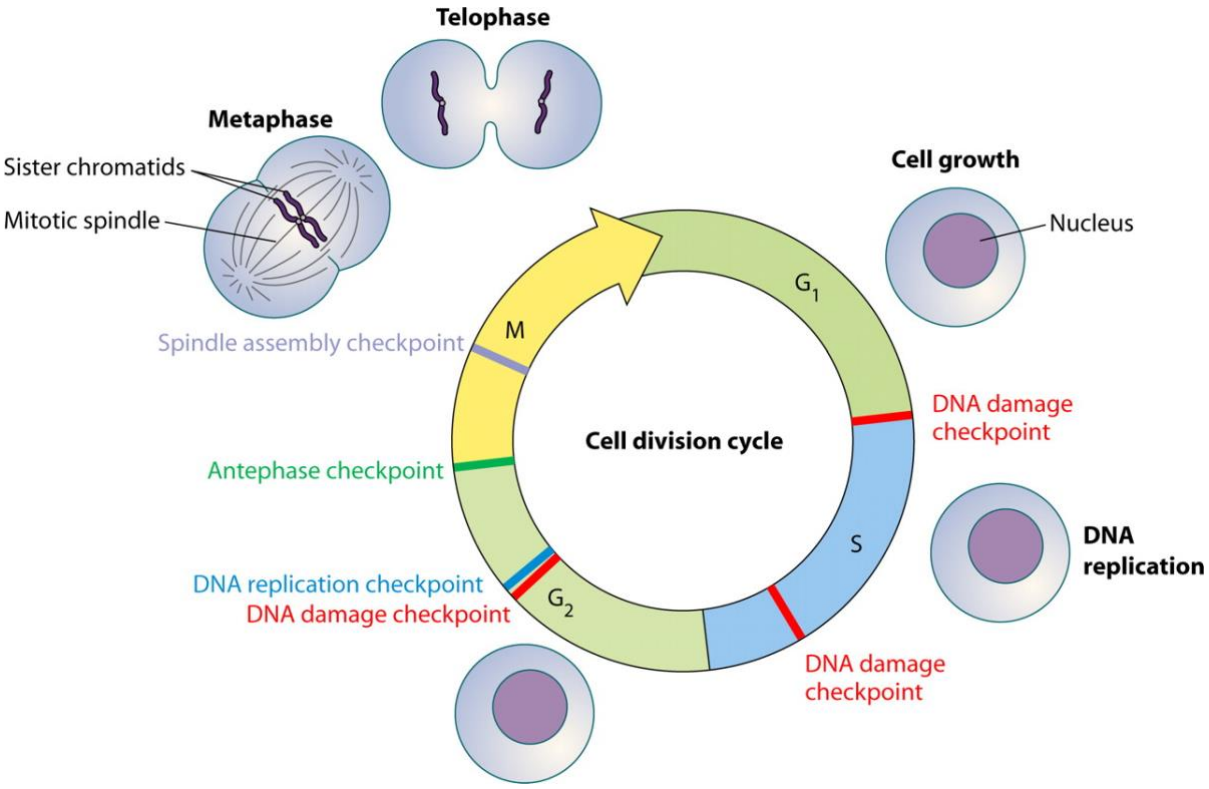


Figure 1.16. Cell cycle checkpoints.
Schematic representation of the four cell cycle steps and positioning of the various checkpoints along its length. There are multiple quality control checkpoints through the cell cycle that serve to ensure that the cell can proceed to the next cell cycle step and eventually move to mitosis. These checkpoints include DNA damage (depicted in *red*), DNA replication (depicted in *blue*) and a spindle formation (depicted in *grey*) that serves as the last roadblock prior to mitosis. Finally, the antephase checkpoint (depicted in *green*) makes sure that the cell is not under different types of stress prior to entering mitosis. Figure adapted from: Chin and Yeong, 2010.

Intracellular responses to external stimuli regulating the cell cycle, come in the form of cascading rounds of phosphorylation that lead to the upregulation of cyclin expression, in turn, associated with cyclin-dependent kinases (CDKs). CDKs are protein kinases that require a cyclin subunit in order to obtain their enzymatic activity thereby establishing cyclin as the regulatory component of the CDKs' catalytic properties (Ding et al., 2020). Among other contributions, they play a central role in the propagation of the cell cycle and division, with different CDKs being activated at different cell cycle stages and fluctuating as those stages come and go (Ding et al., 2020). There are 29 cyclins and 20 CDKs that have been identified in humans so far (Cao et al., 2014). Amongst them, the ones directly involved in cell cycle progression and cell division are CDK1-4, CDK6 and CDK7 while CDK7-11 are involved in the regulation of transcriptional responses (Ding et al., 2020). A great example of how CDKs are involved in the regulation of cell cycle progression is provided by CDK4 and CDK6. With the majority of cells found in adult tissues being arrested in the quiescence characterising the

G_0-phase, it is incumbent upon the mitogenic stimulation of CDK4 and CDK6 to trigger cell cycle progression out of G_0 or G_1-phase, and into the S-phase (Ding et al., 2020). CDK4 and 6 activity is promoted through their association with D-type cyclins, including cyclins D1-3, and conversely, inhibited through the binding of inhibitors belonging to the inhibitors of CDK4 (INK4) protein family, including $p15^{INKB}$, $p16^{INK4A}$, $p18^{INK4C}$, and $p19^{INK4D}$ (Malumbres and Barbacid, 2001). The formation of an activated complex containing cyclin D in addition to CDK4/6 leads to the phosphorylation of retinoblastoma (RB), a protein with tumour suppressing properties that is encoded by the *RB1* gene and plays a central role in the negative regulation of the cell cycle (Giacinti and Giordano, 2006). Sequential RB phosphorylation, leads to its inactivation and permits the cell to transition from G_1-phase to S-phase.

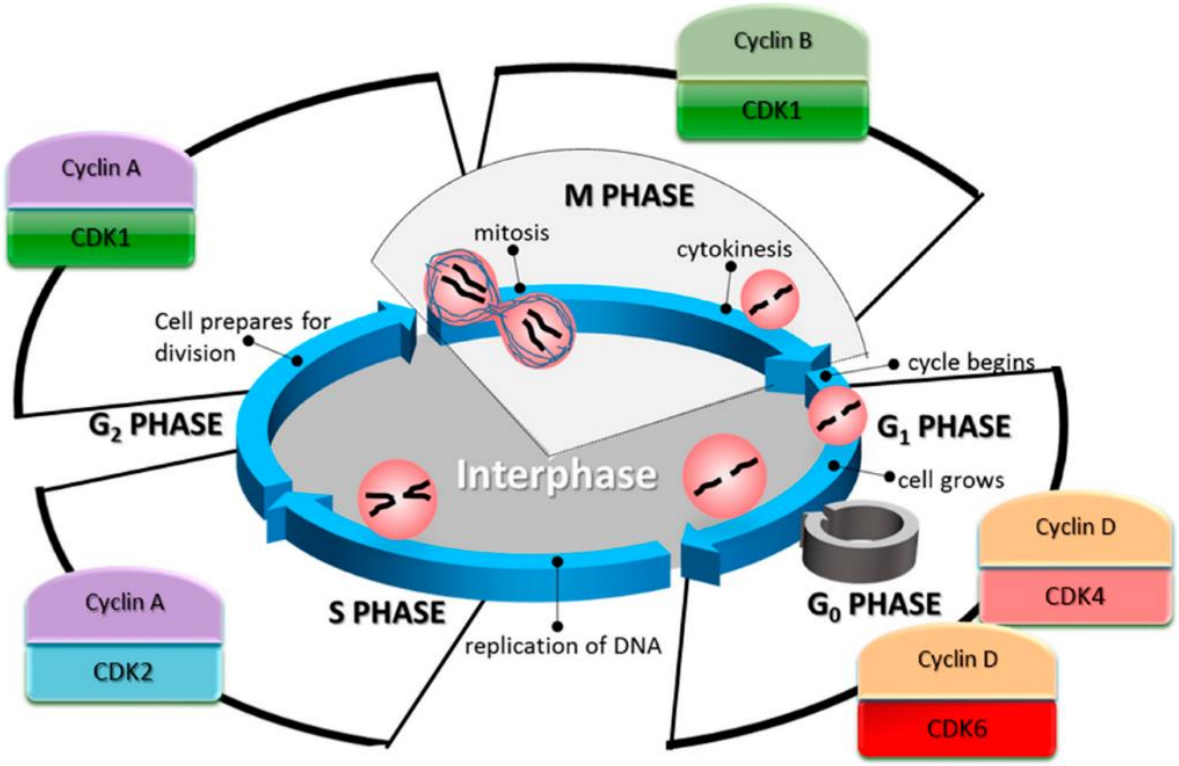


Figure 1.17. The role of Cyclin-Dependent Kinases in cell cycle progression.
Schematic representation of CDK (Cyclin-Dependent Kinase)/cyclin complexes and the cell cycle steps they regulate. Cyclin D activity is increased during the G_1 phase. Interacting with CDK4/6, it functions to promote cell exit from G_0 and subsequent entry and progression through G_1 as well as the G_1-S transition. Proteins involved in DNA replication during the S-phase get phosphorylated by the CDK2-cyclin A complex. This complex remains active until the late stages of G_2-phase where the primary cell cycle regulator becomes the CDK1-cyclin A complex. Figure adapted from: Garcia-Reyes et al., 2018.

So far, we have just scratched the surface of the complex molecular pathways that govern a cell's journey through the cell cycle up to the point of cytokinesis. However, it is clear that controlling the cell's fitness throughout this process is of paramount importance. This is the reason behind the presence of multiple checkpoints and a restriction point (R). The cell cycle's R, can be defined as the developmental stage during a cell's G_1-phase where growth factors are no longer required in order to drive the cell through the cell cycle (Israels and Israels, 2001). It has been experimentally demonstrated that starving cells of serum before they have reached R, leads to them entering G_0, while subjecting them to the same conditions after they have gone through R, has no effect on their progression through the rest of the cell cycle (Pardee, 1974). Further down the cell cycle exist other quality control checkpoints that also ensure the orderly sequencing of cell cycle steps (Hartwell and Weinert, 1989). These checkpoints respond to nDNA damage at either the G_1-S-phase or the G_2-M-phase transition. In case where nDNA damage is detected, the cell cycle will be arrested either at G_1-phase or G_2-phase and the cell will be given the chance to correct its damaged nDNA and proceed to the next cell cycle phase (You and Bailis, 2010). Key regulators of nDNA damage signaling during these checkpoints include ataxia telangiectasia mutated (ATM) protein, ataxia telangiectasia and Rad3-related (ATR) protein, phosphatidylinositol 3-kinase (PI3K)-like protein kinases (PI3KKs) and the checkpoint kinases 1 and 2 (CHK1 and 2) (Zhao and Piwnica-Worms, 2001). ATM and ATR are responsible for responding to damaged nDNA by phosphorylating and activating CHK2 and CHK1 respectively (Santo et al., 2015). The activated for of CHK2 is in turn involved in the activation of tumour protein 53 (p53) which leads to the p53-dependent arrest of the cell cycle at early G_1-phase (Matsuoka et al., 2000). Taken together, the presence of a restriction point and checkpoints guarantees that any given cell is equipped with both the necessary nutritional load and unimpaired genetic information necessary for the generation of functional daughter cells.

1.7. Aims

Inherited mtDNA disorders affect approximately 1:5000 humans (Zeviani and Carelli, 2003). While mutations in 300 nuclear genes are implicated in mitochondrial disorders (Abicht et al., 2018), the majority are caused by mutations that impact the multi-copied mtDNA (Gorman et al., 2015). Most mtDNA mutations exist in a heteroplasmic state (Gorman et al., 2016) that describes the coexistence of mutant mtDNA molecules alongside WT counterparts, within the confines of a single cell (Schon et al., 2020). As the percentage of heteroplasmy increases, so does the proportion of mutant to WT mtDNA molecules. Once a certain heteroplasmy threshold is exceeded, decrease in OXPHOS activity and dysregulation of mitochondrial processes bring forth dysfunction and disease phenotype. The complicated genetic nature characterising mitochondrial disorders leads to cellular responses to mtDNA mutations are both mutation and cell-type specific, creating a landscape that has yet to be completely understood. My primary aim in this dissertation is:

To investigate the principles and mechanisms governing mtDNA heteroplasmy segregation by examining its pattern and impact on models of mitochondrial dysfunction.

In order to accomplish this aim, I set out to work on different projects, each addressing a distinct but related aspect of mtDNA heteroplasmy:

Chapter 3: Optimisation and validation of a pyrosequencing protocol for carrying out accurate high-throughput mtDNA heteroplasmy measurements at the single cell level.

Chapter 4: Uncovering the dynamics governing mtDNA heteroplasmy segregation at the single cell level throughout mouse development and ageing.

Chapter 5: Understand the severity and significance of embryo-wide as well as cell-lineage specific transcriptomic responses that are brought about during murine organogenesis as a result of pathogenic mtDNA mutations.

Chapter 6: Determining the extent to which mitochondrial activity influences cell cycle progression.

Chapter 2. Materials and Methods

2.2. Mouse Models and related procedures

2.2.1. Colony Maintenance

In this study, three different mouse lines were used, the heteroplasmic m.5024C>T and m.5019A>G in addition to the WT control. All three mouse lines used in this study were bred on and maintained on a C57BL/6 nuclear background. At this stage it should be noted that for all experiments that are part of this thesis, with the exception of the liver and cortex proteomics analysis (**Results Chapter, Section 5**), the substrain of the animals was not examined in a more detailed manner. Consequently, in the case of both the m.5019A>G and m.5024C>T colonies, mice should be thought of as being either homoplasmic for C57BL/6J/N or heterozygrous unless specified otherwise. C57BL/6J is the most common inbred mouse straina and the first one to have its entire genome sequenced, it differs from its C57BL/6N conunterpart due to to presence of 34 SNPs and 2 indels present on coding genes, in addition to 15 structural variants overlapping a gene (Ronchi et al., 2013). Most notably for expetriments examining mitochondrial activity, there are two mutations on the nicotinamide (NAD) nucleotide transhydrogenase (*Nnt*) gene, which codes for a mitochondrial protein catalysing the conversion of NAD^+ to NADPH (Ronchi et al., 2013). One should take into account that these mutations, the first, a methionine-to-threonine missense at the MTS and the second, an in-frame deletion that leads to the removal of exons 7-11, are present in the C57BL/6J strain but not the C57BL/6N (Ronchi et al., 2013). All experimentation performed on the animals was carried out in accordance with the UK Animals (Scientific Procedures) Act of 1986, under the Home Office Project Licence (PPL): P6C97520A, as well as the EU Directive 2010/63/EU. Additionally, all experiments were carried out following the Animal research: reporting of *in vivo* experiments (ARRIVE) guidelines following approval by the University of Cambridge animal welfare ethical review body (AWERB).

Mice were maintained at Tecniplast cages that are individually ventilated and maintained at temperatures ranging between 20 and 24ºC, and humidity levels between 45 and 65%. All animals were subject to a 12-hr light/dark cycle and were provided with water and SAFE 105 universal diet (Safe diets) that they had access to *ad libitum*. When planning timed mattings, breeding females were chosen to be between 8 and 12 weeks of age. In the case of males used in timed mattings, they were first placed in empty cages to be given time to acclimatise before

being moved in with the female. Subsequently, females were checked for vaginal plug each morning and noon, with the day when the plug was found being designated as E0.5. Mouse culling was carried out using schedule 1 procedures, either through cervical dislocation or by exposure to CO_2. Death was then confirmed though cessation of circulation.

The m.5019A>G mutation was first identified and characterised at the Max Plank Institute of Biology and Ageing. The entirety of animal work carried out there, was done so in accordance with the guidelines and recommendations set out by the Federation of European laboratory animal science associations (FELASA). All experimentation carried out at the Max Plank was permitted and approved by the Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen according to regulations set out by Germany and the European Union.

2.2.2. Mouse Tissue Dissociation

Upon culling the mouse using the schedule 1 procedures described above, tissues that were meant for dissociation, namely brain, spleen and the uterine horns containing the embryos, were placed within Eppendorf tubes containing Leibovitz's L-15 Medium (Sigma) + 2% foetal bovine serum (FBS) (Gibco) and then on ice to be transported back to the lab.

2.2.2.1. Spleen

Once at the lab, an adult mouse spleen was cut into two pieces and washed with phosphate-buffered saline (PBS) (Gibco™). Then, the two pieces were placed in a plastic 6cm tissue culture (TC) dish (ThermoFisher Scientific) containing PBS and were minced mechanically using the flat ended thumb-rest of a syringe plunger. This process was repeated until no pieces of tissue were visible. The resulting cell suspension was then filtered through a CellTrics® 50µm cell strainer (sysmex) into a 15ml tube placed in ice in order to get rid of tissue remnants and cell clumps. The filtrate was then spun down at 300xg for 10min at 4ºC in order for the cells to precipitate and be separated from lighter debris. After being placed back on ice, the supernatant was aspirated and the remaining cells were once again suspended in solution. This time however, in order to get rid of the substantial population of red blood cells residing in the spleen, 100µl of Red Blood Cell (RBC) Lysis Solution (10x) (Miltenyi Biotec, PN: 130-094-183) diluted 1:10 in distilled but not deionised water was used. Once the cell pellet had been

thoroughly re-suspended, the tube was placed in the fridge (at 4ºC) for 10min. Subsequently, 10ml of freshly prepared PB buffer (D-PBS (Gibco™) + 0.5% bovine serum albumin (BSA) (PAA)) were added on top of the cells, which were then spun again at 300g for 10min at 4ºC. The resulting cell pellet was then re-suspended in 500μl of PBS and placed on ice where it awaited staining in preparation for fluorescence-activated cell sorting (FACS).

2.2.2.2. Brain

The dissociation process involved in neonate mouse brain, where the mouse is less than a week old (<P6) differs from that of an older mouse, due to the difference in myelin content within the tissue. As a result, when neonate mouse brains were dissociated using the Neural Tissue Dissociation Kit (P) (Miltenyi Biotec, PN: 130-092-628). Once the tissue was brought back into the lab it was washed with PBS and then placed in a gentleMACS™ tube (Miltenyi Biotec) containing papain enzyme mix, pre-warmed for 10-15min at 37ºC. Then, the gentleMACS tube (Miltenyi Biotec) was mounted on the gentleMACS™ Dissociator (Miltenyi Biotec) and program m_brain_01 was selected. Following the first run on the gentleMACS™ Dissociator (Miltenyi Biotec), a 15min incubation within a shaking incubator, set at 37ºC was performed. Right after the first incubation, two more spinning steps were carried out using gentleMACS™ Dissociator (Miltenyi Biotec) programs m_brain_02 and m_brain_03. In-between those spins, 10min incubations within the shaking incubator set at 37ºC were carried out. Following the last incubation, the sample was briefly centrifuged for 7min at 300xg and 4ºC, in order for the supernatant to be aspirated. The pellet was then re-suspended in 10ml of Hank's balanced salt solution (HBSS) (Gibco™) and filtered through a 70um MACS Smart Strainer (Miltenyi Biotec). The strainer was then discarded and the newly filtered sample was spun at 300xg for 10min at room temperature (RT), before the supernatant was once again aspirated completely. Finally, the cells were re-suspended in 500μl of D-PBS (Gibco™) and placed on ice, prior to antibody staining.

In the case of adult mouse brains, once the tissue was brought into the lab, it was cut into smaller pieces and washed using PBS. Then, the protocol outlined in the Adult Brain Dissociation Kit mouse and rat (Miltenyi Biotec, PN: 130-107-677), was carried out after being subjected to slight modifications and adjustments in order to accommodate for the lack of the gentleMACS™ Octo Dissociator (Miltenyi Biotec) machine in our lab. Instead of the slow and

steady 37-ABDK-01 spin recommended by the protocol that lasts for 1hr and takes place at 37ºC, the 3 spins outlined in the neonate mouse brain were carried out along with the incubation steps that followed each one. Having borrowed the gentleMACS™ Octo Dissociator machine from Miltenyi Biotec to carry out test runs, we were able to compare the eventual proportion of live cells derived using either technique, and was confident that the results we got with the gentleMACS™ Dissociator (Miltenyi Biotec) machine was of equal, if not better, quality **(Figure 2.2.)**. Following the completion of all 3 runs, the cell suspension was placed on ice and spun at 300g for 10min at 4ºC, in order for cells to precipitate and the separate from lighter debris. The tube was then placed back on ice and the supernatant was discarded. The cell pellet was resuspended in D-PBS and filtered through a 70µm MACS SmartStrainer (Miltenyi Biotec) into a 15ml tube, in order for tissue pieces and cell clumps to be removed. Another spin at 300g for 10min at 4ºC then followed resulting once again in the cells precipitating at the bottom of the tube. After being placed on ice and having the supernatant drained, the next step in the dissociation of the adult mouse brain, involved the removal of myelin. To that effect, the cell pellet was thoroughly re-suspended in 900µl of Debris Removal Solution (Miltenyi Biotec, PN: 130-109-398) mixed with 3100µl of D-PBS (Gibco™), while and additional 4ml of D-PBS (Gibco™) was overlaid, thus creating two distinct layers. When spun at maximum speed (4255xg) and acceleration/deceleration for 10min, using an Allegra X-30R centrifuge (Beckman Coulter). Myelin that previously resided within the cell pellet was isolated as part of the interphase separating the two layers and was aspirated along with the supernatant. Subsequently, the remaining cell pellet underwent the red blood cell lysis step as it was res-uspended in 100µl RBC Lysis Solution (10x) (Miltenyi Biotec, PN: 130-107-677) mixed in 900µl of distilled but not deionized H_2O. The suspension was then incubated in the fridge (4ºC) for 10min and then had 10ml of PB buffer added to it. A final spin at 300g for 10min at 4ºC followed and the resulting pellet was re-suspended in 500µl of D-PBS.

2.2.2.3. Embryonic Stage 8.5 (E8.5)

Once dissected and brought into the lab, mouse uterine horns containing the embryos were placed in a 3.5cm TC dish (ThermoFisher Scientific) containing PBS (Gibco™) and observed under a Leica M205 FA fluorescence dissecting microscope. The muscular wall of the uterine horn, myometrium, was then pulled apart using dissection forceps to reveal the oval embryo-containing structures termed *decidua* **(Figure 2.1A and B)**. Each *deciduum* was then carefully

split open and the embryo within isolated **(Figure 2.1C)**. At this stage it was important for embryos that were used for scRNA-seq to have the majority of the extraembryonic tissue removed in order to enrich the proportion of embryonic cell lineages captured by the 10x platform. Subsequently, the embryo was placed on a separate PBS-containing 3.5cm TC dish (ThermoFisher Scientific) where it waited for downstream processing. Once all embryos were dissected and pulled together in the same TC dish, a phenotypic examination was performed in order to make sure that all embryos matched the expected embryonic structure at E8.5. Any embryos that did not mirror and were either at a slightly earlier or later stage, were discarded. Each selected embryo was then placed within a 1.5ml Eppendorf that contained 200μl of 0.5% Trypsin- Ethylenediaminetetraacetic acid (EDTA) (10x) (Gibco^{TM}) diluted 1:10 in PBS (Gibco^{TM}). All Eppendorf tubes were then mounted on ThermoMixer F1.5 (Eppendorf) and left to incubate there for 10min at 37^{o}C and 800rpm. Subsequently, the Eppendorf tubes were taken off the heat-block and 400μl of PBS (Gibco) + 10% FBS (Gibco) were added to each one. The resulting suspension was mixed using a 200μl pipette, in order to apply some mechanical force that sped up the breakdown of intercellular matrix. The Eppendorf tubes were then placed on the heat-block for a final 5min incubation at 37^{o}C and 800rpm, which was followed by another round of pipetting to ensure that the final product was a single cell suspension. At this stage, the suspension was transferred and into a 5ml polystyrene round bottom tube with cell strainer cap 12x75mm style (Flacon) in preparation for flow sorting.

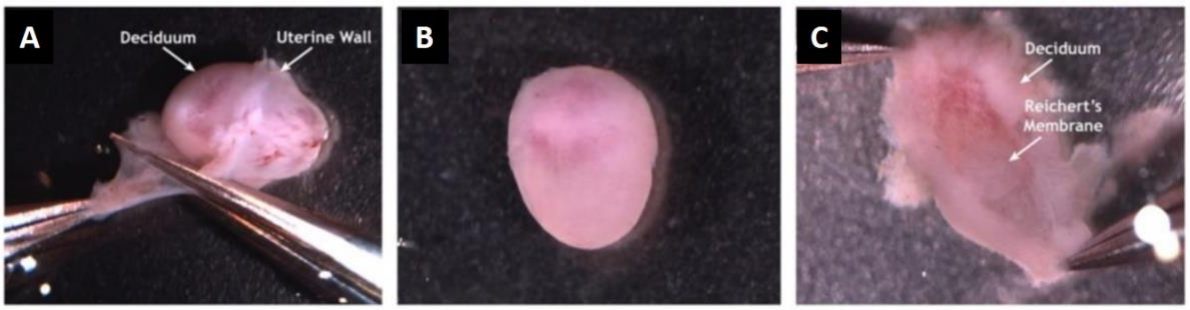


Figure 2.1. Mouse embryo dissection at E8.5.
(A and B), Peeling the uterine wall away and isolating the now clean deciduum. (**C),** Exposing the Reichert's membrane encapsulating the embryo proper by opening up the deciduum. Figure adapted from: Lopez et al., 2020.

2.3. Single Cell Analysis Approaches

2.3.1. Fluorescence-activated cell sorting

All experiments involving FACS were carried out using BD FACS Melody™ Cell Sorter (BD Biosiences) while the data analysis was performed using FlowJo software (BD Biosiences). Surface antibody staining in preparation for FACS analysis was performed in small volumes (50-200μl) at 4ºC, with the concentration of each cellular stain, antibody as well as related incubation times being displayed in **Table 2.1.**. Samples were placed in 5ml polystyrene round bottom tube with cell strainer cap 12x75mm style (Flacon) and kept on ice prior to being mounted on the flow cytometer and returned there once the run was complete. In every case, a sample tube was dedicated to contain an unstained version of the sample that served as the negative control against which all antibodies and dyes were compared, while another tube was stained with a combination of DRAQ5™ (ThermoFisher Scientific, PN: 62254) and DAPI (Invitrogen, PN: D21490) that helped separate intact, DNA-containing cells (positive for DRAQ5™ and negative for DAPI), from fragmented ones and debris.

Work aimed at isolating single cells from E8.5 embryos as well as tissues, was carried out with immediate storage at -80ºC in mind that made possible subsequent measurements of mtDNA heteroplasmy. In order to achieve this, FACS analysis was performed using a combination of DRAQ5™ and DAPI in addition to the antibodies of choice corresponding to each tissue. However, in cases where E8.5 embryo pools were destined for scRNAseq, immediately prior to flow sorting, embryos were stained with DRAQ7™ viability dye (ThermoFisher Scientific) that allowed for the exclusion of cells with fragmented cell membranes. Live cells that appeared negative for DRAQ7 were sorted into a 1.5ml Eppendorf tubes, containing 31μl of L-15 medium (Sigma) + 2% FBS (Gibco), using a BD FACS Melody™ Cell Sorter (BD Biosiences), for a final volume of 47μl. At the same time, 40 live single cells were sorted into a 996-well polymerase chain reaction (PCR) plate and immediately placed on dry ice before being stored at -80ºC.

Table 2.1. Dyes and antibodies used for sample staining in preparation for fluorescence-associated cell sorting.

Dye	Company	Product Number	Concentration (vol/vol)
DRAQ5™	ThermoFisher Scientific	62254	1:1000
DRAQ7™	ThermoFisher Scientific	D15106	1:1000
DAPI	Invitrogen	D21490	1:2.000
Antibody			
Prominin-1-PE-Vio770 anti-mouse	Miltenyi Biotec	130-102-153	1:200
GLAST (ACSA-1)-PE anti-human/mouse/rat	Miltenyi Biotec	130-118-344	1:200
CD19-FITC anti-mouse	abcam	ab86904	1:200
PSD95 anti-human/mouse/rat	abcam	ab18258	1:200
Briliant Violet 421™	BioLegend	406410	1:200
CD3-PE/Cy5® anti-mouse	abcam	ab25531	1:1000

2.3.2. 10x Chromium Single Cell RNA Sequencing

Once the intact and DNA-containing cells derived from E8.5 embryos were sorted into a 1.5ml Eppendorf tube, they were placed on ice and submitted to Cancer Research UK Cambridge Institute Genomics Core Facility, where scRNA-seq library preparation took place using the Chromium Single Cell 3′ Library & Gel Bead Kit v3, Chromium Chip B Kit and Chromium Single Cell 3' Reagent Kits v3 User Guide (Manual Part CG000183 Rev C; 10X Genomics). The 47µl cell suspensions were loaded into Chromium, the equipment responsible for the formation of single cell-containing, gel-bead emulsions. Reverse-transcription was then carried out using the newly barcoded RNA derived from each single cells, using a C1000 Touch Thermal cycler (BioRad). Reverse-transcription and all downstream steps were conducted following the manufacturer's specifications, without the application of any modifications, in order to generate the required single-cell libraries. Quality and quantity control of the resulting cDNA was performed using an Agilent TapeStation 4200 (High Sensitivity 5000 ScreenTape). After making sure that the samples were up to the required standards, 25% of the generated material was used for gene expression library preparation.

Gene expression library quality was confirmed using an Agilent TapeStation 4200 (High Sensitivity D1000 ScreenTape) that allowed for the quantification of library sizes, while a Qubit 4.0 Fluorometer (ThermoFisher Qubit™ dsDNA HS Assay Kit) was used in order to quantify the levels of dsDNA. Sample normalisation and pooling at equimolar concentrations followed and, in order to be certain about the final concentration, the library of pooled samples underwent qPCR prior to sequencing, using the KAPA Library Quantification Kit (Sigma) on a QuantStudio 6 Flex (ThermoFisher Scientific).

Initially, 2 samples were pooled together (each sample containing a single cell mix of 3 E8.5 embryos, derived either from a WT or a m.5024C>T mother) and subsequently sequenced on a 1 lane of SP flowcell. This was then followed by a second pool of 3 samples (each sample containing a single cell mix of 3 E8.5 embryos, derived either from a WT, m.5019A>G or m.5024C>T mother) that was also sequenced on a 1 lane of SP flowcell. Illumina NovaSeq6000 was the platform of choice for the sequencing, set up according to the following parameters: 28 bp, read 1; 8 bp, i7 index; and 91 bp, read 2.

2.3.3. Pyrosequencing

Pyrosequencing in this project was mainly carried out on single cells but in cases where bulk samples derived from mouse tissues or cultured cell lines were sequenced using this technique, DNA extraction preceded the pyrosequencing PCR step. DNA isolation from such samples was carried out using the DNeasy® Blood and Tissue Kit (QIAGEN, PN: 69506) and following the manufacturer's specifications. When single cells sorted in 96-well PCR plates were used for pyrosequencing it was important to spin them at maximum speed (2204xg) at 4ºC for 10min prior to carrying out the PCR step in order to make sure that every cell on the plate was positioned at the bottom of its respective well.

Once the samples to be pyrosequenced, either single cells or bulk DNA, were prepared, they were placed on ice while the PCR mix was being prepared. The impact of DNA contamination when dealing with single cell measurements and the importance of having a consistently contamination-free workspace were central to this project and are discussed in-depth in **Results Section Chapter 3.**. In order to achieve this, a TC hood equipped with UV light had been dedicated to PCR work. Prior to every experiment, the interior of the hood was thoroughly sprayed with 70% ethanol as well as a combination of DNAzap™ solutions 1 and 2 (Invitrogen, PN: 9891G and 9892G), in order to get rid of both living contaminants and lingering DNA fragments. At the same time, dedicated pipettes, tips and sterile tubes were also sprayed with DNAzap™ solutions 1 and 2 (Invitrogen, PN: 9891G and 9892G) prior to their placement inside the hood. Once everything was set in place, arranged so that the number of movements required to prepare the reaction mix was kept to a minimum, the hood cover was closed shut and the interior was placed under UV light for at least 30min. Subsequently, the pyrosequencing PCR master mix was put together using reagents of the PyroMark® PCR Kit (QIAGEN, PN: 978703), the volumes are outlined on **Table 2.2.**. In the case of single cell pyrosequencing, the PCR mix was pipetted directly on top of the cell within the 96-well PCR plate that was used during cell sorting. When bulk DNA samples were used, they were diluted down to 50ng/µl and 1µl was added to the reaction, this was subtracted from the total volume of water accounted for in the reaction mix. In both cases, the plate was then quickly sealed and spun for 2min at 300g in order to make sure that the PCR reaction mix was pulled down, at the bottom of each well. Once spun down, the plate was placed on a heatblock and the PCR was run following the steps presented in **Figure 2.3.**.

Once the pyrosequencing PCR was complete, the plate was spun for 2min at 300g and 4ºC in order to make sure that no reaction mix droplets remained lingering on the plate seal. Pyrosequencing of the single cell samples was performed using the PyroMark® Q48 Autoprep (QIAGEN). First, the appropriate program was selected, depending on the SNP of interest **(Figure 2.4.)**. Then, the required amount of sequencing reagents was loaded into the appropriate cartridges **(Table 2.3.)** and the Q48 was primed. 3µl of PyroMark® Q48 Magnetic Beads (QIAGEN, PN: 1096484) were then placed into each well of a 48-well pyrosequencing plate and mixed with 10µl of sample. The plate we then placed into the Q48 and the pyrosequencing was initiated. The distribution of samples described above, means that in order for an entire 96-well plate to be sequenced, two disks have to be mounted on Q48.

Table 2.2. Pyrosequencing PCR master mix reagents.

Pyrosequencing PCR Reagent	**Volume (µl)**	
	1rxn	**102rxns**
PyroMark PCR Master Mix	12.5	1275
Coral Load	2.5	255
Q-solution	5	510
$MgCl_2$	1.5	153
Forward Primer: 5019- /5biosg/TTCCACCCTAGCTATCATAAGC 5024- ATACTAGTCCGCGAGCCTTCAAA	0.05	5
Reverse Primer: 5019- CGTAGGTTTAATTCCTGCCAATCT 5024- /5Biosg/AATTGCAAATTCGAAGGTGTAGA	0.05	5
H_2O	2.5	350

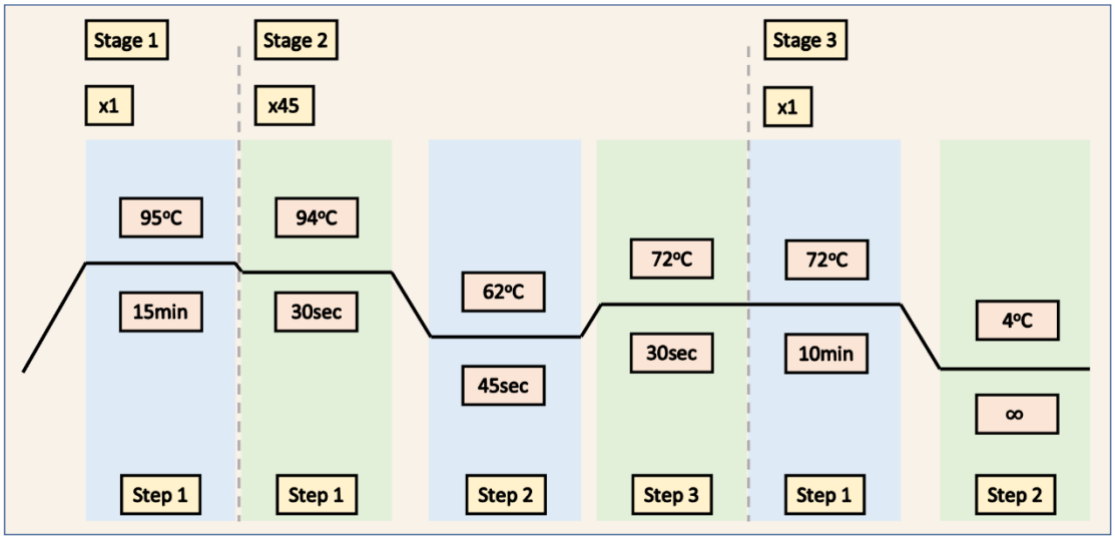


Figure 2.3. Pyrosequencing PCR conditions.
Schematic representation of the pyrosequencing PCR steps followed for the amplification of mtDNA fragments derived from either bulk DNA or single cell samples. Created with BioRender.com.

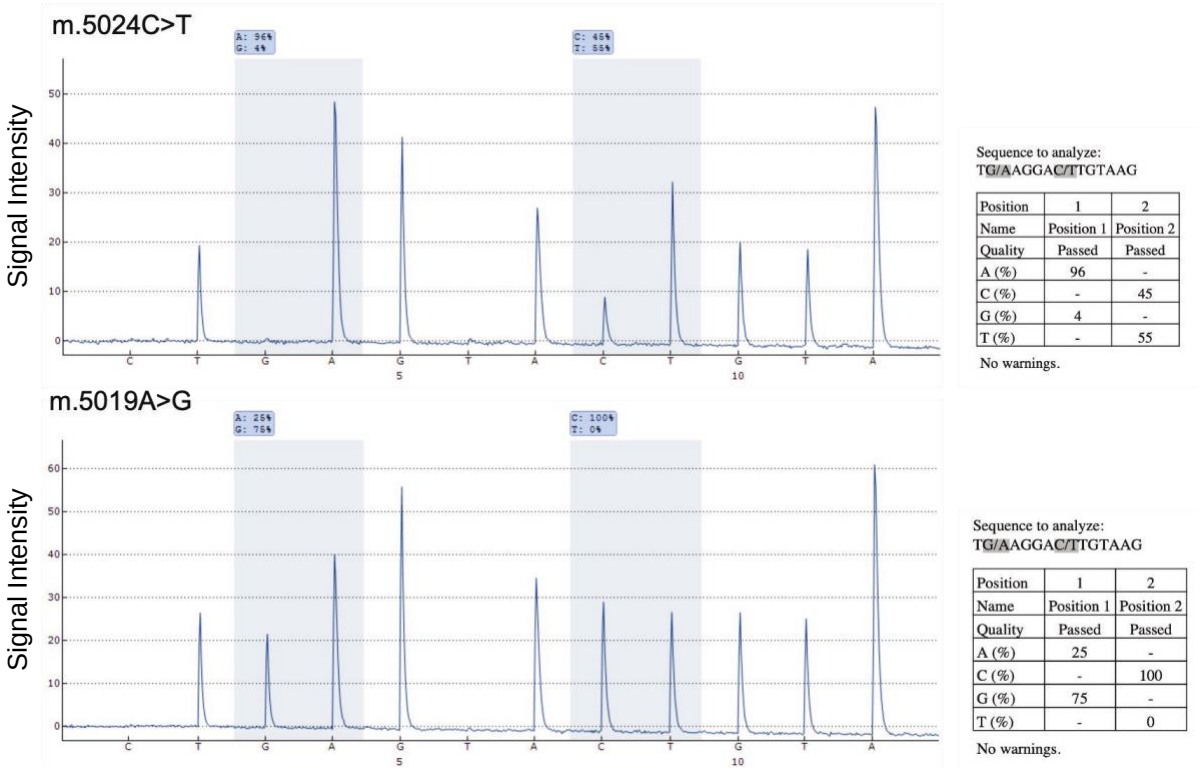


Figure 2.4. Sequences of interest.
Schematic representation of the mitochondrial DNA regions belonging to mt-Ta being pyrosequenced in the case of samples derived from either m.5024C>T or m.5019A>G mouse models. Heteroplasmic regions are highlighted in grey and the percentage of each base is automatically estimated based on the quantification of signal intensity.

Table 2.3. Pyrosequencing reagents.

Reagent	Volume (µl) (48 reactions)
dATPaS	90
dCTP	80
dGTP	70
dTTP	80
Denaturation solution	580
Enzyme	320
Substrate	320
Annealing buffer	2680
Binding buffer	320
Sequencing Primer 5019- GATGTAGGATGAAGTCTTA 5024- CACACAAGTTTAACTTCTGA	160*

*6.4µl of stock in 153.6µl of Annealing Buffer

2.3.4. Agarose Gel Electrophoresis

The product of PCR reactions as well as negative controls used to verify the absence of contamination in 96-well plates containing single cells, were examined using agarose gel electrophoresis. For the vast majority of applications, and unless otherwise stated, 2% agarose gels were used. Plastic containers and removable combs were used to cast the gels and formulate the appropriate number of wells. The 2% weight per volume (w/v) was achieved by mixing 2gr of Agarose I (ThermoFisher Scientific, PN: 17852) in 100ml of 1xTris-borate-EDTA (TBE) buffer [89mM tris base, 89mM boric acid, 2mM EDTA (pH 8.0)]. In order for the agarose to be dissolved in the TBE buffer, the mixture was placed in a heat-resistant flask and microwaved at 100^{o}C for 4 short 1min intervals that were followed up manual mixing. Once the solution became homogeneous, it was cooled down by pouring tap water on the

flask's outer surface. At that stage, SYBR Safe dye (Invitrogen™) was added to the solution at 1:10 concentration (10μl) and more manual mixing was performed in order to make sure that the dye was evenly distributed. The solution was subsequently poured within the cast and the gel was allowed to settle for around 45min. Once gel had settled, the comb was removed and 1xTBE was poured on top of the gel, the QuickLoad® 100bp DNA Ladder (BioLabs, PN: N04675) and samples were then loaded. The samples were electrophoresed at 90V and 60mA for 30-50min (depending on the band separation required). Band visualisation is made possible thanks to the use of the SYBR Safe dye (Invitrogen™) that allows for the detection of ultraviolet (UV) by the transilluminator (Gel DocTM Imaging System, Bio Rad, UK).

2.3.5. Digital Droplet PCR

Digital droplet PCR (ddPCR) was a technique used in this work in order to make exact measurements of the mtDNA copy number that was present within single cells. The first step involved single cell lysis and was carried out within a UV cabinet where prior to the lysis step, all equipment, including pipettes, tips, plates and Eppendorf tubes were subjected to sterilisation under UV light for 30-45min. Subsequently, the lysis buffer [Nuclease-Free Water (Ambion, PN: AM9937), TWEEN® 20 (Sigma, PN: 9416), Tris-HCl (pH8.5), Proteinase K (Ambion, PN: AM2546)] was made up inside the hood and mixed thoroughly by vortexing **(Table 2.4.)**. The 96-well PCR plate containing the flow-sorted single cells was then spun at maximum speed (2204xg) at 4ºC for 10min in order to make sure that all cells were at the bottom of their respective wells. Once the spinning down was complete, the plate was placed inside the cabinet where the seal was removed and 4μl of the lysis buffer were pipetted in each well. The plate was then re-sealed, briefly vortexed and spun for 2min at 300g and 4ºC in order for the lysis buffer to be puller down on top of the cells. Then, the plate was placed in the thermocycler where it incubated for 30min at 37ºC. Once that incubation was complete, the plate was placed back into the cabinet, unsealed and 8ul of water were added to each well. The plate was once again re-sealed, vortexed briefly and spun down for 2min at 300g and 4ºC. A second incubation then followed at 80ºC for 15min. When completed, the plate was once again spun down for 2min at 300g and 4ºC and then placed on ice awaiting the preparation of the ddPCR master mix.

The ddPCR master mix was made up using ddPCR Supermix for Probes (no deoxyuridine triphosphate (dUTP)), (Bio-Rad, PN: 1863024) in addition to the appropriate combination of primers and probes **(Table 2.5.)**, depending on the genes of interest in addition to whether cells were of human or mouse origins. A detailed list of the primers and probes used in the experiments described here can be found on **Table 2.6.**. Once the ddPCR master mix was complete, 18ul were pipetted on each well of a Semi-Skirted 96-well plate (Bio-Rad). Then, the plate containing the recently lysed single cells was brought in the hood and unsealed. From the 12µl of lysate present in each well, 4 were added to each corresponding well of the Semi-Skirted 96-well plate (Bio-Rad), making up the 22µl required to the next step of ddPCR. Both plates were then sealed and briefly spun down for 2min at 300g and 4^{o}C, the plate containing the single cell lysates was then placed in the freezer (-4^{o}C), where it could be preserved for later use, while the Semi-Skirted 96-well plate (Bio-Rad) was placed on ice and taken to the Automated Droplet Generator (Bio-Rad).

Alongside the plate containing the ddPCR reaction mix and single cell lysate, a chilled block (stored in the freezer at -4^{o}C), two boxes of pre-sterilised 120µl filter tips and 3 sets of DG32 Automated Droplet Generator Cartridges (Bio-Rad) were taken to the Automated Droplet Generator (Bio-Rad). Once every item was placed on its appropriate spot in the hood, the plate was unsealed and mounted on the chilled block, the lid of the hood was the shut and the droplet generation begun. Once, complete, the plate was once again sealed, this time using Pierceable Foil Heat Seals (Bio-Rad, PN: 1814040) that were attached on the plate using the PX1 PCR Plate Sealer (Bio-Rad). Subsequently, the empty rip boxes and cartages were discarded and the block was returned to the freezer. The sealed plate was placed on ice before being mounted on a C1000 Touch^{TM} Thermal Cycler (Bio-Rad), which run the PCR program described in **Figure 2.5.**. Having completed the PCR, the plate was then taken out of the thermocycler and mounted on the QX200^{TM} Droplet Reader (Bio-Rad), where the fluorescence emitted by the copies of mtDNA entrapped and amplified within droplets of oil could be detected and quantified, giving an exact measurement of the total number of copies for the genes of interest.

Table 2.4. Single cell lysis buffer.

Reagent	Volume (µl) (150 reactions)
dH2O	845
Tween	5
Trsi-HCl	10
Proteinase K	50

Table 2.5. Digital droplet PCR reaction mix.

Reagent	Volume (µl) (100 reactions)
PCR Supermix	1155
Forward primer A	9.45
Reverse primer A	9.45
Forward primer B	9.45
Reverse primer B	9.45
Probe A-HEX	5.25
Probe B-FAM	5.25
H_2O	792

Table 2.6. Primers and probes used for digital droplet PCR.

Human Primers	
Gene	**Sequence**
mND1 (Fr)	CAGCCGCTATTAAAGGTTCG
mND1 (Rev)	AGAGTGCGTCATATGTTGTTC
mCO3 (Fr)	ATCGCTCACACCTCATATCC
mCO3 (Rev)	TAGGTCTGTTTGTCGTAGGC

Human Probes		
Gene	**5' probe**	**Sequence**
mND1	HEX	ACCCGCCACATCTACCATCACCCT C
mND4	FAM	CAACCAGCCAGAACGCCTGAACGCA

Mouse Primers	
Gene	**Sequence**
mND1 (Fr)	GAGCCTCAAACTCCAAATACTCACT
mND1 (Rev)	GAACTGATAAAAGGATAATAGCTATGGTTACTTCA
mCO3 (Fr)	CCTCGTACCAACACATGATCTAGG
mCO3 (Rev)	AGTGGGACTTCTAGAGGGTTAAGTG

Mouse Probes		
Gene	**5' probe**	**Sequence**
mND1	FAM	CCGTAGCCCAAACAAT
mCO3	HEX	ACCTCCAACAGGAATTTCA

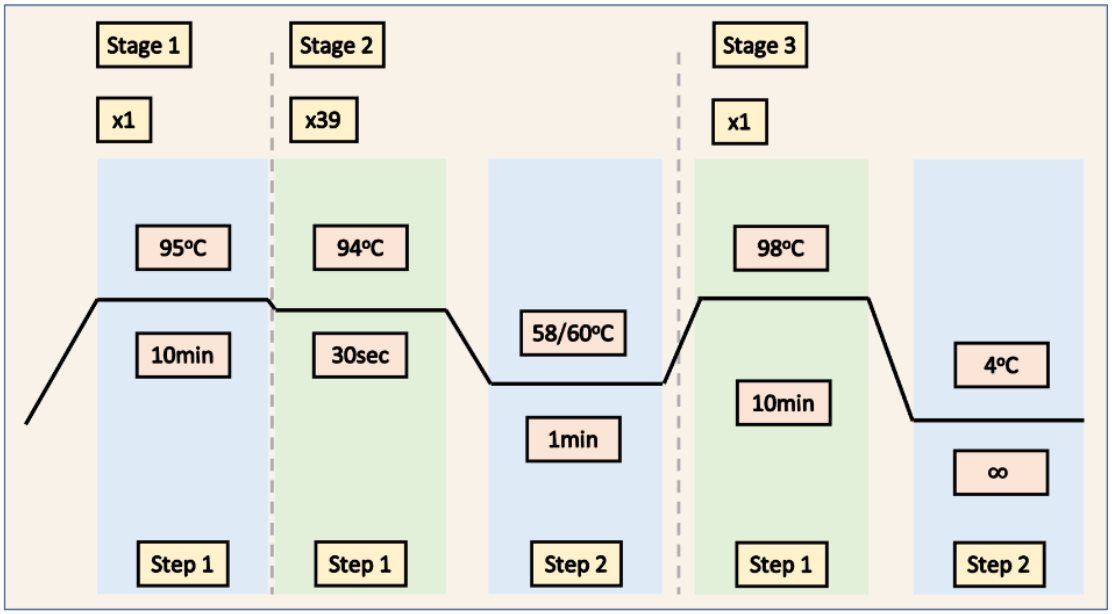


Figure 2.5. Digital droplet PCR conditions.
Schematic representation of the digital droplet PCR (ddPCR) steps. Annealing and extensions steps are combined into a 1min Step 2 carried out in Stage 2 at 58ºC in the case of mouse-derived samples and at 60ºC in the case of human-derived ones. Created with BioRender.com.

2.4. *In vitro* Culture Approaches

2.4.1. Tissue Culture

All cell types were cultured within T75 flasks (ThermoFisher Scientific) or TC plates (ThermoFisher Scientific) placed within incubators set at 37ºC and 5% (vol/vol) CO_2. Processing of the cell lines including splitting or seeding for experiments was carried out in TC hood that underwent thorough cleaning with 70% ethanol and 1% HD_4L disinfectant (Chemgene) prior to every use. All cell lines, unless stated otherwise, were grown in Dulbecco's Modified Eagle Medium (4.5 g/L glucose 2 mM glutamine, 110 mg/ml sodium pyruvate) (DMEM) (Gibco), supplemented with 10% FBS. No use of antibiotics such as penicillin or streptomycin was made. When cells are described as grown in low glucose media, their media includes DMEM (1000mg/L D-glucose, 100mg/L sodium pyruvate in addition to 584mg/L L-glutamine) (Gibco) supplemented with 10% FBS as well as 50μg/ml uridine. In cases where cells are described as being grown in galactose, they underwent culture in DMEM (no glucose or sodium pyruvate, 584mg/L L-glutamine) (Gibco) supplemented with 10mM galactose, 110mg/L sodium pyruvate, 10% FBS as well as 50μg/ml uridine. In cases where sub-cloning was performed from an original population of cells, single cells were placed into single wells of a 96-well plate through FACS. These cells were allowed to divide and once the well was confluent (usually in about 3-4 weeks), the clone was moved to 24- and eventually to a 6-well plate at which stage, bulk samples were taken in order to measure mtDNA heteroplasmy. In the experiments outlined here, use of the following cell lines was made: HeLa, 293T and MEFs derived from both the m.5019A>G and m.5024C>T mt-Ta mutant mouse strains.

2.4.2. Lentiviral Transductions

All lentiviral work was carried out in a Containment Level 2 (CL2) laboratory, having first completed the appropriate risk assessments and following the instructions of given by the unit's Biological Safety Officer.

Our first aim was to follow the Jackson Lab Viral Packaging Protocol in order to acquire the virus required in order to carry out the transductions. To that effect, the first day of the experiment, Lenti-X™ 293T cells (ClonTech, PN: 632180) were plated at a density of $6x10^6$ cells, within 10mm dishes using 17ml of complete growth medium (DMEM High Glucose (Gibco) + 10% FBS (Gibco)), aiming to achieve a 50-70% confluence level on the day of transfection. On the second day, 1.5ml of the plasmid mix solution in OptiMEM (ThermoFisher Scientific) was generated by preparing 3rd generation Addgene packaging plasmids as follows: Lentiviral vector: pMDLg/RRE : pRSV-Rev : pMD2.G in a 5 : 2 : 2 : 1 ratio. As a result, in order to complete the preparation, 13.6μg of total DNA had to be added to 1445μl of OptiMEM (ThermoFisher Scientific), that contained the plasmid mix solution. The solution was created following the directions presented below **(Table 2.7)**. Once the DNA was added, 41μl of TransIT-LT1reagent (Addgene) were placed in the solution, followed by careful mixing by pipetting 10 times, always being careful to avoid foaming. The resulting solution was then incubated for 30min at RT. Finally, the 1500μl of lentiviral complexes were added to in a dropwise manner onto the cells.

On the third day of the experiment, the cells were allowed to incubate with the virus and on the fourth day, the virus was harvested for the first time. The supernatant was collected from the plate and replaced with fresh culture media, while the cells were returned to the incubator. The virus-containing supernatant was then filtered through a 0.2μm filter. The filtered virus was then divided into 250μl aliquots, snap-frozen on dry ice and placed at -80ºC for long-term storage. A second harvest took place on the fifth day of the experiment and the same procedure was carried out in order to store more viral aliquots (this time of lower titre).

In order to transduce the cell lines of interest with the harvested virus, cells were seeded in 6-well plates at $3x10^5$ cells per well. Immediately after, 250μl of the viral supernatant was added to each well and topped up to 500μl with standard complete growth medium (DMEM High

Glucose (Gibco) + 10% FBS (Gibco)). The cells were allowed to incubate for 3hrs before 1ml of standard complete growth medium (DMEM High Glucose (Gibco) + 10% FBS (Gibco)) was added to each well in preparation for an overnight incubation. The cells were expanded 24hrs post transfection. At 72hrs post transfection, antibiotic selection was applied using both blasticidin and puromycin at 10mg/ml. We expect that the selection will kill all cells that are either untransduced or transduced with mock controls, while leaving unscathed only the ones where the transduction with the constructs of interest has been successful. In order for cells to be removed from CL2, 1 week has to pass post transduction and at least 5 changes of culture medium.

The single guide RNA (sgRNA) used in TFAM CRISPR was adapted from Castellani et all. (Castellani et al., 2020) and had the following sequence:

GCGTTTCTCCGAAGCATGTG

The lentiviral vector used was LentiCRISPR v2 (Addgene, PN: 52961) and cloning was performed following the directions of the "LentiCRISPRv2 and lentiGuide-Puro: lentiviral CRISPR/Cas9 and single guide RNA" protocol created in the Feng Zhang lab (Shalem et al., 2014).

Table 2.7 Components required for the 3rd generation plasmid mix solution.

Ratio	Role	Component	Amount (μg)	Supplier
5	Transfer	CRISPR library	8.5	-
2	Packaging	pRSV-Rev	3.4	Addgene #12253
2	Packaging	pMDLg/pRRE	3.4	Addgene #12251
1	Envelope	pMD2.G-VSVg	1.7	Addgene #12259

2.4.3. Live Cell Analysis

Quantitative real-time live cell analysis was performed using Incucyte® (Sartorius). Each well of a 6-well plate was seeded with 300.000 of cells, and allowed to grow for 4 days within the Incucyte® (Sartorius) which was in turn placed within a standard tissue culture incubator set at 37ºC and 5%CO_2. Once the recording was over, the collected data underwent Basic Analysis and then was ready for manual quantification. We opted to view the cells with red, green and pink masks placed over the channels in order to better visualise red and green fluorescence as well as the overlap of the two in accordance with the PIP-FUCCI reporter, discussed in more detail in **Results Section Chapter 6**. Viewing the cells in Phase enabled us to better examine their structure and morphology through the cell cycle. As a result, we were able to trace 10 cells from each well through the different stages of the cell cycle. Such measurements allowed for comparisons to be drawn across different culture conditions and, in the case of DH2.1 cybrids, across populations having different levels of mtDNA heteroplasmy.

2.4.4. Respirometry

High resolution respirometry was carried out using the O2k-Respirometer (Oroboros Instruments). First, the two chambers were cleaned by being rinsed with water which was then aspirated, a process that was repeated twice, always carefully in order not to disrupt the membrane at the bottom of the two chambers. At every stage of the instrument's use, it was imperative to not let the chambers dry out. The instrument was then calibrated by pipetting 2ml of DMEM+FBS (Gibco) in each chamber and creating a bubble that allowed for the machine to calibrate against the atmospheric air. During the calibration process the activity of the membranes was tested and the machine was allowed to reach a plateau which would be used as the baseline O_2 concentration.

Following the instrument's calibration, the media was aspirated from the chambers and they were washed with water. Following-up from the last wash, 2ml of sample were added to each chamber, after making sure that each sample contains the same number of cells (5 million in our case). Once the oxygen consumption rate (OCR) of each sample reached a plateau, three drugs were added sequentially using dedicated and cleaned glass syringes. First, 5µl of oligomycin A (MERCK, PN: 75351) were added, blocking the proton channel of ATP synthase

and along with it, its ability to produce ATP. Then, 2µl of carbonyl cyanide m-chlorophenyl hydrazone (CCCP) (MERCK, PN: C2759) followed up by successive doses of 1µl until a plateau was reached. CCCP being a strong uncoupling agent of the mitochondrial ETC, leads to an increase in the permeability of the IMM to protons (H^+) thus, depolarizing the mitochondria by dissipating the membrane potential ($\Delta\Psi m$) and forcing the ETC to work at its maximal capacity in an attempt to restore the $\Delta\Psi m$. Finally, 1µl of rotenone (MERCK, PN: R8875) and 2µl of antimycin-a (MERCK, PN: A8674) were added in quick succession in order to inhibit the ability of Complex I to accept electrons from NADH as well as the flow of electrons from cytochrome b in Complex III to cytochrome c.

Having completed the run of the first two samples, the cells were aspirated from the chambers and 3x10min washes were performed with 100% ethanol. Then, the chambers were rinsed once using 70% ethanol and finally rinsed two more times using water. At this stage, the chambers were ready to accept the next two samples.

When it came to analysing the data collected from the respirometer, 4 values were of interest for each sample. First, routine or baseline respiration was calculated by subtracting the background (respiration still present after the addition of rotenone (MERCK, PN: R8875) and antimycin-a (MERCK, PN: A8674)) from the basal respiration of the sample, prior to the addition of any drugs. Then, the proton-leak was measured by subtracting the background from the respiration still present when oligomycin A (MERCK, PN: 75351) was added. The ATP-linked respiration was estimated by subtracting the proton-leak from the basal respiration. Finally, the maximal respiratory capacity of the sample was calculated when the background was subtracted from respiration in the presence of CCCP (MERCK, PN: C2759) **(Figure 2.7.)**.

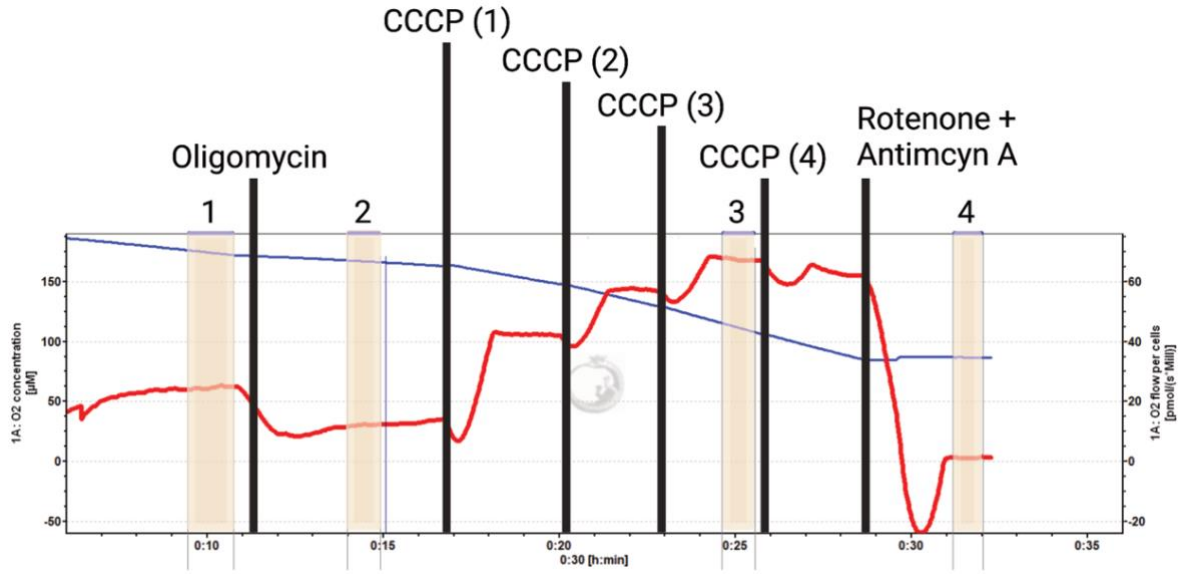


Figure 2.7. Measuring OXPHOS capacity using the O2k-Respirometer.
Representative graph of high-resolution respirometry carried out on the O2k-Respirometer. The blue line represents the O_2 concentration within the chamber, while the red line represents O_2 flow per cell. Black lines denote the timepoints at which each drug was added. Sections 1-4 represent the sampling performed once the O_2 flow rate has stabilised post-drug treatment.

2.4.5. Western Blot

Cells cultured in T75 flasks (ThermoFisher Scientific) were taken out of the incubator and placed in the TC hood. The tissue culture media was aspirated and the cells were washed using 5ml of PBS per flask. Once the PBS was aspirated, 1ml of trypsin was added to each flask. Subsequently, they were placed within the incubator (37ºC, 5% CO_2) for 5min. After making sure that the cells had detached from the bottom of the flask, 500µl of PBS were added and the cells were removed from the flask and placed in a 1.5ml Eppendorf tube. Tubes were spun down at 10,000g for 5min at 4ºC and then had their supernatant removed before being spun down again, under the same conditions. The supernatant was then removed and the cell pellet was snap frozen in liquid nitrogen. At this stage, samples were either placed on ice in order for protein extraction to take place or, were kept frozen at -20ºC. Protein extraction from the cell pellet was performed using 500ul of PathScan® Sandwich ELISA Lysis Buffer (1X) (Cell Signalling Technology®) per sample. The lysis buffer was placed on top of the cells and mixed well with the pellet through pipetting. Then, each sample was left to incubate on ice for 2min before it was spun down for 1min at 14,000g at 4ºC in order to remove any lingering cell debris. At this stage, protein concentration was measured or each sample using the Pierce™ BCA Protein Assay Kit (ThermoFisher Scientific, PN: 23227).

In order to run the samples, 1L of 1x running buffer was prepared by mixing 50ml of the 20x stock with 950ml of H_2O. NuPAGE gels were then removed from their plastic holders and had any bubbles created by the extraction of the comb removed through the use of tap water. The white strip was then taken off before the gels were placed in the running tank. Prior to filling the chambers with 1x running buffer, it was made sure that the gels were facing the correct way and that the pin electrodes were properly aligned. 8ul of the protein ladder SeeBlue™ Plus2 Pre-stained Protein Standard (ThermoFisher Scientific, PN: LC5925) were pipetted in the first well of each gel, prior to pipetting the samples in the remaining wells, always using gel loading tips and making sure that no bubbles were created in the process. The tank's lid was then closed and the gels were run at 165V for the selected duration (usually, around 45min or until the resolution of the proteins of choice is satisfactory).

Once the gel run is complete, the transfer iBlot machine (Life Technologies) was turned on. The gels were removed from the tank and the plastic cassettes were carefully cracked open using a spatula in order to avoid breaking the gel. The spatula was then used to gently transfer the gel into a plate containing freshly prepared 20% ethanol solution. The gel was allowed to soak in the solution for 5min before the transfer stack was unsealed and the gel placed on the bottom portion of the stack. The bottom part of the gel was cut off and disposed of. The remaining gel was quickly rinsed with H_2O before the transfer membrane was placed on top of it. A roller was used in order to make sure that no bubbles were present between the gel and the membrane. The top stack was then placed on top of the membrane and rolled again to once again prevent the formation of bubbles. The absorbent pad was then placed on top of the stack, assembling the transfer tray and making sure that its electrical contacts were aligned with the machine's corresponding nodes. The newly assembled transfer tray was then mounted onto the machine, the lid was shut and the desired protocol was selected. In the case of our work, Protocol 0 was the one of choice, lasting for 7min. After the protocol's completion, the lid was opened and the absorbent pad, the top stack, the filter paper and the gel were all discarded, leaving the membrane to be transferred in a plate containing 5% milk in TBS-Tween 0.1% in order to be blocked. Once the membrane was removed, the bottom half of the transfer stack was discarded.

The blocking step lasted for 1hr and was carried out at RT under slow rocking. Once completed, the membrane was cut according to our experimental requirements and each strip was incubated within a plate containing the primary antibody diluted in 5% milk in TBS-Tween 0.1% at 4ºC, overnight. Following this incubation, three 10min washes were performed with TBS-Tween 0.1%. Once the washes were complete, the membranes were incubated with the secondary antibody, that was diluted 1:1000 in 5% milk in TBS-Tween 0.1%, for 1hr at RT. Following the incubation with the secondary antibody, the membranes underwent three 10min washes in TBS-Tween 0.1% and placed in a plate containing the same buffer while waiting for imaging.

In preparation for imaging, *Luminol Enhancer Solution* Detection Reagent 1 (ThermoFisher Scientific, PN: 1859701) was mixed with Detection Reagent 2 (ThermoFisher Scientific, PN: 1859698) within a foil-covered 15ml tube at a 1:1 ratio. The membrane was reconstructed on a piece of stretch film wrap and any excess buffer was carefully removed. Then, the Detection Reagent solution was poured on top of the membranes which were immediately covered with a foil-covered lid for 5min. Following this incubation away from the light, the membranes were drained from any excess solution and placed on the imager tray. Imaging was conducted using the Amersham Imager 600 (General Electric) with the protocol of choice being subject to experimental needs.

Table 2.8. Antibodies for western blot.

	Target	Species	Concentration	Supplier
1	anti-Rabbit (2º)	Swine	1:1000	Agilent PN: P0217
2	anti-Mouse (2º)	Rabbit	1:1000	Agilent PN: P0260
3	TFAM (1º)	Rabbit	1:1000	Cell Signalling PN: 8076S
4	Vinculin (1º)	Mouse	1:1000	Sigma PN: V4505

2.5. Sample preparation for proteomic analysis

2.5.1. Liver

Liver samples from the mice of choice were taken out of the -80ºC freezer and placed on dry ice. While 30-50mg of tissue is enough for the preparation of the sample, we opted to go for 100mg in order to simplify the grinding of the tissue. The pieces of liver were powdered in liquid nitrogen using a pestle and mortar and 50µl of the aliquoted liver samples was then moved into a new 1.5ml Eppendorf. 200µl of lysis buffer (1% SDS on 50mM of HEPES buffer with 1 tablet of protease inhibitor cocktail (Abcam, PN: ab271306)) were added to each Eppendorf. The samples were then vortexed briefly and spun down. Sonication took place in the cold room (4ºC) at 20% power for each sample, lasting 10sec. A second round as performed in cases where the sample was not completely dissolved. Between sonication intervals the samples were kept on ice. Subsequently, all samples were spun down at 16,000g for 10min at 4ºC. The supernatant was then carefully separated and placed into a new Eppendorf which was subsequently spun again at 16,000g for 10min at 4ºC. Once more, the supernatant was separated and transferred into a new Eppendorf. At this stage the protein concentration of the sample was estimated through a Pierce™ BCA Protein Assay Kit (ThermoFisher Scientific, PN: 23227), and the samples were made up at 2µg/µl for 100µl. All samples were stored at -80ºC prior to submission for analysis.

2.5.2. Brain

In order to prepare samples for proteomic analysis derived from mouse cortex we first chilled a dounce homogeniser on ice alongside lysis buffer (10mM Tris-HCl (pH7.4), 0.15mM NaCl, 1mM EDTA, PBS containing 1% beta-OG (Sigma) (w/v), 1 tablet of protease inhibitor cocktail (Abcam, PN: ab271306)). The tissue was removed from the -80^{o}C freezer and very briefly weighted before being returned into an Eppendorf placed on dry ice. After all the samples had been weighted, 20ul of ice-cold lysis buffer were poured into the chilled dounce vessel for each 5mg of sample tissue. Gel loading tips were used to reach the bottom of the vessel and remove the homogenate and place it into a chilled 1.5ml Eppendorf. Subsequently, all samples were incubated at 4^{o}C for 1hr under mild agitation. Centrifugation followed at 10,000g for 30min at 4^{o}C. The supernatant was removed and placed into a new Eppendorf. Protein quantification followed using the Bradford assay, and the samples were made up at 2µg/µl for 100µl. All samples were stored at -80^{o}C prior to their submission for analysis.

Following their preparation, both brain and liver samples were submitted to the proteomics facility at the University of Cambridge Department of Biochemistry, in order for tandem was tagging (TMT)-labelled proteomic analysis to be conducted.

2.6. Statistical Analyses

Single cell mtDNA heteroplasmy data is typically presented as either violin plots showing median and interquartile range (IQR) or as mean ±SD unless specified otherwise. MtDNA copy number measurements are presented as mean ±SD. In order for single cell mtDNA heteroplasmy measurements to be comparable across mice that had different bulk tissue heteroplasmies, transformation of the data had to take place prior to any statistical test. Where appropriate, normalised heteroplasmy shift $\Delta h'$ was calculated as:

$$\Delta h' = log\left(\frac{h(h_0 - 1)}{h_0(h - 1)}\right)$$

Where h was the single-cell heteroplasmy fraction and h_0 was the mean bulk-tissue heteroplasmy fraction, averaged across all sampled tissues (Burgstaller et al., 2014; Johnston and Jones, 2016). In the above equation, $logit(h)$ denotes the log-odds function.

Comparison across datasets comprised of single cell mtDNA measurements was carried out using the Kolmogorov–Smirnov (KS) test, which is a non-parametric test for the equality of distributions of continuous variables. Bonferoni multiple testing corrections were applied were necessary. Results of statistical analyses throughout are presented throughout the text and in the corresponding figure legends, as appropriate. Statistical tests were carried out at the 5% significant level, with the following annotation being used throughout:

$p < 0.05 \rightarrow$ *, $p < 0.01 \rightarrow$ **, $p < 0.005 \rightarrow$ *** and $p < 0.0001 \rightarrow$ ****

In order to extend our conclusions regarding trends around single cell heteroplasmy distributions beyond observations of mean and variance, we made use of the adapted Kimura distribution as described by Passorn Wonnapinij and colleagues (Wonnapinij et al., 2018). In their work, the Kimura distribution was adjusted in order to predict the presence of random genetic drift in mother-to-offspring datasets. The initial assumption was that the mean will remain constant across generations while the variance will follow the Sewell-Wright formula (Wright, 1942):

$$V(t) = p_0(1 - p_0)\left[1 - e^{-\frac{t}{N_{eff}}}\right]$$

$$\approx p_0(1 - p_0)\left[1 - \left(1 - \frac{t}{N_{eff}}\right)^t\right]$$

V= Heteroplasmy variance in a population of individuals who share the same maternal ancestor, in our case represents the variance in the population of single cells that arise from the same oocyte.

t= The number of generations between the common maternal ancestor and our current population of individuals, in our case represents the number of cell divisions between the oocyte and the single cell population of interest.

p_0= Initial gene frequency (heteroplasmy) of the common maternal ancestor, in our case represents the heteroplasmy of the oocyte.

N_{eff} = The number of individuals in the population, in our case represents the number of single cells in the population of interest.

While the mean and variance of a heteroplasmy distribution are indeed important parameters, they do not inform us of the distribution itself. This is particularly true when dealing with populations whose mean is either at very high or very low heteroplasmy values, resulting in the distribution being asymmetric. The work of Motoo Kimura in 1955 provided a solution when it comes to the probability of gene frequencies in diploid population (Kimura, 1955), his theory was then applied to mitochondrial DNA heteroplasmy (Wonnapinij et al., 2018). The probability distributions functions derived by Kimura in order to define the gene frequency distributions of a certain population under the sole influence of random genetic drift, include the lack of selection, no overlapping generations, no migration, the lack of de novo mutations and a constant population size (Kimura, 1955). The latter was put in place in order to simplify the mathematics as follow-up work has shown that this assumption is unnecessary (Maruyama and Kimura, 1980). Solving this model requires three equations:

$$f(0,t) = (1-p_0) + \sum_{i=1}^{\infty}(2i+1)p_0(1-p_0)(-1)^i F(1\ -\ \mathrm{i}, \mathrm{i}\ +\ 2, 2, 1-p_0)e^{-\left(i(i+1)/2N_{eff}\right)t}$$

$$\phi(x,t) = \sum_{i=1}^{\infty}(i+1)(2i+1)p_0(1-p_0)F(1\ -\ \mathrm{i}, \mathrm{i}\ +\ 2, 2, \mathrm{x})F(1\ -\ \mathrm{i}, \mathrm{i}\ +\ 2, 2, p_0)e^{-\left(i(i+1)/2N_{eff}\right)t}$$

$$f(1,t) = p_0 + \sum_{i=1}^{\infty}(2i+1)p_0(1-p_0)(-1)^i F(1\ -\ \mathrm{i}, \mathrm{i}\ +\ 2, 2, p_0)e^{-\left(i(i+1)/2N_{eff}\right)t}$$

The three equations presented above constitute the Kimura distribution:

f(0,t)= The probability of fixing the WT allele in the population.

f(1,t)= The probability of fixing the mutant allele in the population.

ϕ(x,t)= The probability that the allele of interest is present at a frequency x in the population.

F(1 − i, i + 2, 2, z)= The hypergeometric function

Chapter 3. Results:

High Throughput Quantification of Mitochondrial DNA Heteroplasmy at the Single Cell Level

3.1. Introduction

The endosymbiotic relationship between a primitive proteobacterium and an anaerobic prokaryote over 2 billion years ago set the stage for the formation of the organelles we now know as mitochondria (Kowallik and Martin, 2021). Over the millennia, the overwhelming majority of the prokaryotic genome was transferred to the cell's nucleus and the organelle was left with only a small, much restricted genome termed mtDNA, further reinforcing the co-dependency between the mitochondria and the nucleus. While around 10% of the cell's proteome is found within mitochondria, mtDNA codes only for 37 genes that include 13 polypeptides, 22 tRNAs and 2 rRNAs (Pagliarini et al., 2008). The copy number of mtDNA ranges widely from around one hundred copies in sperm to several hundred of thousand copies found in mature oocytes (Stewart and Chinnery, 2015). Inherited mutations of the mtDNA were once thought of to be extremely rare but more recent work has established that they are present in around 1:5000 individuals (DiMauro et al., 2013; Vafai and Mootha, 2012). mtDNA mutations can range from point mutations to large scale deletions and are often present within a cell alongside WT copies of mtDNA, in a state known as heteroplasmy (Chinnery et al., 2000). Taken together, mtDNA mutations exhibit a wide range of clinical presentations (Stewart and Chinnery, 2015). Their heterogeneity is further compounded by the heteroplasmic state, characteristic to most mtDNA mutations, that renders clinical diagnosis particularly challenging (Stewart and Chinnery, 2015).

The nature of mtDNA prevents the use of transfection as a way of introducing mutations and developing mouse models of mitochondrial disease. In 2016, a novel, phenotype-driven approach was employed by Kaupplia and her colleagues as an alternative to cytoplasmic fusion (Kauppila et al., 2016). Their work, resulted in the generation of mouse models carrying pathogenic mtDNA mutations, using a method that was far less laborious but also allowed for the incorporation of mutations that were not limited to those present in somatic tissues or cell lines (Kauppila et al., 2016). This new approach was based on the detection of mtDNA mutations clonally expanding within the colonic crypts of founder mice originating from heterozygous animals of the mtDNA mutator strain (Kauppila et al., 2016). The result was the generation of a mouse model of mitochondrial disease carrying a heteroplasmic *mt-Ta* (m.5024C>T) mutation (Kauppila et al., 2016).

As mutations on mt-tRNAs are relatively well tolerated, they have a better chance of evading strong purifying selection in the maternal germline, and can be propagated down future generations (Stewart and Larsson, 2014; Freyer et al., 2012). Since the publication of the m.5024C>T *mt-Ta* mouse model in 2016, a second strain has been generated, this time carrying a more pathogenic mutation, also on *mt-Ta*, the heteroplasmic m.5019A>G. This mouse model and the ways in which it differs from the first *mt-Ta* mutant will be covered in more detail in **Results Section, Chapters 4** and **5**. In the interest of simplicity and to avoid confusion, for the remainder of this thesis, I will be referring to these two mice as the m.5024C>T and the m.5019A>G mouse models.

3.2. Project Aims

The heteroplasmic state of most mtDNA mutations is complicated, dynamic and plays a crucial role in whether or not a mutation will reach the threshold required for a pathological phenotype to be presented (Zhang et al., 2018). We know that this threshold can differ greatly between tissues and this might be a key factor driving the organ-specific manifestation of mitochondrial diseases, as they are known to affect multiple tissues within the body while leaving others unscathed (Stewart and Chinnery, 2015). Until today, work on this field has been done at the organ and tissue-wide level. However, the vast majority of the human organs, that are most heavily affected by mitochondrial disorders, bring together a wide range of different cell types in order to perform their function. A great example of the multisystem nature of mitochondrial disorders that can arise from the point of birth until the later stages of life, is provided by the year-long yet unresolved debate of 'lumpers' vs 'splitters' (DiMauro et al., 2013; Vafai and Mootha, 2012). 'Lumpers' were clinicians who considered mtDNA disorders to be an amalgam of multisystem disorders with ill-defined characteristics, while 'splitters' attempted to carefully define well-characterised syndrome such as MELAS, CPEO and KSS (DiMauro et al., 2013). Either way one looks at mitochondrial disorders, it is clear that disruption of the nervous system's proper function is central to such pathologies. With brain being the most heterogeneous organ in mammals, 133 distinct cell types were identified within the mouse neocortex through single cell RNA sequencing (scRNA-seq) (Tasic et al., 2018). As a result, it is not inconceivable that within such a complicated landscape, heavily impacted by mitochondrial dysfunction, certain cells or cell-types would be disproportionately impacted. Indeed, we do not have to rely solely on speculation, as evidence supporting the mosaic nature of mitochondrial disease can be found in histological examinations of affected tissues belonging to the m.5024C>T mouse (Kauppila et al., 2016). More specifically, mice carrying the m.5024C>T mutation at 45% heteroplasmy displayed the presence of colonic crypts that were made-up from COX-deficient epithelial cells, while others appeared unaffected (Kauppila et al., 2016). In the same fashion, when the m.5024C>T heteroplasmy reached 68% in one-year-old mice, some colonic smooth muscle fibres started to exhibit COX-deficiency (Kauppila et al., 2016) **(Figure 3.1.)**.

This innate heterogeneity has implication on the long-established and clinically relevant practice of tissue-sampling. While sampling a section of a tissue in order to get an estimate of the organ's average heteroplasmy might be informative, it risks overshadowing minute nuances that can point towards cell-type specific selection for or against the mutation. Such a pattern, only becomes observable through the employment of high-throughput single cell approaches. Once such techniques become mainstream, they carry with them the potential to revolutionise our understanding of mitochondrial disorders and shed light into the underlying forces driving mtDNA heteroplasmy segregation.

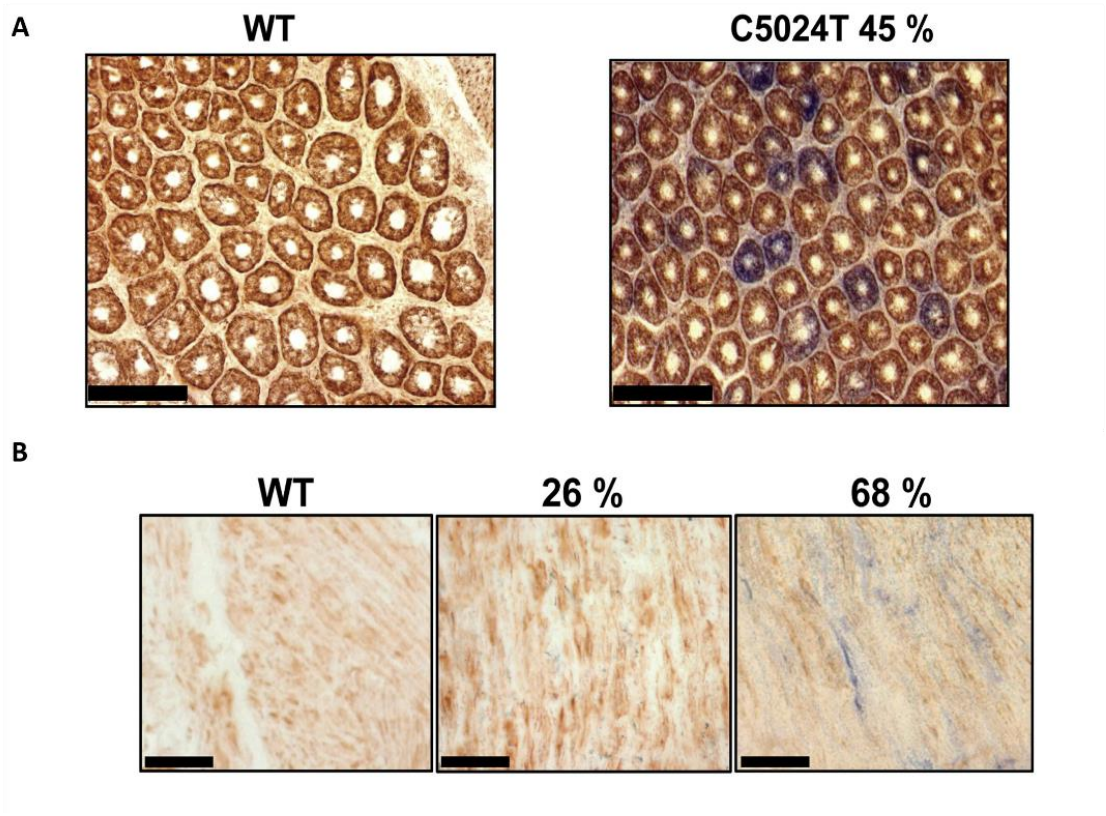


Figure 3.1. COX deficiency in the *mt-Ta* m.5024C>T mouse model.
(A) Representative sections of mouse colonic crypts where COX/SDH staining has been applied. Tissue derived from a wild-type (WT) mouse exhibits unhindered COX activity, while cell stained in blue indicate the presence of COX-deficiency in the case of the *mt-Ta* m.5024C>T mouse model carrying 45% heteroplasmy. Black scale bars represent 100μm. **(B)** At about 1 year of age, high heteroplasmy animals exhibit COX-negative colonic smooth muscle fibres. Black scale bars represent 50μm. Figure adapted from: Kauppila et al., 2016.

In this chapter, I present my work on optimising the method of pyrosequencing for the measurements of mouse mtDNA heteroplasmy both in bulk samples and at the single cell level, in a high-throughput manner. Having discussed what drove us towards conducting single cell heteroplasmy measurements, I study the principles governing pyrosequencing, a technique that allows for high-throughput single cells sequencing and thus, plays a central role in our work. Then, I present evidence that make a case on the discrepancy in heteroplasmy measurements observed when each DNA strand is used as template for pyrosequencing. I demonstrate which strand should be used as a template for each of the two mouse strains. Subsequently, I delve into the issue of pyrosequencing PCR contamination we faced and our attempts to identify the cause and adjust our protocol and practices in order to address it. Finally, I examine whether as a method, pyrosequencing is reliable in measuring mtDNA heteroplasmy, both at the bulk and single cell level, demonstrating that the measurement error in question is within expected parameters, having little to no effect on our conclusions. Statistical modelling of the data in this chapter has been conducted by Dr Lyuba Bozhilova.

3.4. Principles of Pyrosequencing

Since the invention of Sanger sequencing in 1977, DNA sequencing has played an increasingly central role in our effort to understand the function of biological systems (Gharizadeh et al., 2003). Pyrosequencing was the first alternative to Sanger sequencing when it comes to *de novo* sequencing of DNA (Sanger et al., 1977). The idea was conceived by Pål Nyrén in 1986, while he was working as a postdoc in the lab of Sir John Walker (Nyren, 2015). It required ten years of hard work, overcoming both technical and financial hurdles before this idea was finally put into practice (Nyren, 2015). The concept behind this novel sequencing technique was to trace the incorporation of nucleotides by DNA polymerase while monitoring the release of pyrophosphate, in a sequencing-by-synthesis approach (Harrington et al., 2013). The complementary nature of Watson and Crick base pairing would lead to the gradual deciphering of the strand's sequence as known nucleotides would be added in successive rounds (Nyren, 2015). Pyrosequencing is carried out without the use of labelled nucleotides, primers or the need for gel electrophoresis and can be used for the identification of single nucleotide polymorphisms (SNPs) as well as short spans of DNA, of up to one hundred bases in length (Gharizadeh et al., 2003). The nature of pyrosequencing carries distinct advantages as it can be easily automated, while at the same time allowing for higher flexibility, accuracy and parallel processing (Harrington et al., 2013).

Based on the principles of sequencing-by-synthesis (Hyman et al., 1988), pyrosequencing makes use of four enzymes to accurately discern the sequence of nitrogenous bases during the process of DNA synthesis **(Figure 3.2.)** (Fakruddin and Chowdhury, 2012). A biotinylated template of single-stranded DNA is used to as the basis on which the sequencing primer is bound (Gharizadeh et al., 2007). Then, the four different enzymes; luciferase, apyrase, DNA polymerase and ATP sulfurylase are added alongside the substrates luciferin and adenosine 5′ phosphosulphate (APS) (Gharizadeh et al., 2007). Four deoxynucleotide triphosphates (dNTPs) are then added to the reaction in a sequential and repetitive manner (Fakruddin and Chowdhury, 2012). The reaction cascade initiates when a nucleotide gets incorporated by DNA polymerase onto the newly synthesised strand. Nucleic acid polymerisation that takes place as a result of this incorporation, brings about the release of inorganic pyrophosphate (PPi) (Fakruddin and Chowdhury, 2012). The amount of PPi released corresponds to the number of nucleotides that have been successfully incorporated into the newly synthesised DNA strand.

In due course, ATP sulfurylase in the presence of APS drives the conversion of PPi into ATP (Harrington et al., 2013). In every case, the amount of ATP produced will be proportional to the levels of PPi released. In turn, ATP then drives the conversion of luciferin into oxiluciferin, a reaction catalysed by the enzyme luciferase, resulting in the generation of visible light, once more, proportional to the levels of PPi initially released during dNTP incorporation (Harrington et al., 2013). The light produced by this reaction has a maximum wavelength of 560nm and is captured by photon detection device in the form of a photomultiplier or a charge coupled device (CCD) (Fakruddin and Chowdhury, 2012). The role of apyrase as a nucleotide-degrading enzyme is to break down ATP and dNTPs that have not been incorporated and are still lingering around the reaction mixture (Fakruddin and Chowdhury, 2012). Consequently, between each dNTP dispensation, there is a short break of about 65sec that allows for apyrase to perform its function (Fakruddin and Chowdhury, 2012). The structure of this procedure means that one known type of dNTP has to be added to the reaction mixture at each given time point (Gharizadeh et al., 2007), in order for the light produced from each reaction to be linked directly to a single nitrogenous base. At the same time, capturing and recording the intensity of the light produced points us towards the number of bases that have been incorporated, with intensity of a certain level implying the incorporation of a single base while doubling or tripling that intensity signifies that double or triple the number of bases has been successively incorporated. Of course, knowing the exact sequence of nucleotides added to the complementary strand, allows us to infer the sequence of the template based on the principles of complementary base pairing (Ronaghi et al., 2001).

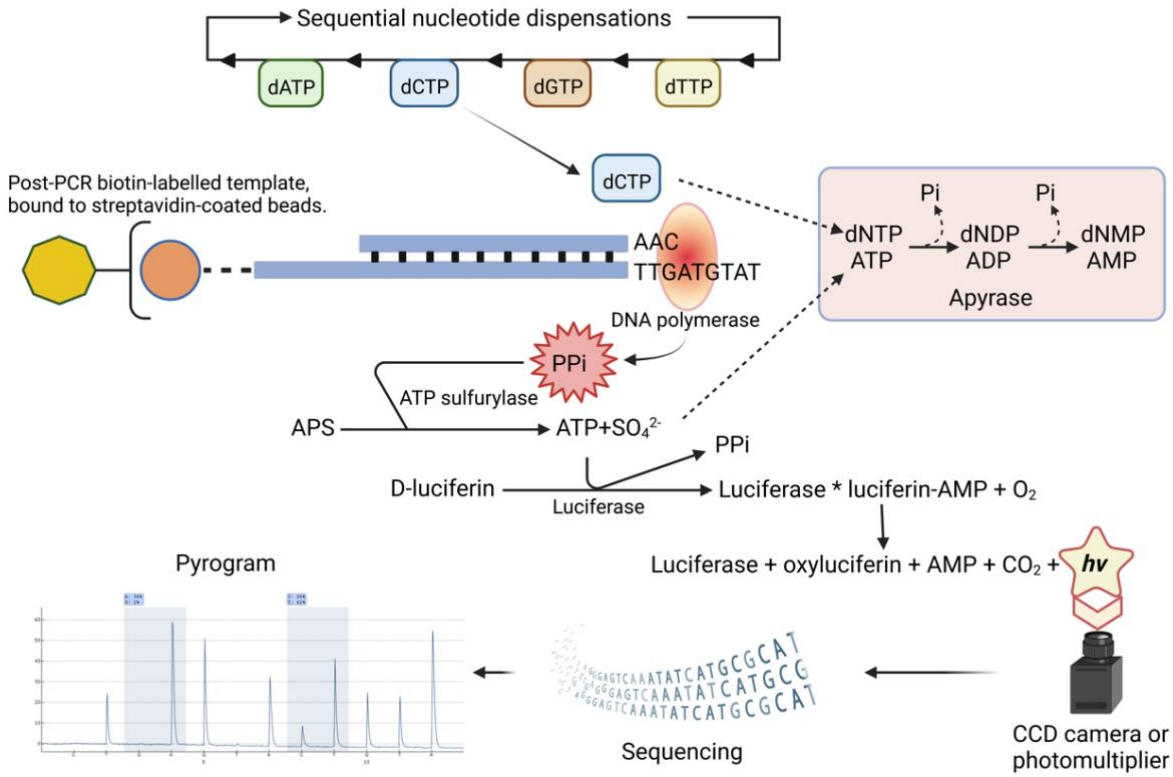


Figure 3.2. The process of pyrosequencing.
Schematic representation of the reaction steps taken when pyrosequencing a sample that has been prepared using a biotinylated PCR primer and has subsequently been bound to a streptavidin coated beads. The light emitted through the conversion of D-luciferin to luciferase after the addition of each subsequent dNTP gets recorded as a peak in the pyrogram that enables the user to accurately read the sequence of choice. Created with BioRender.com.

In the experiments to follow, the PyroMark Q48 Autoprep platform (Qiagen) was used in order to carry out in-house, high-throughput pyrosequencing analysis. More information on the technical details surrounding the use of this platform can be found in **Methods Section, Paragraph 2.3.3**. Pyrosequencing as a technique played a central role in collection of many datapoints presented in the **Results Section, Chapters 4-6**. Consequently, it was imperative that prior to making extensive use of this technique, we first optimised the protocols for measuring heteroplasmy at the 5019 and 5024 positions of mouse mtDNA.

3.4. Optimisation of the Pyrosequencing Protocol

The very first challenge we faced when starting to make large-scale use of pyrosequencing for the measurement of mouse mtDNA heteroplasmy, came with the design of PCR primers. It was observed that while we focused on a single pathogenic SNP for each mouse strain, the strand used as a template, influenced the peak height recorded by the PyroMark Q48 Autoprep platform and as a result, the heteroplasmy percentage measurement on that particular base. Initially this came into our attention when comparing our pyrosequencing results to next generation sequencing (NGS) data obtained using MiSeq (Illumina). Aiming to examine this initial observation in greater detail, mtDNA that was known to be mutant homoplasmic for the m.5019A>G mutation, derived from mouse embryonic fibroblasts (MEFs), was used. Serial dilutions were made and samples were prepared, each containing different percentages of heteroplasmy, ranging from mutant homoplasmic to WT mtDNA. 11 samples of known heteroplasmy were tested using PCR primers that used either the forward or the reverse strand as template **(Figure 3.3)**. When sequencing samples that ranged between 0% and 50%, both sets of primers gave comparable results, as estimates of sample heteroplasmy showed little to no discrepancy from the expected value. However, once the samples exceeded 50% heteroplasmy, the set of primers that used the forward strand as a template seemed to consistently underestimate the heteroplasmy percentage of the sample while its counterpart, remained largely consistent. A similar experiment was carried out in the past by previous members of the lab on the m.5024C>T mutation and the reverse trend was observed, with the forward strand providing consistent results across the whole spectrum while the reverse strand exhibiting a discrepancy at heteroplasmies lower than 50%. These results left us with two options, either to manually adjust the pyrogram peak height generated when dATP was added, or design primers that are measuring heteroplasmy on the complementary strand to avoid adenine incorporation in the heteroplasmic region. The possible cause behind this discrepancy will be analysed in the **Discussion Section**.

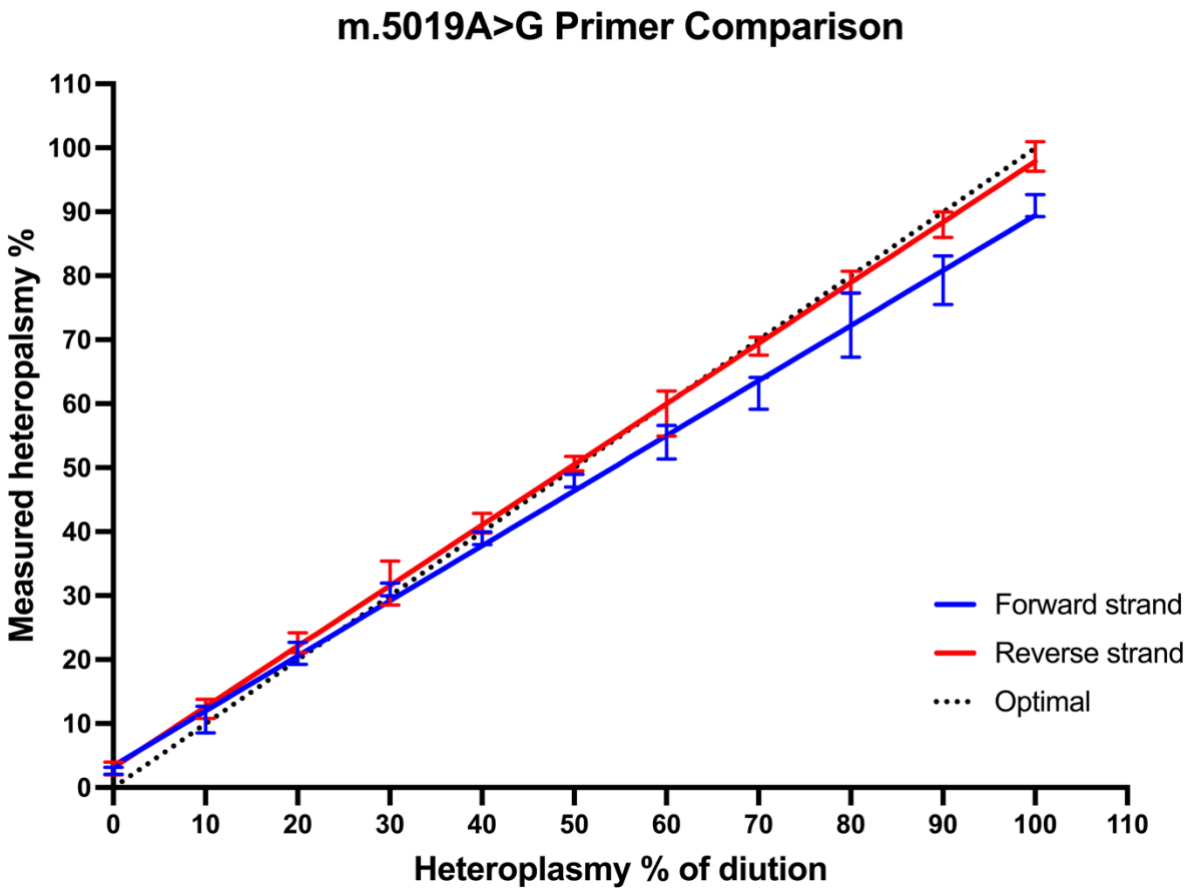


Figure 3.3. Measuring m.5019A>G heteroplasmy on the forward vs. reverse strand.
The dotted line $y = x$ represents optimal measured heteroplasmy levels in the absence of any form of experimental error. The blue and red colours correspond to heteroplasmy values obtained from the forward and reverse strands being used as the PCR template, respectively. Measured data points are represented as whiskers denoting mean ±SD (n=3 at each dilution). The curves are generated using simple linear regression.

3.4.1. Combating Pyrosequencing PCR Contamination

For as long as PCR amplification of DNA has been around, the ever-present danger of contamination is something that needs to be taken into consideration. The very nature of PCR means that the presence of exogenous DNA molecules within the reaction mixture can result in skewing or even overshadowing the readings originating from sample DNA (Kwok and Higuchi, 1989). Of course, the smaller amount of starting DNA present in a reaction, the greater the impact of contamination is, as even a small number of exogenous molecules subject to amplification can skew the result of the reaction (Kwok and Higuchi, 1989). It has always been the case that the best way to deal with PCR contamination is preventing it from occurring in the first place (Kwok and Higuchi, 1989). In our case, tackling this problem begun when trying to understand why the mean heteroplasmy value calculated by averaging out all single cell measurements was so divergent from bulk tissue heteroplasmy **(Figure 3.4.)**. We hypothesised that a single-cell population representative of the tissue of choice, should have a mean heteroplasmy value that matches that of the bulk tissue estimate.

The mismatch between mean single-cell and bulk-tissue heteroplasmy values suggested that PCR contamination was present. While its effect on bulk samples appeared limited due to the sheer amount of mtDNA contained within a piece of tissue, it was detrimental to any single cell heteroplasmy measurements. At the same time, working on two different mutant strains that carry mutations positioned so close to each other, allowed us to devise two pyrosequencing protocols, each measuring heteroplasmy in both the m.5019 and m.5024 position on their respective strands (see **Methods Chapter, Figure 2.4.**).

This protocol enabled us to make sure that when dealing with cells derived from an animal carrying the m.5019A>G mutation, heteroplasmy measurements at the position of interest were being carried out as planned but at the same time, the WT homoplasmic nature of the m.5024 position remained unchanged. Using this protocol, we were able to detect the different levels of m.5024C>T heteroplasmy in cells derived from a mouse carrying the m.5019A>G mutation **(Figure 3.5.)**. In response, we went ahead and applied thresholds in order to test whether the mean single cell heteroplasmy average at the 5019 position, matched the bulk tissue heteroplasmy average, the closer the m.5024 position was to WT homoplasmy **(Figure 3.6.)**. As expected, the lower the heteroplasmy at the m.5024 position, the lower the level of contamination was, and the more accurate heteroplasmy measurements on the m.5019 position were. In that way, establishing a more analytical pyrosequencing program allowed us to make sure there was no skewing of the measurements as a result of cross-contamination between the two strains.

However, these protocols could not detect the presence of contamination arising from WT mouse mtDNA or mtDNA of the same strain. As a result, the incorporation of negative controls in the form of wells left empty during cell sorting was employed. As long as the negative controls came up as blank **(Figure 3.7.)** and the rest of the wells showed heteroplasmic readings only at the expected positions (either m.5019A>G or m.5024C>T), we could assume that there was no widespread mtDNA contamination at the level of the plate.

To conclude, during these experiments, carryover contamination from one sample to the other was not a problem, owing to the use of dispensable pipette tips. However, the presence of exogenous mouse mtDNA, most likely airborne within the premises of our lab appears to contribute to the skewed nature of early single cell measurements when compared to bulk tissue results. The nature of the pyrosequencing PCR protocol (discussed in depth in **Methods Chapter, Section 2.3.3.**), means that even a miniscule amount of exogenous contaminant, over 48 cycles of amplification is enough to heavily influence any single cell heteroplasmy measurement. Since we came to that realisation, all single cell pyrosequencing PCR reactions were carried out in a dedicated TC hood that was meticulously cleaned before and after every experiment using both 70% ethanol and DNA*Zap*™ (ThermoFisher) (as described in **Methods Chapter, Section 2.3.3.**).

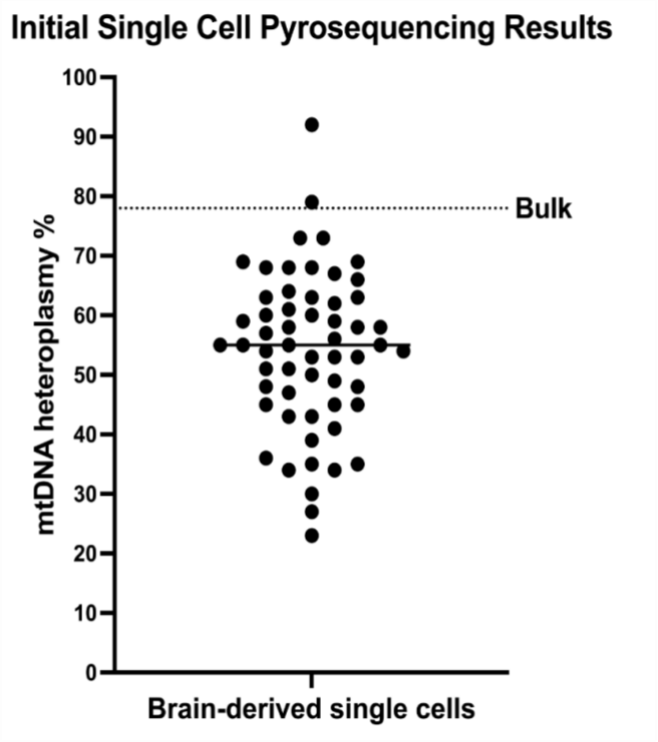


Figure 3.4. The impact of contamination on single cell heteroplasmy measurements.
Heteroplasmy measurements on brain-derived single cells that demonstrate the effect of contamination during pyrosequencing PCR preparation. Each dot represents a single cell; the dotted line represents the bulk brain heteroplasmy measurement; the solid line represents mean single cell heteroplasmy.

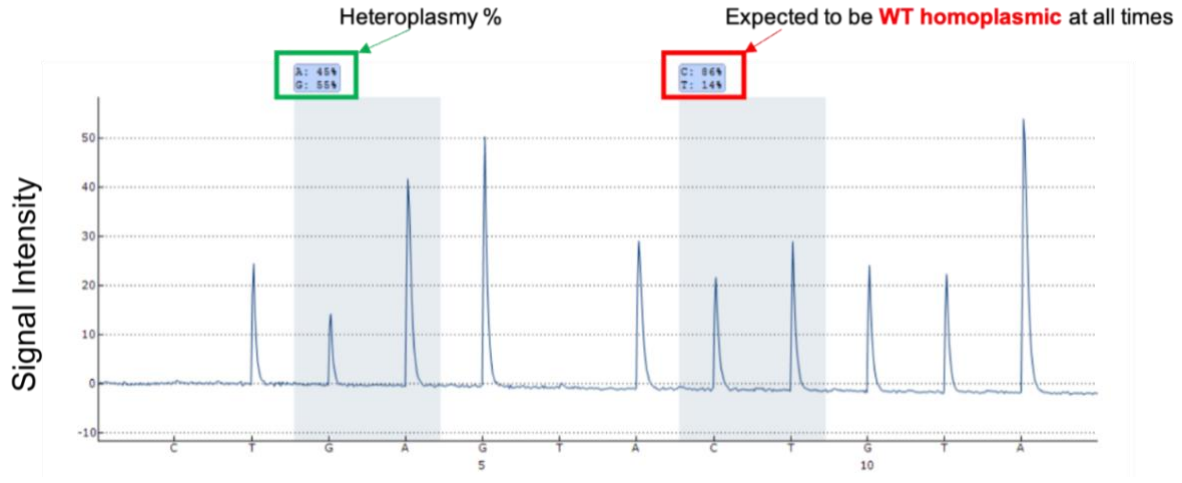


Figure 3.5. Detection of cross-strain contamination during pyrosequencing.
Representative example of a single cell sample derived from a mouse carrying the m.5019A>G mutation showing signs of contamination (heteroplasmy at the m.5024 position).

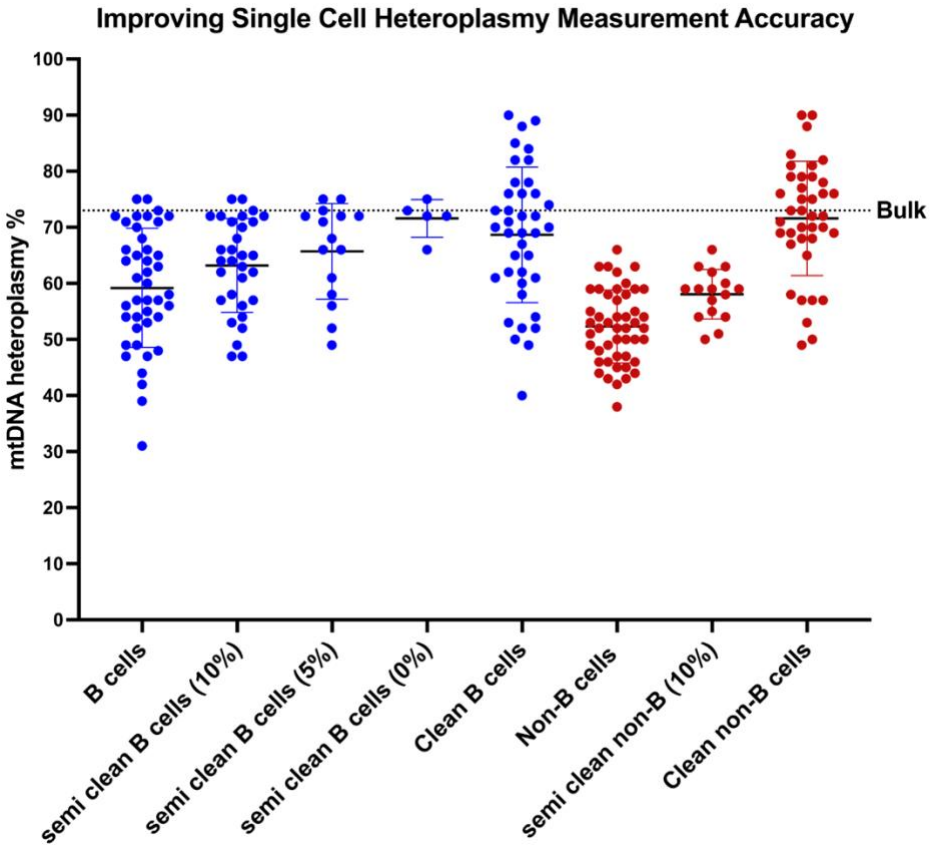


Figure 3.6. Establishing the impact of contamination.

Original spleen-derived cell populations that exhibited high levels of m.5024C>T heteroplasmy, even though the mouse was wild-type (WT) homoplasmic at that position. Semi clean population represent the same dataset with a m.5024C>T threshold applied to it (10%, 5% and 0% in B cells and 10% in non-B cells). In each case, cells which displayed heteroplasmy levels above designated cut-off for m.5024C>T were excluded. Clean populations represent cells of the same origin and animal where no heteroplasmy was present in the m.5024 position. In blue are single cell measurements on CD19+ve (B cells) while in red are CD19-ve (non-B cells).

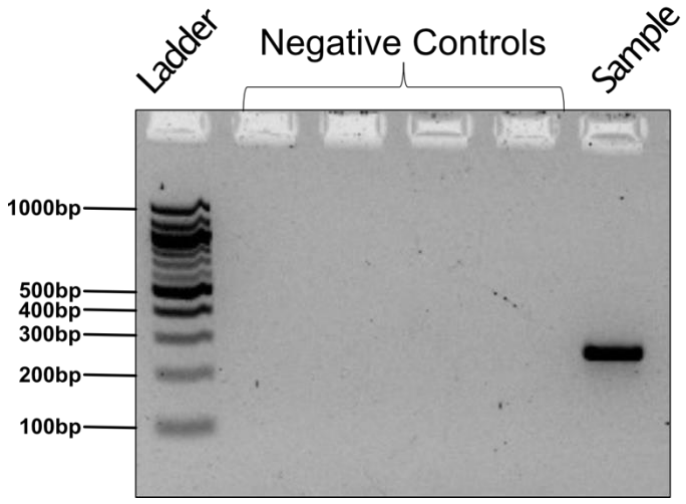


Figure 3.7. Testing negative controls for contamination.
Representative example of an agarose gel where PCR product derived from four wells that did not contain a cell (negative controls) was run besides that of a single cell-containing well. Ladder = 100bps.

3.5. Validation of the Pyrosequencing Assay

At this stage, we had established a pyrosequencing protocol accurate in the measurement of mtDNA heteroplasmy at the single cell level, with the added advantage of being able to detect cross-strain mtDNA contamination. However, it was important to determine the accuracy of pyrosequencing as a technique when it comes to bulk as well as single cell heteroplasmy measurements. At the same time, we were interested in examining the ways in which experimental error is introduced to our pipeline and how it could be accounted for. The first step in this direction was to study whether performing multiple pyrosequencing measurements on the same bulk DNA sample would yield significantly different results. To that end, DNA was extracted from 8 tissues (brain, liver, spleen, gut, muscle, heart, gonads and skin) derived from a P100 mouse carrying the m.5024C>T mutation. The extracted DNA was used to run 16 reactions per tissue and recordings were made of the resulting heteroplasmy measurements **(Figure 3.8.)**. It was clear that repeated measurements of mtDNA heteroplasmy performed on bulk DNA samples obtained from the same tissue yielded consistent results, with a variance of 5.2%.

Having established the heteroplasmy measurement variance we would expect in the presence of large amounts of mtDNA extracted from tissue samples, we decided to take a closer look at the single cell level and examine the effect that single cell lysis can have on mtDNA heteroplasmy quantification. In order to test this, we compared single cells, sorted in the absence of any antibody, derived from an E8.5 mouse embryo carrying the m.5024C>T mutation **(Figure 3.9.)**. Estimating the bulk embryo heteroplasmy through the pyrosequencing of six 14-cell wells, it was not surprising to see that when pyrosequencing entire cells the mean single cell heteroplasmy matched that of the bulk. Subsequently, we decided to lyse the same number of single cells and perform a separate pyrosequencing experiment. This time, we used only one third of the total lysate generated from each single cell. This approach was performed in triplicate, consuming the entirety of the single cell lysate. Each reaction (containing 1/3 of the total cell lysate) was pyrosequenced twice in order to examine whether the measurement variability introduced during sequencing is as minimal at the single cell level as it is in bulk samples **(Figure 3.9.)**. Repeating pyrosequencing measurements on the same single cell sample PCR product, resulted in the production of perfectly correlated heteroplasmy values **(Figure 3.10.A and C)**, in accordance with what we observed in bulk DNA samples **(Figure 3.8.)**.

However, the correlation between pyrosequencing different PCR products derived from sections of the same lysate, appeared to be much lower (in some cases as low as 0.4) **(Figure 3.10.B and C)**. There were not any clear trends that could explain this disparity, as both low and high heteroplasmy cells exhibited the same behaviour. It was interesting however to see that when averaging out the mtDNA heteroplasmy measurements of the 3 lysates that corresponded to each cell, we were left with a population that had smaller variance compared to the initial measurements performed on a fraction of the lysate.

At this stage, we have established that it is the PCR amplification of target sequences and not the pyrosequencing step that is responsible for the introduction of most of the measurement error. This observation held true when it comes to the quantification of mtDNA heteroplasmy in both bulk and single cell samples. Seeing the impact that lysis of a single cell had on mtDNA heteroplasmy reading, we decided to take a closer look at the role that copy number plays in the introduction of pyrosequencing measurement error. Theoretically, the three main factors contributing to the observed variance that characterizes heteroplasmy measurements include experimental error, sampling error and sampling variance. Also, rounding error is introduced as mtDNA heteroplasmy measurements coming out of pyrosequencing are rounded up as whole percentage points, without taking into account the copy number of the sample. While we know that the lower the copy number of the initial sample, the greater the sampling variance will be (always dependent on the true heteroplasmy measurement of the sample), we were interested in testing whether the experimental error is affected by the mtDNA copy number and heteroplasmy percentage. In this section, when mentioning copy number, we refer to the number of mtDNA molecules used at the onset of a PCR reaction and not necessarily the number of molecules found within a cell.

In order to address this question, we decided to make dilutions using a culture day 10 oocyte harvested and grown to this stage by Dr Haixin Zhang. From her work and that of others in the field, oocytes of that stage are expected to possess around 100,000 copies of mtDNA (Zhang et al., 2021). The oocyte, previously stored at -80^{o}C, was lysed and serial dilutions were performed on the lysate in order to run pyrosequencing PCR reactions with templates consisting of increasing mtDNA copy number each time, all deriving from the same single cell **(Figure 3.11.)**. The serial dilutions were carried out on an initial lysate of 10μl in order to obtain samples of varying mtDNA copy number. Labelled as bulk are measurements made on

that initial lysate. In every case, exact measurements of mtDNA copy number were made using ddPCR.

The mean mtDNA heteroplasmy across the different groups remained constant, once again underlining that any deviations from the overall mean are not the product of copy number but rather, that of external contaminants as previously discussed. However, we did observe that intra-group variability was higher the lower the mtDNA copy number of the measured sample was **(Figure 3.11.)**. Interestingly enough, the variability seen in bulk measurements was comparable to the one observed in the samples that had undergone 1:10 dilution, suggesting that the exact mtDNA copy number used at a template for the pyrosequencing PCR may not have such a huge impact on the variance. Importantly, this demonstrates that in cells with a copy number of 500, the pyrosequencing technique is able to provide measurements as accurate as those seen in bulk DNA.

Finally, we asked the question: how much of the variance observed in mtDNA heteroplasmy measurements can be explained as a product of sampling effects and low levels of copy number. In order to address this question, a computational model was developed by Dr Lyuba Bozhilova, where the observed heteroplasmy of each sample was viewed as a function of the true heteroplasmy of the oocyte, the copy number for that specific measurement and the experimental error involved. In this model, it is important to keep in mind that heteroplasmy measurements are being reported as percentage points (two significant digits) which introduces noise to the calculations. Reaffirming our previous conclusions, the observed variance was greater when lower levels of mtDNA were used as template in the PCR reaction **(Figure 3.12.A)**. At the same time, the observed variance was strongly correlated with both the theoretical and experimental variance (Pearson's ρ = 0.94 and 0.84 respectively) **(Figure 3.12.B)**. Lastly, the greater the mtDNA copy number of the sample, the lower the experimental variance was **(Figure 3.12.C)**.

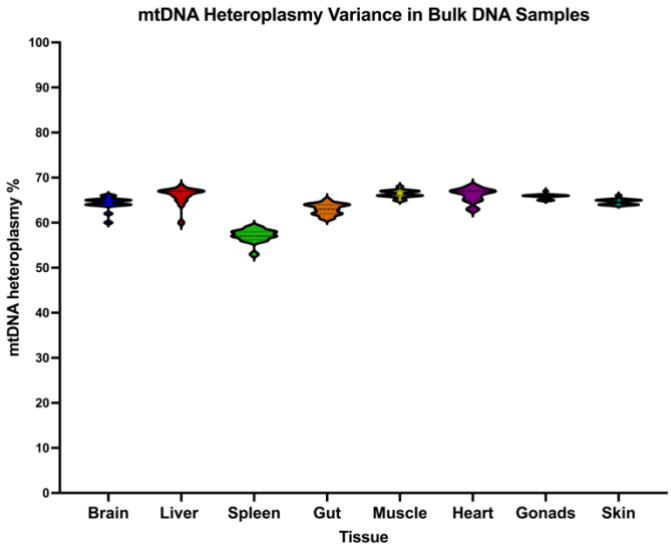


Figure 3.8. Mitochondrial DNA heteroplasmy measurements on bulk tissue samples.
Violin plots of repeated pyrosequencing measurements of mitochondrial DNA (mtDNA) heteroplasmy on tissues derived from adult (P100) mice carrying the m.5024C>T mt-Ta mutation. Each run consists of one measurement conducted on DNA extracted 8 different tissues derived from the same animal, including: brain, liver, spleen, gut, skeletal muscle, heart, gonads and skin. n=16; data presented as median and IQR.

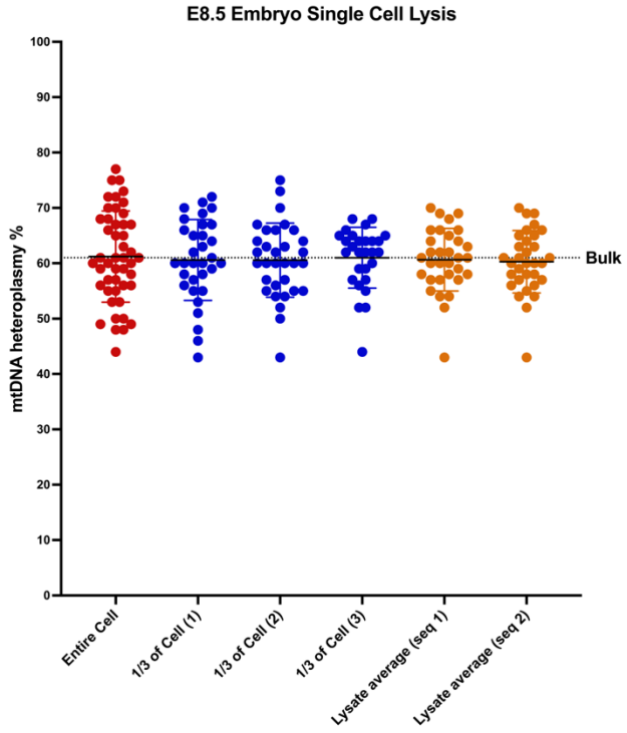


Figure 3.9. The effect of single cell lysis and repeated sequencing on mitochondrial DNA heteroplasmy quantification.
Bee swarm plots of mitochondrial DNA heteroplasmy measurements on single cells derived from an E8.5 embryo carrying the m.5024C>T mutation. Heteroplasmy measurements were conducted on either entire single cells or a lysate fragment (1/3) of a cell, all obtained from the same embryo. At the same time, each of the 3 lysate repeats was pyrosequenced twice (seq1 and 2) and averaged out. Highlighted in red are heteroplasmy measurements on entire single cells; in blue are measurements on lysed cells; in orange are averaged pyrosequencing measurements of the single cell lysates.

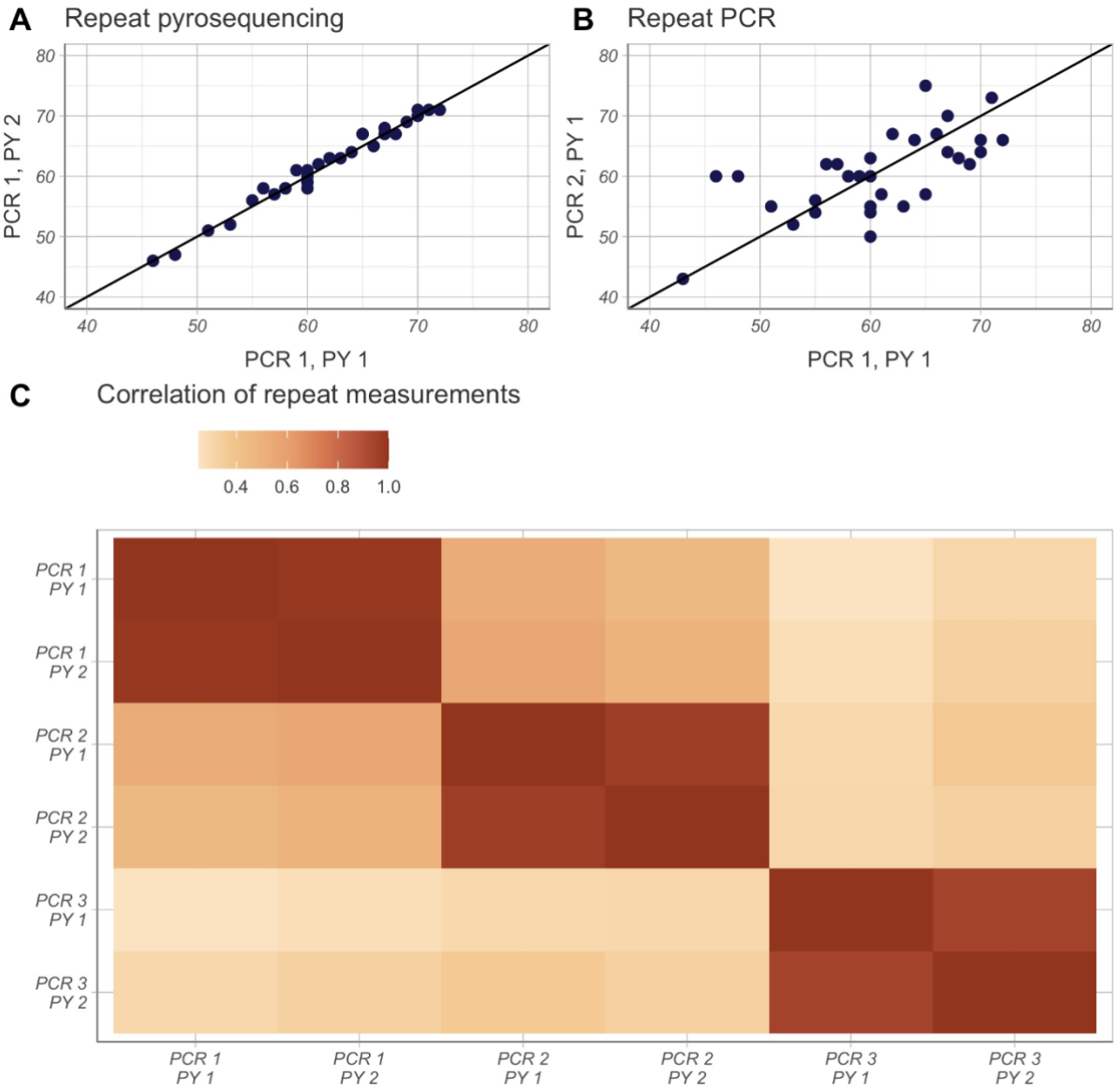


Figure 3.10. Relationship between repeated pyrosequencing and PCR measurements.
(A) Scatter plot of mitochondrial DNA (mtDNA) heteroplasmy values obtained through repeated pyrosequencing performed on the sane pyrosequencing PCR product. **(B)** Scatter plot of mtDNA heteroplasmy values obtained by carrying out multiple pyrosequencing PCR reactions on the lysate obtained from a single cell. **(C)** Heat map showing the correlation between mtDNA heteroplasmy measurements obtained through repeated pyrosequencing as well as PCR. All data presented in here are derived from repeated PCR runs and pyrosequencing measurements on samples obtained from the lysis of single cells derived from E8.5 embryos carrying the m.5024C>T mutation. Figure generated by Dr Lyuba Bozhilova.

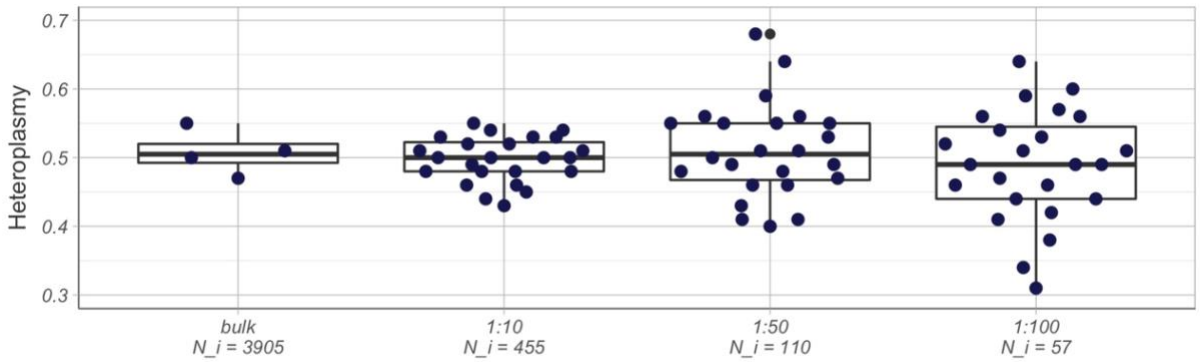


Figure 3.11. The effect of copy number on heteroplasmy measurements.
Repeated pyrosequencing measurements of mitochondrial DNA (mtDNA) heteroplasmy performed on serial dilutions of lysate deriving from a single culture day 10 oocyte carrying the m.5024C>T mutation. Each dilution is matched to digital droplet PCR (ddPCR) measurements of the corresponding mtDNA copy number. Each dot represents a single repeat of the pyrosequencing PCR reaction. In the case of bulk, measurements were made on the original, undiluted oocyte lysate. N_i=mtDNA copy number. Figure generated by Dr Lyuba Bozhilova.

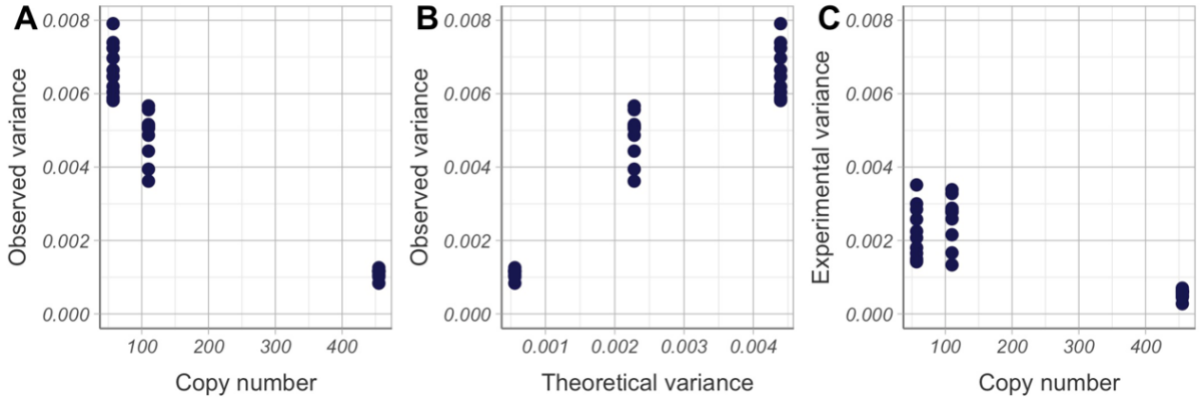


Figure 3.12. Estimates of variance in single cell mitochondrial DNA heteroplasmy measurements.
(A) Copy number of mitochondrial DNA (mtDNA) used as a template at the start of the pyrosequencing PCR amplification step (x-axis) and observed variance in mtDNA heteroplasmy measurements (y-axis). **(B)** Scatter plot of the theoretical variance in mtDNA heteroplasmy measurements that we would expect to see if sampling and rounding errors are taken into account (x-axis), and the observed variance that we see in these samples (y-axis). **(C)** Scatter plot of the copy number used as a template for the pyrosequencing PCR reaction and the experimental variance of the mtDNA heteroplasmy results. Figure generated by Dr Lyuba Bozhilova.

3.6. Key Findings

- Pyrosequencing employed as a technique to measure mtDNA heteroplasmy at the single cell level in-house at a high throughput rate, exhibits sequence-dependent disparities.
- Even though pyrosequencing PCR contamination is insignificant when dealing with the amplification of mtDNA derived from bulk tissue samples, it becomes a significant concern when it comes to mtDNA heteroplasmy measurements at the single cell level.
- Steps taken to eliminate pyrosequencing PCR contamination when working with single cell samples include stringent working practices set in place during reaction preparation, the inclusion of empty wells in each plate that serve as negative controls, as well as the use of pyrosequencing programs that measure heteroplasmy at both the m.5019 and m.5024 position simultaneously.
- When it comes to the accuracy of pyrosequencing as an approach to measuring mtDNA heteroplasmy at the single cell level, lysing a single cell and averaging out the mtDNA heteroplasmy measurement of each fraction of the lysate, while impractical, gives a slightly more accurate representation of the cell's mtDNA heteroplasmy.
- Repeated pyrosequencing of the same PCR product appears to not provide significantly different results in either bulk-DNA or single cell samples. As a result, we can assume that any measurement error involved in mtDNA heteroplasmy measurement is introduced during the sampling and amplification steps.
- The lower the copy number used as a template for the pyrosequencing PCR, the more variable the mtDNA heteroplasmy reading are. Suggesting that lysing the cell and using a portion of the lysate to carry out the reaction will result in more inaccurate results that using the entirety of the cell.
- Pyrosequencing measurements of mtDNA heteroplasmy derived from cells containing high copy numbers are more reliable than ones coming from cells containing low copy number.
- Since about half of the observed variance is a product of sampling and rounding effects, using the entirety of a cell will give us a more accurate estimate of its mtDNA heteroplasmy.

Chapter 4. Results:

Progressive Increase of mtDNA Mosaicism Through Life Uncovered by High-Throughput Single Cell Analysis

4.1. Introduction

Respiration resides at the epicentre of cellular energy production. Hence, the presence of mutations in either the nDNA or mtDNA that result in the disruption of cellular respiration have far-reaching consequences throughout the human body (Gorman et al., 2016). The complex genetic nature of mitochondrial diseases means that different types of defects can disproportionally affect certain organs over others, eventually resulting in a diverse pathological landscape (Pacheu-Grau et al., 2018). While mitochondrial diseases make up one of the most prevalent types of inherited neurological pathologies, their heterogeneous nature makes diagnosis particularly challenging (Gorman et al., 2016). At the heart of this issue lies a topic that we have already been extensively discussed: the need for interaction between the two distinct and largely independent genomes that guarantees much-needed consistency in mitochondrial bioenergetic output. Tissue-specific energetic demands are widely hypothesised to be the reason why certain tissues are preferentially affected over others by mitochondrial dysfunction (Wang et al., 2010). Unsurprisingly, tissues with high energetic demands tend to be the first ones to suffer from conditions that directly hinder energetic output at the cellular level (Gorman et al., 2016).

When considering mitochondrial disorders arising as a result of pathogenic mtDNA mutations, the concept of heteroplasmy suddenly becomes central to the discussion, as the proportion of mutant to wild type copies of mtDNA will often dictate the severity of the phenotype. When heteroplasmy levels are lower than the threshold level of mutation, there is no disease manifestation. However, once this threshold is exceeded, biochemical defects begin to appear and the clinical phenotype manifests (Taylor and Turnbull, 2005). The majority of late-onset mitochondrial disorders are brought about by a gradual accumulation of the pathogenic mutation along the processes of development and ageing (van den Ameele et al., 2020). It is well documented that mutations such as the mt.tRNA$^{Leu(UUR)}$ m.3243A>G accumulates in patient muscle tissue while being selected against in the blood (tHart et al., 1996; Chinnery et al., 1999; Rahman et al., 2001; Pyle et al., 2007; Rajasimha et al., 2008). The same segregation pattern has been observed in other point mutations (Howell et al., 2000; Jacobi et al., 2001), including mt.tRNA$^{Leu(CUN)}$ m.12320A>G (Weber et al., 1997), as well as large-scale deletions (Larsson et al., 1990). In the case of the m.3243A>G mutation, the decrease observed in blood over time is exponential rather than linear in nature (Rajasimha et al., 2008), suggesting that

there is a consistent clearance of defective mitochondria from within the stem cell pool (van den Ameele et al., 2020). As conclusive the observations on m.3243A>G are, one must always keep in mind that the relationship between heteroplasmy segregation across different tissues and ageing is not always that straightforward. Indeed, is has been suggested that the behaviour seen in m.3243A>G and large-scale deletions might not be the norm, but rather, an exception. Other point mutations such as mt.tRNA^{Lys} mt.8344A>G as well as ATP6 m.8993T>C are not selected against in the haematopoietic lineage during aging (Rajasimha et al., 2008; White et al., 1999).

Shifts in the segregation of mtDNA heteroplasmy in various tissues in development and aging are subject to the control of different cellular mechanisms. These mechanisms act on different levels, from that of mtDNA replication and repair, to the degradation of entire organelles and the downstream impact this process has of cellular viability (van den Ameele et al., 2020). When it comes to the preferential selection in favour of a mutation as a result of differential mtDNA replication or repair, the concept of selfish replication emerges. While one might expect neutral variants that have no effect on mitochondrial performance to escape the influence of any selective forces (Samuels et al., 2013), there is evidence demonstrating that variants detrimental to mitochondrial activity have a replicative advantage compared to their WT counterparts (Ma and O'Farrell, 2016). In the case of the mt.3243A>G SNP seen in MELAS, the selfish replication is thought to be a product of lower levels of mTERF binding that in turn limits the amount of replication-pausing on the molecule (Hess et al., 1991; Hyvarinen et al., 2007). Selfish replication has been also noted in mtDNA molecules that carry large-scale deletions. An observation that was initially explained by the shorter replication time that would be required to copy a largely reduced genome (Moraes et al., 1999; Diaz et al., 2002; Russell et al., 2018). However, the fact that a similar benefit was bestowed by more limited deletions (Campbell et al., 2014) as well as the long half-life of the mtDNA molecule (Kowald et al., 2014) called that theory into question. mtDNA repair pathways such as base-excision repair employed in response to oxidative damage (Thorslund et al., 2002) as well as microhomology-mediated end-joining used in the correction of DSBs (Ma, and O'Farrell, 2015; Tadi et al., 2016) have been shown to operate at different levels across different tissues or even regions of the same organ (Imam et al., 2006).

In addition to differences in mtDNA replication and repair that might drive the somatic segregation of heteroplasmy, mitochondrial heterogeneity can arise by forces acting on the process of their biogenesis. In more detail, the replicative advantage described above is not a characteristic shared by most pathogenic mtDNA variants. Instead, difference concepts have been described over the years to explain the accumulation of pathogenic mtDNA variants in both quiescent and rapidly dividing tissues (Chinnery and Samuels, 1999; Johnston et al., 2015). First, vegetative segregation means that as cells divide, the random partitioning of mtDNA molecules over generations could result in certain cell lineages inheriting different heteroplasmy levels compared to others. At the same time, relaxed replication, a process describing the continuous synthesis and degradation of mtDNA independent of the cell cycle, can over time, become a driving force in heteroplasmy dynamics even in the absence of cell divisions **(Figure 4.1.)**. When modelled *in silico*, both of these mechanisms have been predicted to be major contributors in driving heteroplasmy shifts through random genetic drift during the lifespan of an organism (Stewart and Chinnery, 2015). The process of mtDNA segregation can be further influenced or even accelerated by selective forces against or in favour of a certain haplotype are introduced.

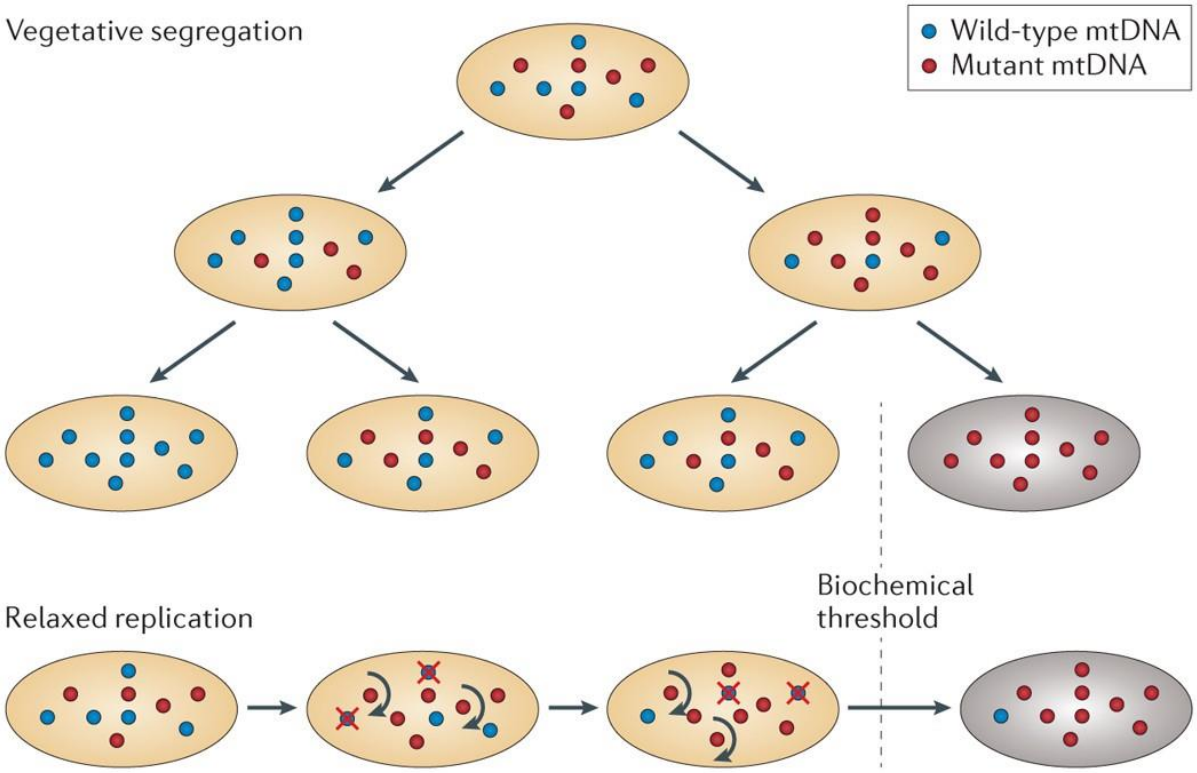


Figure 4.1. Forces behind mitochondrial DNA random genetic drift.
Upon cell division, the random partitioning of mutant (red) and wilt-type (WT) (blue) mitochondrial DNA (mtDNA) molecules takes place. This results in daughter cells inheriting different proportions of mutant to WT copies of mtDNA, thus shifting heteroplasmy either up or down compared to that of the parent cell. This process is termed vegetative segregation. At the same time, throughout a cell's lifespan, mtDNA molecules are continuously and randomly being selected to undergo replication, in a process termed relaxed replication. If these processes lead to the preferential amplification of one variant over the other, shifts in mtDNA heteroplasmy will manifest. It has been computationally predicted that these two mechanisms are active throughout an individual's lifetime and that they take place at the same time. Figure adapted from: Stewart and Chinnery, 2015.

Selective mitochondrial degradation is another mechanism that can potentially alter the mtDNA segregation dynamics. This takes place through the activation of intracellular QC pathways responsible for the degradation of old or dysfunctional organelles in order to remove them from the network. Mitophagy, a specialised form of autophagy is responsible for the clearance of dysfunctional mitochondria alongside their mtDNA (Pickles et al., 2018). Recent work on both fruit fly (*D. melanogaster*) and mouse models has shown that mitophagy is a common process not restricted to pathological states, that is employed consistently across different tissues throughout the animal's lifespan (Sun et al., 2015; McWilliams et al., 2016 Lee et al., 2018). Furthermore, when discussing mechanisms that bring about the clearance of mutant mtDNA, one has to keep in mind the impact that mitochondrial dynamics have on the mtDNA heteroplasmy segregation. More specifically, the fact that mitochondrial do not exist

in isolation but are part of a highly complex and dynamic network means that fusion events can bring about mitochondrial functional complementation (Nakada et al., 2009), a compensatory response describing a heteroplasmic state where WT molecules can complement the lacking activity of their mutant counterparts. Of course, such activity is promoted by the organisation of organelles in a fused network (Nakada et al., 2009). Finally, evidence derived from the *in vitro* culture of myoblasts derived from KSS patients suggests that there is a rapid decrease in the levels of mtDNA carrying the deletion through passages (Lee et al., 2018). This observation was corroborated *in vivo* as damaged muscle tissue belonging to KSS patients showed signs of regional rescue once recovery was complete (Clark et al., 1997). Such observations can be explained either through the out-competing of cells with higher mutational loads by their healthy counterparts or through the presence of organellar and developmental bottlenecks that favour the replication of WT molecules over mutant ones (Mitra et al., 1009; Mandal et al., 2010). In every case, it is understood that multiple selective forces shape the pattern of heteroplasmy segregation at the molecular, organellar and cellular level.

4.2. Project Aims

In an attempt to better understand the dynamics of mtDNA heteroplasmy and the ways in which variants are selected for or against in certain tissues, Ana Victoria Lechuga-Vieco and her colleagues employed engineered heteroplasmic mice that carried two non-pathogenic mtDNA variants, C57BL/6N and NZB on a C57BL/6N nuclear background (Lechuga-Vieco et al., 2020). Their work, carried out at the level of the tissue, displayed the presence of tissue-specific selective forces that were either driving for the accumulation of the C57BL/6N (eg. blood, gut and spleen) or NZB (eg. brain, liver and kidney) haplotype, with certain tissues (eg. tail, eye, skeletal muscle and heart) displaying no selection during the course of development (Lechuga-Vieco et al., 2020). This dynamic process of non-random mtDNA haplotype segregation was shown to be the product of intracellular organelle selection (Lechuga-Vieco et al., 2020). The decision regarding which haplotype is more favourable for the tissue is influenced both by the cell-type specific impact of OXPHOS restriction, the cell's response to environmental impact as well as the strain's nuclear background (Lechuga-Vieco et al., 2020). Further work on the heterozygous mouse model carrying both the C57BL/6N and NZB mtDNA haplotypes was carried out by Katiane Tostes and colleagues (Tostes et al., 2022). Their results, in line with those of Lechuga-Vieco et. al., showed an age-dependent accumulation of NZB in the liver, enhancing the respiratory capacity per mtDNA molecule in the tissue (Tostes et al., 2022). Liver-specific autophagy related 7 (*atg7*) and parkin RBR E3 ubiquitin protein ligase (*prkn*) knockouts abolished the selection for NZB and instead, established a neutral drift more akin to that seen in tissues such as the eye and tail (Tostes et al., 2022). Overall, Tostes et al., provide further evidence in support of organelle selection at the cellular level being the driving force behind haplotype selection in a heteroplasmic context, with macroautophagic clearance of the less fit mtDNA haplotype being the mechanism in the liver (Tostes et al., 2022).

As informative as these two studies are in expanding our understanding of mtDNA heteroplasmy dynamics and the driving forces that govern its segregation in different tissues, it is important to keep in mind that both the C57BL/6N and NZB mtDNA haplotypes are non-pathogenic (Lechuga-Vieco et al., 2020). Consequently, in order to study the dynamics governing heteroplasmy segregation in the case of pathogenic mitochondrial mutations, we made use of the m.5019A>G and m.5024C>T mouse models generated in the lab of James B. Stewart and introduced in more detail in **Result Section, Chapter 5**. Both mouse models carry

a mutation in the acceptor stem of mt-Ta. Even though the two mutations are just 5 base pairs apart, their effects on the mt-Ta molecule are quite distinct. While high levels of m.5019A>G led to an increase proportion of uncharged mt-Ta molecules, the m.5024C>T mutation has been shown to cause molecular instability leading to degradation of the mt-tRNA molecule (Kauppila et al., 2016). The heteroplasmy selection pattern reported in the m.5024C>T mouse model while being quite different from what was seen in the heterozygous C57BL/6N and NZB mouse, bore close resemblance to ones characterising human mitochondrial DNA disorders. In more detail, heteroplasmy measurements performed in 7 different tissues derived from young mice (~20 weeks old) were very close to each other as well as to the ear biopsy samples that were collected upon weaning (~3 weeks old) (Kauppila et al., 2016). However, selection was observed in the rapidly dividing tissues, such as blood and colonic epithelium, of older mice (>40 weeks old) that had an ear punch biopsy heteroplasmy of >60% **(Figure 4.2)** (Kauppila et al., 2016). Even though the tissue-specific nature of heteroplasmy segregation seen in mouse carrying both the C57BL/6N and NZB haplotypes was not present, decrease in mutation burden of rapidly dividing tissues, namely peripheral blood, is not without precedent. Such selection characterises certain human mtDNA disorders that include the MELAS-causing m.3243A>G mutation (Ciafaloni et al., 1991) and CPEO (Larsson and Clayton, 1995).

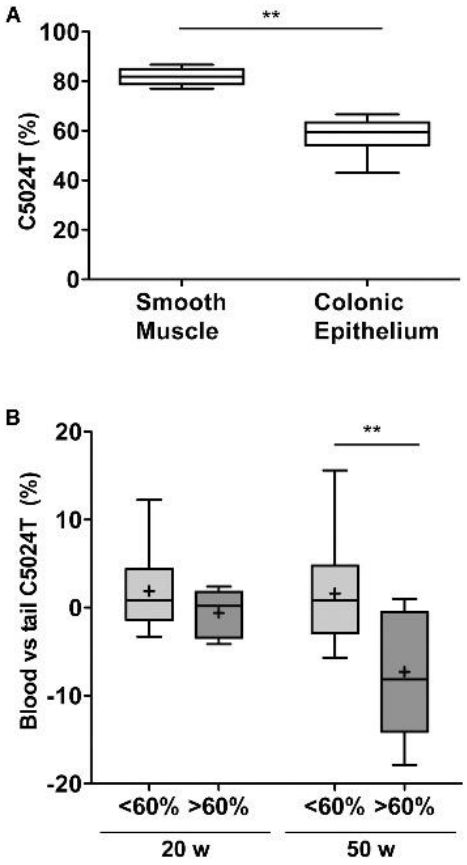


Figure 4.2. Blood and colonic epithelium of aged mice display a selective loss of the mt-Ta 5024C>T mutation.
(A) Laser capture dissection was employed to separate the colonic epithelium from the smooth muscle of the same mouse. The relative levels of the m.5024C>T mutation were measured and a clear loss of the mutation was identified in the colonic epithelium. n = 8. $^{**}p < 0.01$ (Wilcoxon matched-pairs signed rank test). **(B)** Relative levels of the mt-Ta m.5024C>T mutation in blood samples normalised to tail. While at 20 weeks of age there is no difference between low (<60% in other tissues) and high (>60% in other tissues) heteroplasmy animals, at 50 weeks, high heteroplasmy ones show a significant decrease in blood mutation burden. $^{**}p < 0.01$ (Mann-Whitney *U* test). Figure adapted from: Kauppila et al., 2016.

Another aspect of the aforementioned studies we sought to address is that all heteroplasmy measurements were carried out at the level of the tissue, thus potentially masking any cell-type specific differences in heteroplasmy selection. For many years, histopathological examination using COX/SDH staining has been used to provide a cytochemical examination of respiratory complex IV activity at the level of individual cells **(Figure 4.3)** (Taylor and Turnbull, 2005). This methodology demonstrates that when the mtDNA heteroplasmy threshold specific for each mutation is exceeded, a subsequent drop in complex IV activity is brought about (Ross, 2011). At the same time, the activity of SDH, or respiratory complex II, which is in its entirety encoded within the cell's nucleus, typically remains unscathed (Ross, 2011). Similar observations were made by lab of James B. Stewart when working on both the m.5024C>T

(see **Results Section, Figure 3.1**) and m.5019A>G mouse model. COX/SDH staining reveals a mosaic pattern in most tissues that suggests that mitochondrial disease heterogeneity spans beyond the tissue level, down to the single cell level, with certain cells of the same tissues displaying complex IV deficiency while their neighbour does not.

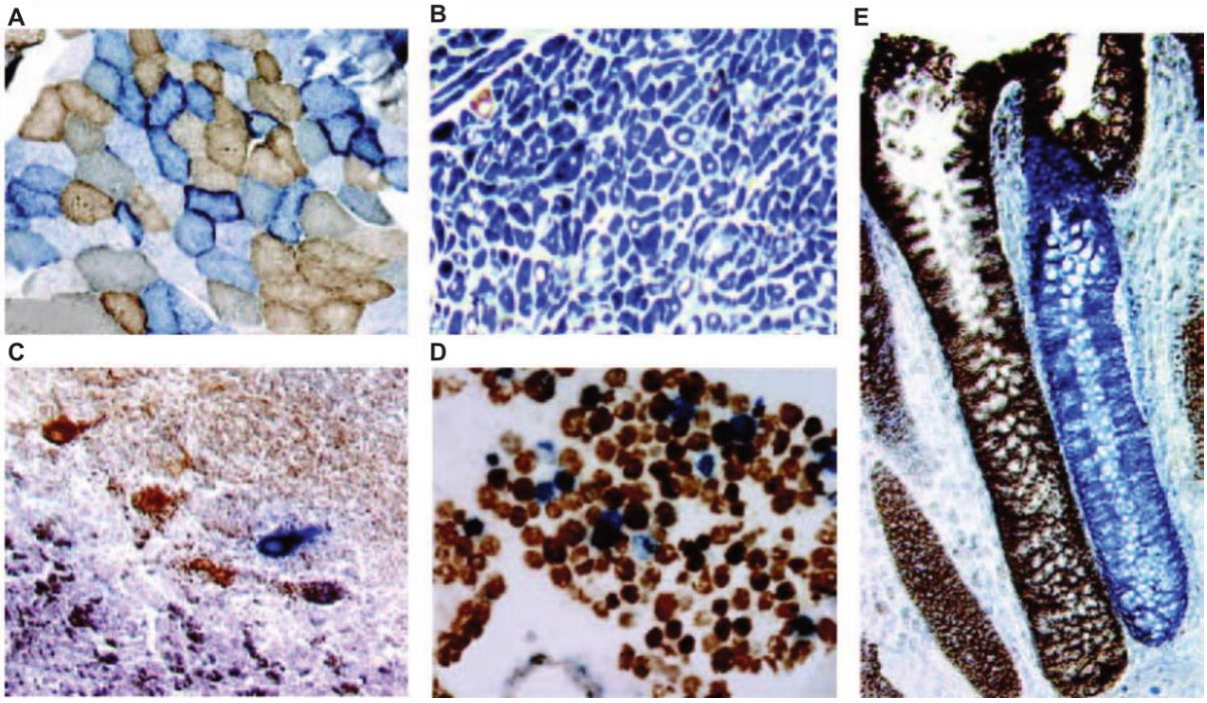


Figure 4.3. Mitochondrial DNA deficient in cytochrome c oxidase is associated with ageing and disease.
Tissue cross-sections derived from human patients, treated to exhibit both COX and SDH activity. Cells with unaffected COX activity are stained brown while those with COX deficiency are stained blue. **(A)** Skeletal muscle sample derived from a patient with a heteroplasmic point mutation affecting mt-tRNA. The typical pattern of 'mosaic' COX activity is clearly visible in this section, with muscle fibers that exceed the heteroplasmy threshold exhibiting the expected respiratory complex IV deficiency as a result of mutant mtDNA accumulation. **(B)** Cardiac tissue cross section displaying the left ventricle of a patient who suffered from hypertrophic cardiomyopathy as a result of a homoplasmic mt-tRNA mutation that is further supported by the lack of COX activity in the vast majority of cells in this section. **(C)** Cerebellum section derived from a patient carrying a rearrangement of the mtDNA. The presence of COX-deficient neurons can be spotted. **(D and E)** Tissue cross-sections that make the case for COX deficiency being a product of the clonal expansion of mtDNA mutations within somatic cells at the single cell level. This phenomenon is observed both in post-mitotic quiescent cells such as extraocular muscles (**D**) as well as rapidly dividing tissues such as the colonic epithelium (**E**) in ageing individuals. Figure adapted from: Taylor and Turnbull, 2005.

The presence of such a heterogeneous pattern of cytochemical activity within affected tissues hints towards either differences in the susceptibility of different cells to mitochondrial damage or, exceptional heterogeneity in heteroplasmy levels between cells belonging to the same tissue. Up to date, the work being done towards uncovering the principles and pathways governing heteroplasmy dynamics focuses solely at the tissue level, overlooking cell-type specific differences as well as the possibility that a minority of cells with high heteroplasmy might skew any bulk measurements taken. At the same time, there is strong evidence suggesting that mtDNA mutations that impact a small portion of cells can be a contributing factor to common late-onset multifactorial diseases, as seen in quiescent (non-dividing) neurons of the *Substatia nigra* in Parkinson's disease (Coxhead et al., 2016; Arthur et al., 2009). Thus, understanding heteroplasmy dynamics at the single cell level could paint a more nuanced picture of the bioenergetic reality within the affected tissues, opening the door to new, more effective, approaches to treatment.

In this chapter, I present work on obtaining different single cell populations from the brain and spleen of m.5019A>G and the m.5024C>T mouse models of mitochondrial disease belonging to different age groups, in an attempt to characterise heteroplasmy dynamics in a high throughput approach. First, I describe my work in obtaining different cell population from the mouse brain and spleen using surface epitope staining with fluorescent antibodies followed by FACS. Then, I detail the single cell heteroplasmy measurements performed on single cell population derived from adult mouse brain and spleen, using the protocol described and optimised in **Results Section, Chapter 3.**. Moreover, I demonstrate how both pathogenic mutations follow the principles of random genetic drift as variance of single cell heteroplasmy measurements increases with age while the mean still matches bulk tissue measurements performed at weaning. I highlight the discrepancies in single cell variance with age and how they differ across the two mutations. Finally, I compare the different cell populations obtained from each tissue through different timepoint in development, showing how there are no significant differences in single cell heteroplasmy between populations derived from the same tissue.

4.3. Obtaining Different Cell Populations from Mouse Brain and Spleen

The difference in heteroplasmy level of pathogenic mtDNA mutations observed between mitotic and post mitotic tissues derived from both aged humans and mice has been well documented. Previous work performed by Kauppila et al., provides further support that indeed the m.5024C>T mutation undergoes negative selection in the blood and colonic crypts of aged mice (~50 weeks old). As a result, when we sought to take a closer look at the single cell heteroplasmy dynamics of both m.5019A>G and m.5025C>T mouse models, we wanted to make sure that both dividing and post-mitotic cell types were included in our analysis. To that end, we first aimed at isolating single immune cells from the spleen. Being the largest secondary lymphoid organ, the spleen is responsible for a wide host of immunologic activities in addition to its contribution to red blood cell clearance and heamatopoesis (Lewis et al., 2019). As a result, it has long been used for the isolation of immune cells that are inherently mitotic in nature (Lim et al., 2016; Coquery et al., 2012). At the same time, we selected brain as an example of a tissue that primarily consists of post mitotic cells, mainly in the form of terminally differentiated (TD) neurons (Kruman, 2004; Nouspikel and Hanawalt, 2002). Settling on these two tissues that provided us with a much-needed contrast between cells that were both the mitotic and postmitotic in nature, while deriving from the lateral plate mesoderm in the case of spleen and neural ectoderm in the case of the brain. The outline of the pipeline established from sample collection to measuring heteroplasmy at the single cell level can be found in **Figure 4.4.**.

Dissociation of both organs was performed according to the respective protocols and using the reagents described in **Methods Section, Chapter 2.2.2.**. FACS analysis was used for the isolation of events that landed within the expected parameters of forward scatter (FSC-A) and side scatter (SSC-A) each being a rough estimate of event size and granularity respectively. At the same time, we employed a strategy to ensure that the selected and sorted events included a defined and substantial amount of DNA and were surrounded by an intact cell membrane, using a combination of live dyes named DRAQ5™ and DAPI dyes. DRAQ5 is an intercalating small molecule that fluoresces when bound to DNA, while DAPI only enters cells with a ruptured plasma membrane (see **Methods Section, Chapter 2.3.1.**). Events that met these criteria were termed ‘live cells’ and selected for downstream processing. In the case of dissociating the adult mouse brain, the protocol modification described in **Methods Section, Chapter 2.2.2.2.** had

to be employed in order to account for the lack of a gentleMACS™ Octo Dissociator (Miltenyi Biotec). Evaluation of the 'live cells' obtained following each approach can be found in **Figure 4.5.**.

Having optimised the isolation of single live cells from the brain and spleen of adult mice (**Figure 4.6.A** and **B** and **Figure 4.7.A** and **B**), we were keen to test whether different cell populations, key to each tissue could be isolated using FACS in order for their heteroplasmy percentages to be evaluated independently. Thus, we decided to add conjugated antibodies to the cell suspension of each tissue, targeting established cell surface markers, expressed by the cell populations of interest. In the case of the spleen, we employed an antibody targeting CD-19 **(Figure 4.6.D)**, a very well-established cell surface marker ubiquitous in all B-cells (Wang et al., 2012). Two populations were visible, one positive and one negative for CD-19. A breakdown of the cell populations found in the mouse spleen reveals that besides B-cells, T-cells are the second most abundant cell type within the tissues (21-25%), with monocytes, granulocytes, dendritic cells, natural killer cells and macrophages being found at significantly lower numbers (Bronte and Pittet, 2013). The significant presence of T-cells in the CD-19 negative cell population was confirmed **(Figure 4.7.D)**.

When it comes to the brain however, selecting cell surface markers was more complicated. Eventually, we decided to make use of ACSA-1 and employ a conjugated antibody specifically targeting GLAST, an astroglial-specific epitope (Yuan et al., 2011), in order to identify astrocytes in both the adult and neonate mouse brain. At the same time, a conjugated antibody against Prominin-1 (CD133) was also used in neonate mouse brain samples in order to identify neural stem and progenitor cells (Olausson et al., 2014; Walker et al., 2013) **(Figure 4.8.D)**. However, this population is particularly scarce in the adult mouse brain and instead, we opted to select for events positive for postsynaptic density protein 95 (PSD95). PSD95 a central component to synaptic maturation as it interacts directly with, stabilizes and traffics N-methyl-D-aspartic acid receptors (NMDARs) and α-amino-3-hydroxy-5- methyl-4-isox-azoleproprionic acid receptors (AMPARs) directly to the postsynaptic cell membrane (Yoo et al., 2019; Coley and Gao, 2018; Leeman et al., 2018). As a result, PSD95 was used in order to select for neurons **(Figure 4.8.E)**. More details regarding the antibodies used and the staining protocol followed can be found in **Methods Section, Chapter 2.3.1.**.

While the cell populations selected could undergo further subdivision into specific cell types, the selection does provide a good indication that broad cell populations abundant and central to the function of our tissues of interest have indeed survived the dissociation process and can be reliably sorted. Additionally, the isolation of specific cell populations allows us to perform intra and inter-tissue comparisons. Overall, by selecting populations that are both abundant and instrumental within their respective tissues we make sure to preserve the image of the tissue as a whole, while still being able to evaluate cell type-specific differences.

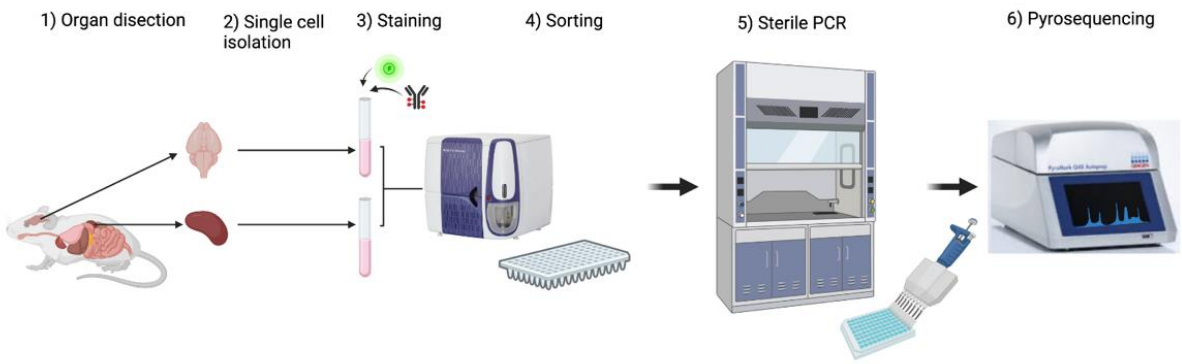


Figure 4.4. Single cell mitochondrial DNA heteroplasmy quantification pipeline.
Schematic representation of the pipeline set up from the point of animal culling and tissue extraction to dissociation, sample labelling, single cell sorting, storage to the final steps of PCR amplification and pyrosequencing. Created with BioRender.com.

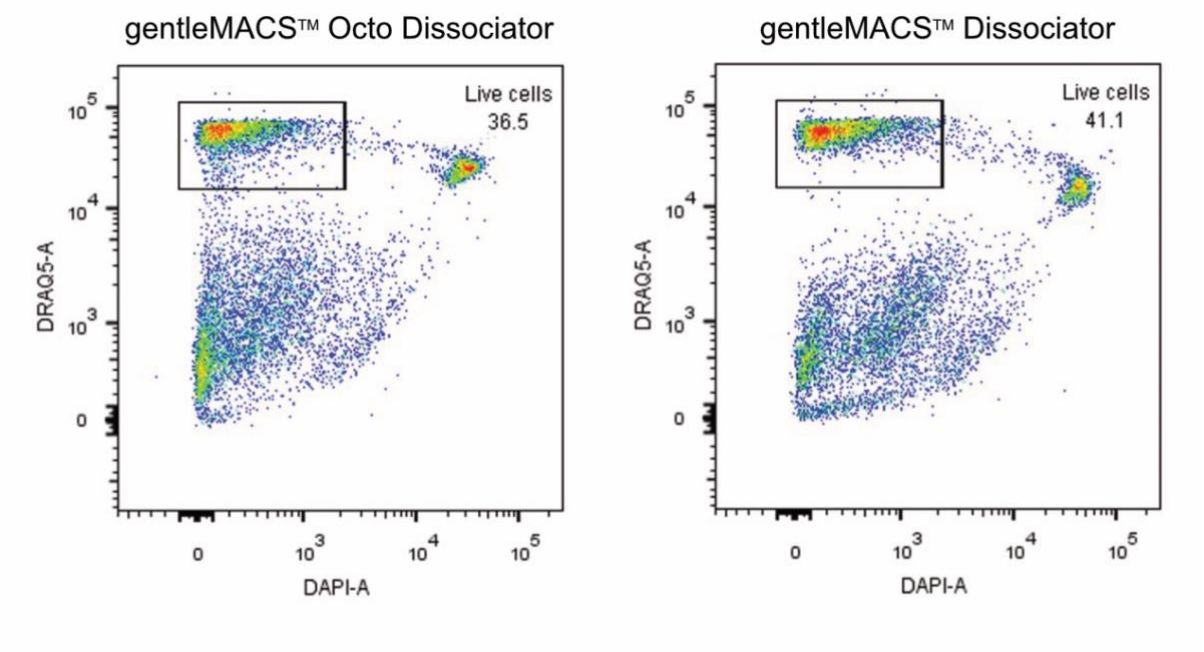


Figure 4.5. Validating adaptations to the adult mouse brain dissociation protocol.
Flow cytometry results comparing the percentage of live cells derived from adult mouse brain samples prepared when the dissociation was conducted using the gentleMACS™ Octo Dissociator (Miltenyi Biotec) versus the newly adapted protocol making use of the standard gentleMACS™ Dissociator (Miltenyi Biotec). Events that contain DNA (positive for DRAQ5™) and that, at the same time, possess an intact cell membrane (negative for DAPI) represent the cells of interest, positioned at the top-left of each figure. Live cells = intact, DNA-containing cells.

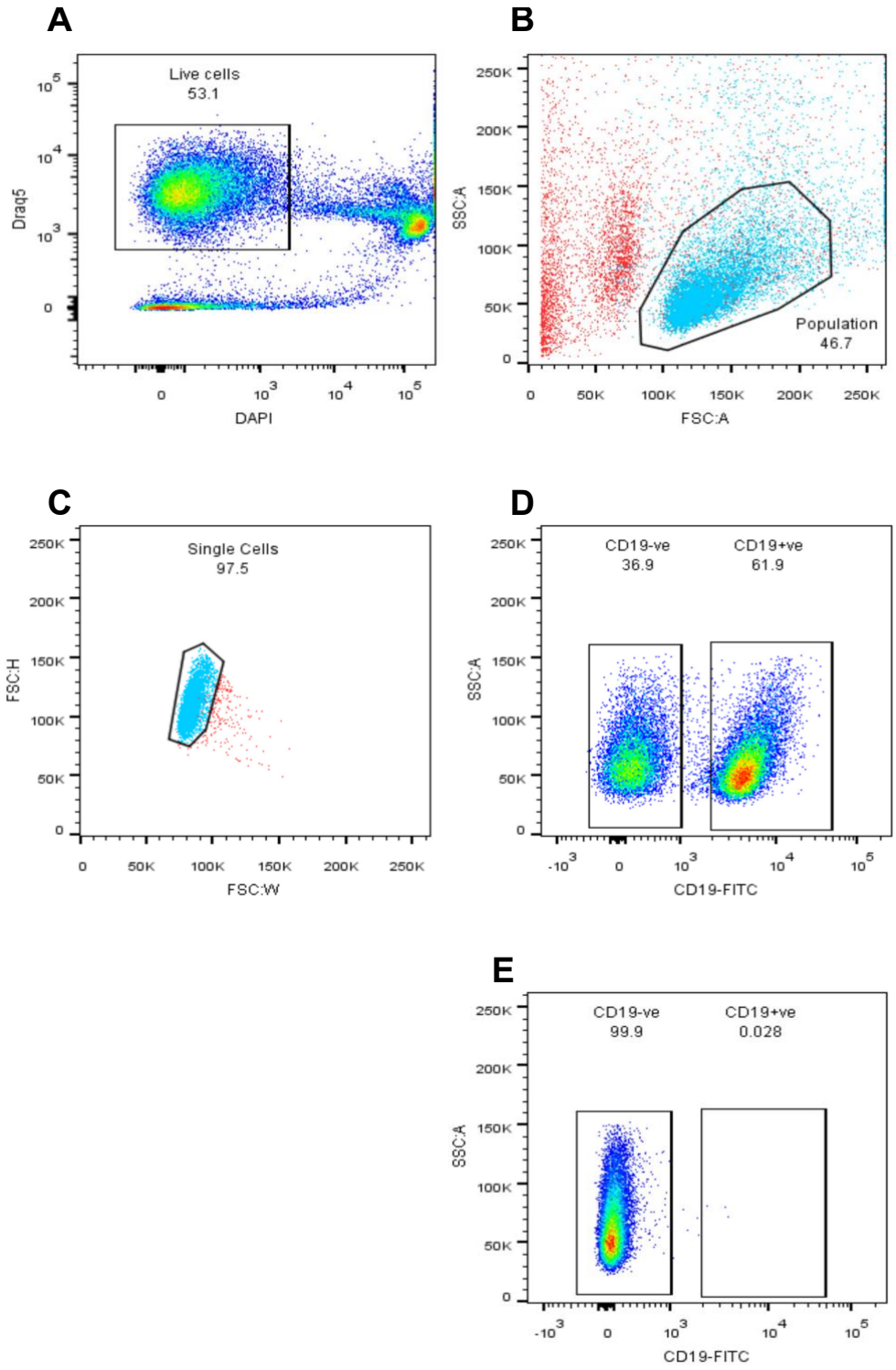


Figure 4.6. B-cell isolation from mouse spleen.
(A) Application of DRAQ5™ and DAPI double staining in order to select for live cells (DNA-containing cells with intact cell membrane). **(B)** Overlay of the live cell population onto the forward over side scatter plot area (FSC-A/SSC-A). **(C)** Doublet exclusion by overlaying forward scatter height over width (FSC-H/FSC-W) to side scatter height over width (SSC-H/SSC-W). **(D)** Discrimination of the cell population positive for CD19 (B-cells) on the basis of Fluorescein isothiocyanate (FITC) fluorescence. **(E)** Unstained spleen sample serving as negative control for FITC fluorescence.

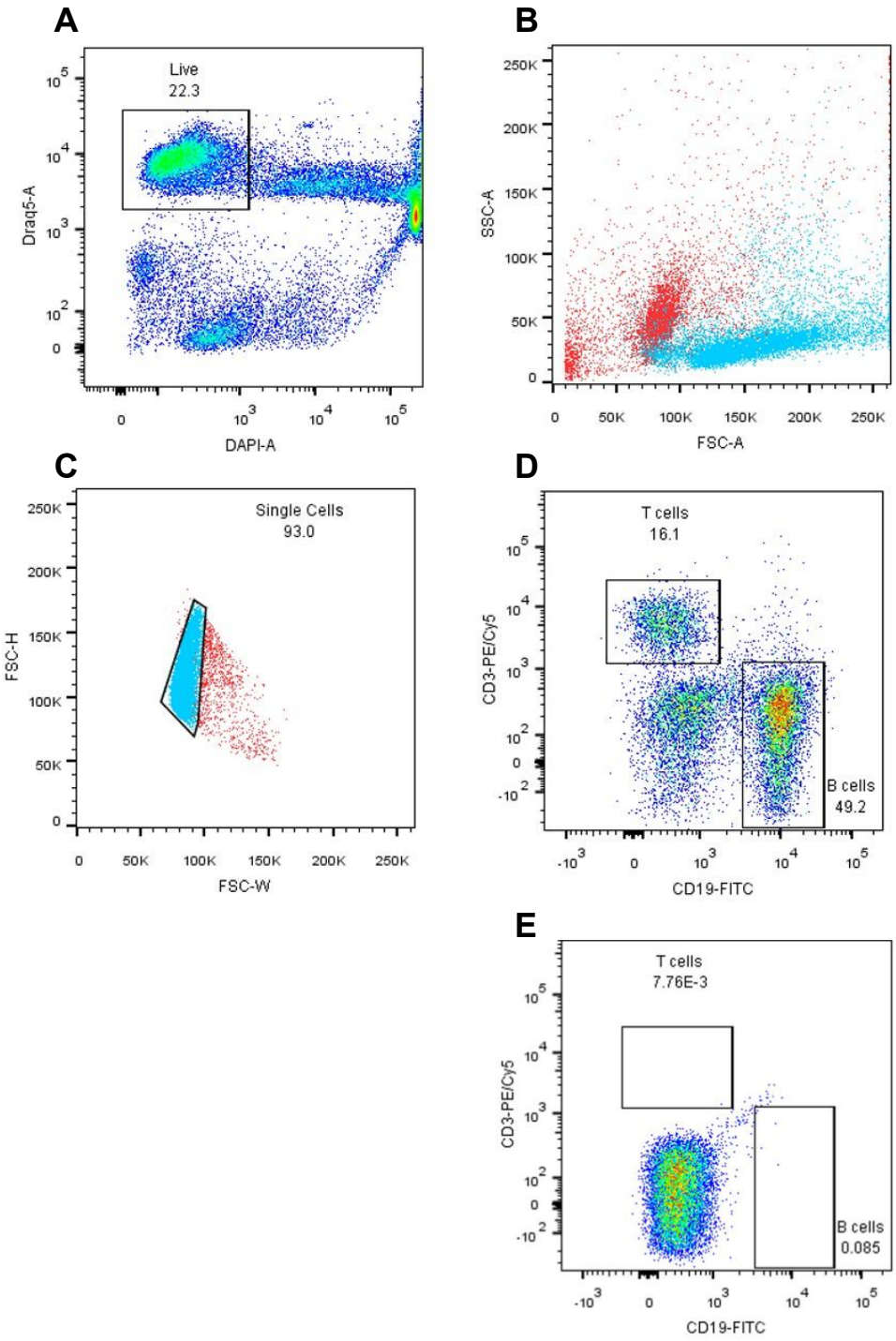


Figure 4.7. Examining the presence of T-cells in the adult mouse spleen.
(A) Application of DRAQ5™ and DAPI double staining in order to select for live cells (DNA-containing cells with intact cell membrane). **(B)** Overlay of the live cell population onto the forward over side scatter plot area (FSC-A/SSC-A). **(C)** Doublet exclusion by overlaying forward scatter height over width (FSC-H/FSC-W) to side scatter height over width (SSC-H/SSC-W). **(D)** Discrimination of cell populations positive for CD19 (B-cells) and CD3 (T-cells) on the basis of Fluorescein isothiocyanate (FITC) and Phycoerythrin (PE)/Cy5 fluorescence. **(E)** Unstained adult mouse spleen sample serving as negative control for FITC and PE/Cy5 fluorescence.

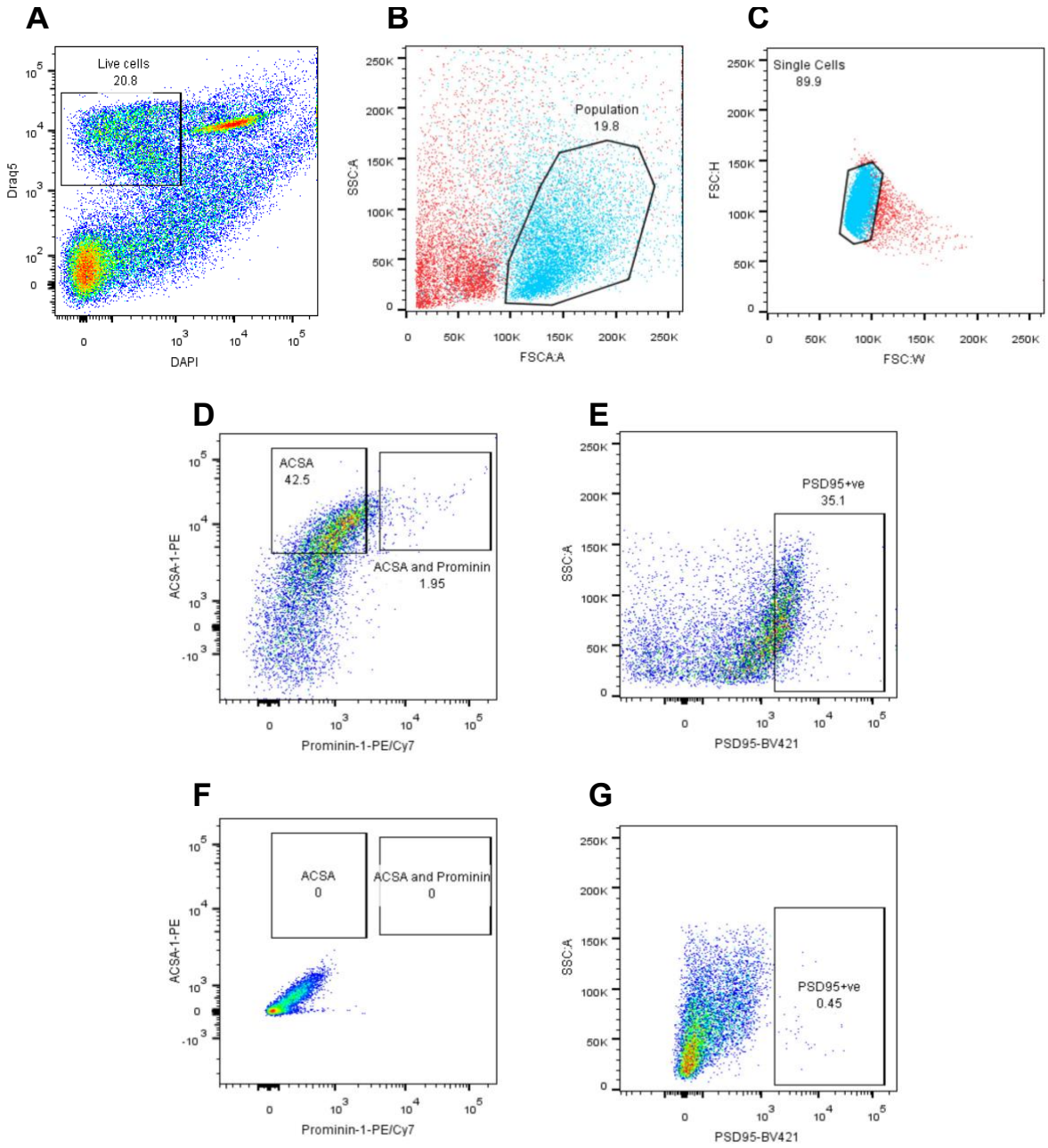


Figure 4.8. Isolation of different cell populations from mouse brain.
(A) Application of DRAQ5^{TM} and DAPI double staining in order to select for live cells (DNA-containing cells with intact cell membrane). **(B)** Overlay of the live cell population onto the forward over side scatter plot area (FSC-A/SSC-A). **(C)** Doublet exclusion by overlaying forward scatter height over width (FSC-H/FSC-W) to side scatter height over width (SSC-H/SSC-W). **(D)** Discrimination of target cell populations from the adult mouse brain, including cells positive for ACSA-1 (astrocytes) on the basis of Phycoerythrin (PE) fluorescence and double positive for ACSA-1 and Prominin-1 (neural stem cells) on the basis of PE/Cy7 fluoresence. **(E)** Discrimination of cells positive for PSD95 (neurons) in the neonate mouse brain, on the basis of BV421 fluoresence. **(F)** Unstained adult brain sample serving as negative control for PE and PE/Cy5 fluorescence. **(G)** Unstained neonate brain sample serving as negative control for BV421 fluorescence.

4.4. Adult Mouse Tissues Exhibit Striking Heteroplasmy Variation at the Single Cell Level

Having optimized protocols enabling us to isolate select cell populations from mouse spleen and brain samples as well as a technique allowing for the quantification of mtDNA heteroplasmy measurement at the single cell level (see **Results Section, Chapter 3**), we proceeded to analyse single cell samples derived from adult (P100) mice. Six mice were selected, half of which carried the heteroplasmic m.5019A>G mutation while the other half carried the m.5024C>T one, both present on *mt-Ta*. After filtering out samples that showed signs of contamination, single cell mtDNA heteroplasmy data was obtained for 122 CD19+ve cells, 118 CD19-ve cells, 158 ACSA+ve cells and 125 PSD95+ve cells **(Figure 4.9.)**. At the same time, bulk heteroplasmy measurements performed on 8 tissues (skin, liver, muscle, gonads, heart, gut, brain and spleen) derived from the mouse on the day it was culled plus the ear biopsy collected on weaning (~3 weeks), painted a picture identical to the one seen by Kauppila et al. Even though heteroplasmy levels in the brain were always greater than the spleen, heteroplasmy measurements across the 8 tissues were very close to the ear biopsy measurements corresponding to each animal **(Figure 4.10.)**. Consequently, there was no indication of selective forces either increasing or decreasing the heteroplasmy of m.5019A>G and m.5024C>T mutations, in any of the examined tissues at P100.

Taken together, measurements on bulk-tissue samples were in line with previously published data. Quantification of mtDNA heteroplasmy at the single cell level on the other hand, revealed a much more complicated picture than one hinted at by bulk estimates. Both spleen- and brain-derived populations coming from m.5019A>G and m.5024C>T mice exhibited striking heteteroplasmic variability, while the mean single cell heteroplasmy was largely in line with the bulk-tissue average **(Figure 4.9)**. It should be noted that the high bulk heteroplasmy that was present in tissues deriving from the three m.5019A>G mice, meant that the single cell variance was limited due to the tissue's proximity to the state of mutant homoplasmy. In more detail, from the 571 single cells analysed across 3 mice, 50% were above 90% heteroplasmy, while 15% were trapped at the 100% absorption point. In the case of mutant homoplasmic cells, we would expect each cell division to produce daughter cells that are also homoplasmic for the m.5019A>G mutation, as there are no copies of the WT haplotype still present in the cell. Towards the lower end of the distribution however, we were able to detect cells that contained

mtDNA heteroplasmy values as low was 18%. In the case of the 3 mice carrying the m.5024C>T mutation, the bulk heteroplasmy of each animal was lower (50-65%), consequently, the spread observed in single cell heteroplasmy was more substantial, ranging in certain cases from 0-96%. As a result of their middling mtDNA bulk tissue heteroplasmy levels, and in further contrast to their m.5019A>G counterparts, mice carrying the m.5024C>T mutation had very few cells being trapped at either absorption point. Overall, these initial observations were enough to spark our interest in the driving forces controlling mtDNA heteroplasmy dynamics, and we sought to expand our study by looking into single cell samples derived from both mutant mouse strains at more timepoints through development.

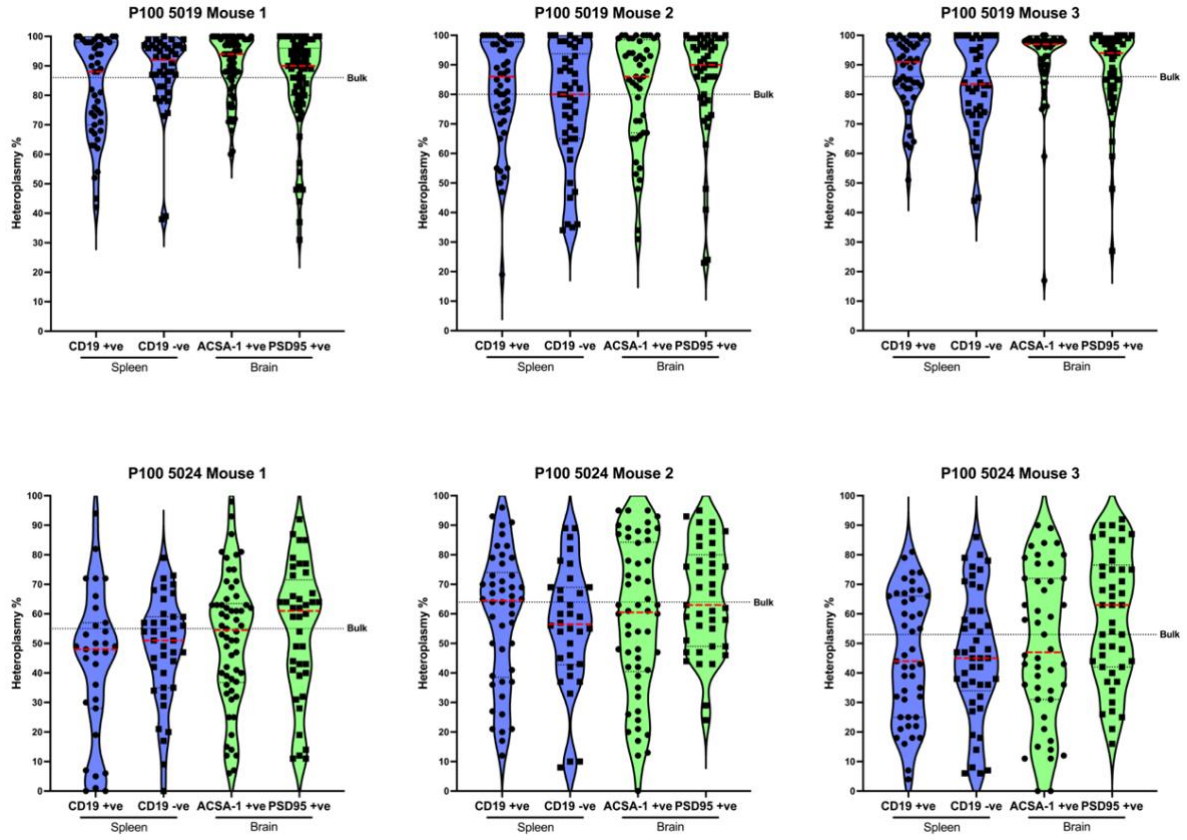


Figure 4.9. Single cell heteroplasmy of spleen- and brain-derived populations at P100.
Single cell pyrosequencing measurements of mitochondrial DNA (mtDNA) heteroplasmy performed on cell populations derived from spleen and brain samples belonging three adult (P100) mice carrying either the m.5019A>G or the m.5024C>T mt-Ta mutation. Dotted lines represent heteroplasmy estimates calculated by averaging out the 8 bulk tissue measurements previously performed on various tissues belonging to the same animal **(Figure 4.10.)**. Violin plots display median and interquartile range (IQR). Highlighted in blue are spleen-derived populations while in green are brain-derived ones.

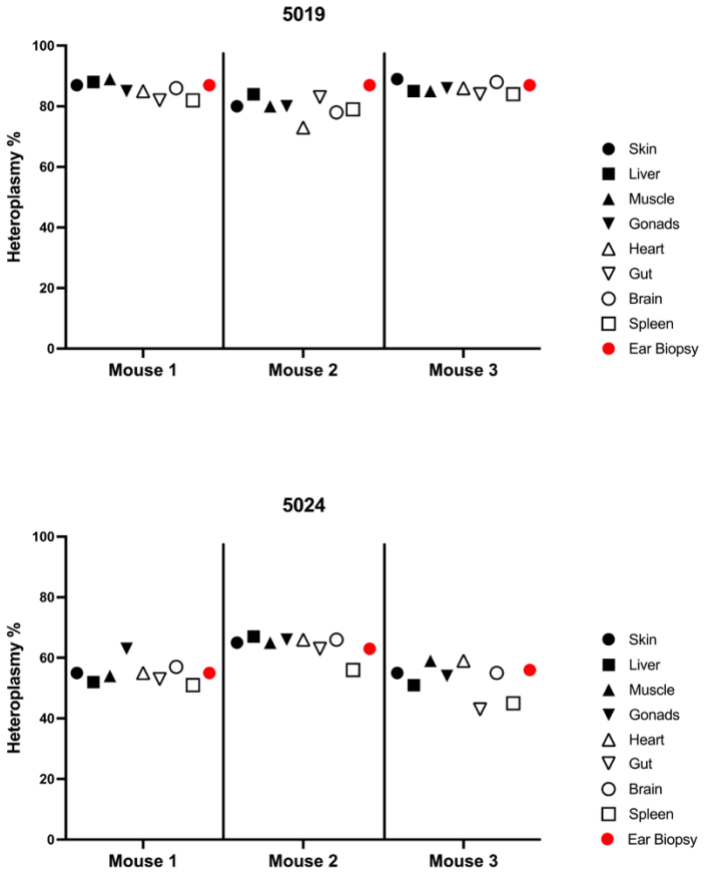


Figure 4.10. Bulk tissue heteroplasmy measurements at P100.
Pyrosequencing measurements of mitochondrial DNA (mtDNA) heteroplasmy, carried out on 8 tissues (skin, liver, muscle, gonads, heart, gut, brain and spleen) derived from 6 adult (P100) mice, half carrying the m.5019A>G while the other half carrying the m.5024C>T mutation. The heteroplasmy reading of the ear biopsy taken during weaning (~3 weeks of age) is highlighted in red.

4.5. mtDNA Heteroplasmy Variability Increases with Age in Both Quiescent and Dividing Tissues

Having observed a striking range in single cell mtDNA heteroplasmy measurements in adult mice of both the m.5019A>G and m.5024C>T strain, we decided to dive further into the dynamics controlling the segregation of mtDNA heteroplasmy, by examining both younger and older mice carrying the m.5024C>T mutation. Single cell heteroplasmy measurements were performed on single-cell populations derived from both the brain and spleen of 3 new-born (P0) pups (n = 474 cells across all populations) as well as 3 six-day old (P6) pups (n = 694 cells across all populations). Finally, 3 animals were collected at one year of age (P365, n=540 cells across all populations).

In a similar fashion to the P100 mice, bulk tissue samples were collected from the culled animals and DNA was extracted in order to provide us with a picture of the mtDNA heteroplasmy percentage governing the mouse as a whole. Predictably, and in line with the P100 data **(Figure 4.10.)**, bulk heteroplasmy values did not appear to differ in any striking or consistent way across the different tissues tested or, in fact with the ear biopsy in the case of P365 mice **(Figure 4.11.)**. It should be mentioned however, that tissues derived from P0 and P6 mice seemed closer to each other compared to ones at P365 but, in order for this observation to be presented as conclusions on the age-dependent nature of inter-tissue heteroplasmy variance, higher numbers of animals will have to be studied. Overall, we can conclude that bulk tissue mtDNA heteroplasmy remains largely stable, with the only selective forces present impacting rapidly dividing tissues of aged mice belonging to the m.5024C>T strain.

Looking at single cell measurements, detailed breakdown of which can be found in detail in **Figures S1-3**, we can see that there is a wide range of heteroplasmies recorded for both organs tested, in every timepoint **(Figure 4.12.B)**. However, when taking a closer look at the data collected, there were only 2 WT homoplasmic cells present in the P0 dataset, which was comprised of 474 cells, with very few having extreme heteroplasmy values. Conversely, in the case of P365 mice, there were a lot more cells that possessed extreme mtDNA heteroplasmy values. In more detail, there were 49 WT homoplasmic cells deriving from a dataset of 540 single cells. The increase in the number of extreme heteroplasmic and homoplasmic cells with age, while definitely influenced by the mouse's bulk heteroplasmy, hint towards a constant

postnatal segregation of mtDNA heteroplasmy throughout life (Burgstaller et al., 2014). In addition to single cell heterplasmy measurements in postnatal mice, embryos were collected at the stage succeeding organogenesis, E8.5. Three embryos were isolated, deriving from the m.5024C>T heteropasmic mouse strain, and single cells were isolated without being subject to any further discrimination (n = 174 cells across all embryos). The bulk heteroplasmy of each embryo was estimated by averaging out the heteroplasmy measurements deriving from 8 wells, each containing 14 cells, again there was no downstream labelling of the cells sorted in this experiment **(Figure 4.12.A)**.

As previously discussed, all heteroplasmy measurements are limited by the 0% and 100% absorption points which inevitably skew the distribution as more and more cells get trapped at either point. The number of cells that will reach these points is dependent amongst other factors, on the bulk heteroplasmy of the animal (Samuels et al., 2010; Wonnapinij et al., 2018). In more detail, we would expect, in the absence of selective pressure, to observe distributions with extremely high or low mean heteroplasmy to have distributions with smaller variance compared to distributions that have a mean heteroplasmy close to 50%. In order to account for these discrepancies in bulk tissue heteroplasmy that inevitably arise when working with animal models, we decided to calculate the heteroplasmy log-odds and normalize each one of those values to the mean bulk heteroplasmy of the animal. In our case, the mean heteroplasmy of each animal is derived by averaging out the bulk measurements obtained from the collected tissues. Transforming heteroplasmy values in this way, allowed us to compare single cell mtDNA measurements derived from mouse embryos in the period following organogenesis (E8.5) **(Figure 4.13.)**. This comparison revealed a consistent increase in single cell heteroplasmy variance beginning at prenatal life and extending into early postnatal stages of development and eventually reaching adulthood and later life. Having transformed all heteroplasmy values, such comparisons can be made independent of the mean heteroplasmy levels that are present in each mouse.

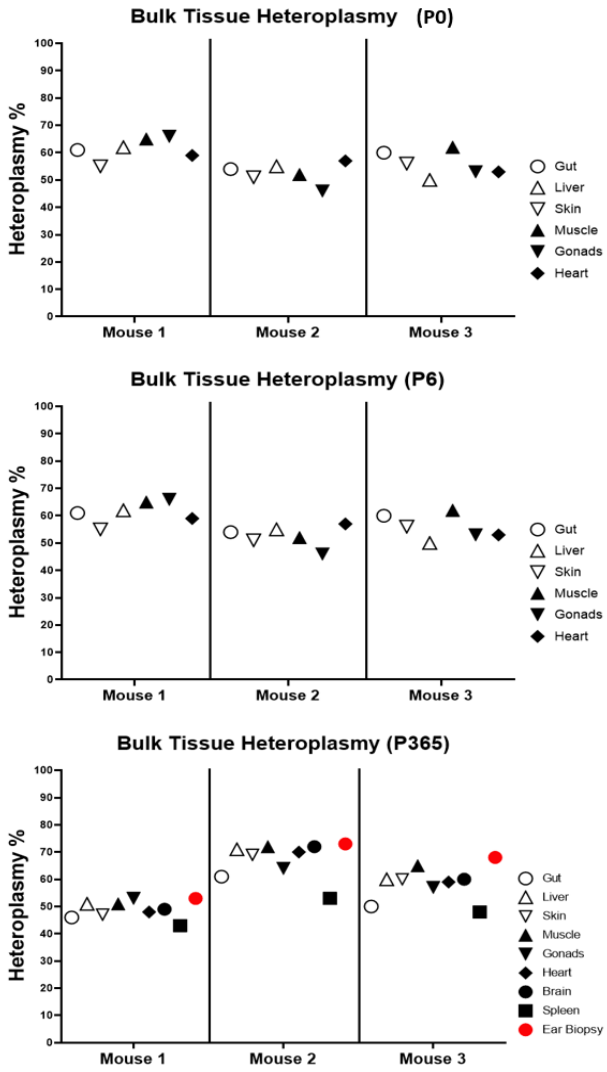


Figure 4.11. Bulk tissue heteroplasmy measurements at P0, P6 and P365 in m.5024C>T mice. Pyrosequencing measurements of mitochondrial DNA (mtDNA) heteroplasmy, carried out on 8 tissues (skin, liver, muscle, gonads, heart, gut, brain and spleen) derived from three neonate (P0), three 1-week -old (P6) and three aged (P365) mice. Where available, the heteroplasmy reading of the ear biopsy taken during weaning (~3 weeks of age) is highlighted in red.

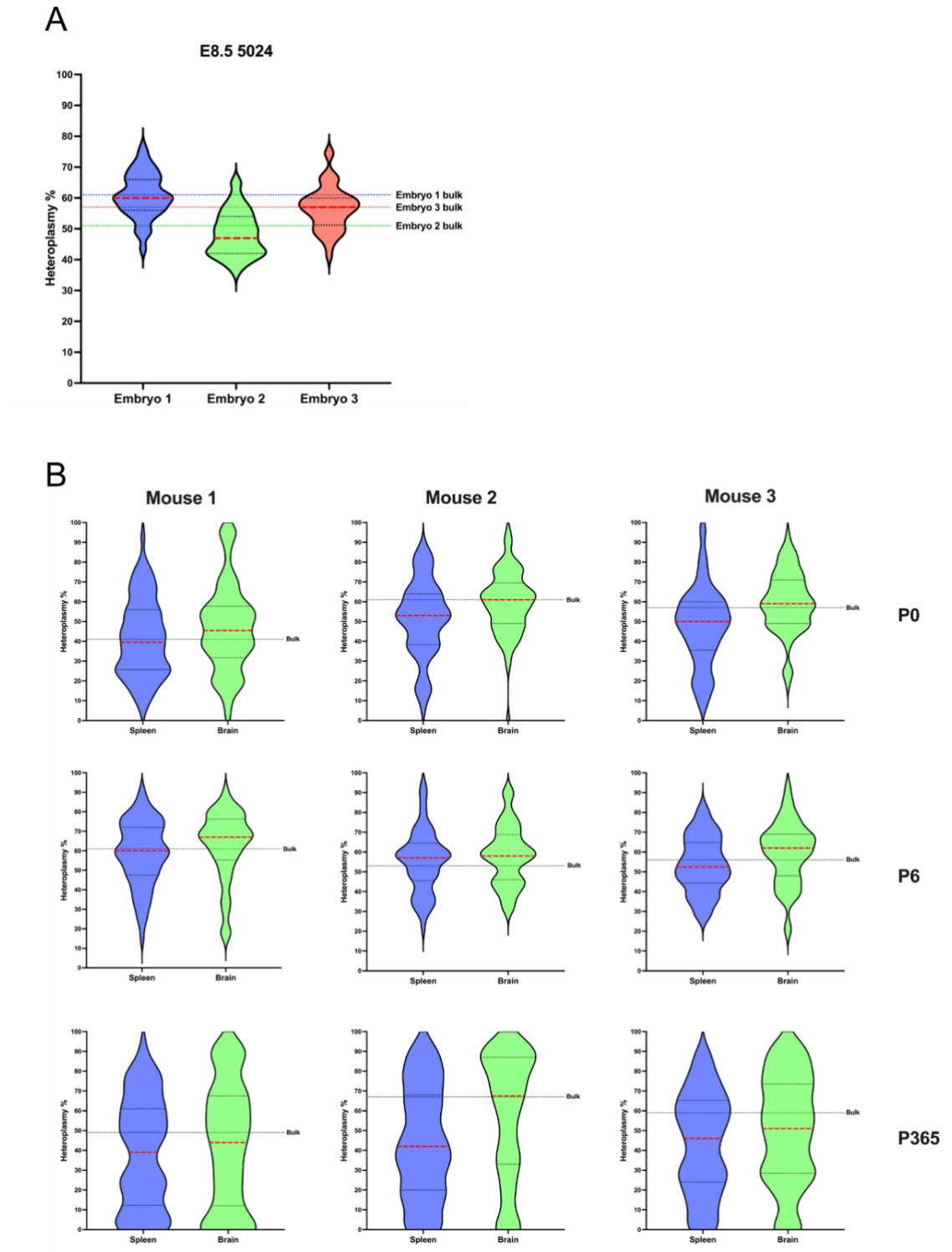


Figure 4.12. Single cell heteroplasmy measurements at E8.5, P0, P6 and P365 in m.5024C>T mice. **(A)** Pyrosequencing measurements of mtDNA heteroplasmy, carried out on single cells isolated from E8.5 mouse embryos. Dotted lines represent bulk mtDNA heteroplasmy values for each embryo. **(B)** Single cell pyrosequencing measurements of mtDNA heteroplasmy, performed on cell populations derived from spleen and brain samples belonging 3 neonate (P6), 3 six-day-old (P6) and three older (P365) mice. Violin plots display median and interquartile range (IQR). Dotted lines represent heteroplasmy estimates calculated by averaging out either the 6 or 8 bulk tissue measurements previously performed on various tissues belonging to the same animal **(Figure 5.10.)**.

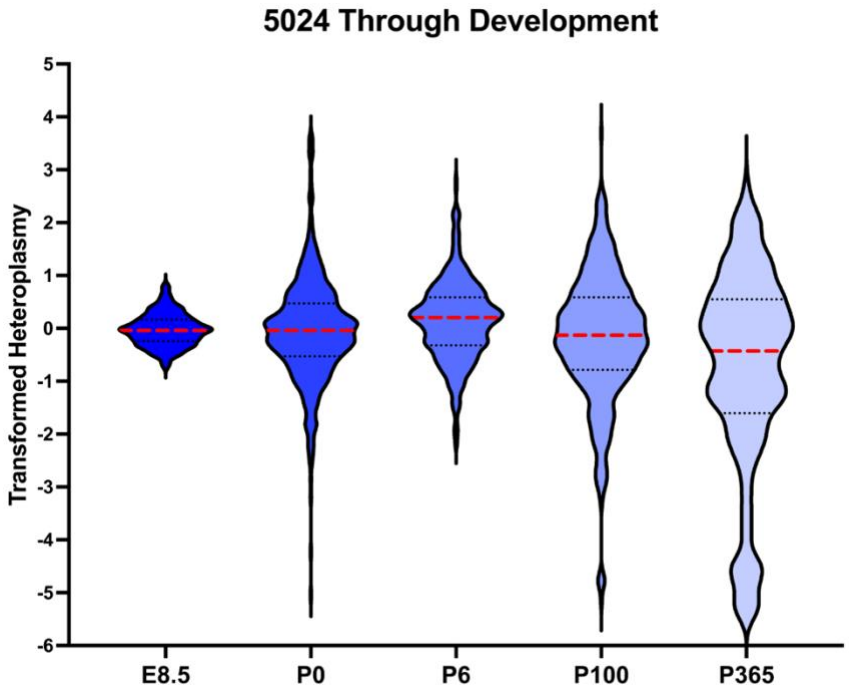


Figure 4.13. Increase in single cell variance with age in m.5024C>T mice.
Transformed single cell mtDNA heteroplasmy values deriving from mice at different stages though life (E8.5, P0, P6, P100 and P365). All mice included carry the heteroplasmic m.5024C>T mt-Ta mutation. Violin plots display median and interquartile range (IQR).

4.6. Accumulation of Mutant Homoplasmic Cells in Both Dividing and Post-Mitotic Tissues in a Mutation-Dependent Manner

Having provided *in vivo* evidence in support of theories that have predicted the increase of single cell heteroplasmy through life in m.5024C>T mice, we decided to extend our study even further and examine whether a similar observation could be made in the m.5019A>G model. So far, we have seen that the two strains have similarly consistent bulk tissue DNA measurements at P100 **(Figure 4.10.)**, with the 851 cells tested for m.5019A>G exhibiting a wide heteroplasmy range for both tissues of interest (Brain IQR: 81.25-98%, Spleen IQR: 74-97%) **(Figure 4.9.)**. However, at P100 there was one striking distinction between the two strains, as populations carrying the m.5019A>G haplotype exhibited a lot of cells that were mutant homoplasmic. Initially, we suspected that this was due to some selective force that was leading cells to favour the accumulation of the pathogenic mutation. However, this hypothesis was quickly debunked in light of further data. First, we had to take into account the fact that all 3 P100 m.5019A>G mice had very high mean heteroplasmy values (80%, 86% and 86%). This meant that in order for the single cell average to match that of the bulk tissue, the majority of the 545 sorted single cells were above the bulk tissue average, with 14% of them being mutant homoplasmic. At the same time, the fact that the single cell mean matched the mean heteroplasmy of the mouse, in addition to the ear biopsy, was another indicator that no selective forces were at work. It is worth noting that the type of selection discussed here is related to but separate from germline selection which in itself is a well-documented phenomenon in the m.5024C>T mouse model (Zhang et al., 2021).

Despite this initial evidence against our hypothesis, we decided to study this further. We collected the same cell populations from 3 m.5019A>G mice that were older than P100 (P239, P318 and P318) and possessed low mtDNA heteroplasmy levels. According to our hypothesis, in the presence of selective forces, the single cell mean would be greater than the bulk average, and we would still observe a significant number of cells accumulating at extremely high heteroplasmy levels. However, this is not the case. Once the data was collected, we observed that the middling bulk heteroplasmy of the mice was in line with the single cell average **(Figure 4.14.)**. While there was a fair number of cells that had accumulated higher levels of the

mutation, the same could be said about lower heteroplasmy levels, especially in mice 2 and 3 where the bulk heteroplasmy was lower than 50%.

Having established that it is random genetic drift that drives the accumulation of cells at extreme heteroplasmy values, including homoplasmy in m.5019A>G mice, we decided to extend our observations to include time points throughout life in the same way we did with the m.5024C>T model. First, bulk tissue heteroplasmy measurements were made, revealing an expected pattern with all tissues being relatively close to each other, and a slight increase in inter-tissue heteroplasmy variability with age (**Figure 4.15.)**. Single cell pyrosequencing measurements of mtDNA heteroplasmy were then performed on a total of 1739 cells, derived from three E8.5 (n = 141 cells across all embryos), three P0 (n = 491 cells across all populations), three P6 (n = 492 cells across all populations) and three P365 mice (n = 535 cells across all populations) **(Figure 4.16.)**. Once again, transformation of single cell datasets revealed an age-dependent increase in mtDNA heteroplasmy variance, revealing that the phenomenon is present in both dividing and non-dividing tissues and not limited to m.5024C>T **(Figure 4.17.)**.

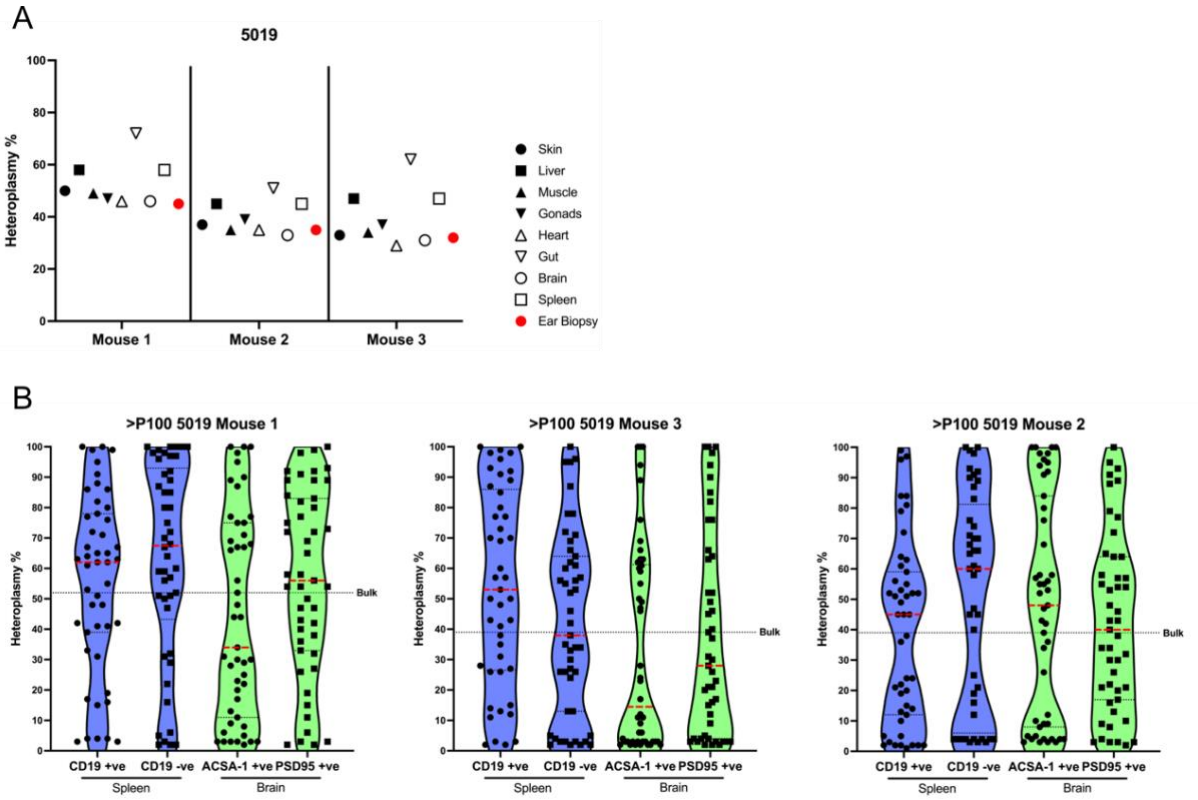


Figure 4.14. Adult mice carrying low levels of the heteroplasmic m.5019A>G mutation.
(A) Pyrosequencing measurements of mitochondrial DNA (mtDNA) heteroplasmy, carried out on eight tissues (skin, liver, muscle, gonads, heart, gut, brain and spleen) derived from three adult (>P100) mice. Heteroplasmy reading of the ear biopsy taken during weaning (~3 weeks of age) is highlighted in red. **(B)** Single cell pyrosequencing measurements of mtDNA heteroplasmy, performed on cell populations derived from spleen and brain samples belonging to three adult (>P100) mice. Dotted lines represent heteroplasmy estimates calculated by averaging out either the 8 bulk tissue measurements previously performed on various tissues belonging to the same animal **(A)**. Violin plots display median and interquartile range (IQR); Mouse 1= P239, Mouse 2 and 3= P318.

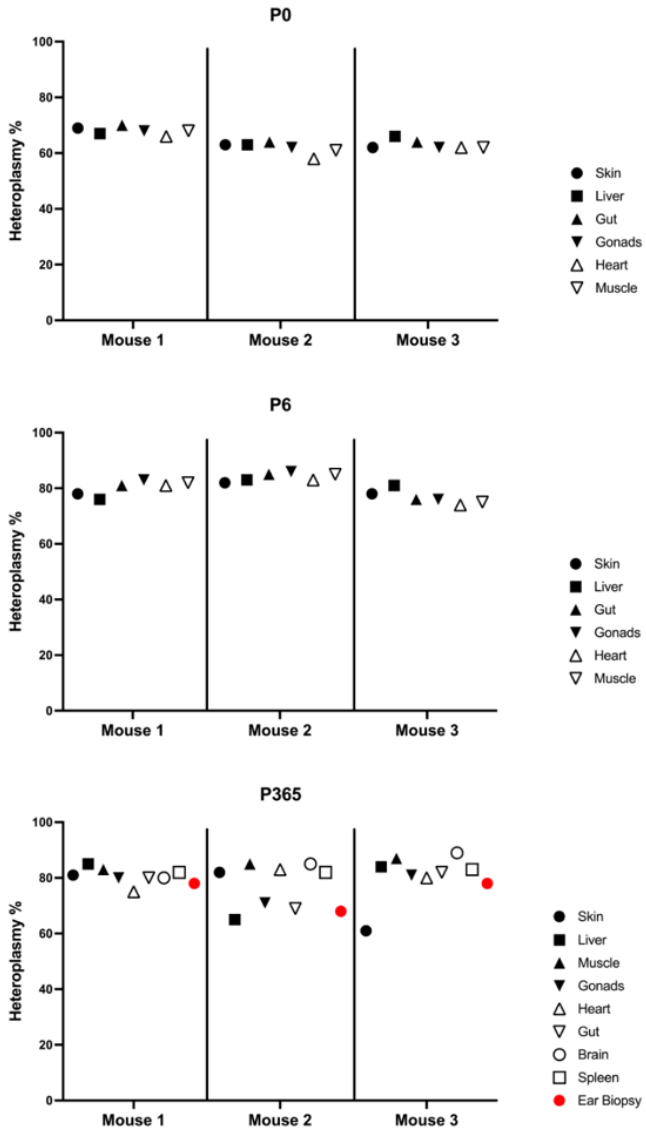


Figure 4.15. Bulk tissue heteroplasmy measurements at P0, P6 and P365 in m.5019C>T mice. Pyrosequencing measurements of mitochondrial DNA (mtDNA) heteroplasmy, carried out on 8 tissues (skin, liver, muscle, gonads, heart, gut, brain and spleen) derived from three neonate (P0), three 1-week -old (P6) and three aged (P365) mice. Where available, the heteroplasmy reading of the ear biopsy taken during weaning (~3 weeks of age) is highlighted in red.

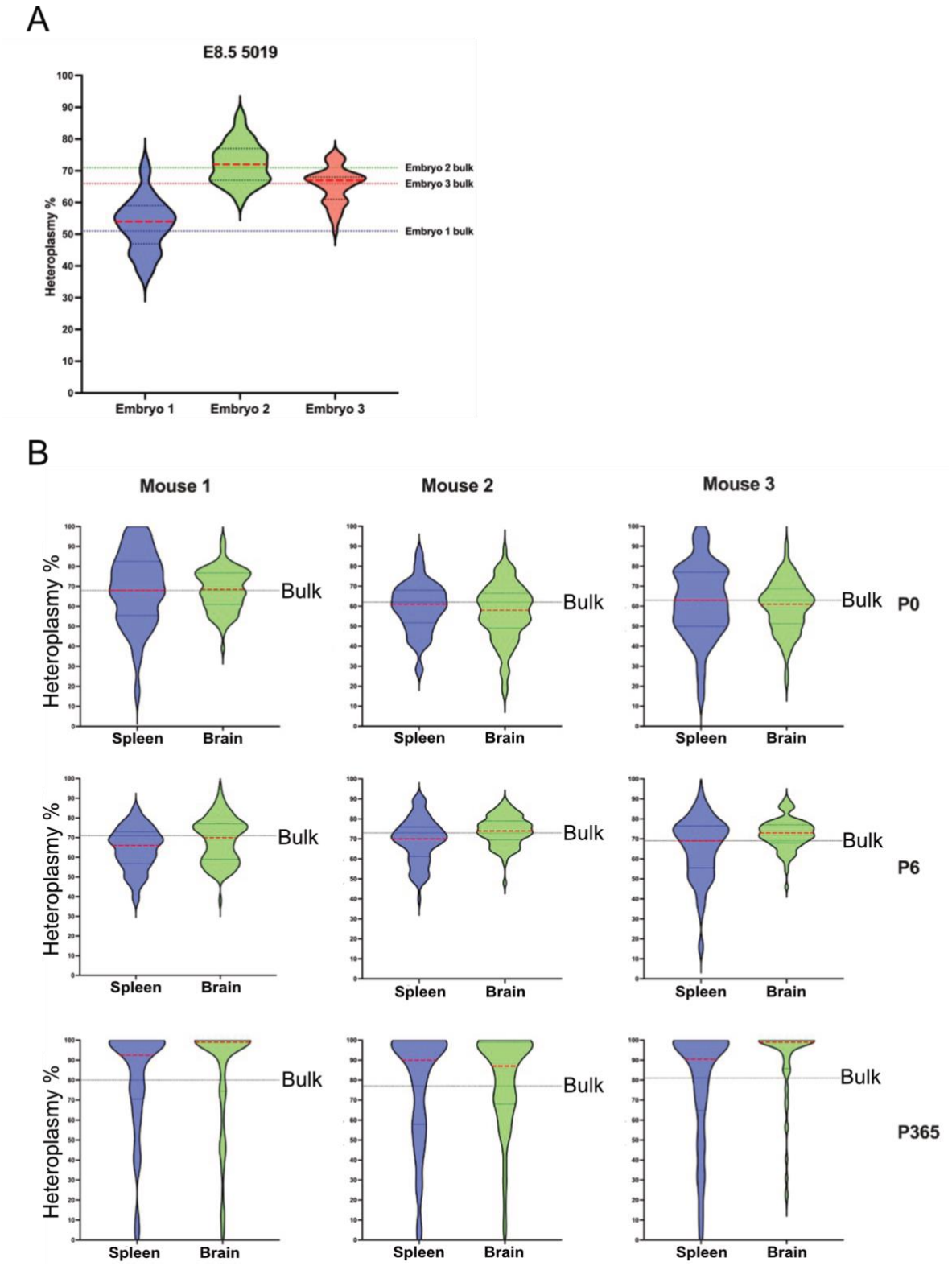


Figure 4.16. Single cell heteroplasmy measurements at E8.5, P0, P6 and P365 in m.5019A>G mice. **(A)** Pyrosequencing measurements of mtDNA heteroplasmy, carried out on single cells isolated from E8.5 mouse embryos. Dotted lines represent bulk mtDNA heteroplasmy values for each embryo. **(B)** Single cell pyrosequencing measurements of mtDNA heteroplasmy, performed on cell populations derived from spleen and brain samples belonging three neonate (P0), three six-day-old (P6) and three aged (P365) mice. Violin plots display median and interquartile range (IQR). Dotted lines represent heteroplasmy estimates calculated by averaging out either the six or eight bulk tissue measurements previously performed on various tissues belonging to the same animal **(Figure 4.15.)**.

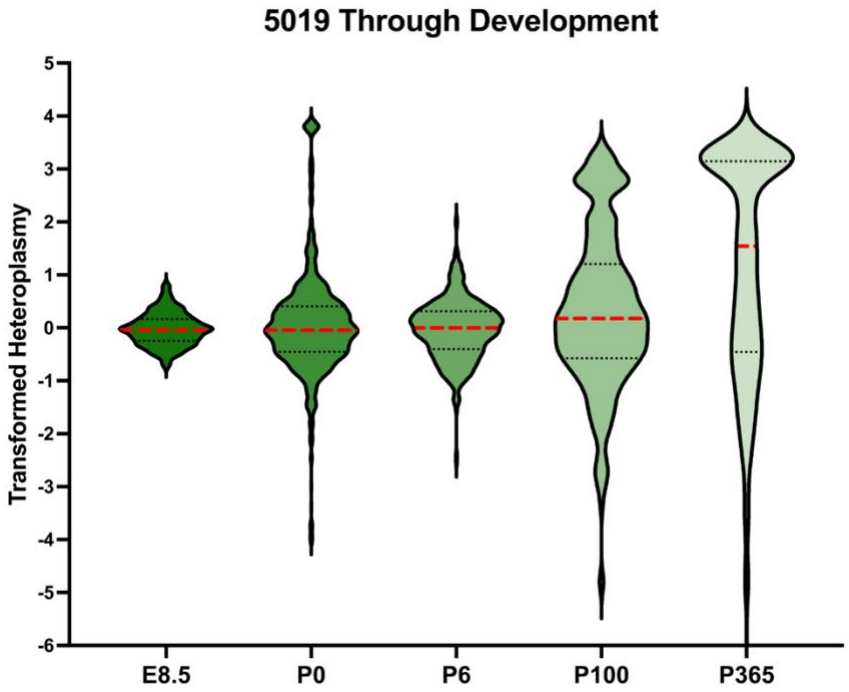


Figure 4.17. Increase in single cell variance with age in m.5019A>G mice.
Transformed single cell mtDNA heteroplasmy values deriving from mice at different stages though life (E8.5, P0, P6, P100 and P365). All mice included carry the heteroplasmic m.5019A>G mt-Ta mutation. Violin plots display median and interquartile range (IQR).

4.7. Random Genetic Drift Drives Tissues Towards Homoplasmy in an Age-Dependent Manner.

In order to better understand the driving forces behind the mechanism previously observed, responsible for mtDNA heteroplasmy segregation, we compared the mean single cell heteroplasmy of each cell population and tissue, to the bulk heteroplasmy value corresponding to each mouse collected from both the m.5019A>G and m.5024C>T heteroplasmic strains **(Figure 4.18.A)**. The resulting strong correlation (Pearson's ρ=0.894) between these values, further supported our hypothesis that it is primarily random genetic drift that accounts for the observed progressive increase in single cell mtDNA heteroplasmy variability. At this point, it is worth noting that some of the spleen derived populations derived from P365 carrying the m.5024C>T mutation deviate from the bulk. This observation is not without precedent and can be explained by looking at the work of Kauppila et al., and the selection that takes place in the blood of mice carrying the m.5024C>T at heteroplasmy at levels greater than 60% **(Figure 4.2.)** (Kauppila et al., 2016).

In keeping with this analysis and our previous work, single cell heteroplasmy values from both the m.5019A>G and the m.5024C>T mouse strains fitted very closely to the Kimura distribution **(Figure 4.18.B)**, which provides a way of testing whether the heteroplasmy values at hand are the product of random genetic drift (Wonnapinij et al., 2008). In more detail, the Kimura distribution can be understood on the basis of two parameters: p0, or starting heteroplasmy, is an assumption that all cells are given rise to from a single progenitor without the application of any selective pressures and *b*, which relates to the level of genetic drift that a certain population has undergone. By estimating $\hat{b}$, we can directly compare the degree of drift that two populations have been subject to. In this analysis, $\hat{b} = 1$, corresponds to the absence of drift (i.e. heteroplasmy remains constant within a population) while the lower the $\hat{b}$ value is, the higher the greater the level of drift (always abounded by $\hat{b} = 0$). Using these parameters, we were able to conclude that the m.5019A>G mutation is subject to a more rapid drift (P365 b = 0.665 ± 0.07, mean ± SD) compared to m.5024C>T (P365 b = 0.573 ± 0.086, mean ± SD; Wilcoxon's rank test p-value = 0.014) **(Figure 4.18.C)**.

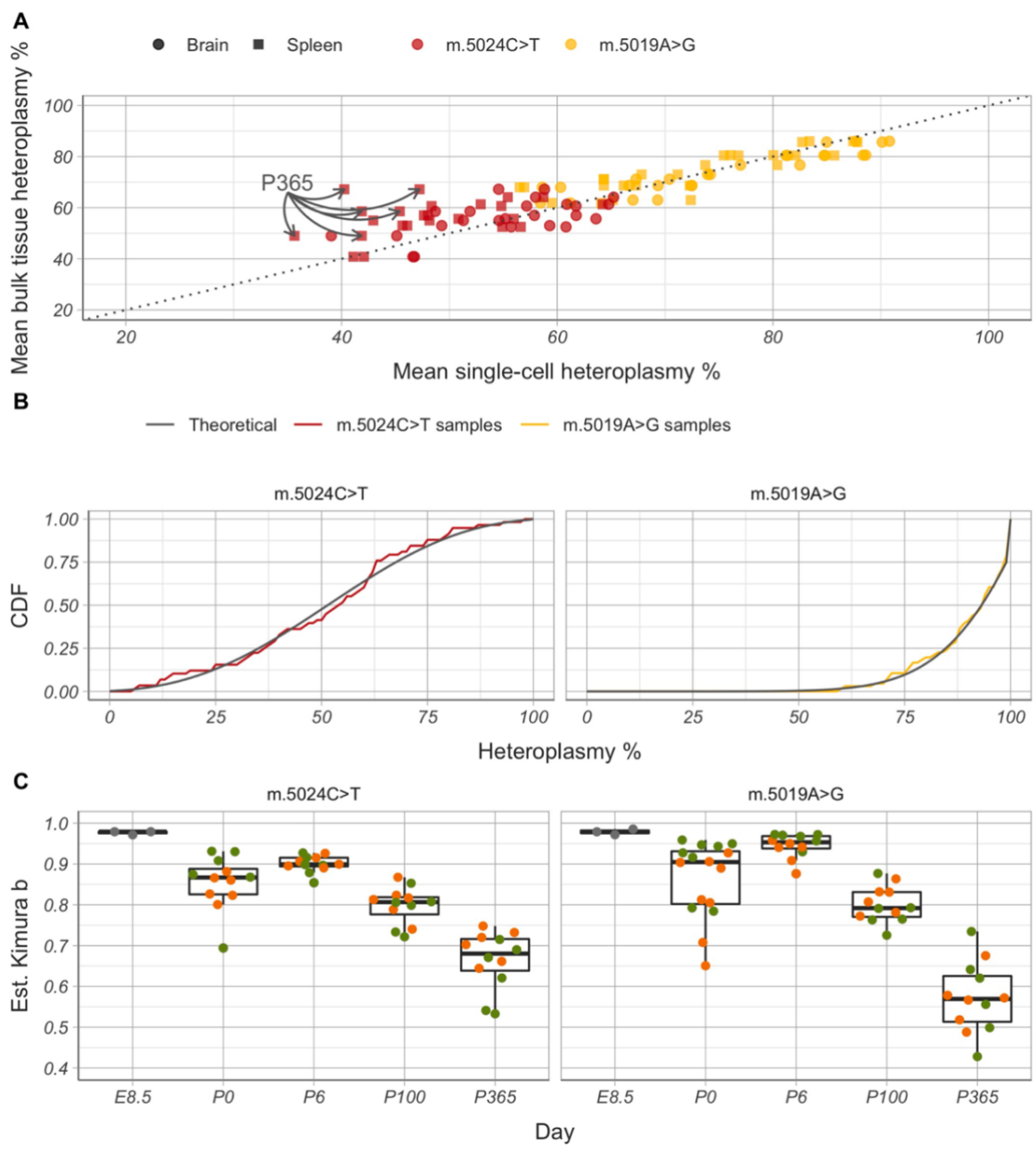


Figure 4.18. Random genetic drift leads to age-dependent accumulation of cells with extreme heteroplasmy values.
(A) Correlation between mean mitochondrial DNA (mtDNA) heteroplasmy of single cell populations and the average of bulk tissue heteroplasmy measurements. Arrows point towards datapoints generated from the spleen of P365 mice carrying the m.5024C>T mutation. **(B)** Representative data fitting the ACSA-1 +ve cells derived from Mouse 1 of the m.5019A>G and m.5024C>T datasets to the Kimura distribution. **(C)** The rate of random generic drift estimated by the Kimura *b* parameter at different stages throughout the lifespan of mice belonging to either the m.5019A>G or m.5024C>T heteroplasmic strains. Brain-derived populations are labelled in green, spleen-derived populations are labelled in orange. Figure generated by Dr Lyuba Bozhilova.

4.8. Cell Populations Deriving from the Same Animal do not Exhibit Significant Differences in Heteroplasmy Distribution

Having established a much more complicated picture when it comes to the distribution of single cell heteroplasmy, we decided to examine whether spleen and brain-derived cell population were different from one another when it comes to their heteroplasmy distribution. Comparisons were drawn at the level of the organism in order to avoid clustering transformed data together. This approach prevented skewing of the statistical analysis as a result of differences in the distributions that would inevitably arise due to the presence of mice with different single cell distributions that are a result of differences in bulk heteroplasmies. Consequently, looking at one mouse at a time, carrying either the m.5019A>G or the m.5024C>T mutation, across all time-points first led us to draw comparisons across the two populations derived from brain and spleen **(Table 4.1.** and **4.2.)**. Then, we went ahead and grouped the two populations derived from each organ, allowing us to create a more complete picture of each tissue while drawing a direct comparison between spleen and brain for each mouse **(Table 4.2.)**.

Drawing comparisons between populations deriving from the same organ using the KS test yielded no significant results, for either the brain or the spleen in mice derived from either the m.5019A>G or the m.5024C>T line. Similar results were obtained when comparing across the two tissues, while the p-values were generally lower, the application of the Bonferroni multiple test correction meant that no significant differences were detected.

Table 4.1. p-values of Kolmogorov–Smirnov tests performed on brain-derived populations. Comparisons made between brain-derived populations (astrocytes vs neural progenitors or neurons) derived from each of the three mice collected at every timepoint (P0, P6, P100 and P365), for both m.5019A>G and m.5024C>T mutant mouse models.

Strain	Age	Mouse Number	Heteroplasmy %	p-value	Significance
m.5019A>G	P0	1	68.7	0.671130	-
		2	61.8	0.761707	-
		3	63.0	0.304901	-
	P6	1	85.7	0.248904	-
		2	80.4	0.450967	-
		3	86.0	0.225550	-
	P100	1	80.4	0.826177	-
		2	76.7	0.185721	-
		3	80.6	0.805584	-
	P365	1	71.2	0.090619	-
		2	73.0	0.451822	-
		3	68.7	0.760827	-
m.5024C>T	P0	1	40.8	0.215313	-
		2	60.7	0.267464	-
		3	57.0	0.053740	-
	P6	1	55.0	0.308532	-
		2	64.1	0.095066	-
		3	53.0	0.308795	-
	P100	1	49.0	0.323294	-
		2	67.2	0.165946	-
		3	58.6	0.389143	-
	P365	1	61.3	0.644522	-
		2	52.5	0.119847	-
		3	55.7	0.068303	-

Table 4.2. p-values of Kolmogorov–Smirnov tests performed on spleen-derived populations. Comparisons made between spleen-derived populations (B vs non-B cells) derived from each of the three mice collected at every timepoint (P0, P6, P100 and P365) , for both m.5019A>G and m.5024C>T mutant mouse models.

Strain	Age	Mouse Number	Heteroplasmy %	p-value	Significance
m.5019A>G	P0	1	68.7	0.487244	-
		2	61.8	0.945570	-
		3	63.0	0.059390	-
	P6	1	85.7	0.027766	-
		2	80.4	0.531362	-
		3	86.0	0.236661	-
	P100	1	80.4	0.107252	-
		2	76.7	0.100251	-
		3	80.6	0.348106	-
	P365	1	71.2	0.995591	-
		2	73.0	0.417152	-
		3	68.7	0.116104	-
m.5024C>T	P0	1	40.8	0.740659	-
		2	60.7	0.147653	-
		3	57.0	0.966972	-
	P6	1	55.0	0.485469	-
		2	64.1	0.367366	-
		3	53.0	0.793116	-
	P100	1	49.0	0.633775	-
		2	67.2	0.289089	-
		3	58.6	0.315733	-
	P365	1	61.3	0.007247	-
		2	52.5	0.945193	-
		3	55.7	0.173220	-

Table 4.3. p-values of Kolmogorov–Smirnov tests performed on organ-derived populations. Comparisons made between grouped spleen-derived populations (B and non-B cells) and grouped brain-derived populations (astrocytes vs neural progenitors or neurons) derived from each of the three mice collected at every timepoint (P0, P6, P100 and P365), for both m.5019A>G and m.5024C>T mutant mouse models.

Strain	Age	Mouse Number	Heteroplasmy %	p-value	Significance
m.5019A>G	P0	1	68.7	0.380755	-
		2	61.8	0.277637	-
		3	63.0	0.048882	-
	P6	1	85.7	0.672904	-
		2	80.4	0.066696	-
		3	86.0	0.018438	-
	P100	1	80.4	0.032287	-
		2	76.7	0.550909	-
		3	80.6	0.020932	-
	P365	1	71.2	0.012464	-
		2	73.0	0.009226	-
		3	68.7	0.000369	*
m.5024C>T	P0	1	40.8	0.149079	-
		2	60.7	0.008547	-
		3	57.0	0.000474	*
	P6	1	55.0	0.010156	-
		2	64.1	0.270461	-
		3	53.0	0.068490	-
	P100	1	49.0	0.509026	-
		2	67.2	0.002924	-
		3	58.6	0.189060	-
	P365	1	61.3	0.021643	-
		2	52.5	0.681574	-
		3	55.7	0.011553	-

4.9. Key Findings

- Dissociation of neonate and adult mouse spleen and brain preserved cell types that are representative of the function of their respective tissues. All populations of interest were isolated through FACS.
- Bulk tissue heteroplasmy measurements taken at different timepoints revealed a pattern consistent with published data, characterised by the lack of selective forces acting either in favour or against the *mt-Ta* m.5019A>G or m.5024C>T mutations.
- Single cell mtDNA heteroplasmy measurements carried out using pyrosequencing revealed a wide range of heteroplasmies harboured within spleen- and bran-derived cell populations. Starting at E8.5, going into early life (P0 and P6), into adulthood (P100) and old-age (P365), single cell mtDNA heteroplasmy measurements always approximated the mean tissue heteroplasmy values for both the m.5019A>G and m.5024C>T mutant mouse strains.
- In mice carrying either the m.5019A>G or m.5024C>T mt-Ta mutation, an age-dependent increase in the variance of single cell mtDNA heteroplasmy measurements manifested.
- In the case of mice carrying the m.5019A>G mt-Ta mutation, there was an age-dependent increase in cells that reached homoplasmy and were trapped at either the 0% or 100% absorption point. The number of cells reaching those points was age- and bulk tissue heteroplasmy-dependent.
- With the slight exception of spleen samples collected from P365 mice carrying the m.5024C>T mutation at levels equal to or greater than 60%, the mean mtDNA heteroplasmy of single cell datasets collected at different timepoints throughout the mouse lifespan correlated strongly to the corresponding bulk tissue measurements. This relationship was consistent across both mutant mouse strains.
- While a brief look at the single cell datasets collected from P100 and P365 mice carrying the m.5019A>G mutation might suggest the presence of a force driving positive selection, this hypothesis was debunked. Single cell data derived from both mutant mouse strains fitted quite consistently to the Kimura distribution, suggesting that the driving force behind the single cell heteroplasmy segregation that we observed is random genetic drift, taking place in both quiescent and proliferating tissues.
- The magnitude of random genetic drift, estimated using the Kimura *b* parameter, appears to be greater in m.5019A>G mice compared to their m.5024C>counterparts.

Chapter 5. Results:

Cell Lineage Specific Adaptability Compensates for Impaired Mitochondrial Activity During Organogenesis

5.1. Introduction

More than 120 distinct cell types that have been identified within the mouse body (Zhang et al., 2021) and 200 in the human one (Khan and Farhana, 2022). Across these cell types one can identify vastly divergent structures and pathways that enable them to carry out a range of unique roles within our bodies. This functional specialisation is accompanied by great discrepancies when it comes to energy requirements. Thus far, we have touched on the extensive co-dependency between mitochondria and the nucleus. The constant interaction between these two organs is termed nuclear-mitochondrial crosstalk and regulates many of the cell's most important responses, including but not limited to energy production (Dawson and Dawson, 2004). The magnitude and intricacy of the nuclear-mitochondrial crosstalk is such that disruptions are known to lead or contribute to major cellular dysfunctions (Fairbrother-Browne et al., 2021). A characteristic example is provided by neurons in patients suffering from Parkinson's disease (PD), where mutations in the *PINK1* and *PARK2* genes lead to dysfunctional proteins that in turn, disrupt mitochondrial biogenesis and clearance (Hardy et al., 2010; Shin et al., 2011). As a result, understanding the ramifications of nuclear-mitochondrial crosstalk disruptions can not only help us better understand the role of mitochondria as part of the cell structure but will provide valuable insight into many complex primary and secondary mitochondrial disorders.

For the longest time, carrying out transcriptomic studies on bulk tissue samples led to the unavoidable merging of different cell types into one homogeneous sample, erasing the nuances of cell-type-specific gene expression; however, this is no longer the case (Raj and Oudenaarden, 2008; Munsky et al., 2012; Stegle et al., 2015). Advances leading to the generation of high throughput single cell transcriptomics, have provided us with unprecedented insight into the intricate mechanisms governing cellular diversity within tissues of choice or even, entire organisms (Kulkarni et al., 2019). Single cell transcriptomics enable us to study the stochastic nature of gene expression and as a result, if used appropriately, shed light into the subtleties of nuclear-mitochondrial crosstalk across different cell types.

The first step taken when carrying out single cell transcriptomics is isolating said cells and performing the library preparation. The methods used to isolate single cells can vary both in terms of the number of cells (distinguishing between high and low-throughput approaches), as well as in the methods followed when it comes to cell selection (leading to biased and unbiased approaches) (Gross et al., 2015; Cristinelli and Ciuffi, 2018). The most commonly used technologies at the moment are droplet-based ones that are both unbiased and high-throughput (Prakadan et al., 2017). The three most popular single cell library preparation platforms at the moment are inDrop, DropSeq and 10X Genomics Chromium (Klein et al., 2015; Macosko et al., 2015; Zheng et al., 2017; Zhang et al. 2019). Each of these platforms has its own strengths and drawbacks and ultimately, the decision on which one to use depends on one's experimental design. While inDrop's customisable parameters might make it optimal for someone looking to detect gene expression levels in under expressed genes, 10x Genomics Chromium is the most sensitive of the three when it comes to detecting the highest transcript number (Zhang et al. 2019). Conversely, 10x Genomics Chromium's cost might make it inaccessible for one looking to run a high number of samples, and this is where the more cost-effective DropSeq comes into play, bearing of course the drawback of lower sensitivity (Zhang et al. 2019). However, all three of these platforms share some basic principles, as they make use of microfluidics to tag each individual cell with a unique barcode, contained within a single bead (Kulkarni et al., 2019). At the same time, each mRNA transcript is bound by an intricate tag that enables the identification of PCR duplicates by bioinformatic means, called a unique molecular identifier (UMI) (Kulkarni et al., 2019). Ultimately, the result is a matrix within which one can find the absolute number of every transcript that is present in every single cell that has been processed. As an alternative to the tag-based approach of generating single cell sequencing libraries, one can use a full-length approach (Hedlund and Deng, 2018; See et al., 2018; Ziegenhain et al., 2017). The full-length approach will increase sensitivity as the reads will be derived from across the length of the gene but a bias will be introduced in favour of longer genes as more reads will be derived from them as opposed to shorter genes (Ziegenhain et al., 2017). While the tag method does not introduce length bias, it does introduce read bias, as it is either the 3′ or 5′ ends of the transcript that get tagged and subsequently read (Ziegenhain et al., 2017). The prevalence of the tag-based methods comes down to their ability to enable direct quantification of transcripts, especially when dealing with large numbers of cells (Ziegenhain et al., 2017; Hedlund and Deng, 2018; See et al., 2018).

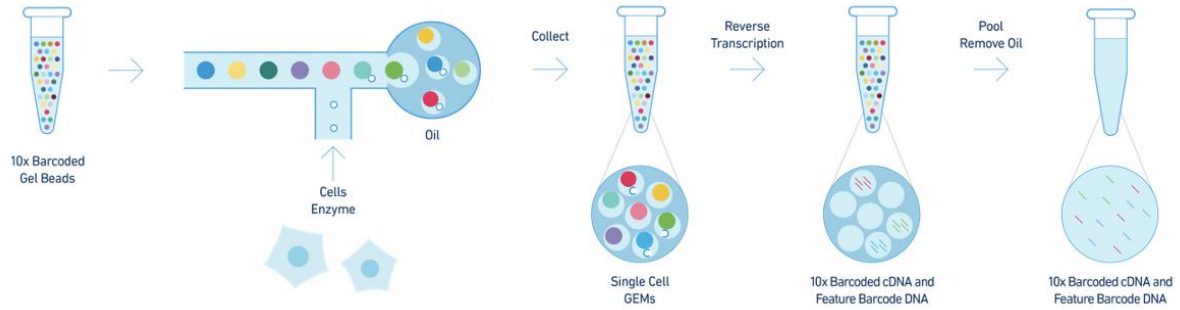


Figure 5.1. 10x Genomics Chromium pipeline.
Schematic representation of the steps taken when carrying out scRNA-seq (single cell RNA sequencing) of a sample through the 10x Genomics pipeline. The platform allows for high-throughput analysis of a range of single cell types as well as isolated nuclei. The single cells submitted by the researchers are handed to trained core staff. The first step is to capture each single cell within a uniquely barcoded gel bead droplet that contains the enzymatic mix necessary to carry out reverse transcription of the RNA. This results in the formation of GEMs (gel bead-in emulsions). Reverse transcription results in the generation of 10x barcoded cDNA and feature barcode DNA that can be traced back to each individual cell post-sequencing. Figure adapted from: Arbor, 2022.

When it comes to evaluating the performance of a high throughput single cell RNA-seq (scRNA-seq) experiment, there are three factors that need to be taken into consideration: sensitivity, precision and accuracy (Saliba et al., 2014; Ziegenhain et al., 2017; Wu et al., 2017). An experiment's sensitivity is largely dependent on sequencing depth as it is a way of evaluating the lowest transcript number that can be detected with consistency (Kulkarni et al., 2019). The importance of this experimental aspect should not be underestimated as it is a direct estimation of the amount of RNA present within a given cell at the start of the experiment. Precision of RNA-seq decreases as variation between technical measurements, also known as technical noise, within a given experiment increases (Kulkarni et al., 2019). Low precision is currently the biggest challenge facing scRNA-seq experiments due to the sparse nature of data being generated (Ziegenhain et al., 2017; Wu et al., 2017). Finally, accuracy is a term referring to the extent of overlap between the abundance of a transcript and the measured expression levels (Ziegenhain et al., 2017; Wu et al., 2017). Protocol-specific factors such as sequencing bias or exponential PCR amplification can have an effect on an experiment's accuracy (Wu et al., 2017).

Overall, understanding a technology such as scRNA-seq, currently reaching its peak potential, can unlock new pathways in the study of nuclear-mitochondrial crosstalk. Comprehending the advantages, drawbacks and risks involved in the experimental design can be valuable when trying to uncover the basis of cellular responses to mitochondrial disease, and might even help us shed light onto transcriptional adaptations that directly affect mtDNA heteroplasmy segregation.

5.2. Project Aims

Thus far, we have discussed at great length the role of mitochondria within the cell and the striking heterogeneity that defines mitochondrial content and mtDNA copy number across different cell types and tissues. However, it is still unknown when it arises and what are the driving forces behind it. Experiments covered within this chapter showcase how essential nuclear mitochondrial genes exhibit expression profiles that are cell lineage-specific while at the same time, tissue-specific isoforms are present prior to organ formation.

After establishing the cell lineage-specific nature that characterises the transcript levels of nuclear-encoded mitochondrial genes during organogenesis, we asked whether perturbations in the expression of mtDNA-encoded genes will have an effect on those transcriptional signatures or, whether they are established independent of mitochondrial activity.

In order to address this question, we made use of both the m.5024C>T (Kauppila et al., 2016) and the m.5019A>G mouse models generated in the lab of James B. Stewart **(Figure 5.2.)**. The m.5019A>G mouse model was created in the same way as the m.5024C>T one, by being identified in the offspring of a heterozygous female that carried a copy of the PolgA^{D257A} allele, impairing the proofreading activity of the sole mtDNA polymerase (Trifunovic et al., 2004). Again, following the phenotype-driven approach described by Kauppila and her colleagues (Kauppila et al., 2016), the colonic crypts of pups were screened for high levels of COX-deficiency. Laser-capture microdissection was then employed and revealed that mitochondrial deficiency co-segregated with a single mtDNA molecule that bore the following three mutations: m.5019A>G, m.15200A>G and m.16232A>T.

Experiments carried out by the Stewart and Minczuk labs provided the evidence required in order to rule out the m.15200A>G and m.16232A>T mutations as major contributors of the mouse's pathological phenotype while at the same time, highlighting the m.5019A>G mutation as the primary suspect. Despite the fact that m.5019A>G is positioned only 5 base pairs away from the m.5024C>T mutation **(Figure 5.2.)**, its presence on mt-Ta is not the cause of instability. Instead, m.5019A>G leads to the conversion of the G·U wobble base pair positioned at the very top of the mt-Ta acceptor stem, into a canonical G-C complementary base pair.

Experimental evidence suggests that the high levels of heteroplasmy of m.5019A>G lead directly to an increase in the proportion of uncharged to charged mt-Ta molecules.

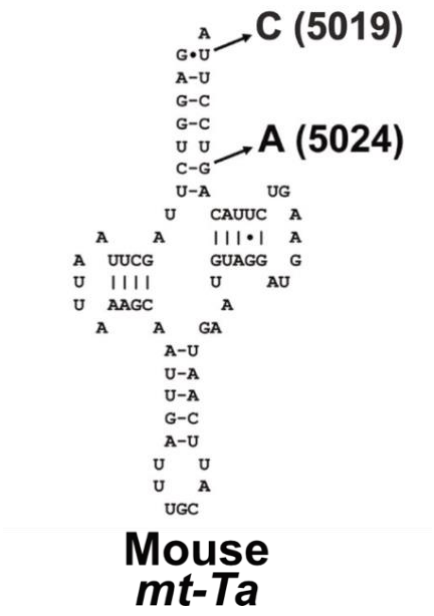


Figure 5.2. Position of the novel m.5019A>G *mt-Ta* mutation.
Positioning of the m.5024C>T and the novel m.5019A>G mutations on the acceptor stem of the mouse mt-Ta molecule.

In this chapter, I present the scRNA-seq work that was carried out using the 10x Genomics platform, using cells from mouse embryos at E8.5. I cover in detail our approach towards sample collection and preparation that was instrumental in generating the scRNA-seq dataset. I also outline differences in nuclear-encoded mitochondrial gene expression patterns that were identified within different cell lineages. Furthermore, I present how *mt-Ta* mutations bring forth an embryo-wide transcriptional response as well as the various mitochondrial and nuclear compensatory responses mounted to account for these pathogenic mutations. Additionally, I demonstrate how many of these transcriptional responses are mutation and cell lineage-specific in nature. Finally, I cover the ways in which gene regulatory networks work to coordinate cell lineage-specific compensatory responses. The project presented in this chapter was a highly collaborative one, and required the joint work of three different labs. Dr Stephen P. Burr and I were responsible for the wet-lab work involving sample collection and preparation, while bioinformatic analysis of the 10x scRNA-seq dataset and generation of the corresponding figures have been the work of Dr Florian Klimm and Dr Malwina Prater, past and current postdoctoral scientists in the labs of Prof Nick Jones and Prof Patrick Chinnery.

5.3. Nuclear-encoded mitochondrial genes exhibit differences in expression patterns during organogenesis

Even though there is plenty of evidence pointing towards differences in mitochondrial activity across different tissues (Johnson et al., 2007a; Johnson et al., 2007b; Fernandez-Vizarra et al., 2011), it has not yet been determined when this variability first surfaces. We sought to determine whether differences in mitochondrial activity are present at the developmental stage of organogenesis, when the organ-specific cell lineages are first established. In pursuit of this question, we scheduled routine mouse matings using C57Bl/6 WT mice and collected embryos at E8.5 post plugging. After the developmental stage of the embryos was morphologically confirmed **(Figure 5.3.)**, three were pooled together and submitted for 10x sequencing. Single cell transcript analysis was then performed on 4,924 cells derived from six of those embryos that were successful in passing transcript quality control. The single cell transcripts were analysed using uniform manifold approximation and projection (UMAP) which led to the distinction of ectoderm, mesoderm and endoderm. The three germ layers were identified based on the expression of Foxa1, Pdgfra, and Pou3F1 (Argelaguet et al., 2019; Pijuan-Sala et al., 2019) **(Figure 5.4.A)** and single cell transcriptional signatures allowed for the identification of 17 distinct tissue lineages (Ibarra-Soria et al., 2018) **(Figure 5.4.B)**.

The overall abundance of mitochondrial transcripts varied significantly across many of the tissue lineages **(Figure 5.5.A)**. A similarly heterogeneous pattern of expression was observed in the transcript levels of multiple nuclear-encoded mitochondrial genes **(Figure 5.5.B)**. Around 60% of nuclear-mitochondrial genes (631/1043), including Ndufb5, Timm44 and Cox10, showed uniform expression across all tissue lineages. However, differential expression in one or more lineages was observed in the remaining 412 genes, that included both genes involved in key mitochondrial functions, such Bnip3, Ldhb and Slc25a4 as well as tissue-specific isoforms of ETC subunits **(Figure 5.5.B)**. A clear example is provided by the cardiac-specific isoforms of complex IV subunits, *Cox6a1* and *Cox6a2* (Kunz, 2003), that were found to be differentially expressed in the cardiac cell lineage **(Figure 5.6.)**. At the same time, both isoforms of *Cox7a* were also primarily expressed within the cardiac cell lineage **(Figure S4.B)**. In a similar fashion, the placental-specific isoform *Cox4i2* was found to be differentially expressed within the extraembryonic-mesoderm, the cell lineage that acts as the precursor to the placenta **(Figure S4.C)**. When taken together, these findings suggest that the expression of

tissue-specific genes does not take place in response to the establishment of a mature organ's physiological function. Instead, the expression of nuclear-encoded mitochondrial genes is subject to modulation in a lineage-specific manner. A process that takes place at the early stages of embryonic development, prior to organ maturation.

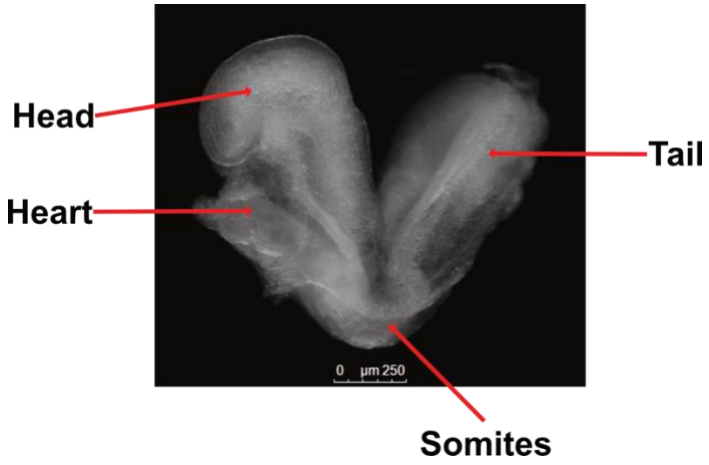


Figure 5.3. Phenotype of mouse embryo at E8.5.
Representative image of a mouse embryo at E8.5, drawing attention to key morphological characteristics. Scale bar: 250µm.

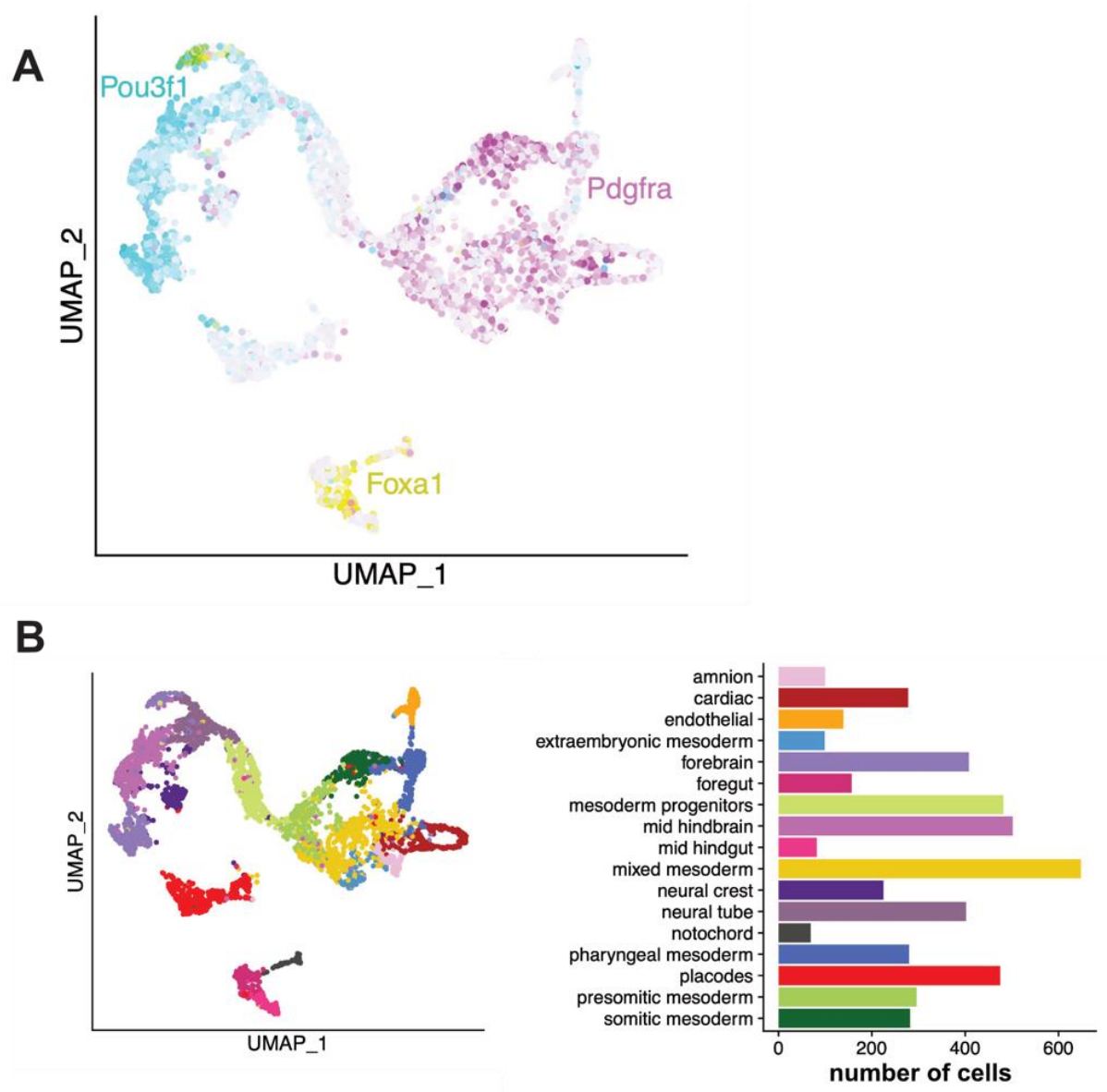


Figure 5.4. Identification of different germ layers and cell lineages at E8.5 mouse embryos.
(A) Uniform Manifold Approximation and Projection (UMAP) displaying the transcripts that were used in order to group the single cells based on the germ layer they originated from (ectoderm = Pou3f1 (blue), endoderm = Foxa1 (yellow) and mesoderm = *Pdgfra* (magenta)). Each dot represents a single cell. **(B)** On the left- UMAP displaying the identification of 17 different cell lines that were defined performing and unbiased analysis of the RNA transcript expression levels of established marker-genes. Each dot represents a single cell. On the right- The number of cells identified as belonging to each cell lineage. Figure generated by Dr Malwina Prater.

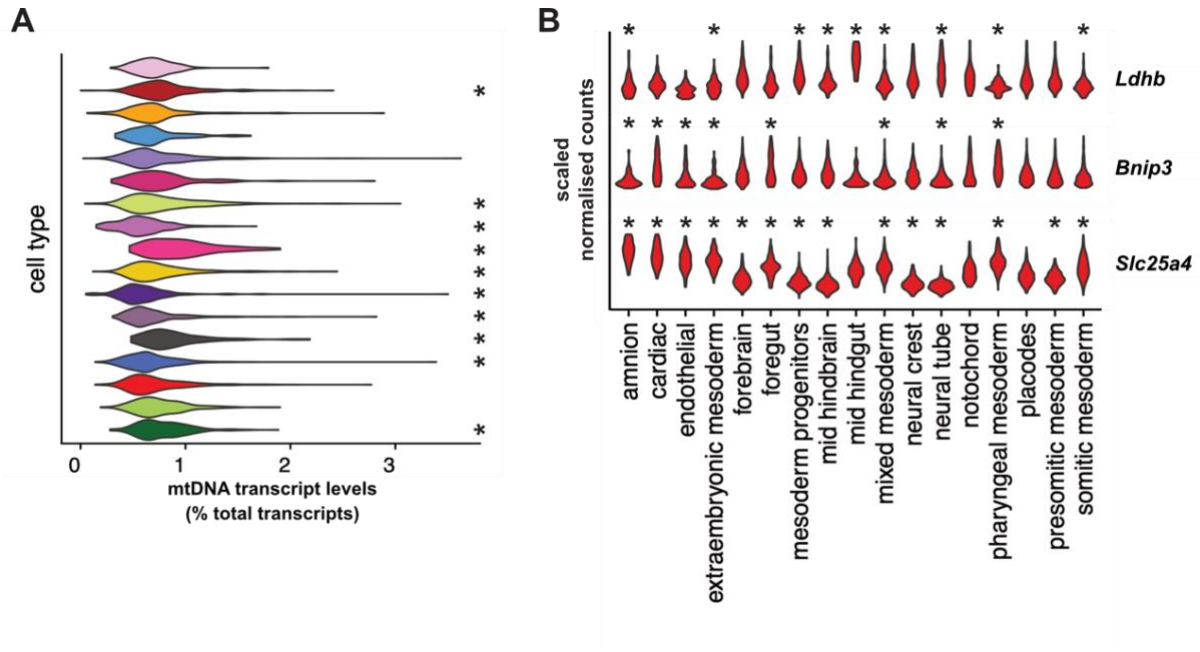


Figure 5.5. Mitochondrial DNA transcript levels in each cell lineage.
(A) Violin plot showing the transcript levels of the mitochondrial DNA (mtDNA) deriving from each of the 17, previously identified, cell lineages. Colour-matched to **(Figure 5.4.B)**. **(B)** The 17 cell lineages identified; each assigned a violin plot showing the expression levels of three select nuclear-encoded mitochondrial genes. An indication is present in lineages where the expression levels of each transcript are significantly different to the scaled normalised average gene expression levels calculated for each of the 17 identified cell lineages. * = $p < 0.05$. Figure generated by Dr Malwina Prater.

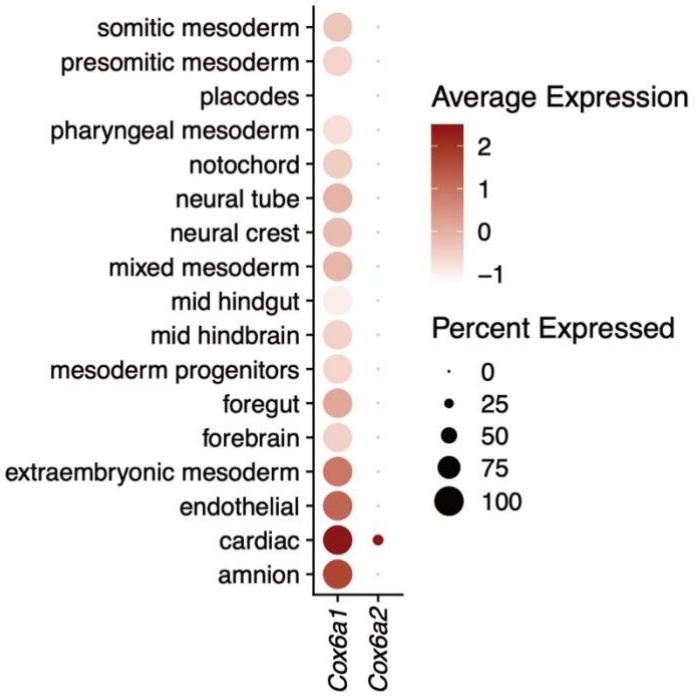


Figure 5.6. Expression of cardiac-specific isoforms of Complex IV subunits.
Dot plot presenting the relative transcript expression levels of the two *Cox6a* isoforms across the 17 cell lineages. Figure generated by Dr Malwina Prater.

5.4. Single Cell RNA Sequencing on Heteroplasmic Mouse Models of Mitochondrial Disease

Based on the work of Kauppila and her colleagues as well as on more recent work from the Stewart and Minczuk labs, we can conclude that once a certain heteroplasmy threshold of either m.5024C>T or m.5019A>G is exceeded, both mutations have a deleterious effect on mt-Ta activity and as a result, on mouse phenotype. Thus, we inferred that any transcriptomic deviations between WT and mutant mice would be more pronounced in mice that carry greater levels of mutant mtDNA. To that effect, when planning mouse mattings, we made sure that high heteroplasmy mothers were used in order to increase our chances of harvesting embryos that fell within the desired heteroplasmy range. At the same time, following dissection, single cells deriving from the embryos of interest were sorted through fluorescence-activated cell sorting (FACS) and pyrosequenced in order to inform us of each sample's mean single cell heteroplasmy **(Figure 5.7.)**.

This process was carried out for high heteroplasmy E8.5 embryos derived from both m.5024C>T and m.5019A>G mothers. Six m.5024C>T embryos were used in order to isolate around 10,000 single cells. These cells had a mean heteroplasmy level of 69% and ranged between 41-93% **(Figure 5.8.A)**. 4,447 of these were passed quality control following sequencing on the 10x genomics platform. Tissue lineage heteroplasmy at E8.5 was determined by looking at allele-specific transcript levels of m.13715C>T, a mutation linked with m.5024C>T (Kauppila et al., 2016) **(Figure 5.8.B)**. In keeping with work done on single E8.5 embryos, as well as with measurements performed of different tissues of neonate mice **(Figure 5.9.)**, both of which were discussed in greater detail in **Results Section, Chapter 4**, similar levels of heteroplasmy were recorded across the 17 tissue lineages present at E8.5. Three embryos carrying the m.5019A>G mutation were collected and around 5,000 cells were isolated from them. In this case, the mean single cell heteroplasmy measured by pyrosequencing was 86%, ranging between 72-94% **(Figure 5.8.A)**. From these, 1,649 cells passed quality control and the m.15200A>G mutation on Cytb linked with 5019 was used as a secondary estimate of mean single cell heteroplasmy (Zhang et al., 2021) **(Figure 5.8.B)**. Again, similar levels of heteroplasmy were recorded across the 17 different lineages as seen in comparisons across tissues of neonate mice **(Figure 5.9.)**.

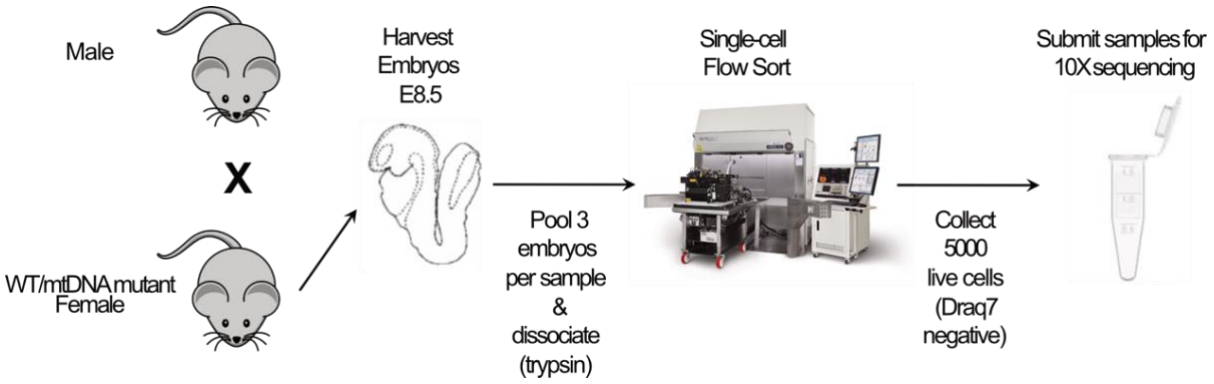


Figure 5.7. Single cell RNA sequencing sample preparation pipeline.
Schematic representation of the workflow pipeline involved in embryonic (E8.5) sample collection, preparation and submission for library preparation and single cell RNA sequencing (scRNA-seq). Schematic created by Dr Stephen Burr.

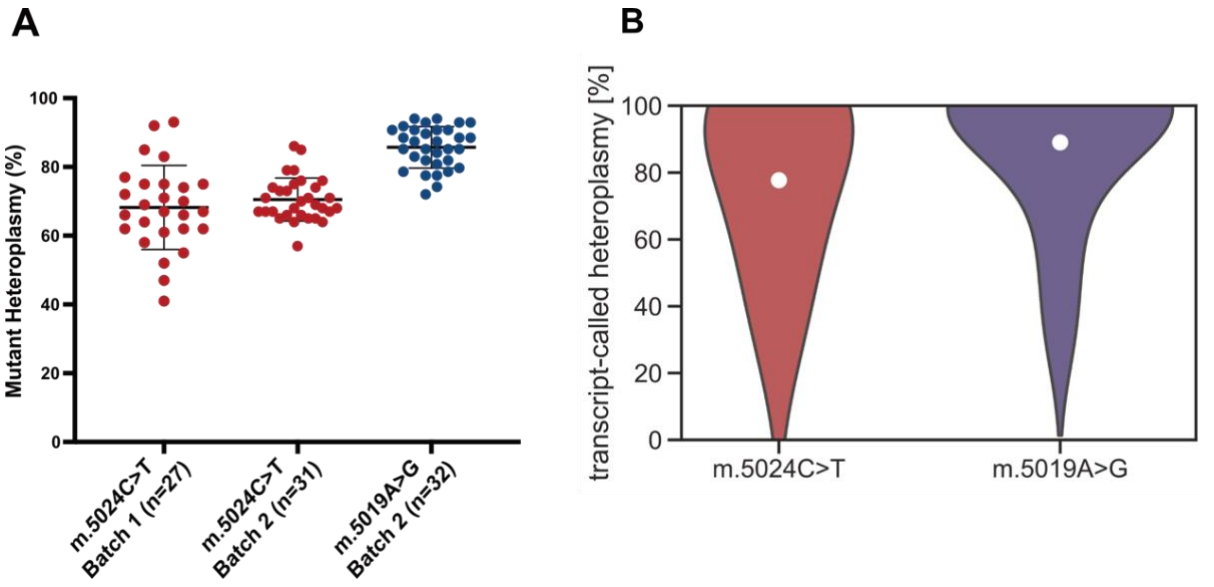


Figure 5.8. Estimating the mitochondrial DNA heteroplasmy of each sample.
(A) Representative pyrosequencing measurements of mtDNA heteroplasmy in cell derived from the 3 grouped m.5019A>G and m.5024C>T E8.5 embryos that were submitted for 10x single cell RNA sequencing (scRNA-seq). Horizontal lines represent mean heteroplasmy value of the population. **(B)** Calls of heteroplasmy percentages originating from scRNA-seq data, based on the transcript variants of m.5019A>G and m.5024C>T E8.5 mouse embryos. **Figure 5.8.B** was generated by Dr Florian Klimm.

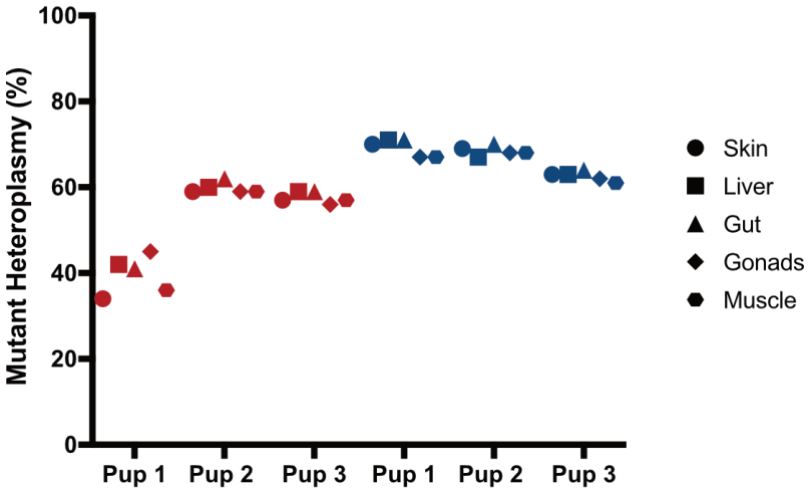


Figure 5.9. Bulk heteroplasmy measurements in neonate mouse tissues
Bulk tissue mitochondrial DNA (mtDNA) heteroplasmy measurements performed using pyrosequencing, on five tissues derived from neonate (P0) mice carrying either the m.5019A>G (seen in blue) or the m.5024C>T (seen in red) mutation.

5.5. Mutations in mt-Ta Provoke an Embryo-Wide Transcriptional Feedback

Having completed the scRNA-seq for WT, m.5019A>G and m.5024C>T mouse embryos at E8.5, we set out to uncover whether the perturbation of mitochondrial translation by the two pathogenic point mutations on *mt-Ta* affects in some way the transcriptional profile of each cell lineage. To that effect, a UMAP was generated, this time including the entirety of the 11,020 cells that underwent RNAseq using the 10x Genomics platform. Cells derived from WT, m.5019A>G and m.5024C>T embryos all showed almost identical clustering patterns **(Figure S5.)**. All 17 cell lineages first identified in the WT embryos were also seen in the two mutant strains. Pseudo-bulk analysis, which involves the clustering of single cell data in order to resemble a single pseudo-sample, demonstrated that embryos derived from both the m.5019A>G and m.5024C>T mouse lines, displayed dysregulated expression of nuclear-encoded mitochondrial genes (Calvo et al., 2016). In accordance with the phenotype of mitochondrial translation dysregulation observed in both the m.5019A>G and m.5024C>T mouse models (Kaupplia et al., 2016), genes involved in many key mitochondrial pathways were downregulated **(Figure 5.10.)**. In more detail, downregulation compared to cells deriving from WT embryos was observed in transcripts involved in tRNA aminoacylation, ATP synthesis and ETC formation. At the same time, an upregulation of mtDNA transcripts was recorded, signifying a compensatory response to the defect in mitochondrial translation across all cell lineages **(Figure 5.11.)**.

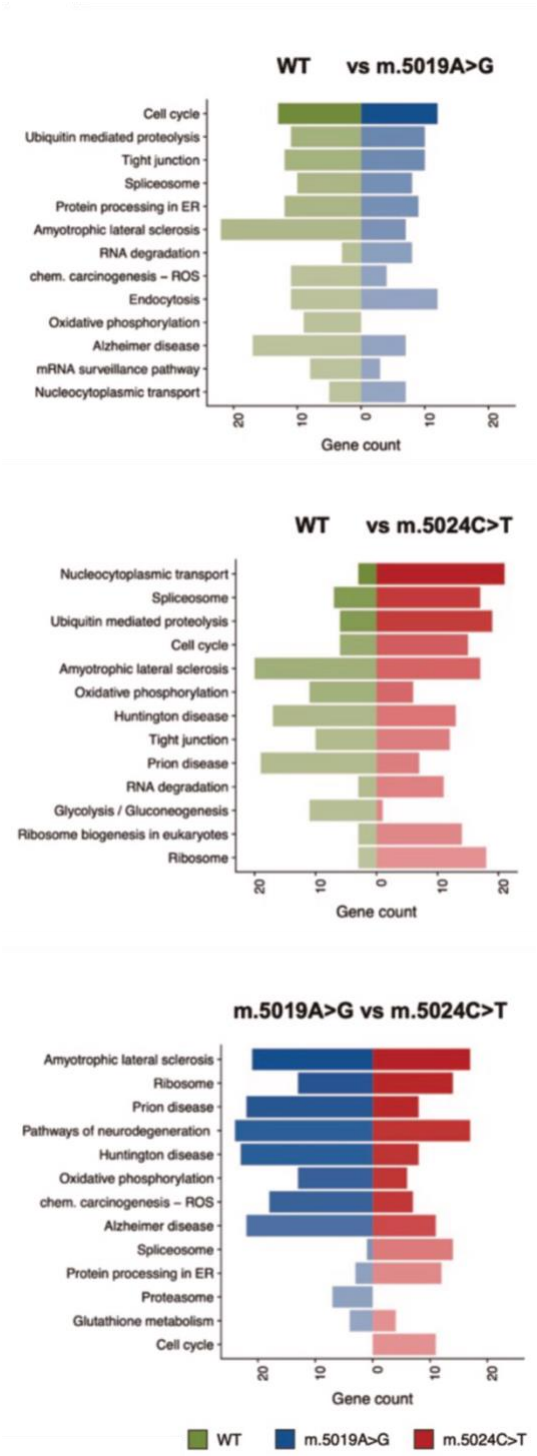


Figure 5.10. Pathogenic variants of *mt-Ta* bring forth a transcriptional response throughout the embryo.
Sideways bar charts showing pseudobulk pathway comparisons across the three different mouse lines (wild-type (WT), m.5019A>G and m.5024C>T). Cellular pathways showing a significant number of differentially expressed genes are ordered from most to least significant. Each coloured bar represents the number of differentially expressed genes being either up or downregulated in each pathway for each of the comparisons. Figure generated by Dr Malwina Prater.

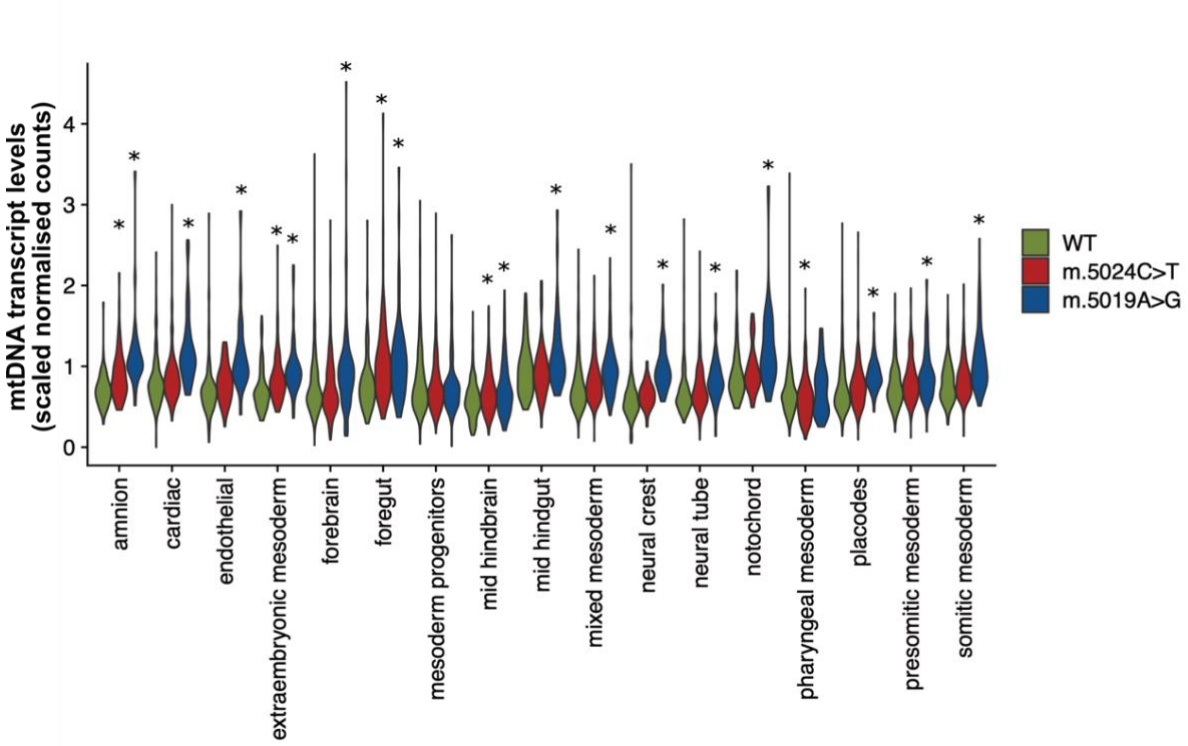


Figure 5.11. Mitochondrial DNA transcript levels in *mt-Ta* mutants compared to controls. Violin plots depicting the transcript levels of mitochondrial DNA (mtDNA) that were detected in each of the 17 cell lineages of WT, m.5024C>T and m.5019A>G mouse embryos. * = $p < 0.05$ (comparing each mutant to the WT). Figure generated by Dr Malwina Prater.

5.6. Nuclear and Mitochondrial Adaptations in Response to mt-Ta Pathogenic Mutations

While the impact of both the m.5019A>G and m.5024C>T mutations on mitochondrial performance and the activity of the ETC has been well documented by the Stweart and Minczuk labs (data not shown). Given the known tolerability of mt-tRNA mutations (Freyer et al., 2012; Stewart and Larsson, 2014), we fail to see any clear developmental abnormalities surfacing during organogenesis. At the same time, there are only minor difference in the number of cells that are committed to each cell lineage **(Figure 5.12.A)** and no difference when it comes to the genes expression profiles that define each of the 17 cell lineages **(Figure 5.12.B)**. As a result, we had to look past the absence of any obvious developmental defects and into the compensatory mechanisms that inevitably kick in, preventing the development of a more severe pathological phenotype.

In his work published in 2019, To and his colleagues performed genome-wide clustered regularly interspaced short palindromic repeats (CRISPR) screens that employed single molecule inhibitors of OXPHOS complexes (To et al., 2019). They discovered 191 unique genetic modifiers that included 91 knockouts that functioned as epistatic buffers of the OXPHOS-targeting toxins, 63 that led to suppression of their toxic effect and 38 that appeared to be synthetic lethal (To et al., 2019). The *in vitro* knockout of the genes labelled as synthetic lethal genes led to cell death prior to the exposure of the cells to a specific toxin. On the other hand, the knockout of genes that functioned as epistatic buffers mitigated the toxicity of an OXPHOS inhibitor. Finally, knocking out genes that were of the suppressor category, led to the acceleration of cell growth in the presence of the toxin. When taking a closer look at this dataset, we observed that from the 191 unique genetic modifiers identified by To et al., 63 were found to be dysregulated in the m.5024C>T mutant and 40 in the m.5019A>G mutant, with 22 of them found to be dysregulated in both mutant mouse strains **(Figure 5.13.A)**. Observing such an extensive overlap, we postulated that since there was no observable effect in the development of any of the 17 identified cell lineages in either the m.5019A>G and m.5024C>T mutant mice, genes that were identified as synthetic lethal when knocked out by To et al., would either have their expression levels increased or remain constant. Extending this thought process further, we hypothesised that genes acting as suppressors or epistatic buffers would be

under-expressed in the two mutant mouse strains as their knockout *in vitro* mitigated the effect of the toxins on the cell's OXPHOS activity **(Figure 5.13.B)**.

Indeed, analysis of the scRNA-seq data, revealed that when comparing the mouse strains carrying mt-Ta mutations, to their WT counterparts, a significant enrichment for these epistatic buffer/suppressor genes underlined their decreased expression levels **(Figure 5.14.A)**. Epistatic buffer/suppressor genes predominantly encoded for proteins that are predicted to be localised in the mitochondria. These proteins form a consistent network of interactions, centred around Mrpl12 and Atp5d and involving proteins that are vital to the cell's OXPHOS activity as well as more distantly related components (Szklarczyk et al., 2019) **(Figure 5.14.B)**. At the same time, we observed the downregulation of nuclear-encoded subunits of ATP-synthase in addition to respiratory complexes I and IV as well as that of the assembly factor Surf1 and mitochondrial ribosomes alongside other key factors that are instrumental in the process of mt-DNA translation. Not just limited to intra-mitochondrial or 'intrinsic' factors, 'extrinsic' proteins that are involved in the regulation of nuclear gene expression were also implicated **(Figure 5.14.A)**. More specifically, epigenetic regulators of gene expression Prmt1 and Suds3, intracellular trafficking components Scap and Sec23b, DNA-binding transcription factor Klf16 and Mthfd1, an enzyme involved in 1-carbon metabolism that has also been implicated in mitochondrial pathologies (Suomalainen and Battersby, 2018) **(Figure 5.14.A)**. Our analysis provided further evidence in support of the critical role that these genetic modifiers, identified through *in vitro* CRISPR screens, play in the regulation of cellular responses to OXPHOS perturbation. Moreover, we have provided *in vivo* validation of their role and biological relevance in the process of organogenesis during embryonic development. Further *in vitro* work, carried out by Bonekamp and colleagues, made use of specific inhibitors of mitochondrial transcription (IMTs) to target POLRMT and as a result, impair OXPHOS activity (Bonekamp et al., 2020). They discovered 526 genetic modifiers that are believed to be involved in the cellular responses to OXPHOS dysfunction (Mennuni et al., 2022). From the 178 genetic modifiers uncovered in the *in vitro* screen carried out by To et al., 78 were corroborated through our *in vivo* experimentation as overlap was observed with both the m.5024C>T and the m.5019A>G mouse models. Notable candidates that were commonly identified across the two experiments involved genes participating in oxygen-sensing and the mTORC1 pathway. When the *in vivo* scRNA-seq data we have generated are compared to these two in vitro studies involving OXPHOS disruption, evidence emerges in support of a positive compensation being present within both the m.5024C>T and the m.5019A>G mouse embryos.

Even though a positive compensatory cellular response would override a more passive reaction to the decrease of the cell's mitochondrial activity, at this point we cannot rule out the existence of other mechanisms being involved in parallel to what we have observed so far.

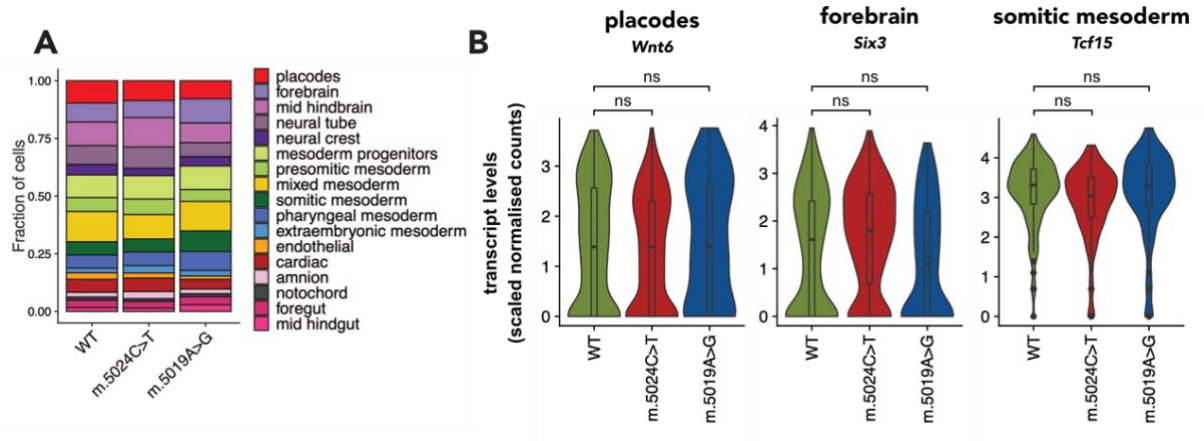


Figure 5.12. Transcriptomic analysis reveals no obvious developmental defects.
(A) Bar chart showing the proportion of cells that can be identified as being part of each of the 17 cell lineages in E8.5 embryos belonging to wild-type (WT), m.5024C>T and m.5019A>G mice. **(B)** Violin plots comparing the expression levels of select lineage-specific marker genes in their corresponding cell types (Wnt6 = placodes, Six3 = forebrain and Tcf15 = somatic mesoderm) belonging to WT, m.5024C>T and m.5019A>G mouse embryos. Figure generated by Dr Malwina Prater.

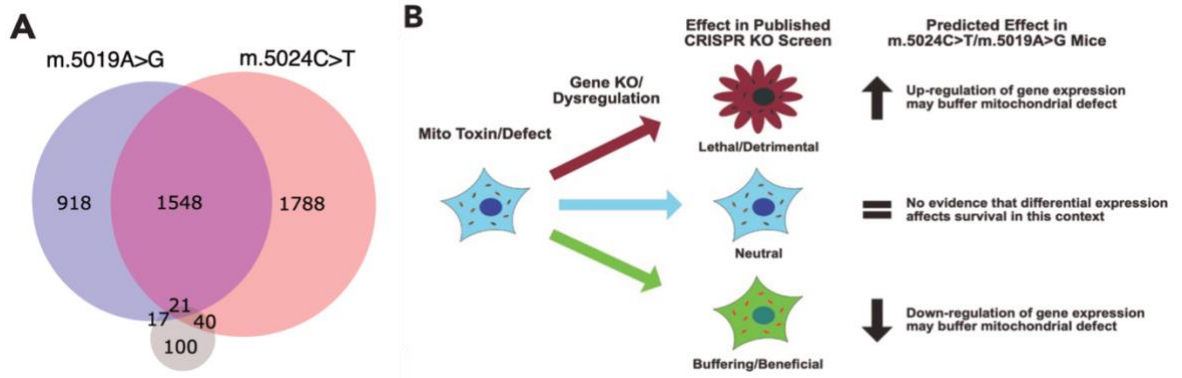


Figure 5.13. Overlap in epistatic regulators and their significance.
(A) Venn diagram showing the overlap between dysregulated genes found in the m.5019A>G (blue) and m.5024C>T (pink) embryos, as well as between genetic modifiers of mitochondrial dysfunction, identified through an *in vitro* screen by To et al. (grey). **(B)** Schematic representation of the proposed relationship between the dataset produced when Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) knock-out (KO) screens accompanied the application of mitochondrial toxins and the predicted effect that m.5024C>T and m.5019A>G will have on E8.5 mouse embryos. Figure generated by Dr Malwina Prater.

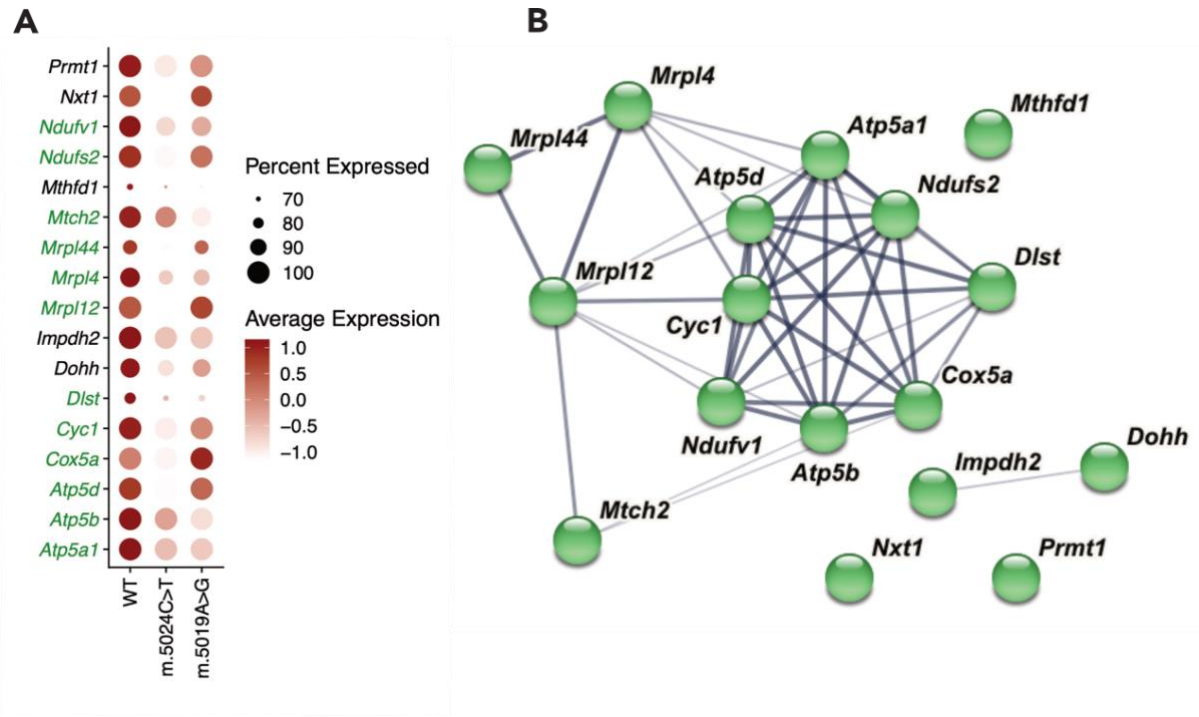


Figure 5.14. A closer look into epistatic buffer/suppressor genes and their interaction.
(A) Dot plot listing 17 buffer/suppressor genes and their corresponding expression levels. Genes included were found to be significantly downregulated in at least one of the two mutant mouse strains when compared to their WT counterparts. Genes that are highlighted in 'green' are included in the Mouse MitoCarta3.0 inventory. **(B)** Genes introduced in **(A)** presented as part of a STRING protein interaction network. The network was generated using STRING v11, with a minimum required score to establish an interaction of 0.4. The degree of confidence in each prediction is represented by the thickness of the line. The thicker the line, the higher the confidence in the interaction. Figure generated by Dr Malwina Prater.

5.7. Analysis of Mutation- and Context-Specific Responses at the Transcript and Protein Level.

When pseudobulk analysis was performed comparing the m.5024C>T to the m.5019A>G mouse strains, there were multiple genes that appeared to be differentially expressed at the level of the embryo **(Figure 5.15.A)**. Further pseudobulk analysis that involved pooling cells stemming from specific lineages together, revealed differentially expressed genes (DEGs) that were arising at the tissue level **(Figure 5.15.B)**. When conducting a WikiPathways analysis **(Figure 5.16.)**, multiple genes involved in cytoplasmic ribosomal proteins, mRNA processing and cell pluripotency were found to be differentially transcribed across one or more cell lineages in both mutant mouse strains, when compared to their control counterparts. At the same time, there were many additional pathways that appeared to be differentially regulated in one or more of the 17 observed cell lineages. Similar results were generated when we made use of Gene Ontology (GO) Biological Processes. Both results lend further credence to a combined mutation-specific as well as lineage-specific cellular response brought about as a result of pathogenic mtDNA mutations.

Having established the strain and cell-lineage-specific nature of transcriptional responses, we set out to examine whether the aforementioned genetic modifiers that were identified by To and colleagues in a genome-wide CRISPR screen, follow a similar expression pattern (To et al., 2019). In both the m.5024C>T and the m.5019A>G mouse lines, we observed an enrichment for epistatic buffer and suppressor genes that were downregulated when compared to their WT expression levels. This behaviour characterised many more embryonic cell lineages belonging to the m.5024C>T mouse strain compared to m.5019A>G, where a strong enrichment was observed only in a limited number of cell lineages **(Figure 5.17.)**. This pattern was also repeated in global pathway and pseudobulk analyses **(Figure 5.16.)**. Overall, this data point towards a stronger compensatory response involving the downregulation of epistatic buffering and suppressor genes **(Figure 5.18.)**, taking place within cell lineages originating from the m.5024C>T mouse strain. This stronger compensatory response might be the answer to why litter sizes in mice of the m.5019A>G mouse strain are significantly lower compared to those of the m.5024C>T strain, averaging 5 pups instead of 6 pups per litter **(Figure S7.)**.

A lineage-specific expression pattern was also observed for most epistatic buffer and suppressor genes **(Figure 5.18.)**. This was brought to light through their scaled expression across multiple cell lineages. A great example is provided by the chicken ovalbumin upstream promoter transcription factor II (COUP-TFII) or Nr2f2, that codes for a ligand-inducible transcription factor, responsible for regulating the expression of multiple genes. It is vital for the development of a number of organs, that include mesoderm-derived blood vessels (Sissaoui et al., 2020), as well as specific types of neurons in the brain (Alzu'bi et al., 2017; Hu et al., 2017). Nr2f2 was identified by To et al. as buffering gene and in our work, it was found to be downregulated more strongly in the neural crest, neural tube and mesodermal cell lineages derived from m.5024C>T mouse embryos when compared to their m.5019A>G counterparts. Conversely, we were able to identify a number of genes, the expression of which was preferentially upregulated in embryos derived from the m.5019A>G mouse strain. Multiple subunits of respiratory complexes were among those genes. More specifically, the expression of CI subunits (Ndufa6/10, Ndufb3/5/7/9/10, Nduf11, Ndufs5, Ndufs8), CIV subunits (Cox5a/14) and CV subunits (Atp5b/d/e) was found to be upregulated in cardiac cell precursors **(Figure S8.)**. At a first glance, upregulation of all these components of respiratory chain complexes might appear beneficial for the organism, similar to the 'sick mitochondrion' or 'ragged red' response staged to counter intra-mitochondrial translational defects (Schon et al., 1997). However, given that all of the aforementioned genes were described as epistatic buffer or suppressors (when knocked out) by To and colleagues, their upregulation is likely to be deleterious for the development of embryos belonging to the m.5019A>G mouse strain.

Despite the established contribution of integrated mitochondrial stress response (ISRmt) in cellular metabolic stress response (Suomalainen and Battersby, 2018), we observed no transcriptional upregulation in either mutant mouse strain when looking into changes impacting genes involved in ISRmt. This observation is in line with results suggesting that Atf4 (a regulator central to ISRmt) transcription is not increased in the heart of adult m.5024C>T mice (Filograna et al., 2019). In conclusion, we can see examples of both mutation and lineage-specific compensatory transcriptional responses **(Figure 5.15.)** that are likely contributing towards the development of phenotypic differences between the strains as well as organ-specific phenotypes.

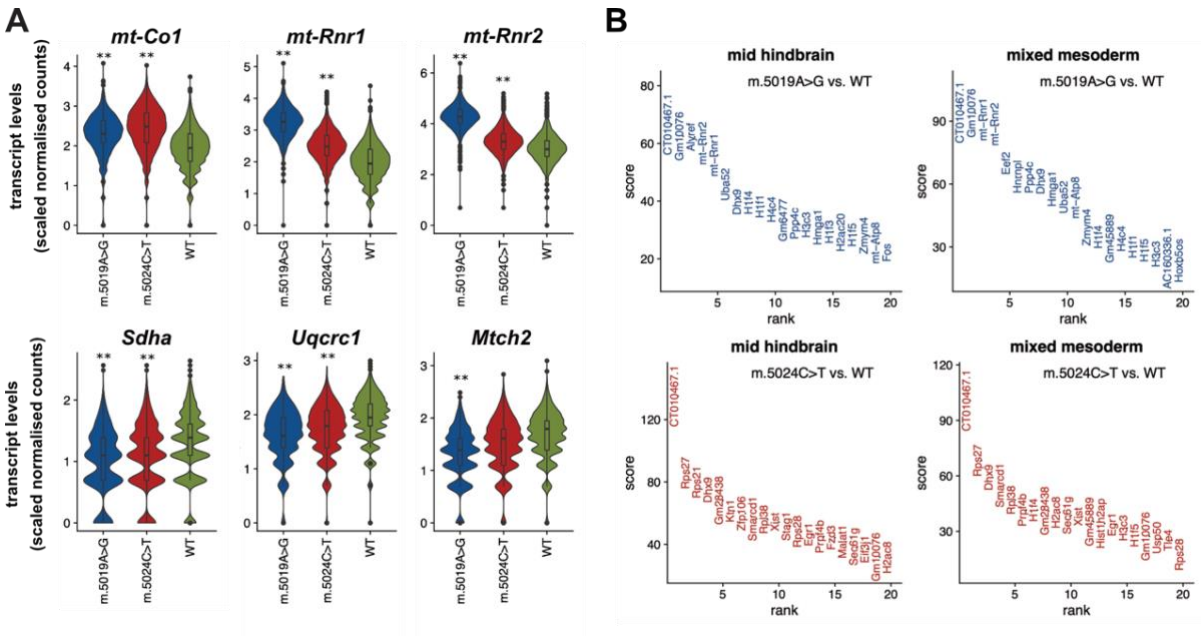


Figure 5.15. Mutation- and cell lineage-specific differentially expressed genes.
(A) Violin plots representing the pseudobulk expression levels of select mitochondrial DNA genes as well as nuclear ones that encode for mitochondrial proteins in wild-type (WT), m.5024C>T and m.5019A>G E8.5 mouse embryos. ** = $p < 0.01$. **(B)** Listing of the top 20 most upregulated genes in mid hindbrain and mixed mesoderm of m.5024C>T and m.5019A>G E8.5 mouse embryos compared to their WT counterparts. Ranking of the genes from left to right is based on relative expression score ($-\log_{10}$(p.adj) * abs(l2fc)) compared to WT. Figure generated by Dr Malwina Prater.

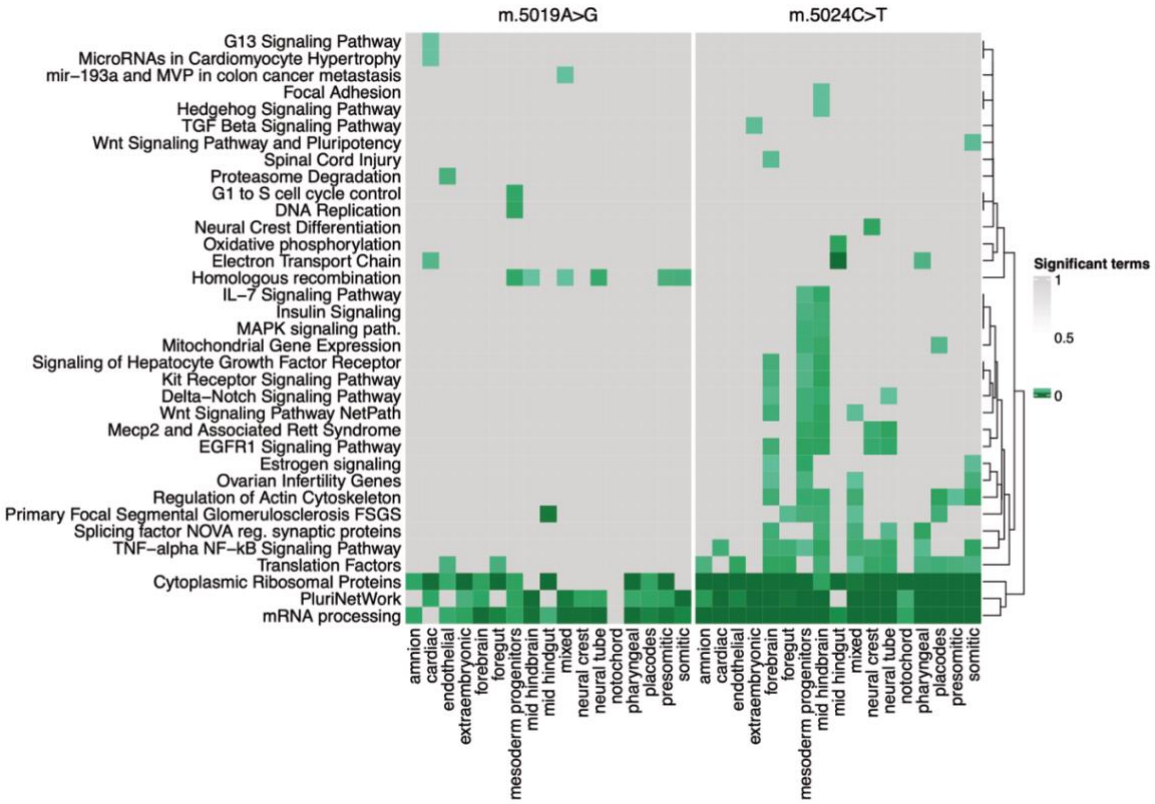


Figure 5.16. Cellular processes that exhibit differentially expressed genes.
Heatmap pointing towards WikiPathways (2019 mouse annotation) that have been significantly enriched. Analysis is based on differentially regulated genes that are present within at least one cell lineage of the m.5024C>T and m.5019A>G E8.5 mouse embryos when compared to their WT counterparts. All green squares exhibit $p<0.05$. Figure generated by Dr Malwina Prater.

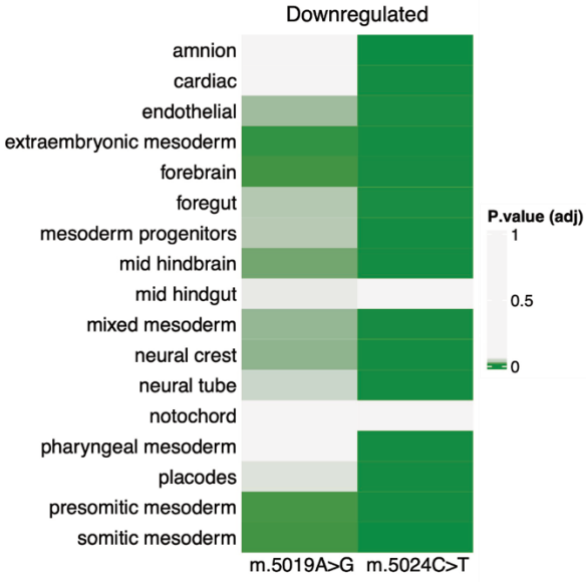


Figure 5.17. Enrichment of epistatic buffer/suppressor gene downregulation across the two mutant mouse strains.
Heatmap showing how epistatic buffers/suppressors are enriched amongst the genes that are downregulated across all cell lineages in embryos carrying the m.5024C>T and m.5019A>G mutations compared to WT. Figure generated by Dr Malwina Prater.

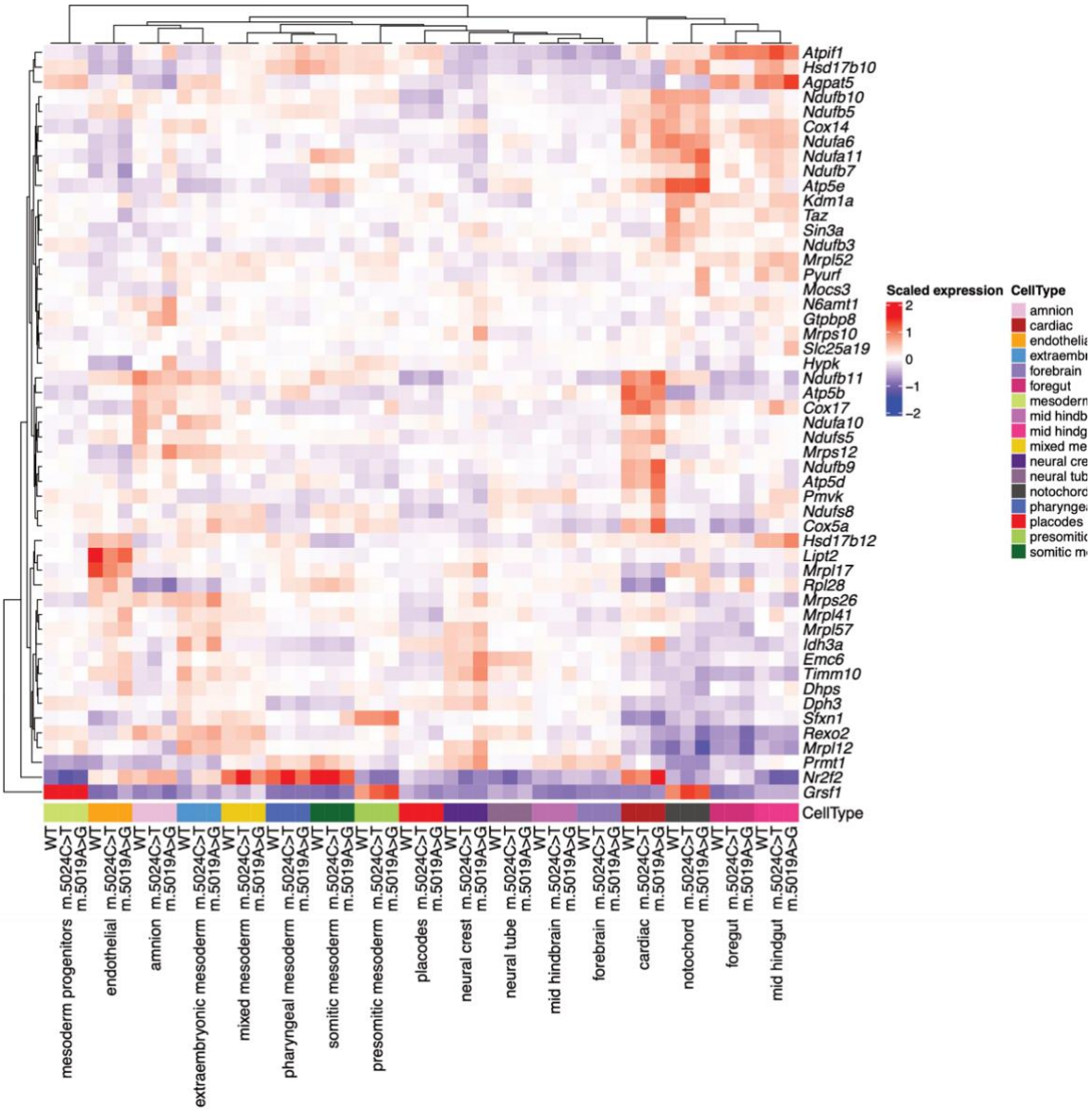


Figure 5.18. The differential expression of epistatic buffer/suppressor genes is both mutation- and cell lineage-specific.

Heatmap demonstrating the relative expression levels of 51 of the epistatic buffer/suppressor genes that were identified as part of the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) knock-out screens. Results have been filtered in order to present genes that are differentially expressed between embryos carrying the m.5024C>T or m.5019A>G mutation and their wild-type (WT) counterparts. Filters were applied on scaled gene expression, with the minimum difference between the mutant and WT being < -0.2 and maximum being > 0.5 for at least one cell type across all 17 cell lineages and the three mouse strains. Figure generated by Dr Malwina Prater.

Work carried out by the Stewart and Minczuk labs has come a long way in examining the extensive molecular and phenotypic impact of the novel m.5019A>G *mt-Ta* mutation (data not shown). Aiming to build on their work, we decided to conduct further experiments on the mutation's impact at the level of the proteome. To that effect, we ran a quantitative proteomics experiment in order to create an in-depth picture of the effect each *mt-Ta* mutation has on the proteome of adult mouse liver and cortex samples. When submitting samples, we aimed at picking mice with high heteroplasmy values from both the m.5019A>G and m.5024C>T strains. Furthermore, we controlled for the age (~20 weeks) and gender (male) of the mice as well as the nuclear background **(Table 5.1.)**. This meant that mice belonging to the m.5024C>T strain and their corresponding controls have a C57BL/6J nuclear background. Although very commonly used amongst mouse strains, this nuclear background has a 17,814bp deletion on the nicotinamide nucleotide transhydrogenase (*Nnt*) gene, between exons 6 and 12 that leads to the mature protein not being detected in these animals (Huang et al., 2006). Conversely, mice belonging to the m.5019A>G strain were bred on a C57BL/6N nuclear background that is WT for *Nnt* and this was kept constant in its respective WT controls.

After generating this dataset, we decided to take a closer look at the effect the impact the novel m.5019A>G *mt-Ta* mutation has on both nuclear- and mitochondrial-encoded subunits belonging to complexes of the ETC **(Figure 5.19.A)**. It was evident that while a mild reduction in the protein levels of ETC complex subunits was visible in the cortex samples of m.5019A>G mutants compared to their WT controls, no significant changes were observed. However, in the case of liver samples that belonged to m.5019A>G mice, we were able to detect quite a few subunits of the ETC as being downregulated. When taking a closer look, we noted an ETC-wide downregulation of subunits belonging to all five complexes with ones belonging to complexes I, III and IV exhibiting significantly lower protein levels compared to their corresponding controls **(Figure 5.19.B)**.

Table 5.1. Outline of mice used in the proteomic analysis.

Number	mtDNA Haplotype	Heteroplasmy %	Gender	Nuclear Background	Age Culled
1	WT	-	Male	C57BL/6J	20w 3d
2	WT	-	Male	C57BL/6J	20w 3d
3	WT	-	Male	C57BL/6J	20w 3d
4	m.5019A>G	76%	Male	C57BL/6J	19w 6d
5	m.5019A>G	85%	Male	C57BL/6J	22w 5d
6	m.5019A>G	74%	Male	C57BL/6J	22w 3d
7	WT	-	Male	C57BL/6N	20w 0d
8	WT	-	Male	C57BL/6N	20w 0d
9	WT	-	Male	C57BL/6N	20w 0d
10	m.5024C>T	68%	Male	C57BL/6N	22w 6d
11	m.5024C>T	73%	Male	C57BL/6N	22w 5d
12	m.5024C>T	68%	Male	C57BL/6N	20w 2d

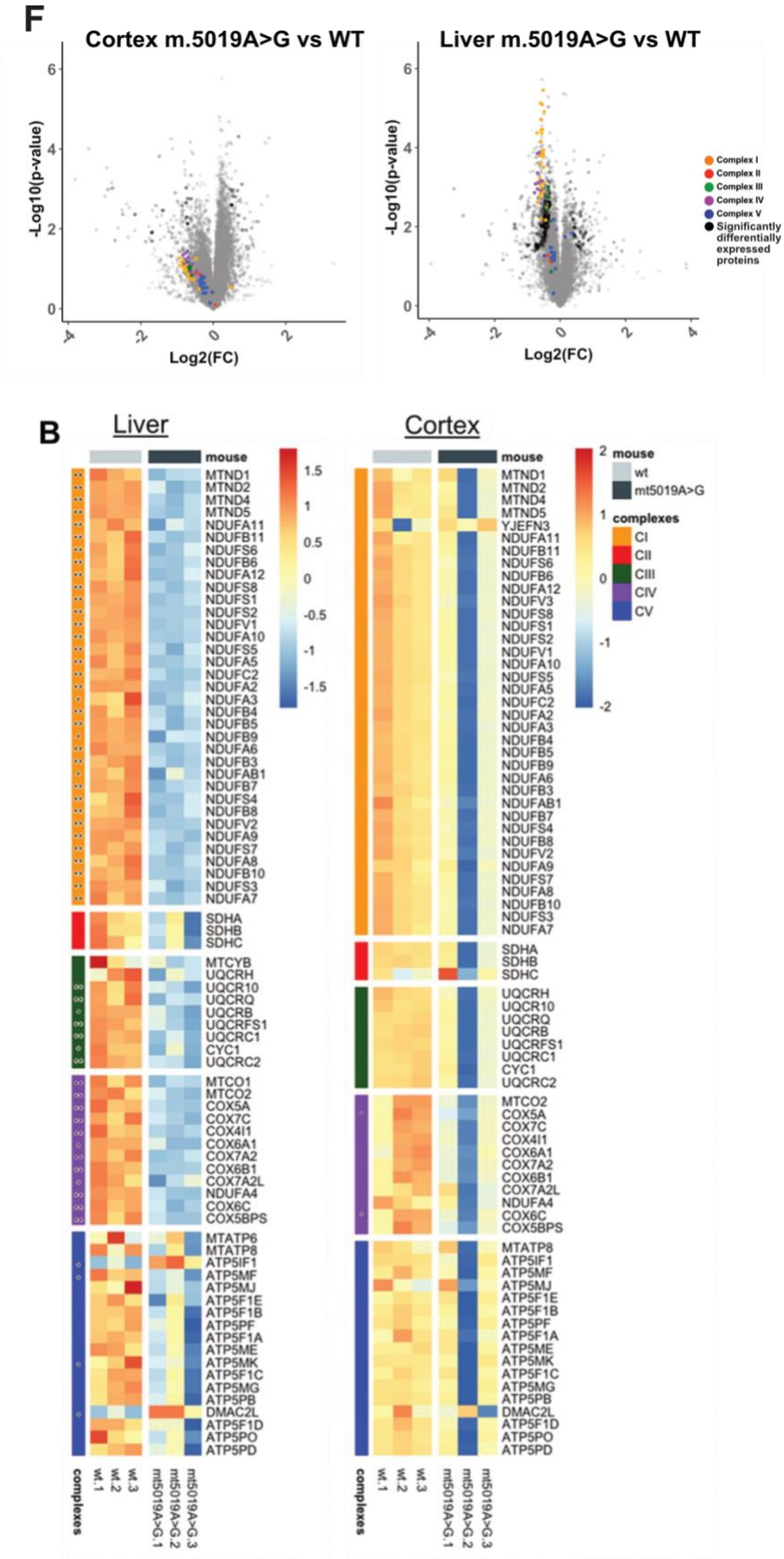

F
Cortex m.5019A>G vs WT
Liver m.5019A>G vs WT
-Log10(p-value)
Log2(FC)
Complex I
Complex II
Complex III
Complex IV
Complex V
Significantly differentially expressed proteins
B
Liver
Cortex
mouse
wt
mt5019A>G
complexes
CI
CII
CIII
CIV
CV
MTND1
MTND2
MTND4
MTND5
YJEFN3
NDUFA11
NDUFB11
NDUFS6
NDUFB6
NDUFA12
NDUFV3
NDUFS8
NDUFS1
NDUFS2
NDUFV1
NDUFA10
NDUFS5
NDUFA5
NDUFC2
NDUFA2
NDUFA3
NDUFB4
NDUFB5
NDUFB9
NDUFA6
NDUFB3
NDUFAB1
NDUFB7
NDUFS4
NDUFB8
NDUFV2
NDUFA9
NDUFS7
NDUFA8
NDUFB10
NDUFS3
NDUFA7
SDHA
SDHB
SDHC
MTCYB
UQCRH
UQCR10
UQCRQ
UQCRB
UQCRFS1
UQCRC1
CYC1
UQCRC2
MTCO1
MTCO2
COX5A
COX7C
COX4I1
COX6A1
COX7A2
COX6B1
COX7A2L
NDUFA4
COX6C
COX5BPS
MTATP6
MTATP8
ATP5IF1
ATP5MF
ATP5MJ
ATP5F1E
ATP5F1B
ATP5PF
ATP5F1A
ATP5ME
ATP5MK
ATP5F1C
ATP5MG
ATP5PB
DMAC2L
ATP5F1D
ATP5PO
ATP5PD
complexes
wt.1
wt.2
wt.3
mt5019A>G.1
mt5019A>G.2
mt5019A>G.3

Figure 5.19. Novel m.5019A>G *mt-Ta* variant leads to a significant decrease of key OXPHOS subunits in adult mouse liver.
(A) Volcano plots showing the protein levels of subunits belonging to the complexes of the electron transport chain (ETC) in the cortex and liver of mice belonging to either the m.5019A>G or m.5024C>T mouse strains compared to their respective wild-type (WT) controls. **(B)** Heat maps showing the protein expression ratio for both mitochondrial- and nuclear-encoded subunits belonging to each ETC complex, derived from the cortex and liver of m.5019A>G and WT mice. Proteins underlined in orange belong to complex I, red to complex II, green to complex III, purple to complex IV and blue to complex V (ATPsynthase). Figure generated by Dr Yu Nie.

5.8. Gene Regulatory Networks Coordinate Compensatory Mechanisms in a Cell Lineage-Specific Manner

In order to shed light into the upstream pathways responsible for driving the compensatory responses to genetically-induced mitochondrial dysfunction, we decided to take a look at gene regulatory networks. Malwina Prater, performed cis-regulatory analysis using single-cell regulatory network inference and clustering (SCENIC) (Aibar et al., 2017). The entire scRNA-seq dataset was used in order to generate 'regulons' and study their expression in the 17 cell lineages that were identified in the E8.5 embryos derived from both the m.5019A>G and m.5024C>T mutant mouse strains in addition to the corresponding WT controls. Regulons can be defined as gene clusters that are jointly regulated, by the same transcription factor that in turn can be acting as either an enhancer or suppressor of gene expression. Unbiased hierarchical clustering was used in order to group all three mouse strains depending on cell lineage and thus, generate a regulon activity matrix. Lineage-specific regulon activation was revealed following the relative scaling of their activity within each cell type **(Figure 5.20.)**. Many of those activated regulons included genes that are directly involved in the development of a specific tissue. A great example is provided by the enhancement of regulon expression that involved GATA transcription factor activity. The expression of regulons controlled by the activity of Gata4/5/6 were found to be increased in cardiac cell progenitors, while the activity of Foxa3 was higher in gut cell progenitors **(Figure 5.21.A)** (Friedman and Kaestner, 2006; Pikkarainen et al., 2004). In an attempt to further substantiate our findings, we compared genes that came up in the SCENIC-derived regulon analysis to *in vitro* RNA-seq data published on mouse and human-derived neural crest cell lines that were subject to siRNA knockdown of the Sox10 transcription factor (Lai et al., 2021; Saur et al., 2021). We were happy to see that the Sox10 knockdown perturbed a lot of genes that were part of the neural crest-specific regulons that were identified as part of our SCENIC analysis. These included Sox9/10 as well as Tfap2a/c. We also managed to detect several transcription factors that were more active in WT embryos compared to their *mt-Ta* mutant counterparts. Mecom is one such transcription factor detected in the endothelial lineage and implicated in the proliferation and differentiation of arterial endothelial cells (Hou et al., 2022) **(Figure 5.21.B)**. On the other hand, cells of the neural crest that belonged to embryos carrying *mt-Ta* mutations, showed higher Sox9/10 activity compared to their WT counterparts **(Figure 5.21.B)**. Both Sox9 and Sox10 play a central role in the differentiation and migration of embryonic neural crest cells (Haldin and

LaBonne, 2010). Their expression was found to be linked to that of Sox5, Zeb2, Ets1, Cdh6/11, Erbb3 and Tfapb2 in embryos derived from both the m.5019A>G and m.5024C>T mouse strains **(Figure 5.22.)**. All of these genes are involved in the transition from epithelium to mesenchyme and play a key role in neural crest specification which begins at the 4-5 somite stage, a developmental phase that corresponds to E8.0-8.5 in the mouse embryo (Trainor et al., 2005; Simoes-Costa and Bronner, 2015). This data comes in support of recent observations that describe mitochondrial activity as being an important factor in driving differentiation and specification of cells belonging to the neural crest lineage (Costa et al., 2021). Taken together, data derived from the SCENIC analysis suggests that genetically-induced mitochondrial dysfunction drives the differential regulation of key developmental pathways involved in lineage-specific cell differentiation and maturation. Transcriptional regulation of such pathways takes place in the form of regulons, networks comprised of co-expressed genes.

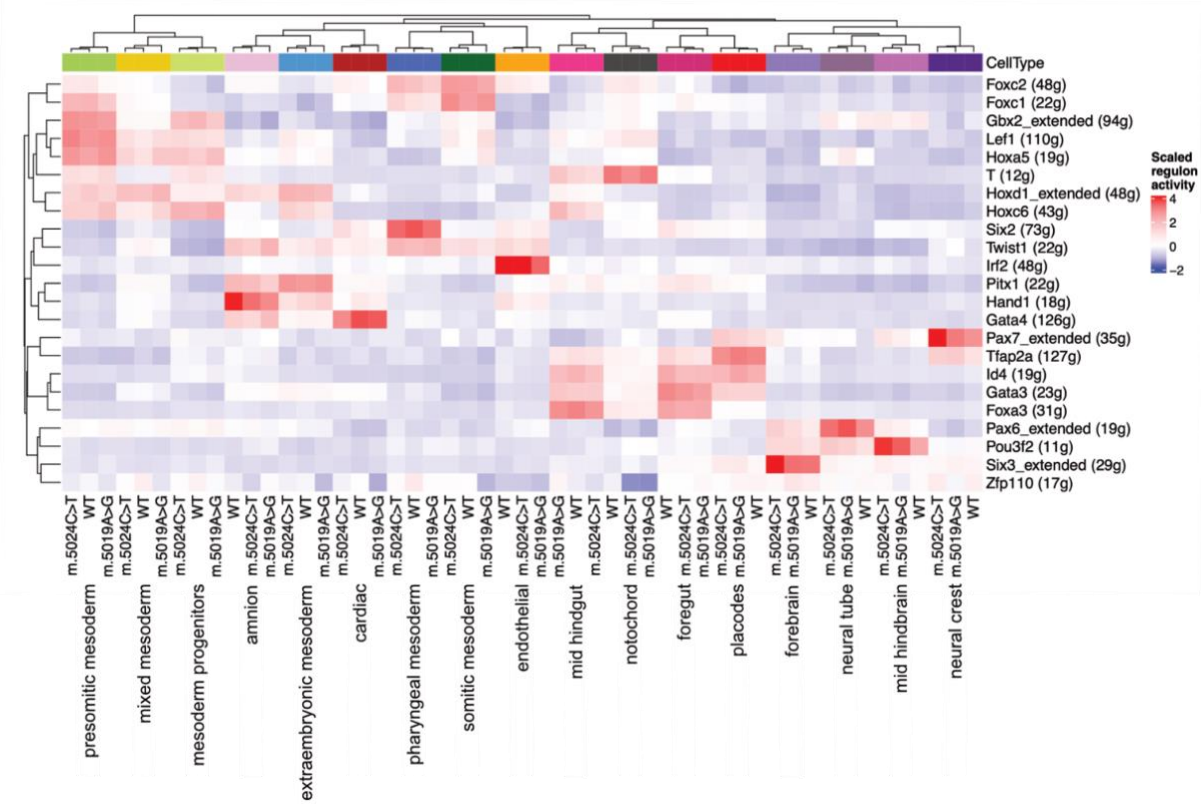


Figure 5.20. Activity of gene regulatory networks across three mouse strains.
Heatmap demonstrating the activity of selected single-cell regulatory network inference and clustering (SCENIC) regulons across all 17 cell lineages in wild-type (WT), m.5019A>G and m.5024C>T mouse embryos. The numbers in brackets next to each regulon followed by the letter (g) represent the number of genes identified with high confidence. Figure generated by Dr Malwina Prater.

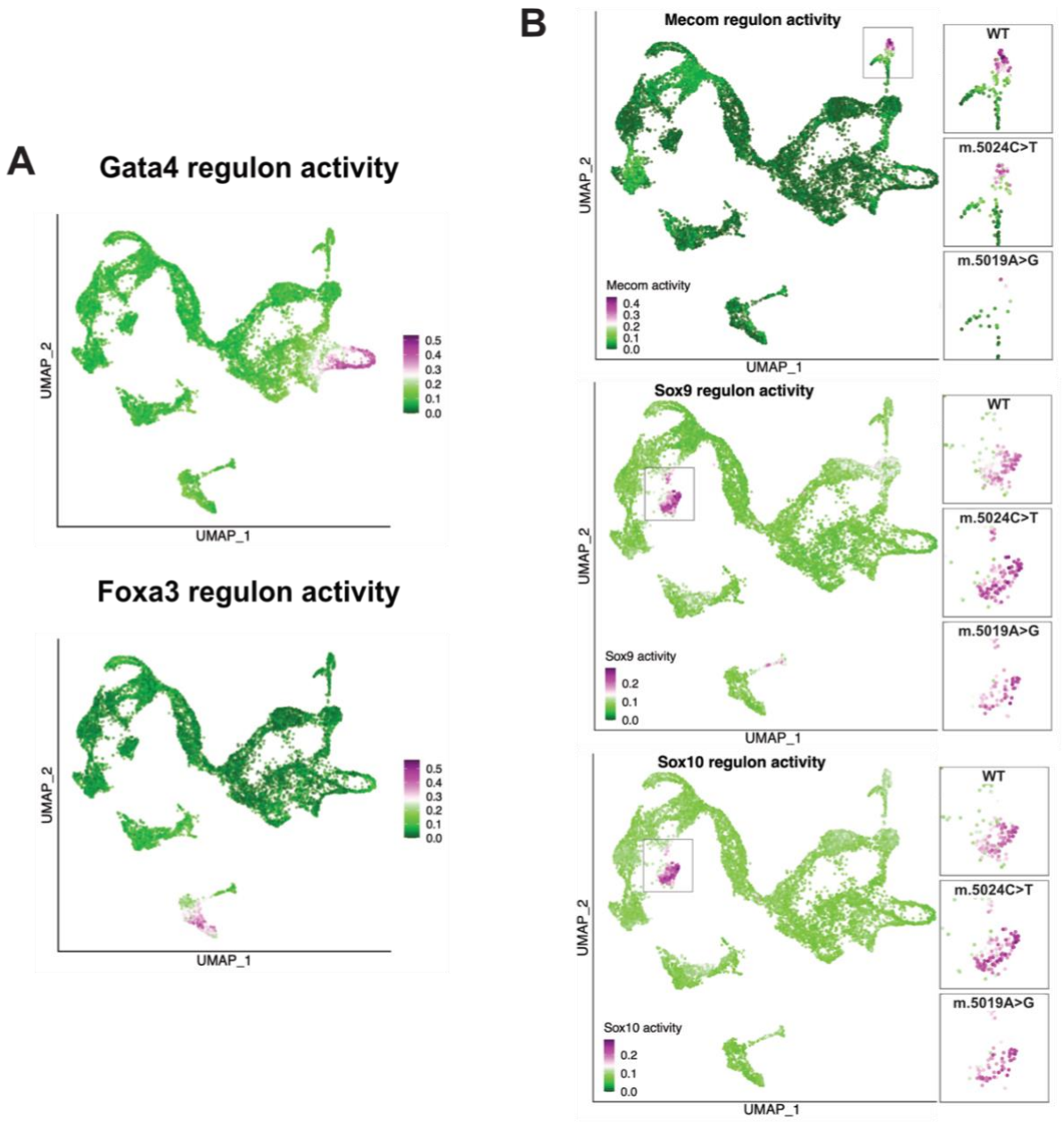


Figure 5.21. Activity patterns of select regulons.
(A) Uniform Manifold Approximation and Projection (UMAP) analyses showcasing the localisation of Gata4 and Foxa3 regulon activity. Each UMAP is an overlay of WT, m.5019A>G and m.5024C>T mouse embryos. Purple = Stronger regulon activity. **(B)** UMAP analyses showcasing the localisation of Mecom, Sox9 and Sox10 regulon activity. Each UMAP is an overlay of WT, m.5019A>G and m.5024C>T mouse embryos. Breakdown of the strains is displayed in boxes on the right-side of each UMAP. Purple = Stronger regulon activity. Figure generated by Dr Malwina Prater.

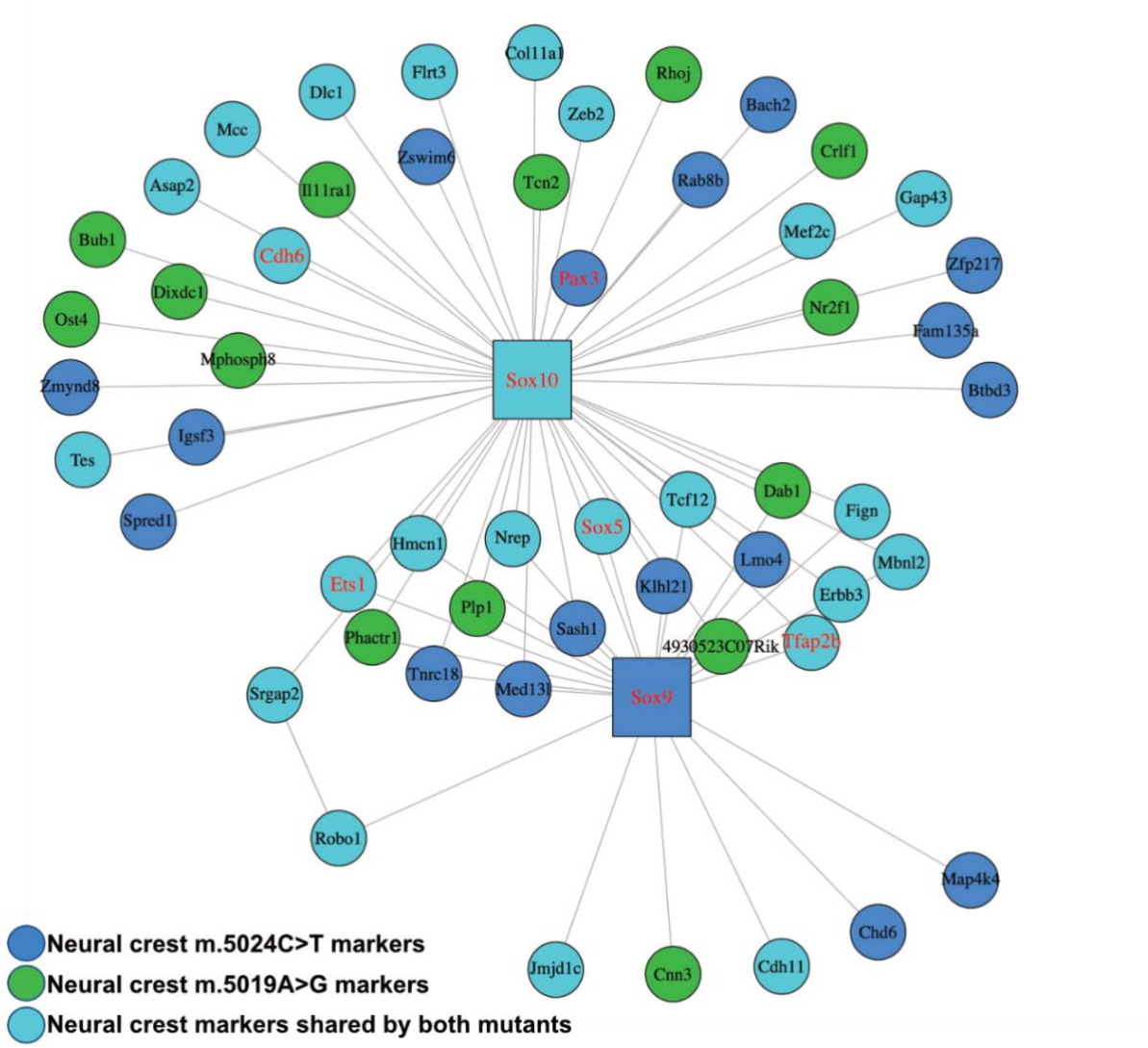


Figure 5.22. Differential expression of marker genes relating to select regulons.
Network consisting of differentially expressed marker genes that were identified as being part of the Sox9 and Sox10 regulons of the neural crest lineage in m.5019A>G and m.5024C>T compared to WT controls. Around half (24/56) of the maker genes could be identified in both mutant mouse strains. Lines connecting the circles to boxes represent regulons associated with each gene, while lines linking circles indicate protein-protein interactions. Red = genes involved in neural crest differentiation. Figure generated by Dr Malwina Prater.

5.10. Key Findings

- Having 4,924 single cells deriving from WT E8.5 embryos go through transcript quality control allowed us to confidently define the 3 germ layers (ectoderm, mesoderm and endoderm) in addition to 17 distinct tissue lineages (amnion, cardiac, endothelial, forebrain, foregut, mesoderm progenitors, mid hindbrain, mid hindgut, mixed mesoderm, neural crest, neural tube, notchcord, pharyngeal mesoderm, placodes, presomitic mesoderm, somitic mesoderm and extraembryonic mesoderm).
- There was a significant variation in mtDNA transcript abundance across different tissue lineages signalling the presence of different levels of mitochondrial activity across different tissues.
- 40% of nuclear-encoded mitochondrial genes also demonstrated differential transcript expression levels across tissue lineages, while the remaining 60% appeared to be uniformly expressed.
- Cardiac-specific isoforms of CIV subunits were found to be overexpressed in the cardiac cell lineage, suggesting that organ-specific transcriptional signatures precede tissue maturation.
- When comparing m.5019A>G and m.5024C>T E8.5 embryos to their WT counterparts, no obvious developmental abnormalities were observed. At the same time, there were no significant differences in the number of cells committed to each cell lineage in m.5019A>G and m.5024C>T mouse embryos compared to WT. This observation was coupled with no significant differences in the transcript levels of lineage-defining genes across the three mouse strains.
- Pseudo-bulk analysis of scRNA-seq data revealed a dysregulation in the transcript levels of nuclear-encoded mitochondrial genes.
- In response to the downregulation of nuclear encoded mitochondrial transcripts observed in embryos carrying the m.5019A>G and m.5024C>T mutations, an upregulation of mtDNA transcripts was detected.
- Out of the 191 unique genetic modifiers outlined by To et al. (To et al., 2019), 63 were found to be dysregulated in m.5024C>T embryos and 40 in m.5019A>G ones, with 22 detected in both strains.

- Epistatic suppressors/buffer genes identified by To et al. (To et al., 2019), were downregulated while genes that were lethal when knocked-out in the presence of an OXPHOS toxin were upregulated in the m.5019A>G and m.5024C>T mouse strains. In both cases acting as buffers, regulating the impact of the mt-Ta mutation.
- Epistatic suppressors/buffer genes form a tight network of interactions centred around Mrpl12 and Atp5d that includes proteins central to the activity of the OXPHOS machinery.
- Pseudobulk analysis revealed genes to be differentially expressed both at the level of the organism and at the level of the cell-lineage.
- Wikipathways analysis revealed that the transcript levels of genes coding for cytoplasmic ribosomal proteins, components of cell pluripotency and mRNA processing pathways were differentially expressed in m.5019A>G and m.5024C>T embryos.
- There is a stronger compensatory response being mounted by cells belonging to m.5024C>T embryos compared to their m.5019A>G counterparts when it comes to the level of downregulation of epistatic buffer and suppressor genes.
- Labelled proteomic analysis of liver samples derived from m.5019A>G mice, revealed significantly lower protein levels of subunits belonging to ETC complexes I, III and IV compared to controls.
- A wide range of epistatic buffer/suppressor genes displayed cell lineage-specific expression patterns.
- There is evidence for a cell lineage-specific activation of gene clusters that are jointly regulated (regulons).
- Many regulons that are differentially expressed across the 3 mouse strains are involved in tissue-specific developmental pathways eg. GATA4 in cardiac cell progenitors and Foxa3 in gut cell progenitors.
- Increased Sox9 and Sox10 regulon activity supports the hypothesis that mitochondrial activity is crucial for neural crest development, underlining the compensatory response seen in m5019A>G and m.5024C>T mutant strains that attempt to ensure that tissue differentiation will take place.

Chapter 6. Results:

The effect of Mitochondrial Activity and DNA Heteroplasmy on Cell Cycle Regulation

6.1. Introduction

A decrease in a cell's aerobic respiration in an environment where oxygen is abundant, is a behaviour that was first observed in cancer cells by the German scientist Otto Warburg (Warburg et al., 1927; Warburg, 1930). More recent work has revealed that a decrease in a cell's rate of aerobic respiration, accompanied by an increase of glycolysis is not a behaviour restricted to tumours but a metabolic change central to the process of cell proliferation (Gatenby and Gillies, 2004; Skeen et al., 2006). Studies placing cellular metabolic changes at the epicentre of cell division renewed our interest in the ways in which the mitochondrial network's dynamics and energy output can influence the cell cycle. Indeed, as we have touched on so far, mitochondrial morphology correlates with cellular responses to external stimuli. A fragmented mitochondrial network is often associated with apoptosis-inducing Cytc release while tubular mitochondrial morphology increases a cell's resistance to apoptotic signalling (Karbowski and Youle, 2003). During the course of the cell cycle, the oxidative capacity of the mitochondrial network shifts, with cells in late G_1-phase having a much greater respiratory capacity compared to their counterparts at early G_1-phase (Schieke et al., 2008). Work carried out on fruit flies (*D. melanogatster*) has revealed that a reduction in the mitochondrial network's ATP production can abolish the G_1-S-phase transition (Mandal et al., 2005). At the same time, cell cycle regulators that control the G_1-S-phase transition, such as p53 and cyclin D1, are known to have an effect on mitochondrial performance and morphology (Matoba et al., 2006; Wang et al., 2006), further implicating the mitochondrial network in the process of cell cycle regulation.

Kasturi Mitra and colleagues, identified the presence of a hyperfused mitochondrial network that was characterised by high ATP producing capacity and required for the G_1-S-phase transition (Mitra et al., 2009). At the same time, depolarisation of the mitochondrial network at this stage of the cell cycle, led to p53/p21-dependent G_1–S-phase transition arrest (Mitra et al., 2009), further underlining the importance of mitochondrial activity in cell cycle regulation. When serum-starved cells arrested at G_0-phase had their mitochondria made to form a hyperfused network, they experienced an increase in cyclin E levels that was followed by an escape from G_0-phase and development beyond the G_1-S-phase transition point (Mitra et al., 2009). The regulation of cyclin E levels by the hyperfused mitochondrial network might explain why Mitofusins 1 and 2 (MFN1/2) knock out cells mimic the behaviour of cyclin E

knockout cells by being unable to restart the cell cycle once they have entered quiescence in the G_0 state (Chen et al., 2003; Geng et al., 2003).

Another cell cycle stage where mitochondrial activity and dynamics are of utmost importance to the generation of healthy and functional daughter cells, is mitosis. During mitosis, in cases of symmetric cell division, the mitochondrial content present within the parent cell has to be evenly partitioned to the two daughter cells. This faithful mitochondrial segregation is possible thanks to the fission that characterises the mitochondrial network during late G_2-phase and M-phase, as well as thanks to the interaction between mitochondrial and cytoskeletal elements in addition to components of the ER that are spread all over the cytosol (Kanfer et al., 2015). Dynamin related protein 1 (Drp1) phosphorylation during M-phase leads to mitochondrial division and the creation of numerous mitochondrial units, each containing at least one nucleoid, that still interact with cytoskeletal components and the ER (Taguchi et al., 2007; Mitra et al., 2009). These interactions allow for mitochondria to be spread out throughout the cellular soma and be partitioned passively and equally to the two daughter cells through Miro-cytoskeletal-associated protein F (Cenp-F) interactions, as cytokinesis takes place (Kanfer et al., 2015). It is worth noting that the segregation of mitochondria in daughter cells can take place in the absence of Drp1-induced fission, as the cytokinetic forces are powerful enough to cleave the fused network into two (Ishihara et al., 2009; Wakabayashi et al., 2009). However, one should consider that forcibly dividing a fused mitochondrial network has been shown to lead to daughter cells inheriting a less uniform mitochondrial network (Ishihara et al., 2009).

Mitochondrial energetic output and network morphology are not the only aspects of the organelle actively interacting with the cell cycle. While for years, generation of new molecules of mtDNA has been thought of as following the principles of relaxed replication, i.e., dividing independently from the nDNA (Chinnery and Samuels, 1999), more recently this concept has been challenged. Laurent Chatre and Miria Ricchetti devised what they termed the mitochondrial Transcription and Replication Imaging Protocol, a novel in situ hybridisation protocol that makes use of fluorescence imaging to detect mtDNA molecules that are in the process of replication initiation (Chatre and Ricchetti, 2013). They observed that both the processes of mtDNA transcription and replication initiation are coordinated with the cell cycle resulting in mtDNA synthesis preceding that of nDNA before being eventually reactivated at the end of S-phase (Chatre and Ricchetti, 2013). As a result, while the concept of relaxed replication is still applicable, as replication of the mtDNA takes place throughout the cell cycle,

a distinct minority of mtDNA structures appears to exhibit increased transcriptional and replicative activities that are coordinated with the synthesis of nDNA (Chatre and Ricchetti, 2013).

Taking into consideration the central role that mitochondria have in energy production and the energy-demanding nature of the cell cycle, it does not come as a surprise that such a core aspect to cellular life is heavily dependent on mitochondrial performance. The cell's capacity to undergo metabolic adaptations, as well as the mitochondrial network's dynamic nature are both factors influencing cell cycle advancement. However, our understanding regarding the pathophysiological aspect of the mitochondria's influence on the cell cycle, remains incomplete.

6.2. Project Aims

The cell cycle is a multi-stage process responsible for driving cell growth and division thus, it is of critical importance to developing tissues. A major priority for any dividing cell is to make sure that the genetic material passed down to the next generation of cells is a faithful copy of that which it possesses. As a result, there are multiple checkpoints through the cell cycle that arrest development when nDNA damage is detected in order to give the cell the time required to repair it (Campos and Clemente-Blanco, 2020). At the same time, it is imperative for the cell to ensure that before committing down the cell cycle path, it has enough resources to see the journey through. Consequently, a regulatory point exists within G_1 where the cell is able to arrest development and be placed at G_0 in the off chance that the required amount of nutrients is not available (Cai et al., 2013). Studying the tightly regulated transitions through the cell cycle phases has for many years depended on chemical components that when introduced into the cell lead to its arrest at a specific stage. Artificially arresting cells at a specific cell cycle stage has been historically used in the study of nDNA damage repair (Singh and Xu, 2016) and the role of certain proteins *in vivo* (Mendenhall et al., 1988; Uchiyama et al., 1997; al-Khodairy and Carr, 1992; Hwang and Murray, 1997). In all of these cases, cell synchronisation takes place as no matter how far along each cell is in the cell cycle, once the drug is applied, all cells will eventually be halted at the exact same cell cycle stage. Cell cycle arrest at G_1-phase can be implemented through the application of alpha factor (α-factor)-induced inhibition of adenylate cyclase (Liao and Thorner, 1980) or though the addition of the plant amino acid mimosine (Park et al., 2012). Mimosine, a zinc and iron chelator, leads to the reduction of intracellular iron levels, which in turn induces double strand breaks (DSBs) that inhibit nDNA replication (Park et al., 2012). At the same time, it has been suggested to bring about the inhibition of iron/zinc-dependent enzymes such as ribonucleotide reductase and serine hydroxymethyltransferase (Perry et al., 2005). Serum deprivation is another technique that can result in cellular arrest prior to entering the S-phase (Howard et al., 1993). Additionally, many cell types can be arrested within G_1-phase as a result of coming in contact with neighbouring cells. This is termed contact inhibition and is thought to be dependent on $p27^{Kip1}$, a cyclin-dependent kinase inhibitor (Li et al., 2000). Cell cycle arrest can also take place during the S-phase, an evolutionarily conserved and tightly regulated cell cycle stage that involves the duplication of the cell's nDNA. The antibiotic aphidicolin isolated from the fungus *Cephalosporium aphidicola* leads to reversible cell cycle arrest during the S-phase through the

inhibition of DNA polymerases A and D (Baranovskiy et al., 2014). The inhibition of ribonucleotide reductase (RNR) by the small molecule Hydroxyurea (HU) prevents the formation of ribonucleotides from deoxyribonucleotides (DNTs) (Koc et al., 2004; Xu et al., 2016). Moreover, p21 overexpression and S-phase arrest can be induced by the small molecule 2[[3-(2,3-dichlorophenoxy) propyl] amino] ethanol (2,3-DCPE) (Zhu et al., 2004). When it comes to arresting the cell at G_2, a phase characterised by rapid growth and increased levels of protein synthesis that key to the cell's progression into mitosis, cell cycle arrest can be achieved by the neutralisation of cyclin mRNA. One such method is the use of antisense oligonucleotides that bind the target mRNA and prevent its translation into protein (Saikawa et al., 2001). Mitosis is the last stage of the cell cycle, where the mitotic spindle is responsible for separating sister chromatids towards the two poles of the cell, a step immediately followed by cytokinesis. As with all previous cell cycle steps discussed here, there are many ways in which a cell can be prevented from undergoing mitosis. The chemical agent nocodazole prevents cells from entering mitosis by disrupting microtubule polymerisation (Kuhn, 1998). Conversely, taxol promotes microtubule stabilisation, preventing their degradation and cause M-phase arrest as the mitotic spindle cannot be disassembled (Choi and Yoo, 2012; Ikui et al., 2005). Finally, in HeLa cells there is evidence that cell cycle progression is temperature-sensitive and can thus be regulated through the introduction of temperature shifts (Rao and Engelberg, 1965). More specifically, cytokinesis was shown to be the most temperature-sensitive stage of the cell cycle and was halted in HeLa cells once temperatures dropped between 24-31^{o}C (Rao and Engelberg, 1965).

Many of these techniques outlined above have been used for years to study mechanisms involved in cell cycle progression. However, cell cycle arrest and synchronisation through such invasive procedures comes with strong perturbations of numerous intracellular processes and increases in cellular stress that many cases can lead cell death (Krek and DeCaprio, 1995). As a result, new methods have been devised as of late to monitor cell cycle stage. One such technique was published by Asako Sakaue-Sawano and her colleagues in 2004, with the introduction of the fluorescent ubiquitination-based cell cycle indicator (FUCCI) system (Sakaue-Sawano et al., 2008). Until the publication of the FUCCI system, transition from M-phase to G_1-phase could be easily monitored by following morphological changes that characterise cell division. The morphologically subtle G_1- to S-phase transition however, could only be monitored by employing either bromodeoxyuridine (BrdU) staining of the cell nucleus or one of the cell synchronisation techniques outline above. FUCCI development provides a

tool that overcomes the need for pharmacological intervention as it presents a way of monitoring the cell cycle through shifts in the expression of nuclear-localised fluorescent proteins (Sakaue-Sawano et al., 2008). The use of genetically-encoded fluorescent probes resulted in labelling nuclei of cells that were going through G_1-phase red, while ones that were at either S-, G_2- or M-phase green (Sakaue-Sawano et al., 2008). Nuclear-specific fluorescence allows for the parallel monitoring of other forms of signalling and changes that take place in the cytoplasm as the cell cycle progresses. Gavin Grant and colleagues took the FUCCI system a step further in 2018 as they sought to enhance its capability to detect transitions between cell cycle phases (Grant et al., 2018). To that effect they created PIP-FUCCI, which efficiently marks the G_1-S-phase and S-G_2-phase transitions thanks to the proliferrating cell nuclear antigen (PCNA)-Interacting Protein (PIP) degron fusion protein (Grant et al., 2018) that directly couples the oscilation of reporter expression to DNA replication, as a result of the PCNA reporter activity. This new reporter system was proven to accurately identify phenotypes associated with cell cycle arrest while at the same time being resistant to moderate levels of nDNA damage (Grant et al., 2018). Overall, Gavin Grant and colleagues provide a comprehensive new tool based on the widely used FUCCI construct that allows for clear identification of the cell's movement into and out of S-phase **(Figure 6.1.)**.

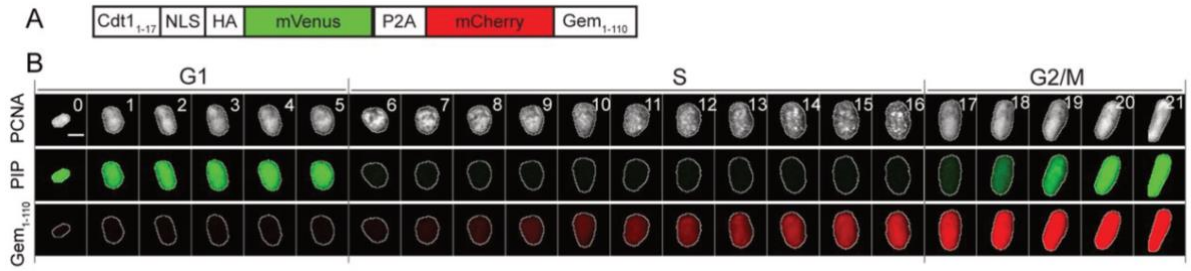


Figure 6.1. PIP-FUCCI, a novel cell cycle indicator.
(A) Schematic of the PIP-FUCCI dual reporter expression construct. $Cdt1_{1–17}$ = human Cdt1 amino acids 1-17 including the PIP degron, NLS = SV40 nuclear localization signal, HA = epitope tag, P2A = self-cleaving peptide, Gem_{1-110} = human Geminin amino acids 1-110 including both the D box and KEN motif. **(B)** Selected images from wide field time-lapse imaging of a U2OS cell expressing mTurq2-PCNA and the PIP-FUCCI reporters PIP-mVenus and mCherry-Gem_{1-110}. One frame from each hour beginning after cytokinesis is shown with the nucleus outlined; numbers in the upper strip indicate hours since mitosis. Scale bar 10 µm. Adapted from: Grant et al., 2018.

In this chapter, I present our work so far as we make use of the PIP-FUCCI construct in order to determine the ways in which changes in mitochondrial activity influence cell cycle progression. I first take a look at the mtDNA copy number during different cell cycle stages in HeLa cells. Furthermore, I explore the effect different culture conditions of glucose availability have on copy number levels and cell cycle progression of PIP-FUCCI HeLa cells. Furthermore, I establish the impact a drastic mtDNA copy number reduction has on PIP-FUCCI HeLa cell cycle progression. Finally, I examine the impact of heteroplasmic pathogenic mtDNA mutations have on cell cycle progression, within the context of ΔH2.1 deletion cybrids. The project I am presenting in this chapter was carried out in close collaboration with Dr Stephen Burr.

6.3. Mitochondrial DNA Copy Number Increases as the Cell Cycle Progresses

Our first step in studying the relationship between mitochondrial DNA and cell cycle progression was to examine fluctuations in mitochondrial DNA copy number through the different stages of the cell cycle. Previous work carried out on FUCCI-HeLa cells by Taeko Sasaki and colleagues uses quantification of fluorescence intensity to suggest that mtDNA replication takes place throughout the cell cycle, with a distinct peak of activity during S-phase (Chatre and Ricchetti, 2013). We decided to study this relationship using HeLa cells that are expressing the PIP-FUCCI construct and monitor mtDNA copy number during the different cell cycle stages. Sub-cloning the initial cell population to which the PIP-FUCCI construct was inserted, we obtained clones with high levels of PIP-FUCCI expression, where cell cycle stages were readily distinguishable using FACS **(Figure 6.2.A)**. At the same time, we confirmed that cells with higher mtDNA copy number, that was a consequence of either larger cell size or just a product of the cell being in a later stage of the cell cycle, had a greater SSC-A **(Figure 6.2.B)**, a parameter that indirectly measures granularity. While a greater mitochondrial network will contribute to a more granular cell, it is not the only factor in play. It is expected that cells approaching the later stages of the cell cycle will contain increased numbers of most organelles in preparation for mitosis. Introducing the PIP-FUCCI construct into both HeLa and 293T cells and sorting them based on cell cycle stage, we observed a very similar behaviour, with progressive cell cycle stages being characterised by increases in mtDNA copy number (**Figure 6.2.C** and **D**). This observation is partly in line with what was described by Taeko Sasaki and colleagues in terms of mtDNA replication taking place throughout the course of the cell cycle. However, we did not observe a particular spike in mtDNA copy number as the cell entered or exited the S-phase. In an attempt to further investigate that point, we decided to go back to the FACS analysis of PIP-FUCCI HeLa cells and try to further subdivide cell cycle stages based on the fluorescence intensity of the construct. As demonstrated in **Figure 6.1.**, cells at G_1-phase, uniformly exhibit mVenus fluorescence, however, as they enter S-phase there is an abrupt loss of mVenus and a gradual gain of mCherry fluorescence. This allows us to subdivide S-phase into three sub-sections based on mCherry fluorescence intensity **(Figure 6.2.E)**. The same principle is applied to the G_2/M-phase step but in this case, it is the mVenus fluorescence that gradually resurfaces and it is used to subdivide the G_2/M-phase into two subsections. With these subdivisions in mind, we performed another analysis in order to detect any possible spikes

in mtDNA copy number along the cell cycle. In this case, while cells in the G_1-phase were still grouped together, ones in S-phase were divided into 3 subgroups while others that were in the G_2/M phase were subdivided into 2 subgroups **(Figure 6.2F)**. The result was similar to the one seen in **Figures 6.2C** and **D**, with no peak in mtDNA copy number present. Instead, a more gradual increase in mean copy number was observed as the cell cycle progressed.

Having established the stepwise nature of mtDNA increase through the cell cycle, we examined whether depriving PIP-FUCCI HeLa cell of glucose would result in long-term changes of their mtDNA copy number. Indeed, even though cells grown in low glucose behaved similarly to the ones grown under normal condition **(Figure 6.4.A** and **B)**, ones grown in galactose exhibited a significant increase in mtDNA copy number after week 2 **(Figure 6.4.C)**. Finally, we sought to examine the progression of PIP-FUCCI HeLa cells through the cell cycle. To that effect we decided to conduct live cell analysis using Incucyte® (Sartorius) **(Figure 6.3.)**. This technique allowed us to compare the cell cycle duration between PIP-FUCCI HeLa cells grown in media with high glucose concentration to ones grown in media with low glucose concentration **(Figure 6.4.B)**. Having seen that the adaptations cells made when placed in a low-glucose environment were minimal, we did not expect to see any radical changes when it came to their cell cycle duration. However, when the same population of PIP-FUCCI HeLa cells was tested in the incucyte, we detected a significant increase in the time required for cells grown in low-glucose to complete G_1, when compared to their counterparts grown in high-glucose media **(Figure 6.4.D)**

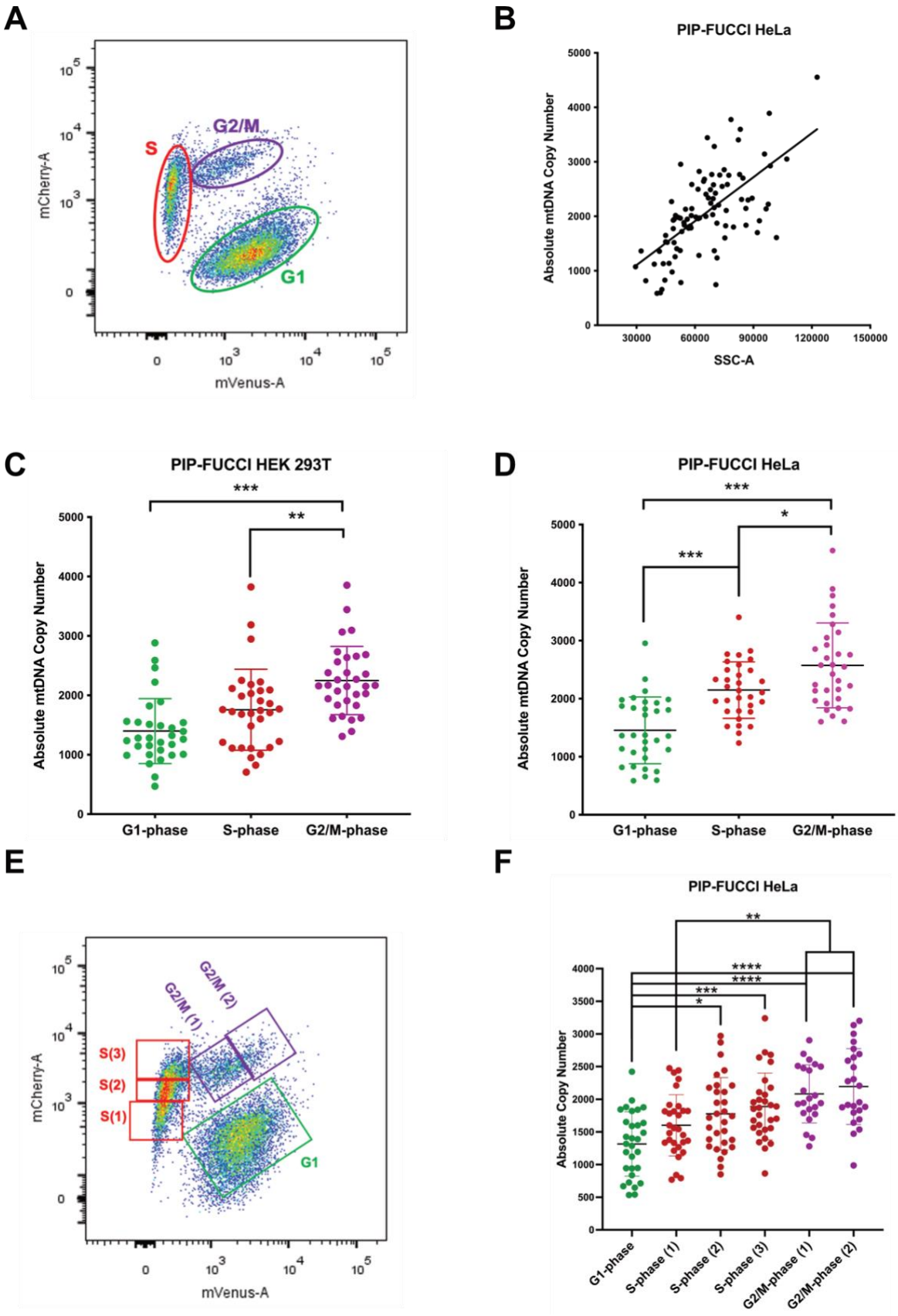

A
S
G2/M
G1
mCherry-A
mVenus-A
B
PIP-FUCCI HeLa
Absolute mtDNA Copy Number
SSC-A
C
PIP-FUCCI HEK 293T
Absolute mtDNA Copy Number
G1-phase
S-phase
G2/M-phase
D
PIP-FUCCI HeLa
Absolute mtDNA Copy Number
G1-phase
S-phase
G2/M-phase
E
G2/M (2)
G2/M (1)
S(3)
S(2)
S(1)
G1
mCherry-A
mVenus-A
F
PIP-FUCCI HeLa
Absolute Copy Number
G1-phase
S-phase (1)
S-phase (2)
S-phase (3)
G2/M-phase (1)
G2/M-phase (2)

Figure 6.2. Mitochondrial DNA copy number changes along the cell cycle.
Cells expressing the PIP-FUCCI reported were isolated at different stages of the cell cycle through fluorescence activated cell sorting (FACS). **(A)** FACS analysis of PIP-FUCCI HeLa. Cells at G_1-phase appear positive for mVenus, while cells at S-phase appear positive for mCherry. Finally, cells at the G_2/M-phase are positive for both mVenus and mCherry. **(B)** The side scatter (SSC) of events sorted using FACS and the copy number of those same events are positively correlated. **(C** and **D)** Both HeLa and 293T cells expressing the PIP-FUCCI reporter show a gradual increase in copy number in as they progress from G_1- to S-phase and from S- to G_2/M-phase. **(E)** Further subdividing the cell cycle phases based on the fluorescence intensity introduced by the PIP-FUCCI construct. **(F)** Copy number measurements in single cells derived from the subdivision of cell cycle steps presented in **(E)**. Statistical analysis involved unpaired t-tests followed by Bonferroni multiple test correction.

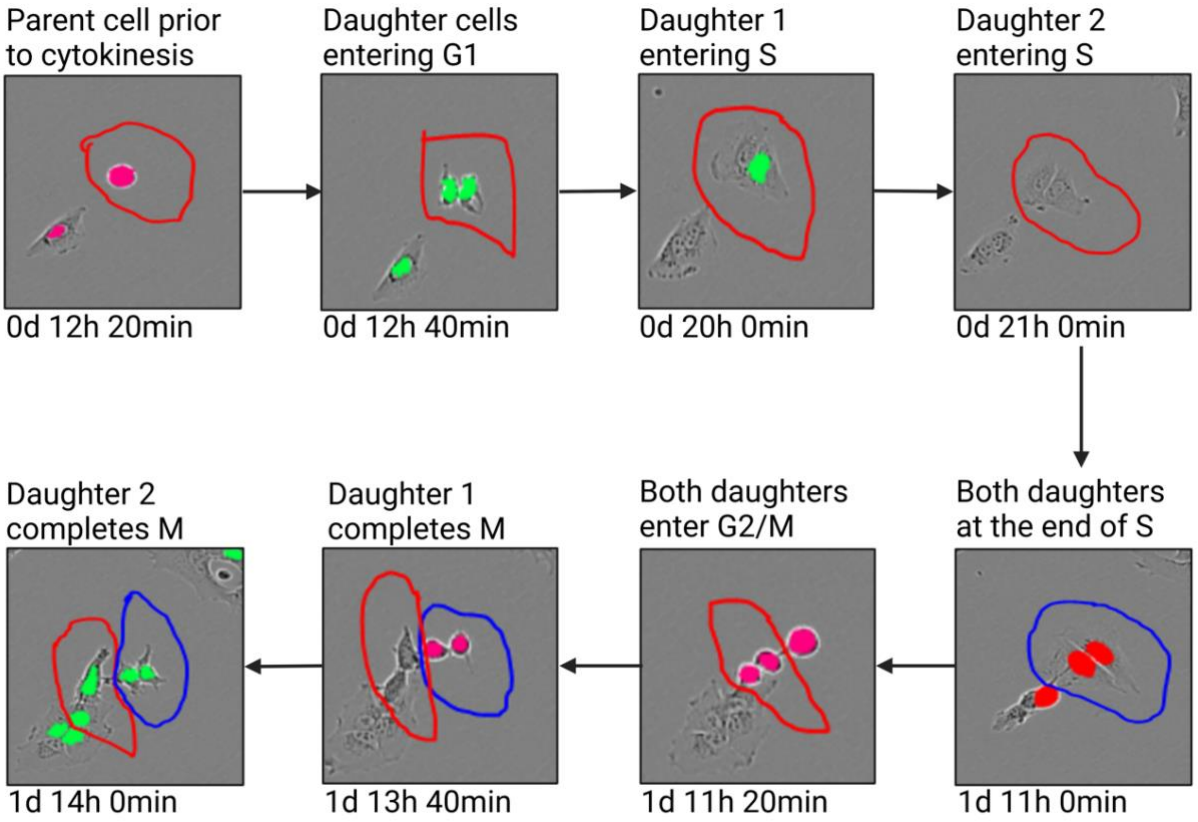


Figure 6.3. Manual tracing of the cell cycle during live cell analysis using the PIP-FUCCI reporter.
Starting from the **Top-Left** picture, we initially spot the parent cell at late G_2/M, immediately prior to cytokinesis. The first instance that the daughter cells are visibly divided as they both enter G_1 (green) at the same time. monitoring the now independent progression of the two daughter cells out of G_1 and into the S-phase (lack of green and progressive accumulation of red). As the two daughter cells progress through the S-phase, entering G_2/M, the green colour reappears alongside red (red-green coexistence denoted as pink). As the two cells divide, their respective daughter cells will enter G_1 (green) and the cell cycle start anew. Green- cells at G_1 and G_2/M; Red- progressively in cells going through S and in G_2/M; Pink- denotes coexistence of red and green. Created with BioRender.com.

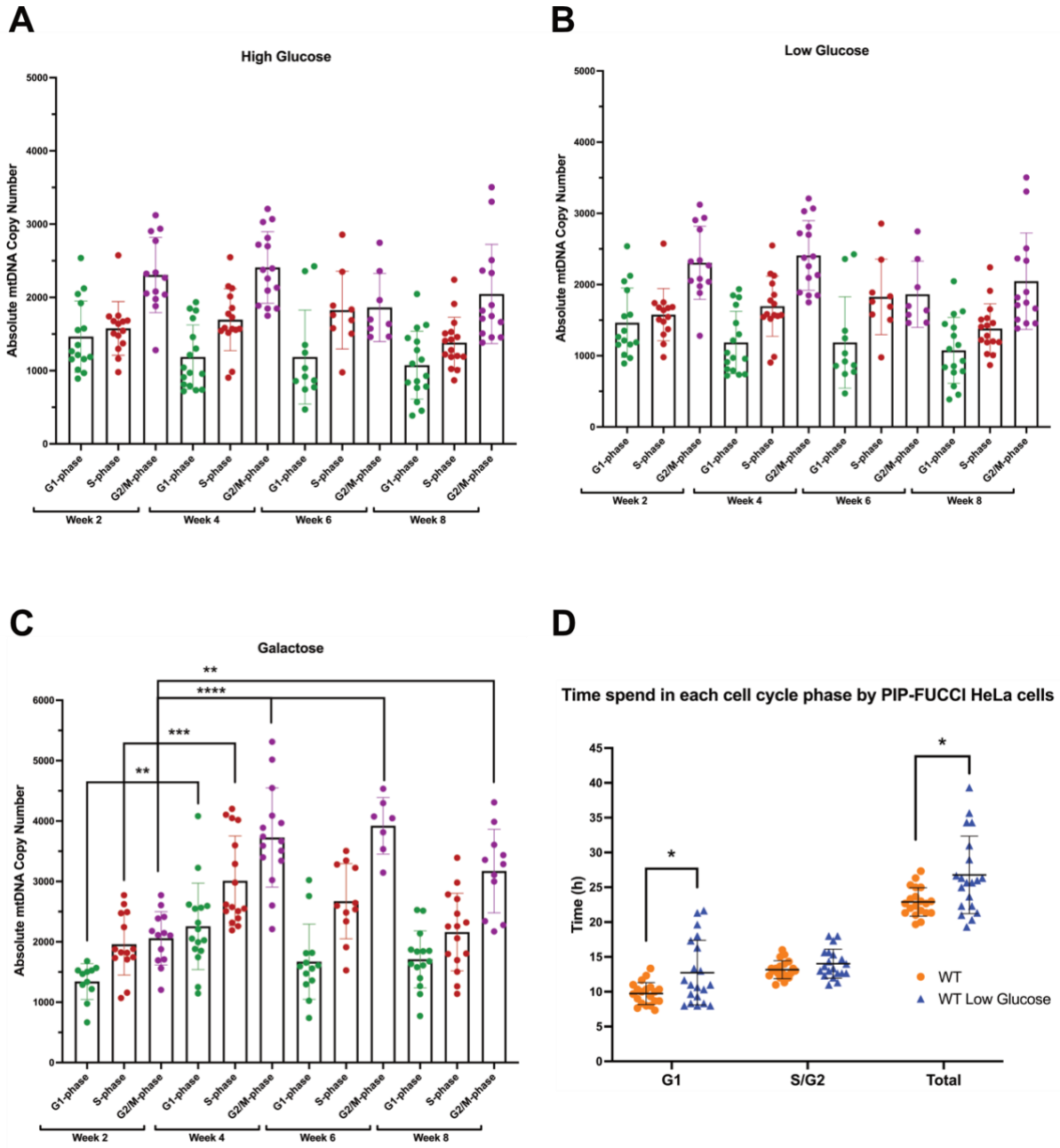


Figure 6.4. The effect of glucose deprivation on mitochondrial DNA copy number during the cell cycle.
(A-C) Mitochondrial DNA copy number measurements performed on PIP-FUCCI HeLa cells during different stages of the cell cycle over the course of 8 weeks. Cells were either cultured under standard conditions, with growth media containing low glucose or with media containing only galactose. **(D)** Time in hours (h) required by cells grown in high and low glucose media to complete a cell cycle. Statistical analysis involved unpaired t-tests followed by Bonferroni multiple test correction.

6.4. The Effect of Mitochondrial DNA Depletion on Cell Cycle Progression

After establishing the progressive increase of mtDNA copy number along the cell cycle as well as the impact glucose deprivation has on mtDNA copy number and cell cycle progression, we decided to investigate whether mtDNA depletion has an effect on cell cycle progression. To that effect, we decided to knock-out (KO) TFAM in HeLa cells carrying the PIP-FUCCI construct and observe the effect this would have on mtDNA copy number. Fist we confirmed that successful TFAM-KO by both measuring the protein levels through western blotting and by examining the copy numbers of the same clones through ddPCR (**Figure 6.5.A** and **B**). Having established that the extent of TFAM-KO was sufficient, we went ahead and compared the copy number of PIP-FUCCI HeLa cells with and without the KO **(Figure 6.5.C)**. The result, was a significantly lower copy number at every stage of the cell cycle from the cells where the TFAM-KO was applied. At the same time, the stepwise increase in mtDNA copy number as the cell cycle progressed was abolished in cell were TFAM was knocked-out. At this point, we decided to examine whether the drastic reduction in mtDNA copy number we introduced with the TFAM-KO had any effect in cell cycle progression. As demonstrated by **Figure 6.5.D**, cell cycle progression of PIP-FUCCI HeLa cells was significantly delayed as a result of mtDNA depletion.

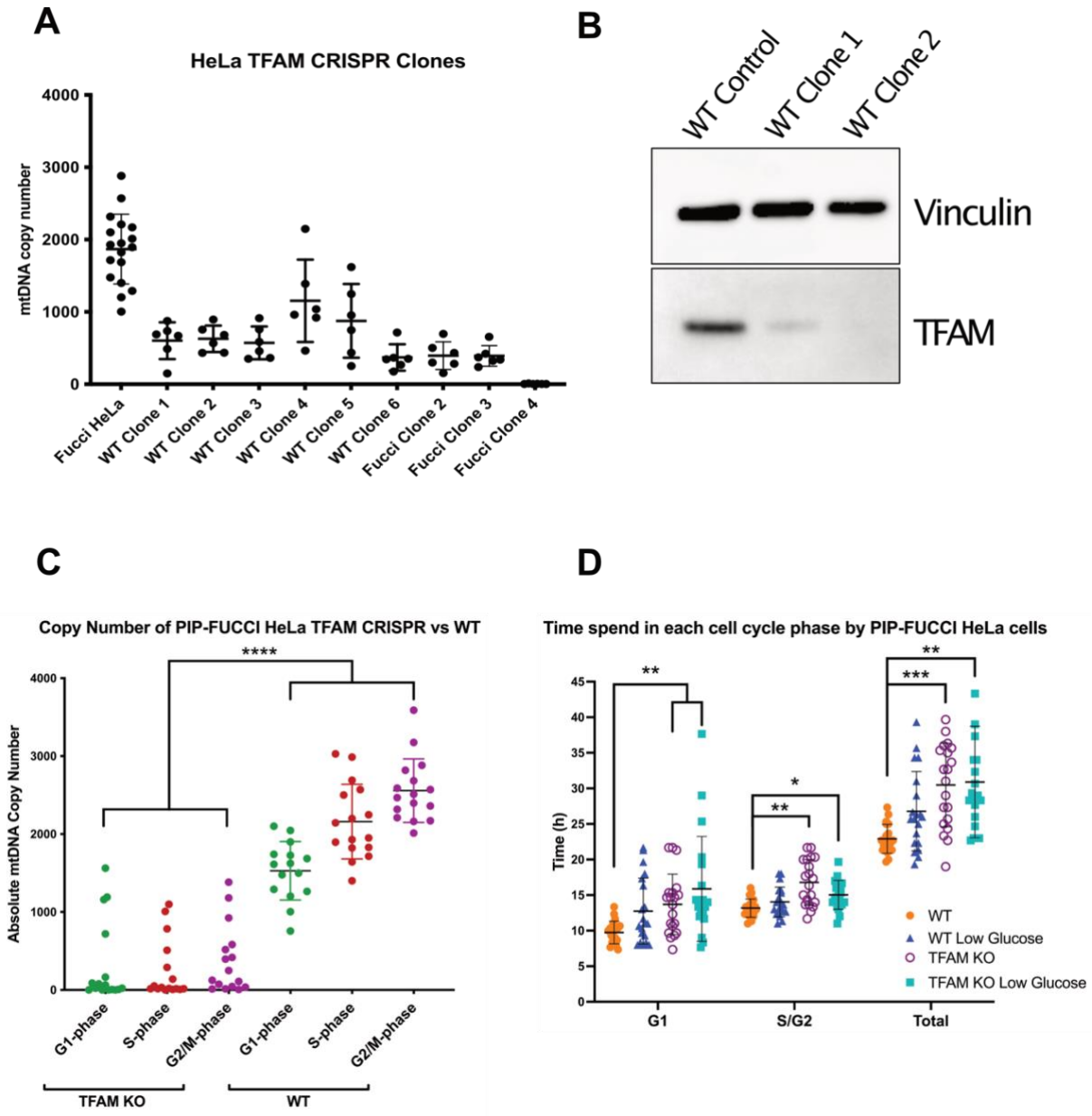


Figure 6.5. The effect of mitochondrial DNA copy number depletion on cell cycle progression. **(A)** Copy number measurements in PIP-FUCCI HeLa clones that have been subjected to TFAM CRISPR knock-out. **(B)** Western blot results showing TFAM protein levels in WT and CRISPR knock-out cell lines. **(C)** Copy number measurements from WT Clone 2 of the PIP-FUCCI HeLa cells after 10 days of standard culture with and without the TFAM CRISPR knock-out, at different stages of the cell cycle. **(D)** Time in hours (h) spent by PIP-FUCCI HeLa cells, with and without the TFAM CRISPR knock-out, at each cell cycle stage. Statistical analysis involved unpaired t-tests followed by Bonferroni multiple test correction.

6.5. Under Standard Culture Conditions, Pathogenic Mitochondrial DNA Mutations have no Effect on Cell Cycle Progression

Having established that HeLa cell cycle progression is significantly impaired as a result of mtDNA depletion, we decided to delve deeper into that relationship by making use of a different cell model. To that effect, we employed the ΔH2.1 cybrid cell line that is heteroplasmic for a large scale mtDNA deletion that encompasses around half the molecule **(Figure 6.6.A)**. Our first aim was to estimate the heteroplasmy level of this mutation in the ΔH2.1 cell line, alongside the mtDNA copy number (**Figure 6.6.B** and **C**). Having an original population with ~65% heteroplasmy allowed us to perform sub-cloning experiments and create cell lines that spanned the entire spectrum, from WT homoplasmic to ~90% heteroplasmy **(Figure 6.6.D)**. At this stage, we decided to investigate what the impact of such a large-scale deletion is on the ability of cells to carry out OXPHOS. In addressing this question, clones ranging from 8.2% to 80% heteroplasmy were analysed using high-resolution respirometry **(Figure 6.7.A)**. It is evident from this analysis that the effect of the mutation is minimal when present between 8.2% and 40%. However, in line with the threshold effect governing the expression of all mtDNA mutations, the 55% heteroplasmy clone appears to have strikingly lower baseline respiration, ATP-linked respiration and maximal respiratory capacity compared to the 40% heteroplasmy one. This suggests that the heteroplasmy threshold for the large-scale deletion carried by the ΔH2.1 cell line lies somewhere between 40% and 55%. Clones that exceeded 55% heteroplasmy continued the downwards trend with the clone having 80% heteroplasmy demonstrating extremely limited ability to respire. Having established the severity of the large-scale deletion carried out by the ΔH2.1 cell line and the ways in which it affects OXPHOS when present at different heteroplasmy levels, we decided to examine whether clones with different mutations levels take longer to complete the cell cycle **(Figure 6.7.B)**. Incucyte data suggested that, despite the huge differences in mutational load between the clones and the established effect that the mutation has on OXPHOS activity, there is no trend that points towards cell cycle progression being influenced by heteroplasmy levels. In order to further investigate this observation, we decided to compare the cell cycle progression of ΔH2.1 high and low heteroplasmy clones grown in high glucose media to ones grown in galactose **(Figure 6.8.)**. The result was the same as glucose availability did not appear to have a disproportional impact on the high-heteroplasmy clone.

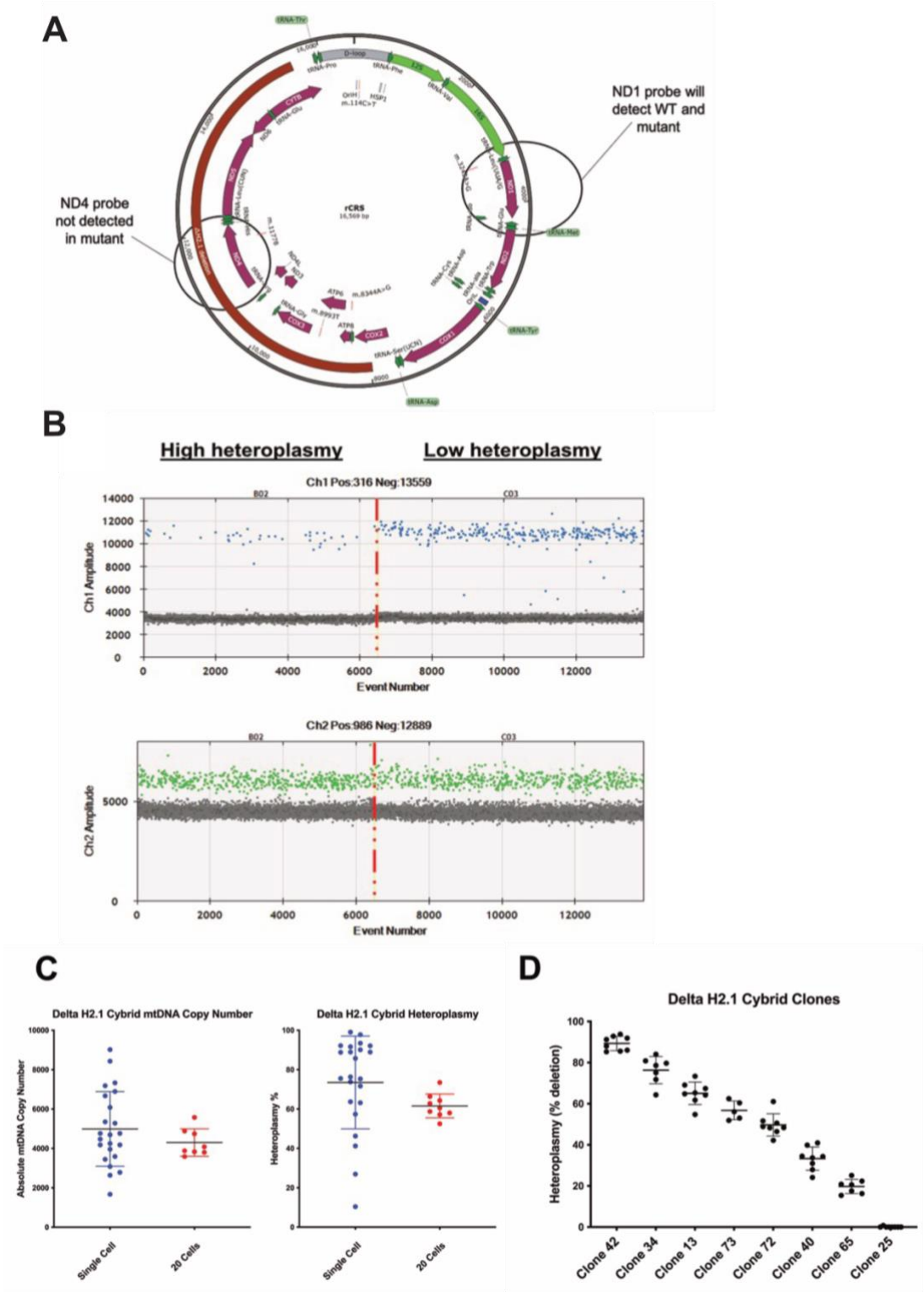


Figure 6.6. Generating ΔH2.1 cybrid populations with a broad range of mtDNA heteroplasmy values.

A) Schematic representation of the human mitochondrial DNA, denoting the region being deleted in the ΔH2.1 cybrid cell line. The position and fate of ND1 and ND4 primers is also highlighted. **(B and C)** Measuring copy number and heteroplasmy in the ΔH2.1 cybrid cell line using the ND1 and ND4 primers respectively. **(D)** ΔH2.1 cybrid cell lines of different mitochondrial DNA (mtDNA) heteroplasmy levels, generated by sub-cloning the original population.

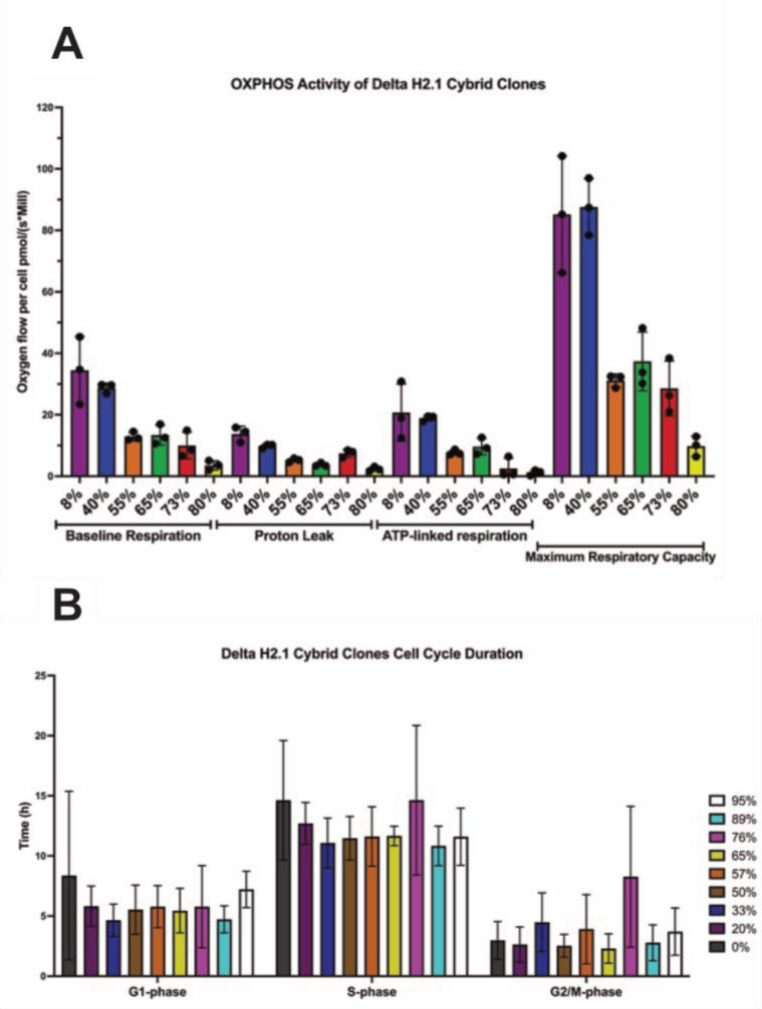


Figure 6.7. The effect of mitochondrial DNA heteroplasmy on cell cycle progression.
(A) High resolution respirometric analysis measuring basal respiration, proton leak, ATP-linked respiration and maximum respiratory capacity in ΔH2.1 clones of varying heteroplasmy. **(B)** Time in hours (h) spend by ΔH2.1 clones of varying heteroplasmy percentages at different stages of the cell cycle. Statistical analysis involved unpaired t-tests followed by Bonferroni multiple test correction.

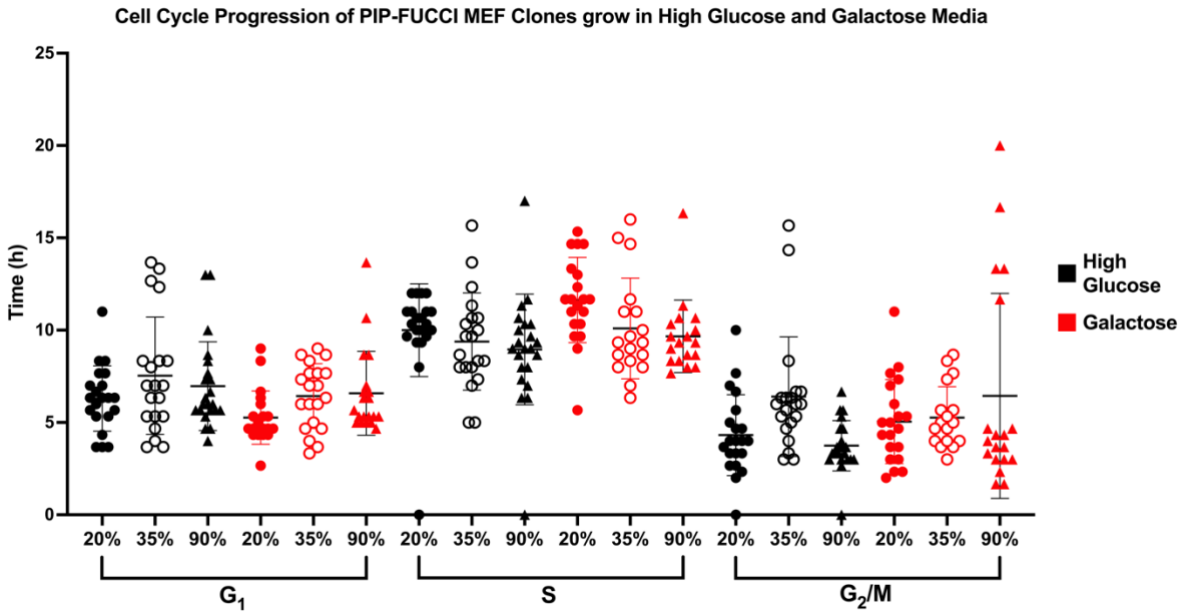


Figure 6.8. Glucose availability has no effect in the cell cycle progression of ΔH2.1 cybrids. Time (h) spent by three ΔH2.1 cybrid clones of different heteroplasmy (20%, 35% and 90%) at each stage of the cell cycle (G_1, S, G_2/M) while grown in high-glucose and galactose media. Statistical analysis involved unpaired t-tests followed by Bonferroni multiple test correction.

6.5. Key Findings

- As cell mtDNA copy number increases, so does the cell's granularity (estimated on the basis of SSC-A). This can be explained as a reflection of an increase in overall mitochondrial content.
- There is a clear increase in mtDNA copy number as cells progress from G_1-phase to S-phase and eventually into the G_2/M-phase.
- Using the progressive increase in mVenus fluorescence provided by the PIP-FUCCI reporter, we are able to subdivide S-phase and G_2/M-phase into shorter sub-phases, revealing a more gradual increase in mtDNA copy number.
- Forcing cells to perform OXPHOS by culturing them in galactose-containing media, leads to a universal mtDNA copy number increase at all 3 cell cycle stages by week 4.
- Restricting the ability of cells to perform glycolysis by culturing them in low glucose has a significant impact on the time required to complete the cell cycle.
- TFAM CRISPR KO leads to a drastic reduction in mtDNA copy number at every stage of the cell cycle.
- TFAM CRISPR KO leads to a significant increase in cell cycle duration when the cells are encouraged to perform more OXPHOS by being cultured in media containing low levels of glucose.
- The higher the heteroplasmy level of ΔH2.1 deletion cybrids is, the greater loss of their capacity to perform OXPHOS. However, this relationship is not linear. Instead, a primary heteroplasmy threshold is observed between 40% and 55% and a secondary one between 74% and 83%, significant drops in maximum respiratory capacity are observed between clones possessing these heteroplasmy levels.
- Cell cycle progression in ΔH2.1 deletion cybrids was unaffected by increases in mtDNA heteroplasmy, an observation further validated in cells cultured in both high-glucose or galactose media.

Chapter 7. Discussion and Future Trends

The concept of heteroplasmy lies at the epicentre of every pathology that involves mutations of the mtDNA, from rare multisystem mitochondrial diseases that affect ~1 in 5000 individuals, to common multi-factorial conditions that bear the hallmark of mitochondrial dysfunction. Consequently, understanding the pathways involved in nDNA response to mtDNA perturbations as well as the principles that govern heteroplasmy dynamics is of utmost importance. For many years, study of those principles has been hampered due to the lack of informative heteroplasmic animal models of mitochondrial disease, as well as our inability to conduct in-depth single cell evaluation of the affected tissues. This is the reason why in **Results Chapter 3**, we sought to establish a reliable protocol for the high-throughput measurement of mtDNA heteroplasmy at the single-cell level. More specifically, we aimed to optimize a protocol for the analysis of samples obtained from two novel mouse models of mitochondrial disease, each carrying either the m.5019A>G or the m.5024C>T mutation on *mt-Ta*. To achieve this, we made use of pyrosequencing, a technique based on the "sequencing by synthesis principle" that relies on the quantification of light that is emitted as nucleotides complementary to the amplified sequence get incorporated. Pyrosequencing, allowed us to evaluate in-house in a cost-effective manner the proportion of mutant to WT mtDNA presented in single cell samples.

This endeavour was not without its complications, as we soon realised that heteroplasmy readings varied depending on the strand that was used as a template during the PCR amplification step. Trying to understand the reason behind this discrepancy, we toook a closer look at the DNA fragment that we were sequencing in each case. We observed that it was the presence of adenine (A) at the variable (heteroplasmic) position that was being overestimated during pyrosequencing. In more detail, when sequencing the forward strand of mtDNA in order to quantify the m.5019A>G mutation, we saw a consistent underestimation of the sample's heteroplasmy, when that value exceeded 50%. This was due to a consistent artificial increase of the A-peak seen in the pyrogram. Conversely, when pyrosequencing the reverse strand of the mtDNA trying to quantify the heteroplasmy percentage in m.5024C>T, we saw an overestimation of the heteroplasmy percentage when the sample measured had a value lower than 50%. Again, the reason was the overestimation of the A-peak that in this protocol represented the mutant haplotype. The cause behind this discrepancy is very likely to be the fact that the PyroMark® Q48 Advanced Reagents (QIAGEN, PN: 974002) instead of using dATP makes use of the slightly modified version, deooxyadenosine alpha thio triphosphate (dATPaS) which while still able to be incorporated in the newly synthesized DNA strand, it

cannot interact with luciferase to produce a false positive signal. However, it is possible that the incorporation of dATPaS causes a slight deviation in the pyrogram that when present at the heteroplasmic site skews the mutant-to-WT ration of mtDNA molecules present within a given sample.

Consequently, moving forward, we chose to avoid including adenine as a base upon which the pyrosequencing readings are based, and thus, we opted to use the forward strand as a template for evaluating m.5024C>T heteroplasmy and the reverse strand when measuring heteroplasmy at the m.5019A>G position. Having established a pyrosequencing protocol, we went ahead with the testing of single cell samples and soon found that they were exceptionally prone to PCR contamination. Speculating that the external source of this contamination was most likely the presence of mitochondrial DNA in the air and on the surfaces of the equipment used, we decided to adapt our practices in order to minimize it. At the same time, quality control steps were introduced in the pipeline intended to detect cross-contamination between the two mutant strains as well as the presence of external mtDNA within each of the 96-well plates that were used. While contamination-testing at the level of the plate and PCR-reaction-mix was feasible and thus carried out consistently, well-specific contamination events could never be ruled-out completely. In line with the stringent criteria introduced, wells that gave suboptimal readings or exhibited cross strain contamination were excluded.

Satisfied with the purity of and consistency of single cell samples that were being analysed, we decided to evaluate accuracy of pyrosequencing as a method. We were able to conclude that pyrosequencing in isolation is a very accurate method of quantifying mtDNA heteroplasmy. Interestingly, the majority of experimental error is not introduced at the level of the sequencing but instead, during PCR amplification of the target sequences. Consequently, we demonstrated that averaging out multiple measurements of mtDNA heteroplasmy on lysate derived from the same cell could be the way of getting the best estimation possible. Unfortunately, due to the sheer number of single cells used in most of our experiments and the cost and time required to implement such an approach, proved to be impractical. Instead, we concluded that the most accurate and cost-efficient way of measuring mtDNA heteroplasmy, would be to perform pyrosequencing on the entirety of the cell.

Overall, when considering single cell heteroplasmy measurements, there are certain confounders that do not dramatically skew the final result but one has to take into consideration.

First, the lower the mtDNA copy number of the cell being measured, the less reliable the copy number measurement is. This discrepancy gets amplified further in cases where the cell has been lysed and only a portion of the lysate has been used to carry out the reaction. While carrying out the reaction with around 455 mtDNA copies did not seen to give a significantly different result compared to bulk measurements, when the copy numbers were decreased even further, the measurement variance increased. Future work should also take into account that the theoretical variance of the sample depends on baseline heteroplasmy and in order to calculate mean-independent variance, similar results would have to be gathered on the variance of heteroplasmy log-odds. Further experiments can also be done in order to establish the mtDNA copy number threshold up to where the variance of repeated measurements does not differ significantly from that of bulk samples. However, in every situation one should keep in mind that when dealing with single cells, confounders such as cell-type specific restrictions in mtDNA copy number as well as variability within cell populations as a result of cell processes such as cell cycle progression will always be present. Seeing that the measurement variability accounted at most for one third of the variability seen in cell populations of interest and that half of the observed variance is a product of sampling and rounding effects, we decided go forward with the protocol while making use of the entirety of each cell in order to eradicate any sampling error. In conclusion, the methodology developed in **Results Chapter 3** was used extensively in the generation of data that was included in the following chapters.

In **Results Chapter 4**, we focused on the m.5019A>G and m.5024C>T mouse models of mitochondrial disease putting into practice the methods previously developed. Our first priority was to successfully isolate abundant cell populations that would be representative of the tissues of interest, namely the brain and spleen. Choosing these tissues was predicated upon the premise that in adult mice the brain would primary be a hub for cells that are quiescent (non-dividing), while the spleen would contain a large population of immune cells that are rapidly dividing throughout the lifespan of the organism. In the case of the spleen, CD19 was the surface epitope of choice that allowed for the isolation of B-cells. The rest of the spleen cells, primarily comprised of T-cells, were clustered together. In the case of the brain, ACSA-1 was used to select for astrocytes and its combination with Prominin-1 allowed for the isolation of double-positive neural stem cells. Unfortunately, this population becomes quite scarce in the adult mouse brain and, as a result, PSD95 was selected as a neuronal synaptic marker. At this stage, it should be mentioned that although informative when it comes to generating a broader view of the tissue, these populations should not be viewed as cell-type specific. When

considering the spleen, CD19 is expressed in B-lymphocytes, while of course, the population negative for the epitope except for T cells, is expected to contain a variety of cell types, including T-cells, granulocytes, dendritic cells, natural killer cells and macrophages. Future work that might be focused on mtDNA heteroplasmy dynamics and whether cell-type-specific patterns exist, a more detailed FACS panel will have to be generated, that employs more stringent selections steps. At the same time, consistency between cell populations at different time points through ageing is expected to also be a concern as seen in our selection of brain populations as well as when it comes to the presence of T-cells in the spleen of neonate mice that only becomes detectable by P2 (Le Campion et al., 2002). Despite these caveats, our FACS analysis provided us with the confidence that major cell populations that are central to the tissue activity survived the dissociation process and provided further credence to our attempt to differentiate between quiescent and dividing cells, while at the same time providing a basic level of cell type-specificity to our work.

Having chosen and isolated our populations of interest from each tissue, we sought to determine the single cell heteroplasmy range in both spleen and brain-derived populations from P100 mice carrying either the m.5019A>G or the m.5024C>T mutation. We were surprised to find out that all populations tested, derived from either mutant mouse strains, shared a wide single cell heteroplasmy range that was masked by the bulk tissue measurements. Moving forward, we decided to have a look at different time points, throughout mouse development and ageing. To that effect, tissues from animals belonging to both strains were collected at P0, P6 and P365 in addition to E8.5. Single cell heteroplasmy distributions at all these time points revealed a variance that was surprisingly broad but, in every case, the mean single cell heteroplasmy matched that of bulk tissue measurements. Transformation of the single cell data allowed for comparisons to be drawn across mice that had different bulk heteroplasmies and thus permitted us to group together all data that originated at a specific time point for each mutation. Having 3 mice from each strain per time point, allowed for conclusive evidence to be generated that pointed towards an increase in single cell variability with age.

Analysis of single cell data originating from mice carrying the m.5019A>G mutation on *mt-Ta*, revealed an age-dependent accumulation of mutant homoplasmic cells. Initially suspected to be a result of a positive selection in favour of the mutation, this hypothesis was quickly debunked when looking at the consistency with which the mean single cell heteroplasmy values correlated to the bulk tissue measurements. Further evidence against the initial hypothesis came

when m.5019A>G mice older than P100 carrying mid-levels of heteroplasmy were examined. These mice revealed that there was no drive towards mutant homoplasmy and, in the case of animals with heteroplasmies lower than 50%, revealed that more cells were driven towards a WT homoplasmic state. Further evidence in support of random genetic drift being the driving force behind the observed heteroplasmy segregation came when the data was fitted to the Kimura distribution. Confirming the null hypothesis, the observed age-dependent increase in mtDNA variance was once again proven to be the result of random genetic drift. Analysis of the Kimura *b* parameter took our observations a step further by establishing that the rate of drift is greater in mice carrying the 5019A>G mutation compared to their 5024C>T counterparts. This was a fascinating observation as it meant that despite their proximity on the mt-Ta molecule, the two mutations are being differentially regulated either at the cellular or tissue level. There are two main mechanisms driving random genetic drift, vegetative segregation, as mtDNA molecules get randomly segregated upon cytokinesis, and relaxed replication which refers to the replication of mtDNA that takes place in a semi-independent manner to the nDNA. Seeing that quiescent cells in the brain mimic the behaviour of the rapidly-dividing immune cells residing in the spleen, points towards vegetative segregation not being a driving factor. Instead, relaxed replication appears to account for most of the random genetic drift that we see in the populations derived from the brain of P100 and P365 mice. At the same time, P365 datasets showed that the rate of drift was increased in mice carrying the m.5019A>G mt-Ta mutation compared to their m.5024C>T counterparts. Future work can be aimed at uncovering why are the two mutations differentially regulated. Having built an understanding around the molecular profiles of the two mutants, we can ask the question: what mechanistic pathway is involved in restricting the drift of m.5024C>T compared to that of m.5019A>G? Do these pathways involve nuclear regulation of relaxed replication with age? And, is it the tighter control we see in tissues derived from m.5024C>T mice related to the hard 80% heteroplasmy threshold seen in neonates of that strain but not in the m.5019A>G one? While the scRNA-seq presented in **Results Chapter 5** was performed on mouse embryos belonging to each strain, we were granted an initial glimpse as to the compensatory mechanisms that were activated in response to each mutation. Looking through the differentially expressed pathways and linking them to data at the level of the proteome might underline molecular pathways that are worth following-up on. Ultimately, answering these questions will provide us with new insights into the mechanisms driving mtDNA heteroplasmy segregation as well as the nuclear control of mtDNA replication.

In **Results Chapter 5**, we address defects in mtDNA translation that have for years been recognized as the primary cause behind inherited mitochondrial disorders, affecting around 1:8000 people (Gorman et al., 2015). Although defective translation of mitochondrial-encoded components of the ETC will inevitably lead to OXPHOS defects, the biochemical characterisation of these mutations does not tell the whole story. Indeed, it is not uncommon to come across patients who suffer from closely related mutations that affect mt-tRNA genes but exhibit vastly different pathological phenotypes (Koenig, 2008). The reasons behind this seemingly common deficiency bringing about such divergent clinical symptoms is yet to be fully understood. A long-standing theory that tries to tackle this question makes the case for different cell types having strikingly different mitochondrial content while, at the same time, exhibiting different properties when it comes to their mitochondrial activity (Kim and Mook-Jung, 2019). This explanation goes a long way in addressing the aforementioned heterogeneity central to many mitochondrial disorders, but fails to adequately answer at which stage in organ development do these differences first emerge.

In **Results Chapter 5**, we present a project that was carried out in close collaboration other members of the Chinnery lab as well as collaborators from the labs of James B. Stewart and Michal Minczuk, as we attempted to expand our understanding around the role of pathogenic mt-tRNA mutations during embryonic development. We made use of two mouse strains carrying pathogenic mutations on *mt-Ta*, the m.5019A>G and the m.5024C>T mouse models. Despite the two mutations being positioned only 5bps apart on the same molecule, we demonstrate that they lead to the generation of distinct transcriptional responses during the process of organogenesis that is central to the development of the mouse embryo at E8.5. Analysis of transcriptomic profiles derived from single cells lead us to define 17 separate cell lineages. Further analysis showed that the differences we observed in nuclear gene transcription also appeared to be cell-lineage-specific in nature. Such observations raise the possibility that the tissue selectivity characterising mitochondrial disorders in humans first emerges during embryonic development *in utero*. Mitochondrial biogenesis and quality control is dependent on the expression of nuclear-encoded genes that are in turn subject to the principles of transcriptional regulation (Mears, 2014). Inter-tissue differences in the transcriptional regulation of those genes can as a result account for at least part of the aforementioned heterogeneity. Indeed, being able to pick apart lineage-specific transcriptional differences in WT mouse embryos at E8.5, reflect the cell-type specific differences in mitoproteome. This

observation falls in line with results suggesting that the presence of organ-specific transcripts precedes organ maturation.

As with human mitochondrial disorders, describing the two mutant mouse strains as just having OXPHOS defects does not do justice to the complexity of the underlying mechanisms and masks the pathology's origin. Consequently, understanding the molecular origins of the mitochondrial dysfunction in each of the two *mt-Ta* mutant mouse strains is of primary importance. While the m.5024C>T mutation has been described as detrimental to the stability of the mt-Ta molecule (Kaupplia et al., 2016), work from the Minczuk lab showed that the m.5019A>G mutation impairs tRNA charging. Proteomic analysis of liver samples derived from adult 5019A>G mice further validated observations that described significant decreases in protein levels of OXPHOS subunits belonging to respiratory complexes I, III and IV. A result in line with COX deficiency observations made by the Stewart lab. A more comprehensive look at this dataset is required in order to identify pathways of interest in the brain and liver of the two mutant mouse strains. Identification of such pathways can point us towards tissue-wide adaptations as well as finer consequences that result from the *mt-Ta* mutations in question.

The exact mechanism in which these mutations alter nuclear gene transcription is yet unclear but we have strived to demonstrate that they influence pathways both intrinsic and extrinsic to mitochondria, in a manner that is cell-lineage-specific in nature. The overlap between genes that were downregulated in embryos carrying the two pathogenic *mt-Ta* mutations with those enriched in *in vitro* CRISPR screens for buffering or rescuing the effects of mitochondrial toxins is telling (To et al., 2019). Similar overlap was noted when comparing our dataset with that generated following the *in vitro* application of mtDNA transcription inhibitors (Mennuni et al., 2022). Having the same molecular pathways being implicated across these three independent studies, points towards the presence of strong, tissue-specific compensatory responses that arise in response to mitochondrial dysfunction during embryonic development. This observation might have the answer as to why we see some severe OXPHOS disorders in humans arise after birth while others do not surface until adulthood. We have also shown how cell-lineage-specific compensatory responses are under the control of certain transcription factors, thus opening the door for tissue-specific therapies that target known druggable targets. Finally, given our still expanding understanding of secondary mitochondrial disorders and the role that the organelle plays in the emergence of common late-onset diseases, such as

Parkinson's disease and diabetes mellitus (Stewart and Chinnery, 2021), it is not a stretch to argue that our findings might be of therapeutic value in conditions where mitochondrial dysfunction is only part of the pathology. With advances is the field of scRNA-seq being developed at a rapid pace, the natural step forwards from this work would be to use the protocols described by Mimitou and colleagues (Mimitou et al., 2021) in order to be able to link the changes in transcriptome observed at the single cell level, to mtDNA heteroplasmy levels. In their work, Eleni Mimitou and colleagues demonstrate the use of single cell systematic multi-omic approaches that combine assay for transposase accessible chromatin (ATAC) with select antigen profiling by sequencing (ASAP-seq) and an adapted from of cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq), termed DOGMA-seq. This approach manages to simultaneously monitor changes in chromatin accessibility, RNA and cell surface protein expression (Mimitou et al., 2021). Such an analysis performed either at the embryonic stage or within a tissue that is known to be heavily impacted by mitochondrial disorders will go a long way in informing us about the exact relationship between the level of mitochondrial disorder that a certain cell type can tolerate, the pathways involved as well as changes that occur as certain molecular thresholds are exceeded. Future work in this field can also be geared towards investigating the impact that age-dependent trends in heteroplasmy segregation (discussed in more detail in **Results Chapter 4**) might have on different cell types residing within tissues of interest.

When taken together, the introduction of both m.5019A>G and m.5024C>T mouse models paved the way to exciting new discoveries, as we were able to study genetically impaired mtDNA translation. Work on these two models, revealed that transcriptional signatures of nuclear-encoded mitochondrial genes are impacted in a cell-specific manner once the organelle's activity is undermined. More specifically, we were able to observe the down-regulation of genes that were previously implicated in the exacerbation of cellular response to exogenous toxins (To et al., 2019). This apparent transcriptional adaptation to genetically-induced mitochondrial dysfunction implicated molecular pathways the involvement of which in mitochondrial maintenance was previously unknown. Furthermore, cellular adaptations in response to these insults to mitochondrial activity appeared to be not only tissue and cell-type specific but also mutation-specific in nature, revealing fascinating intricacies in the way in which different mutations have a disparate impact on nuclear-mitochondrial crosstalk. The tissue-specific nature of compensatory pathways seen as promoters of mitochondrial resilience

hints to the nature of mitochondrial disorders and opens the door to exploring organ-centric approaches to treatment.

Finally, in **Chapter 6** we have taken the first steps into a project aimed at uncovering the complicated relationship between cell cycle control and mtDNA copy number as well as heteroplasmy. Initially, we sought to establish a relationship between cell cycle progression and mtDNA copy number. Making use of the PIP-FUCCI construct, that led to the expression of nuclear-localised, cell cycle-dependent proteins that were fluorescently tagged, we were able to sort cells at different stages of the cell cycle, without having to subject the cells to chemical synchronisation. The result was a step-wise increase as cells progressed from G_1-phase to S-phase and into G_2/M-phase. While there has been evidence of S-phase being the point at which most of mitochondrial DNA synthesis takes place in preparation for cytokinesis, we did not observe such behaviour. Wanting to take a closer look at the intermediate steps we took advantage of the gradual increase in red-fluorescence when transitioning from G_1-phase to S-phase and in green-fluorescence when moving from S-phase into G_2/M-phase. This gradual transition, allowed us to subdivide the cell cycle into 3 sub-phases for S-phase and two for G_2/M-phase. Measuring the copy number of cells belonging to each of the sub-phases revealed a more gradual increase in mtDNA copy number that takes placed throughout the cell cycle. Then, we decided to examine whether restriction of the cell's reliance on OXPHOS, through culture in low glucose and galactose-containing media, would have an effect on the levels of mtDNA copy number present within the cell. Eventually we saw that HeLa cells grown in galactose did exhibit a significant increase in mtDNA copy number by week 4 while culturing them in low glucose-containing media did not have an impact on copy number but did significantly impact the time required for them to complete a cell cycle. It is possible that an increase of mtDNA levels similar to that observed when the same cells were cultured in galactose would compensate glycolytic restriction and allow the cells to progress through the cell cycle at a normal rate. Recognising that HeLa cells being a cancer cell line are quite robust when it comes to shifts in culture conditions, we decided to perform a TFAM CRISPR KO and examine what its effects would be on cell cycle progression. Having confirmed that the KO was successful, we went ahead and tested for changes in cell cycle progression in cells that were cultured in high and low glucose media once having had their mtDNA all but eliminated. Indeed, there was a significant reduction when TFAM KO HeLa cells were grown in normal and low glucose, when compared to their WT controls. This is probably a sign of a universal growth delay that is brought about as a result of the lower levels of OXPHOS taking place.

Having gathered some initial insight into the relationship between mtDNA copy number and the cell cycle using the PIP-FUCCI construct in HeLa cells, we decided to take our research a step further and make use of the ΔH2.1 cybrid cell line that carried a large-scale mtDNA deletion which was responsible for more than half of the molecule being deleted. Having devised a ddPCR protocol that could be used to measure both mtDNA copy number and heteroplasmy at the same time, we proceeded to sub-clone the original ΔH2.1 population in order to acquire uniform cell populations of varying heteroplasmy percentages. We then sought to establish the impact that different heteroplasmy levels of the deletion had on the cell's ability to effectively perform OXPHOS, and found that while low heteropalsmies cause little to no changes, mid-to-high levels brought about a rapid drop in cell respiratory capacity. With this result in mind, we tested the time required for a range of ΔH2.1 clones to complete the cell cycle, and observed that when grown in standard culture media, no significant differences appeared with respect to the difference heteroplasmy values. Culturing high and low heteroplasmy clones in high-glucose and galactose media also had no effect on the time required by ΔH2.1 cells to complete the cell cycle. Overall, **Chapter 6** presents a side project that was birthed out of our interest in the interactions between mtDNA copy number and heteroplasmy and the impact they might have on a cell's ability to grow and divide. While relatively preliminary, out results have shown that in order for the impact of mitochondrial dysfunction to be impactful *in vitro*, cells have to be encouraged to depend on OXPHOS in order to meet their energy demands.

At the same time, the introduction of heteroplasmic cell lines carrying severely pathogenic mtDNA mutations, has opened the door into a new model that could bridge the gap between our previous work on heteroplasmy and our venture into the role of mitochondria in cell cycle control. Future experiments involving cell cycle progression and heteroplasmy could include imaging of the mitochondrial network at different cell cycle stages in high and low-heteroplasmy ΔH2.1 clones in order to determine whether large-scale pathogenic mtDNA deletions have an effect of network morphology. At the same time, the PIP-FUCCI construct can be introduced to the MEFs deriving from the heteroplasmic mouse lines carrying the m.5019A>G and m.5024C>T *mt-Ta* mutations that are already a tool being used extensively in our lab. Finally, once a discrepancy in cell cycle progression is detected, bulk RNA-seq on the clones of interest and follow-up western blotting experiments could reveal a correlation

between the heteroplasmy levels of pathogenic mtDNA mutations and the abundance of proteins involved in cell cycle control and progression.

References

Abicht, A., et al. (2018). "Mitochondrial and nuclear disease panel (Mito-aND-Panel): Combined sequencing of mitochondrial and nuclear DNA by a cost-effective and sensitive NGS-based method." Mol Genet Genomic Med **6**(6): 1188-1198.

Adams, J. M. (2003). "Ways of dying: multiple pathways to apoptosis." Genes & Development **17**(20): 2481-2495.

Ahier, A., et al. (2018). "Affinity purification of cell-specific mitochondria from whole animals resolves patterns of genetic mosaicism (vol 20, pg 352, 2018)." Nature Cell Biology **20**(3): 361-361.

Aibar, S., et al. (2017). "SCENIC: single-cell regulatory network inference and clustering." Nat Methods **14**(11): 1083-1086.

Aiken, C. E. M., et al. (2008). "Variations in mouse mitochondrial DNA copy number from fertilization to birth are associated with oxidative stress." Reproductive Biomedicine Online **17**(6): 806-813.

al-Khodairy, F. and A. M. Carr (1992). "DNA repair mutants defining G2 checkpoint pathways in Schizosaccharomyces pombe." EMBO J **11**(4): 1343-1350.

Allen, J. F. (2017). "The CoRR hypothesis for genes in organelles." J Theor Biol **434**: 50-57.

Alston, C. L., et al. (2017). "The genetics and pathology of mitochondrial disease." J Pathol **241**(2): 236-250.

Altmann, R. 1894. Die Elementarorganismen und ihre Beziehungen zu den Zellen.

Alzu'bi, A., et al. (2017). "The Transcription Factors COUP-TFI and COUP-TFII have Distinct Roles in Arealisation and GABAergic Interneuron Specification in the Early Human Fetal Telencephalon." Cerebral Cortex **27**(10): 4971-4987.

Ameur, A., et al. (2011). "Ultra-Deep Sequencing of Mouse Mitochondrial DNA: Mutational Patterns and Their Origins." Plos Genetics **7**(3).

Amunts, A., et al. (2015). "The structure of the human mitochondrial ribosome." Science **348**(6230): 95-98.

Anderson, S., et al. (1981). "Sequence and organization of the human mitochondrial genome." Nature **290**(5806): 457-465.

Anderson, S., et al. (1982). "Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome." J Mol Biol **156**(4): 683-717.

Andreoli, C., et al. (2004). "MitoP2, an integrated database on mitochondrial proteins in yeast and man." Nucleic Acids Research **32**: D459-D462.

Andrews, R. M., et al. (1999). "Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA." Nat Genet **23**(2): 147.

Antonicka, H. and E. A. Shoubridge (2015). "Mitochondrial RNA Granules Are Centers for Posttranscriptional RNA Processing and Ribosome Biogenesis." Cell Reports **10**(6): 920-932.

Arbor, A. (2022). "10X Genomics Single-cell Processing." Retrieved June 10th, 2022, from https://brcf.medicine.umich.edu/cores/advanced-genomics/technologies/single-cell/10x-genomics/.

Argelaguet, R., et al. (2019). "Multi-omics profiling of mouse gastrulation at single-cell resolution." Nature **576**(7787): 487-+.

Arnaudo, E., et al. (1991). "Depletion of Muscle Mitochondrial-DNA in Aids Patients with Zidovudine-Induced Myopathy." Lancet **337**(8740): 508-510.

Arnoult, D., et al. (2011). "Mitochondria in innate immunity." EMBO Rep **12**(9): 901-910.

Arthur, C. R., et al. (2009). "Parkinson's disease brain mitochondria have impaired respirasome assembly, age-related increases in distribution of oxidative damage to mtDNA and no differences in heteroplasmic mtDNA mutation abundance." Molecular Neurodegeneration **4**.

Bai, Y., et al. (2011). "Structural Basis for Dimerization and Activity of Human PAPD1, a Noncanonical Poly(A) Polymerase." Molecular Cell **41**(3): 311-320.

Baranovskiy, A. G., et al. (2014). "Structural basis for inhibition of DNA replication by aphidicolin." Nucleic Acids Research **42**(22): 14013-14021.

Barchiesi, A. and C. Vascotto (2019). "Transcription, Processing, and Decay of Mitochondrial RNA in Health and Disease." International Journal of Molecular Sciences **20**(9).

Barrientos, A. (2015). "Mitochondriolus: assembling mitoribosomes." Oncotarget **6**(19): 16800-16801.

Bayona-Bafaluy, M. P., et al. (2003). "Revisiting the mouse mitochondrial DNA sequence." Nucleic Acids Res **31**(18): 5349-5355.

Becker, T., et al. (2012). "Mitochondrial protein import: from transport pathways to an integrated network." Trends in Biochemical Sciences **37**(3): 85-91.

Benda, C. (1898). Ueber die spermatogenese der vertebraten und höherer evertebraten, II. Theil: Die histiogenese der spermien. Arch. Anat. Physiol, 393-398.

Benkhalifa, M., et al. (2014). "Mitochondria: Participation to infertility as source of energy and cause of senescence." International Journal of Biochemistry & Cell Biology **55**: 60-64.

Berg, O. G. and C. G. Kurland (2000). "Why mitochondrial genes are most often found in nuclei." Molecular Biology and Evolution **17**(6): 951-961.

Bergman, O. and D. Ben-Shachar (2016). "Mitochondrial Oxidative Phosphorylation System (OXPHOS) Deficits in Schizophrenia: Possible Interactions with Cellular Processes." Can J Psychiatry **61**(8): 457-469.

Bezawork-Geleta, A., et al. (2017). "Mitochondrial Complex II: At the Crossroads." Trends in Biochemical Sciences **42**(4): 312-325.

Birky, C. W. (1994). "Relaxed and Stringent Genomes - Why Cytoplasmic Genes Dont Obey Mendels Laws." Journal of Heredity **85**(5): 355-365.

Boczonadi, V. and R. Horvath (2014). "Mitochondria: impaired mitochondrial translation in human disease." Int J Biochem Cell Biol **48**: 77-84.

Boczonadi, V., et al. (2015). "Reversible infantile mitochondrial diseases." Journal of Inherited Metabolic Disease **38**(3): 427-435.

Boczonadi, V., et al. (2018). "Mitochondrial DNA transcription and translation: clinical syndromes." Mitochondrial Diseases **62**(3): 321-340.

Bonekamp, N. A., et al. (2020). "Small-molecule inhibitors of human mitochondrial DNA transcription." Nature **588**(7839): 712-+.

Borowski, L. S., et al. (2013). "Human mitochondrial RNA decay mediated by PNPase-hSuv3 complex takes place in distinct foci." Nucleic Acids Research **41**(2): 1223-1240.

Bronte, V. and M. J. Pittet (2013). "The spleen in local and systemic regulation of immunity." Immunity **39**(5): 806-818.

Brown, A., et al. (2014). "Structure of the large ribosomal subunit from human mitochondria." Science **346**(6210): 718-722.

Brown, D. T., et al. (2001). "Random genetic drift determines the level of mutant mtDNA in human primary oocytes." American Journal of Human Genetics **68**(2): 533-536.

Budd, S. L. and D. G. Nicholls (1996). "A reevaluation of the role of mitochondria in neuronal Ca2+ homeostasis." Journal of Neurochemistry **66**(1): 403-411.

Burgstaller, J. P., et al. (2014). "mtDNA Segregation in Heteroplasmic Tissues Is Common In Vivo and Modulated by Haplotype Differences and Developmental Stage." Cell Reports **7**(6): 2031-2041.
Bykhovskaya, Y., et al. (2004). "Missense mutation in pseudouridine synthase 1 (PUS1) causes mitochondrial myopathy and sideroblastic anemia (MLASA)." American Journal of Human Genetics **74**(6): 1303-1308.
Cai, X., et al. (2013). "Metformin suppresses hepatocellular carcinoma cell growth through induction of cell cycle G1/G0 phase arrest and p21CIP and p27KIP expression and downregulation of cyclin D1 in vitro and in vivo." Oncol Rep **30**(5): 2449-2457.
Calvo, S. E., et al. (2016). "MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins." Nucleic Acids Research **44**(D1): D1251-D1257.
Calvo, S., et al. (2006). "Systematic identification of human mitochondrial disease genes through integrative genomics." Nature Genetics **38**(5): 576-582.
Campbell, G., et al. (2014). "Dissecting the mechanisms underlying the accumulation of mitochondrial DNA deletions in human skeletal muscle." Human Molecular Genetics **23**(17): 4612-4620.
Campos, A. and A. Clemente-Blanco (2020). "Cell Cycle and DNA Repair Regulation in the Damage Response: Protein Phosphatases Take Over the Reins." International Journal of Molecular Sciences **21**(2).
Cann, R. L., et al. (1987). "Mitochondrial-DNA and Human-Evolution." Nature **325**(6099): 31-36.
Cao, L. H., et al. (2014). "Phylogenetic analysis of CDK and cyclin proteins in premetazoan lineages." Bmc Evolutionary Biology **14**.
Cao, L. Q., et al. (2007). "The mitochondrial bottleneck occurs without reduction of mtDNA content in female mouse germ cells." Nature Genetics **39**(3): 386-390.
Cao, L. Q., et al. (2009). "New Evidence Confirms That the Mitochondrial Bottleneck Is Generated without Reduction of Mitochondrial DNA Content in Early Primordial Germ Cells of Mice." Plos Genetics **5**(12).
Carrodeguas, J. A., et al. (1999). "The accessory subunit of Xenopus laevis mitochondrial DNA polymerase gamma increases processivity of the catalytic subunit of human DNA polymerase gamma and is related to class II aminoacyl-tRNA synthetases." Mol Cell Biol **19**(6): 4039-4046.
Castellani, C. A., et al. (2020). "Mitochondrial DNA copy number can influence mortality and cardiovascular disease via methylation of nuclear DNA CpGs." Genome Medicine **12**(1).
Cavalier-Smith, T. (2006). "Origin of mitochondria by intracellular enslavement of a photosynthetic purple bacterium." Proceedings of the Royal Society B-Biological Sciences **273**(1596): 1943-1952.
Cecchini, G. (2003). "Function and structure of complex II of the respiratory chain." Annual Review of Biochemistry **72**: 77-109.
Chatre, L. and M. Ricchetti (2013). "Prevalent coordination of mitochondrial DNA transcription and initiation of replication with the cell cycle." Nucleic Acids Research **41**(5): 3068-3078.
Chen, H. C., et al. (2003). "Mitofusins Mfn1 and Mfn2 coordinately regulate mitochondrial fusion and are essential for embryonic development." Journal of Cell Biology **160**(2): 189-200.
Chen, J. and J. R. Patton (1999). "Cloning and characterization of a mammalian pseudouridine synthase." RNA **5**(3): 409-419.
Chin, C. F. and F. M. Yeong (2010). "Safeguarding Entry into Mitosis: the Antephase Checkpoint." Molecular and Cellular Biology **30**(1): 22-32.

Chinnery, P. F. (1993). Primary Mitochondrial Disorders Overview. GeneReviews((R)). M. P. Adam, D. B. Everman, G. M. Mirzaa et al. Seattle (WA).
Chinnery, P. F. and D. C. Samuels (1999). "Relaxed replication of mtDNA: A model with implications for the expression of disease." American Journal of Human Genetics **64**(4): 1158-1165.
Chinnery, P. F., et al. (1999). "Nonrandom tissue distribution of mutant mtDNA." American Journal of Medical Genetics **85**(5): 498-501.
Chinnery, P. F., et al. (2000). "The inheritance of mitochondrial DNA heteroplasmy: random drift, selection or both?" Trends in Genetics **16**(11): 500-505.
Chinnery, P. F., et al. (2004). "Risk of developing a mitochondrial DNA deletion disorder." Lancet **364**(9434): 592-596.
Chinnery, P. F., et al. (2012). "Epigenetics, epidemiology and mitochondrial DNA diseases." Int J Epidemiol **41**(1): 177-187.
Choi, Y. H. and Y. H. Yoo (2012). "Taxol-induced growth arrest and apoptosis is associated with the upregulation of the Cdk inhibitor, p21(WAF1/CIP1) in human breast cancer cells." Oncology Reports **28**(6): 2163-2169.
Ciafaloni, E., et al. (1991). "Widespread Tissue Distribution of a Transfer-Rna Leu (Uur) Mutation in the Mitochondrial-DNA of a Patient with Melas Syndrome." Neurology **41**(10): 1663-1665.
Clark, K. M., et al. (1997). "Reversal of a mitochondrial DNA defect in human skeletal muscle." Nature Genetics **16**(3): 222-224.
Claros, M. G., et al. (1995). "Limitations to in-Vivo Import of Hydrophobic Proteins into Yeast Mitochondria - the Case of a Cytoplasmically Synthesized Apocytochrome-B." European Journal of Biochemistry **228**(3): 762-771.
Clayton, D. A. (1982). "Replication of Animal Mitochondrial-DNA." Cell **28**(4): 693-705.
Cohen, T., et al. (2016). "Ancient Out-of-Africa Mitochondrial DNA Variants Associate with Distinct Mitochondrial Gene Expression Patterns." Plos Genetics **12**(11).
Coley, A. A. and W. J. Gao (2018). "PSD95: A synaptic protein implicated in schizophrenia or autism?" Prog Neuropsychopharmacol Biol Psychiatry **82**: 187-194.
Coquery, C. M., et al. (2012). "Optimized protocol for the isolation of spleen-resident murine neutrophils." Cytometry Part A **81a**(9): 806-814.
Costa, R., et al. (2021). "Mitochondrial dysfunction interferes with neural crest specification through the FoxD3 transcription factor." Pharmacological Research **164**.
Cotter, D., et al. (2004). "MitoProteome: mitochondrial protein sequence database and annotation system." Nucleic Acids Research **32**: D463-D467.
Cotterill, M., et al. (2013). "The activity and copy number of mitochondrial DNA in ovine oocytes throughout oogenesis in vivo and during oocyte maturation in vitro." Molecular Human Reproduction **19**(7): 444-450.
Coxhead, J., et al. (2016). "Somatic mtDNA variation is an important component of Parkinson's disease." Neurobiology of Aging **38**.
Cree, L. M., et al. (2008). "A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes." Nature Genetics **40**(2): 249-254.
Cristinelli, S. and A. Ciuffi (2018). "The use of single-cell RNA-Seq to understand virus-host interactions." Current Opinion in Virology **29**: 39-50.
D'Souza, A. R. and M. Minczuk (2018). "Mitochondrial transcription and translation: overview." Mitochondrial Diseases **62**(3): 309-320.
Dawson, V. L. and T. M. Dawson (2004). "Deadly conversations: Nuclear-mitochondrial cross-talk." Journal of Bioenergetics and Biomembranes **36**(4): 287-294.

Deak, A. T., et al. (2014). "IP3-mediated STIM1 oligomerization requires intact mitochondrial Ca2+ uptake." Journal of Cell Science **127**(13): 2944-2955.
Decatur, W. A. and M. J. Fournier (2002). "rRNA modifications and ribosome function." Trends in Biochemical Sciences **27**(7): 344-351.
deGrey, A. D. N. J. (1997). "A proposed refinement of the mitochondrial free radical theory of aging." Bioessays **19**(2): 161-166.
Diaz, F., et al. (2002). "Human mitochondrial DNA with large deletions repopulates organelles faster than full-length genomes under relaxed copy number control." Nucleic Acids Research **30**(21): 4626-4633.
DiMauro, S., et al. (2013). "The clinical maze of mitochondrial neurology." Nat Rev Neurol **9**(8): 429-444.
Ding, J., et al. (2015). "Assessing Mitochondrial DNA Variation and Copy Number in Lymphocytes of similar to 2,000 Sardinians Using Tailored Sequencing Analysis Tools." Plos Genetics **11**(7).
Ding, L., et al. (2020). "The Roles of Cyclin-Dependent Kinases in Cell-Cycle Progression and Therapeutic Strategies in Human Breast Cancer." International Journal of Molecular Sciences **21**(6).
Dolezal, P., et al. (2006). "Evolution of the molecular machines for protein import into mitochondria." Science **313**(5785): 314-318.
Duchen, M. R. (2004). "Mitochondria in health and disease: perspectives on a new mitochondrial biology." Molecular Aspects of Medicine **25**(4): 365-451.
Dudek, J., et al. (2013). "Mitochondrial protein import: Common principles and physiological networks." Biochimica Et Biophysica Acta-Molecular Cell Research **1833**(2): 274-285.
Duechler, M., et al. (2016). "Nucleoside modifications in the regulation of gene expression: focus on tRNA." Cell Mol Life Sci **73**(16): 3075-3095.
Dunn, J. and M. H. Grider (2022). Physiology, Adenosine Triphosphate. StatPearls. Treasure Island (FL).
Eimon, P. M., et al. (1996). "Age-associated mitochondrial DNA deletions in mouse skeletal muscle: comparison of different regions of the mitochondrial genome." Dev Genet **18**(2): 107-113.
Elias-Wolff, F., et al. (2019). "Curvature sensing by cardiolipin in simulated buckled membranes." Soft Matter **15**(4): 792-802.
Embley, T. M. and W. Martin (2006). "Eukaryotic evolution, changes and challenges." Nature **440**(7084): 623-630.
Ernster, L., et al. (1959). "Enzymic activities of human skeletal muscle mitochondria: a tool in clinical metabolic research." Nature **184**: 1851-1854.
Fairbrother-Browne, A., et al. (2021). "Mitochondrial-nuclear cross-talk in the human brain is modulated by cell type and perturbed in neurodegenerative disease." Communications Biology **4**(1).
Fakruddin, M. and A. Chowdhury (2012). "Pyrosequencing-An Alternative to Traditional Sanger Sequencing." American Journal of Biochemistry and Biotechnology **8**(1).
Falkenberg, M. and C. M. Gustafsson (2020). "Mammalian mitochondrial DNA replication and mechanisms of deletion formation." Critical Reviews in Biochemistry and Molecular Biology **55**(6): 509-524.
Falkenberg, M., et al. (2002). "Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA." Nature Genetics **31**(3): 289-294.
Fan, W. W., et al. (2008). "A mouse model of mitochondrial disease reveals germline selection against severe mtDNA mutations." Science **319**(5865): 958-962.
Farge, G., et al. (2014). "In Vitro-Reconstituted Nucleoids Can Block Mitochondrial DNA Replication and Transcription." Cell Reports **8**(1): 66-74.

Farr, C. L., et al. (1999). "Functional interactions of mitochondrial DNA polymerase and single-stranded DNA-binding protein - Template-primer DNA binding and initiation and elongation of DNA strand synthesis." Journal of Biological Chemistry **274**(21): 14779-14785.
Fernandez-Vizarra, E. and M. Zeviani (2021). "Mitochondrial disorders of the OXPHOS system." Febs Letters **595**(8): 1062-1106.
Fernandez-Vizarra, E., et al. (2011). "Tissue-specific differences in mitochondrial activity and biogenesis." Mitochondrion **11**(1): 207-213.
Filograna, R., et al. (2019). "Modulation of mtDNA copy number ameliorates the pathological consequences of a heteroplasmic mtDNA mutation in the mouse." Sci Adv **5**(4): eaav9824.
Filograna, R., et al. (2021). "Mitochondrial DNA copy number in human disease: the more the better?" Febs Letters **595**(8): 976-1002.
Florentz, C., et al. (2003). "Human mitochondrial tRNAs in health and disease." Cellular and Molecular Life Sciences **60**(7): 1356-1375.
Floros, V. I., et al. (2018). "Segregation of mitochondrial DNA heteroplasmy through a developmental genetic bottleneck in human embryos (vol 20, pg 144, 2018)." Nature Cell Biology **20**(8): 991-991.
Fontanesi, F., et al. (2020). Chapter 2 - Human mitochondrial transcription and translation. The Human Mitochondrial Genome. G. Gasparre and A. M. Porcelli, Academic Press**:** 35-70.
Fothergill-Gilmore, L. A. and P. A. M. Michels (1993). "Evolution of Glycolysis." Progress in Biophysics & Molecular Biology **59**(2): 105-235.
Frey, T. G., et al. (2002). "Insight into mitochondrial structure and function from electron tomography." Biochimica Et Biophysica Acta-Bioenergetics **1555**(1-3): 196-203.
Freyer, C., et al. (2012). "Variation in germline mtDNA heteroplasmy is determined prenatally but modified during subsequent transmission." Nature Genetics **44**(11): 1282-1285.
Friedman, J. R. and K. H. Kaestner (2006). "The Foxa family of transcription factors in development and metabolism." Cellular and Molecular Life Sciences **63**(19-20): 2317-2328.
Fu, Y., et al. (2020). "Safeguarding mitochondrial genomes in higher eukaryotes." Nat Struct Mol Biol **27**(8): 687-695.
Fukuoh, A., et al. (2014). "Screen for mitochondrial DNA copy number maintenance genes reveals essential role for ATP synthase." Molecular Systems Biology **10**(6).
Gabaldon, T. and M. A. Huynen (2003). "Reconstruction of the proto-mitochondrial metabolism." Science **301**(5633): 609.
Gaignard, P., et al. (2013). "Mitochondrial Infantile Liver Disease due to TRMU Gene Mutations: Three New Cases." Jimd Reports, Vol 11 **11**: 117-123.
Garcia-Reyes, B., et al. (2018). "The Emerging Role of Cyclin-Dependent Kinases (CDKs) in Pancreatic Ductal Adenocarcinoma." International Journal of Molecular Sciences **19**(10).
Gatenby, R. A. and R. J. Gillies (2004). "Why do cancers have high aerobic glycolysis?" Nature Reviews Cancer **4**(11): 891-899.
Gaziev, A. I., et al. (2014). "Mitochondrial function and mitochondrial DNA maintenance with advancing age." Biogerontology **15**(5): 417-438.
Geng, Y., et al. (2003). "Cyclin E ablation in the mouse." Cell **114**(4): 431-443.
Gharizadeh, B., et al. (2003). "Viral and microbial genotyping by a combination of multiplex competitive hybridization and specific extension followed by hybridization to generic tag arrays." Nucleic Acids Research **31**(22).

Gharizadeh, B., et al. (2007). "Pyrosequencing Technology for Short DNA Sequencing and Whole Genome Sequencing." Seibutsu Butsuri **47**(2): 129-132.
Ghezzi, D., et al. (2012). "Mutations of the Mitochondrial-tRNA Modifier MTO1 Cause Hypertrophic Cardiomyopathy and Lactic Acidosis." American Journal of Human Genetics **90**(6): 1079-1087.
Giacinti, C. and A. Giordano (2006). "RB and cell cycle progression." Oncogene **25**(38): 5220-5227.
Gnaiger, E., et al. (1995). "Control of mitochondrial and cellular respiration by oxygen." Journal of Bioenergetics and Biomembranes **27**(6): 583-596.
Gorman, G. S., et al. (2015). "Prevalence of Nuclear and Mitochondrial DNA Mutations Related to Adult Mitochondrial Disease." Annals of Neurology **77**(5): 753-759.
Gorman, G. S., et al. (2016). "Mitochondrial diseases." Nature Reviews Disease Primers **2**: 1-22.
Grant, G. D., et al. (2018). "Accurate delineation of cell cycle phase transitions in living cells with PIP-FUCCI." Cell Cycle **17**(21-22): 2496-2516.
Gray, M. W., et al. (1999). "Mitochondrial evolution." Science **283**(5407): 1476-1481.
Greber, B. J., et al. (2015). "The complete structure of the 55S mammalian mitochondrial ribosome." Science **348**(6232): 303-308.
Green, D. R. (2005). "Apoptotic pathways: Ten minutes to dead." Cell **121**(5): 671-674.
Green, D. R. and G. I. Evan (2002). "A matter of life and death." Cancer Cell **1**(1): 19-30.
Gross, A., et al. (2015). "Technologies for Single-Cell Isolation." International Journal of Molecular Sciences **16**(8): 16897-16919.
Gustafsson, C. M., et al. (2016). "Maintenance and Expression of Mammalian Mitochondrial DNA." Annual Review of Biochemistry, Vol 85 **85**: 133-160.
Hagstrom, E., et al. (2014). "No recombination of mtDNA after heteroplasmy for 50 generations in the mouse maternal germline." Nucleic Acids Research **42**(2): 1111-1116.
Haldin, C. E. and C. LaBonne (2010). "SoxE factors as multifunctional neural crest regulatory factors." International Journal of Biochemistry & Cell Biology **42**(3): 441-444.
Halliwell, B. and C. E. Cross (1994). "Oxygen-Derived Species - Their Relation to Human-Disease and Environmental-Stress." Environmental Health Perspectives **102**: 5-12.
Hance, N., et al. (2005). "Mitochondrial DNA polymerase gamma is essential for mammalian embryogenesis." Human Molecular Genetics **14**(13): 1775-1783.
Hardy, J. (2010). "Genetic Analysis of Pathways to Parkinson Disease." Neuron **68**(2): 201-206.
Harmel, J., et al. (2013). "The leucine-rich pentatricopeptide repeat-containing protein (LRPPRC) does not activate transcription in mammalian mitochondria." J Biol Chem **288**(22): 15510-15519.
Harrington, C. T., et al. (2013). "Fundamentals of pyrosequencing." Arch Pathol Lab Med **137**(9): 1296-1303.
Hartmann, J. and A. Verkhratsky (1998). "Relations between intracellular Ca2+ stores and store-operated Ca2+ entry in primary cultured human glioblastoma cells." Journal of Physiology-London **513**(2): 411-424.
Hartwell, L. H. and T. A. Weinert (1989). "Checkpoints - Controls That Ensure the Order of Cell-Cycle Events." Science **246**(4930): 629-634.
Hauswirth, W. W. and P. J. Laipis (1982). "Mitochondrial-DNA Polymorphism in a Maternal Lineage of Holstein Cows." Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences **79**(15): 4686-4690.

Hedlund, E. and Q. L. Deng (2018). "Single-cell RNA sequencing: Technical advancements and biological applications." Molecular Aspects of Medicine **59**: 36-46.
Hellebrekers, D. M. E. I., et al. (2012). "PGD and heteroplasmic mitochondrial DNA point mutations: a systematic review estimating the chance of healthy offspring." Human Reproduction Update **18**(4): 341-349.
Helm, M., et al. (2000). "Search for characteristic structural features of mammalian mitochondrial tRNAs." RNA **6**(10): 1356-1379.
Herbst, A., et al. (2021). "Skeletal muscle mitochondrial DNA copy number and mitochondrial DNA deletion mutation frequency as predictors of physical performance in older men and women." Geroscience **43**(3): 1253-1264.
Hess, J. F., et al. (1991). "Impairment of Mitochondrial Transcription Termination by a Point Mutation Associated with the Melas Subgroup of Mitochondrial Encephalomyopathies." Nature **351**(6323): 236-239.
Hillen, H. S., et al. (2017). "Structural Basis of Mitochondrial Transcription Initiation." Cell **171**(5): 1072-+.
Hinkle, P. C., et al. (1991). "Mechanistic Stoichiometry of Mitochondrial Oxidative-Phosphorylation." Biochemistry **30**(14): 3576-3582.
Hirst, J. (2013). "Mitochondrial complex I." Annual Review of Biochemistry **82**: 551-575.
Holt, I. J. and A. Reyes (2012). "Human Mitochondrial DNA Replication." Cold Spring Harbor Perspectives in Biology **4**(12).
Holt, I. J., et al. (2000). "Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA." Cell **100**(5): 515-524.
Hoppins, S., et al. (2007). "The machines that divide and fuse mitochondria." Annual Review of Biochemistry **76**: 751-780.
Hori, H. (2014). "Methylated nucleosides in tRNA and tRNA methyltransferases." Front Genet **5**: 144.
Hoth, M., et al. (2000). "Mitochondrial control of calcium-channel gating: A mechanism for sustained signaling and transcriptional activation in T lymphocytes." Proceedings of the National Academy of Sciences of the United States of America **97**(19): 10607-10612.
Hou, S. Y., et al. (2022). "Heterogeneity in endothelial cells and widespread venous arterialization during early vascular development in mammals." Cell Research **32**(4): 333-348.
Howard, A. and S. R. Pelc (1986). "Synthesis of Deoxyribonucleic-Acid in Normal and Irradiated-Cells and Its Relation to Chromosome Breakage." International Journal of Radiation Biology **49**(2): 207-218.
Howard, M. K., et al. (1993). "Cell Cycle Arrest of Proliferating Neuronal Cells by Serum Deprivation Can Result in Either Apoptosis or Differentiation." Journal of Neurochemistry **60**(5): 1783-1791.
Howell, N., et al. (2000). "Longitudinal analysis of the segregation of mtDNA mutations in heteroplasmic individuals." Journal of the Neurological Sciences **172**(1): 1-6.
Huang, T. T., et al. (2006). "Genetic modifiers of the phenotype of mice deficient in mitochondrial superoxide dismutase." Hum Mol Genet **15**(7): 1187-1194.
Hu, J. S., et al. (2017). "Coup-TF1 and Coup-TF2 control subtype and laminar identity of MGE-derived neocortical interneurons." Development **144**(15): 2837-2851.
Humble, M. M., et al. (2013). "Polg2 is essential for mammalian embryogenesis and is required for mtDNA maintenance." Human Molecular Genetics **22**(5): 1017-1025.
Hurd, T. R., et al. (2015). "Curly Encodes Dual Oxidase, Which Acts with Heme Peroxidase Curly Su to Shape the Adult Drosophila Wing." Plos Genetics **11**(11).

Hussein, M. R. (2005). "Apoptosis in the ovary: molecular mechanisms." Human Reproduction Update **11**(2): 162-178.
Hutchison, C. A., 3rd, et al. (1974). "Maternal inheritance of mammalian mitochondrial DNA." Nature **251**(5475): 536-538.
Hwang, L. H. and A. W. Murray (1997). "A novel yeast screen for mitotic arrest mutants identifies DOC1, a new gene involved in cyclin proteolysis." Mol Biol Cell **8**(10): 1877-1887.
Hyman, E. D. (1988). "A New Method of Sequencing DNA." Analytical Biochemistry **174**(2): 423-436.
Hyvarinen, A. K., et al. (2007). "The mitochondrial transcription termination factor mTERF modulates replication pausing in human mitochondrial DNA." Nucleic Acids Research **35**(19): 6458-6474.
Ibarra-Soria, X., et al. (2018). "Defining murine organogenesis at single-cell resolution reveals a role for the leukotriene pathway in regulating blood progenitor formation." Nature Cell Biology **20**(2): 127-+.
Ikui, A. E., et al. (2005). "Low concentrations of taxol cause mitotic delay followed by premature dissociation of p55CDC from mad2 and BubR1 and abrogation of the spindle checkpoint, leading to aneuploidy." Cell Cycle **4**(10): 1385-1388.
Imam, S. Z., et al. (2006). "Mitochondrial and nuclear DNA-repair capacity of various brain regions in mouse is altered in an age-dependent manner." Neurobiology of Aging **27**(8): 1129-1136.
Imbert, N., et al. (1995). "Abnormal Calcium Homeostasis in Duchenne Muscular-Dystrophy Myotubes Contracting in-Vitro." Cell Calcium **18**(3): 177-186.
Ishihara, N., et al. (2009). "Mitochondrial fission factor Drp1 is essential for embryonic development and synapse formation in mice." Nature Cell Biology **11**(8): 958-U114.
Israels, E. D. and L. G. Israels (2001). "The cell cycle." Stem Cells **19**(1): 88-91.
Jacobi, F. K., et al. (2001). "Segregation patterns and heteroplasmy prevalence in Leber's hereditary optic neuropathy." Investigative Ophthalmology & Visual Science **42**(6): 1208-1214.
Jazek, P. and A. Dlaskova (2019). "Dynamic of mitochondrial network, cristae, and mitochondrial nucleoids in pancreatic beta-cells." Mitochondrion **49**: 245-258.
Jenuth, J. P., et al. (1996). "Random genetic drift in the female germline explains the rapid segregation of mammalian mitochondrial DNA." Nature Genetics **14**(2): 146-151.
Jenuth, J. P., et al. (1997). "Tissue-specific selection for different mtDNA genotypes in heteroplasmic mice." Nature Genetics **16**(1): 93-95.
Jiang, S., et al. (2019). "TEFM regulates both transcription elongation and RNA processing in mitochondria." Embo Reports **20**(6).
Johnson, D. T., et al. (2007a). "Functional consequences of mitochondrial proteome heterogeneity." American Journal of Physiology-Cell Physiology **292**(2): C698-C707.
Johnson, D. T., et al. (2007b). "Tissue heterogeneity of the mammalian mitochondrial proteome." American Journal of Physiology-Cell Physiology **292**(2): C689-C697.
Johnston, I. G. and N. S. Jones (2016). "Evolution of Cell-to-Cell Variability in Stochastic, Controlled, Heteroplasmic mtDNA Populations." American Journal of Human Genetics **99**(5): 1150-1162.
Johnston, I. G., et al. (2015). "Stochastic modelling, Bayesian inference, and new in vivo measurements elucidate the debated mtDNA bottleneck mechanism." Elife **4**.
Jouaville, L. S., et al. (1999). "Regulation of mitochondrial ATP synthesis by calcium: Evidence for a long-term metabolic priming." Proceedings of the National Academy of Sciences of the United States of America **96**(24): 13807-13812.

Jourdain, A. A., et al. (2013). "GRSF1 Regulates RNA Processing in Mitochondrial RNA Granules." Cell Metabolism **17**(3): 399-410.
Kaila, V. R. I., et al. (2011). "A combined quantum chemical and crystallographic study on the oxidized binuclear center of cytochrome c oxidase." Biochimica Et Biophysica Acta-Bioenergetics **1807**(7): 769-778.
Kamer, K. J. and V. K. Mootha (2015). "The molecular era of the mitochondrial calcium uniporter." Nature Reviews Molecular Cell Biology **16**(9): 545-553.
Kanfer, G., et al. (2015). "Mitotic redistribution of the mitochondrial network by Miro and Cenp-F." Nature Communications **6**.
Kanki, T., et al. (2004). "Mitochondrial nucleoid and transcription factor A." Mitochondrial Pathogenesis: From Genes and Apoptosis to Aging and Disease **1011**: 61-68.
Karbowski, M. and R. J. Youle (2003). "Dynamics of mitochondrial morphology in healthy cells and during apoptosis." Cell Death Differ **10**(8): 870-880.
Karsenti, E. and I. Vernos (2001). "Cell cycle - The mitotic spindle: A self-made machine." Science **294**(5542): 543-547.
Kasamatsu, H., et al. (1971). "Novel Closed-Circular Mitochondrial DNA with Properties of a Replicating Intermediate." Proceedings of the National Academy of Sciences of the United States of America **68**(9): 2252-+.
Kauppila, J. H. K., et al. (2016). "A Phenotype-Driven Approach to Generate Mouse Models with Pathogenic mtDNA Mutations Causing Mitochondrial Disease." Cell Reports **16**(11): 2980-2990.
Kaushal, P. S., et al. (2014). "Cryo-EM structure of the small subunit of the mammalian mitochondrial ribosome." Proceedings of the National Academy of Sciences of the United States of America **111**(20): 7284-7289.
Kawai, T., et al. (2005). "IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction." Nat Immunol **6**(10): 981-988.
Kazak, L., et al. (2013). "Alternative translation initiation augments the human mitochondrial proteome." Nucleic Acids Research **41**(4): 2354-2369.
Khan, S. A. (1997). "Rolling-circle replication of bacterial plasmids." Microbiology and Molecular Biology Reviews **61**(4): 442-+.
Kiledjian, M. (2018). "Eukaryotic RNA 5 '-End NAD(+) Capping and DeNADding." Trends in Cell Biology **28**(6): 454-464.
Kim, D. K. and I. Mook-Jung (2019). "The role of cell type-specific mitochondrial dysfunction in the pathogenesis of Alzheimer's disease." BMB Rep **52**(12): 679-688.
Kimura, M. (1955). "Solution of a Process of Random Genetic Drift with a Continuous Model." Proc Natl Acad Sci U S A **41**(3): 144-150.
Kleele, T., et al. (2021). "Distinct fission signatures predict mitochondrial degradation or biogenesis." Nature **593**(7859): 435-439.
Klein, A. M., et al. (2015). "Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells." Cell **161**(5): 1187-1201.
Koc, A., et al. (2004). "Hydroxyurea arrests DNA replication by a mechanism that preserves basal dNTP pools." Journal of Biological Chemistry **279**(1): 223-230.
Koenig, M. K. (2008). "Presentation and diagnosis of mitochondrial disorders in children." Pediatr Neurol **38**(5): 305-313.
Kogelnik, A. M., et al. (1997). "MITOMAP: An update on the status of the human mitochondrial genome database." Nucleic Acids Research **25**(1): 196-199.
Konarikova, E., et al. (2020). "Current progress in the therapeutic options for mitochondrial disorders." Physiol Res **69**(6): 967-994.

Kopajtich, R., et al. (2014). "Mutations in GTPBP3 cause a mitochondrial translation defect associated with hypertrophic cardiomyopathy, lactic acidosis, and encephalopathy." American Journal of Human Genetics **95**(6): 708-720.
Korhonen, J. A., et al. (2003). "TWINKLE has 5 '-> 3 ' DNA helicase activity and is specifically stimulated by mitochondrial single-stranded DNA-binding protein." Journal of Biological Chemistry **278**(49): 48627-48632.
Korhonen, J. A., et al. (2004). "Reconstitution of a minimal mtDNA replisome in vitro." Embo Journal **23**(12): 2423-2429.
Kowald, A., et al. (2014). "Mitochondrial mutations and ageing: Can mitochondrial deletion mutants accumulate via a size based replication advantage?" Journal of Theoretical Biology **340**: 111-118.
Kowallik, K. V. and W. F. Martin (2021). "The origin of symbiogenesis: An annotated English translation of Mereschkowsky's 1910 paper on the theory of two plasma lineages." Biosystems **199**.
Krek, W. and J. A. DeCaprio (1995). "Cell synchronization." Methods Enzymol **254**: 114-124.
Kroemer, G., et al. (2007). "Mitochondrial membrane permeabilization in cell death." Physiological Reviews **87**(1): 99-163.
Kruman, I. I. (2004). "Why do neurons enter the cell cycle?" Cell Cycle **3**(6): 769-773.
Kucej, M. and R. A. Butow (2007). "Evolutionary tinkering with mitochondrial nucleoids." Trends in Cell Biology **17**(12): 586-592.
Kuhn, M. (1998). "The microtubule depolymerizing drugs nocodazole and colchicine inhibit the uptake of Listeria monocytogenes by P388D(1) macrophages." Fems Microbiology Letters **160**(1): 87-90.
Kujoth, G. C., et al. (2005). "Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging." Science **309**(5733): 481-484.
Kukat, C., et al. (2011). "Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA." Proceedings of the National Academy of Sciences of the United States of America **108**(33): 13534-13539.
Kukat, C., et al. (2015). "Cross-strand binding of TFAM to a single mtDNA molecule forms the mitochondrial nucleoid." Proceedings of the National Academy of Sciences of the United States of America **112**(36): 11288-11293.
Kulkarni, A., et al. (2019). "Beyond bulk: a review of single cell transcriptomics methodologies and applications." Current Opinion in Biotechnology **58**: 129-136.
Kunji, E. R. S., et al. (2016). "The transport mechanism of the mitochondrial ADP/ATP carrier (vol 1863, pg 2379, 2016)." Biochimica Et Biophysica Acta-Molecular Cell Research **1863**(12): 3169-3169.
Kunz, W. S. (2003). "Different metabolic properties of mitochondrial oxidative phosphorylation in different cell types--important implications for mitochondrial cytopathies." Exp Physiol **88**(1): 149-154.
Kwok, S. and R. Higuchi (1989). "Avoiding False Positives with Pcr." Nature **339**(6221): 237-238.
Lai, X. Q., et al. (2021). "SOX10 ablation severely impairs the generation of postmigratory neural crest from human pluripotent stem cells." Cell Death & Disease **12**(9).
Lane, N. and W. Martin (2010). "The energetics of genome complexity." Nature **467**(7318): 929-934.
Langen, P. and F. Hucho (2008). "Karl Lohmann and the discovery of ATP." Angewandte Chemie-International Edition **47**(10): 1824-1827.

Larsson, N. G. and D. A. Clayton (1995). "Molecular genetic aspects of human mitochondrial disorders." Annual Review of Genetics **29**: 151-178.
Larsson, N. G., et al. (1990). "Progressive Increase of the Mutated Mitochondrial-DNA Fraction in Kearns-Sayre Syndrome." Pediatric Research **28**(2): 131-136.
Latorre-Pellicer, A., et al. (2016). "Mitochondrial and nuclear DNA matching shapes metabolism and healthy ageing." Nature **535**(7613): 561-+.
Latorre-Pellicer, A., et al. (2019). "Regulation of Mother-to-Offspring Transmission of mtDNA Heteroplasmy." Cell Metab **30**(6): 1120-1130 e1125.
Le Campion, A., et al. (2002). "Naive T cells proliferate strongly in neonatal mice in response to self-peptide/self-MHC complexes." Proc Natl Acad Sci U S A **99**(7): 4538-4543.
Lechuga-Vieco, A. V., et al. (2020). "Cell identity and nucleo-mitochondrial genetic context modulate OXPHOS performance and determine somatic heteroplasmy dynamics." Science Advances **6**(31).
Lee, J. J., et al. (2018). "Basal mitophagy is widespread in Drosophila but minimally affected by loss of Pink1 or parkin." Journal of Cell Biology **217**(5): 1613-1622.
Lee, S., et al. (2007). "Cell cycle-dependent mitochondrial biogenesis and dynamics in mammalian cells." Biochem Biophys Res Commun **357**(1): 111-117.
Leeman, D. S., et al. (2018). "Lysosome activation clears aggregates and enhances quiescent neural stem cell activation during aging." Science **359**(6381): 1277-1283.
Legros, F., et al. (2004). "Organization and dynamics of human mitochondrial DNA." Journal of Cell Science **117**(13): 2653-2662.
Leister, D. (2005). "Origin, evolution and genetic effects of nuclear insertions of organelle DNA." Trends in Genetics **21**(12): 655-663.
Leuenberger, D., et al. (1999). "Different import pathways through the mitochondrial intermembrane space for inner membrane proteins." Embo Journal **18**(17): 4816-4822.
Lewis, S. M., et al. (2019). "Structure and function of the immune system in the spleen." Science Immunology **4**(33).
Li, J., et al. (2000). "Overexpression of p27(KIP1) induced cell cycle arrest in G(1) phase and subsequent apoptosis in HCC-9204 cell line." World J Gastroenterol **6**(4): 513-521.
Li, M. K., et al. (2015). "Extensive tissue-related and allele-related mtDNA heteroplasmy suggests positive selection for somatic mutations." Proceedings of the National Academy of Sciences of the United States of America **112**(8): 2491-2496.
Li, R., et al. (2003). "Identification and characterization of mouse MTO1 gene related to mitochondrial tRNA modification." Biochim Biophys Acta **1629**(1-3): 53-59.
Liao, H. and J. Thorner (1980). "Yeast mating pheromone alpha factor inhibits adenylate cyclase." Proc Natl Acad Sci U S A **77**(4): 1898-1902.
Liesa, M., et al. (2009). "Mitochondrial Dynamics in Mammalian Health and Disease." Physiological Reviews **89**(3): 799-845.
Lim, J. F., et al. (2016). "Isolation and Activation of Murine Lymphocytes." Jove-Journal of Visualized Experiments(116).
Lin, M. T. and M. F. Beal (2006). "Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases." Nature **443**(7113): 787-795.
Lin, Y. F., et al. (2016). "Maintenance and propagation of a deleterious mitochondrial genome by the mitochondrial unfolded protein response." Nature **533**(7603): 416-+.
Liu, X. S., et al. (1996). "Induction of apoptotic program in cell-free extracts: Requirement for dATP and cytochrome c." Cell **86**(1): 147-157.
Longchamps, R. J., et al. (2020). "Evaluation of mitochondrial DNA copy number estimation techniques." PLoS One **15**(1): e0228166.

Longley, M. J., et al. (1998). "Identification of 5'-deoxyribose phosphate lyase activity in human DNA polymerase gamma and its role in mitochondrial base excision repair in vitro." Proc Natl Acad Sci U S A **95**(21): 12244-12248.
Longley, M. J., et al. (2001). "The fidelity of human DNA polymerase gamma with and without exonucleolytic proofreading and the p55 accessory subunit." Journal of Biological Chemistry **276**(42): 38555-38562.
Lopez, A. L., et al. (2020). "Embryonic Mouse Cardiodynamic OCT Imaging." Journal of Cardiovascular Development and Disease **7**(4).
Lopez, J. V., et al. (1994). "Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat." J Mol Evol **39**(2): 174-190.
Lopez, M. F., et al. (2000). "High-throughput profiling of the mitochondrial proteome using affinity fractionation and automation." Electrophoresis **21**(16): 3427-3440.
Ma, H. and P. H. O'Farrell (2015). "Selections that isolate recombinant mitochondrial genomes in animals." Elife **4**.
Ma, H. S. and P. H. O'Farrell (2016). "Selfish drive can trump function when animal mitochondrial genomes compete." Nature Genetics **48**(7): 798-+.
Ma, H. S., et al. (2014). "Transmission of mitochondrial mutations and action of purifying selection in Drosophila melanogaster." Nature Genetics **46**(4): 393-+.
Macosko, E. Z., et al. (2015). "Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets." Cell **161**(5): 1202-1214.
Madsen, C. S., et al. (1993). "Protein binding to a single termination-associated sequence in the mitochondrial DNA D-loop region." Mol Cell Biol **13**(4): 2162-2171.
Mai, N., et al. (2017). "The process of mammalian mitochondrial protein synthesis." Cell and Tissue Research **367**(1): 5-20.
Maier, D., et al. (2001). "Mitochondrial single-stranded DNA-binding protein is required for mitochondrial DNA replication and development in Drosophila melanogaster." Molecular Biology of the Cell **12**(4): 821-830.
Malumbres, M. and M. Barbacid (2001). "To cycle or not to cycle: a critical decision in cancer." Nat Rev Cancer **1**(3): 222-231.
Mancuso, M., et al. (2015). "Redefining phenotypes associated with mitochondrial DNA single deletion." J Neurol **262**(5): 1301-1309.
Mandal, S., et al. (2005). "Mitochondrial regulation of cell cycle progression during development as by the tenured mutation revealed in Drosophila." Developmental Cell **9**(6): 843-854.
Mandal, S., et al. (2010). "Metabolic control of G1-S transition: cyclin E degradation by p53-induced activation of the ubiquitin-proteasome system." Journal of Cell Biology **188**(4): 473-479.
Martin, N. C. (1994). Organellar tRNAs: Biosynthesis and Function. tRNA**:** 127-140.
Martin, W. and M. Muller (1998). "The hydrogen hypothesis for the first eukaryote." Nature **392**(6671): 37-41.
Martinez-Cayuela, M. (1995). "Oxygen-Free Radicals and Human-Disease." Biochimie **77**(3): 147-161.
Martinez-Reyes, I. and N. S. Chandel (2020). "Mitochondrial TCA cycle metabolites control physiology and disease." Nature Communications **11**(1).
Maruyama, T. and M. Kimura (1980). "Genetic-Variability and Effective Population-Size When Local Extinction and Recolonization of Sub-Populations Are Frequent." Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences **77**(11): 6710-6714.
Matoba, S., et al. (2006). "p53 regulates mitochondrial respiration." Science **312**(5780): 1650-1653.

Matsuoka, S., et al. (2000). "Ataxia telangiectasia-mutated phosphorylates Chk2 in vivo and in vitro." Proceedings of the National Academy of Sciences of the United States of America **97**(19): 10389-10394.
McFarland, R., et al. (2010). "A neurological perspective on mitochondrial disease." Lancet Neurol **9**(8): 829-840.
McKinney, E. A. and M. T. Oliveira (2013). "Replicating animal mitochondrial DNA." Genet Mol Biol **36**(3): 308-315.
McWilliams, T. G., et al. (2016). "mito-QC illuminates mitophagy and mitochondrial architecture in vivo." Journal of Cell Biology **214**(3): 333-345.
Mears, J. A. (2014). Mitochondrial Biogenesis and Quality Control. The Structural Basis of Biological Energy Generation. M. F. Hohmann-Marriott. Dordrecht, Springer Netherlands**:** 451-476.
Meirelles, F. V. and L. C. Smith (1997). "Mitochondrial genotype segregation in a mouse heteroplasmic lineage produced by embryonic karyoplast transplantation." Genetics **145**(2): 445-451.
Melton, T. (2004). "Mitochondrial DNA Heteroplasmy." Forensic Sci Rev **16**(1): 1-20.
Mendenhall, M. D., et al. (1988). "Dominant Negative Protein-Kinase Mutations That Confer a G1 Arrest Phenotype." Proceedings of the National Academy of Sciences of the United States of America **85**(12): 4426-4430.
Mennuni, M., et al. (2022). "Metabolic resistance to the inhibition of mitochondrial transcription revealed by CRISPR-Cas9 screen." Embo Reports **23**(1).
Meyer, R. R. and P. S. Laine (1990). "The Single-Stranded DNA-Binding Protein of Escherichia-Coli." Microbiological Reviews **54**(4): 342-380.
Meylan, E., et al. (2005). "Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus." Nature **437**(7062): 1167-1172.
Mimitou, E. P., et al. (2021). "Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells." Nat Biotechnol **39**(10): 1246-1258.
Minczuk, M., et al. (2011). "TEFM (c17orf42) is necessary for transcription of human mtDNA." Nucleic Acids Research **39**(10): 4284-4299.
Mitchell, P. (1961). "Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism." Nature **191**: 144-148.
Mitra, K., et al. (2009). "A hyperfused mitochondrial state achieved at G(1)-S regulates cyclin E buildup and entry into S phase." Proceedings of the National Academy of Sciences of the United States of America **106**(29): 11960-11965.
Mitra, K., et al. (2009). "A hyperfused mitochondrial state achieved at G1-S regulates cyclin E buildup and entry into S phase." Proc Natl Acad Sci U S A **106**(29): 11960-11965.
Monnot, S., et al. (2011). "Segregation of mtDNA Throughout Human Embryofetal Development: m.3243A > G as a Model System." Human Mutation **32**(1): 116-125.
Montier, L. L. C., et al. (2009). "Number matters: control of mammalian mitochondrial DNA copy number." Journal of Genetics and Genomics **36**(3): 125-131.
Montoya, J., et al. (1982). "Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial DNA." Proc Natl Acad Sci U S A **79**(23): 7195-7199.
Montoya, J., et al. (1983). "The pattern of transcription of the human mitochondrial rRNA genes reveals two overlapping transcription units." Cell **34**(1): 151-159.
Moraes, C. T., et al. (1999). "Mechanisms of human mitochondrial DNA maintenance: The determining role of primary sequence and length over function." Molecular Biology of the Cell **10**(10): 3345-3356.
Muller, H. J. (1964). "The Relation of Recombination to Mutational Advance." Mutation Research **1**(1): 2-9.

Munsky, B., et al. (2012). "Using Gene Expression Noise to Understand Gene Regulation." Science **336**(6078): 183-187.
Murphy, M. P. (2009). "How mitochondria produce reactive oxygen species." Biochemical Journal **417**: 1-13.
Murray, A. (1994). "Cell-Cycle Checkpoints." Current Opinion in Cell Biology **6**(6): 872-876.
Nagaike, T., et al. (2005). "Human mitochondrial mRNAs are stabilized with polyadenylation regulated by mitochondria-specific poly(A) polymerase and polynucleotide phosphorylase." Journal of Biological Chemistry **280**(20): 19721-19727.
Nakada, K., et al. (2009). "Mitochondrial functional complementation in mitochondrial DNA-based diseases." International Journal of Biochemistry & Cell Biology **41**(10): 1907-1913.
Nass, M. M. (1966). "The circularity of mitochondrial DNA." Proc Natl Acad Sci U S A **56**(4): 1215-1222.
Nesbitt, V., et al. (2014). "A national perspective on prenatal testing for mitochondrial disease." European Journal of Human Genetics **22**(11): 1255-1259.
Newmeyer, D. D. and S. Ferguson-Miller (2003). "Mitochondria: Releasing power for life and unleashing the machineries of death." Cell **112**(4): 481-490.
Nicholls, T. J. and M. Minczuk (2014). "In D-loop: 40 years of mitochondrial 7S DNA." Exp Gerontol **56**: 175-181.
Nikoletopoulou, V., et al. (2013). "Crosstalk between apoptosis, necrosis and autophagy." Biochimica Et Biophysica Acta-Molecular Cell Research **1833**(12): 3448-3459.
Niyazov, D. M., et al. (2016). "Primary Mitochondrial Disease and Secondary Mitochondrial Dysfunction: Importance of Distinction for Diagnosis and Treatment." Mol Syndromol **7**(3): 122-137.
Nouspikel, T. and P. C. Hanawalt (2002). "DNA repair in terminally differentiated cells." DNA Repair **1**(1): 59-75.
Nyren, P. (2015). "The History of Pyrosequencing (R)." Pyrosequencing: Methods and Protocols, 2nd Edition **1315**: 3-15.
Ojala, D., et al. (1981). "Transfer-Rna Punctuation Model of Rna Processing in Human Mitochondria." Nature **290**(5806): 470-474.
Olausson, K. H., et al. (2014). "Prominin-1 (CD133) Defines Both Stem and Non-Stem Cell Populations in CNS Development and Gliomas." PLoS One **9**(9).
Osawa, S., et al. (1992). "Recent evidence for evolution of the genetic code." Microbiol Rev **56**(1): 229-264.
Otten, A. B. C., et al. (2016). "Differences in Strength and Timing of the mtDNA Bottleneck between Zebrafish Germline and Non-germline Cells." Cell Reports **16**(3): 622-630.
Pacheu-Grau, D., et al. (2018). "Mitochondrial dysfunction and its role in tissue-specific cellular stress." Cell Stress **2**(8): 184-199.
Pagliarini, D. J. and J. Rutter (2013). "Hallmarks of a new era in mitochondrial biochemistry." Genes & Development **27**(24): 2615-2627.
Pagliarini, D. J., et al. (2008). "A mitochondrial protein compendium elucidates complex I disease biology." Cell **134**(1): 112-123.
Palade, G. E. (1952). "The Fine Structure of Mitochondria." Anatomical Record **114**(3): 427-451.
Palade, G. E. (1953). "An Electron Microscope Study of the Mitochondrial Structure." Journal of Histochemistry & Cytochemistry **1**(4): 188-211.
Pande, S. V., et al. (1993). "Carnitine-Acylcarnitine Translocase Deficiency with Severe Hypoglycemia and Atrioventricular-Block - Translocase Assay in Permeabilized Fibroblasts." Journal of Clinical Investigation **91**(3): 1247-1252.

Pardee, A. B. (1974). "A restriction point for control of normal animal cell proliferation." Proc Natl Acad Sci U S A **71**(4): 1286-1290.
Park, C. B., et al. (2007). "MTERF3 is a negative regulator of mammalian mtDNA transcription." Cell **130**(2): 273-285.
Park, S. Y., et al. (2012). "Mimosine arrests the cell cycle prior to the onset of DNA replication by preventing the binding of human Ctf4/And-1 to chromatin via Hif-1alpha activation in HeLa cells." Cell Cycle **11**(4): 761-766.
Payne, B. A. I., et al. (2013). "Universal heteroplasmy of human mitochondrial DNA." Human Molecular Genetics **22**(2): 384-390.
Pearce, S. F., et al. (2017). "Maturation of selected human mitochondrial tRNAs requires deadenylation." Elife **6**.
Pearce, S. F., et al. (2017). "Regulation of Mammalian Mitochondrial Gene Expression: Recent Advances." Trends in Biochemical Sciences **42**(8): 625-639.
Perry, C., et al. (2005). "Mimosine attenuates serine hydroxymethyltransferase transcription by chelating zinc. Implications for inhibition of DNA replication." Journal of Biological Chemistry **280**(1): 396-400.
Pestryakov, P. E. and O. I. Lavrik (2008). "Mechanisms of Single-Stranded DNA-Binding Protein Functioning in Cellular DNA Metabolism." Biochemistry-Moscow **73**(13): 1388-1404.
Peter, B. and M. Falkenberg (2020). "TWINKLE and Other Human Mitochondrial DNA Helicases: Structure, Function and Disease." Genes **11**(4).
Pickles, S., et al. (2018). "Mitophagy and Quality Control Mechanisms in Mitochondrial Maintenance." Current Biology **28**(4): R170-R185.
Piekna-Przybylska, D., et al. (2008). "The 3D rRNA modification maps database: with interactive tools for ribosome analysis." Nucleic Acids Research **36**: D178-D183.
Pijuan-Sala, B., et al. (2019). "A single-cell molecular map of mouse gastrulation and early organogenesis." Nature **566**(7745): 490-+.
Pikkarainen, S., et al. (2004). "GATA transcription factors in the developing and adult heart." Cardiovascular Research **63**(2): 196-207.
Piko, L. and K. D. Taylor (1987). "Amounts of Mitochondrial-DNA and Abundance of Some Mitochondrial Gene Transcripts in Early Mouse Embryos." Developmental Biology **123**(2): 364-374.
Pinton, P., et al. (2008). "Calcium and apoptosis: ER-mitochondria Ca2+ transfer in the control of apoptosis." Oncogene **27**(50): 6407-6418.
Pelc, S. R., and Howard, A. (1955). Effect of Various Doses of X-Rays on the Number of Cells Synthesizing Deoxyribonucleic Acid. Radiation Research, **3**(2), 135–142.
Posse, V., et al. (2015). "TEFM is a potent stimulator of mitochondrial transcription elongation in vitro." Nucleic Acids Research **43**(5): 2615-2624.
Potter, J. A., et al. (2008). "Crystal structure of human IPS-1/MAVS/VISA/Cardif caspase activation recruitment domain." Bmc Structural Biology **8**.
Powel, C. A., et al. (2015). "TRMT5 Mutations Cause a Defect in Post-transcriptional Modification of Mitochondrial tRNA Associated with Multiple Respiratory-Chain Deficiencies." American Journal of Human Genetics **97**(2): 319-328.
Prakadan, S. M., et al. (2017). "Scaling by shrinking: empowering single-cell 'omics' with microfluidic devices." Nature Reviews Genetics **18**(6).
Protasoni, M. and M. Zeviani (2021). "Mitochondrial Structure and Bioenergetics in Normal and Disease Conditions." International Journal of Molecular Sciences **22**(2).
Pyle, A., et al. (2007). "Depletion of mitochondrial DNA in leucocytes harbouring the 3243A -> G mtDNA mutation." Journal of Medical Genetics **44**(1): 69-74.

Pyle, A., et al. (2015). "Extreme-Depth Re-sequencing of Mitochondrial DNA Finds No Evidence of Paternal Transmission in Humans." Plos Genetics **11**(5).
Rahman, S., et al. (2001). "Decrease of 3243 A-->G mtDNA mutation from blood in MELAS syndrome: a longitudinal study." American Journal of Human Genetics **68**(1): 238-240.
Raj, A. and A. van Oudenaarden (2008). "Nature, nurture, or chance: stochastic gene expression and its consequences." Cell **135**(2): 216-226.
Rajasimha, H. K., et al. (2008). "Selection against pathogenic mtDNA mutations in a stem cell population leads to the loss of the 3243A-->G mutation in blood." American Journal of Human Genetics **82**(2): 333-343.
Rambold, A. S. and E. L. Pearce (2018). "Mitochondrial Dynamics at the Interface of Immune Cell Metabolism and Function." Trends in Immunology **39**(1): 6-18.
Rao, P. N. and J. Engelberg (1965). "Hela Cells: Effects of Temperature on the Life Cycle." Science **148**(3673): 1092-1094.
Reyes, A., et al. (2005). "Bidirectional replication initiates at sites throughout the mitochondrial genome of birds." Journal of Biological Chemistry **280**(5): 3242-3250.
Reyes, A., et al. (2011). "Actin and myosin contribute to mammalian mitochondrial DNA maintenance." Nucleic Acids Research **39**(12): 5098-5108.
Reyes, A., et al. (2013). "Mitochondrial DNA replication proceeds via a 'bootlace' mechanism involving the incorporation of processed transcripts." Nucleic Acids Research **41**(11): 5837-5850.
Rinaldo, P. and D. Matern (2002). "Fatty acid oxidation disorders." Annual Review of Physiology **64**: 477-502.
Romero-Garcia, S. and H. Prado-Garcia (2019). "Mitochondrial calcium: Transport and modulation of cellular processes in homeostasis and cancer (Review)." International Journal of Oncology **54**(4): 1155-1167.
Ronaghi, M. (2001). "Pyrosequencing sheds light on DNA sequencing." Genome Research **11**(1): 3-11.
Ronchi, J. A., et al. (2013). "A spontaneous mutation in the nicotinamide nucleotide transhydrogenase gene of C57BL/6J mice results in mitochondrial redox abnormalities." Free Radic Biol Med **63**: 446-456.
Rorbach, J., et al. (2014). "Polyadenylation in Bacteria and Organelles." Polyadenylation: Methods and Protocols **1125**: 211-227.
Ross, J. M. (2011). "Visualization of Mitochondrial Respiratory Function using Cytochrome C Oxidase/Succinate Dehydrogenase (COX/SDH) Double-labeling Histochemistry." Jove-Journal of Visualized Experiments (57).
Rotig, A., et al. (1990). "Pearson's marrow-pancreas syndrome. A multisystem mitochondrial disorder in infancy." Journal of Clinical Investigation **86**(5): 1601-1608.
Ruhanen, H., et al. (2011). "Involvement of DNA ligase III and ribonuclease H1 in mitochondrial DNA replication in cultured human cells." Biochimica Et Biophysica Acta-Molecular Cell Research **1813**(12): 2000-2007.
Rusecka, J., et al. (2018). "Nuclear genes involved in mitochondrial diseases caused by instability of mitochondrial DNA." Journal of Applied Genetics **59**(1): 43-57.
Russell, O. M., et al. (2018). "Preferential amplification of a human mitochondrial DNA deletion in vitro and in vivo." Scientific Reports **8**.
Saikawa, Y., et al. (2001). "Cyclin D1 antisense oligonucleotide inhibits cell growth stimulated by epidermal growth factor and induces apoptosis of gastric cancer cells." Japanese Journal of Cancer Research **92**(10): 1102-1109.
Sakaue-Sawano, A., et al. (2008). "Visualizing spatiotemporal dynamics of multicellular cell-cycle progression." Cell **132**(3): 487-498.

Saliba, A. E., et al. (2014). "Single-cell RNA-seq: advances and future challenges." Nucleic Acids Research **42**(14): 8845-8860.
Sallevelt, S. C. E. H., et al. (2017). "De novo mtDNA point mutations are common and have a low recurrence risk." Journal of Medical Genetics **54**(2): 114-124.
Samuels, D. C., et al. (2010). "Reassessing evidence for a postnatal mitochondrial genetic bottleneck." Nature Genetics **42**(6): 471-472; author reply 472-473.
Samuels, D. C., et al. (2013). "Recurrent Tissue-Specific mtDNA Mutations Are Common in Humans." Plos Genetics **9**(11).
Sanger, F., et al. (1977). "DNA Sequencing with Chain-Terminating Inhibitors." Proceedings of the National Academy of Sciences of the United States of America **74**(12): 5463-5467.
Santo, L., et al. (2015). "Targeting Cyclin-Dependent Kinases and Cell Cycle Progression in Human Cancers." Seminars in Oncology **42**(6): 788-800.
Santos, J. M., et al. (2011). "Mitochondrial biogenesis and the development of diabetic retinopathy." Free Radical Biology and Medicine **51**(10): 1849-1860.
Sato, M. and K. Sato (2012). "Maternal inheritance of mitochondrial DNA: degradation of paternal mitochondria by allogeneic organelle autophagy, allophagy." Autophagy **8**(3): 424-425.
Saur, A. L., et al. (2021). "Formation of the node of Ranvier by Schwann cells is under control of transcription factor Sox10." Glia **69**(6): 1464-1477.
Schapira, A. H. (2006). "Mitochondrial disease." Lancet **368**(9529): 70-82.
Schara, U., et al. (2011). "Acute liver failure with subsequent cirrhosis as the primary manifestation of TRMU mutations." Journal of Inherited Metabolic Disease **34**(1): 197-201.
Schieke, S. M., et al. (2008). "Coordination of mitochondrial bioenergetics with G1 phase cell cycle progression." Cell Cycle **7**(12): 1782-1787.
Schlame, M. (2021). "Protein crowding in the inner mitochondrial membrane." Biochimica Et Biophysica Acta-Bioenergetics **1862**(1).
Schon, E. A., et al. (1997). "Mitochondrial DNA mutations and pathogenesis." Journal of Bioenergetics and Biomembranes **29**(2): 131-149.
Schon, K. R., et al. (2020). "Mitochondrial Diseases: A Diagnostic Revolution." Trends in Genetics **36**(9): 702-717.
Schurr, T. G., et al. (1999). "Mitochondrial DNA variation in Koryaks and Itel'men: Population replacement in the Okhotsk Sea Bering Sea region during the Neolithic." American Journal of Physical Anthropology **108**(1): 1-39.
See, P., et al. (2018). "A Single-Cell Sequencing Guide for Immunologists." Frontiers in Immunology **9**.
Seth, R. B., et al. (2005). "Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3." Cell **122**(5): 669-682.
Shalem, O., et al. (2014). "Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells." Science **343**(6166): 84-87.
Sharpley, M. S., et al. (2012). "Heteroplasmy of Mouse mtDNA Is Genetically Unstable and Results in Altered Behavior and Cognition." Cell **151**(2): 333-343.
Shereda, R. D., et al. (2008). "SSB as an Organizer/Mobilizer of Genome Maintenance Complexes." Critical Reviews in Biochemistry and Molecular Biology **43**(5): 289-318.
Shi, L. and B. P. Tu (2015). "Acetyl-CoA and the regulation of metabolism: mechanisms and consequences." Current Opinion in Cell Biology **33**: 125-131.
Shin, J. H., et al. (2011). "PARIS (ZNF746) repression of PGC-1alpha contributes to neurodegeneration in Parkinson's disease." Cell **144**(5): 689-702.

Shokolenko, I., et al. (2009). "Oxidative stress induces degradation of mitochondrial DNA." Nucleic Acids Research **37**(8): 2539-2548.
Shoshan-Barmatz, V., et al. (2010). "VDAC, a multi-functional mitochondrial protein regulating cell life and death." Molecular Aspects of Medicine **31**(3): 227-285.
Shuster, R. C., et al. (1988). "Mitochondrial DNA in anucleate human blood cells." Biochem Biophys Res Commun **155**(3): 1360-1365.
Sickmann, A., et al. (2003). "The proteome of Saccharomyces cerevisiae mitochondria." Proceedings of the National Academy of Sciences of the United States of America **100**(23): 13207-13212.
Siekevitz, P. (1957). "Powerhouse of the Cell." Scientific American **197**(1): 131-&.
Sies, H. and D. P. Jones (2020). "Reactive oxygen species (ROS) as pleiotropic physiological signalling agents." Nature Reviews Molecular Cell Biology **21**(7): 363-383.
Silva-Pinheiro, P. and M. Minczuk (2022). "The potential of mitochondrial genome engineering." Nat Rev Genet **23**(4): 199-214.
Simoes-Costa, M. and M. E. Bronner (2015). "Establishing neural crest identity: a gene regulatory recipe." Development **142**(2): 242-257.
Singh, A. and Y. J. Xu (2016). "The Cell Killing Mechanisms of Hydroxyurea." Genes **7**(11).
Sissaoui, S., et al. (2020). "Genomic Characterization of Endothelial Enhancers Reveals a Multifunctional Role for NR2F2 in Regulation of Arteriovenous Gene Expression." Circulation Research **126**(7): 875-888.
Sjostrand, F. S. (1956). "The Ultrastructure of Cells as Revealed by the Electron Microscope." International Review of Cytology-a Survey of Cell Biology **5**: 455-533.
Skeen, J. E., et al. (2006). "Akt deficiency impairs normal cell proliferation and suppresses oncogenesis in a p53-independent and mTORC1-dependent manner." Cancer Cell **10**(4): 269-280.
Smeitink, J. A., et al. (2006). "Mitochondrial medicine: a metabolic perspective on the pathology of oxidative phosphorylation disorders." Cell Metab **3**(1): 9-13.
Sokol, A. M., et al. (2014). "Mitochondrial protein translocases for survival and wellbeing." Febs Letters **588**(15): 2484-2495.
Solignac, M., et al. (1983). "Mitochondrial-DNA Heteroplasmy in Drosophila-Mauritiana." Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences **80**(22): 6942-6946.
Starkov, A. A. (2008). "The Role of Mitochondria in Reactive Oxygen Species Metabolism and Signaling." Mitochondria and Oxidative Stress in Neurodegenerative Disorders **1147**: 37-52.
Steffann, J., et al. (2006). "Analysis of mtDNA variant segregation during early human embryonic development: a tool for successful NARP preimplantation diagnosis." Journal of Medical Genetics **43**(3): 244-247.
Stegle, O., et al. (2015). "Computational and analytical challenges in single-cell transcriptomics." Nature Reviews Genetics **16**(3): 133-145.
Stenton, S. L. and H. Prokisch (2018). "Advancing genomic approaches to the molecular diagnosis of mitochondrial disease." Mitochondrial Diseases **62**(3): 399-408.). "Advancing genomic approaches to the molecular diagnosis of mitochondrial disease." Mitochondrial Diseases **62**(3): 399-408.
Stewart, J. B. and N. G. Larsson (2014). "Keeping mtDNA in Shape between Generations." Plos Genetics **10**(10).
Stewart, J. B. and P. F. Chinnery (2015). "The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease." Nature Reviews Genetics **16**(9): 530-542.

Stewart, J. B. and P. F. Chinnery (2021). "Extreme heterogeneity of human mitochondrial DNA from organelles to populations." Nature Reviews Genetics **22**(2): 106-118.
Stewart, J. B., et al. (2008). "Strong purifying selection in transmission of mammalian mitochondrial DNA." Plos Biology **6**(1): 63-71.
Sturman, J. A., et al. (1970). "Absence of Cystathionase in Human Fetal Liver - Is Cystine Essential." Science **169**(3940): 74-&.
Sun, N., et al. (2015). "Measuring In Vivo Mitophagy." Molecular Cell **60**(4): 685-696.
Suomalainen, A. and B. J. Battersby (2018). "Mitochondrial diseases: the contribution of organelle stress responses to pathology." Nat Rev Mol Cell Biol **19**(2): 77-92.
Susin, S. A., et al. (1999). "Molecular characterization of mitochondrial apoptosis inducing factor." Nature **397**(6718): 441-446.
Suzuki, T., et al. (2011). "Human mitochondrial tRNAs: biogenesis, function, structural aspects, and diseases." Annu Rev Genet **45**: 299-329.
Szczesny, R. J., et al. (2010). "Human mitochondrial RNA turnover caught in flagranti: involvement of hSuv3p helicase in RNA surveillance." Nucleic Acids Research **38**(1): 279-298.
Szklarczyk, D., et al. (2019). "STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets." Nucleic Acids Research **47**(D1): D607-D613.
Tadi, S. K., et al. (2016). "Microhomology-mediated end joining is the principal mediator of double-strand break repair during mitochondrial DNA lesions." Molecular Biology of the Cell **27**(2): 223-235.
Taguchi, N., et al. (2007). "Mitotic phosphorylation of dynamin-related GTPase Drp1 participates in mitochondrial fission." J Biol Chem **282**(15): 11521-11529.
Takamatsu, C., et al. (2002). "Regulation of mitochondrial D-loops by transcription factor A and single-stranded DNA-binding protein." Embo Reports **3**(5): 451-456.
Tarnopolsky, M. A., et al. (2017). "Trmt5 Mutations Are Associated with Features of Complex Hereditary Spastic Paraparesis." Neurology **89**(21): 2210-2211.
Tasic, B., et al. (2018). "Shared and distinct transcriptomic cell types across neocortical areas." Nature **563**(7729): 72-+.
Tauber, J., et al. (2013). "Distribution of mitochondrial nucleoids upon mitochondrial network fragmentation and network reintegration in HEPG2 cells." International Journal of Biochemistry & Cell Biology **45**(3): 593-603.
Taylor, R. W. and D. M. Turnbull (2005). "Mitochondrial DNA mutations in human disease." Nature Reviews Genetics **6**(5): 389-402.
Taylor, R. W. and D. M. Turnbull (2005). "Mitochondrial DNA mutations in human disease." Nature Reviews Genetics **6**(5): 389-402.
Temperley, R. J., et al. (2010). "Human mitochondrial mRNAs-like members of all families, similar but different." Biochimica Et Biophysica Acta-Bioenergetics **1797**(6-7): 1081-1085.
tHart, L. M., et al. (1996). "Heteroplasmy levels of a mitochondrial gene mutation associated with diabetes mellitus decrease in leucocyte DNA upon aging." Human Mutation **7**(3): 193-197.
Thompson, K., et al. (2016). "Recurrent De Novo Dominant Mutations in SLC25A4 Cause Severe Early-Onset Mitochondrial Disease and Loss of Mitochondrial DNA Copy Number (vol 99, pg 860, 2016)." American Journal of Human Genetics **99**(6): 1405-1405.
Thorslund, T., et al. (2002). "Repair of 8-oxoG is slower in endogenous nuclear genes than in mitochondrial DNA and is without strand bias." DNA Repair **1**(4): 261-273.

Tilly, J. L. and K. I. Tilly (1995). "Inhibitors of Oxidative Stress Mimic the Ability of Follicle-Stimulating-Hormone to Suppress Apoptosis in Cultured Rat Ovarian Follicles." Endocrinology **136**(1): 242-252.
Tilokani, L., et al. (2018). "Mitochondrial dynamics: overview of molecular mechanisms." Mitochondrial Diseases **62**(3): 341-360.
To, T. L., et al. (2019). "A Compendium of Genetic Modifiers of Mitochondrial Dysfunction Reveals Intra-organelle Buffering." Cell **179**(5): 1222-+.
Tomecki, R., et al. (2004). "Identification of a novel human nuclear-encoded mitochondrial poly(A) polymerase." Nucleic Acids Research **32**(20): 6001-6014.
Torroni, A., et al. (1997). "Haplotype and phylogenetic analyses suggest that one European-specific mtDNA background plays a role in the expression of Leber hereditary optic neuropathy by increasing the penetrance of the primary mutations 11778 and 14484." American Journal of Human Genetics **60**(5): 1107-1121.
Tostes, K., et al. (2022). "Autophagy deficiency abolishes liver mitochondrial DNA segregation." Autophagy.
Trainor, P. A. (2005). "Specification of neural crest cell formation and migration in mouse embryos." Seminars in Cell & Developmental Biology **16**(6): 683-693.
Trifunovic, A., et al. (2004). "Premature ageing in mice expressing defective mitochondrial DNA polymerase." Nature **429**(6990): 417-423.
Tsai, F. C., et al. (2014). "A polarized Ca2+, diacylglycerol and STIM1 signalling system regulates directed cell migration." Nature Cell Biology **16**(2): 133-144.
Tu, Y. T. and A. Barrientos (2015). "The Human Mitochondrial DEAD-Box Protein DDX28 Resides in RNA Granules and Functions in Mitoribosome Assembly." Cell Reports **10**(6): 854-864.
Uchiyama, M., et al. (1997). "A novel mutant allele of Schizosaccharomyces pombe rad26 defective in monitoring S-phase progression to prevent premature mitosis." Molecular and Cellular Biology **17**(6): 3103-3115.
Uusimaa, J., et al. (2011). "Reversible infantile respiratory chain deficiency is a unique, genetically heterogenous mitochondrial disease." Journal of Medical Genetics **48**(10): 660-668.
Vafai, S. B. and V. K. Mootha (2012). "Mitochondrial disorders as windows into an ancient organelle." Nature **491**(7424): 374-383.
van den Ameele, J., et al. (2020). "Mitochondrial heteroplasmy beyond the oocyte bottleneck." Semin Cell Dev Biol **97**: 156-166.
van der Giezen, M. (2011). "Mitochondria and the Rise of Eukaryotes." Bioscience **61**(8): 594-601.
Van Haute, L., et al. (2015). "Mitochondrial transcript maturation and its disorders." Journal of Inherited Metabolic Disease **38**(4): 655-680.
Van Haute, L., et al. (2016). "Deficient methylation and formylation of mt-tRNA(Met) wobble cytosine in a patient carrying mutations in NSUN3." Nature Communications **7**: 12039.
Vinograd, J., et al. (1963). "The bouyant behavior of viral and bacterial DNA in alkaline CsCl." Proc Natl Acad Sci U S A **49**: 12-17.
Vogtle, F. N., et al. (2009). "Global Analysis of the Mitochondrial N-Proteome Identifies a Processing Peptidase Critical for Protein Stability." Cell **139**(2): 428-439.
von Heijne, G. (1986). "Why mitochondria need a genome." Febs Letters **198**(1): 1-4.
Wai, T., et al. (2008). "The mitochondrial DNA genetic bottleneck results from replication of a subpopulation of genomes." Nature Genetics **40**(12): 1484-1488.
Wai, T., et al. (2010). "The role of mitochondrial DNA copy number in mammalian fertility." Biol Reprod **83**(1): 52-62.

Wakabayashi, J., et al. (2009). "The dynamin-related GTPase Drp1 is required for embryonic and brain development in mice." Journal of Cell Biology **186**(6): 805-816.
Walberg, M. W. and D. A. Clayton (1981). "Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA." Nucleic Acids Res **9**(20): 5411-5421.
Walker, J. E. (2013). "The ATP synthase: the understood, the uncertain and the unknown." Biochem Soc Trans **41**(1): 1-16.
Walker, T. L., et al. (2013). "Prominin-1 Allows Prospective Isolation of Neural Stem Cells from the Adult Murine Hippocampus." Journal of Neuroscience **33**(7): 3010-3024.
Wang, C. G., et al. (2006). "Cyclin D1 repression of nuclear respiratory factor 1 integrates nuclear DNA synthesis and mitochondrial function." Proceedings of the National Academy of Sciences of the United States of America **103**(31): 11567-11572.
Wang, K., et al. (2012). "CD19: a biomarker for B cell development, lymphoma diagnosis and therapy." Exp Hematol Oncol **1**(1): 36.
Wang, Y. S. and D. F. Bogenhagen (2006). "Human mitochondrial DNA nucleoids are linked to protein folding machinery and metabolic enzymes at the mitochondrial inner membrane." Journal of Biological Chemistry **281**(35): 25791-25802.
Wang, Z. M., et al. (2010). "Specific metabolic rates of major organs and tissues across adulthood evaluation by mechanistic model of resting energy expenditure." American Journal of Clinical Nutrition **92**(6): 1369-1377.
Wanrooij, S., et al. (2008). "Human mitochondrial RNA polymerase primes lagging-strand DNA synthesis in vitro." Proceedings of the National Academy of Sciences of the United States of America **105**(32): 11122-11127.
Wanrooij, S., et al. (2012). "In vivo mutagenesis reveals that OriL is essential for mitochondrial DNA replication." Embo Reports **13**(12): 1130-1137.
Warburg, O. (1930). "Note on the metabolism of tumours." Biochemische Zeitschrift **228**: 257-258.
Warburg, O., et al. (1927). "The metabolism of tumors in the body." Journal of General Physiology **8**(6): 519-530.
Watanabe, Y., et al. (1994). "Higher-Order Structure of Bovine Mitochondrial Trna(Ser)Uga - Chemical Modification and Computer Modeling." Nucleic Acids Research **22**(24): 5378-5384.
Weber, J. and A. E. Senior (2000). "ATP synthase: what we know about ATP hydrolysis and what we do not know about ATP synthesis." Biochimica Et Biophysica Acta-Bioenergetics **1458**(2-3): 300-309.
Weber, K., et al. (1997). "A new mtDNA mutation showing accumulation with time and restriction to skeletal muscle." American Journal of Human Genetics **60**(2): 373-380.
Wedatilake, Y., et al. (2016). "TRNT1 deficiency: clinical, biochemical and molecular genetic features." Orphanet J Rare Dis **11**(1): 90.
Wells, R. D. and J. E. Larson (1972). "Buoyant density studies on natural and synthetic deoxyribonucleic acids in neutral and alkaline solutions." Journal of Biological Chemistry **247**(11): 3405-3409.
West, A. P., et al. (2011). "Mitochondria in innate immune responses." Nat Rev Immunol **11**(6): 389-402.
White, S. L., et al. (1999). "Mitochondrial DNA mutations at nucleotide 8993 show a lack of tissue- or age-related variation." Journal of Inherited Metabolic Disease **22**(8): 899-914.
Williams, A. J. and L. S. Kaguni (1995). "Stimulation of Drosophila Mitochondrial-DNA Polymerase by Single-Stranded DNA-Binding Protein." Journal of Biological Chemistry **270**(2): 860-865.

Wilson, I. J., et al. (2016). "Mitochondrial DNA sequence characteristics modulate the size of the genetic bottleneck." Human Molecular Genetics **25**(5): 1031-1041.
Wilson, W. C., et al. (2014). "A human mitochondrial poly(A) polymerase mutation reveals the complexities of post-transcriptional mitochondrial gene expression." Human Molecular Genetics **23**(23): 6345-6355.
Wisniewski, J. R., et al. (2016). "In-depth quantitative analysis and comparison of the human hepatocyte and hepatoma cell line HepG2 proteomes." Journal of Proteomics **136**: 234-247.
Wolstenholme, D. R. (1992). "Animal Mitochondrial-DNA - Structure and Evolution." International Review of Cytology-a Survey of Cell Biology **141**: 173-216.
Wonnapinij, P., et al. (2008). "The Distribution of Mitochondrial DNA Heteroplasmy Due to Random Genetic Drift." American Journal of Human Genetics **83**(5): 582-593.
Wright, S. (1942). "Statistical genetics and evolution." Bulletin of the American Mathematical Society **48**(4), 223-246.
Wu, A. R., et al. (2017). "Single-Cell Transcriptional Analysis." Annual Review of Analytical Chemistry, Vol 10 **10**: 439-462.
Xu, Y. J., et al. (2016). "Hydroxyurea Induces Cytokinesis Arrest in Cells Expressing a Mutated Sterol-14 alpha-Demethylase in the Ergosterol Biosynthesis Pathway." Genetics **204**(3): 959-+.
Yang, M. Y., et al. (2002). "Biased incorporation of ribonucleotides on the mitochondrial L-strand accounts for apparent strand-asymmetric DNA replication." Cell **111**(4): 495-505.
Yang, S. Y. and X. Y. Huang (2005). "Ca2+ influx through L-type Ca2+ channels controls the trailing tail contraction in growth factor-induced fibroblast cell migration." Journal of Biological Chemistry **280**(29): 27130-27137.
Yarham, J. W., et al. (2010). "Mitochondrial tRNA mutations and disease." Wiley Interdiscip Rev RNA **1**(2): 304-324.
Yarham, J. W., et al. (2014). "Defective i6A37 modification of mitochondrial and cytosolic tRNAs results from pathogenic mutations in TRIT1 and its substrate tRNA." PLoS Genet **10**(6): e1004424.
Yoboue, E. D. and A. Devin (2012). "Reactive oxygen species-mediated control of mitochondrial biogenesis." Int J Cell Biol **2012**: 403870.
Yoo, K. S., et al. (2019). "Postsynaptic density protein 95 (PSD-95) is transported by KIF5 to dendritic regions." Molecular Brain **12**(1).
You, Z. S. and J. M. Bailis (2010). "DNA damage and decisions: CtIP coordinates DNA repair and cell cycle checkpoints." Trends in Cell Biology **20**(7): 402-409.
Youle, R. J. and D. P. Narendra (2011). "Mechanisms of mitophagy." Nature Reviews Molecular Cell Biology **12**(1): 9-14.
Young, J. C., et al. (2003). "Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70." Cell **112**(1): 41-50.
Young, M. J., et al. (2015). "POLG2 disease variants: analyses reveal a dominant negative heterodimer, altered mitochondrial localization and impaired respiratory capacity." Human Molecular Genetics **24**(18): 5184-5197.
Yuan, S. H., et al. (2011). "Cell-surface marker signatures for the isolation of neural stem cells, glia and neurons derived from human pluripotent stem cells." PLoS One **6**(3): e17540.
Zeharia, A., et al. (2009). "Acute Infantile Liver Failure Due to Mutations in the TRMU Gene." American Journal of Human Genetics **85**(3): 401-407.
Zeviani, M. (2001). "The expanding spectrum of nuclear gene mutations in mitochondrial disorders." Seminars in Cell & Developmental Biology **12**(6): 407-416.

Zeviani, M., et al. (2003). "Nuclear genes in mitochondrial disorders." Current Opinion in Genetics & Development **13**(3): 262-270.
Zhang, H. X., et al. (2018). "The mitochondrial DNA genetic bottleneck: inheritance and beyond." Mitochondrial Diseases **62**(3): 225-234.
Zhang, H. X., et al. (2021). "Mitochondrial DNA heteroplasmy is modulated during oocyte development propagating mutation transmission." Science Advances **7**(50).
Zhang, M. J., et al. (2021). "Mouse aging cell atlas analysis reveals global and cell type-specific aging signatures." Elife **10**.
Zhang, X. N., et al. (2019). "Comparative Analysis of Droplet-Based Ultra-High-Throughput Single-Cell RNA-Seq Systems." Molecular Cell **73**(1): 130-+.
Zhang, Z. L., et al. (1998). "Electron transfer by domain movement in cytochrome bc(1)." Nature **392**(6677): 677-684.
Zhao, H. and H. Piwnica-Worms (2001). "ATR-mediated checkpoint pathways regulate phosphorylation and activation of human Chk1." Molecular and Cellular Biology **21**(13): 4129-4139.
Zheng, G. X. Y., et al. (2017). "Massively parallel digital transcriptional profiling of single cells." Nature Communications **8**.
Zhu, H., et al. (2004). "Induction of S-phase arrest and p21 overexpression by a small molecule 2[[3-(2,3-dichlorophenoxy)propyl] amino]ethanol in correlation with activation of ERK." Oncogene **23**(29): 4984-4992.
Ziegenhain, C., et al. (2017). "Comparative Analysis of Single-Cell RNA Sequencing Methods." Molecular Cell **65**(4): 631-+.

Supplementary Material

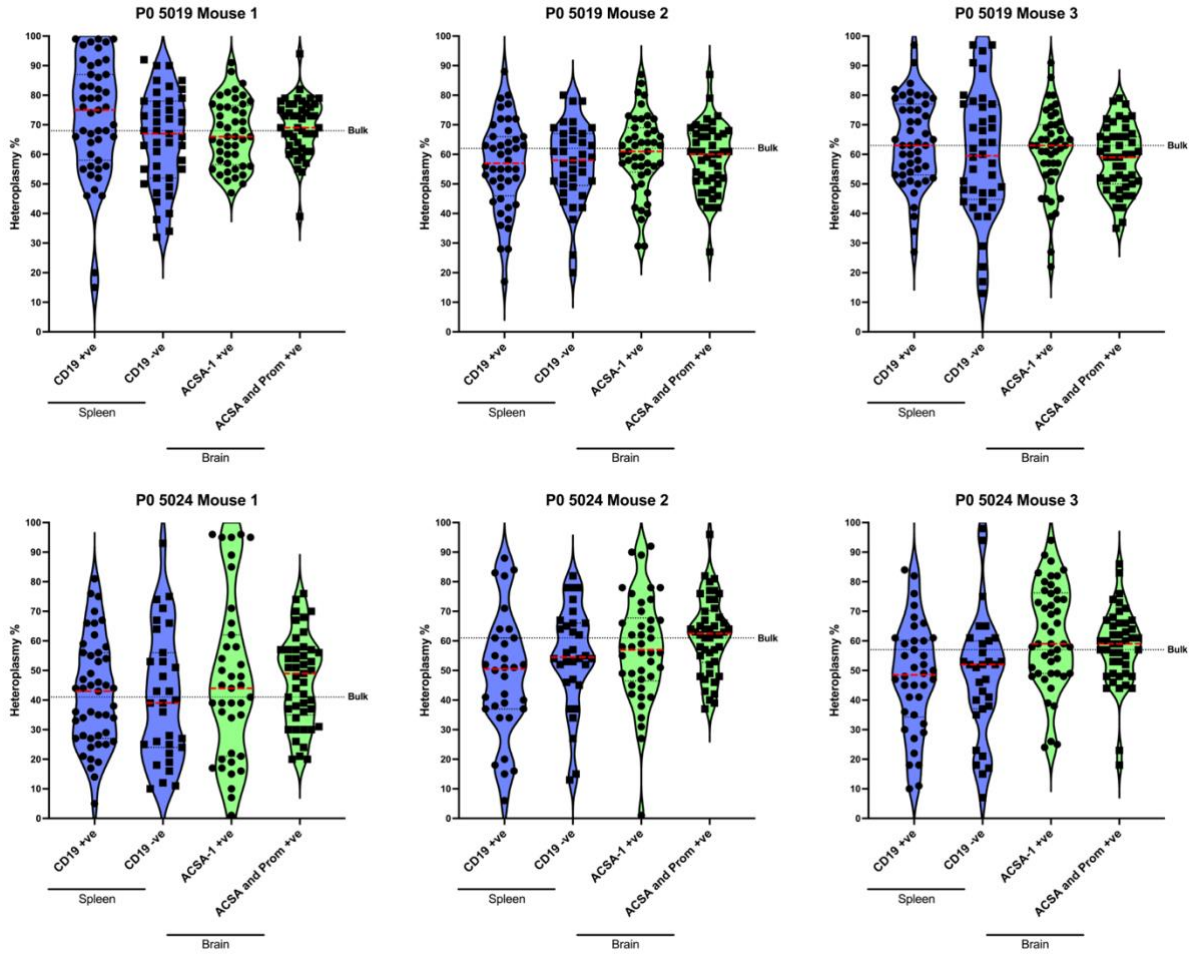


Figure S1. Single cell heteroplasmy of spleen and brain-derived populations at P0.
Single cell pyrosequencing measurements of mitochondrial DNA (mtDNA) heteroplasmy performed on cell populations derived from spleen and brain samples belonging three adult (P0) mice carrying either the m.5019A>G or the m.5024C>T mt-Ta mutation. Dotted lines represent heteroplasmy estimates calculated by averaging out the 8 bulk tissue measurements previously performed on various tissues belonging to the same animal (Figures 4.11. and 4.15.). Violin plots display median and interquartile range (IQR).

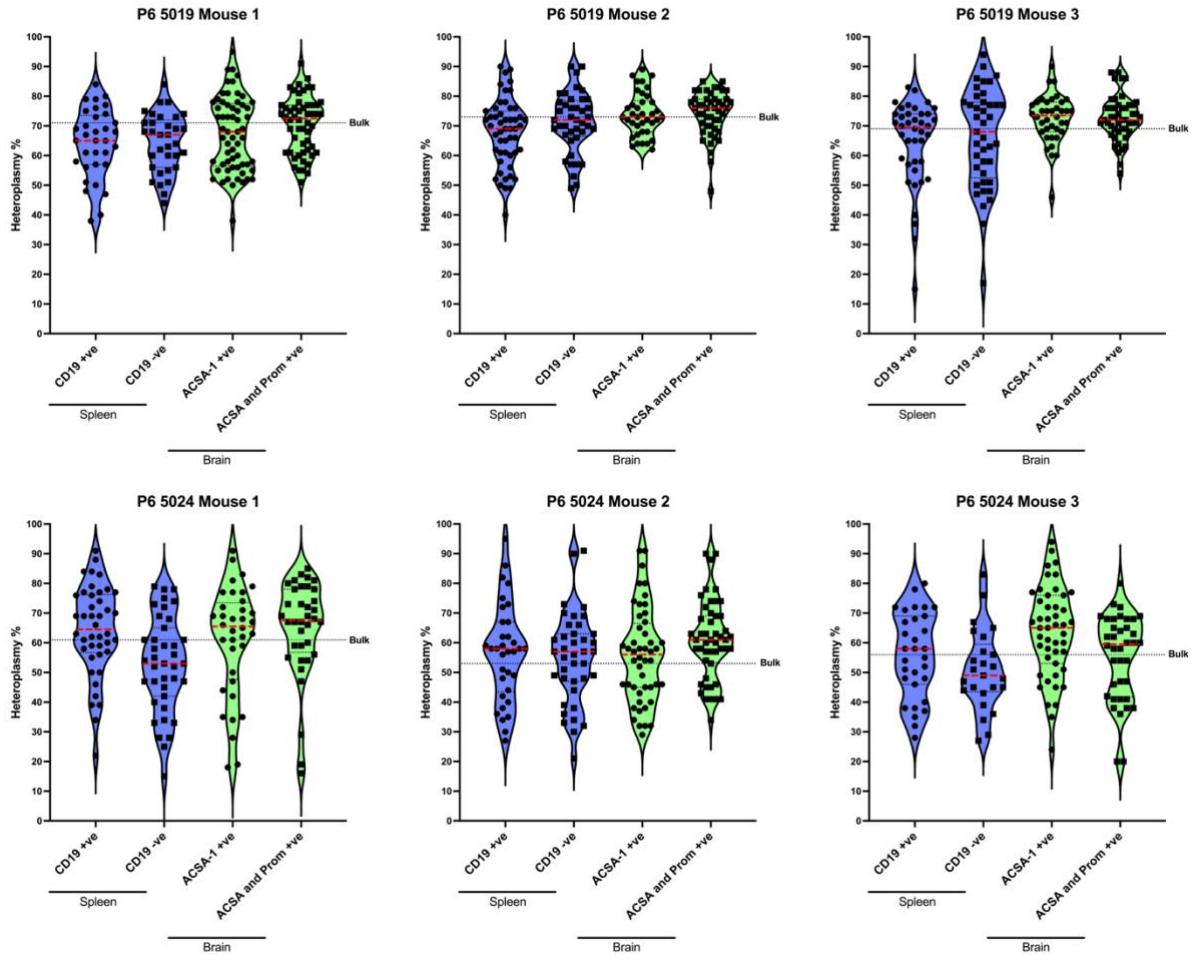


Figure S2. Single cell heteroplasmy of spleen and brain-derived populations at P6.
Single cell pyrosequencing measurements of mitochondrial DNA (mtDNA) heteroplasmy performed on cell populations derived from spleen and brain samples belonging three adult (P6) mice carrying either the m.5019A>G or the m.5024C>T mt-Ta mutation. Dotted lines represent heteroplasmy estimates calculated by averaging out the 8 bulk tissue measurements previously performed on various tissues belonging to the same animal (Figure Figures 4.11. and 4.15.). Violin plots display median and interquartile range (IQR).

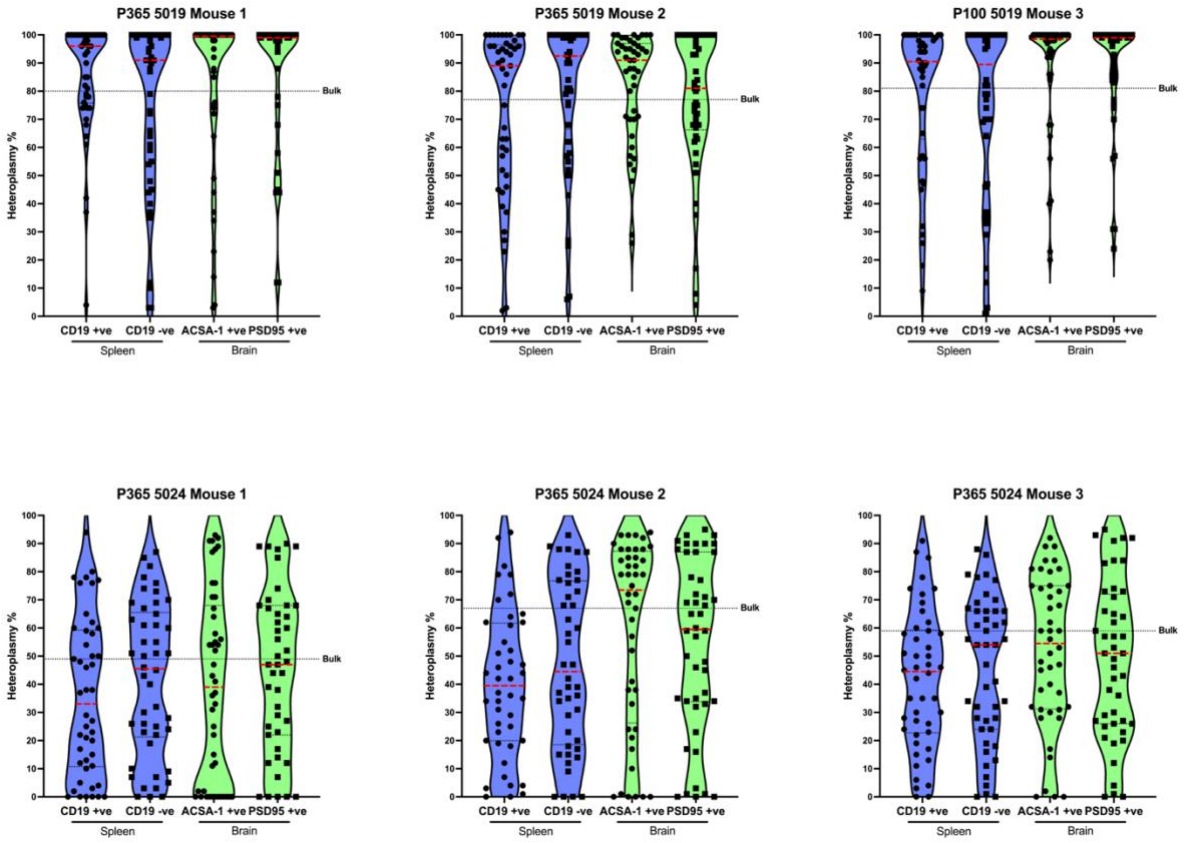


Figure S3. Single cell heteroplasmy of spleen and brain-derived populations at P6.
Single cell pyrosequencing measurements of mitochondrial DNA (mtDNA) heteroplasmy performed on cell populations derived from spleen and brain samples belonging three adult (P6) mice carrying either the m.5019A>G or the m.5024C>T mt-Ta mutation. Dotted lines represent heteroplasmy estimates calculated by averaging out the 8 bulk tissue measurements previously performed on various tissues belonging to the same animal (Figure Figures 4.11. and 4.15.). Violin plots display median and interquartile range (IQR).

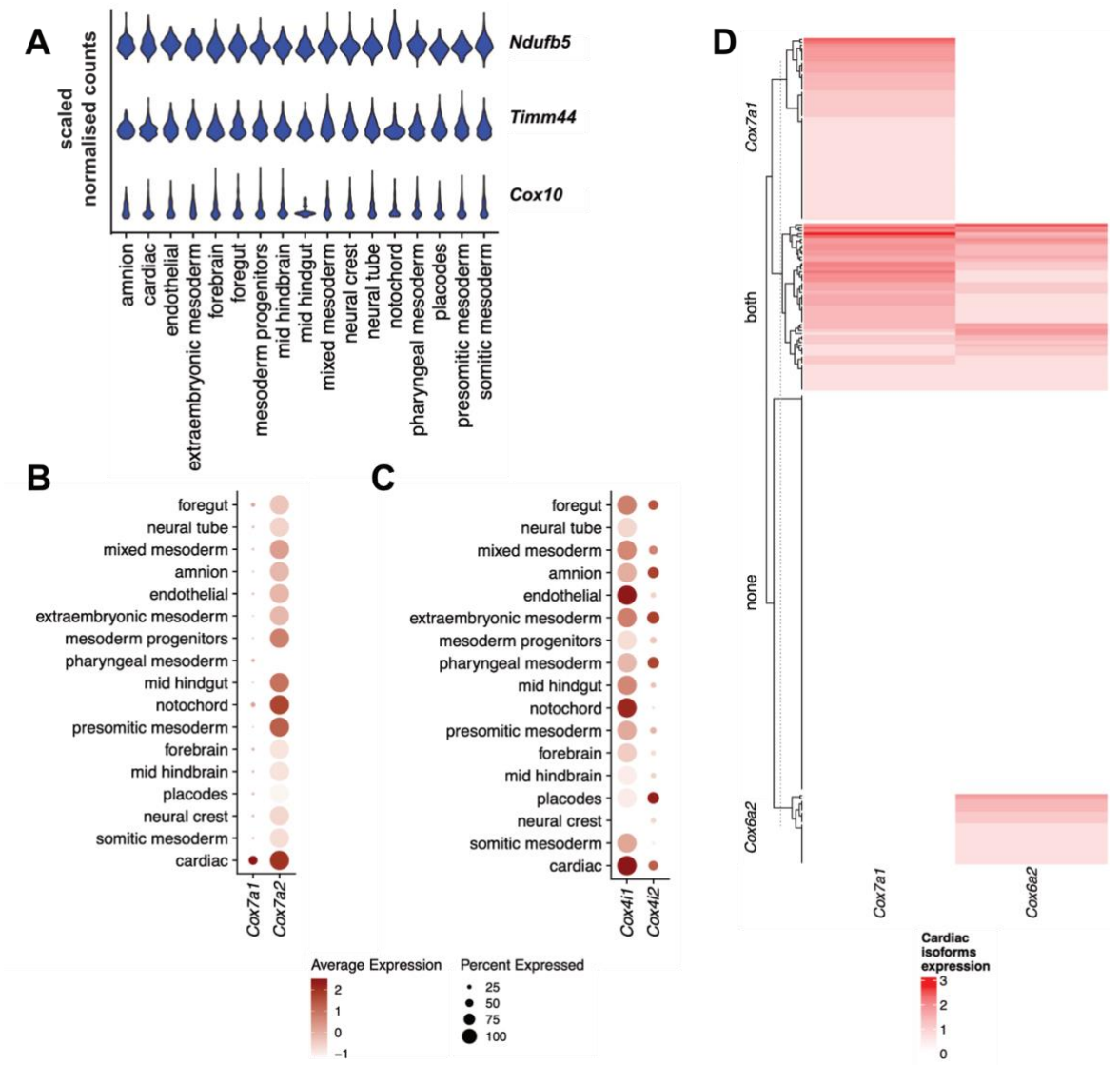


Figure S4. Nuclear-encoded mitochondrial genes are differentially expressed during organogenesis.
(A) Violin plots exhibiting the differential expression profiles of 3 select nuclear-encoded mitochondrial genes, revealing no significant differences across all 17 cell lineages. **(B)** Dot plots depicting the relative expression levels of tissue-specific Complex IV subunit *Cox7a* isoforms 1 and 2, across all 17 cell lineages. **(C)** Dot plots showing the relative expression levels of the two *Cox4* isoforms across all 17 cell lineages. **(D)** Heatmap depicting the relative expression levels of cardiac- and muscle-specific isoforms *Cox7a1* and *Cox6a2* belonging to Complex IV, in cells derived from the cardiac lineage. Figure generated by Dr Malwina Prater.

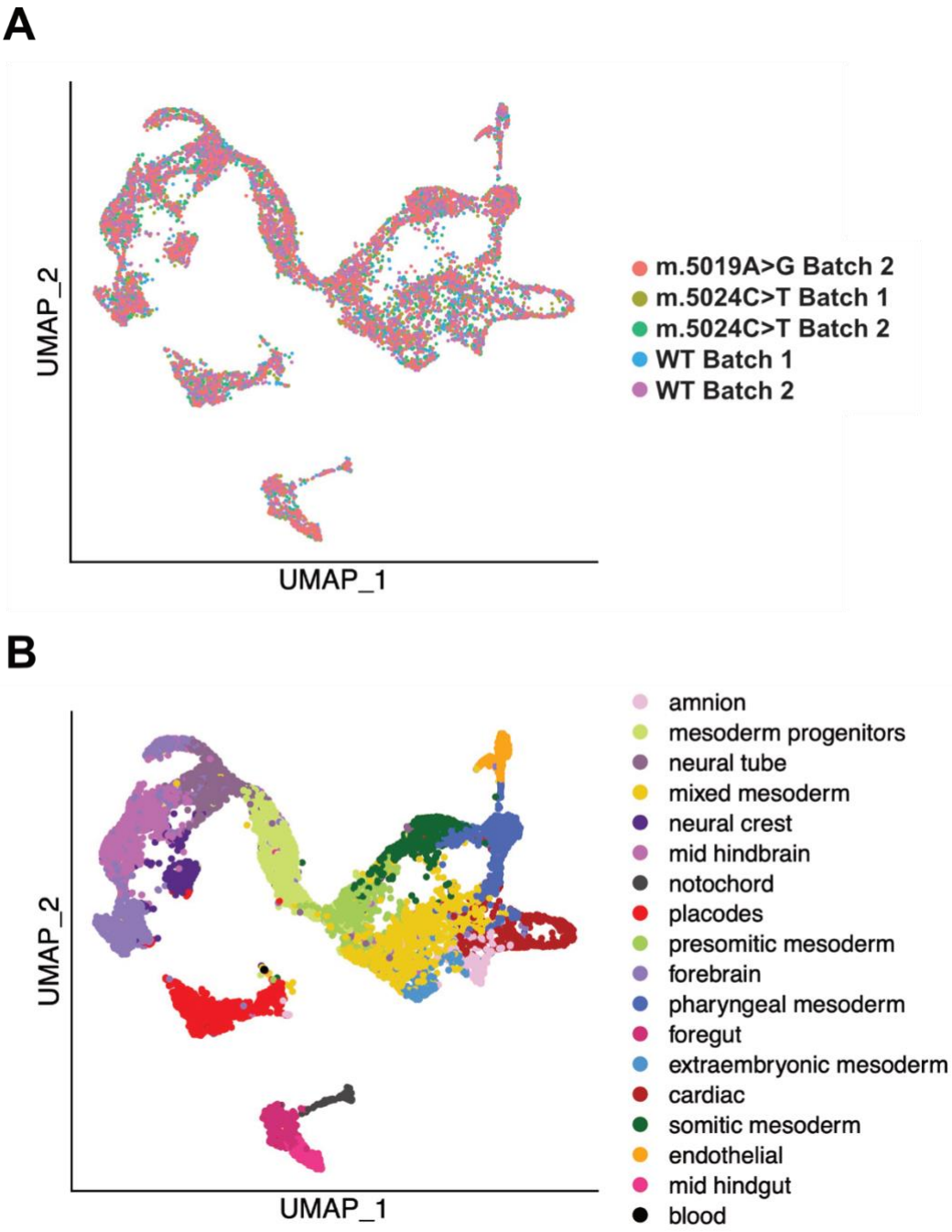


Figure S5. Cells passing quality control.
(A) Uniform manifold approximation and projection (UMAP) showing totality of single cells passing quality control (QC) from wild-type (WT), m.5019A>G and m.5024C>T E8.5 mouse embryos, belonging to ether batch 1 or 2. **(B)** UMAP depicting the entirety of single cells passing QC coloured based on their cell lineage. Batches 1 and 2 have been grouped; the blood cell lineage was detected only in m.5019A>G embryos. Figure generated by Dr Malwina Prater.

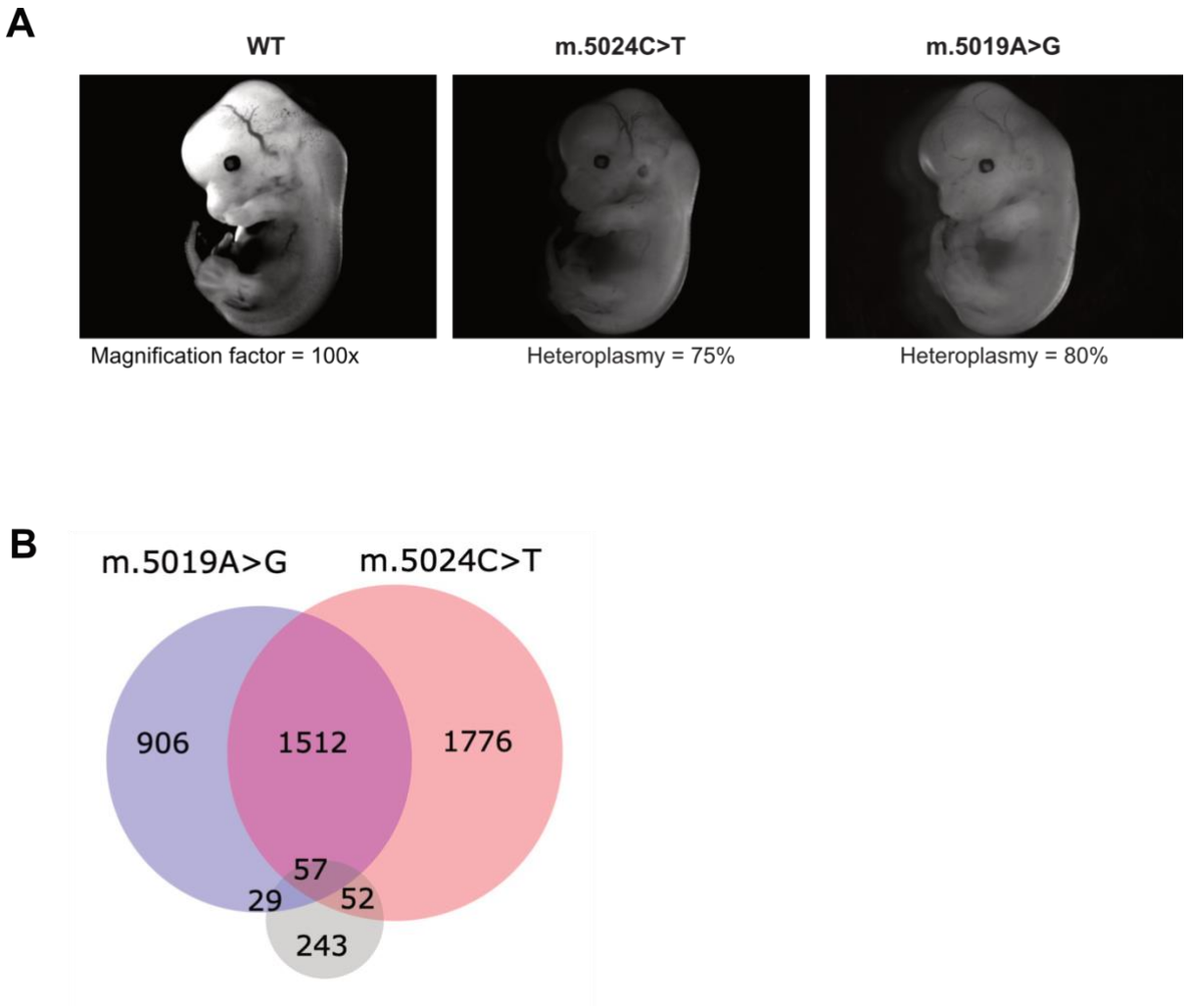
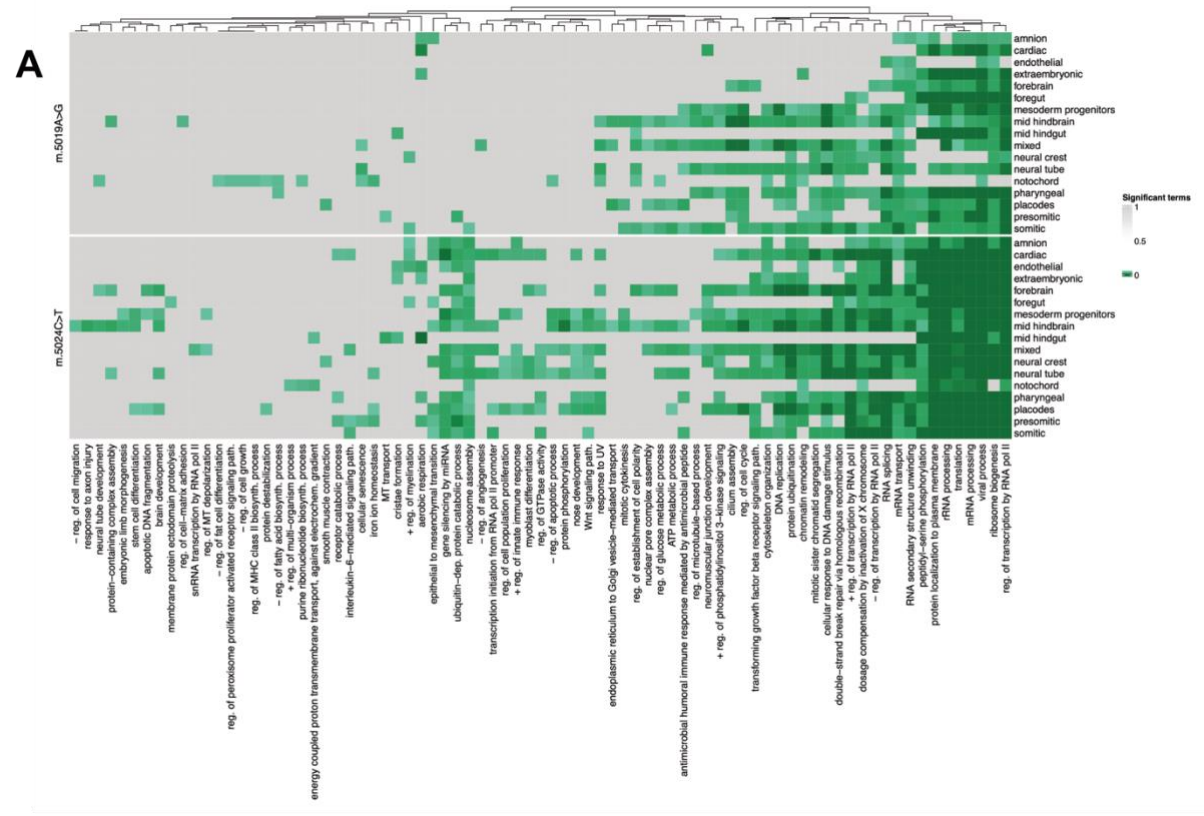


Figure S6. Compensatory response mounted against *mt-Ta* mutations.
(A) Representative images of E13.5 mouse embryos belonging to the wild-type (WT), m.5019A>G and 5024C>T lines, demonstrating comparable morphological characteristics post-organogenesis. Images were captures under 100x combined magnification (10x objective, 10x eyepiece). **(B)** Venn diagram depicting the overlap between established genetic modifiers of mitochondrial dysfunction detected through the application of an *in vitro* Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) screen (Mennuni et al., 2022) and differentially expressed genes belonging to embryos carrying the m.5019A>G and m.5024C>T *mt-Ta* mutations (highlighted in blue and red respectively). Figure **(A)** was generated by Stephen Burr, figure **(B)** was generated by Dr Malwina Prater.


A
m.5019A>G
m.5024C>T
amnion
cardiac
endothelial
extraembryonic
forebrain
foregut
mesoderm progenitors
mid hindbrain
mid hindgut
mixed
neural crest
neural tube
notochord
pharyngeal
placodes
presomitic
somitic
Significant terms
1
0.5
0
– reg. of cell migration
response to axon injury
neural tube development
protein–containing complex assembly
embryonic limb morphogenesis
stem cell differentiation
apoptotic DNA fragmentation
heart development
membrane protein ectodomain proteolysis
reg. of cell–matrix adhesion
snRNA transcription by RNA pol II
reg. of MT depolarization
– reg. of fat cell differentiation
reg. of peroxisome proliferator activated receptor signaling path.
reg. of cell growth
reg. of MHC class II biosynth. process
protein destabilization
– reg. of fatty acid biosynth. process
+ reg. of multi–organism process
purine ribonucleotide biosynth. process
energy coupled proton transmembrane transport, against electrochem. gradient
receptor catabolic process
interleukin–6–mediated signaling path.
cellular senescence
iron ion homeostasis
MT transport
cristae formation
+ reg. of myelination
aerobic respiration
epithelial to mesenchymal transition
gene silencing by miRNA
ubiquitin–dep. protein catabolic process
nucleosome assembly
– reg. of cell cycle
transcription initiation from RNA pol II promoter
reg. of cell population proliferation
+ reg. of innate immune response
myoblast differentiation
– reg. of GTPase activity
– reg. of apoptotic process
protein phosphorylation
nose development
Wnt signaling path.
response to UV
endoplasmic reticulum to Golgi vesicle–mediated transport
mitotic cytokinesis
reg. of establishment of cell polarity
nuclear pore complex assembly
reg. of glucose metabolic process
ATP metabolic process
antimicrobial humoral immune response mediated by antimicrobial peptide
reg. of microtubule–based process
neuromuscular junction development
+ reg. of phosphatidylinositol 3–kinase signaling
cilium assembly
reg. of cell cycle
transforming growth factor beta receptor signaling path.
cytoskeleton organization
DNA replication
protein ubiquitination
chromatin remodeling
mitotic sister chromatid segregation
cellular response to DNA damage stimulus
double–strand break repair via homologous recombination
+ reg. of transcription by RNA pol II
dosage compensation by inactivation of X chromosome
– reg. of transcription by RNA pol II
RNA export
mRNA transport
RNA secondary structure unwinding
peptidyl–serine phosphorylation
protein localization to plasma membrane
rRNA processing
translation
mRNA processing
viral process
ribosome biogenesis
reg. of transcription by RNA pol II

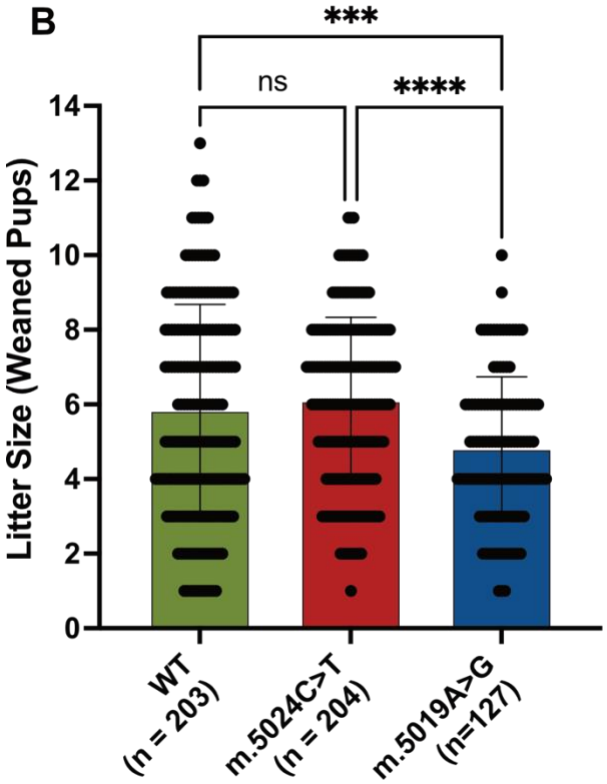

B
***
ns
****
Litter Size (Weaned Pups)
0
2
4
6
8
10
12
14
WT
(n = 203)
m.5024C>T
(n = 204)
m.5019A>G
(n=127)

Figure S7. Transcriptional responses that are mutation- and cell lineage-specific in nature
(A) Heatmap depicting Gene Ontology (GO) Biological Processes (annotation 2018) that are significantly enriched in each of the 17 cell lineages for both the m.5019A>G and m.5024C>T mouse embryos compared to controls. The establishment of biological processes is dependent on the expression levels of differentially regulated genes. **(B)** Litter sizes calculated in terms of number of pups weaned for animals belonging to the wild-type (WT), m.5019A>G and m.5024C>T mouse colonies. p-values were calculated by one way ANOVA with Tuckey's multiple test correction. standard deviation (SD) is shown by error bars; ns = not significant; **** = $p<0.0001$. Figure **(A)** was generated by Dr Malwina Prater, figure **(B)** was generated by Dr Stephen Burr.

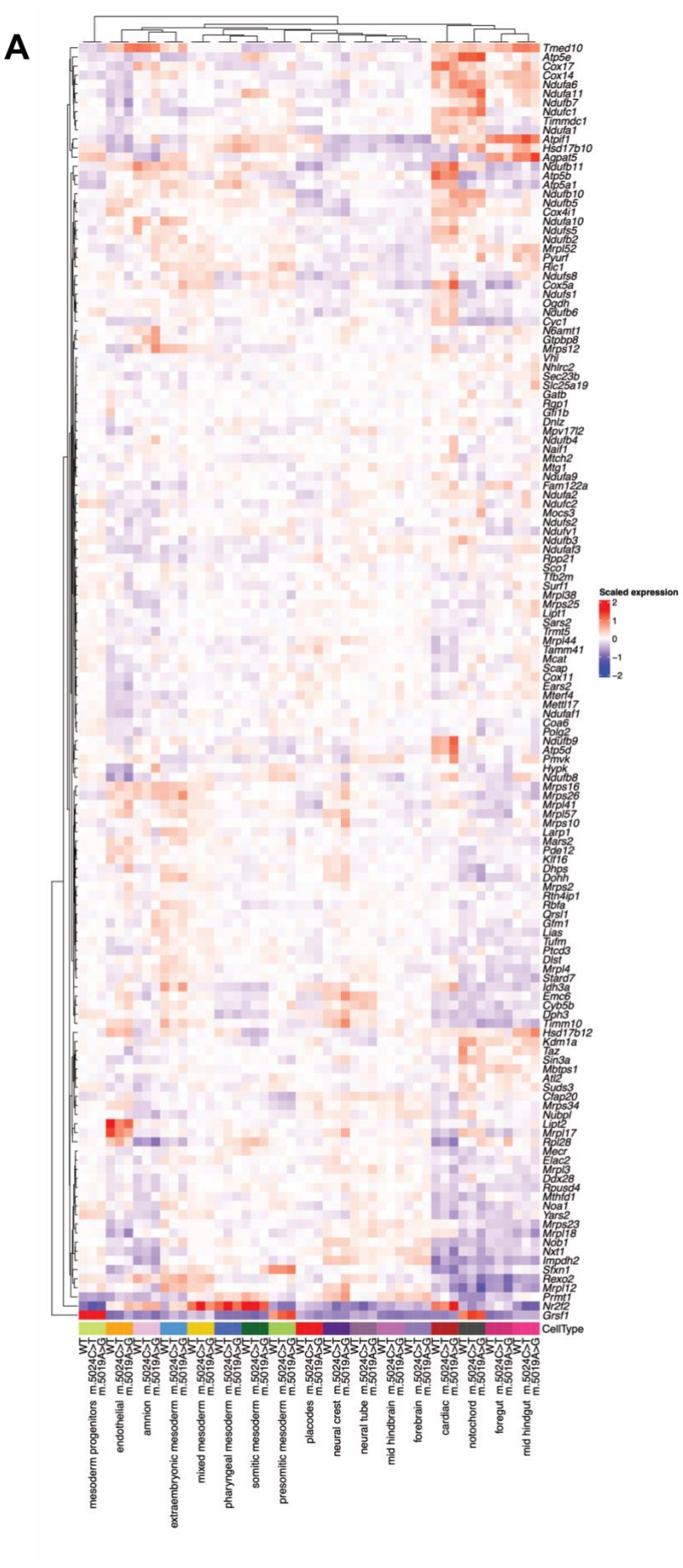

A
Tmed10
Atp5e
Cox17
Cox14
Ndufa6
Ndufa11
Ndufb7
Ndufc1
Timmdc1
Ndufa1
Atpif1
Hsd17b10
Agpat5
Ndufb11
Atp5b
Atp5a1
Ndufb10
Ndufb5
Cox4i1
Ndufa10
Ndufs5
Ndufb2
Mrpl52
Pyurf
Ric1
Ndufs8
Cox5a
Ndufs1
Ogdh
Ndufb6
Cyc1
N6amt1
Gtpbp8
Mrps12
Vhl
Nhlrc2
Sec23b
Slc25a19
Gatb
Rgp1
Glt1b
Dnlz
Mpv17l2
Ndufb4
Naif1
Mtch2
Mtg1
Ndufa9
Fam122a
Ndufa2
Ndufc2
Mocs3
Ndufs2
Ndufv1
Ndufb3
Ndufaf3
Rpp21
Sco1
Tfb2m
Surf1
Mrpl38
Mrps25
Lipt1
Sars2
Trmt5
Mrpl44
Tamm41
Mcat
Scap
Cox11
Ears2
Mterf4
Mettl17
Ndufaf1
Coa6
Polg2
Ndufb9
Atp5d
Pmvk
Hypk
Ndufb8
Mrps16
Mrps26
Mrpl41
Mrpl57
Mrps10
Larp1
Mars2
Pde12
Klf16
Dhps
Dohh
Mrps2
Rtn4ip1
Rbfa
Qrsl1
Gfm1
Lias
Tufm
Ptcd3
Dlst
Mrpl4
Stard7
Idh3a
Emc6
Cyb5b
Dph3
Timm10
Hsd17b12
Kdm1a
Taz
Sin3a
Mbtps1
Atl2
Suds3
Clpp20
Mrps34
Nubpl
Lipt2
Mrpl17
Rpl28
Mecr
Elac2
Mrpl3
Ddx28
Rpusd4
Mthfd1
Noa1
Yars2
Mrps23
Mrpl18
Nob1
Nxt1
Impdh2
Sfxn1
Rexo2
Mrpl12
Prmt1
Nr2f2
Grsf1
CellType
Scaled expression
2
1
0
−1
−2
WT
m.5024C>T
m.5019A>G
mesoderm progenitors
endothelial
amnion
extraembryonic mesoderm
mixed mesoderm
pharyngeal mesoderm
somitic mesoderm
presomitic mesoderm
placodes
neural crest
neural tube
mid hindbrain
forebrain
cardiac
notochord
foregut
mid hindgut

Figure S8. Epistatic buffer genes exhibit mutation- and cell lineage-specific expression patters as a response ton *mt-Ta* mutation.
(A) Unfiltered heatmap depicting the established epistatic buffer/suppressor genes that have been identified through a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) screen and their relative expression levels across all 17 cell lineages belonging to either wilt-type (WT), m.5019A>G or m.5024C>T mouse strains. Figure generated by Dr Malwina Prater.

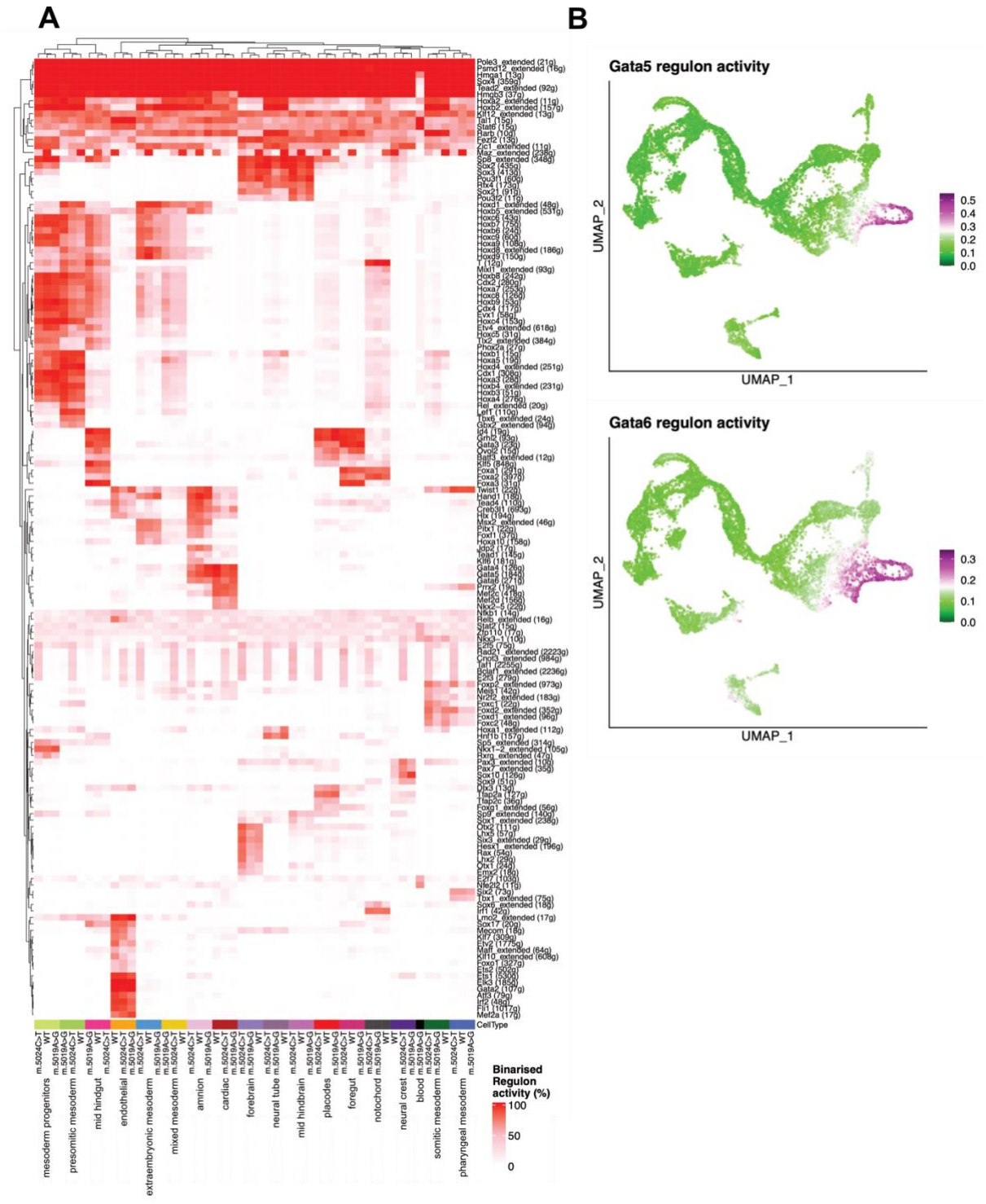


Figure S9. Gene regulatory networks are responsible for mounting the cell-lineage specific compensatory responses to *mt-Ta* mutations.

(A) SCENIC heatmap outlining the binary activity of the identified regulons. In order to define a regulon as 'active,' a minimum of 50% of cells have to show activity in at least one group. (B) Uniform manifold approximations and projections (UMAPs) showing the cell lineage-specific localisation of Gata5 and 6 regulons in wild-type (WT), m.5019A>G and m.5024C>T mouse lines, all overlayed on the same graph. Purple = increased regulon activity. Figure generated by Dr Malwina Prater.