Molecular mechanisms underlying the rate of progeria onset in mtDNA mutator mouse strains

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

SARA RUTH ALBARRAN GUTIERREZ

aus Estado de México, Mexiko

Berichterstatter:

Prof. Dr. **Aleksandra Trifunovic**

Prof. Dr. **Rudolf Wiesner**

Tag der mündlichen Prüfung:

26.05.2020

"You cannot hope to build a better world without improving the individuals. To that end, each of us must work for our own improvement and, at the same time, share a general responsibility for all humanity, our particular duty being to aid those to whom we think can be most useful."

—Marie Curie, *Autobiographical Notes, 1923*

ACKNOWLEDGEMENTS

I will always treasure all the moments I spend doing this research. Although science can be an obscure territory, the sparks I was capable to see have illuminated my world.

Table of Content

ABSTRACT

The mtDNA mutator mouse model contains a D257A mutation in the exonuclease domain of the POLG gene, which generates a mouse with an increased number of mtDNA mutations. The mouse shows several characteristics of premature ageing. To date, three mutators with the same POLG mutation have been described in the literature. However, they show remarkable differences in lifespan and the time when the ageing phenotype appears.

This study has demonstrated that although the genetic background influences to some degree the ageing phenotype, the main cause of the phenotypical differences exhibited by two of the mutator mice strains is due to the POLG allele expressed (*Lar* allele versus the *Pro* allele) and not the nuclear genetic background.

In addition, my work has revealed that the *Pro* allele generates a variant protein product of POLG, and as a result, the mutator mouse with this variant shows a slightly different phenotype than the mutator with the *Lar* allele. *Pro* mutator mouse phenotype differences include male fertility, increased lifespan, slower onset of the ageing phenotypes, reduced mutation rate, reduced mitochondrial dysfunction and different replicative activity.

ZUSAMMENFASSUNG

Das mtDNA-Mutator-Mausmodell enthält eine D257A Mutation in der Exonuklease-Domäne der POLG, die eine Maus mit einem mitochondrialen POLG-defizienten Protein erzeugt. Die Maus weist mehrere Charakteristika vorzeitigen Alterns auf. Bisher wurden drei Mutator-Mausmodelle mit der gleichen POLG Mutation beschrieben. Allerdings weisen diese bemerkenswerte Unterschiede in ihrer Lebenserwartung und dem Zeitpunkt, an dem der Alterungsphänotyp in Erscheinung tritt, auf.

Obwohl der genetische Hintergrund zu einem gewissen Grad den Alterungsphänotyp beeinflusst, hat meine Arbeit gezeigt, dass die Hauptursache für die phänotypischen Unterschiede bei den Mutator-Mäusen die exprimierten Allele sind, und nicht der genetische Hintergrund.

Darüber hinaus hat meine Arbeit offenbart, dass das *Pro*-Allel eine Variante der POLG generiert und, als Ergebnis, die Mutator-Maus mit dieser Variante einen geringfügig anderen Phänotyp aufweist als die Mutator-Maus mit dem *Lar*-Allel. Die phänotypischen Unterschiede der *Pro*-Mutator-Maus umfassen die männliche Fruchtbarkeit, eine erhöhte Lebensdauer, einen milderen Alterungsphänotyp, eine verminderte Mutationsrate, reduzierte mitochondriale Dysfunktion und eine andere replikative Aktivität.

1. INTRODUCTION

1.1. Mitochondria: an overview

1.1.1. History of mitochondrial research

The first descriptions of intracellular structures go back to the 1840s (Ernster and Schatz, 1981). Nevertheless, it was not until 1890 that Richard Altmann, a German pathologist, reported in his book *Die Elementarorganismen Und Ihre Beziehungen Zu Den Zellen* (The elementary organisms and their relations to the cells) some structures he named "bioblasts" (life germs) (Figure 1-1A). He described these structures as elementary organisms living inside the cells, and he postulated that they carry out vital metabolic functions for the cell (Altmann, 1890). However, his book was received with criticism and scepticism, and it took almost 20 years until his ideas were taken into consideration (Cowdry, 1953).

In the meantime, Carl Benda identified some structures during his research on spermatogenesis. He called the unit "mitochondrion", from the Greek *mitos* (thread) and *chondros* (granule), coining the word in 1898. Two years later, Michaelis was able to stain specific mitochondria using Janus Green B dye. This technique allowed researchers (like Meves, Cowdry, among others) to discovered that cells from plants, protozoa, and animals contain mitochondria, leading to speculation that mitochondria may play an important role in cells (Cowdry, 1953; Ernster and Schatz, 1981).

Over the following years, several theories about the role of mitochondria flourished. For instance, the role of mitochondria in genetic inheritance by Payne in 1909 (albeit maternal), the beginning of endosymbiotic theory by Wallin in 1927, and several energy metabolism studies (Reviewed by Ernster and Schatz, 1981).

However, it was not until the development of methods to isolate mitochondria in 1934 and 1948 that the stage for a more thorough mitochondrial analysis was set. These methods included the isolation of mitochondria by cell fractionation (Bensley and Hoerr, 1934) and

the addition of sucrose solution during fractionation (Hogeboom *et al.*, 1948). In 1952, the first electron micrograph of the mitochondria was obtained (Figure 1-1B) (Palade, 1952).
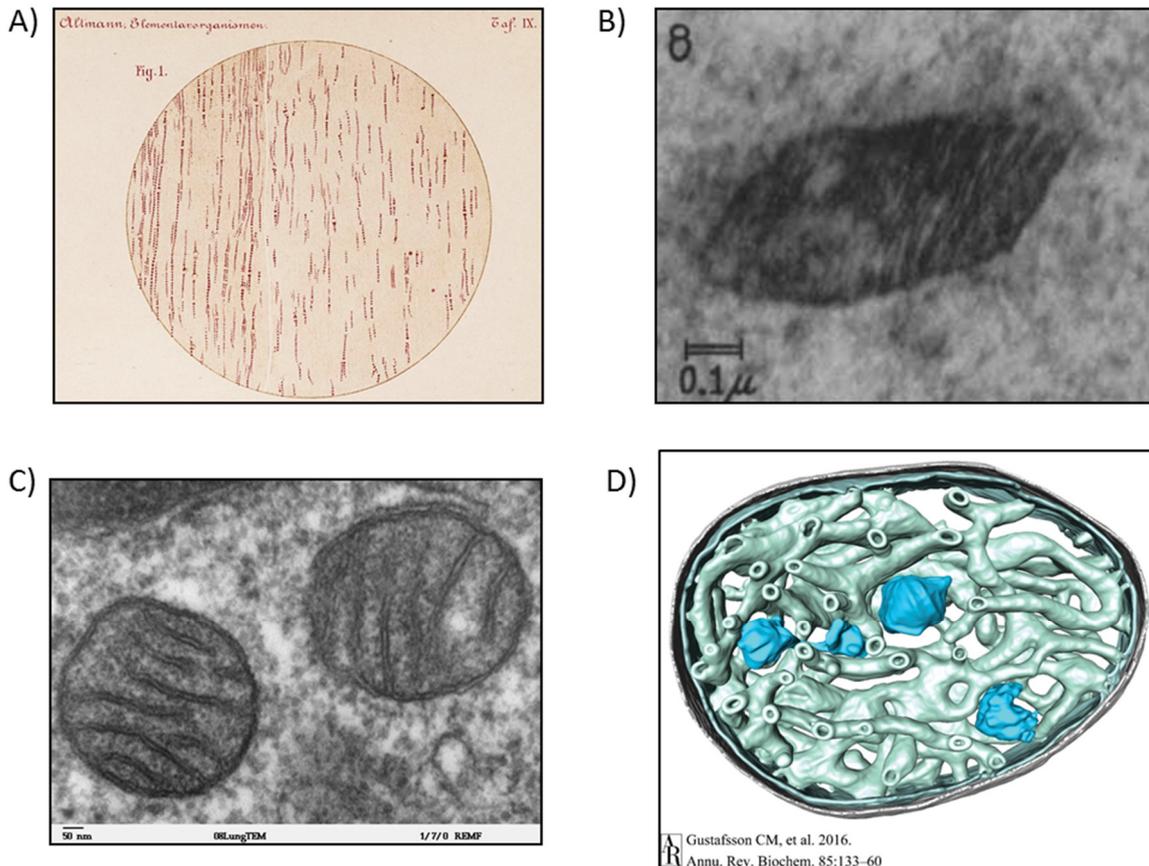
Figure 1-1. Mitochondria through time. A) The first representation of structures (bioblasts) by Richard Altmann in 1890. Sketch of a longitudinal section through the muscle of the adult frog. Staining using mercury mixture with formic acid (Altmann, 1890). B) First Electron micrograph of the mitochondrion from nephron rat embryo by Palade in 1952. Mitochondrion from nephron rat embryo (Palade, 1952). Figure license Number 4735620866487 by John Wiley & Sons and Copyright Clearance Center. C) and D) Contemporary view of mitochondria. C) Transmission electron microscopy of mitochondria from a thin section cut of mammalian lung tissue. Author: Louisa Howard, year 2008. Figure of public domain at: https://commons.wikimedia.org/wiki/File:Mitochondria, mammalian lung - TEM.jpg. D) Computer reconstructed Cryo-electron tomography. The figure is a representation of bovine heart mitochondrion. Nucleoids are shown in blue and green structures are cristae (Gustafsson *et al.*, 2016). Figure license Number 1018149-1 by John Wiley & Sons and Copyright Clearance Center.

In the biochemical field, some big steps were made. For example, during the 1950s it was demonstrated that mitochondria contain enzymes for the citric acid cycle, fatty acid oxidation and phosphorylation. The membrane-localization of the respiratory chain complexes was also reported (Reviewed by Ernster and Schatz, 1981). Because of research during these years, mitochondrion was termed the "Powerhouse of the cell" in a paper of the same name (Siekevitz, 1957). In 1963 an incredible discovery was reported. The electron micrographs from chick embryo cells evidenced that mitochondria contained threadlike structures. Moreover, these structures were sensitive to digestion with DNases, implying that mitochondria (and chloroplasts) contained their own DNA (Nass and Nass, 1963). Months later, DNA from yeast mitochondria was purified using biochemical procedures (Schatz *et al.*, 1964). This research finally confirmed that mitochondria contain their own DNA (mtDNA). All these discoveries formed the basis for a completely new and exciting field of cell biology, which, with the use of new technologies has grown rapidly.

1.1.2. Origin and evolution of mitochondria

To date, genomic and cell biology research has revealed that all eukaryotes are derived from a single common ancestor, termed LECA (last eukaryotic common ancestor). It is generally accepted that LECA originates approx. 1.2 billion years ago (but the estimates can vary twofold), when an archaeon cell entered into an endosymbiotic relationship with an α-proteobacterium. Nevertheless, controversy persists about the precise lineage of the proteobacterium (Gray, 2012; Martijn *et al.*, 2018), how closely related they are to mitochondria (Boussau *et al.*, 2004), and the nature of the cell that engulfed the proteobacterium (Martin *et al.*, 2001).

Since not all eukaryotes followed the same evolutionary pattern, mitochondria have drastic differences in terms of function among organisms. In accordance with their energy metabolism, mitochondria have been classified as aerobic mitochondria, anaerobic mitochondria, hydrogen-producing mitochondria, hydrogenosomes and mitosomes (Muller

et al., 2012). Despite the single exception of *Monocercomonoides sp.* where mitochondria appear to have been lost (Karnkowska *et al.*, 2016), all known eukaryotes to date possess mitochondria or mitochondria-related organelles (MROs) (Roger *et al.*, 2017).

The endosymbiotic process, where an autonomous α-proteobacterium evolved to a dependent organelle, has attracted much attention in the last decades. In this regard, it has been proposed that in order to achieve endosymbiosis, complex changes must have had occurred. These changes during evolution must have included: (1) acquisition of transporters/carriers into the mitochondrial membrane, (2) evolution of protein-import mitochondrial machinery, (3) integration of a communication system between host and symbiont, (4) re-targeting of proteins to mitochondria, (5) genome reduction through loss of unnecessary genes in the mitochondria, and (6) gene transfer to the nucleus (Roger *et al.*, 2017).

The mitochondrial genome transfer to the nucleus has raised many questions and debates. The major debate focuses on understanding why most genes were transferred from the symbiont to the nuclear genome, while others remained (like OXPHOS components and tRNAs) (Blanchard and Lynch, 2000). Several theories have been thus far proposed, for example, the Co-location for Redox Regulation (CoRR) hypothesis, postulates that genes from the electron transport chain (ETC) are maintained in the mtDNA because it allows a rapid regulation and gene expression during changes in the redox state (Allen, 2003). A second theory proposes that codon usage variations between the nucleus and organelles have prevented the gene transfer (Race *et al.*, 1999). A third hypothesis postulates that some large hydrophobic proteins remained encoded in the mtDNA, with multiple transmembrane domains, because they are difficult to import across membranes (Adams and Palmer, 2003; Popot and Vitry, 1990). Moreover, if they were expressed in the cytosol they could be mis-targeted to the endoplasmic reticulum (Björkholm *et al.*, 2017; von Heijne, 1986). Also, it has been discussed that mitochondrial small and large subunit rRNA genes remained encoded in the mtDNA since translocating large structured RNAs through mitochondrial membranes might be difficult (Roger *et al.*, 2017).

The second theory that proposes codon usage variations as the limiting step for gene transfer have some evidence against. For example, several genes from the α-proteobacterium have

already been passed to the nuclear genome (Martin, 2003; Mourier *et al.*, 2001). Moreover, the universal genetic code is used by the mitochondria of green plants (Jukes and Osawa, 1990). Therefore, this hypothesis has been largely discarded. To date, the CoRR theory has a strong experimental evidence. For instance, it has been demonstrated with *in organello* assays that protein synthesis is responsive to redox reagents (Allen *et al.*, 1995; Escobar Galvis *et al.*, 1998) and the transcription of chloroplast DNA can be redox regulated (Pfannschmidt *et al.*, 1999; Wilson *et al.*, 1996). In addition, these theories are not necessarily mutually exclusive. For instance, mtDNA genome analysis of 2015 species suggest that GC content and hydrophobicity are two characteristics that strongly predict if a gene will be retained in the mtDNA, and redox stress is known to be linked to adenine pools depletions, which could explain the preference to use GC-rich codons (Johnston and Williams, 2016).

1.1.3. Structural organisation and functions of mitochondria

There is a strong mitochondria variation among eukaryotes (Roger *et al.*, 2017); therefore, in this thesis, we will just focus on mammalian mitochondria. Although mammalian mitochondria maintain some characters from their protobacterium ancestors, these organelles have undergone to several changes, transforming them into highly specialised structures (Roger *et al.*, 2017).

Mitochondria are not isolated structures, in fact, they are dynamic organelles that continuously change shape through fission (division of one mitochondrion into two daughter mitochondria) and fusion (union of two mitochondria into one mitochondrion), and change their position by the movement along cytoskeletal tracks (van der Bliek *et al.*, 2013). Mitochondrial movement was first described more than 100 years ago (Lewis and Lewis, 1915), but our knowledge in regard to mitochondrial dynamics has dramatically increased in the last 40 years because of the development of live cell imaging technology

(Liesa *et al.*, 2009). Now we know that mitochondrial dynamics have a clear relevance for processes like cell growth and death (Westermann, 2010), immune response (Rambold and Pearce, 2018), mitochondrial re-distribution and mitochondrial quality control (Ni *et al.*, 2015).

Mitochondria have two lipid bilayer membranes, the inner membrane (IM) and the outer-membrane (OM), which form different compartments. Membranes and compartments exhibit specific functions (Friedman and Nunnari, 2014; McCarron *et al.*, 2013). The outer-mitochondrial membrane (OM) forms a barrier to the cytosol, mediates the import of proteins and lipids, and the passage of small molecules (Yamashita *et al.*, 2016). The inter-membrane space (IMS) contains proteins implicated in mitochondrial energetics and apoptosis, like cytochrome c (CytC) (McCarron *et al.*, 2013). The inner-membrane contains a large amount of transporters for carrying proteins to the matrix (Wohlrab, 2009), and contains double-membrane interdigitations called cristae (which are the principal site of oxidative phosphorylation) (Cogliati *et al.*, 2016). Finally, the matrix is the site where the tricarboxylic acid (TCA) cycle occurs, and where mtDNA is located in structures called nucleoids (Osellame *et al.*, 2012) (Figure 1-2).

Mitochondria are best known to be responsible for harnessing the energy released in oxidative metabolism of the cell, in the form of ATP. In general terms, mitochondrial energetic metabolism begins with the internalisation of pyruvate (generated mainly from glucose) and acyl-CoA (generated from fatty acids) in to the matrix. Briefly, during beta-oxidation, fatty acids are broken down to generate acetyl-CoA (which enters the TCA cycle), NADH and $FADH_2$. Pyruvate, on the other hand, will be oxidised by the pyruvate dehydrogenase complex and form acetyl-CoA (which will enter the TCA cycle) and NADH. The acetyl-CoA then enters the TCA cycle, generating more NADH and $FADH_2$. Then, the electrons carried by NADH and $FADH_2$, are transported to the oxidative phosphorylation system (OXPHOS), forming an electrochemical gradient that is harnessed to produce ATP from ADP. In the end, the ATP is transported to the cytoplasm by the ADP/ATP translocase (Spinelli and Haigis, 2018). Additionally, mitochondria have other critical metabolic functions, for instance Fe-S cluster biogenesis, metabolism of fatty acids, and is implicated in the metabolism of some amino acids (Spinelli and Haigis, 2018).
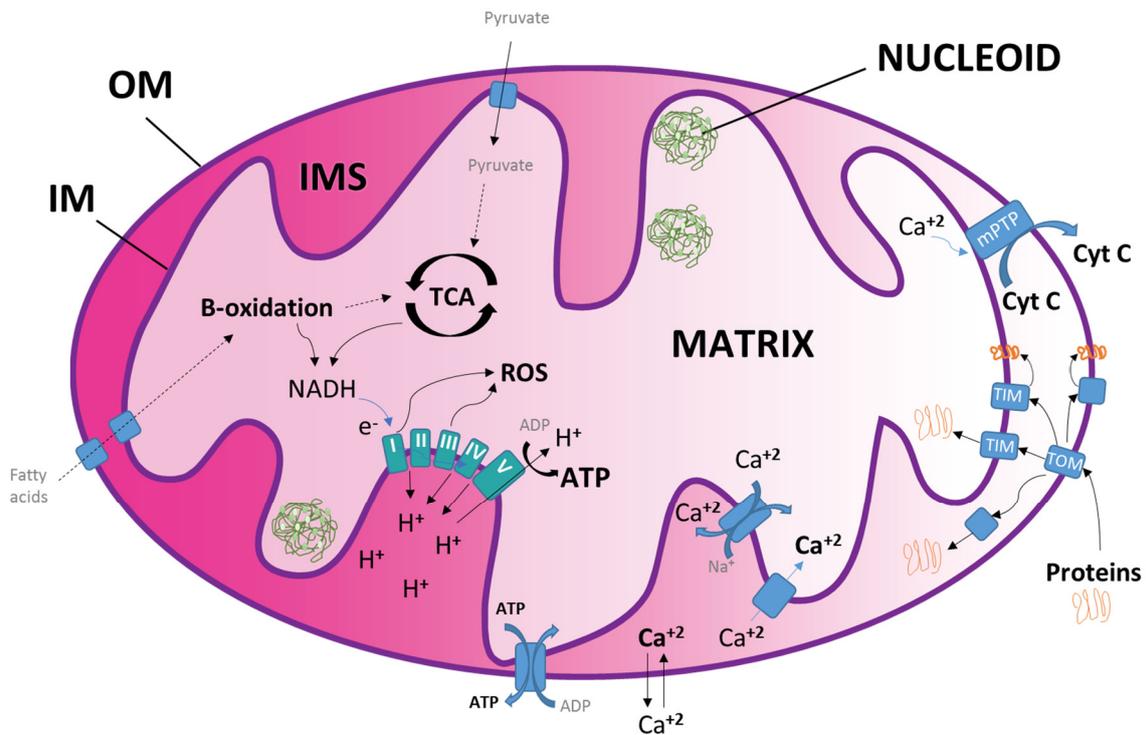
Figure 1-2. Schematic representation of mitochondrial main functions. The illustration shows the different functions discussed in the text. The main structures/parts in the mitochondria are IM (inner-membrane), OM (outer-membrane), IMS (inter-membrane space), matrix and nucleoids. Important in oxidative respiration is the generation of ATP through the OXPHOS system (complexes I, II, III, IV, V). The OXPHOS system is fed by the electron transporters NADH or FADH2, which are produced by b-oxidation and the TCA (tricarboxylic acid) cycle. Mitochondria can also generate ROS (reactive oxygen species), as a side product of OXPHOS. Mitochondria are also involved in apoptosis, through the release of calcium and CytC (cytochrome c) by the mPTP (mitochondrial permeability transition pore). The majority of mitochondrial proteins are synthesized as precursors in the cytosol and later imported by the protein import mitochondrial machinery. In this schema, the proteins TIM and TOM represent the mitochondrial protein transport; however, mitochondrial machineries to import proteins are much more complex. Mitochondria have an important role in calcium signalling by sequestering and releasing Ca2+ (calcium ion), for example through calcium diffusion, calcium uniporter, sodium-calcium exchanger and the mPTP.

Moreover, this amazing organelle interacts actively with other parts of the cell and has a much wider range of functions. There is a vast amount of literature demonstrating the important role of mammalian mitochondria. For instance, in apoptosis activation through the release of proteins from the intermembrane space (e.g. CytC and caspases) (Tait and

Green, 2013), calcium signalling through the sequester and release of Ca^{2+} (calcium ion) (Giorgi *et al.*, 2018), and ROS signalling pathway through the production of ROS during the OXPHOS process (Shadel and Horvath, 2015).

1.2. Mammalian mitochondrial DNA

1.2.1. Genome organisation

The size of the mtDNA varies among the species (Kolesnikov and Gerasimov, 2012). In humans, the mtDNA length is ∼16.5 kb (Andrews *et al.*, 1999) while in mice it is ∼16.3 kb (Bayona-Bafaluy *et al.*, 2003). Although nuclear DNA encodes 99% of the proteins required for mitochondrial function, mitochondrial DNA (mtDNA) is indispensable. The mammalian mtDNA encodes 13 polypeptides, which are necessary for the assembly and function of the OXPHOS system. In addition, the mtDNA encodes 22 tRNAs and 2 ribosomal RNAs, all necessary for mtDNA translation (Gammage and Frezza, 2019).

Mammalian mtDNA is a double-stranded circular molecule. Each strand is named either heavy (H) or light (L) strand; these names were given to them due to their guanine content differences when human mtDNA are separated in $CsCl_2$ gradients (Borst, 1972; Borst *et al.*, 1967). Specifically, from the 37 genes encoded by the mtDNA, the majority are encoded in the light strand and transcribed from the heavy strand: 10 mRNAs (two bi-cistronic, ATP8/ATP6 and Nd4l/Nd4), 2 rRNAs (12S and 16S) and 14 tRNAs. Only 1 mRNA and 8 tRNAs are encoded in the heavy strand and transcribed from the light strand (Figure 1-3) (Chinnery and Hudson, 2013).

Most of mtDNA (almost 93%) consists of coding regions. This feature, unlike nuclear DNA, is because mtDNA lacks intron regions and additionally genes are separated only by one or two base pairs (Chinnery and Hudson, 2013) (Figure 1-3B). The only large non-coding region in the mammalian mtDNA is the named non-coding control region (NCR) and is ∼1.1 kb of size in human mtDNA (Nicholls and Minczuk, 2014). However, the non-coding

control region (NCR) (See Figure 1-3A and chapter 1.2.2) contains several important regulatory sequences: (1) promoters for transcription of light and heavy strands, (2) the replication origin of the heavy strand, (3) the regulatory conserved sequence blocks (CSB 1-3), (4) the termination-associated sequence TAS, and (5) the displacement loop (D-loop) structure (Nicholls and Minczuk, 2014).
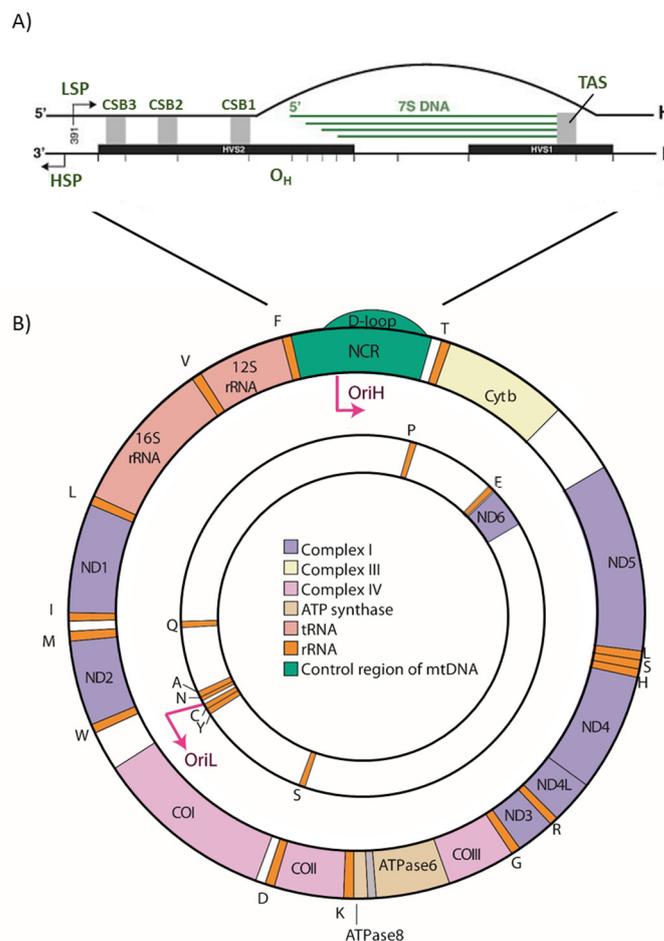
Figure 1-3. Mammalian mitochondrial DNA. A) Schematic representation of the non-coding region (NCR) area. This area contains the LSP (light strand promoter) and HSP (heavy strand promoter), the conserved sequence blocks (CSB1, CSB2, CSB3), the origin of replication of the heavy strand (O_H) and the termination-associated sequences (TAS). In this area, the displacement loop (D-loop) is formed by the replication of the 7S DNA. B) Annotated genetic features of the mtDNA molecule. Here is shown the heavy-strand encoding majority of genes and also the light strand, which only encodes for ND6 and 8 tRNAs. Genes are coloured by type of gene-encoded. Figure modified from Nicholls and Minczuk, 2014. License Number 4737910785177 by Elsevier and Copyright Clearance Center.

1.2.2. mtDNA structural organisation and characteristics

MtDNA is tightly packed in structures named nucleoids. These structures have an ellipsoidal shape, with a diameter of ~100nm, and on average contain 1.4 copies of mtDNA per nucleoid. In this regard, mitochondrial transcription factor A (TFAM) of mammals has been identified as the basic and minimum component for the compaction of these structures. In detail, TFAM has been proved not only to be abundant in the mtDNA (~1000 molecules of TFAM per mtDNA molecule) but also to be capable of bending mtDNA (even 180°) (Bonekamp and Larsson, 2018; Kukat *et al.*, 2011, 2015). Additionally, TFAM has been demonstrated to regulate mtDNA copy number (Ekstrand *et al.*, 2004; Farge *et al.*, 2014).

The mtDNA exhibits exceptional characteristics. For instance, unlike in the nucleus, the codon code used by the mitochondria involves a minimalistic set of tRNAs. In humans, for example, the tRNA^{Met} in the nucleus only recognises AUG codons but in mitochondria binds also AUA codons (and AUU for the start codon of the NADH dehydrogenase subunit 2). In the case of the isoleucine, only one single species of tRNA^{Ile} is necessary to translate both isoleucine codons (Boos *et al.*, 2016; Chinnery and Hudson, 2013). A remarkable feature of the mitochondrial genome is the multi-copy nature of it. It has been estimated that each mitochondrion can contain ~1-3 molecules of mtDNA (Wiesner *et al.*, 1992). However, mitochondria are dynamic organelles and during fusion and fission processes, mitochondria can exchange mtDNA (Arimura *et al.*, 2004). Therefore, a more reliable quantification is to use mtDNA copy number per cell. In that regard, each nucleated cell (excluding the germline) can contain between 100 and 10,000 copies depending on the cell type. For example, in humans, leukocytes contain ~350 (Liou *et al.*, 2010), while myocardial cells contain ~6,000 copies per cell (Miller, 2003). Although the mechanisms regulating copy number are still not fully understood, one of the mechanisms to maintain and regulate mtDNA copy number is mtDNA replication (Shadel and Horvath, 2015).

1.3. mtDNA replication

1.3.1. Replisome

I*n vitro* studies suggest that in mammals, the minimum machinery necessary for mtDNA replication consists of three proteins: single-stranded DNA binding protein (mtSSB), TWINKLE, and the core of the machinery, the mitochondrial DNA polymerase gamma (Pol γ). In this complex, normally named the replisome, the Pol γ is incapable of using annealed, double-stranded DNA (dsDNA) as a template. Therefore, it is required that the helicase TWINKLE unwinds the mtDNA and leads to the exposure of single-stranded DNA. The single-stranded binding protein mtSSB is also involved in stimulating this reaction and then binds to the displaced single-stranded DNA (Korhonen *et al.*, 2004). However, *in vivo*, other proteins may also play an important role like mitochondrial genome maintenance exonuclease 1 (MGME1) that processes the 5' ends, enabling efficient ligation of newly replicated mtDNA (Uhler *et al.*, 2016), or the ribonuclease H1 (RNaseH1), that removes the RNA primer used for initiating replication (Cerritelli *et al.*, 2003).

1.3.2. Replication model

Mitochondrial replication research studies dates back to 1981 (Anderson *et al.*, 1981); however, the exact replication mechanism is still under investigation and many models remain controversial. In the literature, three replication models have been proposed: (1) the strand displacement model (SDM) (Brown *et al.*, 2005; Clayton, 1982, 2000; Shadel and Clayton, 1997), (2) the strand-coupled model (Holt *et al.*, 2007), and (3) the mtDNA replication model via ribonucleotide incorporation throughout the lagging strand (RITOLS) (Yasukawa *et al.*, 2006).

To date, the model with major experimental support is the strand displacement model (SDM), which through the years has been improved. According to this model, DNA synthesis occurs in both strands (each strand with their own replication origin) but

replication of the strands does not start at the same time (Gustafsson *et al.*, 2016). According to the SDM, mtDNA replication initiates with the synthesis of a primer at the O_H (heavy strand origin), probably by POLRMT (mitochondrial RNA polymerase) (Fusté *et al.*, 2010). Then, Pol γ begins the DNA synthesis with TWINKLE unwinding the DNA. During this first step of the synthesis, the L-strand acts as a template, and the H-strand is displaced and coated with mtSSB proteins. mtSSB proteins. The latter not only cover the displaced strand, but they also help to block transcription by preventing random initiation of RNA primer synthesis (Gammage and Frezza, 2019).
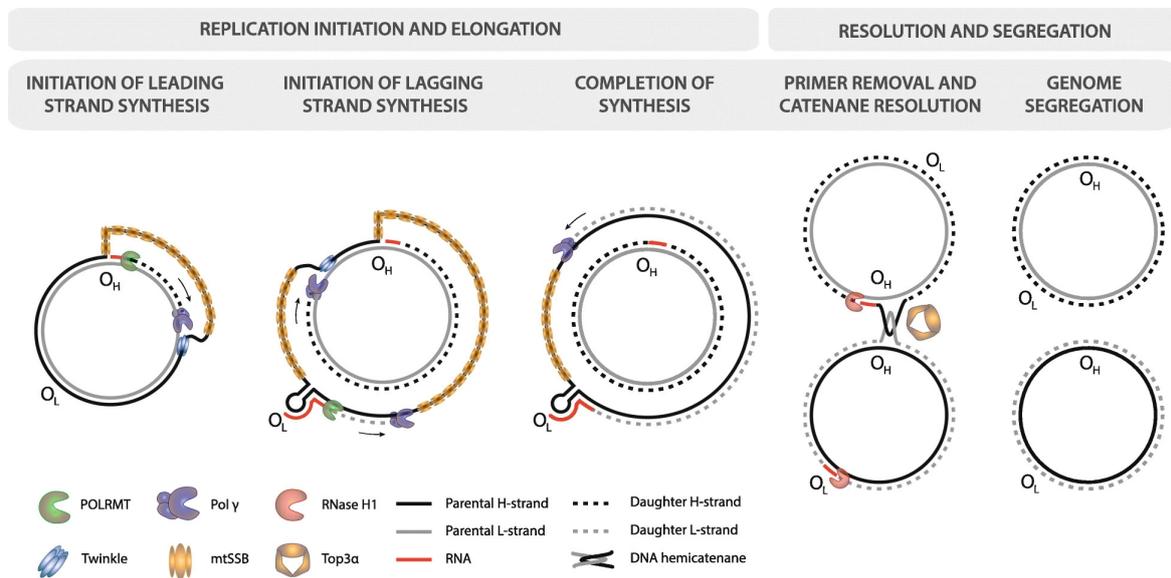
Figure 1-4. Strand-displacement mtDNA replication model. Replication initiates with the synthesis of a RNA primer by POLRMT, then Pol ⬚ begins the DNA synthesis, with TWINKLE unwinding the DNA. The L-strand acts as a template; the H-strand is displaced and coated with mtSSB proteins. When replication reaches O_L, a stem-loop is formed. The stem-loop allows the creation of an RNA primer by POLRMT, which starts the synthesis of the L-strand using the H-strand as a template. When DNA synthesis is completed, hemicatenated DNA is resolved by the topoisomerase 3a. In the end, the RNaseH1 and FEN1 remove the RNA primers. Figure from Gammage and Frezza, 2019. Figure license under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/).

When the replication machinery reaches O_L (light strand origin), which is about two-thirds of the way around the mtDNA molecule, a stem-loop is formed. This structure avoids the

binding of mtSSB proteins, allowing POLRMT to initiate the synthesis of ~25 nucleotides of an RNA primer. Then, POLRMT is replaced by Pol γ; and the synthesis of the new L-strand begins in the opposite direction, using the H-strand as a template (Gammage and Frezza, 2019; Gustafsson *et al.*, 2016). When mtDNA synthesis is completed, hemicatenated DNA is resolved by the topoisomerase 3α (TOP3α) (Nicholls *et al.*, 2018), and the RNA primers are removed, most likely by RNaseH1 or MGME1 (Uhler and Falkenberg, 2015).

1.3.3. D-loop

Previous studies indicate that more than 95% of all replication events are prematurely terminated (Bogenhagen and Clayton, 1978). This produces a small DNA product of ~650 nucleotides named 7S DNA. The 7S DNA fragment starts to be synthesized around the origin O_H until the termination associated sequences (TAS). The short 7S DNA fragment remains bound to the parental L-strand, forming in the NCR a D-loop shape structure (triple-stranded structure) (Figure 1-3A) (Falkenberg, 2018).

In mammals, the proportion of mtDNA molecules with D-loop is variable. For instance, depending on the tissue, it can range from 1% to 82% (Nicholls and Minczuk, 2014). It has been proposed that the D-loop is important for the regulation of mtDNA replication and mtDNA maintenance. For example, it has been suggested that it is a binding centre for proteins that regulate mitochondrial nucleoids dynamics, such as Pol γ or TWINKLE (He *et al.*, 2007; Jemt *et al.*, 2015). However, the precise function and regulating mechanisms at the D-loop have not been elucidated yet (Nicholls and Minczuk 2014).

1.4. Mitochondrial polymerase

1.4.1. Mitochondrial polymerase holoenzyme functions

All proteins required for mtDNA replication are encoded by nuclear DNA, including the Pol γ (polymerase gamma), the key protein for replication (Capps *et al.*, 2003). From the 16 polymerases in mammals, only the Pol γ has demonstrated to be responsible for the mtDNA synthesis and proofreading. More specifically, several studies confirm that Pol γ is localised in the mitochondria (Krasich and Copeland, 2017), and Pol γ is the only mitochondrial DNA polymerase demonstrated to be embryonically lethal when it is depleted (Hance *et al.*, 2005). Moreover, mutations in this protein have shown to increase mtDNA mutations (Chan and Copeland, 2009). The Pol γ holoenzyme is a heterotrimeric protein composed of one catalytic subunit POLG and two accessory subunits POLGB (Graziewicz *et al.*, 2006). In mammals, the catalytic subunit POLG (alternative symbols: Pol γA, POLG1 or POLGA) of ~140 kDa, is encoded by the *Polg* gene (or *Polg1*). In humans, the gene localises in chromosome 15 (current reference NC_000015.10) while in mice it localises in chromosome 7 (current reference NC_000073.6). The catalytic subunit has three main functions: DNA polymerase, 3'-5' exonuclease and 5' dRP lyase activity. Each function is encoded in a specific domain (Figure 1-5) (Graziewicz *et al.*, 2006); and sequence alignment has revealed that the DNA polymerase and the 3'-5' exonuclease domains are well conserved across species (Ito and Braithwaite, 1991).

The accessory subunit POLGB (alternative symbols: Pol γB or POLG2) of ~55 kDa, encoded by the *Polg2* gene, is also important. POLGB increases the processivity of the enzyme almost100-fold because it increases the enzyme affinity for the template DNA (Carrodeguas *et al.*, 2002; Lim *et al.*, 1999) and also, it has been suggested, because suppresses the exonuclease activity (Graziewicz *et al.*, 2006).

In 2009, the crystal structure and modelling of the human Pol γ showed that the POLG catalytic subunit is closely related to the polymerase from the bacteriophage T7, and adopts a canonical "right-hand" configuration. In this configuration, a positively charged channel formed by the thumb, palm, and fingers of the POLG surrounds the mtDNA. Because the POLG spacer domain (between the *pol* and *exo* domains) interacts only with one POLB

monomer, it has been proposed that the DNA-binding affinity is specific of the catalytic subunit POLG (Lee *et al.*, 2009).
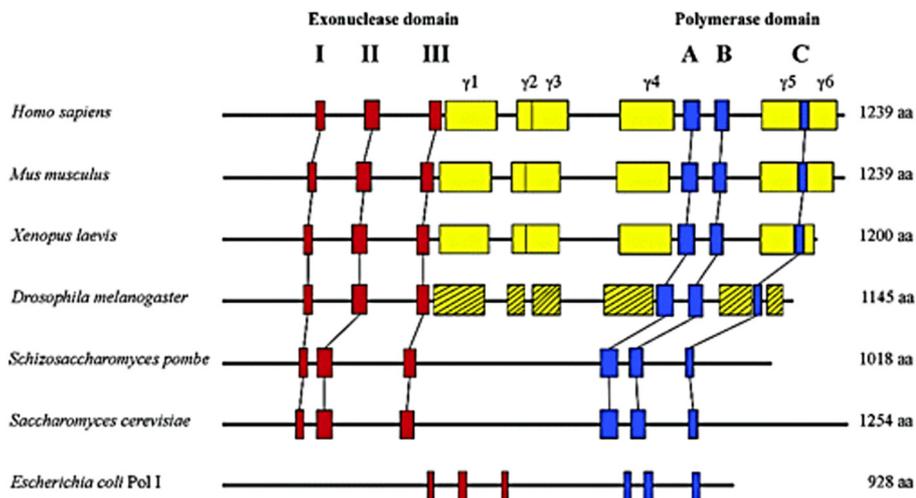
Figure 1-5. Scheme of the conserved domains in the catalytic subunit POLG. In red, the 3'-5' exonuclease domains encoded by I, II and III motifs. In blue, the polymerase domain, encoded by the A, B and C motifs. In yellow, the conserved sequences in vertebrates (weakly conserved in insects). To the right, the number of amino acids predicted for the PolG protein of some species. Figure from Graziewicz 2006. Figure not subject of U.S. Copyright but adapted from Kaguni 2004.

1.4.2. POLG catalytic subunit

The catalytic subunit POLG is part of the A polymerases family, had a broad pH range (7.5-8) and requires a divalent metal cation (Mg^{2+} cations are preferred when processing DNA templates) (Kaguni, 2004).

POLG contains several functions, for example, the 5' dRP lyase activity. This function removes the 5'-terminal dRP sugar moiety generated by glycosylases and AP endonucleases during DNA damage repair (therefore important for the base-excision repair in mitochondria) (Kaguni, 2004). Additionally, *in vitro* studies showed that Pol γ also possess

a reverse transcriptase activity and other catalytic RNA-associated activities (Murakami *et al.*, 2003). POLG has also been recently implicated in the degradation of linearized mtDNA (Nissanka *et al.*, 2018; Peeva *et al.*, 2018).

However, the major function of POLG is during replication. POLG displays accurate polymerisation (Lee and Johnson, 2006; Lynch, 2011) on the one hand by discriminating the correct nucleotides and on the other hand by correcting the errors with the 3'-5' exonuclease activity. So, the overall fidelity is the result of the sum of both reactions (Johnson and Johnson, 2001). Polymerase selectivity is complex, as each polymerase needs to select the correct one from a pool of four similar desoxynucleotide triphosphates (dNTPs) to pair in a Watson-Crick base manner. Additionally, the identity of the correct dNTP changes with each step along the template strand (Showalter and Tsai, 2002). However, the exonuclease activity can increase the fidelity from 5-34 fold depending on the substrate and the mis-pair (Kaguni, 2004) and despite being a complex reaction, it seems not to be significantly affected by the polymerisation process (Johnson and Johnson, 2001).

1.4.3. POLG related disorders

Currently, ~180 unique disease-causing point mutations in the gene encoding the catalytic subunit POLG have been reported in humans. These mutations have been associated with a number of different diseases (Figure 1-6, Human DNA Polymerase Gamma Mutation Data Base: https://tools.niehs.nih.gov/polg/) (Nurminen *et al.*, 2017) like Progressive External Ophthalmoplegia (PEO), Alpers-Huttenlocher syndrome (AHS), childhood myocerebrohepatopathyspectrum (MCHS), myoclonic epilepsymyopathy sensory ataxia (MEMSA) and Ataxia neuropathy spectrum (ANS) (Chan and Copeland, 2009).

Symptoms usually appear in compound heterozygous patients bearing two distinct recessive alleles, or in autosomal dominant syndromes (Nurminen *et al.*, 2017). In general, symptoms can be variable like, epilepsy, myopathy, ataxia; and the time when the symptoms appear is very variable, from early neonatal to middle age. Therefore, the clinical identification of

POLG-related disorders includes the combination of symptoms, different tempos of disease progression, affected organs, neurological signs, and the sudden appearance of symptoms during stressful illnesses (like a viral infection) (Cohen and Naviaux, 2010). For example, PEO is a mitochondrial disorder transmitted mainly in an autosomal dominant way, but autosomal recessive alleles are also known). The disorder is characterised by progressive weakness of the external eye muscle responsible for eye movements, exercise intolerance, hearing loss, hypogonadism, dysphagia, and, neuromuscular problems. Symptoms often appear when the patient is between 18 and 48 years old. Another example is Alpers syndrome, characterised by progressive spastic quadriplegia, with progressive cerebral degeneration, cortical blindness, deafness and early death. In this case, the disorder is an autosomal recessive disease, affecting children of only a few years old (Chan and Copeland, 2009; Graziewicz *et al.*, 2006).

According to the crystal structure, Lee *et al.* propose that depending on the domain where the mutations are located, they can affect either the catalytic enzyme activity (like polymerase activity, exonuclease activity or dRP lyase activity), alter the affinity for the mtDNA or disrupt the interaction with the accessory subunits and decrease the processivity (Lee *et al.*, 2009). For instance, substitutions in the polymerase domain like H932Y, R943H and Y955C, which have been related to PEO disorders, show that mutant POLG's have a severe polymerase activity reduction in *in vitro* assays (Graziewicz *et al.*, 2006; Stumpf *et al.*, 2010). Moreover, *in vitro*, Y955C POLG mutant cells are 2-fold less accurate for base pair substitutions and can even increase 10-100 fold when exonuclease activity is inactivated (Ponamarev *et al.*, 2002).

The identification of mutations associated to the POLG exonuclease domain could lead to the idea that these mutations generate a mutator phenotype, due to a deficient proof reading activity. Surprisingly, to date, mutations in the exonuclease domain of the POLG that have been found in humans, do not cause a mitochondrial mutator phenotype *in vivo* nor *in vitro* (Baruffini *et al.*, 2006; Stumpf *et al.*, 2010; Szczepanowska and Foury, 2010). On the contrary, mutations in the exonuclease domain seem to increase the activity generating a severe imbalance between mtDNA polymerisation and degradation (Szczepanowska and Foury, 2010). It is unknown if some exonuclease-deficient POLG variants exist within the

healthy populations. However, a mutator phenotype is not a clinical factor displayed in POLG-related disorders (Stumpf *et al.*, 2013).
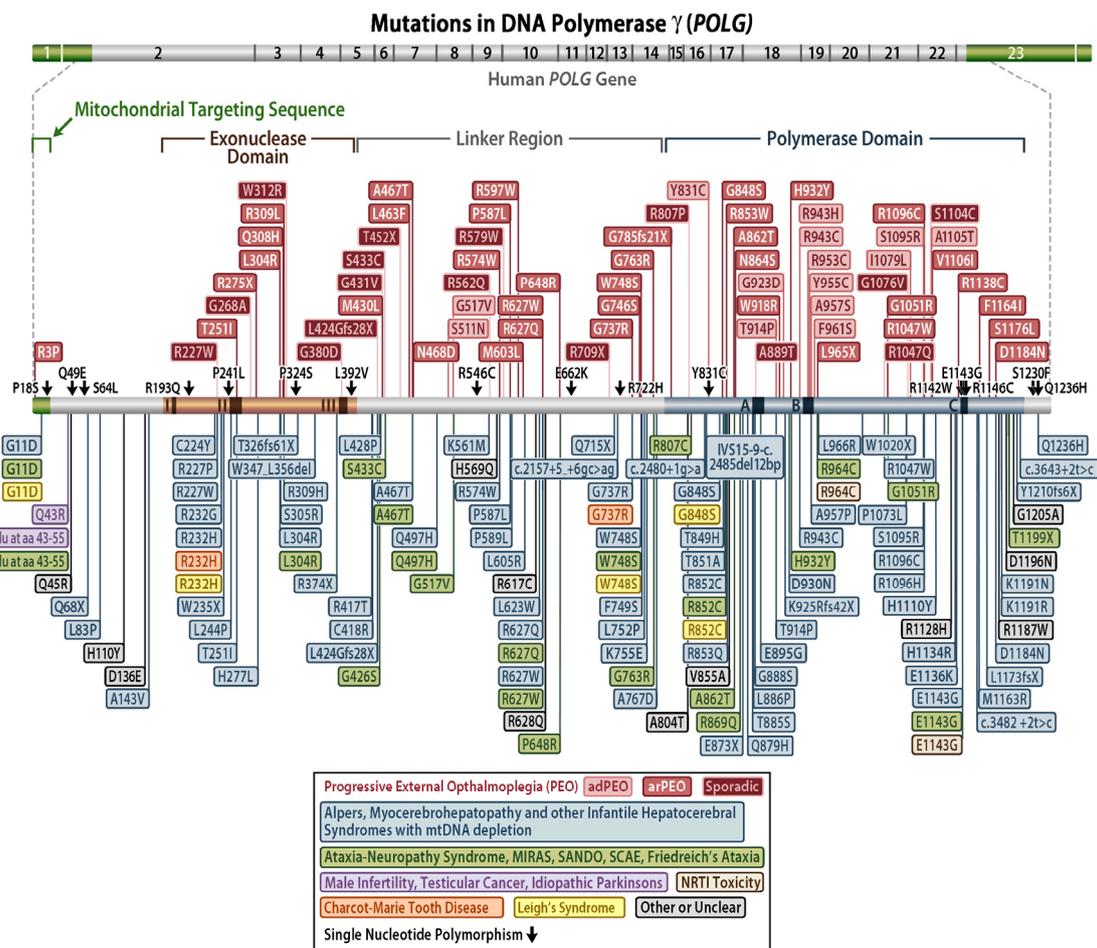
Figure 1-6. Human mutations reported to be associated to diseases. The figure shows the reported mutations in the POLG gene associated with different diseases until 20.02.2019. Figure from the National Institute of Environmental Health Sciences, https://tools.niehs.nih.gov/polg/. Rights from Copeland PhD.

POLG point mutations have not only been associated with disorders but also repeated sequences and pathogenic variants. For instance, in humans, near the N-terminus of the POLG gene, there is a CAG(n) repeat sequence encoding 13 glutamine residues (Figure 1-

6). These POLG gene CAG repeat variants have been linked to idiopathic sporadic Parkinson's disease (Luoma *et al.*, 2007), risk of testicular cancer (Nowak *et al.*, 2005) and male infertility (Rovio *et al.*, 2001). In addition, POLG pathogenic variants have been associated with MNGIE-like illness (Tang *et al.*, 2012) and Charcot-Marie-Tooth neuropathy type 2 (Harrower *et al.*, 2008; Naess *et al.*, 2009; Taanman *et al.*, 2009).

1.5. Mitochondrial POLG mutator mouse model

1.5.1. History of mitochondrial POLG mutator mouse models

Studies of model organism engineered with mitochondrial polymerase mutations date back to 1992 when several mutations where introduced in the gene *MIP1* that codes for the mitochondrial DNA polymerase of yeast (Foury and Vanderstraeten, 1992). Results from those studies concluded that specific mutations in the exo1, exo2, and exo3 of MIP1 reduce the fidelity of replication by reducing the exonuclease activity (Foury and Vanderstraeten, 1992; Vanderstraeten *et al.*, 1998). To date, several transgenic mice with variants in the POLG gene have been constructed. Of those, the majority of mutant mice contain a deficient proofreading exonuclease. There is only one reported mouse model with a POLG ectopically expressed mutation in the polymerase domain. This mouse carries a POLG Y955C point mutation, targeted at the heart, and resembles the same mutation that causes PEO (progressive external ophthalmoplegia). The mice displayed cardiomyopathy, with signs of oxidative stress and premature death (Lewis *et al.*, 2007).

In 2000, Kasahara and colleagues constructed a transgenic mouse with the ectopic expression of a POLG mutant (D181A), under a mouse cardiac-specific α-myosin heavy chain (α-MHC) promoter. This mouse displays signs of cardiomyopathy with no apparent deleterious effects on mitochondrial protein content, gene expression or respiratory functions. This model also showed mtDNA mutations and deletions (Zhang *et al.*, 2000).

Another mouse with the same D181A proofreading-deficient POLG expressed ectopically, but neuronal-specific targeted under the promoter of mouse Ca^{2+}/calmodulin-activated protein kinase II α (CaMKIIα), was published in 2006. This mouse exhibited forebrain-specific defects in mtDNA. As the main phenotype, the mouse displays altered intra-day wheel-running activity rhythm without any detectable abnormalities in sensorimotor functions, like learning or memory (Kasahara *et al.*, 2006).

A breakthrough occurred when two independent research groups (Larsson and Prolla groups) created a constitutive knock-in mouse (D257A) with a disrupted POLG proofreading function. These mice, named mtDNA mutator mice, display a remarkable characteristic: they recapitulate several signs of premature ageing (Kujoth *et al.*, 2005; Trifunovic *et al.*, 2004). To date, the mutator mouse is the most studied POLG mutant mouse model.

1.5.2. Phenotype of mitochondrial POLG mutator mice

The knock-in mitochondrial mutator mouse carries a missense mutation with a single aspartate mutated to alanine in the residue 257 (D257A). This mutation, in the exonuclease domain of POLG, drastically reduce the 3'-5' proof-reading activity of the polymerase. As a result, the homozygous mutator mouse shows high levels of point mutations, reaching 20-30 mutations per molecule of mtDNA (Ross *et al.*, 2013).

The literature also indicates that mtDNA mutator mice show an increased amount of unusual linearised mtDNA deletions. It has been reported that the homozygous mutator mouse displays an increased amount of linear mtDNA deletions (25-30% of mtDNA), encompassing the region between the origins of replication O_H and O_L (Trifunovic *et al.*, 2004). These deletions are possibly the result of impaired ligation, which cause nicks near to the replication origin of the heavy strand. These nicks generate that during the subsequent round of replication, linear fragments form due to double-strand breaks (Macao *et al.*, 2015). Additionally, the accumulation of circular mtDNA molecules with large deletions have been

reported whereas, some authors have not detected these type of deletions (Ameur *et al.*, 2011; Williams *et al.*, 2010), others estimate, that from the total mtDNA only ~1% or fewer molecules contain large deletions (Kraytsberg *et al.*, 2009; Vermulst *et al.*, 2008). Finally, it has also been found, that mutator mice accumulate control region multimers (CRMs) in the brain, heart and skeletal muscle. These CRMs are repeated sequences around the 5' end of the mtDNA control region (Williams *et al.*, 2010).

At the biochemical level, the homozygous mutator mouse cells shows decreased oxygen consumption, reduction in ATP content (Trifunovic *et al.*, 2005) and severely impaired stability of the respiratory chain complexes (Edgar *et al.*, 2009). Consequently, the mutator mouse displays mitochondrial dysfunction which has been proposed as the primary inducer of the premature aging phenotype (Edgar and Trifunovic, 2009). Additionally, this mouse model has shown no increased level of ROS (Hiona and Leeuwenburgh, 2008; Kujoth, 2007; Trifunovic *et al.*, 2005), which argues against the role of oxidative stress as an inductor of aging in the mouse model (Alexeyev, 2009).

To date, three mtDNA mutator mouse strains with the same D257A mutation have been described in the literature. In 2004, Larsson and colleagues (Trifunovic *et al.*, 2004) generated the first D257A mutator mouse and have maintained it on the C57BL/6NCrl mouse substrain. They observed premature ageing-related phenotypes such as weight loss, alopecia, kyphosis (curvature of the spine), reduced subcutaneous fat, osteoporosis, anaemia, reduced fertility, reduced testis size, heart enlargement and drastically reduced life span. To generate the mutation, they used a FLP-FRT recombination system to produce a change at the position 257 from an aspartic acid (GAC) to an alanine (GCC). In detail, for the construct, the PolgA gene clone derive from the 129/SvJ l Fix II phage library (Stratagene). This construct was used for targeting knock-in 129/SvJ embryonic stem cells, which were introduced into blastocyst to create chimeric mice that transmitted the target mutation. The line was subsequently maintained on the C57Bl/6NCrl substrain. As a result of the methodology used, the final allele maintains two *loxP* sites flanking the mutation, and one *Frt* site upstream of it.

In 2005, Prolla and colleagues (Kujoth *et al.*, 2005) generated a second D257A mutator mouse and maintained it on the C57BL/6J mouse substrain (Kujoth, 2007). They observed

the same premature age-related phenotypes than previously reported by Larsson's group, in addition to presbycusis (hearing loss), sarcopenia, thymic involution and loss of duodenum crypts. In the report (Kujoth, 2007), they also showed the induction of apoptotic markers. To generate the mutation, they used a *Cre-lox* recombination system to generate a change at the position 257 from an aspartic acid (GAC) to an alanine (GCT). Genomic manipulation was carried out on a clone from the same 129/SvJ l Fix II phage library (Stratagene), and the manipulated construct was also introduced into 129Sv ES cells. The final allele contains a *loxP* site downstream of the D257A mutation. In 2013, Hayashi and colleagues (Mito *et al.*, 2013) published a third D257A mutator generated in the C57BL/6NCrl substrain showing the same premature ageing phenotype. In this report, they used the same methodology as in Trifunovic *et al.* to generate the mutation (Mito *et al.*, 2013).

The three mutator mouse models show the same premature ageing phenotype but they differ in time of symptoms onset. Larsson and colleagues reported that symptoms in the homozygous mutator start to appear at 25 weeks (Trifunovic *et al.*, 2004). In comparison, Prolla and colleagues reported the beginning of symptoms at 36 weeks (Kujoth *et al.*, 2005). Moreover, the lifespan reported it is variable. Initial research with the mutator mice from Larsson laboratory have reported a median lifespan of 48 weeks (Trifunovic *et al.*, 2004); in contrast, reports from Prolla and colleagues reported a median lifespan of 60 weeks (Kujoth *et al.*, 2005) and 65 weeks (Someya *et al.*, 2017) for the homozygous mutator.

It is not yet known if the differences in lifespan and the onset of the phenotype are due to variation of the genetic background (C57BL/6J and C57BL/6NCrl), or if such differences are due to the specific allele used for each mutator mice. Even so, these studies do not take into account the details of the breeding. In 2013, Ross *et al.* reported that maternally transmitted mutations can reduce lifespan. Taking as a reference the mutator mice from Larsson research group, they determined that the standard breeding (with low level maternally transmitted mutations) show a median lifespan of 42 weeks, but using a breeding scheme which avoids maternally transmitted mutations, the lifespan can increase up to 50 weeks. Therefore, it has been difficult to compare previous reports side by side.

1.5.3. Mutator mouse as a research model

Ageing is a process of fitness decline observed over time (Kauppila *et al.*, 2017). Although ageing is more evident among humans and animals maintained in captivity, a large number of studies have shown that the majority of metazoans in the wild show signs of ageing (Kauppila *et al.*, 2017; Nussey *et al.*, 2013). The main cause of ageing is unknown, however, a number of hallmarks that could contribute to the ageing process are: (1) genomic instability, (2) telomerase attrition, (3) epigenetic alterations, (4) loss of proteostasis, (5) deregulated nutrient sensing, (6) altered intracellular communication, (7) cellular senescence, (8) stem cell exhaustion and (9) mitochondrial dysfunction (López-Otín *et al.*, 2013).

Alterations of mitochondrial function as important players in the ageing process are well documented in the literature (Haas, 2019). For instance, mitochondrial dysfunction has been widely related to the aging of the immune system (McGuire, 2019), skin ageing (Stout and Birch-Machin, 2019), and ageing-related illnesses like cardiomyopathies (Duran *et al.*, 2019), diabetes (Montgomery, 2019), Parkinson (Chen *et al.*, 2019) and Alzheimer disease (Weidling and Swerdlow, 2019). Mutations of mtDNA are hypothesised to cause mitochondrial dysfunction in ageing (Kauppila *et al.*, 2017). Since ~93% of the mitochondrial genome encodes important mitochondrial genes (Andrews *et al.*, 1999; Kolesnikov and Gerasimov, 2012), the probability that random mtDNA mutations alters an important sequence is high. Mutations will then be clonally expanded during the years (Elson *et al.*, 2001; Greaves *et al.*, 2012; Wanrooij *et al.*, 2012) leading to mitochondrial dysfunction (Edgar *et al.*, 2009). Moreover, evidence that supports this theory is that ageing in humans has been associated with an increase of mtDNA deletions and point mutations. During ageing in post-mitotic tissues the predominant mtDNA mutations are large-scale mtDNA deletions (Cortopassi and Arnheim, 1990; Cortopassi *et al.*, 1992; Kraytsberg *et al.*, 2006), and in mitotic tissues mtDNA point mutations are more commonly identified (Greaves *et al.*, 2006, 2010; Shin *et al.*, 2004). In addition, maternally transmitted mtDNA mutations can reduce lifespan in animals with a wild-type nuclear background (Ross *et al.*, 2014). However, it is unknown if these mtDNA mutations are the cause or a consequence of ageing (Payne and Chinnery, 2015).

In that regard, the mutator mouse model supports the hypothesis that mtDNA mutations could be implicated in ageing, as the mouse exhibit a premature progeria phenotype that resembles in some extent human clinic symptoms of ageing and display mitochondrial dysfunction with increased number of point mutations and deletions (Kujoth *et al.*, 2005; Trifunovic *et al.*, 2004). Evidence that supports the hypothesis that mtDNA point mutations are the driving force of the ageing phenotype in the mutator mouse is that the accumulation of mtDNA point mutations impairs the stability of the respiratory complexes, by amino acid substitutions leading to mitochondrial dysfunction (Edgar *et al.*, 2009). In addition, it has been demonstrated that patterns of clonally expanded mtDNA mutations in colons of heterozygous mutator mice, resemble the patterns exhibited in ageing humans (Baines *et al.*, 2014).

However, ageing research using the mutator mouse as an ageing model should be interpreted with caution. Like any other mouse model of ageing, the mutator mouse phenotype does not fully represent normal ageing and the mechanisms of premature ageing could be different from the normal ageing process (Kõks *et al.*, 2016). For example, mtDNA deletions found in the mutator mouse and in humans are different. Most of the mtDNA deletions that are found during normal ageing affect the major arc of mtDNA (Dong 2014, Samuels 2004) and are classified into two types. Type I, deletions flanked by homologous or near-homologous repeats; and type II, deletions not flanked by homologous repeats and co-localised in secondary DNA structures like hairpins, or G-quadruplex structures (Damas 2012 and Dong 2014). In contrast, the majority of mtDNA deletions in the mutator mouse affect the minor arc, and in some tissues has been described low levels of deletions between homologous sequences (Kraytsberg *et al.*, 2009; Vermulst *et al.*, 2008). In addition, in a recent publication, Hämäläinen *et al.* show that errors during mtDNA replication in the mutator mouse might induce nuclear DNA replication stress. In the publication, they show evidence that in mutator mouse induced pluripotent stem cells (iPSC) there is increased replication stress and DNA damage in the nucleus. They also show that the dNTP sequestering generated in the mitochondria leads to an imbalance in the dNTP pools used for the nucleus (Hämäläinen *et al.*, 2019).

The mtDNA mutator mouse model has been used to study not only the effects of a deficient exonuclease POLG on mtDNA mutations and ageing; but also to decipher the molecular mechanisms of the distribution and transmission of mtDNA mutations (Stewart and Chinnery, 2015; Stewart and Larsson, 2014; Stewart *et al.*, 2008). Moreover, the mutator mouse has been proved to be an essential tool to generate novel mouse models that carry pathogenic mtDNA mutations which are important for the study of mitochondrial diseases (Kauppila *et al.*, 2016).

2. RESEARCH AIMS

The first mtDNA mutator mouse was generated almost 16 years ago (Trifunovic *et al.*, 2004). Since that moment, it has become a powerful tool to study and investigate mtDNA mutations in general, and mtDNA mutations in one of the most intriguing scientific topics: ageing.

This mutator mouse model carries an altered mtDNA polymerase with a mutation in the exonuclease domain that results in increased mtDNA mutations and, additionally, recapitulates several signs of premature ageing. To date, three mtDNA mutator mouse strains with the same D257A mutation have been described in the literature. In general terms, the first mutator mouse was maintained in the C57BL/6NCrl mouse substrain, in the Larsson research group (Trifunovic *et al.*, 2004). One year later, using another recombination strategy to generate the mutation, the Prolla research group published a second mutator mouse, which was maintained in the C57BL/6J mouse substrain (Kujoth *et al.*, 2005). The reports showed that both mutator mouse models contain the same D257A *POLG* mutation but with genetically allele variations in non-coding sequences. Recently, a third mutator was generated in the C57BL/6J substrain with a POLG allele like the one described by Larsson research group (Mito *et al.*, 2013).

Interestingly, all of them show a similar premature ageing phenotype. However, they differ drastically in the time when the symptoms appear and, even more remarkable, exhibit striking differences regarding lifespan. For instance, mutator mice reported by Trifunovic *et al.*, (2004) and Mito *et al.*, (2013), showed a median lifespan of 48 and 40 weeks respectively, while the mutator reported by Kujoth *et al.*, (2005), shows a higher median lifespan of 59 weeks. In addition, there was a reported 11 weeks difference in the onset of symptoms between Trifunovic *et al.*, (2004) and Kujoth *et al.*, (2005) (25 weeks vs 36 weeks).

Therefore, the aim of my Ph.D. project was:

1) By evaluating and characterising different mutator mouse lines, determine if previously reported variations among the mtDNA mutator mouse models were because of:

a) Differences in the genetic background

b) Differences in the allele expressed

2) Investigate the molecular mechanisms responsible for the phenotypical variation between the mtDNA mutator mouse models.

3. RESULTS

3.1. Description, generation, and maintenance of mtDNA mutator mouse lines.

To date, three mtDNA mutator mouse strains with the same D257A mutation have been described in the literature by independent groups using two different recombinant mutation strategies and two different mice sub-strains to maintain the mice lines (Kujoth *et al.*, 2005; Mito *et al.*, 2013; Trifunovic *et al.*, 2004). However, the reports showed a drastic variation in lifespan and in the time when the symptoms appear.

Therefore, in order to determine the reasons behind the progeria variation rate among the previously reported mtDNA mutator mouse models, we decided to evaluate two hypotheses 1) Differences are due to the mouse substrain used and 2) Differences are due to the allele expressed. To test our ideas, we used a series of mtDNA mutant mouse lines that allowed us to discriminate between these two hypotheses (Table 3-1).

In-house, we had access to the original mtDNA mutator mice from Trifunovic *et al.*, (2004), that in this study we name B6NPolg$^{Lar/Lar}$. This mouse line contains the D257A mutation, labelled here as *Lar* allele, maintained in the C57BL/6NCrl substrain.

To test our first hypothesis, we wanted to compare mutators with different genetic backgrounds but same allele. Therefore, we generated a new mouse line, named here B6JPolg$^{Lar/Lar}$, which allowed us to compare the allele *Lar* but in the C57BL/6J background.

To generate the B6JPolg$^{Lar/Lar}$ mouse line, we backcrossed B6NPolg$^{wt/Lar}$ males with C57BL/6J wild-type females. Selecting from the offspring heterozygous males B6JNPolg$^{wt/Lar}$ and backcrossing them with C57BL/6J wild-type females for more than 10 generations, we were able to obtain heterozygous B6JPolg$^{wt/Lar}$ mice with 99.99% C57BL/6J genetic background. These heterozygous B6JPolg$^{wt/Lar}$ mice were mated to obtain homozygous B6JPolg$^{Lar/Lar}$ mice (Figure 3-1).

Label	Substrain	Allele 1	Allele 2	Color code
wt	C57BL/6J	Wild-type	Wild-type	⬜
B6NPolg$^{Lar/Lar}$	C57BL/6NCrl	Lar	Lar	🟩
B6JPolg$^{Lar/Lar}$	C57BL/6J	Lar	Lar	🟪
B6JPolg$^{Pro/Pro}$	C57BL/6J	Pro	Pro	🟦
B6JPolg$^{Lar/Pro}$	C57BL/6J	Lar	Pro	
B6JPolg$^{Lar/ko}$	C57BL/6J	Lar	-	
B6JPolg$^{Pro/ko}$	C57BL/6J	Pro	-	
B6JPolg$^{Mut/ko}$	C57BL/6J	Pro or Lar	-	🟥

Table 3-1. Mouse lines used in this study. The table summarises the characteristics of the mouse lines and the colour code used in the graphs of this study.

To test our second hypothesis, we wanted to compare mutators with the same genetic background but different alleles. Therefore, from Jackson Laboratories, we purchased the original mutator mice published by Kujoth *et al.*, (2005). In this study, we name this mouse B6JPolg$^{Pro/Pro}$. This mouse contains the D257A mutation, labelled here *Pro* allele, maintained in the C57BL/6J substrain. Comparing this new line B6JPolg$^{Pro/Pro}$ with the previously generated B6JPolg$^{Lar/Lar}$, allowed us to analyse different alleles (*Lar* and *Pro*) at the same genetic background (C57BL/6J) (See Table 3-1). We did not analyse the allele generated by Mito *et al.*, (2013) due to a lack of access to the mouse.

It is important to note that we always used a breeding scheme lacking maternally transmitted mtDNA mutations, previously described by Ross *et al.*, (2013). In this scheme, the mutation D257A was maintained in heterozygous males, in order to avoid mtDNA mutations inherited maternally by germline transmission.

During this study, we also analysed a compound heterozygous mutator mouse B6JPolg$^{Lar/Pro}$ for splicing analysis, and two hemizygous mutant lines B6JPolg$^{Lar/ko}$ and B6JPolg$^{Pro/ko}$ for mtDNA sequencing. The complete list with characteristics of each line can be found in Table 3-1.
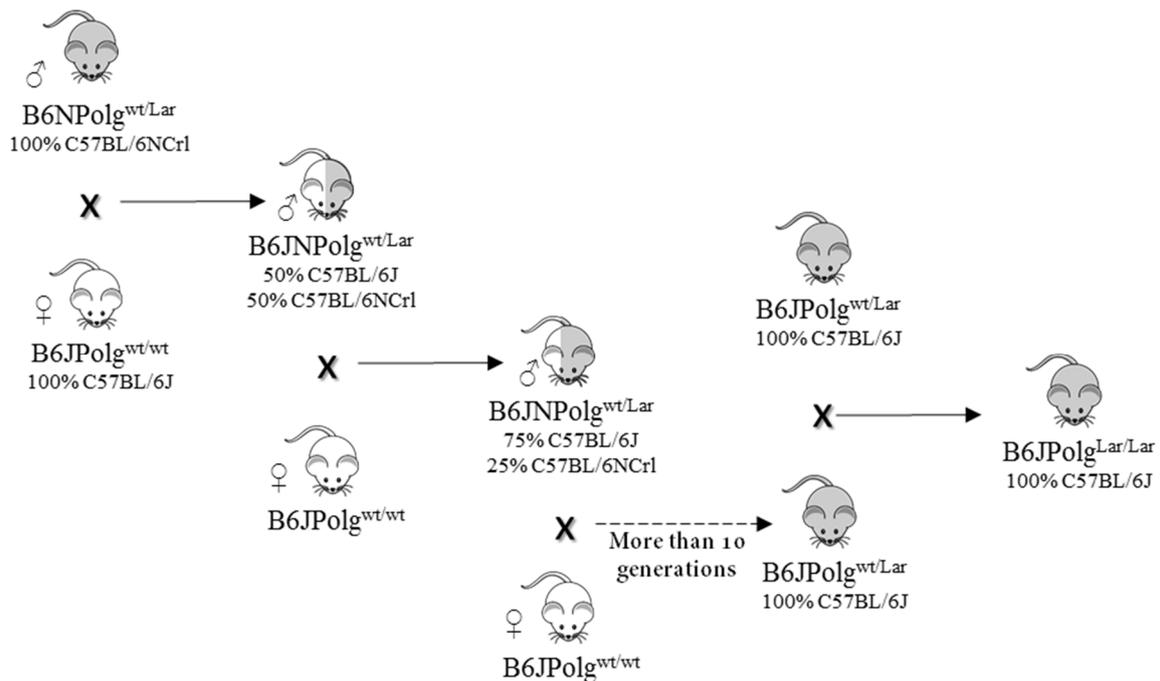
Figure 3-1. Breeding strategy to generate the B6JPolg^Lar/Lar^. A male heterozygous *Lar* mutant was backcrossed with wild-type females C57BL/6J for more than 10 generations, as a result, a heterozygous *Lar* mutant was obtained with 99.9% C57BL6J genetic background. Finally, homozygous mutants were obtained crossing two heterozygous mutants.

3.2. Increased Cardiomyopathy and Anaemia in mice with the *Lar* allele.

Studies have reported that mutations and deletions in the mtDNA can lead to impairment of mitochondria, causing tissue dysfunction (reviewed in Bates *et al.*, 2012; Bratic and Larsson, 2013; Kauppila *et al.*, 2017).

The heart for example, an organ rich in mitochondria, is assumed to be very sensitive to mitochondrial dysfunction due to its high oxygen consumption rate (Dai *et al.*, 2010). In addition, it has been demonstrated that mutations in the mtDNA have a big impact on the cardiac progenitor cell differentiation function (Orogo *et al.*, 2015). Interestingly, cardiomyopathy is one of the best-characterised phenotypes in the mtDNA mutator mouse (Dai *et al.*, 2010; Orogo *et al.*, 2015; Safdar *et al.*, 2011; Trifunovic *et al.*, 2004; Woodall *et al.*, 2019). Therefore, to investigate whether there is a difference in cardiomyopathy

between these three different mutator mouse lines (Table 3-1), we screened for heart enlargement, which is a marker of heart disease.

When we analysed the weight of the heart in relation to the body of mutators at 10 weeks old, we did not find a difference in comparison with wild-type animals. However, at 35 weeks old, when the ageing phenotype is evident in the mutator, the data showed a heart enlargement in the three mutator mouse lines compared to wild-type animals. The B6NPolg$^{Lar/Lar}$ mouse displayed the greatest enlargement, followed by the B6JPolg$^{Lar/Lar}$ mouse, and the B6JPolg$^{Pro/Pro}$ mouse with the least amount of enlargement (Figure 3-1A).

Another well-characterised phenotype in the mtDNA mutator mouse, which is caused by the impairment of mitochondria function, is anaemia (Reviewed by Ahlqvist *et al.*, 2015). Actually, severe anaemia has been described as the main cause of death in the mutator mouse and data suggested it is due to a dysfunctional repopulation of haematopoietic stem cells, myeloid differentiation problems and defective mitophagy during erythrocyte maturation (Ahlqvist, 2015; Ahlqvist *et al.*, 2012; Chen *et al.*, 2009; Norddahl *et al.*, 2011).

Anaemia is defined as a reduction in haemoglobin, haematocrit or red blood cell count. Therefore, we performed blood cell count of erythrocytes and measured haemoglobin concentration in 35 week-old mice. Our data showed, as was previously reported (Dillon *et al.*, 2012; Trifunovic *et al.*, 2004), that mutator mice contain significantly lower haemoglobin and erythrocyte concentrations in peripheral blood, when compared to wild-type mice. Notably, we saw only a mild decrease in B6JPolg$^{Pro/Pro}$ (22% reduction) and a strong decrease in mutator mice with *Lar* allele (B6NPolg$^{Lar/Lar}$ and B6JPolg$^{Lar/Lar}$) (more than 40% reduction) (Figure 3-2C).

Previously, anaemia in the mutator mouse has been described as macrocytic with increased red pulp (Trifunovic *et al.*, 2004), two characteristics that can lead to spleen enlargement (splenomegaly) (Pivkin *et al.*, 2016). Therefore, we analysed the ratio of spleen weight to body weight in 35-week-old mice. Strikingly, B6JPolg$^{Pro/Pro}$ weight was not altered in comparison with wild-type animals, but B6NPolg$^{Lar/Lar}$ and B6JPolg$^{Lar/Lar}$ animals had a spleen increase of almost twice the weight of the wild-type (Figure 3-2B).

Our observations where anaemia in B6JPolg$^{Pro/Pro}$ mice is milder than the anaemia exhibited in B6NPolg$^{Lar/Lar}$ and B6JPolg$^{Lar/Lar}$ animals, led us to suspect that the main factor for the

phenotypical differences is the allele expressed. This idea was supported by the cardiomyopathy data, where the less affected mouse line was the mutator B6JPolg$^{Pro/Pro}$.
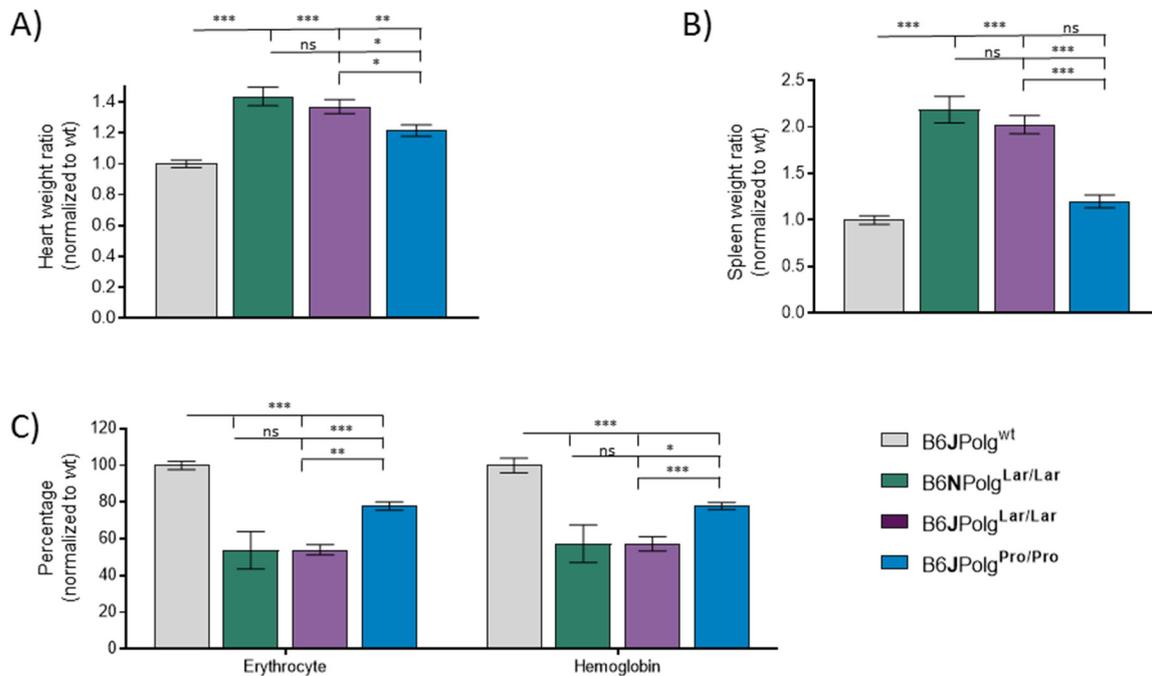
Figure 3-2. Anaemia and cardiomyopathy phenotype in the mtDNA mutator mouse lines. A) Heart weight ratio of 35 week-old mice (normalised to wild-type) (n=3-10). B) Spleen weight ratio of 35 week-old mice ratio (normalised to wild-type) (n=5-11). C) Percentage of erythrocytes and haemoglobin on peripheral blood of mice at 35 week-old (n=4-8). Data are represented as mean ± SEM. Two-way ANOVA *p<0.05; **p<0.01; ***p<0.001.

3.3. Male mutators with *Lar* allele are sterile while mutators with *Pro* allele remain fertile

Studies in humans and mice suggest that male age-related-infertility, in which the main feature is a deficient semen quality, correlates with impaired mitochondrial function (Gao *et al.*, 2016; Nakada *et al.*, 2006; Spiropoulos, 2002; Vaught and Dowling, 2018). In addition, studies have shown a link between male infertility and mtDNA point mutations and deletions, ultimately correlating the severity of the mitochondrial diseases with

infertility severity (Ji *et al.*, 2017; Martikainen *et al.*, 2017). Moreover, it has been described that the Larsson variant mtDNA mutator mouse shows reduced male fertility (Trifunovic *et al.*, 2004). Therefore, we decide to assess fertility by analysing several aspects related to this phenotype.

First, in order to determine the number of pups generated in each mouse line, we mated homozygous *POLG* male mutants to wild-type females. Our results showed that, in line with previous results, B6NPolg[Lar/Lar] were infertile (Ross *et al.*, 2013; Trifunovic *et al.*, 2004), but surprisingly the B6JPolg[Pro/Pro] only exhibited a reduced fertility (reduction of 42.8%) in comparison with wild-type mice (Figure 3-3A).
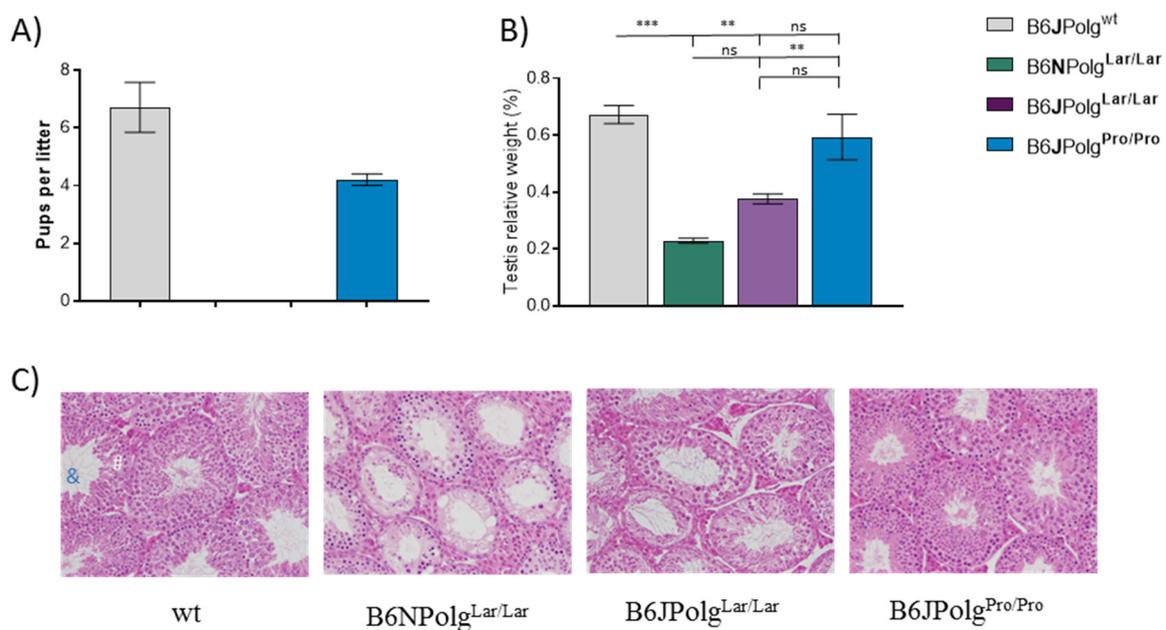
Figure 3-3. Male fertility analysis. A) Number of pups per litter (n=5-7). B) Relative testes weight (%) at 10 weeks (n=3). Data are represented as mean ± SEM. Two-way ANOVA *p<0.05; **p<0.01; ***p<0.001. C) Haematoxylin and eosin staining (H&E) of the seminiferous epithelium in testes, sections from 10 week-old mice. Marked in the wt picture are the germinal epithelium (#) and spermium (&).

Next, we measured the testes weight. We found that B6NPolg[Lar/Lar] and B6JPolg[Lar/Lar] testes were substantially lighter than those of wild-type mice; while B6JPolg[Pro/Pro] showed no significant difference in comparison with wild-type animals (Figure 3-3B).

In addition, we analysed the seminiferous tubule morphology of the testes in sexual mature mice (10 weeks) because tubules display all spermatogenesis stages with high degree of cellular heterogeneity (Ernst *et al.,* 2019; De Rooji and Russell, 2000). The results showed no visible defects present in the B6JPolg$^{Pro/Pro}$, while B6NPolg$^{Lar/Lar}$ and B6NPolg$^{Lar/Lar}$ exhibited a much more deteriorated structure: reduced germinal epithelium, a near-lack of spermium, and complete lack of spermatids (Figure 3-3C). These histological results in addition to the testes weight data suggested that mutator lines with the *Lar* allele (B6NPolg$^{Lar/Lar}$ and B6JPolg$^{Lar/Lar}$) have remarkably deteriorated testes, while B6JPolg$^{Pro/Pro}$ maintains a normal testis morphology.

Since studies have demonstrated that mitochondrial respiratory defects can give rise to sperm abnormalities (Folgerø *et al.*, 1993; Nakada *et al.*, 2006; Vaught and Dowling, 2018), we next decide to compare the sperm quality among the mtDNA mutator mouse lines.

First, we analysed the normal sperm count. In this regard, B6JPolg$^{Pro/Pro}$ maintained normal sperm counts while B6NPolg$^{Lar/Lar}$ and B6JPolg$^{Lar/Lar}$ showed a significant reduction compared to the wild-type. Interestingly B6NPolg$^{Lar/Lar}$ displayed the highest drop with almost a 90% reduction (Figure 3-4A).

Second, we evaluated total sperm motility, which plays a major role in achieving pregnancy. According to with the 2010 WHO (World Health Organization) report (Cao *et al.*, 2011), sperm motility is graded as follows: a) sperm with progressive motility, that are able to swim fast and in a straight line, b) sperm with non-progressive motility, that tend to travel in a curved or crooked motion, c) sperm either move their tail or are immotile. Many sperm have limited motility or even are immotile, and are restricted to the same spot, therefore not able to fertilise an oocyte. Since sperm need to swim forward and fast to fertilise the egg, motility refers to the sum of progressive sperm and non-progressive sperm capable to fertilise the egg, being the progressive sperm the ones with more chances to achieve fertilisation.

Using computer-aided sperm analysis (CASA), we quantify the percentage of total motile sperm, which is sperm that exhibit motility of any form. Our results indicated that the 3 mutator lines had a decreased sperm motility (Figure 3-4B), with B6JPolg$^{Pro/Pro}$ being the least affected, with only a 50% decrease; whereas B6JPolg$^{Lar/Lar}$ was the most affected, with almost no motile sperm. Then, when we quantify the percentage of progressive sperm,

which is sperm that exhibit rapid and linear movement. Our results indicated that all three mutator mouse lines were similar affected in this regard (Figure 3-4B).
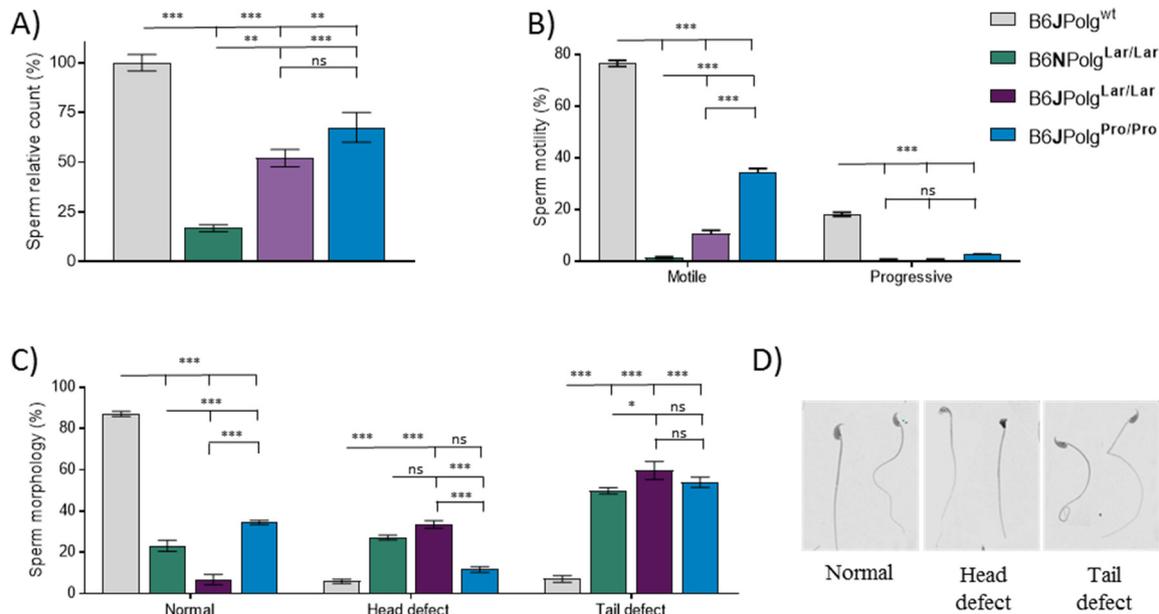
Figure 3-4. Sperm analysis. A) Sperm relative count at 10 weeks (n=3). B) Sperm motility (%) at 10 weeks (n=3), C) Quantification of sperm morphology: normal sperm, sperm with head defects and sperm with tail defects (n=3). D) Images show examples of variable sperm morphology. Data are represented as mean ± SEM. Two-way ANOVA *p<0.05; **p<0.01; ***p<0.001. Relative levels were normalised to the wt average.

Third, we evaluated sperm morphology by visually counting; (i) the number of normal sperm, (ii) sperm with head defects (heads lacking the typical hook structure or with irregular forms) and (iii) sperm with tail defects (folded or broken tails) (Figure 3-4C). Results showed that the three mutator mouse lines exhibited a significant reduction of normal sperm in comparison with the wild-type: B6JPolg^Pro/Pro (60% reduction), B6NPolg^Lar/Lar (73% reduction), and the most severe in B6JPolg^Lar/Lar (89% reduction). We also found a remarkable tail defect increase on the three mutator mouse lines. Interestingly, only mutators with the *Lar* allele (B6JPolg^Lar/Lar and B6NPolg^Lar/Lar) showed an increase in the proportion of sperm with head abnormalities, while B6JPolg^Pro/Pro levels remained similar to the wild-type.

The data clearly showed that despite B6JPolg$^{Pro/Pro}$ exhibiting a slight reduction in sperm quality, mice with this genotype are fertile, with normal testes weight and normal morphological testes structures. In contrast, mutator mouse lines with the *Lar* allele (B6JPolg$^{Lar/Lar}$ and B6NPolg$^{Lar/Lar}$), are infertile and exhibit abnormalities in testes morphology and sperm quality. Among them, the most severe phenotype is showed in the B6NPolg$^{Lar/Lar}$ line, in which the testes weight, sperm count and sperm motility values are highly altered.

Finally, based on the observation that the allele expressed is a determinant factor in the fertility phenotype, we decided to investigate the Mendelian transmission in mutators with the same nuclear genetic background (C57BL6/J). When we crossed wild-type animals with heterozygous mutants, no significant difference from the expected Mendelian transmission rate was found for the *Lar* and *Pro* alleles. However, results from intercrosses of heterozygous mutants, the *Lar* allele produced only 8.5% of Polg$^{Lar/Lar}$ pups, less than expected for Mendelian proportions. In contrast, the *Pro* allele did not significantly differ from normal Mendelian transmission (Table 3-2).

	Total pups	Pups obtained			Pups expected			P-value (Chi-Test)
		mut/mut	wt/mut	wt/wt	mut/mut	wt/mut	wt/wt	
wt/wt X wt/mut								
B6JPolg$^{Lar/wt}$	325		140	185		162.5	162.5	0.1662
B6JPolg$^{Pro/wt}$	292		130	162		146	146	0.2731
wt/mut X wt/mut								
B6JPolg$^{Lar/Lar}$	329	28	192	109	82.25	164.5	82.25	0.0006
B6JPolg$^{Pro/Pro}$	245	57	120	68	61.25	122.5	61.25	0.8006

Table 3-2. Allele transmission and comparison to the expected Mendelian rate. In the table, we show the cross, the number of animals totals counted, the percentage of animals obtained of each genotype and the percentage of animals expected according to Mendelian inheritance. In the last column, the values of Chi-Test.

3.4. Lifespan comparison demonstrates allele hypothesis

One of the most remarkable characteristics of the mtDNA mutator mouse is the short lifespan. To date, this feature has been assessed in different research publications (Dillon *et al.*, 2012; Kujoth *et al.*, 2005; Mito *et al.*, 2013; Ross *et al.*, 2013; Someya *et al.*, 2017; Trifunovic *et al.*, 2004).
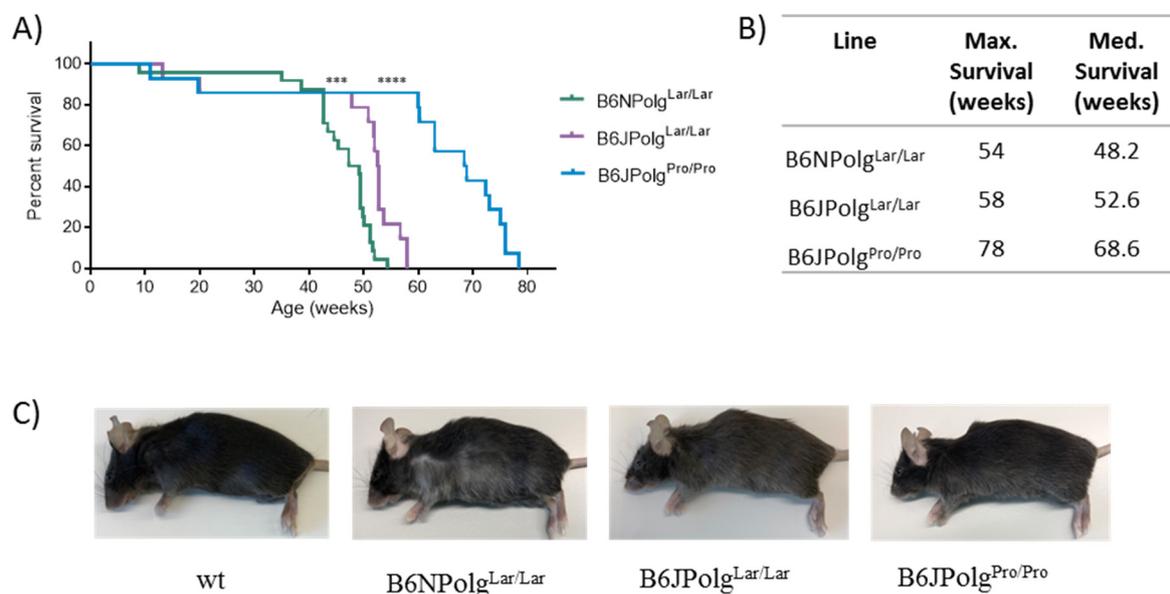
Figure 3-5. Lifespan and Phenotype. A) Lifespan obtained using breeding scheme lacking maternally transmitted mtDNA mutation. Log-rank (Mantel-Cox) test, ****p<0.0001, ***p<0.001. B) Maximum and medium survival data, obtained from the lifespan curve, end points were determinate when mice showed strong signs of sickness. C) Representative photos of the mtDNA mutator mice at 40 week old.

In this regard, it has been shown that maternally transmitted mtDNA mutations can shorten the lifespan in the mtDNA mutator mouse or wild-type animals, even at low levels (Ross *et al.*, 2013, 2014). Therefore, using a breeding scheme lacking maternally transmitted mtDNA mutations, we analysed the lifespan in our three mtDNA mutator mouse lines (Figure 3-5A and 3-5B). The data showed that B6NPolg^Lar/Lar^ exhibits the shortest lifespan, followed by B6JPolg^Lar/Lar^, and with the longest lifespan B6JPolg^Pro/Pro^. In addition, the

median lifespan showed the same pattern. We found that *Pro* mutator has an approximately 38% increased lifespan compared to mutators with the *Lar* allele.

We could visualise the difference in progeria on the mutator mouse lines at 40 weeks old, when the ageing phenotype is evident (Figure 3-5C). As shown in the pictures, mice with the Lar allele exhibit stronger progeria characteristics like alopecia, greasy fur, kyphosis (curvature of the spine); being the B6NPolg$^{Lar/Lar}$ evidently, the more affected line.

3.5. Transcriptome analysis reveals a differential immune response between mutators with different alleles

Because all our data confirmed that the expressed allele is evidently the major factor that gives as a result phenotypical differences among the mutator mouse lines, we performed the next experiments using only animals with the same genetic background (C57BL/6J) but with the different Polg alleles, B6JPolg$^{Pro/Pro}$ and B6JPolg$^{Lar/Lar}$.

To explore whether other biological processes were regulated differentially between the *Lar* and *Pro* allele, we examined global patterns of gene expression by RNA-Seq, using total RNA from the hearts of mice at 35 weeks old. Taking into account only genes with adjusted p-value <0.05, we observed 1,169 differentially expressed genes in B6JPolg$^{Lar/Lar}$ samples compared to wild-type samples. Strikingly, we only observed 312 dysregulated genes in B6JPolg$^{Pro/Pro}$ samples. From these, B6JPolg$^{Pro/Pro}$ and B6JPolg$^{Lar/Lar}$ shared 235 dysregulated genes (Figure 3-6A and 3-6B). Supplementary figures S3, S4 and S5 contain the data of the 40 genes most dysregulated in each mutator mice strain.
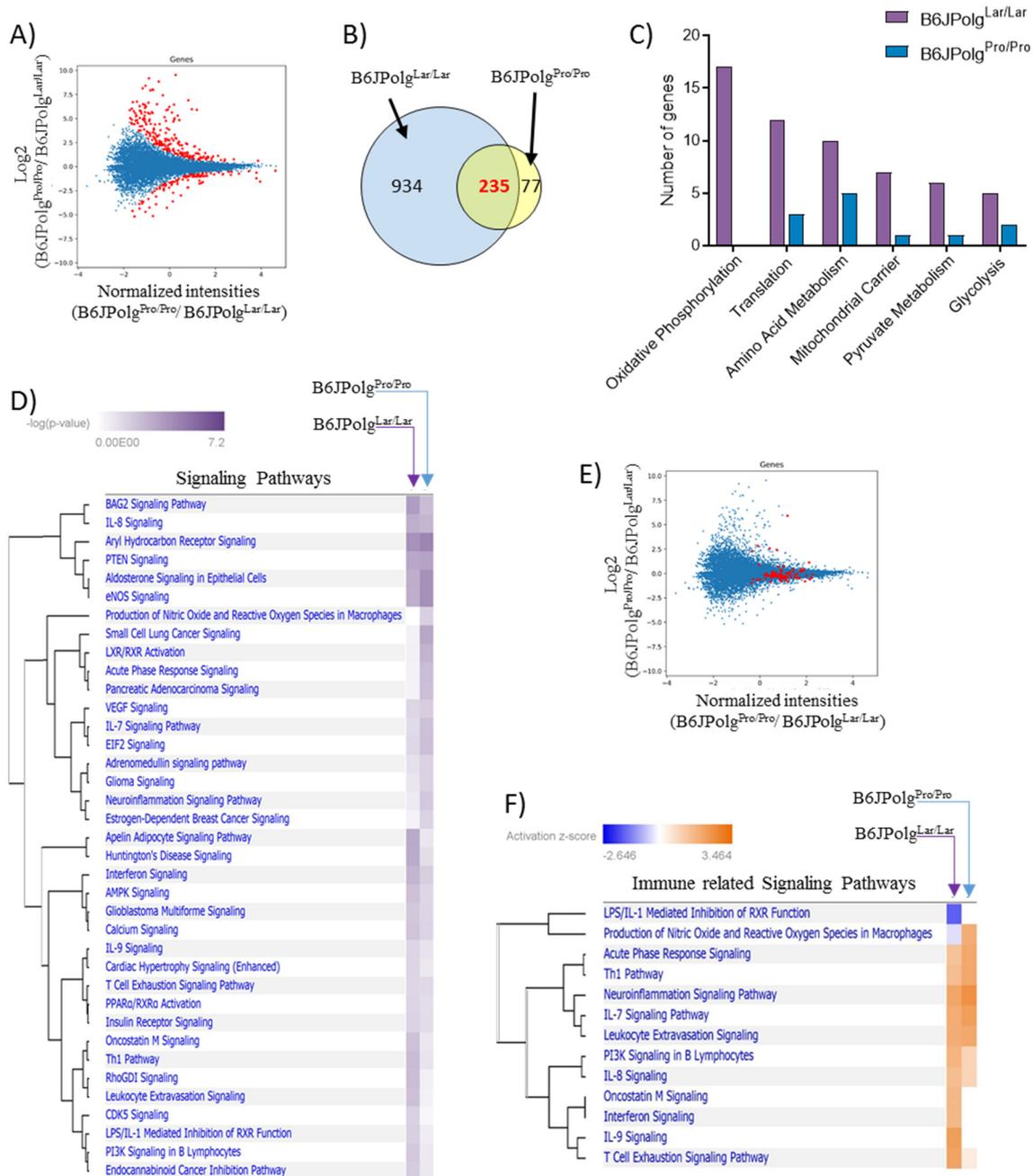
Figure 3-6. Heart-RNA-Seq Analysis. RNA-Seq data obtained from heart RNA of mice at 35 week old. Data using as a reference wild-type values and only genes with an adjusted p-value <0.05 (n=3). A) MA plot showing log2-fold expression ratio *vs* average log2-expression, identified in B6JPolg$^{Pro/Pro}$ compared to B6JPolg$^{Lar/Lar}$. Genes expressed differentially are highlighted in red. B) Venn diagram of a number of differentially expressed genes between B6JPolg$^{Pro/Pro}$ and B6JPolg$^{Lar/Lar}$. C) Genes identified with MitoXplorer, changing in B6JPolg$^{Pro/Pro}$ and B6JPolg$^{Lar/Lar}$. D) Hierarchical cluster dendrogram of signalling pathways between B6JPolg$^{Pro/Pro}$ and B6JPolg$^{Lar/Lar}$ (filter: only signalling pathways, p-value cut-off at 1.5 and z-score cut-off at 1.5). E) MD plot showing log2-fold expression changes *vs* average log2-expression identified in B6JPolg$^{Pro/Pro}$ compared to B6JPolg$^{Lar/Lar}$. Genes implicated in immune signalling pathways are highlighted in red. F) Hierarchical cluster dendrogram of activated or inhibited

pathways between B6JPolg$^{Pro/Pro}$ and B6JPolg$^{Lar/Lar}$ (filter: only immune-related signalling pathways, p-value cut-off at 1.5 and z-score cut-off at 1.5).

Next, we narrowed our focus on genes with established roles in mitochondria (Figure 3-6C). Using the web-platform mitoXplorer v.1.0 (http://mitoxplorer.ibdm.univ-mrs.fr/) (Yim *et al.*, 2019), we identified genes dysregulated in oxidative phosphorylation, translation, amino acid metabolism, mitochondrial carrier pathway, pyruvate metabolism, and glycolysis pathways. Overall, the data showed higher numbers of dysregulated mitochondrial genes in mice with the *Lar* allele compared with *Pro* allele mice (Figure 3-6C) (Complete list of genes in Table 3-3). While we did not detect dysregulated genes related to oxidative phosphorylation in B6JPolg$^{Pro/Pro}$ samples, we detected 17 genes down-regulated in B6JPolg$^{Lar/Lar}$ samples: 7 genes for complex I, 5 genes for complex V, 3 genes for complex IV and 2 genes for complex III.

Finally, in order to identify signalling pathways changing between the B6JPolg$^{Pro/Pro}$ and B6JPolg$^{Lar/Lar}$, we performed a functional enrichment analysis using the QIAGEN's Ingenuity Pathway Analysis software (IPA, https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/). The IPA enrichment algorithm uses two scores to determine different aspects: 1) P-value score, based on Fischer's exact test, where the *P*-value represents the significant overlap between the predicted and the regulated molecules. 2) Activation Z-score, which is a prediction for activation or inhibition of significant pathways (Krämer *et al.*, 2014).

When we analysed our data using the p-value score, we identified 37 biological pathways significantly dysregulated in B6JPolg$^{Pro/Pro}$ and B6JPolg$^{Lar/Lar}$ (Filter: only signalling pathways) (Figure 3-6D). Interestingly, a large number of differently expressed genes were related to immune response signalling pathways (Figure 3-6E). Therefore, we decided to analyse the activation and inactivation of the immune-related signalling pathways using Activation Z-score. Using as a filter for only immune-related signalling pathways, the data indicated that some immune pathways were activated in both mutant mice (Figure 3-6F). However, a differential response was found in pathways like interferon signalling, IL-9 signalling and T-cell exhaustion signalling, where activation was found in B6JPolg$^{Lar/Lar}$ mice but not in B6JPolg$^{Pro/Pro}$. A complete list of these genes is reported in Supplementary Figure S1.
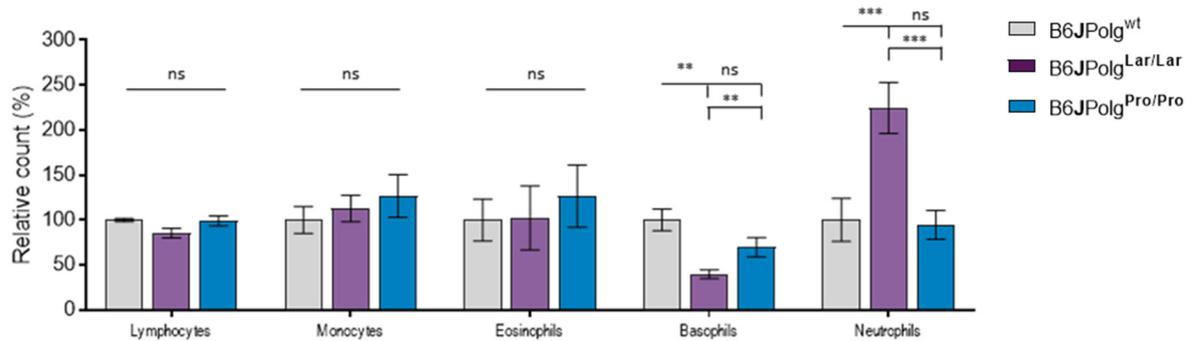
Figure 3-7. White Blood Cell counts on peripheral blood of mice. Data (n=4-8) are represented as mean ± SEM. Relative levels were normalised to the wt average. Two-way ANOVA $*p<0.05$; $**p<0.01$; $***p<0.001$.

The activation of immune-related pathways in the mutator mouse lines prompted us to investigate whether the immune system cells were being mobilised in peripheral blood. Therefore, we analysed the total count of white blood cells in the mouse lines. The results revealed that B6JPolg[Lar/Lar] samples showed a reduction of 60.3% in the number of basophils, while B6JPolg[Pro/Pro] mice only showed a 30.5% reduction. In addition, B6JPolg[Lar/Lar] mice showed a dramatic increase in the number of neutrophils (up to 224%), but the levels were not altered in mice with the B6JPolg[Pro/Pro] allele. We also performed Toluidine staining in heart, skeletal muscle and colon (Supplementary Figure S7) to identify if mast cells (the type of cells implicated in the activation of the immune response by the release of histamine and heparin) were being mobilised to these tissues. However, we did not observe any difference in comparison with wild-type samples.

Oxidative Phosphorilation		Translation		Aminoacid Metabolism		Mitochondrial Carrier		Piruvate Metabolism		Glycolysis	
L	Atp5e	L	Aurkaip1	L,P	Abat	L	Slc16a7	L	Acss1	L	Eno1
L	Atp5j2	L	Lars2	L	Aldh5a1	L	Slc25a20	L	Mpc1	L	Fbp2
L	Atp5k	L	Lrpprc	L,P	Bcat2	L,P	Slc25a22	L	Mpc2	L	Gapdh
L	Cox7a1	L,P	Mrm1	L	Bckdhb	L	Slc25a26	L	Pdk1	L,P	Gck
L	Fmc1	L	Mrpl11	L.P	Hmgcs2	L	Slc25a29	L,P	Pdk4	L	Pfkp
L	Higd1a	L	Mrpl41	L	Kyat3	L	Slc25a30	L	Pdpr	P	Pck1
L	Ndufa1	L	Mrpl42	L	Mpst	L	Ucp2				
L	Ndufaf4	L	Mrps21	L	Oxct1						
L	Ndufaf5	L	Mrps22	L	Ppm1k						
L	Ndufb3	L	Mrps33	L	Sqor						
L	Ndufc1	L	Rida	P	Ass1						
L	Ndufs8	L,P	Rnasel	P	Bckdha						
L	Ndufv2	P	Hemk1								
L	Sdhaf4										
L	Tmem70										
L	Uqcc2										
L	Uqcr10										

Table 3-3. Genes up- or down-regulated obtained from RNA-Seq of hearts at 35 week-old mice. *L,* denote genes dysregulated between B6JPolg[Lar/Lar] and wild-type samples; *P*, denote genes dysregulated between B6JPolg[Pro/Pro] and wild-type samples. *L,P,* denote genes dysregulated in both B6JPolg[Lar/Lar] and B6JPolg[Pro/Pro] in comparison to wild-type. Only genes with an adjusted p-value <0.05.

3.6. SNPs on the *POLG Pro* allele and possible consequences

Lar and *Pro* alleles were generated using two different gene-editing methodologies to generate the same POLG D257A protein mutation (i.e. plasmids, codon sequence, recombination systems); for this reason, they harbour differences at the nucleotide sequence level (See diagram 3-9). Briefly, in the *POLG Lar* allele (Trifunovic *et al.*, 2004), the codon GAC (aspartic acid) at the position 257 (exon 3), was exchanged for codon GCC (alanine). In addition, as a result of the methodology used, the final allele harbours two *loxP* sites flanking the exon 3, and one *Frt* site upstream exon 3. Whereas in the *POLG Pro* allele (Kujoth *et al.*, 2005), the position 257, the aspartic acid codon was exchanged for the codon GCT (alanine) and, as a result of the methodology used, the final allele contains a *loxP* site downstream of the exon 3 (Figure 3-9A). Therefore, to determine if these expected nucleotide sequence differences between the *Lar* and *Pro* allele were affecting the

phenotype, we first sought to confirm that no other mutations were present in these *POLG*^{D257A} alleles.

Unexpectedly, the alignment using RNA-seq data revealed that the *Pro* allele contains several single nucleotide polymorphisms (SNPs) in comparison to the GRCm38.p6 reference genome assembly from the substrain C57BL/6J (Genome reference of *Mus musculus*). As a result of these SNPs, the final protein should contain 4 non-synonymous substitutions (S27A, S308P, S665N, S1023A) and 3 additional synonymous substitutions (G74G, I228I, T525T), in addition to at least 8 nucleotide changes on non-coding regions (Figure 3-8). Meanwhile, the *Lar* allele was found to be identical to the current reference genome assembly GRCm38.p6.

From the SNPs that are in the coding sequences, of high relevance is the SNP that generate the substitution S308P, as is located within the exonuclease domain. The rest of the synonymous and non-synonymous substitutions are located outside of the exonuclease and the polymerase domain (according to Pfam arrangement Q05BB8, https://www.uniprot.org/uniprot/Q05BB8). In addition, SNPs, close to the 3' UTR region and changes at the exon-intron boundary downstream exon 4 may have functional relevance, as splicing can be disrupted.
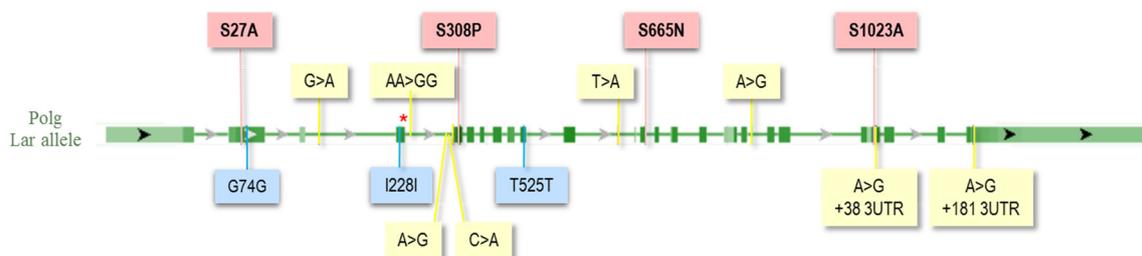
Figure 3-8. Scheme showing the *POLG Pro* allele and polymorphisms localised by RNA-seq-based mapping. In red are denoted non-synonymous substitutions, in blue synonymous substitutions, in yellow are the SNPs on non-coding regions. The D257A mutation is indicated by a red *.

During splicing, exons are removed from primary transcripts at conserved sequences called splicing sites. These splicing sites are critical consensus sequences located at the 5' and 3' ends of the introns (Reviewed in Sibley *et al.*, 2016). Therefore, one of the possible effects

of the nucleotide changes downstream of the exon 4 (AA>GG, A>G, and C>A) is the creation of a pseudo-splice site, which could lead to an altered splicing process (Anna and Monika, 2018). To explore if a splicing problem was occurring between the exon 3 and exon 4, we analysed the transcript levels of *POLG* in skeletal muscle and spleen of 35-week-old mice by RT-qPCR. Using a TaqMan probe spanning the junction exon 2 – exon 3 (figure 3-9A, grey arrows) and a probe spanning the junction exon3-exon4 (figure 3-9, green arrows), we could compared the proportion of mature RNA before and after the exon-intron boundary at risk. We added as a control, samples from hemizygous B6JPolg$^{Mut/ko}$ mice (harbouring a *POLG* knock-out allele and either the *Lar* or *Pro* allele); since they only contain one allele and have been shown to contain half the steady-state levels of *PolG* mRNA (Hance *et al.*, 2005).

The RT-qPCR data in skeletal muscle (Figure 3-9B) and spleen (Figure 3-9C) showed no significant difference in transcript levels when comparing mRNA spanned from exon 2-3 *versus* mRNA spanned from exon 3-4, meaning that with qPCR we could not detect any splicing alterations in the steady-state RNA pools of any of the mutator mice.

To gain more detailed insight about the mRNA processing between *Lar* and *Pro* allele, we decided to generate a compound heterozygous mutator mouse: B6JPolg$^{Lar/Pro}$. We obtained this new mouse line by crossing heterozygous mutants B6JPolg$^{Lar/wt}$ and B6JPolg$^{Pro/wt}$.

In order to detect different RNA species, we generated two pairs of primers: (i) a pair of primers to amplify the exon3-exon4 junction (Figure 3-9A, pink arrows), which will amplify the mature mRNA, and (ii) a pair of primers to amplify from the exon 3 to an intronic region, to detect pre-mRNA (Figure 3-9A, orange arrows). Taking advantage of the codon differences at the third codon of the D257A alleles, and the newly discovered variations between *Lar* and *Pro* allele, we were able to reveal the RNA relative levels produced from each allele.
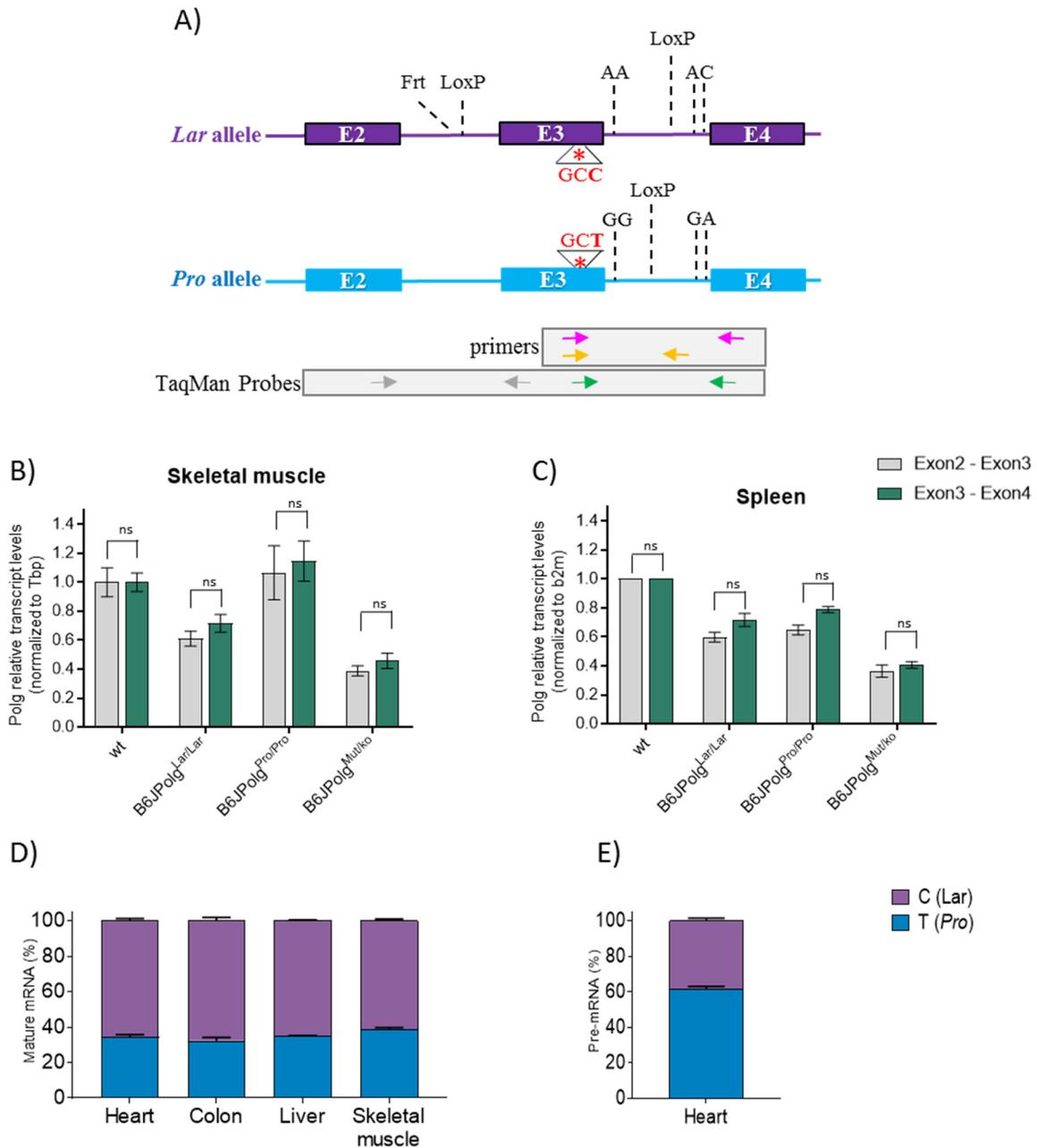
Figure 3-9. RNA processing analysis. A) Draw showing the *Lar* and the *Pro* allele from exon 2 to exon 4. Indicated by a red (*) is the D257A mutation and the codon used to generate the mutant alanine. In addition, dotted lines indicate the *loxP* and *Frt* sites and the location of the nucleotide changes close to the splicing sites. At the bottom are the locations of the primers and probes used to amplify specific fragments of RNA are shown. B) and C) are the *POLG* transcript levels determined by RT-qPCR on total RNA isolated from skeletal muscle and spleen, at 35 weeks old. Normalization to Tbp or B2m (n=2-5). Relative levels were normalised to the wt average. Data are represented as mean ± SEM. Two-way ANOVA *p<0.05; **p<0.01; ***p<0.001. E) Pyrosequencing analysis of the ratio of *Lar* (Purple) and *Pro* (blue) alleles in cDNA from compound heterozygous mice B6JPolg^{Lar/Pro} at 10 weeks old (n=4).

Using cDNA from heart, colon, liver and skeletal muscle of B6JPolg$^{Lar/Pro}$ mice at 10 weeks old, we performed allele-quantification Pyrosequencing. Theoretically, if both alleles were expressing equally, we could expect approximately 50% expression of each allele, either from the pre-mRNA or from the mRNA. We found that in all the analysed tissues, the proportion of mRNA expression of the *POLG* was ~40% *Pro* allele and ~60% *Lar* allele (Figure 3-9D). In contrast, when we analysed the pre-mRNA expression in heart, the proportions were inverted, with RNA derived from the *Pro* allele being more abundant (Figure 3-9E). These results show that there is more *Pro* pre-mRNA in the B6JPolg$^{Lar/Pro}$ mice at the steady-state level, in comparison with *Lar* pre-mRNA, and the proportion of *Pro* mature mRNA is lower than the *Lar* counterpart, implying altered splicing between exon 3 and 4.
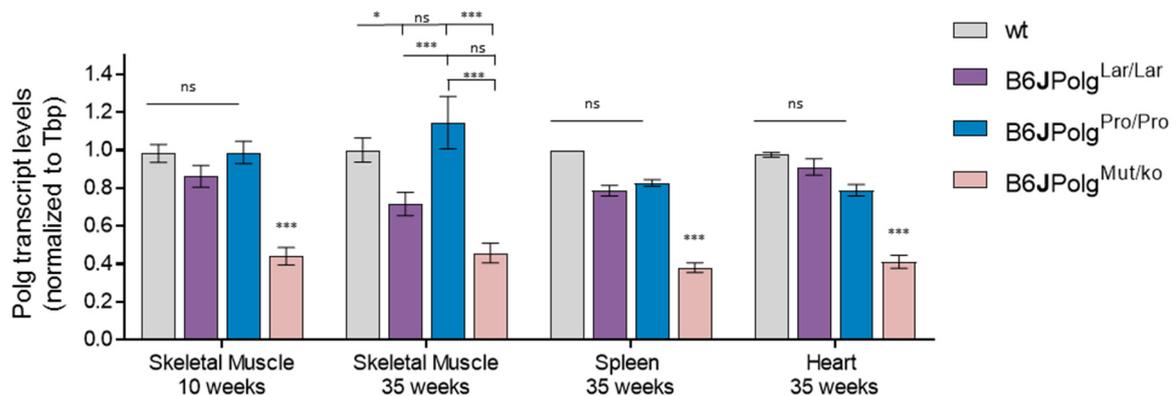
Figure 3-10. *POLG* transcript levels in different tissues. *POLG* transcript levels determined by RT-qPCR on total RNA isolated from mice at different ages (10 or 35 week old) and in different tissues (skeletal muscle, spleen and heart). Hemizygous B6JPolg$^{Mut/KO}$ mice were used as knockdown controls. Normalisation was to the *Tbp* mRNA (n=2-6). Relative levels normalised to wt average. Data are represented as mean ± SEM. Two-way ANOVA *p<0.05; **p<0.01; ***p<0.001.

Our next question was to analyse *POLG* transcript levels across different tissues. To address this question, we compare RT-qPCR data of skeletal muscle, spleen and heart of mice at 35 weeks old. Our data showed no statistical difference in steady-state mRNA levels in spleen and heart samples between mutators and with wild-type samples. However, in skeletal muscle data at 35 weeks old, we detect an increase in *POLG* transcript levels in B6JPolg$^{Pro/Pro}$ and a decrease in B6JPolg$^{Lar/Lar}$ (Figure 3-10). To determine if this variation was also visible in younger animals, we decide to analyse the same tissue in 10 week-old mice. The results indicate that at a young age, the levels were not affected (Figure 3-10).

3.7. POLG protein levels are not affected by the allele expressed

The results above prompted us to investigate the POLG protein levels in the mutator mouse lines. Therefore, we proceeded to assess by western blot the POLG expression. Because POLG is expressed at low levels in the cell, we used protein from mitochondria purified from the heart. In our analysis, we did not detect a difference in the POLG protein levels between genotypes (Figure 3-11A and 3-11B). As before, we used hemizygous B6JPolg$^{Mut/ko}$ animals as controls, to show that our assay was able to detect a 50% decrease in POLG levels. Taking together transcript level and protein level results, the data suggests that no alteration in POLG protein steady state levels could be resolved between the genotypes by western blot/antibody detection.
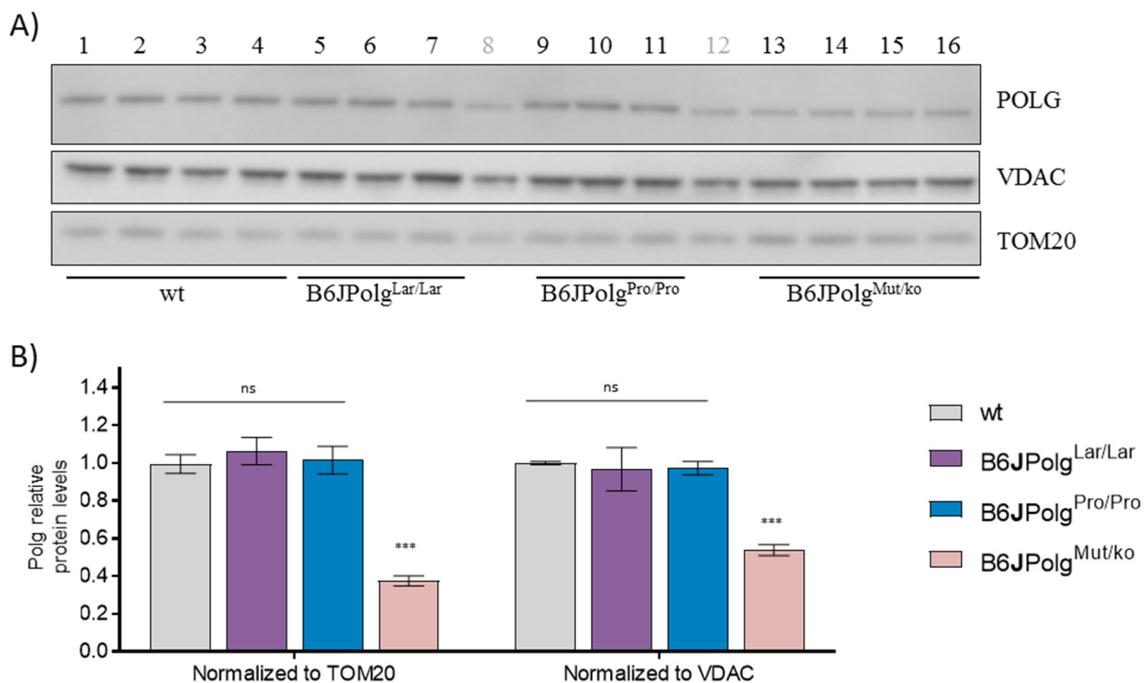
Figure 3-11. POLG protein levels. A) Western blot of POLG levels in mitochondrial extracts from the heart of mice at 35 week old (lines 1-7, 9-11, 13-16) and 10 week old (lines 8 and 12). B) Western blot quantification of POLG protein levels. Normalization to TOM20 and VDAC. Relative levels were normalised to the wt average. Data are represented as mean ± SEM. One-way ANOVA *p<0.05; **p<0.01; ***p<0.001.

3.8. Mutation rates and COX-deficiency differences between *Lar* and *Pro* allele

Once we had determined that *Lar* and the *Pro* alleles were genetically different, we decide to explore if the alleles were showing differences in mtDNA mutation load.

Previous reports have demonstrated that female mice with Polg$^{mut/wt}$ transmit mtDNA mutations through their oocytes, which can reach relative levels above 1% in the offspring (Ross *et al.*, 2013). Therefore, to avoid maternally transmitted mutations in the quantification of mutation load, we decide to generate hemizygous mtDNA mutator mouse lines with only one mutant allele and a knock-out *POLG* allele. By crossing B6JPolg$^{Pro/wt}$ and B6JPolg$^{Lar/wt}$ males with heterozygous knock-out B6JPolg$^{wt/ko}$ females, we generate two new lines of the mice. These new lines were containing the *Lar* and the *Pro* allele, but lacking germline transmitted mutations, here labelled these hemizygous as B6JPolg$^{Lar/ko}$ and B6JPolg$^{Pro/ko}$, respectively

Using high-quality enriched mtDNA samples as previously described by Isokallio *et al.,* (2018), we extracted mtDNA from the liver of young mice (11-14 weeks old), then we sequenced with an Illumina HiSeq3000 in 1 x 150bp read mode. Our results showed the same patterns of per-base coverage depth between the hemizygous mutant samples (Figure 3-12A).

We estimated the total mutation rate (transitions, transversions, and single-nucleotide indels) by dividing the number of mutations between the total of reads per sample. Our data (Figure 3-12B), indicate that both hemizygous mutants had an elevated frequency of mutations compared with the wild-type samples as previously reported (Ross *et al.*, 2013). Interestingly, B6JPolg$^{Pro/ko}$ samples exhibit a slightly lower mean mutation load of 7.5x10^{-4} mutations/bp in contrast to the 9.9x10^{-4} mutation/bp in B6JPolg$^{Lar/ko}$ samples.

One characteristic of the mtDNA mutator mice is that approximately 30% of the mtDNA molecules are short linear truncated pieces of mtDNA containing the major arc between the replication origins OriH and OriL (Figure 3-14) (Ameur *et al.*, 2011; Bailey *et al.*, 2009; Trifunovic *et al.*, 2004). We detected a decrease in the sequence coverage of the minor arc between the two origins of replications, which corresponds to the previously reported deleted region from the truncated linear double-stranded mtDNA molecules (Figure 3-14)

(Ameur *et al.*, 2011; Bailey *et al.*, 2009; Edgar and Trifunovic, 2009; Trifunovic *et al.*, 2004).
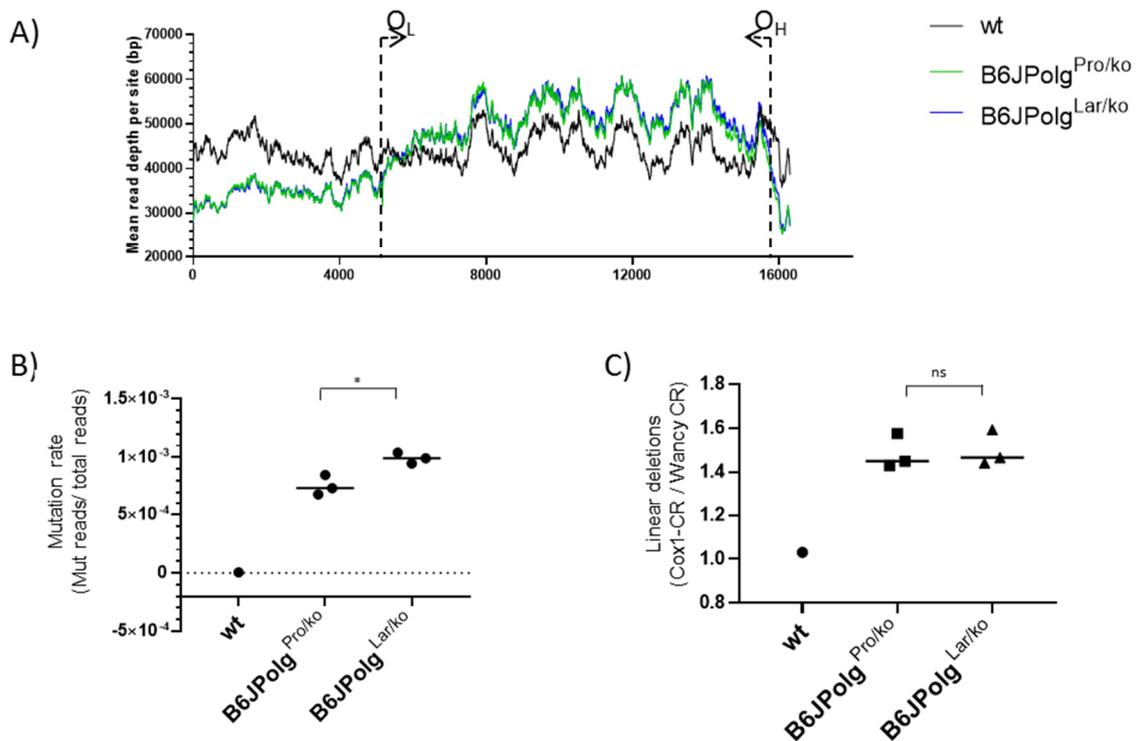
Figure 3-12. mtDNA mutation rates in *POLG* mut/ko mice. A) Sequence coverage across the mtDNA of sequenced samples with Illumina. The mitochondrial genome position is located at the x-axis and the sequence coverage divided by maximum coverage for each sample at the y-axis. Each line corresponds to the mean depth of 3 samples of mtDNA isolated from the liver of 10 week-old mice. The approximate locations of the origins of light-strand (O_L) and heavy-strand (O_H) are also shown. B) Mutation frequency was calculated by dividing the number of mutations (transitions, transversions and single-nucleotide indels) by the total of reads. One-way ANOVA *$p<0.05$. C) Linear deletions were calculated by dividing the number of mutations in the region CR-Cox1 by the Wancy t-RNA-cluster region. One-way ANOVA were ns= no significate.

To investigate if the relative amount of linear deletions differs due to the *POLG* allele expressed, we quantify these deletions dividing the number reads with the major arc region (which is present in linear deletion and wild-type mtDNA molecules) by the minor arc region (which is absent in the linear deleted molecules) (See figure 3-14). As expected, our results showed that hemizygous mutants contain a high level of deletions, which were absent in the wild-type samples. However, we did not find a detectable difference between

B6JPolg$^{Lar/ko}$ and B6JPolg$^{Lar/ko}$ samples, showing both samples an increase in the major arc sequencing coverage of 1.5x (Figure 3-12C). Taking together, our results showed a very mild increase in the mutation rate in mice with *Lar* allele, without alteration in relative levels of linear deletion molecules between hemizygous mutants.
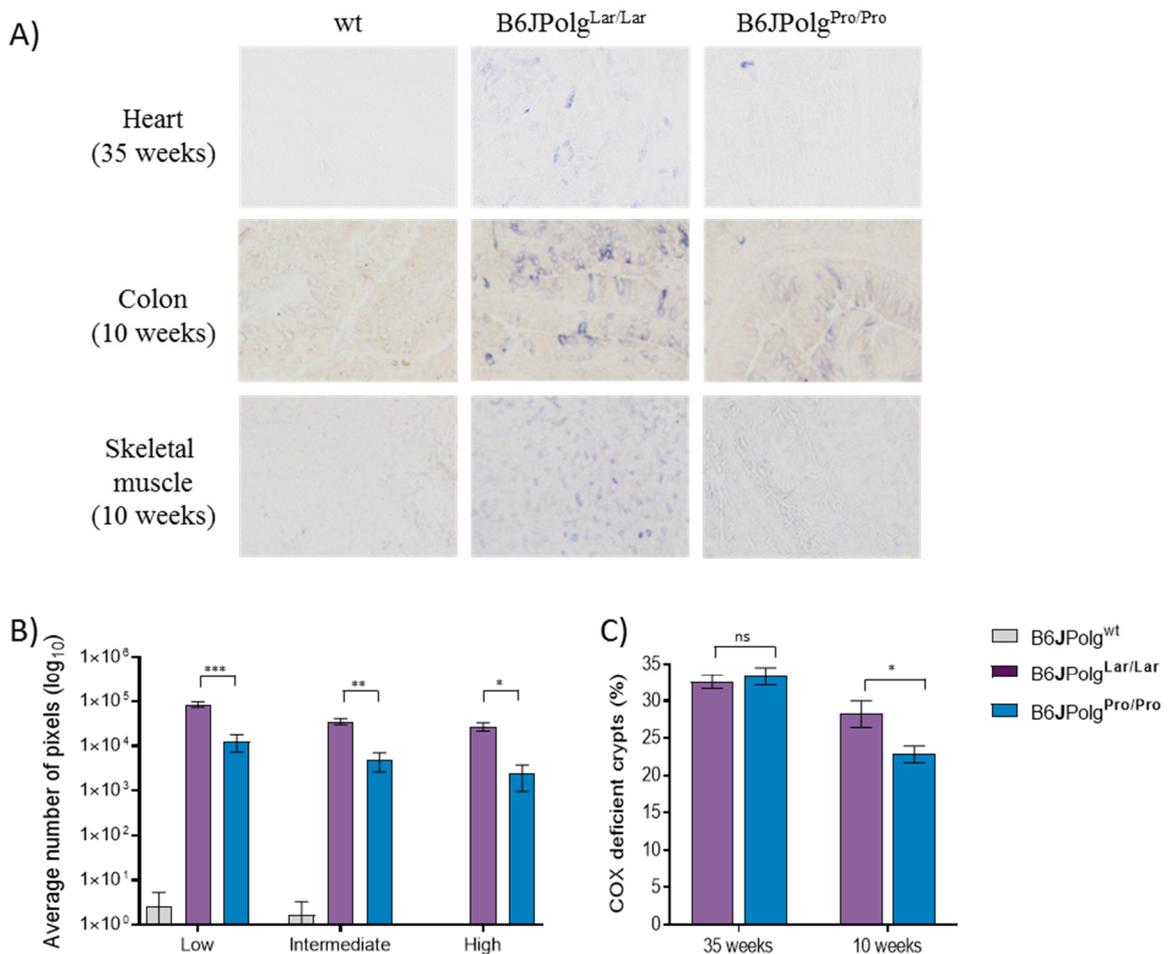
Figure 3-13. Mitochondrial COX deficiency. A) Representative pictures of NBTx staining in heart, colon and skeletal muscle at different ages. Blue colour indicates COX defective cells. B) Quantification by category (low, intermediate and high) of COX deficiency in heart tissues of 35 week-old mice using NBTx staining and computer quantification (n=4-6). C) Quantification of COX deficiency of colon tissues in 35 and 10 week-old mice, using COX/SDH staining and counting manually the number of crypts (n=5-7). In all graphs, data are represented as mean ± SEM. Two-way ANOVA *p<0.05; **p<0.01; ***p<0.001.

Next, we assessed mitochondrial function (Cytochrome c oxidase deficiency) in various postmitotic tissues of homozygous mutators. Cytochrome c oxidase (COX) is a partially mitochondrial encoded complex that can have an impaired function due to mutations (mitochondrial or nuclear) affecting directly COX subunits or mutations affecting the complex assembly (Rak *et al.,* 2016). Moreover, a decreased or lack of COX function can also be the result of mitochondrial protein synthesis impairment (Horvath *et al.,* 2009). However, a direct co-relation has been shown between mtDNA mutations and COX deficiency by the generation of respiratory chain complexes instability (Edgar *et al.*, 2009). Additionally, this mutations can be clonally expanded, creating a mosaic pattern of respiratory chain deficiency in the tissues (reviewed by Alston *et al.*, 2017).

Therefore, we performed in heart, skeletal muscle and colon, histochemical analyses to reveal cytochrome c oxidase (COX) activity variability between mutators and to know if (indirectly) there is a correlation with mitochondrial mutation loads. We decide to use the NBTx technique (Simard *et al.*, 2018) because it allows us to detect and quantify (in the heart) with high sensitivity any mitochondrial COX deficiency. This method is based on the assessment of blue colour in tissue samples, due to formazan crystals from reduced tetrazolium salts in cytrochrome c oxidase-deficient mitochondria. To perform the assay, two main components are added: 1) PMS (phenazine methosulphate), which is a strong electron carrier, 2) NBT (nitro blue tratrazolium), which can be reduced to form blue formazan crystals. In COX-competent mitochondria, no reduction occurs because electrons are passed to molecular oxygen via electron acceptors located in COX subunits; however, in COX-deficient mitochondria, electrons are passed to NBT, triggering the NBT reduction into the formation of blue formazan crystals. Visual inspection indicated that B6JPolg$^{Lar/Lar}$ has more COX deficiency in all tested tissues (heart, skeletal muscle (quadriceps) and colon), while B6JPolg$^{Pro/Pro}$ mice only showed a milder phenotype (Figure 3-13A). To confirm the visual data, we quantify with the software ImageJ, the images obtained from the heart (Figure 3-13B). The results show a dramatic increase in COX deficiency in B6JPolg$^{Lar/Lar}$ compared to B6JPolg$^{Pro/Pro.}$

In the colon, due to the tissue shape (lack of homogeneity and separation of the crypts), we choose to perform COX/SDH staining. Counting manually in the microscope the number of COX deficient crypts and the number of normal crypts, we obtained the percentage of mitochondrial dysfunction (Figure 3-13C). Our results clearly showed that there is a higher

number of mitochondrial COX dysfunction colonic crypts in mutators expressing the *Lar* allele.

3.9. Differences in Replication between *Lar* and *Pro* allele
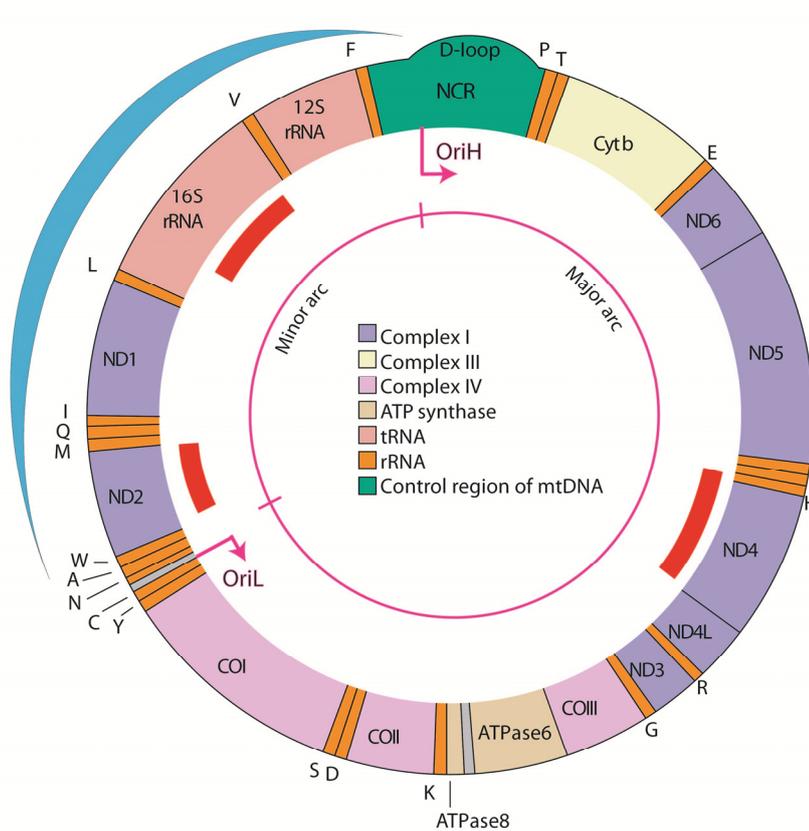
Figure 3-14. Schematic representation of the mtDNA. In red, the probes used to analyse mtDNA copy number, in green the control region and the D-loop, in blue the fragment absent in the linear mtDNA deletions, also of relevance are marked the light strand (OriL) and heavy strand (OriH) replication origins and the minor and major arc.

Finally, we decide to analyse if mtDNA replication could be affected by the *Lar* or *Pro* allele expressed. Therefore, we analysed mtDNA copy number by relative qPCR

quantification. This assay allows us to determine if a difference in mtDNA steady-state was occurring between the *Lar* and *Pro* mutator mice. Additionally, we used two different probes to confirm the previous result about no difference in the level of the mtDNA linear deletions.

We choose a Taqman probe that expands the ND4 gene, which is located in the major arc between the replication origins OriH and OriL. We also used Taqman probes for 16s or ND2, because these are genes from the fragment not present in the linear deleted mtDNA molecules (Figure 3-14). To normalise our data, we used a nuclear-encoded 18s probe. Using DNA samples extracted from liver and colon at 10 weeks old and 35 weeks old respectively, the results (Figure 3-15A) showed that in the 3 lines assessed (wild-type, B6JPolg$^{Lar/Lar}$ and B6JPolg$^{Pro/Pro}$) mtDNA copy number was not altered, either using the 16s or ND2 probe. Suggesting no alterations in the copy number of full-length mtDNA.

In agreement with our previous results (Figure 3-12A), the data using the ND4 probe confirmed the presence of the linear deleted mtDNA molecules in the mutator mice, and this is more evident in liver samples. However, we did not detect a difference between B6JPolg$^{Lar/Lar}$ and B6JPolg$^{Pro/Pro}$ in either liver or colon.
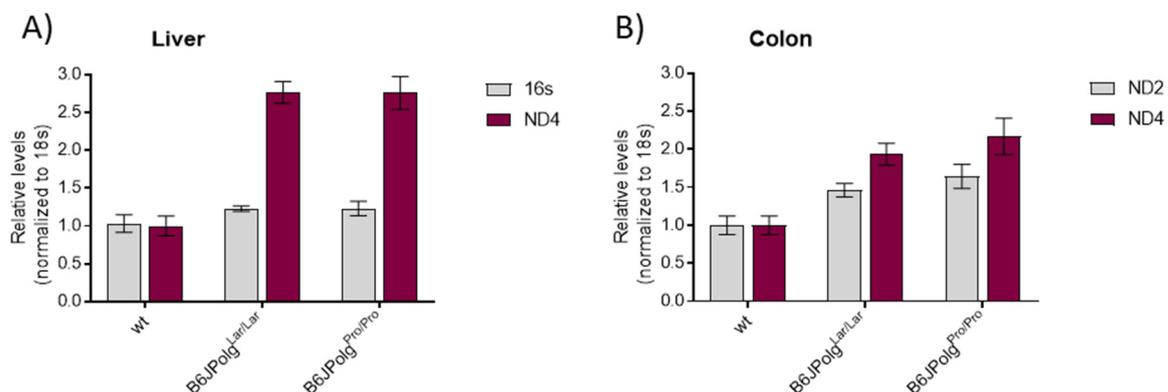
Figure 3-15. Mitochondrial DNA copy number. A) mtDNA copy number determined by qPCR on total liver DNA isolated from mice at 10 week old using 16s and ND4 probes. B) mtDNA copy number determined by qPCR on total colon DNA isolated from mice at 35 week old using ND2 and ND4 probes. Normalization to 18S (n=4-6).

Since several studies have established that TFAM (mitochondrial transcription factor A) protein levels strongly correlates with mtDNA copy number (Larsson *et. al.*, 1998, Ekstrand 2004, Kukat 2015), we carried out TFAM quantification by Western blot. Using mitochondria protein isolated from the heart and as an additional control samples from B6JPolg$^{Mut/ko}$ mice, the data indicated no variation in TFAM protein levels between wild-type, B6JPolg$^{Lar/Lar}$ and B6JPolg$^{Pro/Pro}$ samples (Figure 3-16A and 3-16B).

Taking together mtDNA copy number by qPCR and protein levels of TFAM by Western blot, results suggest no substantial variation in steady state mtDNA copy number in these post-mitotic tissues.
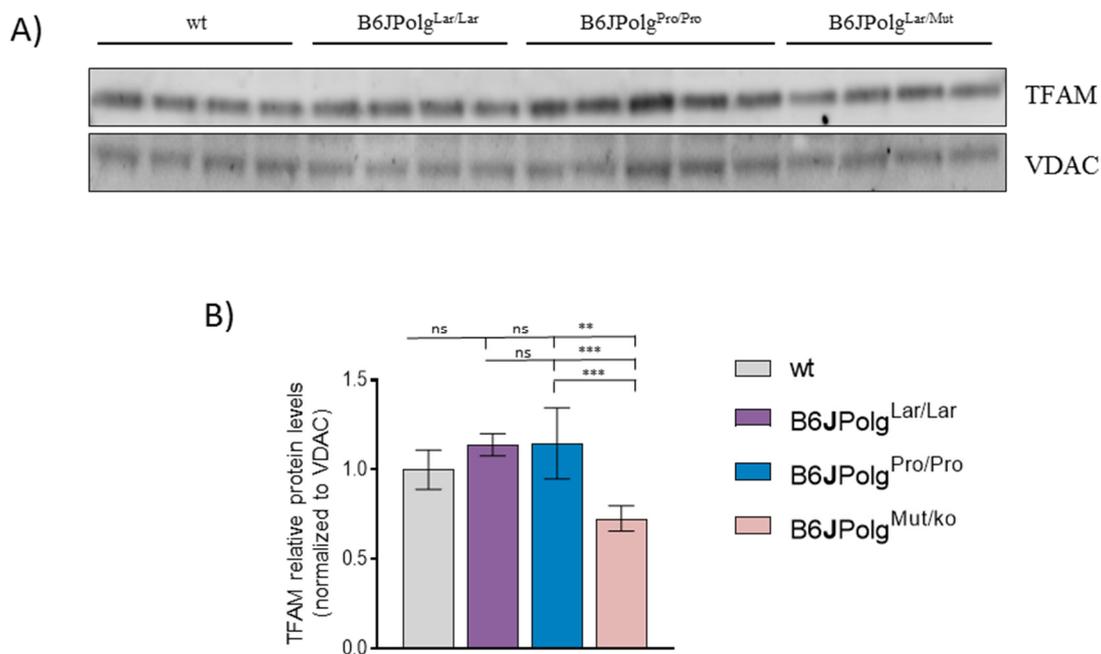
Figure 3-16. TFAM protein levels. A) Western blot on samples from isolated heart mitochondria of 10 week-old mice. Antibody for TFAM. Normalization to VDAC. Relative levels were calculated to the average normalised of wild-type. Data are represented as mean ± SEM. Two-way ANOVA ***p<0.001.

We next proceeded to study *ex vivo* function of the *POLG* coded by each allele. Therefore, we conducted mitochondrial DNA replication *in organello* assay, as previously described in Enríquez *et al.*, (1994) . Briefly, we incubate for 2 hours freshly isolated mitochondria with dNTPs were the dATPs is radioactively labelled with Phosphorus-32 (P32). As a result,

de novo synthesised mtDNA is radioactively labelled. Additionally, to follow the stability of the newly synthesised mtDNA, we performed a subsequent 1-hour chase. Samples were analysed in the native and denatured forms. At the end of the protocol, and once the signal of P32 was no longer detectable, the DNA levels were analysed by performing a Southern blot to detect total mtDNA (PAM1 probe).

First, we analysed *de novo* mtDNA synthesis in mouse samples at 10 weeks old. We detected an increase of newly synthesised mtDNA in B6JPolg$^{Lar/Lar}$ samples, and no difference was found between wild-type and B6JPolg$^{Pro/Pro}$ samples. We also detected that *de novo* mtDNA remains stable in all samples after 1 hour, as an abundant band is present in the native chase gel. The main results are shown in Figure 3-17A, and quantifications by ImageQuant software are shown in Figure 3-17A, 3-17B, 3-17C. It is important to mention that this result is different in samples of mice at 38 weeks old, where we saw a decrease in *de novo* mtDNA (Supplementary Figure S8).

However, intriguing results were observed for the 7S DNA fragment. Is has being described that most of the mtDNA replications events are abortive due to premature termination, generating 7S DNA fragments. The called 7S DNA is a linear DNA strand of approximately 650 nucleotides, which forms a characteristic triple-stranded DNA structure (the D-loop) (Reviewed by Nicholls and Minczuk, 2014)). In our results, the *de novo* canonical 7S DNA fragment is absent in the mutator samples, with a more prominent depletion in the B6JPolg$^{Lar/Lar}$ (Figure 3-17A and 3-17C).

Additionally, we observed that samples from the mutators generate a strong smear, which migrated slower than the 7S DNA band. When we quantify this area, we detected more isotope incorporation in B6JPolg$^{Lar/Lar}$ samples (Figure 3-17A, 3-17 D). Moreover, with the chase experiments, we found that these fragments are stable after one hour (observed in the denatured chase). Because of the size, we hypothesise that this smear is longer 7S DNA fragments, produced by the control region multimers (CRM) (Williams *et al.*, 2010). The CRMs are thought to be partial duplications of the D-loop region, generated by the gain of strand displacement activity in the exo- deficient POLG mutators (Macao *et al.*, 2015). To determine if the smears were longer 7S DNA fragments, and to determine the levels of these fragments at the steady-state level, we decide to performed Southern blot with a 7S probe. The results (Figure 3-18A and 3-18B) showed that the samples obtained from the mutator mice lack a defined band 7S. We also detect a smeared area with the 7S probe, implying

that these fragments are indeed longer 7S fragments due to the generation of control region multimers. Moreover, the quantification of the smeared area indicates that mutator samples with the *Lar* allele contain slightly more smeared 7S fragments, implying a more extensive generation of these aberrant products by the *Lar* POLG.
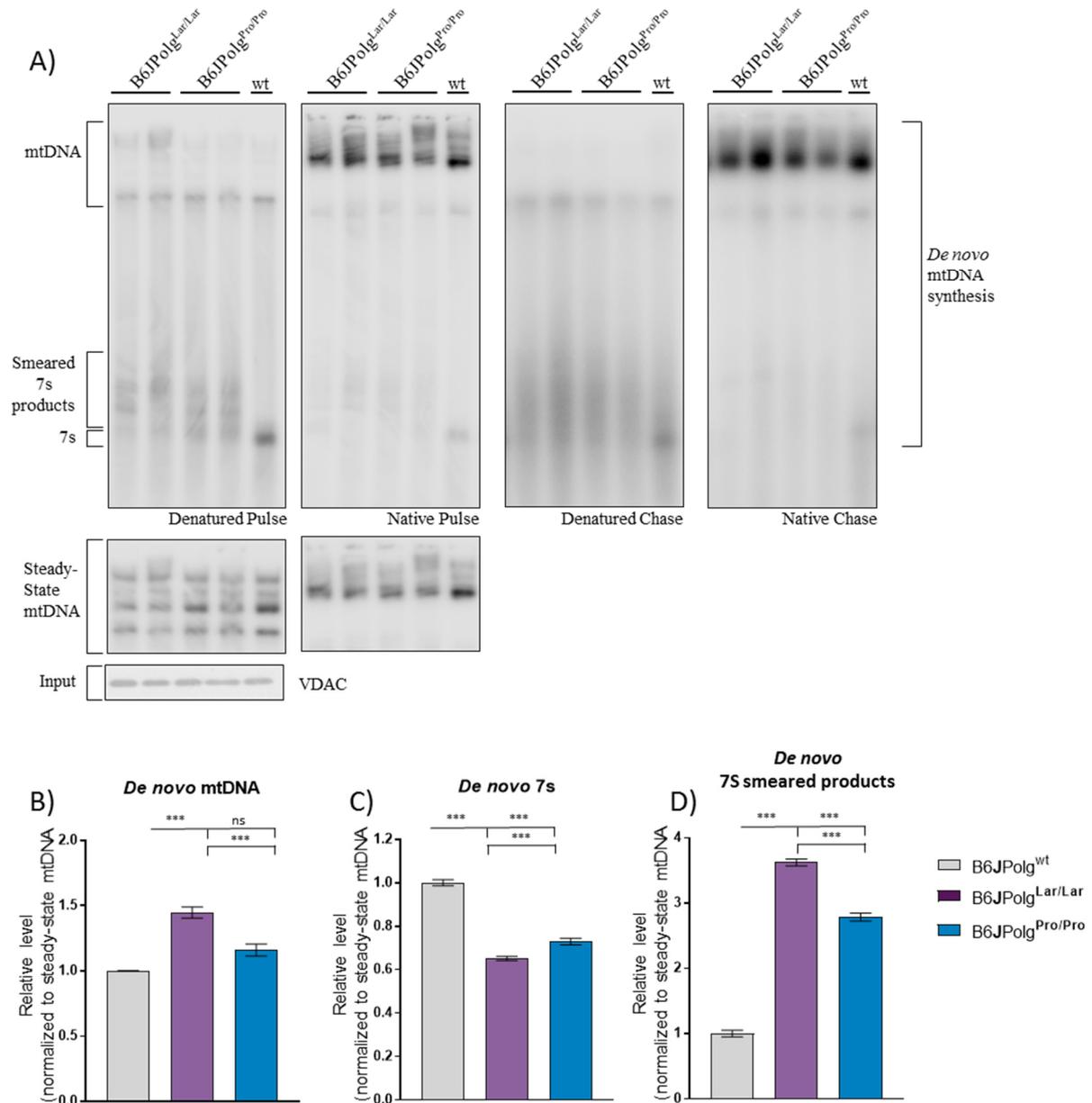
Figure 3-17. Mitochondrial replication analysis. A) *In organello* assay in mitochondria isolated from heart of 10 week-old mice. We show the samples from the denatured pulse and native pulse experiments, with the corresponding mtDNA steady-state levels using the probe PAM1, and protein input by Western blot with VDAC as a control. Denatured chase and native chase are also shown. Pulse was performed for 2 hrs and chase for 1 hour. B), C), D), Quantification of *de novo* mtDNA (B), *de novo* 7s (C), and *de novo* 7s smeared (D). Signals were normalised to the mtDNA steady state levels. In all graphs, data are represented as mean ± SEM. Two-way ANOVA *p<0.05; **p<0.01; ***p<0.001.

72

Taking together *in organello* assays and southern blot analysis, we conclude that there is variability in replication between the POLG generated by the *Lar* allele and the POLG generated by the *Pro* allele. Both alleles exhibit a lack of the canonical 7S DNA, and accumulation of multimers molecules derived from the partially duplicated control region.
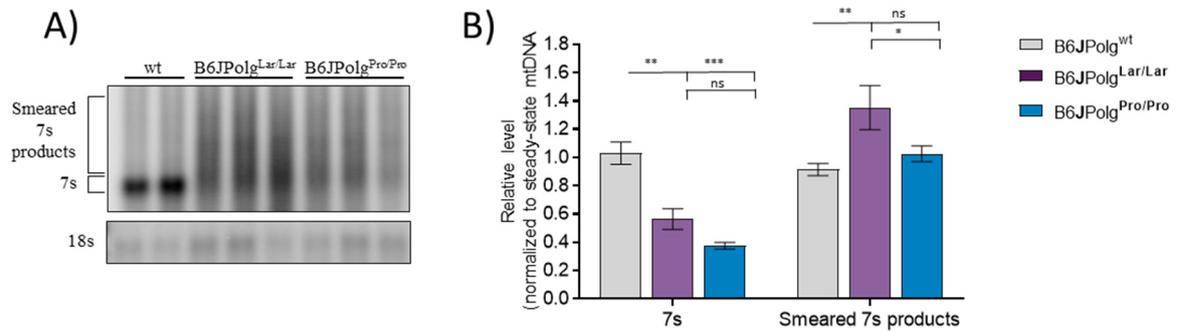
Figure 3-18. Southern blot analysis of SacI-digested mtDNA with 7S probe, and 18S probe. Quantification of steady-state levels of 7S and 7S smeared products. Signals were normalised to 18S. In all graphs data are represented as mean ± SEM. Two-way ANOVA *p<0.05; **p<0.01; ***p<0.001.

4. DISCUSSION

The mtDNA mutator mouse model contains a D257A mutation in the exonuclease domain of the POLG, which generates a mouse with an exonuclease deficient mitochondrial POLG protein. To date, three mutators with the same D257A POLG mutation have been described. These three mice display premature ageing; however, they show remarkable differences in their lifespan and the onset of the ageing phenotype. The first aim of this thesis was to determine, whether these variations were due to variations in the genetic background (a result of genetic differences between the substrains used), or the mutant POLG alleles expressed (which contain variations because of the different approaches to construct the mutants). The second aim of this thesis was to investigate the molecular mechanisms involved in the phenotypical variation.

Mutator mice reports from 2004 and 2005 (Trifunovic *et al.*, 2004 and Kujoth *et al.*, 2005) clearly showed differences in median lifespan and onset time of the ageing phenotype. However, it was not until 2013 that experiments by Ross and colleagues demonstrated that maternally transmitted mutations can directly affect the lifespan (Ross *et al.*, 2013). Therefore, it was highly relevant for us to compare the phenotype of the mutators under the same breeding scheme, lacking excess maternally transmitted mtDNA mutations that can occur by successive breeding of heterozygous $PolG^{D257A/WT}$ females. Our results confirmed the previously reported lifespan difference between the B6NPolg$^{Lar/Lar}$ (mouse published by Trifunovic *et al.*, [2004]) and B6JPolg$^{Pro/Pro}$ (mouse published by Kujoth *et al.*, [2005]) (Figure 3-5B).

To decipher the main reason for this phenotype variability among the mutators we analysed whether the genetic background lead to this variability. Previously, several reports have shown that the substrain C57BL/6J (B6J), used by Kujoth *et al.*, (2005); and the substrain C57BL/6NCrl (B6N), used by Trifunovic *et al.*, (2004) are different. These two substrains have been separated for almost 70 years, which has generated the accumulation of significant genetic variations (Kiselycznyk and Holmes, 2011). An earlier genomic analysis revealed that these two substrains contain at least 34 single nucleotide polymorphisms, 2 indels and 15 structural variants within the coding regions of the mouse genome (Simon *et*

al., 2013). As a result, the substrains exhibit several phenotypical differences in metabolism, behaviour, and susceptibility to diseases (Eisfeld *et al.*, 2019; Matsuo *et al.*, 2010; Simon *et al.*, 2013; Stiedl *et al.*, 1999). For instance, B6N carries a mutation in the Crumbs homolog 1 gene (Crb1) (Mattapallil *et al.*, 2012) and B6J carries a mutation in the NAD(P) transhydrogenase gene (NNT) (Toye *et al.*, 2005). Crb1 codes for a protein highly expressed in the eye and central nervous system of mice (Den Hollander *et al.*, 2002). Loss of Crb1 has shown to lead to retinal degeneration, and in humans is linked to a progressive degenerative disease that causes visual loss (Lotery *et al.*, 2001). In addition, evidence shows that the severity of the symptoms is dependent on additional genetic factors like variations in the genetic background (Luhmann *et al.*, 2015). NNT is a mitochondrial protein that catalyses the reduction of $NADP^+$ by oxidising NADH and coupling proton translocation across the inner membrane to the matrix. Mice with this mutation display a differential response to insulin secretion and glucose tolerance (Freeman *et al.*, 2006; Toye *et al.*, 2005). Therefore, these mutations could be generating additional consequences at the phenotypical level as well as at the behavioural level.

Our results demonstrate that mutators with the same *Lar* allele on different genetic backgrounds (B6NPolg$^{Lar/Lar}$ and B6JPolg$^{Lar/Lar}$) do not show significant differences. Therefore, I conclude that the genetic background is not the main factor for phenotype onset among the mutators. However, in this thesis, I demonstrate by comparing mouse phenotypes, that the main factor implicated in the time variation when the symptoms appear and lifespan differences, is due to allele variation in the *Polg* gene. Our results show that by comparing B6JPolg$^{Lar/Lar}$ and B6JPolg$^{Pro/Pro}$, which contain the same B6J genetic background, all the analysed phenotypes show a significant difference.

For instance, analysing cardiomyopathy, which is one of the phenotypes that changes less between mutators, the data not show a significant difference between B6JPolg$^{Lar/Lar}$ and B6NPolg$^{Lar/Lar}$. However, mice with the *Pro* allele (B6JPolg$^{Pro/Pro}$) show only a small increase in heart size. Data describing the anaemia phenotype shows similar results, with no significant difference between B6JPolg$^{Lar/Lar}$ and B6NPolg$^{Lar/Lar}$ mice that share the *Lar* allele and both showing a severe anaemia (50% erythrocyte reduction). On the other hand, in age-matched mice with the *Pro* allele, the number of erythrocytes shows only a mild reduction (22% reduction) and even more remarkable the spleen size is indistinguishable from wild-type samples, while substantially enlarged in the two lines bearing the *Lar* allele.

In other phenotypes, the allele differences are even more evident. For example, with male fertility, the testis structure and testis relative weight from mice with the *Pro* allele is indistinguishable from wild-type mice, whereas the mutator mice B6JPolg$^{Lar/Lar}$ and B6NPolg$^{Lar/Lar}$ show no difference between them and display substantial pathologies in comparison with wild-type samples. However, sperm analysis reveal variations between all of our mutator mouse strains. The results show that the B6JPolg$^{Pro/Pro}$ mouse has a better sperm quality than B6JPolg$^{Lar/Lar}$ and B6NPolg$^{Lar/Lar}$ mice, but large variations between B6JPolg$^{Lar/Lar}$ and B6NPolg$^{Lar/Lar}$ were detected in sperm quantity and quality (motility and morphology). While the nuclear genetic background may influence to some degree the ageing phenotype, for instance by variations in anaemia and sperm quality, the analyses presented here was unable to uncover any substantive difference.

In the end, lifespan results agreed with the allele hypothesis. Even though the data showed a small, non-significant increase in median survival time between mice with the *Lar* allele (B6JPolg$^{Lar/Lar}$ and B6NPolg$^{Lar/Lar}$), a ~1.4-times lifespan increase was found in mice with the *Pro* allele (B6JPolg$^{Pro/Pro}$). Maximum survival data also showed a 22 week increase in the lifespan of mice with the *Pro* allele, an increase of almost 40% in the maximum survival time. This result confirmed our hypothesis, that the main cause for the lifespan and onset of phenotype differences between the mutator mice is which of the transgenic alleles was expressed.

This conclusion leads us to analyse in detail other phenotypical differences, specifically between the *Lar* and *Pro* allele. For instance, a highly remarkable distinction in immune response was identified in heart samples by RNA-Seq. Although data showed an increase in immune response in B6JPolg$^{Lar/Lar}$ and B6JPolg$^{Pro/Pro}$ samples, they show a remarkable difference in the degree of activation. For instance, data analysis could predict activation in B6JPolg$^{Lar/Lar}$ in pathways like PI3K, IL-8, Oncostatin M (IL-6), IL-9 signalling, but not in B6JPolg$^{Pro/Pro}$. These signalling pathways have been shown to play important roles in differentiation, proliferation, activation of immune cells, and inflammatory pathways in general (Bickel, 1993; Goswami and Kaplan, 2011; Koyasu, 2003; Richards, 2013). In addition to the computer-predicted data, our experimental results also support the hypothesis that a remarkable immune response activation is occurring in mutators with the *Lar* allele. White blood cell count data show an increase in the number of neutrophils (neutrophilia) in

addition to a decrease in the number of basophils (basopenia) only in B6JPolg$^{Lar/Lar}$ mutators.

Neutrophils are a type of leukocyte highly abundant in the blood. Among the main functions of them are phagocytosis of pathogenic microorganisms, degranulation (secretion of antimicrobial proteins, enzymes, and pro-inflammatory substances), the release of the nuclear material to form microbial traps (NETs), and interaction with other immune cells (Mortaz *et al.*, 2018). Neutrophillia has been associated with immune responses to bacterial or fungal infections, inflammation, cancer, etc. It has been described that the increase in the number of circulating neutrophils is due to their release from the bone morrow and other marginated pools by inflammatory mediators like chemokines (Christoffersson and Phillipson, 2018). Our results show an increase of more than double in the number of neutrophils in mice with the *Lar* allele, supporting the idea of an increased immune response in B6JPolg$^{Lar/Lar}$.

Basophils represent less than 1% of the blood leukocytes; however, they play an important role in the development and progression of Th2-type immune responses and in chronic allergic inflammation. For instance, during basophil degranulation (activation) cytokines and chemokines are released (Sokol and Medzhitov, 2010). Basopenia has not often used in diagnosis, because under normal conditions basophil counts are very low in the blood. Causes of basopenia include acute stress, for example during infection and haemorrhage, anaphylaxis, acute urticaria and other acute allergic reactions (Bain, 2015). On the contrary, an increase of basophils has been associated with other immune responses, e.g. chronic inflammatory diseases like rheumatoid arthritis, ulcerative colitis, some leukaemias, helminthiasis, influenza, smallpox, and tuberculosis (Nathan and Orkin, 2009). Our results show that B6JPolg$^{Lar/Lar}$ mice exhibit low basophil counts in comparison with wild-type.

Therefore, here I discuss three ideas to explain basopenia detected in B6JPolg$^{Lar/Lar}$. First, previous literature has described a dysfunctional repopulation of haematopoietic stem cells in the mutator mice (Ahlqvist *et al.*, 2012; Chen *et al.*, 2009; Norddahl *et al.*, 2011; Wahlestedt *et al.*, 2014). Since we did not detect a decrease in other types of blood cells (except erythrocytes), I hypothesise that maybe basophils are a sensitive population to haematopoietic stem cell dysfunction. A second theory is that maybe basophils are releasing their granules in response to inflammation, becoming empty and are therefore not being detected in the blood test. Lastly, it could be possible that basophils are migrating to specific

tissues, and therefore are not visible in the circulatory blood. It has been reported that there is a substantial reduction in the number of basophils in the blood of individuals experiencing allergic reactions (Korosec *et al.*, 2017). However, when we did Toluidine staining (which stain basophils and mast cells) on colon, heart and skeletal muscle, we did not find any difference between wild-type and mutant tissues. However, I cannot discard the possibility that maybe the accumulation of basophils or mast cells is occurring in other tissues.

Overall, our RNA-Seq data suggest an increased immune response in the mutator mice. In both mice B6JPolg$^{Lar/Lar}$ and B6JPolg$^{Pro/Pro}$, RNA-Seq data predicts the activation of pathways like neuro-inflammation signalling, Th1 signalling, IL-7 signalling, leukocyte extravasation signalling, and acute phase response signalling. However, I was unable to elucidate the main factor triggering this response. Here I discuss some ideas. One hypothesis could be that the inflammation is triggered by the ageing process *per se*. Inflammaging is a term used to refer to a chronic low-grade inflammation that occurs in humans due to ageing. During ageing, tissues start to deteriorate. It has been suggested that the accumulation of damaged macromolecules and cells (self-debris) could trigger inflammation (Xia *et al.*, 2016). The second hypothesis is that, as I mentioned before, mtDNA is capable to activate neutrophils, which can indeed activate the immune response (Zhang *et al.*, 2010). In addition, it has been proposed that IL-26, a cytokine derived from Th17 cells, could bind to cytosolic mtDNA triggering an antimicrobial immune response (Meller *et al.*, 2015). Recently, it has been shown that during infection or mtDNA depletion, fragmented mtDNA can be released from the into the cytosol, triggering an antiviral immune response (STING-IRF3-dependent signalling by the expression of interferon-stimulated genes) (West *et al.*, 2015). Also, studies have shown that mtDNA can be released from neutrophils into circulation by the activation of formyl peptide receptor-1 and Toll-like receptor 9 (TLR9) (Zhang *et al.*, 2010). However, our RNA-seq data do not show any dysregulation on the genes reported by West *et al.* (*usp18, Ifi44*, *Ifit1*, *Ifit3*, *Oasl2*, *Rtp4,* etc*.)*. Because I do not know what triggered the putative immune response in the mutators, I cannot discard the idea that inflammation is directly implicated in the lifespan and onset of ageing phenotype difference between the mutator mice.

In this regard, there is a considerable amount of literature suggesting that mtDNA can be released from the mitochondria. For instance, neutrophils (Yousefi *et al.*, 2009) and basophils (Morshed *et al.*, 2014; Yousefi *et al.*, 2015) can release mtDNA from during the

generation of NETs (Neutrophil extracellular traps) and BETs (basophil extracellular traps). In addition, it has been found that aberrant mtDNA packaging (by TFAM deficiency) can release mtDNA to the cytosol. This release can activate DNAsensors, like the cGAS-STING pathway, which activates interferon signalling and in the end activating antiviral signalling (West *et al.*, 2015).

It is important to mention that our RNA-seq data show disagreement with recent data published by Hämäläinen *et al.* (Hämäläinen *et al.*, 2019). In their analysis, they found dysregulated pathways like TGF-β signalling, chemokine signalling, Wnt signalling, activation of the lysosome pathway, p53 signalling and RNA transport. Our data (Supplementary S2) using the same KEEG (Kyoto Encyclopedia of Genes and Genomes) database showed that the pathways linked to both mutator mice (that is B6JPolg$^{Lar/Lar}$ and B6JPolg$^{Pro/Pro}$) were cancer pathways, PI3K-Akt signalling pathway, focal adhesion, Rap1 signalling pathway and AMPK signalling pathway. Probably, this difference between Hämäläinen *et al.* and our data is related to the fact that they used iPSCs and not a tissue from an adult mouse where complex signalling occurs between the immune system and the tissues. Nevertheless, I found something interesting related to nuclear DNA damage. One of the genes with stronger downregulation in both mutator mice (B6JPolg$^{Lar/Lar}$ and B6JPolg$^{Pro/Pro}$) is ERCC2 (Supplementary S3 and S4). This gene codes for a protein involved in the transcription-coupled nucleotide excision repair (NER), as part of the DNA repair factor IIH (TFIIH). In NER, TFIIH opens the DNA around the lesion allowing the repair (Rimel and Taatjes, 2018). However, no other proteins part of the TFIIH were dysregulated in our data. It would be interesting to determine if a downregulation of ERCC2 is also occurring in tissues other than the heart, and if it is detectable at the protein level. Moreover, if the downregulation of ERCC2 has an effect on TFIIH activity.

Another phenotypical variation caused by the allele expressed was detected when we analysed mitochondrial COX deficiency. Previous findings in the literature have shown that the mutator mice display an impaired assembly and stability of the respiratory chain complexes. In general, impaired stability has been demonstrated in complexes I and IV, and reported in complexes III and V (Ahlqvist *et al.*, 2012; Edgar *et al.*, 2009; Hiona *et al.*, 2010; Woodall *et al.*, 2019). Our RNA-Seq data are in agreement with these previous results; on our case, we detect downregulation of genes for complex I, IV and V. Interestingly, such dysregulation was only found in mutators with the *Lar* allele in this study. These data

correlate with our staining results, where we identified in heart, colon and skeletal muscle a higher COX deficiency in mice samples with the *Lar* allele. In this regard, point mutations have been demonstrated to cause respiratory chain deficiency in the mutator mouse (Edgar *et al.*, 2009). As suspected, a slightly higher rate of mutation is detected in mouse samples with the *Lar* allele. All these results confirm a strong variability between the *Lar* and the *Pro* allele.

Contrary to our expectations, using RNAseq data to investigate the genomic sequence of the *Lar* and the *Pro* alleles, the results show that the *Pro* allele contains unexpected polymorphisms. Because of these, the POLG protein variant from the *Pro* allele exhibits four amino acid substitutions, compared to the Polg reference from the B6J substrain. However, these polymorphisms are found in other divergent substrains like 129S, C3H or CBA.It is hard to predict what affects these amino acid variants would have on POLG activity. However, I suspect that the polymorphisms with a major impact in the variability of the phenotype may be the S308P (in the exonuclease domain), S665N (linker region) and S1023A (polymerase domain), since they are located at important functional domains of POLG. To date, almost 200 mutations have been described in the POLG associated with different diseases; however, very little is known about the effect of them. Moreover, when these mutations were tested *in vitro* or in yeast model systems, the results were not always as expected (Stumpf *et al.*, 2013). For instance, the G517V substitution present in several syndromes like Leigh, neuropathy, myopathy and Kearns Sayre (Wong *et al.*, 2008), does not show any decrease of catalytic functions in studies *in vitro* (Kasiviswanathan and Copeland, 2011).

Our phenotypical characterisation of the mutator mice showed that the mice with the *Pro* allele exhibit a milder ageing phenotype. Therefore, one possibility is that one or more of the amino acid variations of the Polg *Pro* variant could somehow diminish the effect of the D257A mutation. Although this seems to be a radical idea, a report by Chan *et al.* shows it is possible. In ataxia patients, the mutation W748S (in the exonuclease domain) often correlates with the E1143G substitution (in the polymerase domain) (Hakonen *et al.*, 2005; Winterthun *et al.*, 2005). Unexpectedly, evidence *in vitro* shows that the E1143G substitution reduces partially the effect of the W348S mutation (Chan *et al.*, 2006). In the same way, the polymorphism S1023A, localised in the polymerase domain, could be reducing the impact of the deficient exonuclease activity in the *Pro* mutator mouse. Another

potentially impactful non-synonymous polymorphism is S665N, which is localised in the liker region. It has been described that the linker region interacts with one of the accessory subunits POLB (Lee *et al.*, 2009) and that the binding of the POLB could suppress the exonuclease activity (Graziewicz *et al.*, 2006). Therefore, I speculate that the S665N might also be modifying the effect of the deficient exonuclease activity in the B6JPolg$^{Pro/Pro}$ mutator mouse.

We also found important SNPs in non-coding sequences of *PolG*. First, there are sequence variations close to the exon 3, which could generate splicing variations; however, splicing changes are difficult to predict, particularly if new splicing sites are created (Adamson *et al.*, 2018). Second, there are SNPs close to the 3' UTRs (Figure 3-8) which are known to regulate mRNA stability, translation, and mRNA localisation (Mayr, 2017). We analyzed if sequence modifications of the 3'-UTRs could have an effect on the total mRNA's expression, but our POLG expression data by qPCR do not show any changes between B6JPolg$^{Pro/Pro}$ and WT mice in any tissue. However, using Pyro sequencing we did find mRNA POLG splicing abnormalities between the exon 3 and exon 4 using compound heterozygous mutator mice (Figure 3-9D and 3-9E). In our results, we detect that tissues were expressing more mature mRNA from the *Lar* allele, but more pre-mRNA from the *Pro* allele. I suspect that nucleotide changes downstream of the exon 4 (AA>GG, A>G, and C>A) could lead to splicing malfunction.

At the protein level, there is not a significant difference in POLG expression between B6JPolg$^{Lar/Lar}$ and B6JPolg$^{Pro/Pro}$. I am are aware that if a small variation in the POLG expression exists, it is plausible that with the Western Blotting technique used in our study, we would not be able to detect it. Additionally, it is important to mention that the detection of the POLG expression in the cell is difficult, since the POLG concentration is very low in the cells.

Finally, our data demonstrate that the POLG expressed by the *Pro* allele shows replicative differences in comparison with the POLG expressed by the *Lar* allele. For instance, analysing mtDNA replication by *in organello assay* in heart samples of 10-week-old mice, we found an increase in *de novo* mtDNA replication in mutators with the *Lar* allele, while no significant increase (in comparison with wild-type) was observed in mutators expressing the *Pro* POLG. According to our Pyro-seq data, we found a differential splicing between alleles in compound heterozygous mice. Therefore, I suggest that if the *Lar* allele is being

more expressed than the *Pro* allele this could lead to major differences in the synthesis of mtDNA.

Our data also show a difference in the *de novo* 7s DNA products, and 7s-like products derived from mtDNA bearing control region multimers. These results clearly show that there is variation in the Polg activity expressed by the respective alleles. Our results show that although both mutant alleles fail to generate the canonical sized 7s DNA fragments, mutants with the *Lar* allele show a lower amount of the canonical 7s fragments than mutants with the *Pro* allele (Figure 3-17C). Moreover, we visualised that mice with the *Lar* allele showed more of a smear of larger-sized DNA fragments that also binds the 7s probe (Figure 3-17-D). I hypothesise that these fragments are products from the control region multimers previously described by Williams *et al.*, (2010). It has been reported that mutator mice display control region multimers, probably created by defective mtDNA replication. It is suggested that when the polymerase, with a defective exonuclease activity, fails to halt at 5' ends and continues the mtDNA polymerisation, it creates a 5'-flap inefficient for ligation. Then, unligated ends are more likely to mis-pair within the control region, leading to small duplications (Macao *et al.*, 2015). These duplicated control regions could produce mis-sized 7s-like DNA products that we detect as smear in the *in organelle* assay. The difference in the multimer control region products, I suspect, is due to one or more of the substitutions generated by the SNPs that are modifying the activity of the Polg and as a result, the generation of these products. However, we cannot discard the possibility that these *in organello* variations are due to different amounts of Polg expression, because of the described splicing variations.

5. CONCLUSIONS AND FUTURE PERSPECTIVES

Engineering mouse models to study mtDNA mutations is problematic. First, mitochondria are double membrane organelles, which makes difficult the access of chemicals and transfection methods. Second, the apparent lack or minimal recombination of mtDNA prevents the successful integration of foreign DNA. Third, the multi-copy and heteroplasmic characteristics make it difficult to analyse single mutations. Fourth, physical or chemical mutagenic agents not targeted specifically at the mitochondria will not only produce mtDNA mutations but also nuclear DNA mutations (Jang and Lim, 2018; Tuppen *et al.*, 2010). Therefore, the mtDNA mutator mice is a very important animal model. It has allowed the generation of mutations in the mtDNA of mice that would not have been possible any other way (Kauppila *et al.*, 2016). However, the main force driving the premature ageing phenotype in these mice, and the relevance as a model of normal human ageing are still under intense debate (Dogan and Trifunovic, 2011; Hämäläinen *et al.*, 2019; Kraytsberg *et al.*, 2009; Vermulst *et al.*, 2008). To date, the mitochondria mutator mouse has already helped to reveal the putative consequences of a deficient proof-reading exonuclease POLG in humans (Dogan and Trifunovic, 2011) and clarify the principles of mtDNA mutated maternal transmission (Stewart *et al.*, 2008).

This study has demonstrated that although the genetic background influence to some degree the ageing phenotype, the main cause of the phenotypical differences exhibited by the different mutator mice strains is due to the Polg allele expressed. Our work reveals SNPs in the *Pro* allele that were derived from a divergent mouse line. In addition, I characterised some molecular mechanism differences exhibited by the different polymerases exhibited by each allele, including differences in the mutation rate, replication, and mitochondrial dysfunction.

I am confident that our results have improved the knowledge of the mutator mice model and we consider it will be helpful to other researchers for future work with the mouse model. I recommend taking into consideration our work during research projects using the mutator

mouse model, to avoid misinterpretation of data due to the existence of these two fundamentally different mtDNA mutator mouse strains. I also think our research improves the knowledge about the effect of POLG mutations and we hope it can be useful during mutation prediction in human diseases.

I think future work needs to concentrate on characterising and confirming an immune response in the mutator mouse model and establishing the cause. I also think it is important to determine the cause of the increased immune response in the *Lar* mutator. Moreover, further investigations need to establish how this relates to humans, and if the mutator mouse can be used as an inflammatory mouse model. In this regard, I consider that a start would be to determine if mtDNA is been released to cytosol or circulation, and the actions of neutrophils and basophils in the mutator immune response.

Future studies should also examine what are the implications of each of the substitutions found in the *Pro* allele, as we cannot discard the idea that some of these polymorphisms are rescuing the ageing phenotype. Therefore, this would be interesting for pathologies associated with the POLG.

6. SUPPLEMENTARY FIGURES

Gene	Log2(fold_change) between B6JPolg$^{Lar/Lar}$ and B6JPolg$^{Pro/Pro}$	significant
Stat1,Stat4	-0.516078	no
Stat1,Stat4	-0.516078	no
Apoa2	5.90351	yes
Irs1	-0.124927	no
Bcl2	-0.192351	no
Fmo2	-0.805905	yes
Atf6	-0.357416	no
Akt3	0.434778	no
Atf3	0.154027	no
Ctnna3	-0.373574	no
Itgb2	-0.518557	no
Timp3	-0.443571	yes
Arhgap9	-0.856039	no
Itgb3	-0.998968	no
Myl7	1.11601	yes
Acox1	-0.439268	yes
Pik3cg	-0.645018	no
Syk	-0.169809	no
Aldh5a1	-0.355875	no
Nfil3	0.147808	no
Vcl	-0.473999	no
Itih4	2.835	yes
Mmp14	-0.19074	no
Clu	-0.611916	yes
Ptk2b	-0.617242	no
Angpt1	0.737157	yes
Ahsg	2.41444	yes
H2-Ab1	0.141043	no
Ezr	-0.330144	no
Notch3	-0.195359	no
H2-Aa	0.190041	no
C3	-0.802584	yes
Hbegf	0.0773722	no
Ndst1	-0.114669	no
Ppargc1b	-0.216397	no
Cpt1a	-0.506431	yes
Cd274	-0.199145	no
Traf6	-0.0267497	no
Plcb1	-0.170333	no
Myl9	0.447524	no
Itih2	2.31215	yes
Fnbp1	-0.106308	no
Rnd3	0.22063	no
Nr1h3	0.298304	no
Cat	-0.498907	yes
Notch2	-0.123028	no

Gene	Log2(fold_change) between B6JPolg$^{Lar/Lar}$ and B6JPolg$^{Pro/Pro}$	significant
Hmgcs2	-0.698248	yes
Gstm5	0.230778	no
Gstm6	0.896989	no
Gstm1	-0.806339	yes
Rap1gap	0.681322	yes
Abca1	-0.326538	no
Jak1	-0.0483821	no
Cpt2	-0.452569	yes
Pla2g5	0.00973307	no
Isg15	-0.0727097	no
Abcb1a	-0.304707	no
Kdr	-0.426287	yes
Actb	-0.207889	no
Flt1	-0.386181	no
Gng11	0.414093	yes
Gstk1	0.367357	no
Itpr1	-0.395955	no
Cxcl12	-0.147058	no
Mgst1	0.432916	no
Timp4	0.0881348	no
Gnb3	-0.604813	no
Ccnd2	-0.237554	no
Il4ra	-0.422694	no
Arhgap35	-0.309265	no
Ifitm3	-0.0963743	no
Hras	0.179285	no
Irf7	-0.271394	no
Ccnd1	-0.0092593	no
Acsl1	-0.437454	yes
Cdh5	-0.32834	no
Nfatc3	-0.0821938	no
Nfat5	-0.000884026	no
Irs2	-0.601181	no
Angpt2	-0.298222	no
Dlc1	-0.280534	no
Hp	2.56149	yes
Icam1	-0.273937	no
Cbl	-0.238238	no
Atm	-0.00405512	no
Xiap	0.0875045	no
Msn	-0.420286	yes
Med14	-0.138759	no

S 1. RNA-Seq data of dysregulated genes associated to immune response signalling. Data from B6JPolg$^{Lar/Lar}$ and B6JPolg$^{Pro/Pro}$. RNA samples were obtained from RNA-Seq of hearts at 35 week-old mice. Only genes with an adjusted p-value <0.05, taking as a reference wild-type samples. Also shows if there is significant variation in expression between B6JPolg$^{Lar/Lar}$ and B6JPolg$^{Pro/Pro}$ (p-value <0.05) and the respective fold-change between B6JPolg$^{Lar/Lar}$ and B6JPolg$^{Pro/Pro}$.

| KEGG ID and Name | B6JPolg$^{Lar/Lar}$ | | B6JPolg$^{Pro/Pro}$ | |
	Hits	Genes	Hits	Genes
mmu05200 **Pathways in cancer**	39	Fzd4, Cdk6, Ctnna3, Ccnd1, Cxcl12, Jak1, AC124712.1,Hif1a,Snapc1, Hif1a, AC124712.1,Hif1a, Snapc1, AC124712.1, Fgf12, Ret, Stat1,Stat4, Stat1, Stat4, Gm28177, Sos1, Igf1r, Adcy4, Pik3cg, Apc, Gnb3, Cdkn1a, Col4a4, Adcy9, Plcb1, Cbl, Dvl3, Lama5, Prkacb, Lamb1, Rbx1, Gsk3b, Epas1, Rasgrp3, Traf6, Pdgfrb, Csf1r, Ptch1, F2r, Ednra, Bcl2, Hsp90aa1, Gng11, Hras	18	Cdk6, Xiap, Ctnna3, Ccnd1, Cxcl12, Jak1, AC124712.1,Hif1a,Snapc1, Hif1a, AC124712.1,Hif1a, Snapc1, AC124712.1, Max, Lama5, Akt3, Fgf12, Igf1r, Rasgrp3, Epas1, Bcl2, Hsp90aa1, Cbl, Plcb1
mmu04151 **PI3K-Akt signaling pathway**	38	Cdk6, Col5a3, Ccnd1, Angpt2, Jak1, Osmr, Thbs1, Fgf12, Thbs4, Syk, Sos1, Insr, Igf1r, Pik3cg, Prkaa2, Flt1, Gnb3, Cdkn1a, Col4a4, Sgk1, Itgb3, Irs1, Il4ra, Comp, Lama5, Lamb1, Gsk3b, Ccnd2, Angpt1, Pdgfrb, Csf1r, Thbs2, F2r, Kdr, Bcl2, Hsp90aa1, Gng11, Hras	16	Cdk6, Pck1, Col5a3, Angpt2, Ccnd1, Jak1, Lama5, Akt3, Fgf12, Insr, Igf1r, Angpt1, Kdr, Agxt2,Gm21973,Prlr, Prlr, Gm21973, Agxt2, Bcl2, Hsp90aa1
mmu04510 **Focal adhesion**	27	Col5a3, Ccnd1, Thbs1, Thbs4, Sos1, Igf1r, Pik3cg, Actb, Vcl, Flt1, Col4a4, Dock1, Itgb3, Arhgap35, Comp, Lama5, Lamb1, Gsk3b, Flnb, Ccnd2, Pdgfrb, Myl7, Myl9, Thbs2, Kdr, Bcl2, Hras	11	Col5a3, Myl9, Myl7, Xiap, Ccnd1, Kdr, Bcl2, Lama5, Akt3, Igf1r, Flnb
mmu04015 **Rap1 signaling pathway**	23	Itgb3, Angpt2, Rap1gap, Thbs1, Fgf12, Itgb2, Insr, Igf1r, Pik3cg, Rasgrp3, Adcy4, Actb, Angpt1, Pdgfrb, Csf1r, F2r, Kdr, Flt1, Hras, Sipa1l2, Pfn2, Adcy9, Plcb1	10	Rasgrp3, Angpt1, Angpt2, Kdr, Sipa1l1, Fgf12, Akt3, Insr, Plcb1, Igf1r
mmu04152 **AMPK signaling pathway**	16	Irs1, Ccnd1, Adra1a, Scd1, Fbp2, Insr, Igf1r, Cpt1a, Scd4, Pik3cg, Irs2, Gm29427,Pfkfb2, Pfkfb2, C4bp-ps1, Gm29427, Prkaa2, Acaca, Pfkp, Acacb	9	Cpt1a, Scd4, Pck1, Gm29427,Pfkfb2, Pfkfb2, C4bp-ps1, Gm29427, Ccnd1, Akt3, Acacb, Insr, Igf1r
mmu04066 **HIF-1 signaling pathway**	15	Pdk1, Angpt2, AC124712.1,Hif1a,Snapc1, Hif1a, AC124712.1,Hif1a, Snapc1, AC124712.1, Nppa, Tfrc, Insr, Rbx1, Igf1r, Pik3cg, Eno1, Angpt1, Bcl2, Flt1, Cdkn1a, Gapdh	9	Angpt1, Angpt2, AC124712.1,Hif1a,Snapc1, Hif1a, AC124712.1,Hif1a, Snapc1, AC124712.1, Bcl2, Akt3, Nppa, Tfrc, Insr, Igf1r

S 2. KEGG pathways. Results from enriched RNA-seq analysis with the software DAVID, showing KEGG pathways shared by B6JPolg$^{Lar/Lar}$ and B6JPolg$^{Pro/Pro}$.

Ensembl_gene_id	gene	sample_1	sample_2	log2 (fold change)	p_value	significant
ENSMUSG00000030400	**Ercc2**	wt	Pro	-9.38509	0.00005	yes
ENSMUSG00000076315	**Ercc2**	wt	Pro	-9.38509	0.00005	yes
ENSMUSG00000094475	Gm11007	wt	Pro	-6.52118	0.00005	yes
ENSMUSG00000026163	**Sphkap**	wt	Pro	-4.42546	0.0003	yes
ENSMUSG00000037263	Aldh3b3	wt	Pro	-4.14503	0.0002	yes
ENSMUSG00000030046	Bmp10	wt	Pro	-3.99561	0.00005	yes
ENSMUSG00000078894	2210418O10Rik	wt	Pro	-3.52553	0.00015	yes
ENSMUSG00000091971	Hspa1a	wt	Pro	-3.31494	0.00005	yes
ENSMUSG00000090877	Hspa1b	wt	Pro	-2.86595	0.00005	yes
ENSMUSG00000027694	Gm8325	wt	Pro	-2.80179	0.0003	yes
ENSMUSG00000000730	Dnmt3l	wt	Pro	-2.41181	0.00135	yes
ENSMUSG00000068745	Mybphl	wt	Pro	-2.29873	0.00005	yes
ENSMUSG00000061086	Myl4	wt	Pro	-2.24657	0.00005	yes
ENSMUSG00000020469	Myl7	wt	Pro	-2.10886	0.00005	yes
ENSMUSG00000042045	Sln	wt	Pro	-1.98314	0.00005	yes
ENSMUSG00000054966	Lmntd1	wt	Pro	-1.96248	0.00005	yes
ENSMUSG00000057880	Abat	wt	Pro	-1.80074	0.00005	yes
ENSMUSG00000029372	Ppbp	wt	Pro	-1.77863	0.00005	yes
ENSMUSG00000097039	Gm27718,Gm27913,Gm27957,Pvt1	wt	Pro	-1.74866	0.00085	yes
ENSMUSG00000031594	Fgl1	wt	Pro	1.95595	0.00005	yes
ENSMUSG00000076441	Ass1	wt	Pro	2.04998	0.00005	yes
ENSMUSG00000030532	Hddc3	wt	Pro	2.0604	0.00005	yes
ENSMUSG00000003585	Sec14l2	wt	Pro	2.08188	0.0001	yes
ENSMUSG00000037254	Itih2	wt	Pro	2.08439	0.00005	yes
ENSMUSG00000027875	**Hmgcs2**	wt	Pro	2.19103	0.00005	yes
ENSMUSG00000073016	**Uprt**	wt	Pro	2.19736	0.00005	yes
ENSMUSG00000010651	Acaa1b	wt	Pro	2.45203	0.00005	yes
ENSMUSG00000021922	Itih4	wt	Pro	2.5358	0.00005	yes
ENSMUSG00000000416	Cttnbp2	wt	Pro	2.57875	0.00135	yes
ENSMUSG00000027513	Pck1	wt	Pro	2.58106	0.00005	yes
ENSMUSG00000025479	Cyp2e1	wt	Pro	2.6205	0.00005	yes
ENSMUSG00000020051	Pah	wt	Pro	2.89706	0.00005	yes
ENSMUSG00000023070	Rgn	wt	Pro	2.93628	0.00005	yes
ENSMUSG00000073424	Cyp4f15	wt	Pro	3.18386	0.0012	yes
ENSMUSG00000067235	H2-Q10	wt	Pro	4.27889	0.00005	yes
ENSMUSG00000039196	Orm1	wt	Pro	4.51785	0.00105	yes
ENSMUSG00000071177	Serpina1d	wt	Pro	5.27271	0.0008	yes
ENSMUSG00000005681	Apoa2	wt	Pro	5.57302	0.00005	yes
ENSMUSG00000071178	Serpina1b	wt	Pro	5.62135	0.00005	yes

S 3. RNA-Seq data comparing wild-type vs B6JPolg$^{Pro/Pro}$. Here are 20 genes with major and minor fold-change, comparing samples from heart of wild-type and B6JPolg$^{Pro/Pro}$ animals.

Ensembl_gene_id	gene	sample_1	sample_2	log2 (fold change)	p_value	significant
ENSMUSG00000030400	**Ercc2**	wt	Lar	-9.43003	0.00005	yes
ENSMUSG00000076315	**Ercc2**	wt	Lar	-9.43003	0.00005	yes
ENSMUSG00000078875	Gm14419	wt	Lar	-5.35112	0.00005	yes
ENSMUSG00000026163	Sphkap	wt	Lar	-5.25188	0.0005	yes
ENSMUSG00000091971	Hspa1a	wt	Lar	-4.05656	0.00005	yes
ENSMUSG00000078894	2210418O10Rik	wt	Lar	-3.94501	0.00005	yes
ENSMUSG00000037145	2210407C18Rik	wt	Lar	-3.63439	0.00005	yes
ENSMUSG00000067786	Nnat	wt	Lar	-3.542	0.0002	yes
ENSMUSG00000030046	Bmp10	wt	Lar	-3.47667	0.00005	yes
ENSMUSG00000090877	Hspa1b	wt	Lar	-3.43165	0.00005	yes
ENSMUSG00000042045	Sln	wt	Lar	-3.38066	0.00005	yes
ENSMUSG00000061086	Myl4	wt	Lar	-3.32889	0.00005	yes
ENSMUSG00000020469	Myl7	wt	Lar	-3.22487	0.00005	yes
ENSMUSG00000024806	Mlana	wt	Lar	-3.18349	0.0002	yes
ENSMUSG00000022523	Fgf12	wt	Lar	-2.75532	0.00005	yes
ENSMUSG00000097174	Gm4890	wt	Lar	-2.70718	0.00095	yes
ENSMUSG00000020676	Ccl11	wt	Lar	-2.5498	0.00005	yes
ENSMUSG00000034656	Cacna1a	wt	Lar	-2.50141	0.00005	yes
ENSMUSG00000076928	Trac	wt	Lar	-2.46066	0.0002	yes
ENSMUSG00000091754	Gm3636	wt	Lar	1.90514	0.00005	yes
ENSMUSG00000033713	Foxn3	wt	Lar	1.92796	0.00005	yes
ENSMUSG00000029752	Asns	wt	Lar	1.94073	0.00005	yes
ENSMUSG00000043671	Dpy19l3	wt	Lar	1.95512	0.00075	yes
ENSMUSG00000048120	Entpd1	wt	Lar	1.95788	0.00005	yes
ENSMUSG00000019577	Pdk4	wt	Lar	2.06959	0.00005	yes
ENSMUSG00000020911	Krt19	wt	Lar	2.07363	0.0002	yes
ENSMUSG00000113626	Gm7240	wt	Lar	2.08693	0.00035	yes
ENSMUSG00000027171	Prrg4	wt	Lar	2.26443	0.00005	yes
ENSMUSG00000073016	Uprt	wt	Lar	2.26658	0.00005	yes
ENSMUSG00000040270	Bach2	wt	Lar	2.28524	0.00005	yes
ENSMUSG00000108621	Gm37494	wt	Lar	2.30048	0.00005	yes
ENSMUSG00000040274	Cdk6	wt	Lar	2.31608	0.00005	yes
ENSMUSG00000050944	Efcab5	wt	Lar	2.54003	0.00005	yes
ENSMUSG00000099693	C130026I21Rik	wt	Lar	2.58907	0.001	yes
ENSMUSG00000052477	C130026I21Rik	wt	Lar	2.58907	0.001	yes
ENSMUSG00000054728	Phactr1	wt	Lar	2.61499	0.00005	yes
ENSMUSG00000045333	Zfp423	wt	Lar	2.64954	0.00005	yes
ENSMUSG00000027875	Hmgcs2	wt	Lar	2.88928	0.00005	yes
ENSMUSG00000006784	Ttc25	wt	Lar	3.09629	0.00005	yes

S 4. **RNA-Seq data comparing wild-type vs B6JPolg^Lar/Lar^.** Here are 20 genes with major and minor fold-change, comparing samples from heart of wild-type and B6JPolg^Lar/Lar^ animals.

Ensembl_gene_id	gene	sample_1	sample_2	log2 (fold_change)	p_value	significant
ENSMUSG00000032172	Olfm2	Lar	Pro	-3.91023	0.00005	yes
ENSMUSG00000054728	Phactr1	Lar	Pro	-3.00076	0.00005	yes
ENSMUSG00000003882	Il7r	Lar	Pro	-2.526	0.00065	yes
ENSMUSG00000006784	Ttc25	Lar	Pro	-2.28463	0.00005	yes
ENSMUSG00000050944	Efcab5	Lar	Pro	-2.26762	0.00125	yes
ENSMUSG00000108621	Gm37494	Lar	Pro	-2.18398	0.00005	yes
ENSMUSG00000020911	Krt19	Lar	Pro	-2.18119	0.0017	yes
ENSMUSG00000027694	Gm8325	Lar	Pro	-2.15679	0.00115	yes
ENSMUSG00000017897	Eya2	Lar	Pro	-1.99926	0.00055	yes
ENSMUSG00000044534	Ackr2	Lar	Pro	-1.85887	0.00115	yes
ENSMUSG00000040270	Bach2	Lar	Pro	-1.80562	0.00005	yes
ENSMUSG00000053093	Myh7	Lar	Pro	-1.75621	0.00005	yes
ENSMUSG00000042985	Upk3b	Lar	Pro	-1.73977	0.00005	yes
ENSMUSG00000033713	Foxn3	Lar	Pro	-1.69733	0.00005	yes
ENSMUSG00000060206	Zfp462	Lar	Pro	-1.61882	0.0004	yes
ENSMUSG00000068606	Gm4841	Lar	Pro	-1.52094	0.0011	yes
ENSMUSG00000063388	BC023105	Lar	Pro	-1.41701	0.00075	yes
ENSMUSG00000037661	Gpr160	Lar	Pro	-1.40325	0.00005	yes
ENSMUSG00000058297	Spock2	Lar	Pro	-1.27648	0.0002	yes
ENSMUSG00000031722	Hp	Lar	Pro	2.56149	0.00005	yes
ENSMUSG00000003585	Sec14l2	Lar	Pro	2.71119	0.00005	yes
ENSMUSG00000076441	Ass1	Lar	Pro	2.72568	0.00005	yes
ENSMUSG00000021922	Itih4	Lar	Pro	2.835	0.00005	yes
ENSMUSG00000046991	Wdr27	Lar	Pro	2.90424	0.0018	yes
ENSMUSG00000037145	2210407C18Rik	Lar	Pro	2.92637	0.00005	yes
ENSMUSG00000023070	Rgn	Lar	Pro	2.94033	0.00005	yes
ENSMUSG00000037686	Aspg	Lar	Pro	2.96896	0.00005	yes
ENSMUSG00000050545	Fam228b	Lar	Pro	2.97243	0.00005	yes
ENSMUSG00000024747	Aldh1a7	Lar	Pro	3.11998	0.0004	yes
ENSMUSG00000035875	AI182371	Lar	Pro	3.12005	0.00005	yes
ENSMUSG00000024039	Cbs	Lar	Pro	3.50781	0.0003	yes
ENSMUSG00000031594	Fgl1	Lar	Pro	3.85985	0.00005	yes
ENSMUSG00000030382	Slc27a5	Lar	Pro	3.92464	0.00005	yes
ENSMUSG00000067235	H2-Q10	Lar	Pro	4.00268	0.00005	yes
ENSMUSG00000005677	Nr1i3	Lar	Pro	4.38082	0.00045	yes
ENSMUSG00000078875	Gm14419	Lar	Pro	4.68799	0.00005	yes
ENSMUSG00000071177	Serpina1d	Lar	Pro	5.87702	0.00085	yes
ENSMUSG00000005681	Apoa2	Lar	Pro	5.90351	0.00005	yes
ENSMUSG00000071178	Serpina1b	Lar	Pro	6.45816	0.00005	yes

S 5. **RNA-Seq data comparing B6JPolg**[Lar/Lar] **and B6JPolg**[Pro/Pro]. Here are 20 genes with major and minor fold-change, comparing samples from heart of B6JPolg[Lar/Lar] and B6JPolg[Pro/Pro]

| Position	dbSNP	RNA-Seq data Pro Allele	RNA-Seq data Lar allele	Ref (GRCm38.p6)	SNPs in other strains	Consequences
79449299	rs254447053	G*	A	A	G*	3_prime_utr_variant nmd_transcript_variant downstream_gene_variant
79449424	rs37134955	T*	C	C	T*	3_prime_utr_variant nmd_transcript_variant downstream_gene_variant
79450709	rs32308962	-	-	T	C*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79451929	rs31834612	G*	A	A	G*	synonymous_variant 3_prime_utr_variant nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79452000	rs31713195	C*	A	A	C*	missense_variant 3_prime_utr_variant nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79453448	rs31806755	-	-	G	A*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79453971	rs31526362	-	-	G	A*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79454512	rs31213966	G*	A	A	G*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79455788	rs32518466	-	-	G	T*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79455892	rs32139899	-	-	A	G*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79455965	rs31429820	-	-	T	C*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79456727	rs31546017	T*	C	C	T*	missense_variant 3_prime_utr_variant nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79456913	rs31899011	-	-	A	G*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79457291	rs32191422	A*	T	T	A*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79457967	rs32327250	-	-	C	A*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant

79457974	rs31335373	-	-	C	T*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79458032	rs32473070	-	-	T	C*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79458034	rs31101233	-	-	G	A*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79459013	rs31334366	-	-	C	T*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79459026	rs32480059	-	-	C	T*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79459254	rs13465959	C*	T	T	C*	synonymous_variant 3_prime_utr_variant nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79459385	rs32158620	-	-	T	C*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79459652	rs31277724	-	-	A	C*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79460563	rs31419114	G*	A	A	G*	missense_variant nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79460717	rs32306699	T*	G	G	T*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79460773	rs31924510	-	-	C	T*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79461274	**rs31160733**	**C***	**T**	**T**	**C***	**nmd_transcript_variant upstream_gene_variant downstream_gene_variant**
79461669	**rs31991119**	**C***	**T**	**T**	**C***	**nmd_transcript_variant upstream_gene_variant downstream_gene_variant**
79461670	**rs32333834**	**C***	**T**	**T**	**C***	**nmd_transcript_variant upstream_gene_variant downstream_gene_variant**
79461818	rs32472336	T*	G	G	T*	synonymous_variant nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79462408	rs31617548	-	-	T	C*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79463291	rs265473164	-	-	G	C*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79463506	rs31943710	A*	G	G	A*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant

79463737	rs49532739	-	-	A	C*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79464290	rs3674989	-	-	A	C*	nmd_transcript_variant upstream_gene_variant downstream_gene_variant
79464990	rs31639173	T*	G	G	T*	synonymous_variant nmd_transcript_variant upstream_gene_variant
79465133	rs31192147	C*	A	A	C*	missense_variant nmd_transcript_variant upstream_gene_variant
79465635	rs31625700	A*	G	G	A*	nmd_transcript_variant upstream_gene_variant
79465778	rs31579544	C*	A	A	C*	nmd_transcript_variant upstream_gene_variant
79466081	rs31555652	C*	T	T	C*	5_prime_utr_variant nmd_transcript_variant upstream_gene_variant

S 6. Table with all the SNPs described in the POLG allele and data from RNA-Seq. The first column shows the position in the POLG gene. The second column shows the ID from the Single Nucleotide Polymorphism Database (dbSNP). The third and fourth column shows the data from the RNA-Seq data obtained from hearts of B6JPolg$^{Lar/Lar}$ and B6JPolg$^{Pro/Pro}$ mice at 35 weeks. The fifth column shows the nucleotide from the reference sequence GRCm38.p6. The sixth column shows the SNPs from the strains 129S, C3H, CBA, DBA, KK_HiJ and SEA_GnJ substrains. Data obtained from the Mouse Genomes Project - Gene: Polg; Chr: 7:79,448,791-79,466,362 (ftp://ftp-mouse.sanger.ac.uk/REL-1505-SNPs_Indels) from the Sanger Institute. The seventh column shows the consequences of these SNPs.
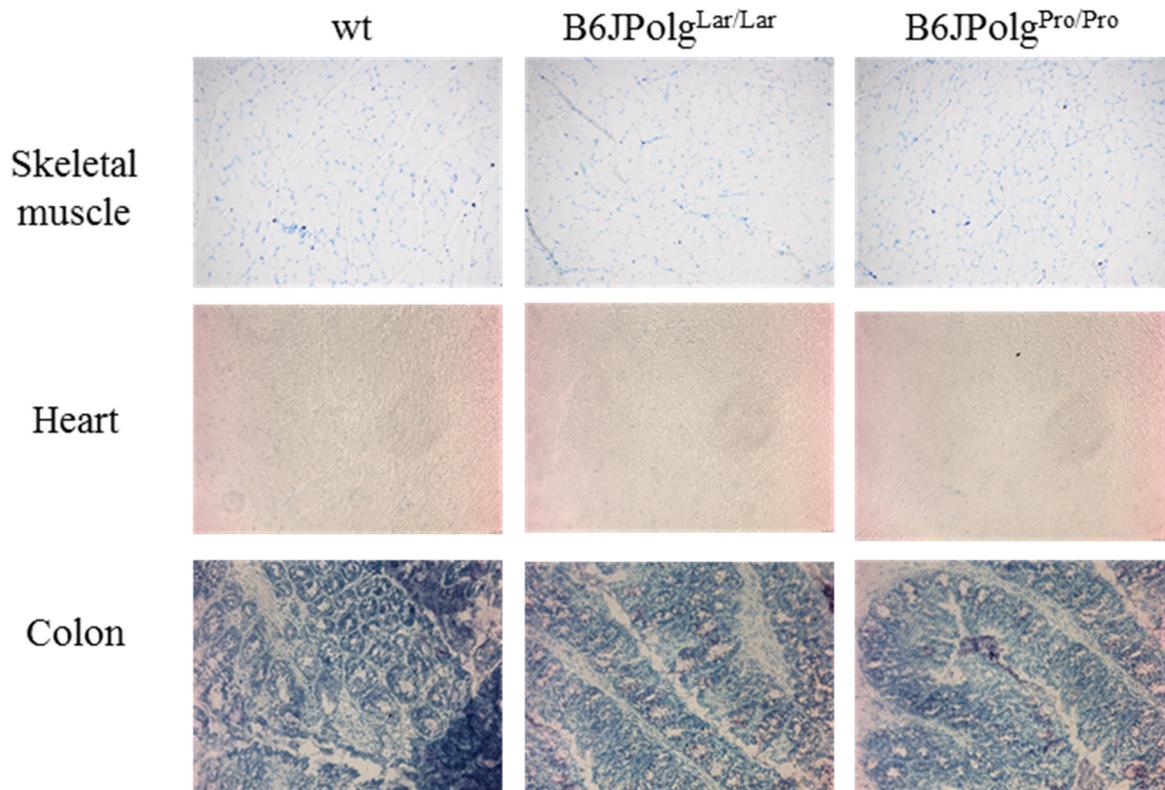
S 7. **Toludine staining.** Toluidine staining in heart, colon and skeletal muscle of mice at 35 weeks old.
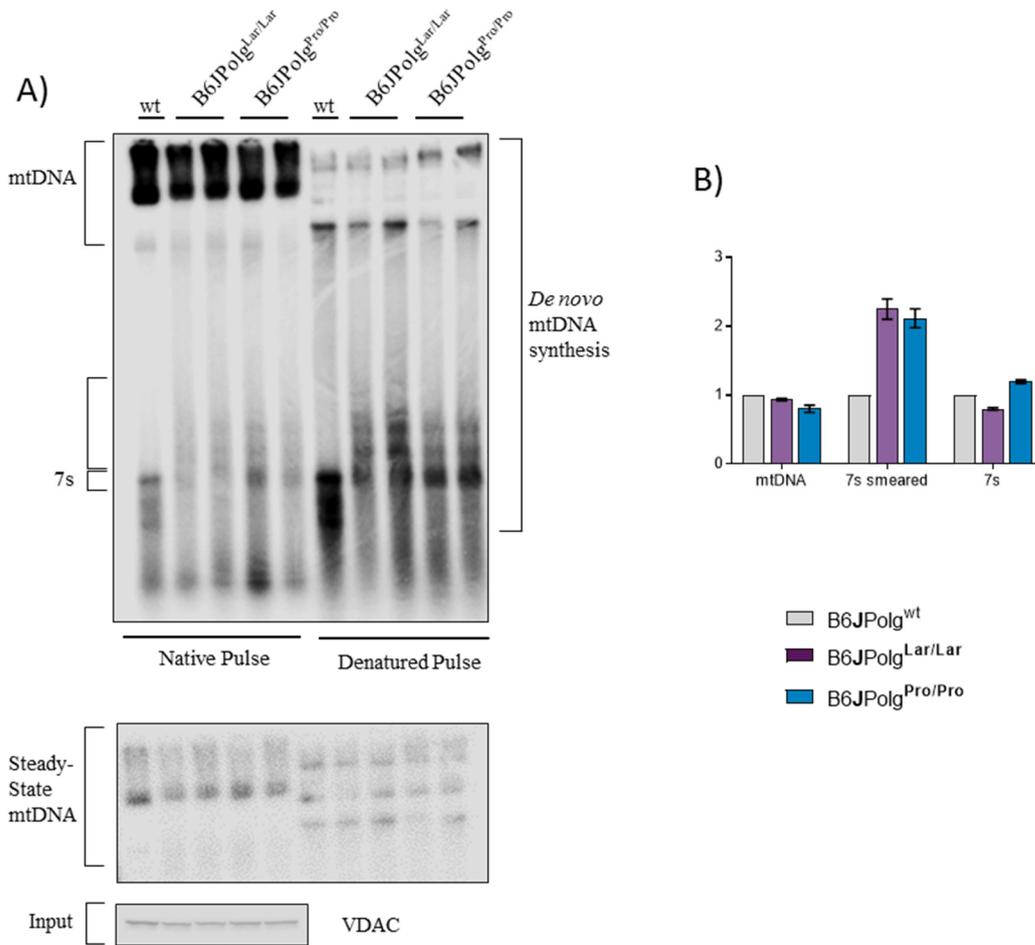
S 8. *In organello* **assay in mitochondria isolated from heart of 38 week-old mice.** We show the samples from the denatured pulse and native pulse, with the corresponding mtDNA steady state using the probe PAM1, and protein input by Western blot with VDAC as a control. B) Quantification of *de novo* mtDNA, *de novo* 7s and *de novo* 7s smear. Signals normalized to the mtDNA steady state with PAM1.

7. MATERIALS AND METHODS

7.1. Generation and Maintenance of Mice

7.1.1. Mouse husbandry

All mice were maintained in individually ventilated cages at the Max Planck Institute for Biology of Ageing (MPI-AGE). Cages conditions were at constant temperature of 21°C, 50 - 60% of humidity and with a 12 h light/dark cycle. Mice were fed *ad libitum*. Experiments were approved by the Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen (LANUV) in accordance with German and European Union regulations (ID 84-02.04.2015.A103). All procedures were performed following the recommendations of the Federation of the European Laboratory Animal Science Association (FELASA) recommendations.

7.1.2. Maintenance and generation of mutant mice

Wild-type C57BL/6J mice and mutator B6JPolg$^{Pro/Pro}$ (strain name B6.129S7(Cg)-*Polg^{tm1Prol}*/J) mice were purchased from Charles River Company in Germany. Wild-type C57BL/6NCrl mice and mutator B6NPolg$^{Lar/Lar}$ were obtained from Professor Nils-Göran Larsson. To avoid maternally transmitted mutations in the mutator mice, each mutant line was maintained backcrossing wild-type females with heterozygous mutant males. To obtain homozygous animals, male and female heterozygous mutants were crossed.

7.2. Molecular biology methods

7.2.1. Mouse genotyping

For genotyping, DNA from tail or ear-clip biopsies of two-three week old pups were extracted using 70 ul of buffer I (25mM NaOH, 0.2mM EDTA) and incubated for 45 min at 96°C. Later, samples were incubated a few minutes on ice and 70 ul of buffer II (40mM Tris pH 7.5 -8) was add it. Samples were storage at 4 °C until use.

Sample DNA	1 ul
dH2O	10 ul
Buffer 5X	4 ul
dNTPs (1.25 mM each)	3.2 ul
Primer Fw (100uM)	0.8 ul
Primer Rev (100uM)	0.8 ul
GoTaq	0.1 ul

Table 7-1. General PCR reaction mix

Stage	Step	Temperature (°C)	Time	Expected Products
1 cycle	1	95°C	1 min	Wild-type: ~520 bp
30 cycles	1	95°C	30 sec	Heterozygous: ~720 + ~520 bp
	2	58 °C	30 sec	Homozygous: ~720bp (mutator)
	3	72°C	45 sec	
1 cycle	1	72°C	2min	
	2	8°C	∞	

Table 7-2. POLG PCR Program to amplify *Lar* allele.

Stage	Step	Temperature (°C)	Time	Expected products
1 cycle	1	94°C	3 min	Wild-type: ~300 bp
10 cycles	1	94°C	20 sec	Heterozygous: ~500 + ~300 bp
	2	65 (°C)	15 sec	Homozygous: ~500 bp (mutator)
	3	68°C	10 sec	
28 cycles	1	94°C	15 sec	
	2	60 (°C)	15 sec	
	3	72°C	10 sec	
1 cycle	1	72°C	2min	
	2	8°C	∞	

Table 7-3. POLG PCR Program to amplify *Pro* allele.

Stage	Step	Temperature (°C)	Time	Expected Products
1 cycle	1	95°C	1 min	Wild-type: <700 bp
30 cycles	1	95°C	30 sec	mut: >700 bp
	2	55 °C	30 sec	KO= ~399 bp
	3	72°C	45 sec	
1 cycle	1	72°C	2min	
	2	8°C	∞	

Table 7-4 POLG PCR Program to amplify *KO* allele.

Allele to detect	Genotyping primers (5'-3')
Pro (genotyping)	CTCGCTTTCTCCGTGACTG
	ATGTGGCCCAGGCTGTAACT
Lar (genotyping)	CTTCGGAAGAGCAGTCGGGTG
	GGGCTGCAAAGACTCCGAAGG
KO (genotyping)	CTTCGTCGATCGACCTCGAATAAC
	CTGCCATTCACCTTACCC
	GGATGGGCAGGAACAGTTAG
WANCY (to detect mtDNA)	AGTATAGTAATGCCTGCGGC
	CCTACCCTAGCCCCCC

Table 7-5 List of primers used for the PCR

7.2.2. Blood cell count

Blood (~500ul) was sent to LABOKLIN (LABOKLIN GMBH & CO.KG, Labor Für Klinische Diagnostik) in heparin coated blood collection tubes. The laboratory performed complete blood count with flow cytometry incl. peroxidase stain.

7.2.3. mtDNA extraction

Protocol from Isokallio *et al.*, (2018) was followed. Briefly, fresh tissue (~300 mg liver) was cut into small pieces and washed with PBS (Thermo Fisher, 14190). All procedures were performed at 4°C, unless another temperature is indicated. Tissue was homogenized with 10 teflon pestle strokes at 200 rpm in buffer MIB1 (320 mM sucrose, 20 ml Tris-HCl, 1 mM EGTA, 0.2% BSA, pH 7.2). Then, mitochondria were differential centrifugation purified (two times at 800 g for 10 min and 1 time at 8,500 g for 10 min). The resulting

mitochondrial pellet was re-suspended in 1.4 mL of Mito-DNase buffer (300 mM sucrose, 10 mM MgCl2, 20 mM Tris HCl, pH 7.5, 0.15% BSA, 0.03 mg/ml DNase I type IV, 170 ng/µl RNase A), and incubated for 1 hr at 37°C. Subsequently, mitochondria were pelleted at 13 000 g, for 15 min and washed with 500 µl of MIB02 (320 mM sucrose, 20 ml Tris-HCl, 1 mM EGTA, 0.2% BSA, pH 7.2), two times. Then, samples were fast-frozen in liquid nitrogen and stored at -80° C until use.

The frozen mitochondrial pellets were re-suspended in 400 µl of lysis buffer (20 mM Tris HCl, 150 mM NaCl, 20 mM EDTA, 1% SDS, pH 8.7, 0.2 mg/mL Proteinase K and 0.2 mg/ml RNase A). Samples were incubated at 56°C overnight. The next day, the samples were cooled down at room temperature. To extract the mtDNA, 100 ul of 6M potassium acetate was added, then 500 µl of chloroform:isoamylalcohol 24:1 (Thermos Fisher, AM9720) was added. The tubes were shaken for 20 sec.

200 µg RNase A (Macherey-Nagel, CAS 9001-99-4) was added and incubated for 45-60 min at 37°C. To precipitate the DNA, 1 mL of cold ethanol was added. Samples were incubated for 1 hr at -80 °C (or overnight at -20 °C). Finally, samples were centrifuged at 16,000g at room temperature, and the pellet was re-suspended in 25 µl of AE buffer (5 mM Tris/HCl pH 8.5, Macherey-Nagel). Quantification was performed with a fluorometric method, using the reagents dsDNA Qubit (ThermoFisher, Q32850). The mtDNA purity (nuclear DNA free) was confirmed by the lack of amplification using *Lar* or *Pro* primers and by the amplification using mtDNA specific "WANCY" primers Table 7-5.

7.2.4. Total DNA extraction

Total DNA was extracted from snap-frozen tissues with Gentra Puregene Tissue Kit (QIAGEN, 51404), following the indicated manufacturer's recommendations. For qPCR, DNA was treated with RNase (Macherey-Nagel, CAS 9001-99-4) during DNA extraction according to the QIAGEN protocol. DNA concentration was determined with the spectrophotometer NanoDrop 2000c (Thermo Fisher).

7.2.5. Total RNA extraction

Ground, snap-frozen mouse tissues were homogenized using lysing matrix D (MPBio) prep machine with two pulses of 30 sec at speed 6. The homogenization step was repeated after 5 min incubation at 4 °C. Next, TRIzol LS reagent (Thermo Fisher, 10296028) was used to extract total RNA, following the manufacturer's instructions. The isolated RNA was DNase treated using the TURBO DNA-*free* kit (Ambion, AM2238), according to manufacturers' instructions. The RNA concentration was determinate with the spectrophotometer NanoDrop 2000c (Thermo Fisher). The RNA integrity was analyzed in agarose-SDS gel (1% agarose, 0.1%SDS). RNA purity (DNA free) was confirmed by the lack of amplification using *Lar* or *Pro* PCR primers.

7.2.6. cDNA Synthesis

cDNA was generated from RNA (DNase treated) using the High Capacity cDNA reverse transcription kit (Thermo Fisher,4368814), following manufacturers' instructions.

7.2.7. qPCR and RT-qPCR

qPCR and RT-qPCR were performed using the Taqman 2x Universal PCR mastermix (Applied Biosystems, 4324020) and commercially available Taqman probes (Life technologies), on a QuantStudio 6 RealTime PCR System (Applied Biosystems). All reactions were made in triplicates with 5 ng of total DNA or cDNA (diluted 1:10 or 1:30), using 1 ul of sample per reaction (final volume 10 μl). The amplification data were analyzed with a standard-curve method using an artificial curve (mix of samples).

Probe	Name	Assay ID
POLG	Mitochondrial Polymerase gamma, exon boundary 2-3	Mm00450527_m1
POLG	Mitochondrial Polymerase gamma, exon boundary 3-4	Mm00450528_m1
TBP	TATA-box binding protein	Mm00446973_m1
B2M	beta2-microglobulin	Mm00437762_m1
18s	18s rRNA-housekeeping gene	Hs99999901_s1
16s	mt 16s rRNA (Rnr2)	Mm04260181_s1

| ND2 | mt NADH dehydrogenase 2 | Mm04225288_s1 |
| ND4 | mt NADH dehydrogenase 4 | Mm04225294_s1 |

Table 7-6 List of Taqman Probes

7.2.8. Southern blotting

5 ug of DNA was digested over night at 37°C with SacI-HF enzyme (New England BioLabs, R3156), following manufacturers' instructions. Then, DNA was precipitated with 8 ul of 5M NaCl and 500 ul of ethanol 95%. Samples were incubated overnight at -80 °C and centrifuged at 13,000 rpms at room temperature. Pellet was washed with ethanol 75% and DNA re-suspended in 20 ul of TE buffer. The DNA was quantified by nanodrop. Subsequently, 5 ul of DNA was loaded with 6X loading dye (New England Biolabs, B7024S) and DNA separated on a 0.8 % agarose gel, running at 40 volts for about 16 hours.

Gel was pre-treated with 0.2 M HCl for 20 minutes, two washes with water (5 minutes each), denaturizing solution (1.5 M NaCl and 0.5 M NaOH) for 20 minutes, two washes with water (5 minutes each), and neutralization solution (1.5 M NaCl and 0.5 M Tris-HCl pH 7.4) for 20 minutes. Hybond-N+ membrane (Amersham, RPN303B) was incubated in water for 10 min. Then, the gel and membrane were incubated for 20 minutes in 20x SSC (3 M Sodium chloride and 300 mM tri-Sodium citrate dihydrate, pH 7.0). The Southern blot was assembled with whatman paper, membrane, gel, whatman paper, and paper tower to transfer the DNA. The transfer was performed for approximately 48 hours. The membrane was UV-cross-linked and labeled with a radioactive probe.

For radioactive labeling, DNA probes were radioactive-labelled with [α-32P] dCTP (Hartmann Analytic, Catalog #srp-205) using Prime-It® II Random Primer Labeling Kit (Agilent, catalog #300385), following manufacturers' instructions. Then the probe was purified with the kit ProbeQuant G-50 Micro Columns (GE Healthcare, 28903408) according to manufacturers' instructions. Next, the membrane was incubated in Perfect Hyb Plus hybridization buffer (Sigma-Aldrich, H7033) for one hour and then, incubated at 65 °C overnight with radioactive probes and random primers (Roche, 11034731001). Membranes were washed in 2X SSC, 0.1% SDS, 15 min at room temperature and 15 min at 65 °C. In the end, the membrane was sealed in a plastic bag, exposed to a phosphor screen,

and developed using the PhosphorImager instrument Typhoon FLA 7000 (GE healthcare). Quantifications were performed using the software MultiGauge (v3.0).

Probe	Sequence
18S	GGTCTACAAGACGCCACATCCCCTATTATAGAAGAGCTAATAAATTTC CATGATCACACACTAATAATTGTTTTCCTAATTAGCTCCTTAGTCCTCT ATATCATCTCGCTAATATTAACAACAAAACTAACACATACAAGCACAA TAGATGCACAAGAAGTTGAAACCATTTGAACTATTCTACCAGCTGTAA TCCTTATCATAATTGCTCTCCCCTCTCTACGCATTCTATATATAATAGAC GAAATCAACAACCCCGTATTAACCGTTAAAACCATAGGGCACCAATGA TACTGAAGCTACGAATATACTGACTATGAAGACCTATGCTTTGATTCAT ATATAATCCCAACAAACGACCTAAAACCTGGTGAACTACGACTGCTAG AAGTTGATAACCGAGTCGTTCTGCCAA
pAM1	Entire mouse mtDNA cloned in pAcyc177
7s DNA	Entire mouse control region (15423..16299) cloned to pCRIITOPO ACR

Table 7-7. Probes for Southern blotting

7.2.9. *In organello* replication assay

In organello replication was performed on freshly isolated mitochondria from mouse hearts by differential centrifugation as described in method 7.2.10. Mitochondria were quantified with DC protein quantification assay (Method 7.3.1). 1 mg of freshly isolated mitochondria were re-suspended in transcription labeling buffer (25 mM sucrose, 75 mM sorbitol, 100 mM KCl, 10 mM K2HPO4, 50 μM EDTA, 5 mM MgCl2, 1 mM ADP, 10 mM glutamate, 2.5 mM malate, 10 mM Tris pH 7.4, 1 mg/ml of bovine serum albumin (BSA), and 1 mM ADP). 50 uM of each dNTP (except dATP) and 20uCi of [α-32P] adenosin-5'-triphoshphate (dATP) was added. Samples were incubated at 37°C for 2 hours. An aliquot was taken for chase and incubated for 1 extra hour. Later, mitochondria were pelleted at 9000 rpms for 4 min at 4 °C, and mitochondria washed two times in washing buffer (10% glycerol, 10 mM Tris-HCl pH 6.8, and 0.15 mM MgCl2). An aliquot was taken for western blot analysis (Method 7.3.2.). The pellet was re-suspended in 150 ul of PK buffer (50 mM TrisHCl pH 7.4, 75 mM NaCl, 6 mM EDTA and 1% SDS). Then, with 200 ul of phenol/chloroform, the samples were shaken for 20 seconds and centrifuged at 16,000 g for 5 minutes. DNA was precipitated adding 50 ul 5M Ammonium acetate and 250 ul cold 100% ethanol. The tubes were placed at -20°C for at least 3 hours and centrifuged for 20 min at 4°C. Then, ethanol was removed and the pellet was re-suspended in 30 ul of TE solution (QIAGEN, 1018499). DNA was loaded in an agarose gel 0.8% and ran at 90V for 160 min.

Subsequently, the gel was transferred to Hybond-N+ membrane (Amersham, RPN303B) as previously described (Method 7.2.8). In the end, the membrane was sealed in a plastic bag and exposed to a phosphor screen and developed using the PhosphorImager instrument Typhoon FLA 7000 (GE healthcare). Quantifications were performed using the program MultiGauge (v3.0).

7.2.10.   Isolation of mitochondria from heart

Mice were euthanized and the heart extracted. The heart was rinsed and cut in ice-cold PBS (Thermo Fisher, 14190). The tissue was placed in a 5 ml Potter S homogenizer containing 5 ml of isolation buffer (320 mM sucrose, 1 mM EDTA, 10 mM Tris-HCl pH 7.4 and 1X Roche complete protease inhibitors). The tissue was homogenized in the Homogen plus homogenizer (Schuett-biotech) with 16 strokes at 300 rpm. 5 mL of isolated buffer was added and samples centrifuged for 10 minutes at 1,000 rcf and 4°C. Supernatants were transferred into a new tube and centrifuged at 1000 rcf. Then, mitochondria pelleted was centrifuged 10,000 rcf for 10 minutes at 4°C. Crude mitochondria were suspended in isolation buffer and used for *in organello* assay or stored at -80°C until use.

7.2.11.   Sperm analysis

Cauda region epididymis of mice was clamped proximally and distally by Dr. Jiang. Tissue was rinsed in pre-warmed (37°C) PBS (Thermo Fisher, 14190). Later, the tissue was placed in pre-warmed (37°C) M2 medium (SIGMA). The cauda epididymis was then unclamped and pierced with a scalpel blade to release the sperm. Then, sperm were allowed to disperse for 10 min at 37°C. Sperm relative number and sperm motility were determinate by CASA detection (Hamilton Thorne Research) at the Transgenesis core facility. To determine sperm morphology, sperm were spread on glass slides and fixed with 4% PFA (X). Subsequently, slides were stained with H&E (Method 7.3.5). Deformities were classified according to (Wyrobek and Bruce, 1975).

7.3. Biochemical methods

7.3.1. DC protein assay

Protein quantification was determinate using DC Protein Assay (BioRad, LIT448), following the standard assay protocol from manufacturers' instructions.

7.3.2. Western blotting

Isolated mitochondria pellet were re-suspended in 4X NuPAGE LDS sample buffer (Invitrogen, NP0007) with 50 mM DTT. Proteins were separated using precast gels 4-12% Bis-Tris (Invitrogen, WG1402BOX) and the electrophoretic NuPAGE system (Invitrogen). The gel was wet-transferred to a PVDF membrane (Merck, IPVH00010), with transfer buffer (38.6 mM Tris, 47.8 mM glycine, 20% methanol), using a Criterion blotter chamber (Bio-Rad). Next, the membrane was blocked with 5% milk in T-TBS (150 mM NaCl, 50 mM Tris-HCl pH 7.6 and 0.05% TWEEN-20). Primary antibodies were also diluted in T-TBS 5% milk, followed by 3 washes with T-TBS. Secondary antibodies with horseradish peroxidase-linked were also diluted in T-TBS 5% milk, followed by 3 washes with T-TBS. Final wash was performed with PBS (Thermo Fisher, 14190). Visualization was obtained with chemioluminescence ECL Western Blotting Detection Reagent (GE Healtcare Life Sciences, RPN2106) and imaging system (LI-COR Biosciences).

Antibody	Company	Cataloge Number	Dilution
POLG	Abcam	ab128899	1:500
VDAC	Abcam	ab14734	1:1,000 or 1:10,000 (isolated mitochondria)
TOM20	Santa Cruz Biotechnology	sc-11415	1:5,000
TFAM	Abcam	ab131607	1:1,000
Secondary Anti-Mouse, HRP	Merck	NXA931	1:5,000
Secondary Anti-Rabbit, HRP	Thermo Fisher	g21234	1:5,000

Table 7-8. List of Antibodies and dilutions.

7.3.3. NBTx staining

For Nitrotetrazolium Blue Exclusion Assay (NBTx), tissue sections of 10 um (colon and skeletal muscle) and 7 um (heart) of tissue was generated in the cryostat at the FACS & Imaging core facility. Samples were incubated 10 min with 1mL PBS, then the PBS was discarded. 1 mL of NBTx solution was added (130 mM sodium succinate, 200 μM phenazinemethosulphate, 1.5 mM nitroblue tetrazolium in 0.2 M phosphate buffer pH 7.0). Then, sections were washed for 5 min with PBS (PBS, 0.1 M, pH 7.0). Dehydration was performed through graded ethanol series: 2 min in 50%, 2 min in 75%, 2 min in 96%, 2 min in 100%, and 5 min in 100%. Finally, slides were incubated in two Xylene changes, 5 min each. Mounting was performed with Cystoseal (Thermo Scientific, 8312-4). Imaging was performed with Nikon Eclipse Ci Microscope at the FACS & Imaging core facility. Image mitochondrial dysfunction quantification was performed with Image J software, following image settings and macro specifications as previously described by Simard *et al*., (2018).

7.3.4. COX-SDH staining

For cytochrome c oxidase (COX)/ succinate dehydrogenase (SDH) enzyme histochemistry (COX-SDH), tissue sections of 10 um of the colon were generated in the cryostat at the FACS & Imaging core facility. Slide samples were incubated at 37 °C in 100 μl of COX staining medium (100 μM cytochrome c, 4 mM diaminobenzidine tetrahydrochloride, 20 μg/ml catalase in 0.2 M phosphate buffer pH 7.0) for 45 min. Subsequently, the sections were washed 3 times with PBS for 5 min. Next, samples were incubated at 37 °C for 20 min in 50 μl of SDH solution (130 mM sodium succinate, 200 μM phenazinemethosulphate, 1 mM sodium azide, 1.5 mM nitroblue tetrazolium in 0.2 M phosphate buffer pH 7.0). Then, sections were washed 3 times for 5 min with PBS (PBS, 0.1 M, pH 7.0). Dehydration was performed through graded ethanol series: 2 min in 50%, 2 min in 75%, 2 min in 96%, 2 min in 100%, and 5 min in 100%. Finally, slides were incubated in two Xylene changes, 5 min each. Mounting was performed with Cystoseal (Thermo Scientific, 8312-4). Imaging was performed with Nikon Eclipse Ci Microscope at the FACS & Imaging core facility.

7.3.5. H&E staining

Sperm and heart sections were fixed with 4% Paraformaldehyde (Sigma-Aldrich, 28908). Haematoxylin and Eosin Y staining was performed using Gemini AS Automated Slider (Thermo Fisher), according to manufactures protocol. Solutions used were Haematoxylin (AppliChem) and Eosin Y (Carl Roth). Imaging was performed with Nikon Eclipse Ci Microscope at the FACS & Imaging core facility.

7.3.6. Toloudine staining

Tissue sections of 10 um (colon and skeletal muscle) and 7 um (heart) thickness were fixed with 4% Paraformaldehyde (Sigma-Aldrich, 28908). Then, 1 mL of Toluidine solution was added (Toluidine blue 0.1% in 1% sodium chloride at pH 2.3) and incubated for 3 minutes at room temperature. Sections were washed 3 times in dH20 and dehydrated (fast dipping a few times, first in 95%, then 100% and 100% again). Finally, slides were incubated in Xylene for 5 minutes, two times. Mounting was performed with Cytoseal (Thermo Scientific, 8312-4). Imaging was performed with Nikon Eclipse Ci Microscope at the FACS & Imaging core facility.

7.4. Sequencing

7.4.1. RNA Sequencing and Data analysis

Total RNA samples (~1.5 ug) from the heart were sent for sequencing at Max-Planck Genome-center Cologne (MP-GC). Sequencing was performed in Illumina HiSeq3000, paired-end mode, using TruSeq compatible directional RNA library. The sequencing libraries were prepared by the MP-GC. Analysis of the row data, determination of fold change expression, KEGG pathways and MD plots were performed by the Bioinformatics Core Facility.

Mitochondrial related genes analysis was performed using the web-platform mitoXplorer v.1.0 (http://mitoxplorer.ibdm.univ-mrs.fr/), using only genes with an adjusted p-value

<0.05. Functional enrichment analysis was performed in software QIAGEN's Ingenuity® Pathway Analysis (IPA®, https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/).

7.4.2. mtDNA sequencing

Enriched mtDNA samples (>100 ng) from the liver were sent for sequencing at Max-Planck Genome-center Cologne (MP-GC). Sequencing was performed in Illumina HiSeq3000 in 1 x 150bp mode, using TruSeq compatible DNA library. The sequencing libraries were prepared by the MP-GC. Data were analyzed by the Bioinformatics Core Facility.

7.4.3. Pyrosequencing

Allele quantification was performed by pyrosequencing PyroMark Q24 Pyrosequencer (QIAGEN) on cDNA from heart, colon, liver and skeletal muscle of compound heterozygous mice B6JPolg$^{Lar/Pro}$. First, a PCR was performed to amplify from exon 3 to exon 4 (see table primers forward and revers), and from exon 3 to intron 3-4 (see table primers forward and revers). Second, PCR products were combined with binding buffer, Streptavidin sepharose TM high performance beads (GE Healthcare, GE17-5113-01) annealing buffer (QIAGEN, 979009) and primer for sequencing (see table, 7-9) in the Pyromark Q24 vacuum workstation (QIAGEN) according to manufacturer's instructions. Finally, sequencing was performed with PyroMark Gold Q24 reagents (QIAGEN, 970802) following the manufacturer's instructions.

Exon 3 to exon 4	Forward	[Btn]TGGTGGGGCACAATGTTTC
	Reverse	AGCTGCTCAGCCCCGAGAT
	For Sequencing	CCTGATATGGGCTCG
Exon 3 to intron 3-4	Forward	AGTTAGTGGTGGGGCACAAT
	Reverse	[Btn]TGGAAACAGTCACGGAGA
	For Sequencing	TGGAAACAGTCACGGAGA

Table 7-9. List of primers used for Pyrosequencing

7.5. Statistical analysis

7.5.1. Data graphic and statistical analysis

Data graphs and statistical analyses were done using GraphPad Prism 5 software. For all experiments, the data were assumed to follow a normal distribution. Data are presented as means ± SEM (standard error of the mean). Statistical comparisons between more than two groups were performed using one-way analysis of variance (ANOVA) in multiple comparisons, with post hoc analysis conducted with Tukey's multiple comparison test. For two groups Student's t test was performed. Statistical significance was considered if at $P < 0.05$ (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$). All statistical parameters are outlined at the figure legends.

8. REFERENCES

Adams, K.L., and Palmer, J.D. (2003). Evolution of mitochondrial gene content: Gene loss and transfer to the nucleus. *Mol. Phylogenet. Evol. 29*, 380–395.

Adamson, S.I., Zhan, L., and Graveley, B.R. (2018). Vex-seq: high-throughput identification of the impact of genetic variation on pre-mRNA splicing efficiency. *Genome Biol. 19*, 71.

Ahlqvist, K.J., Leoncini S., Pecorelli A., *et al. (*2015). MtDNA mutagenesis impairs elimination of mitochondria during erythroid maturation leading to enhanced erythrocyte destruction. *Nat Commun. 6*, 6494

Ahlqvist, K.J., Hämäläinen, R.H., Yatsuga, S., Uutela, M., Terzioglu, M., Götz, A., Forsström, S., Salven, P., Angers-Loustau, A., Kopra, O.H., et al. (2012). Somatic progenitor cell vulnerability to mitochondrial DNA mutagenesis underlies progeroid phenotypes in polg mutator mice. *Cell Metab. 15*, 100–109.

Ahlqvist, K.J., Suomalainen, A., and Hämäläinen, R.H. (2015). Stem cells, mitochondria and aging. *Biochim. Biophys. Acta – Bioenerg. 1847*, 1380–1386.

Alexeyev, M.F. (2009). Is there more to aging than mitochondrial DNA and reactive oxygen species? *FEBS J. 276*, 5768–5787.

Allen, J.F. (2003). Why chloroplasts and mitochondria contain genomes. *Comp. Funct. Genomics 4*, 31–36.

Allen, C.A., Håkansson, G., and Allen, J.F. (1995). Redox conditions specify the proteins synthesised by isolated chloroplasts and mitochondria. *Redox Rep. 1*, 119–123.

Alston, C.L., Rocha, M.C., Lax, N.Z., Turnbull, D.M., and Taylor, R.W. (2017). The genetics and pathology of mitochondrial disease. *J. Pathol. 241*, 236–250.

Altmann, R. (1890). Altmann, Richard: Die Elementarorganismen und ihre Beziehungen zu den Zellen. (Leipzig: Verlag von Veit & Comp.). [Internet*:

http://www.deutschestextarchiv.de/book/view/altmann_elementarorganismen_1890?p=9]

Ameur, A., Stewart, J.B., Freyer, C., Hagström, E., Ingman, M., Larsson, N.G., and Gyllensten, U. (2011). Ultra-deep sequencing of mouse mitochondrial DNA: Mutational patterns and their origins. *PLoS Genet. 7*, 18–21.

Anderson, S., Bankier, A.T., Barrell, B.G., De Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., *et al.* (1981). Sequence and organization of the human mitochondrial genome. *Nature 290*, 457–465.

Andrews, R.M., Kubacka, I., Chinnery, P.F., Lightowlers, R.N., Turnbull, D.M., and Howell, N. (1999). Reanalysis and revision of the cambridge reference sequence for human mitochondrial DNA. *Nat. Genet. 23*, 147.

Anna, A., and Monika, G. (2018). Splicing mutations in human genetic disorders: examples, detection, and confirmation. *J. Appl. Genet. 59*, 253–268.

Arimura, S.I., Yamamoto, J., Aida, G.P., Nakazono, M., and Tsutsumi, N. (2004). Frequent fusion and fission of plant mitochondria with unequal nucleoid distribution. *Proc. Natl. Acad. Sci. U. S. A. 101*, 7805–7808.

Bailey, L.J., Cluett, T.J., Reyes, A., Prolla, T.A., Poulton, J., Leeuwenburgh, C., and Holt, I.J. (2009). Mice expressing an error-prone DNA polymerase in mitochondria display elevated replication pausing and chromosomal breakage at fragile sites of mitochondrial DNA. *Nucleic Acids Res. 37*, 2327–2335.

Bain, B. (2015). Blood Cells: A Practical Guide, 5th Edition (Chichester, UK: Wiley-Blackwell Publishing Ltd). pp. 237

Baines, H.L., Stewart, J.B., Stamp, C., Zupanic, A., Kirkwood, T.B.L., Larsson, N.G., Turnbull, D.M., and Greaves, L.C. (2014). Similar patterns of clonally expanded somatic mtDNA mutations in the colon of heterozygous mtDNA mutator mice and ageing humans. *Mech. Ageing Dev. 139*, 22–30.

Baruffini, E., Lodi, T., Dallabona, C., Puglisi, A., Zeviani, M., and Ferrero, I. (2006). Genetic and chemical rescue of the Saccharomyces cerevisiae phenotype induced by

mitochondrial DNA polymerase mutations associated with progressive external ophthalmoplegia in humans. *Hum. Mol. Genet. 15*, 2846–2855.

Bates, M.G.D., Bourke, J.P., Giordano, C., D'Amati, G., Turnbull, D.M., and Taylor, R.W. (2012). Cardiac involvement in mitochondrial DNA disease: Clinical spectrum, diagnosis, and management. *Eur. Heart J. 33*, 3023–3033.

Bayona-Bafaluy, M.P., Acín-Pérez, R., Mullikin, J.C., Park, J.S., Moreno-Loshuertos, R., Hu, P., Perez-Martos, A., Fernández-Silva, P., Bai, Y., and Enríquez, J.A. (2003). Revisiting the mouse mitochondrial DNA sequence. *Nucleic Acids Res. 31*, 5349–5355.

Bensley, R.R., and Hoerr, N.L. (1934). Studies on cell structure by the freezing-drying method VI. The preparation and properties of mitochondria. *Anat. Rec. 6*, 449–455.

Bickel, M. (1993). The role of interleukin-8 in inflammation and mechanisms of regulation. *J. Periodontol. 64*, 456–460.

Björkholm, P., Ernst, A.M., Hagström, E., and Andersson, S.G.E. (2017). Why mitochondria need a genome revisited. *FEBS Lett. 591*, 65–75.

Blanchard, J.L., and Lynch, M. (2000). Organellar genes. Why do they end up in the nucleus? *Trends Genet. 16*, 315–320.

van der Bliek, A.M., Shen, Q., and Kawajiri, S. (2013). Mechanisms of mitochondrial fission and fusion. *Cold Spring Harb. Perspect. Biol. 5*, a011072.

Bogenhagen, D., and Clayton, D.A. (1978). Mechanism of mitochondrial DNA replication in mouse L-cells: Kinetics of synthesis and turnover of the initiation sequence. *J. Mol. Biol. 119*, 49–68.

Bonekamp, N.A., and Larsson, N.G. (2018). SnapShot: Mitochondrial Nucleoid. *Cell 172*, 388-388.

Boos, F., Wollin, M., and Herrmann, J.M. (2016). Methionine on the rise: how mitochondria changed their codon usage. *EMBO J. 35*, 2066–2067.

Borst, P. (1972). Mitochondrial Nucleic Acids. *Annu. Rev. Biochem. 41*, 333–376.

Borst, P., Van Bruggen, E.F.J., Ruttenberg, G.J.C.M., and Kroon, A.M. (1967). Mitochondrial DNA II. Sedimentation analysis and electron microscopy of mitochondrial DNA from chick liver. *BBA Sect. Nucleic Acids Protein Synth. 149*, 156–172.

Boussau, B., Karlberg, E.O., Frank, A.C., Legault, B.A., and Andersson, S.G.E. (2004). Computational inference of scenarios for α-proteobacterial genome evolution. *Proc. Natl. Acad. Sci. U. S. A. 101*, 9722–9727.

Bratic, A., and Larsson, N. (2013). Review series The role of mitochondria in aging. *J. Clin. Invest. 123*, 951–957.

Brown, T.A., Cecconi, C., Tkachuk, A.N., Bustamante, C., and Clayton, D.A. (2005). Replication of mitochondrial DNA occurs by strand displacement with alternative light-strand origins, not via a strand-coupled mechanism. *Genes Dev. 19*, 2466–2476.

Cao, X.W., Lin, K., Li, C.Y., and Yuan, C.W. (2011). [A review of WHO Laboratory Manual for the Examination and Processing of Human Semen (5th edition)]. *Zhonghua Nan Ke Xue 17*, 1059–1063.

Capps, G.J., Samuels, D.C., and Chinnery, P.F. (2003). A model of the nuclear control of mitochondrial DNA replication. *J. Theor. Biol. 221*, 565–583.

Carrodeguas, J.A., Pinz, K.G., and Bogenhagen, D.F. (2002). DNA binding properties of human pol γB. *J. Biol. Chem. 277*, 50008–50014.

Cerritelli, S.M., Frolova, E.G., Feng, C., Grinberg, A., Love, P.E., and Crouch, R.J. (2003). Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. *Mol. Cell 11*, 807–815.

Chan, S.S.L., and Copeland, W.C. (2009). DNA polymerase gamma and mitochondrial disease: Understanding the consequence of POLG mutations. *Biochim. Biophys. Acta - Bioenerg. 1787*, 312–319.

Chan, S.S.L., Longley, M.J., and Copeland, W.C. (2006). Modulation of the W748S mutation in DNA polymerase γ by the E1143G polymorphism in mitochondrial

disorders. *Hum. Mol. Genet. 15*, 3473–3483.

Chen, C., Turnbull, D.M., and Reeve, A.K. (2019). Mitochondrial dysfunction in Parkinson's disease—cause or consequence?. Biology (Basel) *8(2), 38*.

Chen, M.L., Logan, T.D., Hochberg, M.L., Shelat, S.G., Yu, X., Wilding, G.E., Tan, W., Kujoth, G.C., Prolla, T.A., Selak, M.A., *et al.* (2009). Erythroid dysplasia, megaloblastic anemia, and impaired lymphopoiesis arising from mitochondrial dysfunction. *Blood 114*, 4045–4053.

Chinnery, P.F., and Hudson, G. (2013). Mitochondrial genetics. *Br. Med. Bull. 106*, 135–159.

Christoffersson, G., and Phillipson, M. (2018). The neutrophil: one cell on many missions or many cells with different agendas? *Cell Tissue Res. 371*, 415–423.

Clayton, D.A. (1982). Replication of animal mitochondrial DNA. *Cell 28*, 693–705.

Clayton, D.A. (2000). Transcription and replication of mitochondrial DNA. In Human Reproduction, (Oxford University Press), *pp. 11–17*.

Cogliati, S., Enriquez, J.A., and Scorrano, L. (2016). Mitochondrial Cristae: Where Beauty Meets Functionality. *Trends Biochem. Sci. 41*, 261–273.

Cohen, B.H., and Naviaux, R.K. (2010). The clinical diagnosis of POLG disease and other mitochondrial DNA depletion disorders. *Methods 51*, 364–373.

Cortopassi, G.A., and Arnheim, N. (1990). Detection of a specific mitochondrial DNA deletion in tissues of older humans. *Nucleic Acids Res. 18*, 6927–6933.

Cortopassi, G.A., Shibata, D., Soong, N.W., and Arnheim, N. (1992). A pattern of accumulation of a somatic deletion of mitochondrial DNA in aging human tissues. *Proc. Natl. Acad. Sci. U. S. A. 89*, 7370–7374.

Cowdry, E. (1953). Historical background of research on mitochondria. *J. Histochem. Cytochem. 1*, 183–187.

Curie, M. (1923). Pierre Curie: With Autobiographical Notes by Marie Curie (Dover Pubns: THE SAINT BOOKSTORE).

Dai, D.F., Chen, T., Wanagat, J., Laflamme, M., Marcinek, D.J., Emond, M.J., Ngo, C.P., Prolla, T.A., and Rabinovitch, P.S. (2010). Age-dependent cardiomyopathy in mitochondrial mutator mice is attenuated by overexpression of catalase targeted to mitochondria. *Aging Cell 9*, 536–544.

De Rooij D.G., Russell L.D. (2000). All you wanted to know about spermatogonia but were afraid to ask. *J Androl. 21(6)*:776-798.

Dillon, L.M., Hida, A., Garcia, S., Prolla, T.A., and Moraes, C.T. (2012). Long-Term Bezafibrate Treatment Improves Skin and Spleen Phenotypes of the mtDNA Mutator Mouse. *PLoS One 7*, 1–11.

Dogan, S.A., and Trifunovic, A. (2011). Modelling mitochondrial dysfunction in mice. *Physiol. Res. 60 Suppl 1*, S61-70.

Duran, J., Martinez, A., and Adler, E. (2019). Cardiovascular manifestations of mitochondrial disease. *Biology (Basel). 8(2), 34*.

Edgar, D., and Trifunovic, A. (2009). The mtDNA mutator mouse: Dissecting mitochondrial involvement in aging. *Aging (Albany. NY). 1*, 1028–1032.

Edgar, D., Shabalina, I., Camara, Y., Wredenberg, A., Calvaruso, M.A., Nijtmans, L., Nedergaard, J., Cannon, B., Larsson, N.G., and Trifunovic, A. (2009). Random Point Mutations with Major Effects on Protein-Coding Genes Are the Driving Force behind Premature Aging in mtDNA Mutator Mice. *Cell Metab. 10*, 131–138.

Eisfeld, A.J., Gasper, D.J., Suresh, M., and Kawaoka, Y. (2019). C57BL/6J and C57BL/6NJ Mice Are Differentially Susceptible to Inflammation-Associated Disease Caused by Influenza A Virus. *Front. Microbiol. 9, 3307*.

Ekstrand, M.I., Falkenberg, M., Rantanen, A., Park, C.B., Gaspari, M., Hultenby, K., Rustin, P., Gustafsson, C.M., and Larsson, N.G. (2004). Mitochondrial transcription factor A regulates mtDNA copy number in mammals. *Hum. Mol. Genet. 13*, 935–944.

Elson, J.L., Samuels, D.C., Turnbull, D.M., and Chinnery, P.F. (2001). Random Intracellular Drift Explains the Clonal Expansion of Mitochondrial DNA

Mutations with Age. *Am J Hum Genet 68*, 802–806.

Enríquez, J.A., Ramos, J., Pérez-martos, A., López-pérez, M.J., and Montoya, J. (1994). Highly efficient DNA synthesis in isolated mitochondria from rat liver. *Nucleic Acids Res. 22*, 1861–1865.

Ernst, C., Eling, N., Martinez-Jimenez, C.P., Marioni J.C., Odom D.T. (2019.) Staged developmental mapping and X chromosome transcriptional dynamics during mouse spermatogenesis. *Nat Commun* 10, 1251.

Ernster, L., and Schatz, G. (1981). Mitochondria : A Historical Review. *The Journal of Cell Biology* 1;91*91*.

Escobar Galvis, M.L., Allen, J.F., and Håkansson, G. (1998). Protein synthesis by isolated pea mitochondria is dependent on the activity of respiratory complex II. *Curr. Genet. 33*, 320–329.

Falkenberg, M. (2018). Mitochondrial DNA replication in mammalian cells: Overview of the pathway. *Essays Biochem. 62*, 287–296.

Farge, G., Mehmedovic, M., Baclayon, M., vandenWildenberg, S.M.J.L., Roos, W.H., Gustafsson, C.M., Wuite, G.J.L., and Falkenberg, M. (2014). *In Vitro*-reconstituted nucleoids can block mitochondrial DNA replication and transcription. *Cell Rep. 8*, 66–74.

Folgerø, T., Bertheussen, K., Lindal, S., Torbergsen, T., and Øian, P. (1993). Mitochondrial desease and reduced sperm motility. *Human Reprod. 8*, 1863–1868.

Foury, F., and Vanderstraeten, S. (1992). Yeast mitochondrial DNA mutators with deficient proofreading exonucleolytic activity. *EMBO J. 11*, 2717–2726.

Freeman, H.C., Hugill, A., Dear, N.T., Ashcroft, F.M., and Cox, R.D. (2006). Deletion of nicotinamide nucleotide transhydrogenase: A new quantitive trait locus accounting for glucose intolerance in C57BL/6J mice. *Diabetes 55*, 2153–2156.

Friedman, J.R., and Nunnari, J. (2014). Mitochondrial form and function. *Nature 505*, 335–343.

Fusté, J.M., Wanrooij, S., Jemt, E., Granycome, C.E., Cluett, T.J., Shi, Y., Atanassova, N., Holt, I.J., Gustafsson, C.M., and Falkenberg, M. (2010). Mitochondrial RNA Polymerase Is Needed for Activation of the Origin of Light-Strand DNA Replication. *Mol. Cell 37*, 67–78.

Gammage, P.A., and Frezza, C. (2019). Mitochondrial DNA: The overlooked oncogenome? *BMC Biol. 17*, 53

Gao, Y., Bai, X., Zhang, D., Han, C., Yuan, J., Liu, W., Cao, X., Chen, Z., Shangguan, F., Zhu, Z., et al. (2016). Mammalian elongation factor 4 regulates mitochondrial translation essential for spermatogenesis. *Nat. Struct. Mol. Biol. 23*, 441–449.

Giorgi, C., Marchi, S., and Pinton, P. (2018). The machineries, regulation and cellular functions of mitochondrial calcium. *Nat. Rev. Mol. Cell Biol. 19*, 713–730.

Goswami, R., and Kaplan, M.H. (2011). A Brief History of IL-9. *J. Immunol. 186*, 3283–3288.

Gray, M.W. (2012). Mitochondrial evolution. Cold Spring Harb. *Perspect. Biol. 4,9*

Graziewicz, M.A., Longley, M.J., and Copeland, W.C. (2006). DNA polymerase γ in mitochondrial DNA replication and repair. *Chem. Rev. 106*, 383–405.

Greaves, L.C., Preston, S.L., Tadrous, P.J., Taylor, R.W., Barron, M.J., Oukrif, D., Leedham, S.J., Deheragoda, M., Sasieni, P., Novelli, M.R., et al. (2006). Mitochondrial DNA mutations are established in human colonic stem cells, and mutated clones expand by crypt fission. *Proc. Natl. Acad. Sci. U. S. A. 103*, 714–719.

Greaves, L.C., Barron, M.J., Plusa, S., Kirkwood, T.B., Mathers, J.C., Taylor, R.W., and Turnbull, D.M. (2010). Defects in multiple complexes of the respiratory chain are present in ageing human colonic crypts. *Exp. Gerontol. 45*, 573–579.

Greaves, L.C., Elson, J.L., Nooteboom, M., Grady, J.P., Taylor, G.A., Taylor, R.W., Mathers, J.C., Kirkwood, T.B.L., and Turnbull, D.M. (2012). Comparison of Mitochondrial Mutation Spectra in Ageing Human Colonic Epithelium and Disease: Absence of Evidence for Purifying Selection in Somatic Mitochondrial DNA Point Mutations. *PLoS Genet. 8,11*.

Gustafsson, C.M., Falkenberg, M., and Larsson, N.-G. (2016). Maintenance and Expression of Mammalian Mitochondrial DNA. *Annu. Rev. Biochem. 85*, 133–160.

Haas, R.H. (2019). Mitochondrial dysfunction in aging and diseases of aging. *Biology (Basel). 8(2),48.*

Hakonen, A.H., Heiskanen, S., Juvonen, V., Lappalainen, I., Luoma, P.T., Rantamäki, M., Van Goethem, G., Löfgren, A., Hackman, P., Paetau, A., et al. (2005). Mitochondrial DNA polymerase W748S mutation: A common cause of autosomal recessive ataxia with ancient European origin. *Am. J. Hum. Genet. 77*, 430–441.

Hämäläinen, R.H., Landoni, J.C., Ahlqvist, K.J., Goffart, S., Ryytty, S., Rahman, M.O., Brilhante, V., Icay, K., Hautaniemi, S., Wang, L., et al. (2019). Defects in mtDNA replication challenge nuclear genome stability through nucleotide depletion and provide a unifying mechanism for mouse progerias. *Nat. Metab. 1*, 958–965.

Hance, N., Ekstrand, M.I., and Trifunovic, A. (2005). Mitochondrial DNA polymerase gamma is essential for mammalian embryogenesis. *Hum. Mol. Genet. 14*, 1775–1783.

Harrower, T., Stewart, J.D., Hudson, G., Houlden, H., Warner, G., O'Donovan, D.G., Findlay, L.J., Taylor, R.W., De Silva, R., and Chinnery, P.F. (2008). POLG1 mutations manifesting as autosomal recessive axonal Charcot-Marie-Tooth disease. *Arch. Neurol. 65*, 133–136.

He, J., Mao, C.C., Reyes, A., Sembongi, H., Di Re, M., Granycome, C., Clippingdale, A.B., Fearnley, I.M., Harbour, M., Robinson, A.J., et al. (2007). The AAA+ protein ATAD3 has displacement loop binding properties and is involved in mitochondrial nucleoid organization. *J. Cell Biol. 176*, 141–146.

von Heijne, G. (1986). Why mitochondria need a genome. *FEBS Lett. 198*, 1–4.

Hiona, A., and Leeuwenburgh, C. (2008). The role of mitochondrial DNA mutations in aging and sarcopenia: Implications for the mitochondrial vicious cycle theory of aging. *Exp. Gerontol. 43*, 24–33.

Hiona, A., Sanz, A., Kujoth, G.C., Pamplona, R., Seo, A.Y., Hofer, T., Someya, S., Miyakawa, T., Nakayama, C., Samhan-Arias, A.K., et al. (2010). Mitochondrial

DNA Mutations Induce Mitochondrial Dysfunction, Apoptosis and Sarcopenia in Skeletal Muscle of Mitochondrial DNA Mutator Mice. *PLoS One 5*, e11468.

Hogeboom, G.H., Schneider, W.C., and Pallade, G.E. (1948). Cytochemical studies of mammalian tissues; isolation of intact mitochondria from rat liver; some biochemical properties of mitochondria and submicroscopic particulate material. *J. Biol. Chem. 172*, 619–635.

Horvath R., Kemp J.P., Tuppen H.A., et al. (2009) Molecular basis of infantile reversible cytochrome c oxidase deficiency myopathy. *Brain 132,* 3165-3174.

Den Hollander, A.I., Ghiani, M., De Kok, Y.J.M., Wijnholds, J., Ballabio, A., Cremers, F.P.M., and Broccoli, V. (2002). Isolation of Crb1, a mouse homologue of Drosophila crumbs, and analysis of its expression pattern in eye and brain. *Mech. Dev. 110*, 203–207.

Holt, I.J., He, J., Mao, C.C., Boyd-Kirkup, J.D., Martinsson, P., Sembongi, H., Reyes, A., and Spelbrink, J.N. (2007). Mammalian mitochondrial nucleoids: Organizing an independently minded genome. *Mitochondrion 7*, 311–321.

Isokallio, M.A., Stewart, J.B., Kauppila, J.H.K., Bonekamp, N.A., Mourier, A., Just, A., and Kauppila, T.E.S. (2018). Isolation of high-quality, highly enriched mitochondrial DNA from mouse tissues. *Nucleic Acid Res.* 1–11.

Ito, J., and Braithwaite, D.K. (1991). Compilation and alignment of DNA polymerase sequences. Nucleic Acids Res. *19*, 4045–4057.

Jang, Y. ha, and Lim, K. il (2018). Recent advances in mitochondria-targeted gene delivery. *Molecules 23,9.*

Jemt, E., Persson, Ö., Shi, Y., Mehmedovic, M., Uhler, J.P., López, M.D., Freyer, C., Gustafsson, C.M., Samuelsson, T., and Falkenberg, M. (2015). Regulation of DNA replication at the end of the mitochondrial D-loop involves the helicase TWINKLE and a conserved sequence element. *Nucleic Acids Res. 43*, 9262–9275.

Ji, J., Xu, M., Huang, Z., Li, L., Zheng, H., Yang, S., Li, S., Jin, L., Ling, X., Xia, Y., et al. (2017). Mitochondrial DNA sequencing and large-scale genotyping identifies

MT-ND4 gene mutation m.11696G>A associated with idiopathic oligoasthenospermia. *Oncotarget 8*, 52975–52982.

Jiang, M., Kauppila, T.E.S., Motori, E., Li, X., Atanassov, I., Folz-Donahue, K., Bonekamp, N.A., Albarran-Gutierrez, S., Stewart, J.B., and Larsson, N.G. (2017). Increased Total mtDNA Copy Number Cures Male Infertility Despite Unaltered mtDNA Mutation Load. *Cell Metab. 26*, 429-436.

Johnson, A.A., and Johnson, K.A. (2001). Exonuclease Proofreading by Human Mitochondrial DNA Polymerase. *J. Biol. Chem. 276*, 38097–38107.

Johnston, I.G., and Williams, B.P. (2016). Evolutionary inference across eukaryotes identifies specific pressures favoring mitochondrial gene retention. *Cell Syst. 2*, 101–111.

Jukes, T.H., and Osawa, S. (1990). The genetic code in mitochondria and chloroplasts. *Experientia 46*, 1117–1126.

Kaguni, L.S. (2004). DNA Polymerase γ, The Mitochondrial Replicase. *Annu. Rev. Biochem. 73*, 293–320.

Karnkowska, A., Vacek, V., Zubáčová, Z., Treitli, S.C., Petrželková, R., Eme, L., Novák, L., Žárský, V., Barlow, L.D., Herman, E.K., et al. (2016). A eukaryote without a mitochondrial organelle. *Curr. Biol. 26*, 1274–1284.

Kasahara, T., Kubota, M., Miyauchi, T., Noda, Y., Mouri, A., Nabeshima, T., and Kato, T. (2006). Mice with neuron-specific accumulation of mitochondrial DNA mutations show mood disorder-like phenotypes. *Mol. Psychiatry 11*, 577–593.

Kasiviswanathan, R., and Copeland, W.C. (2011). Biochemical analysis of the G517V POLG variant reveals wild-type like activity. *Mitochondrion 11*, 929–934.

Kauppila, J.H.K., Baines, H.L., Bratic, A., Simard, M.L., Freyer, C., Mourier, A., Stamp, C., Filograna, R., Larsson, N.G., Greaves, L.C., et al. (2016). A Phenotype-Driven Approach to Generate Mouse Models with Pathogenic mtDNA Mutations Causing Mitochondrial Disease. *Cell Rep. 16*, 2980–2990.

Kauppila, T.E.S., Kauppila, J.H.K., and Larsson, N.G. (2017). Mammalian Mitochondria

and Aging: An Update. *Cell Metab. 25*, 57–71.

Kiselycznyk, C., and Holmes, A. (2011). All (C57BL/6) mice are not created equal. *Front. Neurosci. 5*, 5–7.

Kõks, S., Dogan, S., Tuna, B.G., González-Navarro, H., Potter, P., and Vandenbroucke, R.E. (2016). Mouse models of ageing and their relevance to disease. *Mech. Ageing Dev. 160*, 41–53.

Kolesnikov, A.A., and Gerasimov, E.S. (2012). Diversity of mitochondrial genome organization. *Biochem. Biokhimii 77*, 1424–1435.

Korhonen, J.A., Pham, X.H., Pellegrini, M., and Falkenberg, M. (2004). Reconstitution of a minimal mtDNA replisome in vitro. *EMBO J. 23*, 2423–2429.

Korosec, P., Turner, P.J., Silar, M., Kopac, P., Kosnik, M., Gibbs, B.F., Shamji, M.H., Custovic, A., and Rijavec, M. (2017). Basophils, high-affinity IgE receptors, and CCL2 in human anaphylaxis. *J. Allergy Clin. Immunol. 140*, 750-758.e15.

Koyasu, S. (2003). The role of P13K in immune cells. *Nat. Immunol. 4*, 313–319.

Krämer, A., Green, J., Pollard, J., and Tugendreich, S. (2014). Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics 30*, 523–530.

Krasich, R., and Copeland, W.C. (2017). DNA polymerases in the mitochondria: A critical review of the evidence. *Front. Biosci. – Landmark 22*, 692–709.

Kraytsberg, Y., Kudryavtseva, E., McKee, A.C., Geula, C., Kowall, N.W., and Khrapko, K. (2006). Mitochondrial DNA deletions are abundant and cause functional impairment in aged human substantia nigra neurons. *Nat. Genet. 38*, 518–520.

Kraytsberg, Y., Simon, D.K., Turnbull, D.M., and Khrapko, K. (2009). Do mtDNA deletions drive premature aging in mtDNA mutator mice? *Aging Cell 8*, 502–506.

Kujoth, G.C. (2007). Mitochondrial DNA Mutations, Oxidative Stress, and Apoptosis in Mammalian Aging. *Science 481*, 481–484.

Kujoth, G.C., Hiona, a, Pugh, T.D., Someya, S., Panzer, K., Wohlgemuth, S.E., Hofer, T., Seo, a Y., Sullivan, R., Jobling, W. a, *et al.* (2005). Mitochondrial DNA

mutations, oxidative stress, and apoptosis in mammalian aging. *Science 309*, 481–484.

Kukat, C., Wurm, C.A., Spåhr, H., Falkenberg, M., Larsson, N.G., and Jakobs, S. (2011). Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA. *Proc. Natl. Acad. Sci. U. S. A. 108*, 13534–13539.

Kukat, C., Davies, K.M., Wurm, C.A., Spåhr, H., Bonekamp, N.A., Kühl, I., Joos, F., Polosa, P.L., Park, C.B., Posse, V., et al. (2015). Cross-strand binding of TFAM to a single mtDNA molecule forms the mitochondrial nucleoid. *Proc. Natl. Acad. Sci. U. S. A. 112*, 11288–11293.

Lee, H.R., and Johnson, K.A. (2006). Fidelity of the human mitochondrial DNA polymerase. *J. Biol. Chem. 281*, 36236–36240.

Lee, Y.S., Kennedy, W.D., and Yin, Y.W. (2009). Structural Insight into Processive Human Mitochondrial DNA Synthesis and Disease-Related Polymerase Mutations. *Cell 139*, 312–324.

Lewis, M.R., and Lewis, W.H. (1915). Mitochondria (and other cytoplasmic structures) in tissue cultures. *Am. J. Anat. 17*, 339–401.

Lewis, W., Day, B.J., Kohler, J.J., Hosseini, S.H., Chan, S.S.L., Green, E.C., Haase, C.P., Keebaugh, E.S., Long, R., Ludaway, T., et al. (2007). Decreased mtDNA, oxidative stress, cardiomyopathy, and death from transgenic cardiac targeted human mutant polymerase γ. *Lab. Investig. 87*, 326–335.

Liesa, M., Palacín, M., and Zorzano, A. (2009). Mitochondrial dynamics in mammalian health and disease. *Physiol. Rev. 89*, 799–845.

Lim, S.E., Longley, M.J., and Copeland, W.C. (1999). The mitochondrial p55 accessory subunit of human DNA polymerase γ enhances DNA binding, promotes processive DNA synthesis, and confers N- ethylmaleimide resistance. *J. Biol. Chem. 274*, 38197–38203.

Liou, C.W., Lin, T.K., Chen, J.B., Tiao, M.M., Weng, S.W., Chen, S.D., Chuang, Y.C.,

Chuang, J.H., and Wang, P.W. (2010). Association between a common mitochondrial DNA D-loop polycytosine variant and alteration of mitochondrial copy number in human peripheral blood cells. *J. Med. Genet. 47*, 723–728.

López-Otín, C., Blasco, M.A., Partridge, L., Serrano, M., and Kroemer, G. (2013). The hallmarks of aging. *Cell 153*, 1194.

Lotery, A.J., Jacobson, S.G., Fishman, G.A., Weleber, R.G., Fulton, A.B., Namperumalsamy, P., Héon, E., Levin, A. V., Grover, S., Rosenow, J.R., et al. (2001). Mutations in the CRB1 gene cause Leber congenital amaurosis. *Arch. Ophthalmol. 119*, 415–420.

Luhmann, U.F.O., Carvalho, L.S., Holthaus, S.M. kleine, Cowing, J.A., Greenaway, S., Chu, C.J., Herrmann, P., Smith, A.J., Munro, P.M.G., Potter, P., et al. (2015). The severity of retinal pathology in homozygous Crb1rd8/rd8 mice is dependent on additional genetic factors. *Hum. Mol. Genet. 24*, 128–141.

Luoma, P.T., Eerola, J., Ahola, S., Hakonen, A.H., Hellström, O., Kivistö, K.T., Tienari, P.J., and Suomalainen, A. (2007). Mitochondrial DNA polymerase gamma variants in idiopathic sporadic Parkinson disease. *Neurology 69*, 1152–1159.

Lynch, M. (2011). The lower bound to the evolution of mutation rates. *Genome Biol. Evol. 3*, 1107–1118.

Macao, B., Uhler, J.P., Siibak, T., Zhu, X., Shi, Y., Sheng, W., Olsson, M., Stewart, J.B., Gustafsson, C.M., and Falkenberg, M. (2015). The exonuclease activity of DNA polymerase γ is required for ligation during mitochondrial DNA replication. *Nat. Commun. 6*, 1–10.

Martijn, J., Vosseberg, J., Guy, L., Offre, P., and Ettema, T.J.G. (2018). Deep mitochondrial origin outside the sampled alphaproteobacteria. *Nature 557*, 101–105.

Martikainen, M.H., Grady, J.P., Ng, Y.S., Alston, C.L., Gorman, G.S., Taylor, R.W., McFarland, R., and Turnbull, D.M. (2017). Decreased male reproductive success in association with mitochondrial dysfunction. *Eur. J. Hum. Genet. 25*, 1162–1164.

Martin, W. (2003). Gene transfer from organelles to the nucleus: Frequent and in big

chunks. *Proc. Natl. Acad. Sci. U. S. A. 100*, 8612–8614.

Martin, W., Hoffmeister, M., Rotte, C., and Henze, K. (2001). An overview of endosymbiotic models for the origins of eukaryotes, their ATP-producing organelles (mitochondria and hydrogenosomes), and their heterotrophic lifestyle. *Biol. Chem. 382*, 1521–1539.

Matsuo, N., Takao, K., Nakanishi, K., Yamasaki, N., Tanda, K., and Miyakawa, T. (2010). Behavioral profiles of three C57BL/6 substrains. *Front. Behav. Neurosci. 4*, 1–12.

Mattapallil, M.J., Wawrousek, E.F., Chan, C.C., Zhao, H., Roychoudhury, J., Ferguson, T.A., and Caspi, R.R. (2012). The Rd8 mutation of the Crb1 gene is present in vendor lines of C57BL/6N mice and embryonic stem cells, and confounds ocular induced mutant phenotypes. *Invest. Ophthalmol. Vis. Sci. 53*, 2921–2927.

Mayr, C. (2017). Regulation by 3 -Untranslated Regions. *Annual Review of Genetics* 51, 171-194

McCarron, J.G., Wilson, C., Sandison, M.E., Olson, M.L., Girkin, J.M., Saunter, C., and Chalmers, S. (2013). From Structure to Function: Mitochondrial Morphology, Motion and Shaping in Vascular Smooth Muscle. *J. Vasc. Res. 50*, 357–371.

McGuire, P.J. (2019). Mitochondrial dysfunction and the aging immune system. *Biology (Basel). 8(2), 26*

Meller, S., Di Domizio, J., Voo, K.S., Friedrich, H.C., Chamilos, G., Ganguly, D., Conrad, C., Gregorio, J., Le Roy, D., Roger, T., et al. (2015). TH17 cells promote microbial killing and innate immune sensing of DNA via interleukin 26. *Nat. Immunol. 16*, 970–979.

Miller, F.J. (2003). Precise determination of mitochondrial DNA copy number in human skeletal and cardiac muscle by a PCR-based assay: lack of change of copy number with age. *Nucleic Acids Res. 31*, 61e – 61.

Mito, T., Kikkawa, Y., Shimizu, A., Hashizume, O., Katada, S., Imanishi, H., Ota, A., Kato, Y., Nakada, K., and Hayashi, J.I. (2013). Mitochondrial DNA Mutations in Mutator Mice Confer Respiration Defects and B-Cell Lymphoma Development.

PLoS One 8.

Montgomery, M.K. (2019). Mitochondrial dysfunction and diabetes: Is mitochondrial transfer a friend or foe?. *Biology (Basel). 8*(2), 33.

Morshed, M., Hlushchuk, R., Simon, D., Walls, A.F., Obata-Ninomiya, K., Karasuyama, H., Djonov, V., Eggel, A., Kaufmann, T., Simon, H.-U., et al. (2014). NADPH Oxidase–Independent Formation of Extracellular DNA Traps by Basophils. *J. Immunol. 192*, 5314–5323.

Mortaz, E., Alipoor, S.D., Adcock, I.M., Mumby, S., and Koenderman, L. (2018). Update on neutrophil function in severe inflammation. *Front. Immunol. 9*, 2171

Mourier, T., Hansen, A.J., Willerslev, E., and Arctander, P. (2001). The human genome project reveals a continuous transfer of large mitochondrial fragments to the nucleus. *Mol. Biol. Evol. 18*, 1833–1837.

Muller, M., Mentel, M., van Hellemond, J.J., Henze, K., Woehle, C., Gould, S.B., Yu, R.-Y., van der Giezen, M., Tielens, A.G.M., and Martin, W.F. (2012). Biochemistry and Evolution of Anaerobic Energy Metabolism in Eukaryotes. *Microbiol. Mol. Biol. Rev. 76*, 444–495.

Murakami, E., Feng, J.Y., Lee, H., Hanes, J., Johnson, K.A., and Anderson, K.S. (2003). Characterization of novel reverse transcriptase and other RNA-associated catalytic activities by human DNA polymerase γ: Importance in mitochondrial DNA replication. *J. Biol. Chem. 278*, 36403–36409.

Naess, K., Freyer, C., Bruhn, H., Wibom, R., Malm, G., Nennesmo, I., von Döbeln, U., and Larsson, N.G. (2009). MtDNA mutations are a common cause of severe disease phenotypes in children with Leigh syndrome. *Biochim. Biophys. Acta - Bioenerg. 1787*, 484–490.

Nakada, K., Sato, A., Yoshida, K., Morita, T., Tanaka, H., Inoue, S., Yonekawa, H., and Hayashi, J. (2006). Mitochondria-related male infertility. *PNAS 103*, 1–6.

Nass, S., and Nass, M.M. (1963). Intramitochondrial Fibers With Dna Characteristics. *Ii. Enzymatic and. J. Cell Biol. 19*, 613–629.

Nathan, D., and Orkin, S. (2009). The phagocyte system and disorders of granulopoiesis and granulocyte and function. In Hematology of Infancy and Childhood, (*Saunders Elsevier*), pp. 1109–1220.

Ni, H.M., Williams, J.A., and Ding, W.X. (2015). Mitochondrial dynamics and mitochondrial quality control. *Redox Biol. 4*, 6–13.

Nicholls, T.J., and Minczuk, M. (2014). In D-loop: 40 years of mitochondrial 7S DNA. *Exp. Gerontol. 56*, 175–181.

Nicholls, T.J., Nadalutti, C.A., Motori, E., Sommerville, E.W., Gorman, G.S., Basu, S., Hoberg, E., Turnbull, D.M., Chinnery, P.F., Larsson, N.G., et al. (2018). Topoisomerase 3α Is Required for Decatenation and Segregation of Human mtDNA. *Mol. Cell 69*, 9-23.e6.

Nissanka, N., Bacman, S.R., Plastini, M.J., and Moraes, C.T. (2018). The mitochondrial DNA polymerase gamma degrades linear DNA fragments precluding the formation of deletions. *Nat. Commun. 9* (1), 2491

Norddahl, G.L., Pronk, C.J., Wahlestedt, M., Sten, G., Nygren, J.M., Ugale, A., Sigvardsson, M., and Bryder, D. (2011). Accumulating mitochondrial DNA mutations drive premature hematopoietic aging phenotypes distinct from physiological stem cell aging. *Cell Stem Cell 8*, 499–510.

Nowak, R., Zub, R., Skoneeczna, I., Sikora, K., and Ligaj, M. (2005). CAG repeat polymorphism in the DNA polymerase γ gene in a Polish population: An association with testicular cancer risk. *Ann. Oncol. 16*, 1211–1212.

Nurminen, A., Farnum, G.A., and Kaguni, L.S. (2017). Pathogenicity in POLG syndromes: DNA polymerase gamma pathogenicity prediction server and database. *BBA Clin. 7*, 147–156.

Nussey, D.H., Froy, H., Lemaitre, J.F., Gaillard, J.M., and Austad, S.N. (2013). Senescence in natural populations of animals: Widespread evidence and its implications for bio-gerontology. *Ageing Res. Rev. 12*, 214–225.

Orogo, A.M., Gonzalez, E.R., Kubli, D.A., Baptista, I.L., Ong, S.B., Prolla, T.A.,

Sussman, M.A., Murphy, A.N., and Gustafsson, Å.B. (2015). Accumulation of mitochondrial DNA mutations disrupts cardiac progenitor cell function and reduces survival. *J. Biol. Chem. 290*, 22061–22075.

Osellame, L.D., Blacker, T.S., and Duchen, M.R. (2012). Cellular and molecular mechanisms of mitochondrial function. *Best Pract. Res. Clin. Endocrinol. Metab. 26*, 711–723.

Palade, G. (1952). The fine structure of mitochondria. *Anat. Rec. 144*, 427–451.

Payne, B.A.I., and Chinnery, P.F. (2015). Mitochondrial dysfunction in aging: Much progress but many unresolved questions. *Biochim. Biophys. Acta - Bioenerg. 1847*.

Peeva, V., Blei, D., Trombly, G., Corsi, S., Szukszto, M.J., Rebelo-Guiomar, P., Gammage, P.A., Kudin, A.P., Becker, C., Altmüller, J., et al. (2018). Linear mitochondrial DNA is rapidly degraded by components of the replication machinery. *Nat. Commun. 9*, 1727

Pfannschmidt, T., Nilsson, A., and Allen, J.F. (1999). Photosynthetic control of chloroplast gene expression. *Nature 397*, 625–628.

Pivkin, I. V., Peng, Z., Karniadakis, G.E., Buffet, P.A., Dao, M., and Suresh, S. (2016). Biomechanics of red blood cells in human spleen and consequences for physiology and disease. *Proc. Natl. Acad. Sci. U. S. A. 113*, 7804–7809.

Ponamarev, M. V., Longley, M.J., Nguyen, D., Kunkel, T.A., and Copeland, W.C. (2002). Active site mutation in DNA polymerase γ associated with progressive external ophthalmoplegia causes error-prone DNA synthesis. *J. Biol. Chem. 277*, 15225–15228.

Popot, J.L., and Vitry, C. (1990). On the Microassembly of Integral Membrane Proteins. *Annu. Rev. Biophys. Biophys. Chem. 19*, 369–403.

Race, H.L., Herrmann, R.G., and Martin, W. (1999). Why have organelles retained genomes? *Trends Genet. 15*, 364–370.

Rak, M., Bénit, P., Chrétien, D., Bouchereau, J., Schiff, M., El-Khoury, R., Tzagoloff, A., & Rustin, P. (2016). Mitochondrial cytochrome c oxidase deficiency. *Clinical science*

130(6), 393–407

Rambold, A.S., and Pearce, E.L. (2018). Mitochondrial Dynamics at the Interface of Immune *Cell Metabolism and Function. Trends Immunol. 39*, 6–18.

Richards, C.D. (2013). The Enigmatic Cytokine Oncostatin M and Roles in Disease. *ISRN Inflamm. 2013*, 1–23.

Rimel, J.K., and Taatjes, D.J. (2018). The essential and multifunctional TFIIH complex. *Protein Sci. 27*, 1018–1037.

Roger, A.J., Muñoz-Gómez, S.A., and Kamikawa, R. (2017). The Origin and Diversification of Mitochondria. *Curr. Biol. 27*, R1177–R1192.

Ross, J.M., Stewart, J.B., Hagström, E., Brené, S., Mourier, A., Coppotelli, G., Freyer, C., Lagouge, M., Hoffer, B.J., Olson, L., et al. (2013). Germline mitochondrial DNA mutations aggravate ageing and can impair brain development. *Nature 501*, 412–415.

Ross, J.M., Coppotelli, G., Hoffer, B.J., and Olson, L. (2014). Maternally transmitted mitochondrial DNA mutations can reduce lifespan. *Sci. Rep. 4*, 1–3.

Rovio, A.T., Marchington, D.R., Donat, S., Schuppe, H.C., Abel, J., Fritsche, E., Elliott, D.J., Laippala, P., Ahola, A.L., McNay, D., *et al.* (2001). Mutations at the mitochondrial DNA polymerase (POLG) locus associated with male infertility. *Nat. Genet. 29*, 261–262.

Safdar, A., Bourgeois, J.M., Ogborn, D.I., Little, J.P., Hettinga, B.P., Akhtar, M., Thompson, J.E., Melov, S., Mocellin, N.J., Kujoth, G.C., *et al.* (2011). Endurance exercise rescues progeroid aging and induces systemic mitochondrial rejuvenation in mtDNA mutator mice. *Proc. Natl. Acad. Sci. U. S. A. 108*, 4135–4140.

Schatz, G., Haslbrunner, E., and Tuppy, H. (1964). Deoxyribonucleic acid associated with yeast mitochondria. *Biochem. Biophys. Res. Commun. 15*, 127–132.

Shadel, G.S., and Clayton, D.A. (1997). MITOCHONDRIAL DNA MAINTENANCE IN VERTEBRATES. *Annu. Rev. Biochem. 66*, 409–435.

Shadel, G.S., and Horvath, T.L. (2015). Mitochondrial ROS Signaling in Organismal Homeostasis. *Cell 163*, 560–569.

Shin, M.G., Kajigaya, S., McCoy, J.P., Levin, B.C., and Young, N.S. (2004). Marked mitochondrial DNA sequence heterogeneity in single CD34+ cell clones from normal adult bone marrow. *Blood 103*, 553–561.

Showalter, A.K., and Tsai, M.D. (2002). A reexamination of the nucleotide incorporation fidelity of DNA polymerases. *Biochemistry 41*, 10571–10576.

Sibley, C.R., Blazquez, L., and Ule, J. (2016). Lessons from non-canonical splicing. *Nat. Rev. Genet. 17*, 407–421.

Siekevitz, P. (1957). Powerhouse of the cell. *Sci. Am. 197*, 131–140.

Simard, M.L., Mourier, A., Greaves, L.C., Taylor, R.W., and Stewart, J.B. (2018). A novel histochemistry assay to assess and quantify focal cytochrome c oxidase deficiency. J. *Pathol. 245*, 311–323.

Simon, M.M., Greenaway, S., White, J.K., Fuchs, H., Gailus-Durner, V., Wells, S., Sorg, T., Wong, K., Bedu, E., Cartwright, E.J., et al. (2013). A comparative phenotypic and genomic analysis of C57BL/6J and C57BL/6N mouse strains. *Genome Biol. 14*.

Sokol, C.L., and Medzhitov, R. (2010). Emerging functions of basophils in protective and allergic immune responses. *Mucosal Immunol. 3*, 129–137.

Someya, S., Kujoth, G.C., Kim, M.J., Hacker, T.A., Vermulst, M., Weindruch, R., and Prolla, T.A. (2017). Effects of calorie restriction on the lifespan and healthspan of POLG mitochondrial mutator mice. *PLoS One 12*, 4–5.

Spinelli, J.B., and Haigis, M.C. (2018). The multifaceted contributions of mitochondria to cellular metabolism. *Nat. Cell Biol. 20*, 745–754.

Spiropoulos, J. (2002). Can mitochondrial DNA mutations cause sperm dysfunction? Mol. Hum. Reprod. *8*, 719–721.

Stewart, J.B., and Chinnery, P.F. (2015). The dynamics of mitochondrial DNA heteroplasmy: Implications for human health and disease. *Nat. Rev. Genet. 16*.

Stewart, J.B., and Larsson, N.-G. (2014). Keeping mtDNA in Shape between Generations. *PLoS Genet. 10*, e1004670.

Stewart, J.B., Freyer, C., Elson, J.L., and Larsson, N.-G. (2008). Purifying selection of mtDNA and its implications for understanding evolution and mitochondrial disease. *Nat Rev Genet 9*, 657–662.

Stiedl, O., Radulovic, J., Lohmann, R., Birkenfeld, K., Palve, M., Kammermeier, J., Sananbenesi, F., and Spiess, J. (1999). Strain and substrain differences in context- and tone-dependent fear conditioning of inbred mice. *Behav. Brain Res. 104*, 1–12.

Stout, R., and Birch-Machin, M. (2019). Mitochondria's role in skin ageing. *Biology (Basel). 8,13-22*.

Stumpf, J.D., Bailey, C.M., Spell, D., Stillwagon, M., Anderson, K.S., and Copeland, W.C. (2010). Mip1 containing mutations associated with mitochondrial disease causes mutagenesis and depletion of mt DNA in Saccharomyces cerevisiae. *Hum. Mol. Genet. 19*, 2123–2133.

Stumpf, J.D., Saneto, R.P., and Copeland, W.C. (2013). Clinical and molecular features of POLG-related mitochondrial disease. *Cold Spring Harb. Perspect. Biol. 5*.

Szczepanowska, K., and Foury, F. (2010). A cluster of pathogenic mutations in the 3′-5′ exonuclease domain of DNA polymerase gamma defines a novel module coupling DNA synthesis and degradation. *Hum. Mol. Genet. 19*, 3516–3529.

Taanman, J.W., Rahman, S., Pagnamenta, A.T., Morris, A.A.M., Bitner-Glindzicz, M., Wolf, N.I., Leonard, J. V., Clayton, P.T., and Schapira, A.H.V. (2009). Analysis of mutant DNA polymerase γ in patients with mitochondrial DNA depletion. *Hum. Mutat. 30*, 248–254.

Tait, S.W.G., and Green, D.R. (2013). Mitochondrial regulation of cell death. *Cold Spring Harb. Perspect. Biol. 5*.

Tang, S., Dimberg, E.L., Milone, M., and Wong, L.J.C. (2012). Mitochondrial neurogastrointestinal encephalomyopathy (MNGIE)-like phenotype: An expanded clinical spectrum of POLG1 mutations. *J. Neurol. 259*, 862–868.

Toye, A.A., Lippiat, J.D., Proks, P., Shimomura, K., Bentley, L., Hugill, A., Mijat, V., Goldsworthy, M., Moir, L., Haynes, A., et al. (2005). A genetic and physiological study of impaired glucose homeostasis control in C57BL/6J mice. *Diabetologia 48*, 675–686.

Trifunovic, A., Wredenberg, A., Falkenberg, M., Spelbrink, J.N., Rovio, A.T., Bruder, C.E., Bohlooly-Y, M., Gldlöf, S., Oldfors, A., Wibom, R., et al. (2004). Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature 429*, 417–423.

Trifunovic, A., Hansson, A., Wredenberg, A., Rovio, A.T., Dufour, E., Khvorostov, I., Spelbrink, J.N., Wibom, R., Jacobs, H.T., and Larsson, N.-G. (2005). Somatic mtDNA mutations cause aging phenotypes without affecting reactive oxygen species production. *Proc. Natl. Acad. Sci. U. S. A. 102*, 17993–17998.

Tuppen, H.A.L., Blakely, E.L., Turnbull, D.M., and Taylor, R.W. (2010). Mitochondrial DNA mutations and human disease. *Biochim. Biophys. Acta – Bioenerg. 1797*, 113–128.

Uhler, J.P., and Falkenberg, M. (2015). Primer removal during mammalian mitochondrial DNA replication. *DNA Repair (Amst). 34*, 28–38.

Uhler, J.P., Thörn, C., Nicholls, T.J., Matic, S., Milenkovic, D., Gustafsson, C.M., and Falkenberg, M. (2016). MGME1 processes flaps into ligatable nicks in concert with DNA polymerase γ during mtDNA replication. *Nucleic Acids Res. 44*, 5861–5871.

Vanderstraeten, S., Van Den Brûle, S., Hu, J., and Foury, F. (1998). The role of 3'-5' exonucleolytic proofreading and mismatch repair in yeast mitochondrial DNA error avoidance. *J. Biol. Chem. 273*, 23690–23697.

Vaught, R.C., and Dowling, D.K. (2018). Maternal inheritance of mitochondria: Implications for male fertility? *Reproduction 155*, R159–R168.

Vermulst, M., Wanagat, J., Kujoth, G.C., Bielas, J.H., Rabinovitch, P.S., Prolla, T.A., and Loeb, L.A. (2008). DNA deletions and clonal mutations drive premature aging in mitochondrial mutator mice. *Nat. Genet. 40*, 392–394.

Wahlestedt, M., Ameur, A., Moraghebi, R., Norddahl, G.L., Sten, G., Woods, N.B., and Bryder, D. (2014). Somatic cells with a heavy mitochondrial DNA mutational load render induced pluripotent stem cells with distinct differentiation defects. *Stem Cells 32*, 1173–1182.

Wanrooij, S., Miralles Fusté, J., Stewart, J.B., Wanrooij, P.H., Samuelsson, T., Larsson, N.G., Gustafsson, C.M., and Falkenberg, M. (2012). *In vivo* mutagenesis reveals that OriL is essential for mitochondrial DNA replication. *EMBO Rep. 13*, 1130–1137.

Weidling, I., and Swerdlow, R.H. (2019). Mitochondrial dysfunction and stress responses in Alzheimer's disease. *Biology (Basel). 8(2)*.

West, A.P., Khoury-Hanold, W., Staron, M., Tal, M.C., Pineda, C.M., Lang, S.M., Bestwick, M., Duguay, B.A., Raimundo, N., MacDuff, D.A., et al. (2015). Mitochondrial DNA stress primes the antiviral innate immune response. *Nature 520*, 553–557.

Westermann, B. (2010). Mitochondrial fusion and fission in cell life and death. *Nat. Rev. Mol. Cell Biol. 11*, 872–884.

Wiesner, R.J., Rüegg, J.C., and Morano, I. (1992). Counting target molecules by exponential polymerase chain reaction: copy number of mitochondrial DNA in rat tissues. *Biochem. Biophys. Res. Commun. 183*, 553–559.

Williams, S.L., Huang, J., Edwards, Y.J.K., Ulloa, R.H., Dillon, L.M., Prolla, T.A., Vance, J.M., Moraes, C.T., and Züchner, S. (2010). The mtDNA mutation spectrum of the progeroid polg mutator mouse includes abundant control region multimers. *Cell Metab. 12*, 675–682.

Wilson, S.B., Davidson, G.S., Thomson, L.M., and Pearson, C.K. (1996). Redox control of RNA synthesis in potato mitochondria. *Eur. J. Biochem. 242*, 81–85.

Winterthun, S., Ferrari, G., He, L., Taylor, R.W., Zeviani, M., Turnbull, D.M., Engelsen, B.A., Moen, G., and Bindoff, L.A. (2005). Autosomal recessive mitochondrial ataxic syndrome due to mitochondrial polymerase γ mutations. *Neurology 64*, 1204–1208.

Wohlrab, H. (2009). Transport proteins (carriers) of mitochondria. *IUBMB Life 61*, 40–46.

Wong, L.-J.C., Naviaux, R.K., Brunetti-Pierri, N., Zhang, Q., Schmitt, E.S., Truong, C., Milone, M., Cohen, B.H., Wical, B., Ganesh, J., et al. (2008). Molecular and clinical genetics of mitochondrial diseases due to POLG mutations. *Hum. Mutat. 29*, E150–E172.

Woodall, B.P., Orogo, A.M., Najor, R.H., Cortez, M.Q., Moreno, E.R., Wang, H., Divakaruni, A.S., Murphy, A.N., and Gustafsson, Å.B. (2019). Parkin does not prevent accelerated cardiac aging in mitochondrial DNA mutator mice. *JCI Insight 5(10), 127713*.

Wyrobek, A.J., and Bruce, W.R. (1975). Chemical induction of sperm abnormalities in mice. *Proc. Natl. Acad. Sci. U. S. A. 72*, 4425–4429.

Xia, S., Zhang, X., Zheng, S., Khanabdali, R., Kalionis, B., Wu, J., Wan, W., and Tai, X. (2016). An Update on Inflamm-Aging: Mechanisms, Prevention, and Treatment. *J. Immunol. Res. 2016*.

Yamashita, A., Fujimoto, M., Katayama, K., Tsutsumi, N., and Arimura, S. (2016). Mitochondrial outer membrane forms bridge between two mitochondria in *Arabidopsis thaliana. Plant Signal. Behav. 11*, e1167301.

Yasukawa, T., Reyes, A., Cluett, T.J., Yang, M.Y., Bowmaker, M., Jacobs, H.T., and Holt, I.J. (2006). Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand. *EMBO J. 25*, 5358–5371.

Yim, Y., A., A., Koti, K., P., P., Bonnard, B., A., A., Duerrbaum, D., M., M., Mueller, M., C., C., et al. (2019). mitoXplorer, a visual data mining platform to systematically analyze and visualize mitochondrial expression dynamics and mutations. *BioRxiv Bioinforma*. 1–72.

Yousefi, S., Mihalache, C., Kozlowski, E., Schmid, I., and Simon, H.U. (2009). Viable neutrophils release mitochondrial DNA to form neutrophil extracellular traps. *Cell Death Differ. 16*, 1438–1444.

Yousefi, S., Morshed, M., Amini, P., Stojkov, D., Simon, D., Von Gunten, S., Kaufmann, T., and Simon, H.U. (2015). Basophils exhibit antibacterial activity through extracellular trap formation. *Allergy Eur. J. Allergy Clin. Immunol. 70*, 1184–1188.

Zhang, D., Mott, J.L., Chang, S.W., Denniger, G., Feng, Z., and Zassenhaus, H.P. (2000). Construction of transgenic mice with tissue-specific acceleration of mitochondrial DNA mutagenesis. *Genomics 69*, 151–161.

Zhang, Q., Raoof, M., Chen, Y., Sumi, Y., Sursal, T., Junger, W., Brohi, K., Itagaki, K., and Hauser, C.J. (2010). Circulating mitochondrial DAMPs cause inflammatory responses to injury. *Nature 464*, 104–107.

9. SUPPLEMENTARY LISTS

9.1. List of Abbreviations

ADP	Adenosine diphosphate
AHS	Alpers-Huttenlocher syndrome
AMP	Adenosine monophosphate
AMPK	AMP-activated protein kinase
ANOVA	Analysis of variance
ANS	Ataxia neuropathy spectrum
ATP	Adenosine triphosphate
B2M	Beta-2-microglobulin
B6J	C57BL/6J
B6N	C57BL/6NCrl
BN-PAGE	Blue-native polyacrylamide gel electrophoresis
Ca2+	Calcium ion
CaMKIIα	Ca^{2+}/calmodulin-activated protein kinase II α
CASA	Computer-aided sperm analysis
cDNA	Complementary DNA
CI	OXPHOS complex I - NADH:Q oxidoreductase
CII	CII OXPHOS complex II - succinate dehydrogenase
CIII	CIII OXPHOS complex III - ubiquinol:cytochrome c oxidoreductase
CIV	CIV OXPHOS complex IV - cytochrome c oxidase
CoA	Co-enzyme A
CoRR	Co-location for Redox Regulation
COX	Cytochrome c oxidase
Crb1	Crumbs homolog 1
CRMs	Control region multimers
CSBQ	Conserved sequence blocks
CsCl	Cesium chloride
CV	OXPHOS complex V - ATP synthase
CytB	Cytochrome b
CytC	Cytochrome c
DAB	3-3' diaminobenzidine

D-loop	Displacement loop
DNA	Deoxyribonucleic acid
DNA-Seq	DNA sequencing
dNTP	Deoxyribonucleotide
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum
ERCC2	Excision repair cross-complementing rodent repair deficiency, complementation group 2
ETC	Electron transport chain
FAD	Flavin adenine dinucleotide
FADH	Reduced FAD
HSP	Heavy strand promoter
IL	Interleukin
IM	Inner membrane
IMS	Intermembrane space
IPA	Ingenuity Pathway Analysis software
iPSC	Induced pluripotent stem cells
KO	Knockout
LoxP	Locus of crossover x in P1
MCHS	Childhood myocerebrohepatopathyspectrum
MEMSA	Myoclonic epilepsymyopathy sensory ataxia
MgCl2	Magnesium chloride
MGME1	MGME1 Mitochondrial genome maintenance exonuclease - 1
MOPS	3-(N-morpholino) propanesulfonic acid
mPTP	Mitochondrial permeability transition pore
mRNA	Messenger RNA
MROs	Mitochondria-related organelles
mtDNA	Mitochondrial DNA
mtSSB	Single-stranded DNA binding protein
NAD	Nicotinamide adenine dinucleotide
NADH	Reduced NAD
NaOH	Sodium hydroxide
NBT	Nitro blue tratrazolium
NBTx	Nitrotetrazolium blue exclusion assay
NCR	Non-coding region

ND4	NADH:Ubiquinone Oxidoreductase Core Subunit 4
NER	Nucleotide excision repair
NETs	Neutrophil extracelullular traps
NNT	Nicotinamide nucleotide transhydrogenase
O_H	Origin of replication of the heavy strand
O_L	Origin of replication of the light strand
OM	Outer membrane
OXPHOS	Oxidative phosphorylation
PAM1	The entire mtDNA cloned into pACYC177-vector backbone
PEO	Progressive External Ophthalmoplegia
PI3K	Phosphoinositide 3-kinase
PMS	Phenazine methosulphate
Pol γ	polymerase gamma (holoenzyme)
POLB	Polymerase gamma (accesory subunit)
POLG	Polymerase gamma (catalitic subunit)
POLRMT	Mitochondrial RNA polymerase
PVDF	Polyvinylidene difluoride
qPCR	Quantitative real-time PCR
RITOLS	Ribonucleotide incorporation throughout the lagging strand
RNA	Ribonucleic acid
RNA-Seq	RNA sequencing
ROS	Reactive oxygen species
rRNA	Ribosomal RNA
RT-qPCR	Reverse transcription quantitative real-time PCR
SDH	Succinate dehydrogenase
SDM	Strand displacement model
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SNP	Single Nucleotide Polymorphism
TAS	Termination-associated sequences
TAS	TAS Termination associated sequences
TBP	TATA binding protein
TBS	Tris Buffered Saline
TCA	Tricarboxylic acid
TFAM	Mitochondrial transcription factor A
TFIIH	Transcription factor IIH
TH	T helper

TLR9	Toll-like receptor 9
TOM20	Translocase of outer mitochondrial membrane 20
TOP3α	Topoisomerase 3α
tRNA	Transfer RNA
UTR	Untranslated region
VDAC	Voltage-dependent anion-selective channel
WHO	World Health Organization
WT	Wild-type
α-MHC	α- Myosin heavy chain

Abbreviations of amino acids

A	Alanine
R	Arginine
N	Asparagine
D	Aspartic acid
C	Cysteine
E	Glutamic acid
Q	Glutamine
G	Glycine
H	Histidine
I	Isoleucine
L	Leucine
K	Lysine
M	Methionine
F	Phenylalanine
P	Proline
S	Serine
T	Threonine
W	Tryptophan
Y	Tyrosine
V	Valine

9.2. List of Figures

9.3. List of Supplementary figures

9.4. List of Tables

10. ERKLÄRUNG

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzen Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorlegen hat, dass sie – abgesehen von unten angegebenen Teilpublikationen – noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluß des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Aleksandra Trifunovic betreut worden.

Köln, 19.02.2020

_____

Sara Ruth Albarran Gutierrez

11. CURRICULUM VITAE

ALBARRAN GUTIERREZ, SARA RUTH

Personal Data

Current Residence: Cologne Germany
Nationality: Mexican
Email: Sara.AlbarranG@icloud.com
Linkedln: https://www.linkedin.com/in/sara-albarrang/

Education

Ph.D. Candidate at the Max Planck Institute for Biology of Ageing
(2015 - 2020)
Member of the Graduate School for Biological Sciences, University of Cologne.
Thesis: Molecular mechanisms underlying the rate of progeria onset in mtDNA mutator mouse strains.

Master of Science
(2012 - 2014)
Master in Biochemistry at the National Autonomous University of Mexico
Thesis: Study the protein p32 in maintaining the malignant phenotype of colon cancer.
Final grade: 9.27 of 10 (Excellent)

Bachelor in Biology (2007 - 2012)
Biology at the National Autonomous University of Mexico
Thesis: Biscistronic vector construction to measure calcium.
Final grade: 8.75 of 10 (Good)

Scientific collaborations

Increased total mtDNA copy number cures male infertility despite unaltered mtDNA mutation load. 2017. Jiang M., Kauppila T.E.S., Motori E., Li X., Atanassov I., Folz-Donahue K., Bonekamp N., **Albarran-Gutierrez S.**, Stewart J.B. Larsson N.G. *Cell Metabolism*. 26: 429-436.e4
https://www.ncbi.nlm.nih.gov/pubmed/28768180

Languages

Spanish: native speaker, **English:** fluent, **German:** working knowledge.