From DEPARTMENT OF LABORATORY MEDICINE
Karolinska Institutet, Stockholm, Sweden

GENERATION AND TRANSMISSION OF mtDNA MUTATIONS AND THEIR EFFECT ON AGING

Erik Hagström

Stockholm 2014

Published by Karolinska Institutet. Printed by Taberg Media Group

ISBN 978-91-7549-422-7

ABSTRACT

Mutations in mtDNA are known to cause neuromuscular diseases and have been associated with several common age-associated diseases as well as being a contributor to the aging process. Despite their importance there are still important aspects of mtDNA inheritance and mtDNA mutations that remain unknown and need to be addressed. The aim of this thesis is to examine how mtDNA mutations behave in in vivo systems with regard to inheritance and distribution. A further aim is to investigate the phenotypic consequences of mtDNA mutations, specifically in the aging process. For this purpose we used mouse models and different approaches for DNA sequence analysis.

Our group previously demonstrated that increased somatic mtDNA mutagenesis can cause premature aging in the mouse, thus providing the first experimental data for the involvement of mtDNA mutations in aging. This work was based on analysis of the mtDNA mutator mouse, which is a knock-in mouse model with increased mtDNA mutation load due to a proofreading deficiency of mitochondrial DNA polymerase (Polγ). The work presented in this thesis demonstrates that a significant proportion of mtDNA mutations are passed on via the maternal germline and expand clonally in the offspring to contribute significantly to the observed phenotypes. By performing a variety of mouse crosses to generate mice with different combinations of maternally inherited and somatic mtDNA mutations we found that even low levels of inherited mtDNA mutations can lead to premature aging phenotypes. These phenotypes could be rescued by reintroduction of wild-type mtDNA. Our findings show that clonal expansion of maternally transmitted mtDNA mutations can be an important factor in the aging process. This finding is very intriguing in light of recent findings that inherited heteroplasmy is common in humans. Furthermore, a combination of inherited and somatic mtDNA mutations can disturb development to cause stochastic brain malformations.

There is a lack of understanding how pathogenic mtDNA mutations are inherited and studies in the area have been hampered by the lack of appropriate animal models. By backcrossing mtDNA mutator females we were able to isolate a mouse line carrying a single base pair deletion in the mitochondrial $tRNA^{Met}$ gene. Although this mutation was under selection in living offspring, we failed to observe any clear selection in germ cells and primary oocytes. The selection is instead occurring after fertilization as the embryo develops.

Finally, we investigated the clonality of mtDNA inheritance. Mutations in mtDNA are the most commonly used marker in population genetics studies, and much of this popularity is based on the assumed lack of recombination in mtDNA. This assumption has now been questioned in a number of studies that report frequent recombination of mtDNA. In this thesis, I developed a novel direct cloning protocol to trap single mtDNA molecules, allowing for DNA sequence analysis without prior amplification by PCR. Applying this technique, there was no evidence for germline recombination of mtDNA in mice that has been heteroplasmic for >50 generations.

LIST OF PUBLICATIONS

I. Ultra-Deep Sequencing of Mouse Mitochondrial DNA:
Mutational Patterns and Their Origins.
A Ameur, JB Stewart, C Freyer, **E Hagström**, M Ingeman, N-G Larsson and U Gyllensten.
PLoS Genetics 2011, 7(3): e1002028 doi: 10.1371/journal.pgen.1002028

II. Variation in germline mtDNA heteroplasmy is determined prenatally but modified during subsequent transmission.
C Freyer, LM Cree, A Mourier, JB Stewart, C Koolmeister, D Milenkovic, T Wai, VI Floros, **E Hagström**, E Chatzidaki, RF Wiesner, DC Samuels, N-G Larsson and PF Chinnery.
Nature Genetics 2012, 44, 1282-1287

III. Germline mtDNA mutations aggravate aging and impair brain development.
JM Ross, JB Stewart, **E Hagström**, S Brené, A Mourier, G Coppotelli, C Freyer, M Lagouge, BJ Hoffer, L Olson and N-G Larsson.
Nature 2013, doi:10.1038/nature12474

IV. No recombination of mtDNA after heteroplasmy for 50 generations
in the mouse maternal germline.
E Hagström, C Freyer, BJ Battersby, JB Stewart and N-G Larsson.
Nucleic Acids Research 2013, doi:10.1093/nar/gkt969

CONTENTS

LIST OF ABBREVIATIONS

ADP	Adenosine di-phosphate
ATP	Adenosine tri-phosphate
bp	Base pair
COX	Cytochrome c oxidase
CSB	Conserved sequence block
DNA	Deoxyribonucleic acid
d_N	Non-synonymous mutation
d_S	Synonymous mutation
d.p.c	Days post coitum
ETC	Electron transport chain
EM	Electron microscope
FACS	Fluorescent activated cell sorting
FAD	Flavine-adenine dinucleotide
GFP	Green fluorescent protein
kb	Kilo base pair
LD	Linkage disequilibrium
LSP	Light strand promotor
Mbp	Mega base pair
mtDNA	Mitochondrial DNA
mtSSB	Mitochondrial single strand binding protein
NAD	Nicotinamide adenine dinucleotide
ND	NADH dehydrogenase
nDNA	Nuclear DNA
NGS	Next generation sequencing
PCR	Polymerase chain reaction
qPCR	Quantitative PCR
PGC	Primordial germ cell
PolG	Mitochondrial DNA polymerase
POLRMT	Mitochondrial RNA polymerase
RC	Respiratory chain
RFLP	Restriction fragment length polymorphism
rRNA	Ribosomal ribonucleic acid
ROS	Reactive oxygen species
smPCR	Single molecule PCR
tRNA	Transfer ribonucleic acid

INTRODUCTION

MITOCHONDRIA AND ORIGIN

The blue print of life is DNA. The information stored within this molecule is used to create the bewildering variation of living organisms that populate our planet. In animals, DNA is found in two locations of the cell: the nucleus and the mitochondria. The majority of the DNA is in the nucleus and ~1 % in the mitochondria. Mitochondria are cellular organelles that are the generators of ATP, the energy currency of cells. These organelles have a bacterial origin, characterised by their circular genome, double membrane and their close relation to the α-proteobacteria (Burger, Gray & Lang, 1999; Andersson, Karlberg, Canbäck, *et al.*, 2003). The eukaryotic cell evolved only once and as all eukaryotes have mitochondria (a few organisms have lost them secondarily) the organelle must have been present in the ancestral eukaryote (Lane, 2007). The prevalent hypothesis is that the mitochondria entered the proto-eukaryote about 2 billion years ago by endophagocytosis. From this event a symbiotic relationship was established which proved to be extremely beneficial for both parts (Dyall, 2004). However, present benefits probably do not shed any light on the initial benefits. The popular notion is that the endosymbiont provided the host cell with ATP in the exchange for carbohydrates. This explanation does not hold up for closer scrutiny. Export and import of ATP has never been detected in free living bacteria and the ATP/ADP transport system found in present mitochondria has a eukaryotic origin (Andersson, Karlberg, Canbäck, *et al.*, 2003). An alternative explanation has to do with the appearance of atmospheric oxygen. Two billion years ago, atmospheric oxygen content started to rise, creating a highly toxic environment for all anaerobic life. Therefore any possibility to lower local oxygen concentration would be highly adaptive. Similar to mitochondria, α-proteobacteria utilize and consume oxygen during aerobic respiration. Hence, a possible benefit of the endosymbiont could have been that it provided its host with local reduction of oxygen content.

MITOCHONDRIAL FUNCTION AND FORM

Form

A mitochondrion can be divided into four different compartments (Fig. 1): (i) the outer mitochondrial membrane, (ii) the inter-membrane space, (iii) the inner mitochondrial

membrane and (iv) the matrix. The outer membrane consists of a lipid bilayer and is semipermeable, allowing for diffusion of molecules <5Da in size. The inner mitochondrial membrane is impermeable and does not allow diffusion. The five complexes of the mitochondrial oxidative phosphorylation system are located in the inner membrane. The inner membrane is folded into invaginations, creating cristae. This is probably a way to increase its surface, allowing higher respiratory capacity per mitochondrion. Between these two membranes is the inter-membrane space. Enclosed by the inner membrane is the matrix, where mtDNA and mito-ribosomes are located.

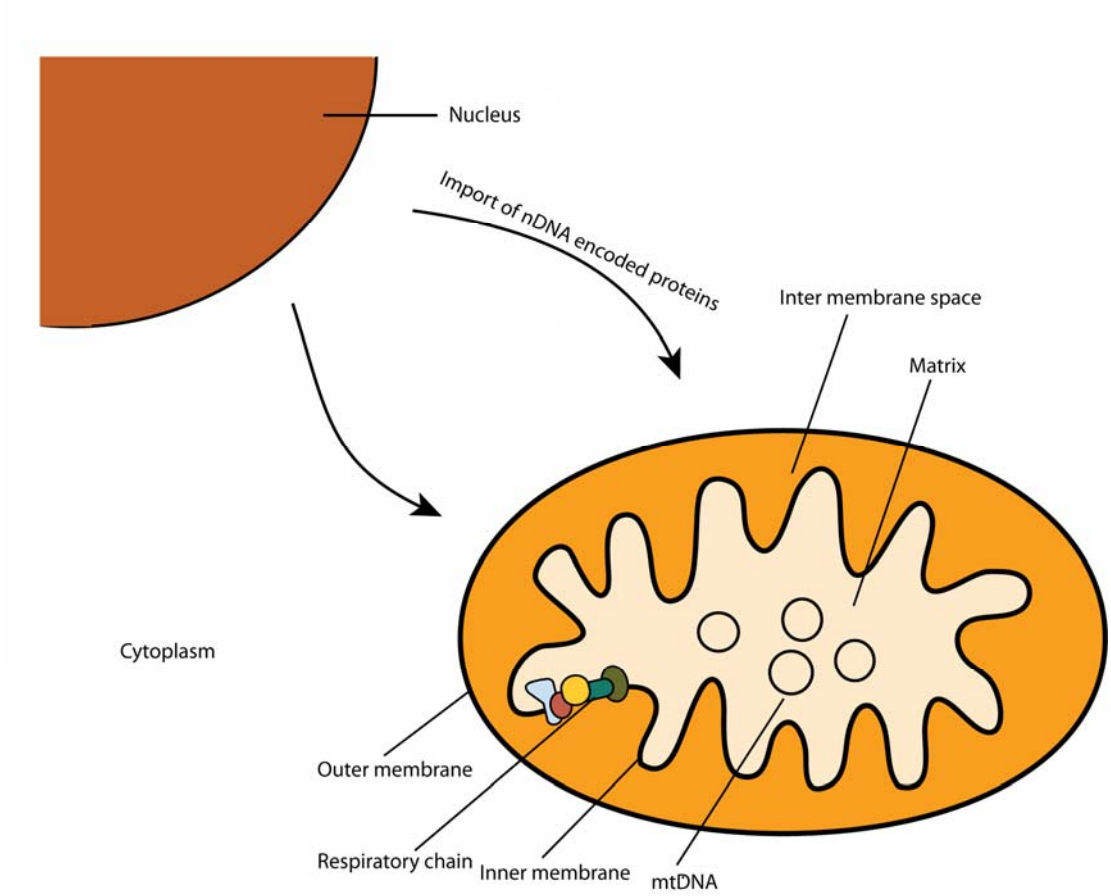

Figure 1. Schematic drawing showing the major components of the mitochondria. The mitochondrial respiratory chain (RC) is located in the inner mitochondrial membrane. The between $10^{\wedge 2}$-$10^{\wedge 5}$ mtDNA molecules (depending on the cell type) are located in the matrix. The majority of the mitochondrial proteins are nuclear encoded and imported into the mitochondria.

Mitochondria used to be regarded as single bean-shaped entities, but this model has been extensively modified in recent years. It is now widely accepted that mitochondria vary greatly in shape and size between different tissues (Kuznetsov, Hermann, Saks, *et*

al., 2009) and that they can alter their organization and morphology due to intra- and extra cellular signals (Seo, Joseph, Dutta, *et al.*, 2010). Furthermore, mitochondria within an organism form a dynamic, constantly dividing and fusing network with a varying morphology (Chen & Chan, 2004). The system of fusion and fission is vital for mitochondrial function. Inhibition of fusion leads to loss of mtDNA (Chen & Chan, 2010) and causes embryonic lethality in animals (Nunnari & Suomalainen, 2012).The functional significance of this system is still debated, but one explanation might be to achieve complementation between different mitochondria (Nakada, Sato & Hayashi, 2009).

Function

Mitochondria take part in many important cellular functions, including controlled cell death by apoptosis (Smith, Ng, Kluck, *et al.*, 2008) and calcium buffering (Rizzuto, de Stefani, Raffaello, *et al.*, 2012). However mitochondria are most widely known for being the "powerhouse" of the cell, converting ADP to ATP by oxidative phosphorylation.

Catabolism of sugars and fats generates reduced forms of the electron carriers NADH and $FADH_2$. They deliver electrons to complex I and II of the mitochondrial electron transport chain (ETC). The ETC consists of four protein complexes, Complex I (NADH:ubiquinone oxidoreductase), III (ubiquinol-cytochrome c reductase) and IV (cytochrome c oxidase) are under dual genetic control (nDNA and mtDNA encoded proteins) while complex II (succinate:ubiquinone reductase) is built up only of nDNA encoded proteins. The electrons received by complex I and II are passed down the ETC in a tightly controlled series of redox reactions finally reducing oxygen to water. The energy released during the passage is used by complex I, III and IV to pump protons from the matrix into the inter-membrane space, creating an electrochemical proton gradient across the inner membrane. The protons are subsequently trans-located back into the matrix by passing through ATP-synthase. ATP-synthase is the final complex of the oxidative phosphorylation system, which consists of the four complexes of the ETC (complex I-IV) and ATP-synthase (complex V). ATP-synthase utilizes the proton gradient created by complex I, III and IV to combine ADP and phosphate into ATP, the energy currency of the cells.

There are currently two competing models for how the complexes constituting the RC are structured in the inner mitochondrial membrane. The "fluid state" model describes the different complexes as autonomous entities that float freely in the membrane,

connected by the electron transporters cytochrome c and ubiquinone. The second model proposes that the complexes are stably associated, forming super-complexes. The support has shifted from the former to the latter as more evidence has been put forward in support for the existence and significance of super-complexes (Vartak, Porras & Bai, 2013).

mtDNA IN EUKARYOTES

The genes of mtDNA are highly conserved among metazoan organisms and encode proteins needed for aerobic respiration and ribosomal RNAs (Burger, Gray & Lang, 2003). The first mitochondrial genome to be investigated was of mammalian origin. It was therefore initially thought that mtDNA always was circular and of modest size (15-20 kb). It has over time become clear that this is not true and that a significant percentage of fungi and plants have huge mtDNA (up to 2Mb) with linear chromosomes (Burger, Gray & Lang, 2003). Interestingly, there is no correlation between mtDNA size and gene content (Andersson, Karlberg, Canbäck, *et al.*, 2003) with the introduction of intragenic elements being the major factor behind the size increase in plants (Burger, Gray & Lang, 2003).

mtDNA in eukaryotes tend to have between 12-20 protein coding genes and an additional number of tRNA and rRNA genes. The sharp reduction in gene number compared to their α-proteobacteria ancestors has been ascribed to at least two different processes: (i) loss of genes from the mtDNA as they were no longer needed in the new intracellular environment and/or became redundant as host genes of similar function existed and (ii) transfer of organelle genes to the nuclear genome. One intriguing question is why the mitochondrial genome has not been lost altogether. There seems to be a perfect co-segregation between the presence of mtDNA and a functional respiratory chain. The mtDNA genes encoding cytochrome b and cytochrome c oxidase subunit I has been present in all organism surveyed so far (Park & Larsson, 2011). Two different processes have been proposed to explain why mtDNA has been maintained: (i) the gene products of cytochrome b and cytochrome c oxidase are hydrophobic and difficult to transport into mitochondria or (ii) mtDNA is required to be in close proximity of the RC to regulate gene expression in relation to cellular energy demands (Park & Larsson, 2011). At present, 99% of the approximately 1500 mitochondrial proteins are nDNA encoded and imported into the mitochondria (Andersson, Karlberg, Canbäck, *et al.*, 2003; Nunnari & Suomalainen, 2012). The mitochondrial proteins are either of bacterial origin (~50%) or of eukaryotic origin

(~50%). These two groups are functionally divided, with genes of bacterial origin often encoding proteins involved in translation and respiration, whereas those of eukaryotic origin often are encoding regulatory factors and membrane proteins (Andersson, Karlberg, Canbäck, *et al.*, 2003).

mtDNA IN MAMMALS

Mammalian mtDNA is a double stranded circular molecule that ranges from 15-20 kb in size (Fig. 2). The number of protein coding genes is well preserved among Metazoans, with variation in total gene number being mostly due to loss of tRNAs. In humans and mice, mtDNA is ~16.3 kb in size and harbors 37 genes. These genes encode 13 of the ~90 proteins of the mitochondrial RC plus 22 tRNAs and 2 rRNAs.

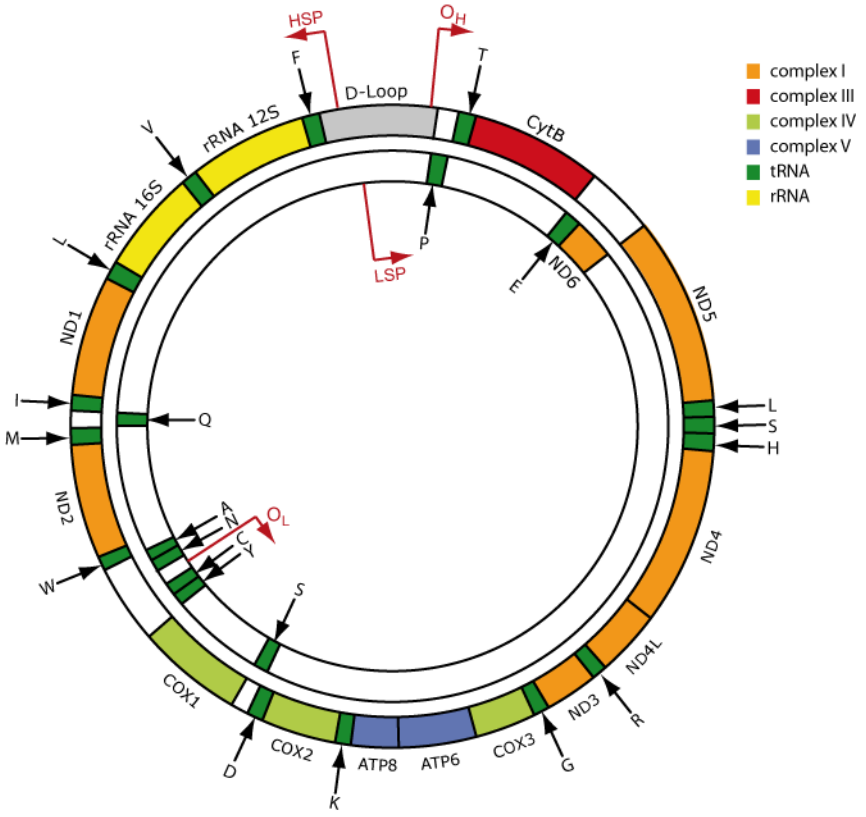

Figure 2. Human/mouse mtDNA is a circular double stranded molecule ~16.5 Kb in size. It harbors 37 genes in total. The remaining ~1500 mitochondrial proteins are nuclear encoded. Complex I (ND1-ND6) is depicted in orange, complex III (CYTB) in red, complex IV (COXI-COXIII) in green and complex V (ATP6 and ATP8) in blue. The ~1 kb non-coding D-loop is depicted in grey, tRNAs in green and the 2 ribosomal genes (12S and 16S rRNA) in yellow. All of the protein coding genes except ND6 are located on the H-strand. Also the majority of the tRNAs and the two rRNAs are located on the H-strand.

The two individual strands of mtDNA are denoted heavy strand (H strand) and light strand (L strand) as they differ in guanine content and hence migrate differently in

alkaline cesium gradients (Larsson, 2010). The H strand harbors all of the protein coding genes (ND1-ND5, ND4L, CYTB, COXI-COXIII, ATP6 and ATP8), except ND6 which is located on the L strand. The majority of the tRNAs and the 2 rRNAs are located on the H strand (Fig. 2). The only major noncoding part of the mammalian mtDNA is the D-loop, also termed the control region. The latter name refers to that the origin of replication of the leading strand and promoters for transcription of both strands are located within this part of the genome.

IMPORTANT ASPECTS OF mtDNA GENETICS

mtDNA differs from the nDNA in some important aspects:

1. mtDNA is polyploid with multiple copies per cell and mitochondrion. The number of copies varies depending on the cell type, ranging from $10^{\wedge 2}$ in sperm to $10^{\wedge 2}$ -$10^{\wedge 4}$ in somatic cells and $10^{\wedge 5}$ in mature oocytes (Shoubridge & Wai, 2007). Some cells such as erythrocytes lack mtDNA altogether.
2. The mtDNA molecules in a mitochondrion, cell or organism can all have the same DNA sequence, a state termed homoplasmy. However, the mtDNA population can also harbor a mixture of mtDNA molecules, a state termed heteroplasmy. Until recently the general notion was that humans were homoplasmic at birth (Taylor & Turnbull, 2005;Parsons, Muniec & Sullivan, 1997). A number of studies have now questioned this model as NGS technology (Payne, Wilson, Yu-Wai-Man, *et al.*, 2013;Li, Schönberg, Schaefer, et al., 2010) and denaturing gels (Tully, Parsons, Steighner, *et al.*, 2000) seems to show a low level heteroplasmy (~1%) universally among humans (Payne, Wilson, Yu-Wai-Man, *et al*., 2012). It is estimated that 25 % of humans carry mtDNA mutations at heteroplasmy levels of 10 % or more (Li, Schönberg, Schaefer, *et al.*, 2010). An alternative explanation to these findings could be the inherent high error rates of the NGS platforms (Schmitt, Kennedy, Salk, *et al.*, 2012). However, the stringent controls applied in these studies together with the fact the other methodological approaches have given similar results (Tully, Parsons, Steighner, *et al.*, 2000) argues against this explanation.
3. There is continuous replication and turnover of mtDNA, even in post mitotic tissues such as heart and skeletal muscle (Larsson, 2010). Replication of mtDNA is not coupled to the cell cycle. This means that some mtDNA molecules may be replicated more than once and some not at all during a

single round of cell division (Bogenhagen & Da Clayton, 1977). This can cause stochastic changes in heteroplasmy level within a cell, tissue and organism over time.

4. mtDNA molecules are packed into protein aggregates called nucleoids within the mitochondrion. Estimates utilizing high resolution microscopy on mammalian cell lines have revealed that each nucleoid contains approximately one mtDNA molecule (Kukat, Wurm, Spåhr, *et al.*, 2011).
5. mtDNA is maternally inherited in mammals, though some exceptions have been reported in the literature.

mtDNA INHERITANCE

Maternal inheritance of mtDNA was first demonstrated by RLFP analysis in humans (Giles, Blanc & Cann, 1980). This is the general inheritance pattern among animals. There are however exceptions. Certain bivalves males have maternal mtDNA in the soma but paternal mtDNA in the germ cells (Ladoukakis & Zouros, 2001). It is sometimes stated that the exclusion of paternal mtDNA during inheritance is due to the fact that mtDNA of sperm do not enter the oocyte during fertilization. This is true for some animals, such as tunicates. In mammals, sperm mitochondria do enter the oocyte, but are then degraded by an incompletely understood mechanism (Rokas, Ladoukakis & Zouros, 2003). It has been observed that mitochondria in sperm become ubiquitinated during spermatogenesis (Sutovsky, DMoreno, Ramalho-Santos, *et al.*, 2000). Ubiquitination is a marker for proteolytic degradation. The strict maternal inheritance can however brake down under some circumstances. In inter-specific crosses of mice the selective destruction of paternally inherited mitochondria seems to fail, resulting in transmission of low amounts of paternal mitochondria to the offspring (Kaneda, Hayashi & Takahama, 1995;Gyllensten, Wharton, Josefsson, *et al.*, 1991). This process has been termed paternal leakage. The rate of paternal leakage has not been studies extensively, but the few studies conducted in this field has reported frequencies ranging between 10^{-3} – 10^{-4} per fertilization (Rokas, Ladoukakis & Zouros, 2003).

During transmission between generations the mtDNA pool in the maternal germline is passing through a physical and/or genetic bottleneck. The bottleneck hypothesis was put forward about thirty years ago after observing rapid shifts of mtDNA variants over just a few generations in cows (Olivo, Walle, Laipis, *et al.*, 1983). The observation that

most mtDNA polymorphisms in organisms tend to be found between individuals and not within individuals also imply rapid segregation. Mitochondria cannot be made *de novo*, but have to be elaborated from existing mitochondria. Hence, all mitochondria in a mammal are derived from those in the mother´s oocyte. Studies using qPCR (Wai, Teoli & Shoubridge, 2008) and mathematical modeling (Jenuth, Peterson, Fu, *et al.*, 1996) have shown that there is a sharp reduction in mtDNA copy number per cell during germline development in mice. A mouse oocyte contains approximately $10^{\wedge 5}$ mtDNA molecules (Shoubridge & Wai, 2007). However, during the first 7 days of embryonic development mtDNA replication is halted (Thundathil, Filion & Smith, 2005;Shoubridge & Wai, 2007). The result is that the mtDNA molecules present in the fertilized oocyte are allocated among the dividing cells, resulting in fewer and fewer mtDNA molecules per cell by each cell division. The germline is established within the developing embryo as a cluster of approximately 50 cells called primordial germ cells (PGCs). The number of mtDNA molecules per PGC is low and has been estimated to range between 180-280 (Jenuth, Peterson, Fu, *et al.*, 1996;Wai, Teoli & Shoubridge, 2008). Hence, during transmission only 0,01 % of the mitochondria in the fertilized oocyte contribute to the next generation (Shoubridge & Wai, 2007) creating a stark subsampling effect. The rapid segregation of mtDNA variants between generations has mostly been ascribed to this physical reduction in mtDNA copy number (Jenuth, Peterson, Fu, *et al.*, 1996;Cree, Samuels, de Lopes, *et al.*, 2008). An alternative explanation has been put forward, ascribing the segregation of mtDNA variants to replication of a subsample of mtDNA molecules during germline development (Wai, Teoli & Shoubridge, 2008).

mtDNA REPLICATION AND REPAIR

Replication

Replication of mtDNA is strictly dependent on nDNA encoded proteins. Hence, all proteins needed for mtDNA replication and maintenance are imported from the cytosol into the mitochondria. The polymerase responsible for mtDNA replication is DNA polymerase γ (Polγ). Of the 16 identified eukaryotic DNA polymerases this is the only one that has been detected in mitochondria (Graziewicz, Longley & Copeland, 2006). Polγ consists of a catalytic subunit - having a 3´-5´exonuclease activity - and an accessory subunit required for processivity. The exonuclease activity can degrade single-stranded DNA and show a preference for 3´mismatches (Kaguni & Olson,

1989). This activity is critically dependent on conserved amino acid residues. Mutating these residues in yeast (Foury, S & Vanderstraeten, 1992) and mice (Trifunovic, Wredenberg, Falkenberg, *et al.*, 2004) resulted in an increase in mtDNA mutations, clearly showing the importance of the exonuclease activity for mtDNA integrity (Longley, 2001).

Other known proteins relevant for mtDNA replication include: mtSSB (binds single stranded DNA and stimulates Polγ activity), mtDNA helicase Twinkle (separates the two DNA strands during replication), topoisomerases (alter the topology of mtDNA by creating reversible breaks in the phosphordiester DNA backbone) and RNase H with a possible role in removal of RNA primers (Kasiviswanathan, Collins, Copel, *et al.*, 2012).

To date three different models of mtDNA replication have been proposed: (i) an asynchronous strand displacement model (Clayton, 1991), (ii) a strand coupled bidirectional model (Holt, Lorimer & Jacobs, 2000), and (iii) a version of the strand coupled model termed RITOLS (RNA Incorporated Through Lagging Strand) (Yang, Bowmaker, Reyes, *et al.*, 2002). In the first proposed model, initiation of mtDNA replication is dependent on the creation of a short RNA primer generated by a prematurely terminated transcript from LSP within the control region. In mitochondria, transcription and replication are tightly coupled processes. However, the models then differ in how they propose that mtDNA replication proceeds.

The asynchronous strand displacement model was proposed by studying replication intermediates by EM and biochemistry. It postulates that after approximately 70% of the H-strand has been replicated, OriL is exposed and L-strand replication can initiate in the opposite direction. The strand coupled bidirectional model is based on findings from 2D gels, depicting replication intermediates with shapes indicative of replication forks (Holt, Lorimer & Jacobs, 2000). These replication intermediates were reported to be resistant to nucleases digesting single stranded DNA, consistent with duplex replication intermediates (Holt, Lorimer & Jacobs, 2000). Further analysis of these replication intermediates - using antibodies specific for RNA/DNA hybrids showed the presence of RNA in mtDNA, leading to the proposal of the RITOLS model of replication (Yang, Bowmaker, Reyes, *et al.*, 2002).

To date, there is no consensus whether these three proposed modes of replication are all true *in vivo* and whether one of them is the dominating mode of mtDNA replication. Three of the proteins involved in mtDNA replication (Polg, Polrmt and Twinkle) are derived from bacteriophages (Shutt & Gray, 2006). In bacteriophages, the mode of

DNA replication switches during different stages of phage infection. It is therefore not unrealistic to presume that switching in mode of DNA replication also takes place in mitochondria. However, findings that it is possible to initiate replication from OriL *in vitro* (Fusté, Wanrooij, Jemt, *et al.*, 2010) and that it is impossible to mutate some nucleotide positions in OriL *in vivo* (Wanrooij, Fusté, Stewart, *et al.*, 2012) lends support to the asynchronous strand displacement model as this is the only model in which OriL has a proposed function . Also, the high levels of mtSSB in mitochondria lends support to the asynchronous strand displacement model, as this is the only model generating longer stretches of single stranded DNA.

Repair

mtDNA used to be regarded as a naked molecule without any physical protection (such as histones in nDNA) and devoid of DNA repair mechanisms. This view has changed during the last decade. mtDNA has been shown to be packaged into protein structures called nucleoids and is covered with TFAM molecules (one TFAM molecule per 15-20 bp, (Kukat, Wurm, Spåhr, *et al.*, 2011). It has also been shown that mitochondria do have DNA repair mechanisms. Both short-patch base excision repair (Bogenhagen, 1999; de Souza-Pinto,Wilson, Stevnsner, *et al.*, 2008) and long-patch base excision repair (Zheng, Zhou, Guo, *et al.*, 2008) have been identified in mammalian mitochondria. Both of these mechanisms include the removal of the damage base(s) by specific glycosylases and subsequent filling of the gap by Polγ. Long patch base excision repair needs the additional help of FEN-1 and DNA2 proteins in the process (Kasiviswanathan, Collins, Copel, *et al.*, 2012). The other DNA repair mechanisms identified in nDNA are either not present in mtDNA or remain to be discovered (Larsson, 2010).

mtDNA MUTATIONS AND DISEASES

Mutations in the mitochondrial genome have been shown to cause different neuromuscular diseases. Pathogenic mtDNA mutations were first identified in 1988 (Holt, Harding & Morgan-Hughes, 1988;Wallace, Singh, Lott, *et al.*, 1988) and since then more than 400 disease causing mutations have been identified (Li, Schönberg, Schaefer, et al., 2010). Identification of disease causing mtDNA mutations came as a bit of a surprise as it was assumed that the polyploid nature of mtDNA should protect the cell against the deleterious effects of *de novo* mutations. Epidemiological studies have shown that diseases caused by mtDNA mutations have a prevalence of 1/5000

(Chinnery, Johnson, Wardell, *et al.*, 2000;Schaefer, McFarland, Blakely, *et al.*, 2008) and that these mutations are present in 1/200 live births and occur de novo in 1/1000 live births (Elliott, Samuels, Eden, *et al.*, 2008). These mutations can show extreme variability in phenotypic expression (McFarland, Taylor & Turnbull, 2010) with the same mtDNA mutation resulting in different phenotypes in different individuals, implying that nuclear encoded genes play an important role in disease development. mtDNA mutations tend to be functionally recessive and need to reach a certain heteroplasmy level (60-90 % depending on the mutation) to cause a biochemical defect and disease (Larsson, 2010). There are however exceptions where low level heteroplasmic mutations can cause diseases (Dubeau, de Stefano, Zifkin, *et al.*, 2000;Sacconi, Salviati, Nishigaki, *et al.*, 2008). Furthermore, due to the mitochondrial bottleneck, pathogenic mutations can segregate to very different heteroplasmy levels in the offspring. Hence, a mother carrying a pathogenic mutation can give birth to both severely sick and healthy children. To date there is no way to counsel prospective mothers carrying pathogenic mtDNA mutations whether their children will be healthy or develop mitochondrial disease (Freyer, Cree, Mourier, *et al.*, 2012).

mtDNA MUTATIONS AND AGING

Aging is commonly defined as a progressive and generalized impairment of cellular function, resulting in an increasing vulnerability to environmental challenge and a growing risk of disease and death of the organism (Kirkwood, 2005). Genetics has a clear role in aging as demonstrated by the different lifespans of animals in captivity, the heritability of human longevity and the possibility to select for long lived mutants in model organisms (Kirkwood, 2005). A positive correlation between increasing age, number of dysfunctional mitochondria and mtDNA mutation load has been found in a variety of organisms, implicating mtDNA mutations in the aging process (Larsson, 2010). It was initially thought that the polyploid nature of mtDNA would protect the mitochondria against deleterious effect of *de novo* mutations. However, it was later shown that there is a decline of mitochondrial respiratory function with age (Trounce, Byrne & Marzuki, 1989) and that heart and muscle fibers in humans develop a mosaic pattern of respiratory chain deficient cells (Müller-Höcker, 1989;Müller-Höcker, 1990) with 30-60 % of skeletal muscle fibers in aged humans and monkeys containing respiratory deficient segments (Bua, Johnson, Herbst, *et al.*, 2006;Pak, Herbst, Bua, *et al.*, 2003). These respiratory deficient cells tend to co-localize with mtDNA deletions in both rats, monkeys and humans (Larsson, 2010). mtDNA point mutations generally

accumulate in colonic crypts and ~15 % of all crypts in aged humans show respiratory dysfunction (Taylor, Barron, Borthwick, *et al.*, 2003). Hence, a great number of correlative studies support the involvement of mtDNA mutations in the aging process.
It was initially hypothesized that the observed increase in mtDNA mutation load with age was due to the accumulation of somatically generated mutations. An alternative explanation proposed is the clonal expansion of somatic mtDNA mutations generated during embryonic development (Larsson, 2010). A further possibility is that these mutations are not generated somatically at all, but rather inherited via the maternal germline to expand clonally during life. This hypothesis seems quite plausible in light of recent findings that heteroplasmy is universal among humans and that a great proportion of these mutations seem to have been passed on via the maternal germline (Payne, Wilson, Yu-Wai-Man, *et al*., 2012).

THE mtDNA MUTATOR MOUSE

The mtDNA mutator mouse was developed to experimentally address the involvement of mtDNA mutations in aging (Trifunovic, Wredenberg, Falkenberg, *et al.*, 2004; Kujoth, Hiona, Pugh, *et al.*, 2005). This is a knock-in mouse where an amino acid residue critical for the exonuclease activity of Polγ has been changed from an aspartic to an alanine. This results in a drastically impaired proof reading function of Polγ. The mutator mice are born in mendelian proportions and have a normal phenotype at birth (Trifunovic, Wredenberg, Falkenberg, *et al.*, 2004; Kujoth, Hiona, Pugh, *et al.*, 2005). However, at 175-270 days of age they begin to develop a progeroid phenotype with alopecia, kyphosis, anemia, reduced fertility, heart disease and decreased spontaneous movement (Trifunovic, Wredenberg, Falkenberg, *et al.*, 2004; Kujoth, Hiona, Pugh, *et al.*, 2005). This occurs concomitantly with an increase in number of respiratory chain deficient cells in liver and heart. Interestingly, no increase in ROS production was detected in these mice, but an increase in apoptosis was observed (Trifunovic, Hansson, Wredenberg, *et al.*, 2005). The same pattern - with increased apoptosis but not ROS production - has been observed in other mouse models with respiratory chain dysfunction (Trifunovic, Wredenberg, Falkenberg, *et al*., 2004). The median and maximal life span of the mice is 336-416 and 430-460 days, respectively (Trifunovic, Wredenberg, Falkenberg, *et al.*, 2004;Kujoth, Hiona, Pugh, *et al.*, 2005). Due to the proof reading deficiency of Polγ, mice homozygous for the mutator allele ($Polg^{mut/mut}$) have a 3-27 fold increase in mutation load compared to WT siblings and approximately 30 % of the mtDNA molecules are linear molecules with deletions (Trifunovic,

Wredenberg, Falkenberg, *et al*., 2004). The linear molecules with deletions seem to be a result of replication stalling (Bailey, Cluett, Reyes, *et al.*, 2009) as the deleted part consists of the minor arc between OriL and OriH. A third group of mutated molecules, in the form of circular molecules with deletions, have also been also been reported (Vermulst, Wanagat, Kujoth, *et al.*, 2008).

The molecular mechanism behind the progeroid phenotype has been hotly debated (Vermulst, Wanagat & Loeb, 2009; Edgar, Larsson & Trifunovic, 2010; Edgar & Trifunovic, 2009) and claims have been put forward that either point mutations (Edgar, Shabalina, Camara, *et al.*, 2009) or circular molecules with deletions (Vermulst, Wanagat, Kujoth, *et al.*, 2008) are the driving molecular force creating the aging phenotype. However, in light of current knowledge about mitochondrial genetics and biochemistry, arguments supporting that circular molecules with deletions are the causative factor are very weak. These molecules are found at extremely low heteroplasmy levels (Edgar, Shabalina, Camara, *et al*., 2009). Point mutations and deletions in mtDNA tend to be functionally recessive and need to reach a certain threshold level to cause a biochemical defect (60-90 % depending on the mutation). With a detection limit of 0,1 %, Edgar, Shabalina, Camara, *et al*., 2009 were unable to detect any circular deleted mtDNA molecules. Genetically modified mice carrying high levels either single (Inoue, Nakada, Ogura, *et al.*, 2000) or multiple (Tyynismaa, Mjosund, Wanrooij, *et al.*, 2005) deletions in mtDNA do not show premature aging. Furthermore, deletions tend to generate translational defects as one or more tRNAs often are removed by the deleted region. In contrast to this expectation, translation is unaffected in the mtDNA mutator mouse (Edgar, Shabalina, Camara, *et al*., 2009). Taken together these findings argue against that circular molecules with deletions have any role in causing the progeroid phenotype.

The steady state levels of assembled mtDNA encoded RC complexes were much lower in the mutator than in WT mice (Edgar, Shabalina, Camara, *et al*., 2009). This deficiency in assembled RC complexes was not due to impaired protein synthesis as no decrease in transcription or translation was detected (Edgar, Shabalina, Camara, *et al*., 2009). However, an increase in turnover of mtDNA encoded respiratory chain complexes was detected. A plausible explanation for the progressive respiratory dysfunction in the mutator mouse is thus that accumulation of point mutations in protein coding mtDNA genes generates mis-folded proteins. The misfolding prevents proper assembly and/or stability of the respirator chain complexes, leading to subsequent degradation and increased turnover.

mtDNA AND METHODOLOGY

To date, there is no universally accepted method to identify and quantify the different types of mtDNA mutations that are associated with diseases and aging. What makes this measurement problematic (compared to nDNA) is the cellular organization of mtDNA. The polyploid nature of mtDNA means that mutations can be found at a wide range of levels. This is in stark contrast to nDNA where a mutation is either present on one or both copies of a gene. To complicate things further, mtDNA mutations can expand clonally in a cell or a cluster of cells. This has been documented during aging in various organisms (Wanagat, Cao, Pathare, *et al.*, 2001; Bua, Johnson, Herbst, *et al.*, 2006; Pak, Herbst, Bua, *et al.*, 2003). The result is that even though some cells carry a mutation at high heteroplasmy level, the mean level of that mutation in the tissue as a whole can be very low (Alex, Kraytsberg, Khrapko, *et al.*, 2007). Hence, mtDNA mutation analysis involves the identification of homoplasmic to very low frequency mutations, which requires a highly stringent and precise method. Application of different methods to estimate the mutation load in the same tissue have systematically generated estimates that differ by 2-3 orders of magnitude (Alex, Kraytsberg, Khrapko, *et al.*, 2007; Greaves, Beadle, Taylor, *et al.*, 2009) clearly showing the current methodological shortcomings in mtDNA mutation analysis. This discrepancy may be due to many different factors, but one significant problem lies in that all methods have an inherent error rate, inducing different types of artifacts to varying degrees. This is a problem for mitochondrial research in general, but maybe most pronounced in mitochondrial aging research, investigating the role of mtDNA mutations in the aging process. A handful of methods are used for mtDNA mutation identification and quantification. Listed below are the most commonly used methods. The pros and cons associated with respective approach are also included.

Sanger sequencing of PCR products

This has been the gold standard technique for mtDNA mutation identification. Due to the initial PCR amplification step very little starting material is needed. Hence, a single cell can be sequenced and analyzed. However, a major limitation with the technique is the detection level. As the sequence data comes out as the mean signal strength per base pair, mutations with a heteroplasmy level of 20 % or less will not be detectable. This is a not a major problem for the identification of diseases causing mtDNA mutations, which tend to expand monoclonally to high heteroplasmy levels. However, it is a

significant limitation in aging research where many different mutations expand polyclonally within the organism. These mutations expand to high heteroplasmy levels within a single cell or a cluster of cells. This generates a low heteroplasmy level of the mutation in the tissue as a whole and unless DNA sequence analysis is carried out on single cells, heteroplasmy levels will be below the detection limit of Sanger sequencing.

Sanger sequencing of cloned PCR products

To circumvent the problems associated with the detection limit of conventional Sanger sequencing, cloned PCR products may instead be analyzed. This will generate sequence data on the single molecule level as each clone is derived from a single PCR amplicon. There are however limitations with this approach. The polymerase used during PCR amplification will introduce artifacts in the form of point mutations, deletions and chimeric molecules (Pääbo, Irwin & Wilson, 1990; Keohavong & Thilly, 1989). High fidelity polymerases have reported error rate of $6*10^{\wedge-7}$ errors per bp per replication (Pavlov, Pavlova, Kozyavkin, *et al.*, 2004). In an average PCR (20 cycles) this equals a final error rate of $1{,}2*10^{\wedge-5}$ per bp. Another study report an error rate of $1{,}4*10^{\wedge-5}$ error per bp (Smigrodzki, Parks & Parker, 2004). This is generally not a problem during standard Sanger sequencing as the frequency of most artificial mutations will be very low and hence not detectable. However, when cloning single PCR amplicons these artifacts will be present in all of the copies and will therefore be indistinguishable from true mutations.

The most common way to test the error rate of a method is to pass a control DNA fragment through the entire experimental procedure and check for novel mutations. As the sequence of the control DNA is known, any novel mutations in the control DNA can be scored as errors introduced by the method. The problem is, however, that control DNA templates differs from the cellular DNA templates in at least two ways that can affect the number of errors that are introduced. First, control DNA is often in the form of a high copy number plasmid. Therefor much less number of PCR cycles is needed to reach PCR plateau than when using a cellular DNA template (Alex, Kraytsberg, Khrapko, *et al*., 2007). This means that fewer artifacts will be introduced in the control DNA than in the cellular DNA. Secondly, cellular DNA often contains chemical modifications (for example 8-oxo-guanine and deaminated cytosine) that can be converted into mutations. The cloned control DNA is most likely devoid of chemical damage (Alex, Kraytsberg, Khrapko, *et al*., 2007). Taken together, these factors might

lead to an underestimation of the number of errors that are introduced in the cellular template.

The error rate of a method will start to confound the data when it is on par with (or higher) than the mutation load being estimated. Hence the PCR and cloning method is well suited for mutation load estimates in organisms with higher mutation loads, but less suitable for estimating lower mutation loads.

Single molecule PCR

To circumvent the problems with artifacts in the PCR and cloning approach single molecule PCR (smPCR) was developed. The principle of this technique is that templates used in the PCR are diluted to such an extent that only each fifth well has a single DNA template (molecule) in it and generates a PRC product (Kraytsberg & Khrapko, 2005). This approach is very intellectually appealing. It would generate sequence data on single molecule level and would give the possibility to tell artificial and true mutations apart. However, there are some critical aspects regarding this method that have not been properly addressed. First, the idea that template dilution generates PCR reactions with a single DNA template has mostly been based on theoretical calculations and has not been experimentally verified in a satisfying way. Furthermore, it does not exclude the possibility of the presence of fragmented DNA templates in the reaction. Secondly, it has been reported that mtDNA exists in the form of catenated networks in some tissues of both mice and humans (Pohjoismäki & Goffart, 2011). So far the smPCR protocol has not reported to include any treatment of the mtDNA templates with endonuclease and/or topoisomerase prior to amplification.

Random mutation capture

One way of avoiding PCR induced artifacts is to conduct the screening for mutations prior to PCR amplification. The Random Mutation Capture (RMC) method is a restriction enzyme based method aimed at screening a large number of specific short DNA sequences for mutations prior to PCR amplification (Bielas & Loeb, 2005) (Vermulst, Wanagat, Kujoth, *et al.*, 2008). In short, the target sequence is cleaved with a restriction enzyme, followed by PCR amplification with primers annealing up- and downstream of the target sequence. As WT sequences will be cut, only mutated molecules will generate a PCR product. Hence, the number of PCR products can be used as an indicator for the number of mutated sites. As the number of mtDNA molecules screened has been quantified by qPCR, the mutation rate can be calculated.

There are however limitations with this method. The most critical, from a mitochondrial aging research perspective, is that it only screens for mutations within a specific target sequence, which is very limited in size (often only 4 bp). It has been documented that mutation accumulation varies between nucleotides positions in mtDNA (Stoneking, 2000) and that mtDNA mutations accumulate in a clonal fashion during aging (Larsson, 2010). Screening for mutations by RMC systematically generates lower mutation load estimates than with other methods (Greaves, Beadle, Taylor, *et al*., 2009; Alex, Kraytsberg, Khrapko, *et al*., 2007). This is most likely a consequence of the method, which misses mutations due to its restricted target sequence. This severely limits its application in mitochondrial aging research.

Direct cloning

The most stringent methodological approach to identify and quantify mtDNA mutations should be direct cloning (Edgar & Trifunovic, 2009; Alex, Kraytsberg, Khrapko, *et al*., 2007). As the cloning is performed by using non-amplified genetic material the problems associated with PCR - in the form of point mutations and artificial recombinant molecules - should be avoided to a great extent. The errors introduced by DNA replication should be few, as the DNA replication error rate in the vectors normally used for direct cloning has been estimated to be very low, $7.7*10^{\wedge-8}$ to $5.4*10^{\wedge-10}$ errors per bp (Drake, Charlesworth, Charlesworth, *et al.*, 1998). Furthermore, it should also be possible to differentiate between artificial mutations and true mutations. As the sequence data is based on a single molecule, artifacts introduced either by vector DNA replication or by any downstream reaction should show up as heteroplasmies in the sequence data and can thus be discarded (Kraytsberg & Khrapko, 2005). Direct cloning should therefore be a suitable methodological approach to estimate mutation load in organisms with a low mutation load. However, direct cloning of human and mouse mtDNA has proven difficult (Yonemura, Nakada, Sato, *et al.*, 2007; Bigger, Tolmachov, Collombet, *et al.*, 2000). The use of bacterial vectors have nevertheless proven successful in direct cloning of mtDNA in a few occasions (Bigger, 2003; Yonemura, Nakada, Sato, *et al.*, 2007) but it demanded a lot of starting material and gave a low yield in the form of few clones In summary, direct cloning, albeit sometimes labor intensive, offers one of the more error free and least biased ways of analyzing mtDNA mutations.

Next generation sequencing

The appearance of next generation sequencing (NGS) platforms offers an opportunity to screen for mtDNA mutations at a depth that has not previously been possible. These platforms are all based on parallel amplification and sequencing of millions of single DNA molecules, generating data in the order of $10^{\wedge 10}$ bp, all on single molecule level. In theory, this should enable identification of single molecules and unique mutations when sequencing at a sufficient depth. However, all of the NGS platforms have a high inherent error rate which has up until recently effectively hindered their application for identification of low level mutations. The unfiltered error rate of the Illumina platform has been reported to be as high as 11 % per sequenced bp (Li, Schönberg, Schaefer, et al., 2010) whereas the error rate of the SOLiD platform has been reported to 0,075 % per sequenced bp (Tolmachov & Collombet, 2012). The error rate after applying different quality filters varies between 0,005-0.01 errors per sequenced bp depending on the platform used, read length, sequence context and the bp position in the amplicon (Kinde, Wu, Papadopoulos, *et al.*, 2011; Li, Schönberg, Schaefer, et al., 2010). One additional step used to control for the high error rate is to sequence a control sample with a known sequence. All variants detected in the control sample can be regarded as method-induced errors. A control sample does not only allow identification of the background error rate, but can also be used to subtract the errors introduced to the actual sample. The generation of artificial mutations tends to be dependent on the DNA sequence and the nucleotide position (Li, Schönberg, Schaefer, *et al*., 2010) as well as the guanine-cytosine content (Payne, Wilson, Yu-Wai-Man, *et al*., 2012). By utilizing a control sample with similar DNA sequence to the actual sample, the mutations found in the former can be subtracted from the latter, and by this control for method-induced errors. A further step to control for errors that has proven useful is to score mutations as true events only if they are verified by reads from both strands. Mutations and errors tend to be sequence specific. Errors will therefore appear on one strand only whereas true mutations will be found on both strands. This approach has been show to lower the error rate 30 fold (Li, Schönberg, Schaefer, *et al*., 2010). An extension of double strand validation is to barcode each molecule with a unique id (Schmitt, Kennedy, Salk, *et al.*, 2012). Here, the double validation can be done both within the strand but also between the two strands of the very same molecule. According to the authors, this allows for identification of mutation with a mutation load of $10^{\wedge -6}$ mutations per bp (Schmitt, Kennedy, Salk, *et al.*, 2012).

In summary, the NGS platforms tend to have a high error rate, but when proper controls and quality filters are applied they can be used to screen for mtDNA mutations with low heteroplasmy levels.

mtDNA AND RECOMBINATION

Mutations in mtDNA are the most frequently used genetic markers depicting the evolution of humans but also other animals. This popularity rests upon three pillars: the high sequence divergence of mtDNA, the almost exclusive maternal transmission and the assumed lack of germline recombination. The two latter points mean that mtDNA is used as a molecular marker showing patterns of clonal inheritance and that an unbroken maternal phylogeny can be constructed. This inheritance pattern reduces the complexity of the analyses as biparental inheritance does not have to be taken into account. However, during the last two decades a number of reports have questioned the assumption that there is no mtDNA recombination based on analyses of human population genetic sequence data (Hagelberg, Goldman, Lió, *et al.*, 1999; Awadalla, Eyre-Walker & Smith, 1999), human patient tissues (Kraytsberg, Schwartz, Brown, *et al.*, 2004; Zsurka, Kraytsberg, Kudina, *et al.*, 2005), tissue culture cells (D'Aurelio, Gajewski & Lin, 2004; Fan, Lin, Potluri, *et al.*, 2012) and tissues from interspecies animal hybrids (Ujvari, Dowton & Madsen, 2007). These findings could have far reaching implications, as even modest frequencies of germline recombination can affect coalescent time estimates and branch lengths in phylogenetic trees (Schierup & Hein, 2000). Hence, if mtDNA does recombine, there might be a need to re-evaluate all studies where this marker has been used.

Today, there is a lack of compelling data either for or against mtDNA recombination in the mammalian germline. This stems from a combination of methodological limitations and the lack of appropriate *in vivo* experimental systems to address this important issue. This has left the research field in uncertainty, with claims both for and against recombination being questioned. The type of tests commonly used to screen for recombination can be divided into direct and indirect tests. Listed below are the methodological approaches commonly used to screen for mtDNA recombination and the pros and cons associated with them.

Indirect tests utilize population genetic data and screen for: (i) a negative correlation between linkage disequilibrium (LD) and physical distance (recombination breaks up linkage between mutations and the chance of this to occur increases with physical distance) or, (ii) an excess of homoplasies within the data set. Homoplasies are

mutations that are identical, but not by descent. Finding a significant excess of homoplasies within different branches of a phylogenetic tree can be an indication of recombination (but alternative explanations exists, see below). The advantage of indirect tests is that one uses data covering many generations, hence investigating germline events. As somatic recombination is irrelevant to evolution and only germline recombination will affect the clonality of mtDNA inheritance, this is a very important point. The problems with indirect tests are, however, their low statistical power to detect recombination (Rokas, Ladoukakis & Zouros, 2003; Ladoukakis & Eyre-Walker, 2004) and the presence of mutational hotspots in mtDNA. Mutational hotspots and uneven distribution of mutations can create the same pattern and distribution of genetic markers as recombination does. LD analysis (Awadalla, Eyre-Walker & Smith, 1999) and homoplasmy tests (Hagelberg, Goldman, Lió, *et al*., 1999) were used in two publications that reported mtDNA recombination in the germline of humans and caused a great stir within the scientific community (Kivisild & Villems, 2000; Elson, Andrews, Chinnery, *et al.*, 2001; Piganeau & Eyre-Walker, 2004; Bandelt, 2005). However, these studies did not hold up to closer scrutiny and the positive results for recombination turned out to be due to a miss alignment (Hagelberg, Goldman, Lió, *et al*., 1999) and the use of a too small sample size (Awadalla, Eyre-Walker & Smith, 1999). Hence, to date indirect tests have failed to produce any convincing evidence either for or against recombination (Elson & Lightowlers, 2006).

Direct tests screen for recombinant molecules within an organism. For recombination to be detectable this needs to occur between mtDNA molecules with different DNA sequence. Hence, direct tests involve the identification of heteroplasmic organisms and subsequent screening for molecules that carry a mixture of the two mtDNA populations, indicating recombination. A number of studies applying direct tests have reported the presence of recombinant mtDNA molecules in humans (Zsurka, Kraytsberg, Kudina, *et al.*, 2005; Kraytsberg, Schwartz, Brown, et al., 2004), model organisms (Sato, Nakada, Akimoto, *et al.*, 2005; Ujvari, Dowton & Madsen, 2007), and cell lines (D'Aurelio, Gajewski & Lin, 2004; Fan, Lin, Potluri, et al., 2012). However, a significant limitation with direct tests is that they mostly cover somatic events which are irrelevant to evolution and the question of mtDNA clonality. A further matter that complicates the application of direct tests is homoplasies and genetic drift. Homoplasies can generate a distribution of mutations that is difficult to tell apart from recombination, especially when the mtDNA populations differ at a limited number of nucleotide positions. The loss of mtDNA haplotypes through genetic drift (termed

lineage extinction) makes identification of recombinant molecules difficult. If one of the ancestral haplotypes that constitutes a recombinant has been lost it will effectively hinder identification of the recombinant (Elson & Lightowlers, 2006). Furthermore, many of the studies applying direct tests have used sequencing of cloned PCR products to screen for recombinants. PCR does produce chimeric PCR products *in vitro* that can be falsely interpreted as evidence for *in vivo* recombination. Creation of artificial recombinants during PCR amplification when using a mixed population of templates has been robustly documented in nuclear encoded genes (Pääbo, Irwin & Wilson, 1990; Keohavong & Thilly, 1989). The most stringent study conducted so far applying direct test is the study by (Sato, Nakada, Akimoto, *et al.*, 2005). In this paper direct cloning and sequencing of mtDNA molecules from a two heteroplasmic mouse lines was carried out. One of the mouse lines was heteroplasmic for a deletion causing respiratory deficiency and the other mouse line was heteroplasmic for neutral mtDNA genotypes. They detected ~1% recombinants in the respiratory deficient mouse line but no recombinants in the mouse line with neutral heteroplasmy. This finding may imply that maybe there is selection for somatic recombinants in the soma under mitochondrial dysfunction. However, one severe limitation of this study is that it almost exclusively investigates somatic events which, as said, are irrelevant for evolution and the question of the clonality of mtDNA inheritance. Hence, to date direct tests have not produced any convincing data either for or against mtDNA recombination in the germline of mammals.

mtDNA AND SELECTION

For a long time mtDNA was considered to be inherited in a neutral fashion, not affected by the forces of natural selection. However, during the last two decades data accumulated indicating negative/purifying selection in mtDNA (Nachman, Brown & Stoneking, 1996; Nachman, 1998; Elson, Turnbull & Howell, 2004; Stewart, Freyer, Elson, *et al.*, 2008; Fan, Waymire, Narula, *et al.*, 2008). The nature of mtDNA sequence evolution has commonly been tested either by: (i) comparing the ratio of non-synonymous (d_N) to synonymous (d_S) mutations per site in a population of organisms (a ratio of 1 equals neutral evolution and a value <1 indicates purifying selection) or (ii) by comparing the ratio of d_N/d_S per site within and between species (McDonald-Kreitman test). If the DNA sequence evolves under neutral conditions this ratio should be the same within and between species. However, if purifying selection is acting upon

the DNA sequence there should be less d_N mutations in the interspecific comparison than in the intraspecific comparison.

Nachman (Nachman, Brown & Stoneking, 1996; Nachman, 1998) applied the McDonald-Kreitman test on mtDNA sequence data from humans, chimpanzees and 25 different animal species to test the neutrality of mtDNA sequence evolution. The ratio of d_N/d_S per site was less between species than within them, indicating purifying selection. A similar approach was used by Elson, Turnbull & Howell (2004) to test the nature of mtDNA sequence evolution in humans. Also here a lack of d_N mutations was observed over time.

Stewart, Freyer, Elson, *et al.*, 2008 obtained results showing signs of strong purifying selection against d_N mutations in the mouse germline. To investigate this, maternal lineages were derived from the mtDNA mutator mouse to pass on a wide range of mtDNA mutations via the germline. In the N2 generation there was an obvious lack of d_N mutations. There were also fewer mutations in first and second codon positions than in third codon positions. Mutations in first and second codon positions tend to generate d_N mutations but mutations in third codon positions tend to generate d_S mutations. Thus, an excess of mutations in the third codon is a universal sign of purifying selection.

Taken together, there is now ample evidence for selection against d_N mutations in mtDNA in the germline of animals. Interestingly, the strength of selection even seems to be stronger in mtDNA than in nDNA. The d_N/d_S ratio can be used to estimate the strength of the selection (the lower the value, the stronger the selection). When d_N/d_S ratios were calculated for protein coding genes located in mtDNA and nDNA the ratios were much lower for the mtDNA encoded genes (Nabholz, Ellegren & Wolf, 2013). This is opposite to what the neutral theory of molecular evolution predicts and indicates that the purifying selection is stronger in mtDNA-encoded genes than in nDNA.

There however seems to be little selection against pathogenic/d_N mutations in the soma (Greaves, Elson, Nooteboom, *et al.*, 2012). When comparing the mutational spectrum in human colonic crypts of aged individuals to the mutational spectrum found in the general population there was a general expansion of d_N mutations in the aged samples (including many more frame shift and termination mutations), indicating relaxed selection pressure in the soma compared to the germline (Greaves, Elson, Nooteboom, *et al.*, 2012).

What has been observed is somatic selection for apparently neutral mtDNA genotypes. Mice artificially heteroplasmic for the NZB and BALB/c mtDNA genotypes (Jenuth,

Peterson, Fu, *et al.*, 1996) showed neutral segregation in the germline but tissue specific selection in the soma (Jenuth, Peterson & Shoubridge, 1997). Liver and kidney selected for the NZB genotype and blood and spleen for the BALB/c genotype. All other tissues showed no signs of selection. This segregation pattern do not depend on mitochondrial function or replication rates, but is modulated by nuclear encoded genes (Battersby, Loredo-Osti & Shoubridge, 2003)

λ-PHAGE CLONING VECTOR

λ-phage has been used as a cloning vector since the early 1970 (Murray & Murray, 1974) and is one of the most well established systems for cloning DNA. The λ-phage is a bacteriophage which proliferates by infecting host bacteria. They attach to the maltoporin receptor of the membrane of the host bacteria and inject their DNA by the use of their tail tube. Upon entering the host the λ-phage can "choose" between two different paths: (i) entering lytic growth and active proliferation by continuous infection and lysing of host cells or (ii) to enter the lysogenic state, incorporating its genome into the host genome. By this pathway the phage genome is replicated together with the host genome and thus transmitted to progeny bacteria. The pathway "chosen" depends on the multiplicity of the infection and the nutritional state of the host bacterial. High multiplicity infection and bad nutritional state of the host bacteria favors the lysogenic state (Boyd, 1951). As lytic growth is required during λ-phage cloning the host bacteria used during cloning are grown on a rich medium.

The λ-phage genome was one of the first to be fully sequenced (Sanger, Coulson, Hong, *et al.*, 1982). The genome of WT λ-phage is a double stranded DNA molecule of 48502 bp in size. The genome can be divided into three functional clusters: (i) the left region which harbors genes necessary for packaging of DNA and assembly of virus particles, (ii) the central region which harbors genes for recombination and lysogenic infection and finally (iii) the right region containing genes for DNA replication and lysis of host bacteria. The central region is not necessary for lytic infection and can thus be removed in λ-phage vectors to make room for a DNA insert. The central region also harbors the genes for recombination. Removing this part effectively makes the λ-phage a rec- cloning vector. In phage particles the genome is linear. However, upon entering the host bacteria the cohesive single stranded termini of the genome is sealed by the host ligase, creating a circular genome (Wu & Taylor, 1971). During lytic infection DNA replication of the λ-phage genome alternates between theta replication and rolling circle mode. After injection into the host cell the now circular genome goes through

approximately 50 rounds of theta replication. Thereafter, it shifts to rolling circle mode creating tandem polymers of λ-phage genomes. These concatemers are then cut up and single genomes packaged into virus particles. The host bacterium is then lysed and the phage particles within it diffuse to infect new bacteria. The lytic cycle takes about 45 minutes and generates approximately 100 virus particles per infected bacterium. When phages infect host bacteria grown as a lawn on an agar plate the clonal expansion of a single phage generates transparent plaques where the host cells have been lysed. These plaques can be picked and DNA extracted for analysis of single clones.

SPECIFIC AIMS

Paper I: To investigate mtDNA mutations in an *in vivo* system by the use of NGS. More specifically, this was done in an effort to elucidate the distribution and origin of mtDNA mutations in mice.

Paper II: To investigate how pathogenic mtDNA mutations are inherited. More specifically, to examine how human equivalent pathogenic tRNA mtDNA mutations are inherited in an in vivo system and whether there is selection against these mutations.

Paper III: To investigate the role of mtDNA mutations in the aging process. More specifically, we wanted to examine the effects of germline inherited mtDNA mutations on the aging process

Paper IV: To investigate whether mtDNA recombines. More specifically, by direct cloning and sequence analysis we wanted to investigate whether recombinant mtDNA molecules are created and passed on via the maternal germline in mammals.

RESULTS AND DISCUSSION

ON SOMATIC mtDNA MUTATIONS AND AGING

In paper I we show that somatic mutations in mtDNA are created by replication errors and not oxidative damage. We also show that mtDNA mutations are passed on via the germline to the offspring and that there is a huge variation in mutation accumulation between different nucleotide positions in mtDNA. Some site accumulated 100-1000 time more mutations than others, clearly showing the presence of mutational hotspots in mtDNA. We failed to detect any increase in mutation load with increasing age in any of the mice used in this study.

To date there is no universally accepted method for accurate quantification of somatic/low heteroplasmy mtDNA mutations. The development of NGS platforms might offer an opportunity to overcome these methodological shortcomings as they generate sequence data with incredible depth. These platforms are all based upon parallel amplification and sequencing of millions of DNA molecules, generating data at the single molecule level. In theory, this should enable identification of single molecules and unique mutations when sufficient sequencing depth is applied. However, NGS platforms tend to have a high inherent error rate, which up until recently has hindered their application for identification of somatic and low frequency mutations.

We hypothesised that by using the SOLiD NGS platform (with the lowest error rate at that time point) in combination with a control for method induced artifacts (in the form of a Lambda cloned mtDNA molecule) that we could control for the high error rates and utilize the SOLiD NGS platform to detect and analyze somatic and low heteroplasmy mtDNA mutations.

To investigate the mtDNA mutation spectrum in mice we used both mtDNA mutator- and wt mice. We set up the standard cross (an intercross between a $PolG^{mut/+}$ male and female, see fig 3) to generate offspring with the following genotypes: (i) mice homozygous for WT PolG ($PolG^{+/+}$) (ii) mice heterozygous for the mutator allele ($PolG^{mut/+}$) and (iii) mice homozygous for the mutator allele ($PolG^{mut/mut}$). DNA from two mice, 30 and 40 weeks of age, of each genotype were sequenced. We also used C57Bl/6 WT mice of 30, 40 and 84 weeks of age, henceforth denoted WT. Finally we used a clonally amplified WT mtDNA molecule (cloned into λ-phage) with a known sequence, henceforth denoted λ^{mtDNA}. This sample was sequenced as a step to control for method induced errors. All novel mutations found in this sample can be regarded as

sequencing errors as we know the input sequence. Thus, all mutations found in λ^{mtDNA} sample were subtracted from the actual samples.

We extracted mtDNA from liver from mice of respective genotypes and age for sequencing on the SOLiD NGS platform. The samples were directly sequenced without any prior amplification. This was done to avoid PCR induced artifacts. We obtained sequence data for all of the samples with coverage of at least 1800x per bp. The coverage was evenly distributed among the bp for all of the samples except for the $PolG^{mut/mut}$ mice. In these two samples there was a sharp reduction in coverage in the minor arc located between the two origins of replication. This decrease in coverage is most likely a consequence of the reported linear deleted molecules that constitute approximately 25-30 % of the mtDNA pool in the $PolG^{mut/mut}$ mice (Trifunovic, Wredenberg, Falkenberg, *et al.*, 2004). No difference in mutation load was observed between the minor arc and the major arc, implying that the mutation load in the deleted mtDNA molecules is similar to the non-deleted mtDNA molecules.

The mutation load was similar in WT and $PolG^{+/+}$ mice, with between 1.3-1.8*10^-4 mutations per bp. The $PolG^{mut/+}$ had a slightly increased mutation load compared to $PolG^{+/+}$ and WT mice, with between 3.3-3.8*10^-4 mutations per bp. The $PolG^{mut/mut}$ had a substantially higher mutation load compared to the other genotypes, with 11.7-12.5*10^-4 mutations per bp. We also investigated the presence of high frequency mutations (defined as mutations with a heteroplasmy level >0.5 %) in the different genotypes. Interestingly there was an equal number of high frequency mutations in $PolG^{+/+}$ and $PolG^{mut/+}$ mice, implying that these were inherited via the germline. WT animals had only half the number of high frequency mutations compared to $PolG^{+/+}$ mice. The mutation load in the $PolG^{mut/mut}$ mice was similar in protein coding genes, tRNAs and rRNAs. There was however a sharp reduction in mutation load within the control region. This was most pronounced in the CSBs, which had an 80 % reduction in mutation load compared to the coding regions. This implies that the control region is very sensitive to mutations. This is maybe not that surprising as OriH, LSP and HSP is located within this region. This finding goes in line with other reports, finding that some nucleotide positions within OriL are impossible to mutate due to their essential function in mtDNA replication (Wanrooij, Fusté, Stewart, *et al.*, 2012).

To elucidate the molecular origin of the mutations we mapped the number and ratio of transitions and transversions in the samples. Mutations generated by oxidative damage are expected to increase the number of transversions, exemplified by the common oxidative product 8-oxo-guanine, which generates G>T transversions when replicated

(Schmitt, Kennedy, Salk, *et al.*, 2012). We could not detect an increase in transversions by increasing age in any of the samples, implying that oxidative damage is not the major mutagenic force in mitochondria. This argues against the vicious cycle that has been proposed to drive somatic mutagenesis (Harman, 1955) and instead suggests that replication errors are the causative factor. There was a great variation in mutation accumulation between different nucleotide positions, with some positions experience 100-1000x more mutations than others, clearly showing the presence of mutational hotspots in mtDNA.

We did not to detect any increase in mutation load with increasing age in any of the samples. This finding could argue against a role of mtDNA mutations in the aging process. However, an alternative explanation is that this increase in mutation load was too low to be detected by our method. This could be due to at least three reasons. First, aging is associated with polyclonal expansion of many different low-level mtDNA mutations (Park & Larsson, 2011).The mean heteroplasmy level of a specific mutation therefore becomes very low in the tissue as a whole, making it difficult to detect by sequencing of an entire tissue. Secondly, we only sequenced liver DNA. The results might have been different with another tissue such as heart or skeletal muscle. Thirdly, the error rate might have been too high, obscuring a subtle increase in mutation load by increasing age. We tried to minimize the error rate by using the SOLiD platform and a control sample. We did however obtain a mutation load estimate for WT mice (1.3-1.8*10^-4 mutations per bp) that was an order of magnitude higher than estimates obtained by direct cloning (1*10^-5 mutations per bp, Sato, Nakada, Akimoto, *et al.*, 2005) which is supposed to be the most error free method. A methodological improvement that could have been applied is to score mutations as true only if they had been found on both strands. This was applied by Li, Schönberg, Schaefer, *et al*., 2010 and was reported to lower the error rate 30 fold.

ON mtDNA MUTATIONS AND INHERITANCE

In paper II we show that the heteroplasmy level of pathogenic mtDNA mutations located in tRNA genes is determined prenatally during early germline development. We also show that there is a skewed heteroplasmy level distribution in the offspring from mothers with high heteroplasmy levels, consistent with purifying selection against these mutations.

tRNAs constitute only 9 % of mtDNA in humans, but 58 % of the known pathogenic human mtDNA mutations are located in these genes (Stewart, Freyer, Elson, *et al.*,

2008). Pathogenic mutations tend to be heteroplasmic and functionally recessive. To cause mitochondrial disease the mtDNA mutation load often has to reach a certain heteroplasmy level (around 70-80 %) and once this threshold is passed the level of the pathogenic mutations determines the severity of the disease. The heteroplasmy level among children from the same mother can vary enormously. Hence, an unaffected mother carrying a pathogenic mutation can give birth to both healthy and severely sick children. To date there is no way to counsel prospective mothers whether their children will be healthy or develop mitochondrial disease. These rapid shifts in heteroplasmy between generations have been explained by the mitochondrial bottleneck (Olivo, Walle, Laipis, *et al.*, 1983). The bottleneck has been suggested to be caused by either a sharp reduction in mtDNA copy number (Cree, Samuels, de Lopes, *et al.*, 2008) or by replication of a subpopulation of mtDNAs during germline development (Wai, Teoli & Shoubridge, 2008). However, the precise timing of these changes remains controversial and it is not clear whether selection has a role during transmission of pathogenic mtDNA mutations. These issues are important for the genetic counseling of prospective mothers and for the development of treatments aimed at disease prevention. Unfortunately this is a difficult topic to study in humans and the lack of appropriate animal models has hindered progress in this field.

To address this important issue we generated a mouse model carrying a human equivalent pathogenic mtDNA mutation. This was done by backcrossing females homozygous for the mtDNA mutator allele ($PolG^{mut/mut}$) to WT males to generate maternal lineages carrying different spectrums of mtDNA mutations. One of the lineages carried two mutations in its mtDNA: a m5245T<C substitution in $tRNA^{Cys}$ and a m3875delC deletion in $tRNA^{Met}$. The m5245T<C mutation segregated to homoplasmy without any obvious phenotypic effects. The m3875delC mutation remained heteroplasmic with no offspring having a heteroplasmy level of >86 %, implying purifying selection against the mutation.

To investigate this in more depth we quantified heteroplasmy levels in 533 offspring born to 44 different mothers. If there is no selection and the mutations are transmitted by random genetic drift the mean heteroplasmy level in the offspring should be the same as the heteroplasmy of the mother and follow a Kimura distribution. However, offspring from mothers with a heteroplasmy level of >40 % showed a significant deviation from the Kimura distribution and neutral transmission. As the high heteroplasmy mothers had the same number of surviving offspring the lack of >86 % offspring was not due to elimination of sick pups. Hence, some sort of purifying

selection against the m3875delC mutation took place. To investigate this further we crossed the mouse line to a reporter mouse (*Stella-GFP*), which expresses GFP in the germ cells. This enables isolation of single germ cells via FACS sorting, generating the possibility to quantify heteroplasmy levels of single germ cells. We measured the heteroplasmy level in 819 PGCs from 18 embryos at 13.5 d.p.c. We found a huge variation in heteroplasmy level among the PGCs, with some cells carrying the m3875delC mutation in homoplasmic state. The distribution of mutations in the PGCs was in line with random genetic drift and was in stark contrast to the pattern observed in the offspring. We then studied the heteroplasmy level in 340 oocytes isolated from 5 different neonate animals at 3.5 days of age. Also here the distribution of heteroplasmy among the cells was in line with random genetic drift, indicating that no selection had taken place. As a final test to address when selection occurs, we calculated the heteroplasmy level difference between the mother and the PGCs, oocytes and the offspring. A negative correlation between the mother´s heteroplasmy level and the heteroplasmy level in the PGC, oocyte or offspring indicates selection. Only the offspring showed such a correlation.

The m3875delC mutation does cause a conformational change of the $tRNA^{Met}$, resulting in amino-acylation defects in heart and muscles of these mice. However, we could not detect any defects in mitochondrial respiration or enzymatic activities. This is might be explained by increased transcription and protein levels in these mice, masking the effects of the mutation. A similar compensatory effect has been seen in other mouse models with translational defects (Metodiev, Lesko, Park, *et al.*, 2009).

The results from this study show us two important things regarding mtDNA mutation inheritance and selection: (i) that the heteroplasmy level variation for tRNA mtDNA mutations is set early during germline development and (ii) that there is no selection against pathogenic tRNA mtDNA mutations in the germline, but rather in the offspring. A previous study (Stewart, Freyer, Elson, *et al.*, 2008) found that tRNAs accumulate a similar amount of mutations as third codon positions in protein coding genes, indicating relaxed selection in tRNA genes. Furthermore, as the majority of pathogenic mtDNA mutations are located in tRNA genes it has been assumed that there is no or very little selection against mutations in these genes. The findings in this study contradict this, showing clear signs of selection against the tRNA mutation. The finding that selection takes place in the offspring argues against the involvement of artresia in the purging of pathogenic mtDNA mutations. If atresia functions as a selective filter (removing

oocytes with pathogenic mtDNA mutations) we should see a skew in heteroplasmy distribution among the oocytes. This was not the case.

ON GERMLINE mtDNA MUTATIONS AND AGING

In paper III we show that low levels of maternally inherited mtDNA mutations can induce premature aging and reduced fertility in mice. These phenotypes can be reversed by introduction of WT mtDNA, clearly showing the role of inherited mtDNA mutations in inducing these phenotypes. Furthermore, we show that a combination of somatic mutagenesis and inherited mtDNA mutations can lead to the development of brain malformations.

Our group has previously shown that an increase in somatic mtDNA mutation load can cause the development of a progeroid phenotype in mice (Trifunovic, Wredenberg, Falkenberg, *et al*., 2004). However, what has been less clear is what role inherited mtDNA mutations have in the aging process. Furthermore, it has recently been shown that low level heteroplasmy is universal in humans and that a significant part of these mutations are inherited via the germline (Payne, Wilson, Yu-Wai-Man, *et al*., 2012). This finding makes elucidating the effect of inherited mtDNA mutations in aging a matter of great interest and importance.

To address this subject we used the mtDNA mutator mouse (Trifunovic, Wredenberg, Falkenberg, *et al.*, 2004; Kujoth, Hiona, Pugh, *et al.*, 2005). This knock in mouse model has an increased mtDNA mutation load due to a proofreading deficiency in *PolG* and develops a progeroid phenotype. We have shown that females heterozygous for the mutator allele ($PolG^{mut/+}$) transmit a significant amount of mtDNA mutations via the germline to the offspring (Ameur, Stewart, Freyer, *et al.*, 2011). We set up matings to generate mice that differed both in their mtDNA mutation load but also to what extent these mutations were inherited, generated somatically or both (Fig 3). We used the standard mutator mating (intercross between $PolG^{mut/+}$ male and female) to generate offspring with three different genotypes: type I mice ($PolG^{+/+}$), type II mice ($PolG^{mut/+}$) and type III mice ($PolG^{mut/mut}$). Type I, II and III mice all carry mtDNA mutations inherited via the germline, in addition to the somatically generated mtDNA mutations in type II and III mice (type I lacks somatic mutations as it carries WT nDNA). (Type I-III animals are the same genotypes as in paper I, but due to the great number of different mice used in this paper we had to rename them). We also set up a second mating (between a $PolG^{mut/+}$ males and a $PolG^{+/+}$ females) to generate offspring with the following genotypes: type IV mice ($PolG^{+/+}$) and type V mice ($PolG^{mut/+}$). Both type

IV and V mice lack germline inherited mtDNA mutations and only carry somatic mutations. Finally we set up a mating between a $PolG^{mut/+}$ male and a $PolG^{KO/+}$ female to generate type VI mice ($PolG^{mut/KO}$). Type VI mice are functionally homozygous for the mutator allele but lack germline inherited mutations.

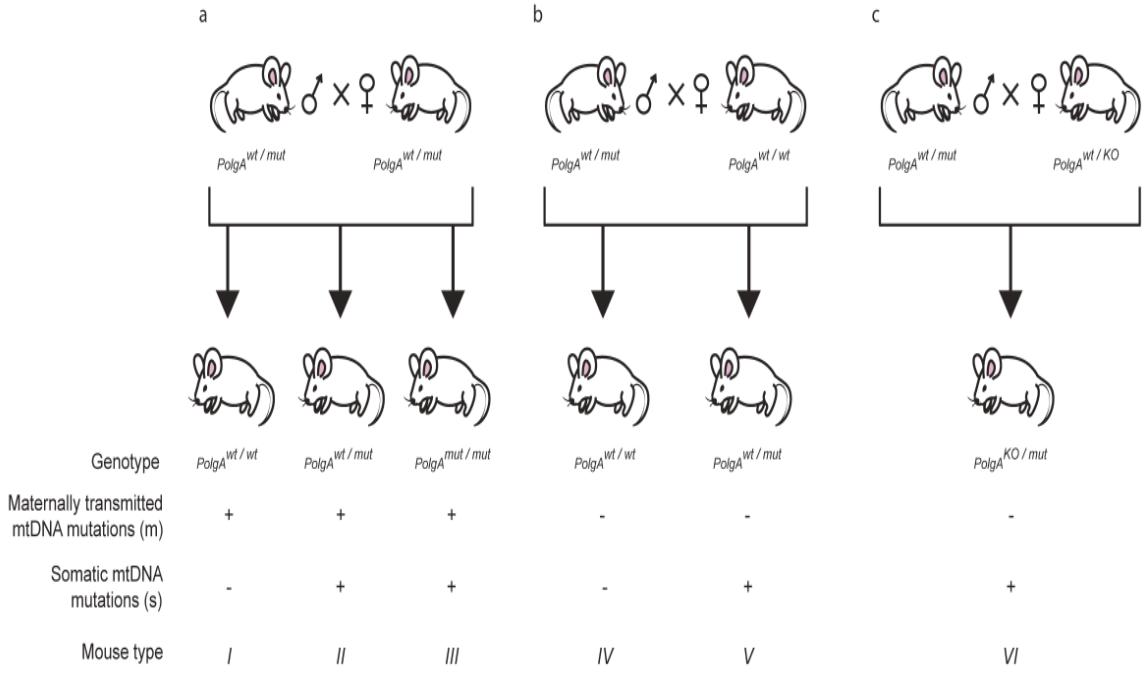

Figure 3. Showing the crosses used to generate mice with different genotypes for PolG. The mice will differ in mtDNA mutation load but also to what extent mutations are inherited via the germline, generated somatically or both (see list in figure). The genotypes generated from cross **a** were used in paper I and the genotypes generated from cross **a, b** and **c** were used in paper III. **a**, intercross of mice heterozygous for the mutator allele ($PolG^{mut/+}$) to generate type I-III mice. These mice all have mutations passed on via the germline but differ in somatic mutagenesis. **b**, cross between $PolG^{mut/+}$ male and $PolG^{+/+}$ female. The offspring from this cross will lack germline inherited mutations. This was the mating strategy used in paper III to reintroduce WT mtDNA. **c**, cross between a $PolG^{mut/+}$ male and $PolG^{wt/KO}$ female generating $PolG^{mut/KO}$ offspring. These mice are functionally homozygous for the mutator allele but lack germline inherited mutations. (Figure from Ross et al. 2013, Nature doi 10.1038/nature12474).

Continuous intercrossing of $PolG^{mut/+}$ animals for several generations resulted in reduced fertility in the form of fewer and smaller litters and less $PolG^{mut/mut}$ mice in the offspring. However, when WT mtDNA was re-introduced to the mice (by using a $PolG^{+/+}$ female, Fig. 3, cross b) these phenotypes was rescued and almost reverted to WT levels. Continuous intercrossing of mice without re-introduction of WT mtDNA thus leads to a progressive decline in fertility. Reintroducing WT mtDNA via the male did not rescue these phenotypes, clearly showing the role of mtDNA mutations in

inducing these phenotypes. To further investigate the role of germline inherited mtDNA mutations in the aging process we compared the phenotypes of three different types of $PolG^{+/+}$ mice: wt^N-1 mice (generated from an intercross using a $PolG^{mut/+}$ female, thus carrying mtDNA mutations inherited via the maternal germline), wt^N-2 (generated from an intercross using a mother with reintroduced WT mtDNA, thus carrying mtDNA with mutation load close to WT levels) and finally wt^N-3 mice (generated from a intercross between a $PolG^{mut/+}$ female and a male with reintroduced WT mtDNA, thus having same mtDNA mutation load levels as the wt^N-1 mice). The wt^N-3 mice function as controls to ensure that any observed phenotypic differences between the wt^N-1 and wt^N-2 mice are due to mutations in mtDNA and not due to the influence of nDNA. Wt^N-1 and wt^N-3 mice showed significant weight reduction and signs of premature aging at 65 weeks of age. At this time point wt^N-2 mice showed virtually no signs of aging. However, continuous intercrossing of type wt^N-2 mice to $PolG^{mut/+}$ mice resulted in a similar phenotype as the wt^N-1 and wt^N-3 mice. The similar phenotypes in wt^N-1and wt^N-3 mice show the role of mtDNA mutations in the development of these phenotypes.

To investigate the genetic and molecular basis of the detected phenotypes we sequenced mtDNA from several animals of the respective genotypes. We applied two different strategies to obtain mtDNA sequence data: (i) PCR-and cloning and (ii) λ-phage direct cloning. The PCR-and cloning approach has been criticized for introducing artifacts as it tends to generate higher mutation load estimates than other methods (Greaves, Beadle, Taylor, *et al*., 2009; Alex, Kraytsberg, Khrapko, *et al*., 2007). The problem with the cloning procedure is that it renders artifacts introduced by the PCR polymerase indistinguishable from true mutations .To correct for this we also used direct cloning by λ-phage to estimate the mutation load and spectrum. Direct cloning is supposed to be the methodological approach with the lowest error rate (Alex, Kraytsberg, Khrapko, *et al*., 2007; Edgar, Larsson & Trifunovic, 2010). The estimates from the two different methods were in good concordance for all genotypes, implying that artifacts do not confound our data. The sequence data showed a significant increase of clonally expanded mtDNA mutations in type I, II and III animals, compared to type IV, V and VI mice. Hence, mutations inherited via the germline expand clonally, most likely as a consequence of passing through the mitochondrial bottleneck. To study this further, we generated a single burst of mutations in the female germline by using $PolG^{mut/+}$ females. We crossed the offspring to wt animals and sequenced mtDNA from several consecutive generations. From generation N2 and onwards mutations were

detectable from bulk sequencing of a tissue (this is a consequence of that the mutations had expanded clonally to higher and detectable levels). It is also from the N2 generation and onwards that mutations start to generate a phenotype in the both the $PolG^{mut/+}$ and wt^{N}-2 mice. This clearly shows the importance of clonal expansion for mtDNA mutations to have a phenotypic effect.

We continued to investigate the role of inherited and somatically generated mtDNA mutations in the aging process by detailed phenotyping of PolG- $^{mut/mut}$ and $PolG^{mut/KO}$ mice. We compared the mean lifespan of $PolG^{mut/mut}$ mice generated from a standard cross and from a cross where the mother carried re-introduced wt mtDNA. The latter mice had a mean lifespan that was approximately 10 weeks longer than the former, clearly showing the influence of inherited mtDNA mutations in aging. We also did a detailed phenotyping of type III and type VI mice. At 35 weeks of age they both showed weight loss, reduced hemoglobin concentration and erythrocyte count. However they also differed in some other important aspects. Type III had more pronounced heart enlargement than type VI mice. Heart cells in type III mice also showed reduced COX enzyme activity. This was also true for colonic crypts, where the number of COX deficient cells was significantly higher in type III mice than in type VI mice. The finding that mice with inherited mtDNA mutations have more COX deficient cells makes sense in light of the results from sequence data. It shows that inherited mtDNA mutations can expand clonally, to create heteroplasmy levels high enough to cause a biochemical defect in cells.

Furthermore, 32 % of type III mice developed stochastic brain malformations. These were not present in any of the other types of mice, showing that a dual hit of inherited mtDNA mutations and somatic mutagenesis is required for these brain malformations to occur.

ON mtDNA AND RECOMBINATION

In paper IV we show that mtDNA in the mammalian germline does not undergo recombination and that mtDNA is inherited clonally. We could also show that a new direct cloning protocol we have developed can trap single whole mtDNA molecules without biasing towards specific molecules or introducing artifacts. By using this direct cloning approach to isolate single mtDNA molecules from a long term heteroplasmic mouse line were could address the important question of mtDNA recombination in the mammalian germline without suffering from the common limitations of direct test for recombination.

mtDNA is assumed to be inherited in a clonal fashion, lacking germline recombination. This trait has made mutations in mtDNA the most widely used marker in population genetic studies. However, clonal inheritance and lack of recombination in mtDNA has now been questioned in a number of studies (Zsurka, Kraytsberg, Kudina, *et al.*, 2005; Kraytsberg, Schwartz, Brown, *et al*., 2004; Fan, Lin, Potluri, *et al*., 2012; D'Aurelio, Gajewski & Lin, 2004; Ujvari, Dowton & Madsen, 2007; Awadalla, Eyre-Walker & Smith, 1999; Hagelberg, Goldman, Lió, *et al*., 1999) reporting either the identification of recombinant mtDNA molecules or signs of recombination in population genetic mtDNA sequence data. These findings, if true, could have profound effects on the interpretation of phylogenetic studies as even a limited amount of recombination has been shown to affect coalescent time estimates and branch lengths in phylogenetic trees (Schierup & Hein, 2000). However, there are two general problems with mtDNA recombination studies: (i) they mostly cover somatic events which are irrelevant to evolution and (ii) they often use PCR, thus introducing the risk of creating artificial recombinants. We were able to circumvent these problems by direct cloning of mtDNA molecules from a longterm heteroplasmic mouse line. This mouse line used was generated by cytoplasmic transfer >20 years ago (Jenuth, Peterson, Fu, *et al.*, 1996) and is artificially heteroplasmic for the NZB and BALB/c mtDNA genotypes. These genotypes differ at 90 nucleotide positions, which are evenly distributed in the mitochondrial genome. As this mouse line has been heteroplasmic for a defined set of mutations for >50 generations, it allows us to use present genetic variation to screen for recombination during this entire time period. Hence, we could use a direct test to screen for mtDNA recombination in the germline without suffering from the limitations of genetic drift and lineage extinction. Furthermore, the even distribution of the 90 SNP where the NZB and BALB/c genotypes differ should make it possible to differentiate between homoplasies and recombination. Homoplasies should result in single base pair similarities and recombination in multiple similarities distributed over a longer stretch of DNA.

Direct tests of recombination demand the isolation and analysis of single mtDNA molecules. This is most often obtained by cloning of PCR products. This approach is however highly unsuitable for recombination studies as PCR has been shown to create artificial recombinant molecules *in vitro* (Pääbo, Irwin & Wilson, 1990; Keohavong & Thilly, 1989) and the cloning step makes the distinction between artefacts and true recombinants impossible. We could in this study also show frequent creation of artificial recombinants by PCR. We have previously applied a standard PCR-Clone-

Sequence protocol to detect mtDNA mutations in C57BL/6N mice (Wanrooij, Fusté, Stewart, *et al.*, 2012). From this available dataset we extracted sequences that had been filtered due to similarity with nuclear mtDNA pseudogenes (NuMts). We identified NuMts in 41 of 3940 clones (~1% of samples). Nine of the clones showed evidence of 8 independent PCR-induced recombination events between NuMts and mtDNA. This high frequency was present despite the very low abundance of nuclear DNA relative to mtDNA in the samples, clearly showing how unsuitable PRC and cloning is for recombination studies. A more stringent methodological approach that has been applied in one study is direct cloning of parts of mtDNA molecules (Sato, Nakada, Akimoto, *et al.*, 2005). Direct cloning omits PCR amplification prior to cloning and thus circumvents the problem of PCR induced artificial recombinants. However, a problem with the Sato, Nakada, Akimoto, *et al.*, 2005 study is that almost exclusively investigates somatic recombination. We wanted to apply direct cloning to screen for mtDNA recombination in the germline of mammals. For this purpose we utilized the λ-phage. After setting up a well functional protocol that enabled us to clone whole mtDNA molecules we carried out method verification experiments. We experimentally tested that the method was unbiased towards which molecules that were being cloned, that it had a low error rate and that it did not create artificial recombinants molecules.

To screen for recombination we extracted mtDNA from liver and kidney from two NZB-BALB/c animals and cloned it into the λ-phage, followed by RFLP analysis of 2992 molecules. Each genotype generates a specific cleavage pattern in the RFLP analysis. We therefore screened for molecules generating a mixed pattern, indicative of recombination. We detected one molecule showing a mixed NZB and BALB/c cleavage pattern. However upon sequencing the entire molecule only one single nucleotide position was changed (m15657C > T) indicating that this change was a homoplasy, rather than a recombination event. In total we screened 70,128 bp of mtDNA by RFLP and identified a single mutation. This gives a mutation load of $\sim 1.43 \times 10^{-5}$ mutations per bp. This mutation load is in good concordance with the mutation load estimate obtained by PCR-clone-sequencing and λ-phage cloning (unpublished data). Furthermore we sequenced 8 NZB and 9 BALB/c molecules. This did not reveal any recombinant molecules.

In summary, we used direct cloning of whole mtDNA molecules from a long term heteroplasmic mouse line to show that mtDNA in the germline of mammals does not undergo recombination under biological and evolutionary relevant conditions and that mtDNA is inherited in a clonal manner.

ACKNOWLEDGEMENTS

MANY THANKS TO:

Lab members

Nils-Göran Larsson for giving me the opportunity to join your group. Being given the opportunity to work in a dynamic and world leading research group has really inspired me a developed me as a scientist. It has also taught me that nothing in research comes easy and that a lot of hard work is behind every finding and paper. So again, many thanks!

Aleksandra Trifunovic for technical advice in lab work.
Claes Gustafsson for nice talks about everything from science, movies to dogs.
Maria Falkenberg for technical advice in lab work and nice chats.
Jim Stewart for discussions and support with projects, but also for manuscript writing and sequencing.
Chris Freyer for support and collaboration with projects, but also for manuscript writing.
Anna Wredenberg for nice chats every now and then.
Fredrik Sterky for company and discussions about everything from diving, fishes to renovation of apartments.
Camilla Koolmeister for good company in the lab and discussions about children and parenting, but also for all your help with paperwork.
Ana Bratic for being a supportive PhD student (now post doc!).
Henrik Spåhr for good office company when the lab was still in Huddinge.
Seungmin Lee for Korean sweets and company.
Adya Misra for always being close to a smile but also help with proof-reading when my English fell short.
Benedetta Ruzzenente for your sunny personality. Will miss your excellent presentations at the retreats =)
Mugen Terzioglu for nice company and chats about literature and politics.
Chris Kukat for always being up for a chat.

People working in the group of Aleksandra Trifunovic, especially **Daniel Edgar** for long interesting discussion about science, skepticism, politics, children and everything between.

People working in the groups of Claes Gustafsson and Maria Falkenberg, especially **Bertil Macao** for our nice chats over a beer(s) at retreats,

In the group of Anna Wredenberg and Chris Freyer; **Paula Clemente** for good company and Spanish lessons. Also many thanks for helping out with the figures. **Alberto Cruz Bermudes** for nice chats every now and then.

Lovisa Case for being a great “corridor mate” to discussion everything from movies, cats, food, wine and the life of being a PhD student with, **Alexander Berg** and **Stefan Plantman** nice coffee breaks.

Family and friends

My son **Alvin** for being the light of my life, making waking up in the morning and coming home from work a pure joy.

My love **Caroline** for love and support during good and hard times.

My **parents** for support during all these years and for always having my back.

My cousin **David Blomqvist** for being my cousin.

My best and childhood friend **Mikael Rödin** for always being there.

My close friends **Ingemar Boström** for amazing fishing trips around Sweden (No guts no glory!) and **Nicklas Edvardsson** for arranging excellent movie nights (Tarkan versus the Vikings) with beer tasting.

REFERENCES

Alex, Y.K., Kraytsberg, Y., Khrapko, E.N.K., Nicholas, A., et al. (2007) Are somatic mitochondrial DNA mutations relevant to our health? A challenge for mutation analysis techniques. Expert Opinion on Medical Diagnostics. 1, 109-116.

Ameur, A., Stewart, J.B., Freyer, C., Hagström, E., et al. (2011) Ultra-Deep Sequencing of Mouse Mitochondrial DNA: Mutational Patterns and Their Origins. PLoS genetics. 7, e1002028.

Andersson, S.G., Karlberg, O., Canbäck, B. & Kurland, C.G. (2003) On the origin of mitochondria: a genomics perspective. Philosophical transactions of the Royal Society of London. Series B, Biological sciences. 358, 165-77

Awadalla, P., Eyre-Walker, A. & Smith, J. (1999) Linkage disequilibrium and recombination in hominid mitochondrial DNA. Science. 286, 2524-2525.

Bailey, L.J., Cluett, T.J., Reyes, A., Prolla, T.A., et al. (2009) Mice expressing an error-prone DNA polymerase in mitochondria display elevated replication pausing and chromosomal breakage at fragile sites of mitochondrial DNA. Nucleic Acids Research. 37, 2327-2335.

Bandelt, H. (2005) More evidence for non-maternal inheritance of mitochondrial DNA? Journal of Medical Genetics. 42, 957-960.

Battersby, B.J., Loredo-Osti, J.C. & Shoubridge, E.A. (2003) Nuclear genetic control of mitochondrial DNA segregation. Nature Genetics. 33, 183-186.

Bielas, J.H. & Loeb, L.A. (2005) Quantification of random genomic mutations. Nature Methods. 2, 285-290..

Bigger, B. (2003) Efficient cloning and engineering of entire mitochondrial genomes in Escherichia coli and transfer into transcriptionally active mitochondria. Nucleic Acids Research. 31, 1407-1415.

Bigger, B., Tolmachov, O., Collombet, J. & Coutelle, C. (2000) Introduction of chloramphenicol resistance into the modified mouse mitochondrial genome: cloning of unstable sequences by passage through yeast. Analytical biochemistry. 277, 236-242.

Bogenhagen, D. & Da Clayton (1977) Mouse L cell mitochondrial DNA molecules are selected randomly for replication throughout the cell cycle. Cell. 11, 719-727.

Bogenhagen, D.F. (1999) Repair of mtDNA in Vertebrates. The American Journal of Human Genetics. 64, 1276-1281.

Boyd, J. (1951) 'Excessive Dose'Phenomenon in Virus Infections. Nature. 167, 1061-1062.

Bua, E., Johnson, J., Herbst, A., Delong, B., et al. (2006) Mitochondrial DNA–deletion mutations accumulate intracellularly to detrimental levels in aged human skeletal muscle fibers. Journal of Human Genetics. 79, 469-480.

Burger, G., Gray, M.W. & Lang, B.F. (2003) Mitochondrial genomes: anything goes. Trends in Genetics. 19 , 709-716.

Chen, H. & Chan, D. (2004) Mitochondrial dynamics in mammals. Current topics in developmental biology. 59, 119-144.

Chen, H. & Chan, D.C. (2010) Physiological functions of mitochondrial fusion. Annals of the New York Academy of Sciences. 1201, 21-25.

Chinnery, P., Johnson, M., Wardell, T., Singh-Kler, R., et al. (2000) The epidemiology of pathogenic mitochondrial DNA mutations. Annals of neurology. 48, 188-193.

Clayton (1991) Replication and transcription of vertebrate mitochondrial DNA. Annual review of cell biology. 7, 453-478.
Cree, L.M., Samuels, D.C., de Lopes, S.C., Rajasimha, H.K., et al. (2008) A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes. Nature Genetics. 40, 249-254.
D'Aurelio, M., Gajewski, C. & Lin, M. (2004) Heterologous mitochondrial DNA recombination in human cells. Human molecular genetics. 13, 3171-3179.
Drake, J.W., Charlesworth, B., Charlesworth, D. & Crow, J.F. (1998) Rates of Spontaneous Mutation. Genetics. 148, 1667-1686.
Dubeau, F., de Stefano, N., Zifkin, B., DL, A., et al. (2000) Oxidative phosphorylation defect in the brains of carriers of the tRNAleu (UUR) A3243G mutation in a MELAS pedigree. Annals of neurology. 47, 179-185.
Dyall, S.D. (2004) Ancient Invasions: From Endosymbionts to Organelles. Science. 304, 253-257.
Edgar, D. & Trifunovic, A. (2009) The mtDNA mutator mouse: Dissecting mitochondrial involvement in aging. Aging. 1, 1028-1032.
Edgar, D., Larsson, N. & Trifunovic, A. (2010) Point mutations are causing progeroid phenotypes in the mtDNA mutator mouse. Cell Metabolism. 11, 1.
Edgar, D., Shabalina, I., Camara, Y., Wredenberg, A., et al. (2009) Random point mutations with major effects on protein-coding genes are the driving force behind premature aging in mtDNA mutator mice. Cell Metabolism. 10, 131-138.
Elliott, H., Samuels, D., Eden, J., Relton, C., et al. (2008) Pathogenic mitochondrial DNA mutations are common in the general population. The American journal of human genetics. 80, 254-260.
Elson, J. & Lightowlers, R. (2006) Mitochondrial DNA clonality in the dock: can surveillance swing the case? Trends in Genetics. 22, 603-607.
Elson, J.L., Andrews, R.M., Chinnery, P.F., Lightowlers, R.N., et al. (2001) Analysis of European mtDNAs for recombination. American journal of human genetics. 68, 145-153.
Elson, J.L., Turnbull, D.M. & Howell, N. (2004) Comparative Genomics and the Evolution of Human Mitochondrial DNA: Assessing the Effects of Selection. The American Journal of Human Genetics. 74, 229-238.
Fan, W., Lin, C.S., Potluri, P., Procaccio, V., et al. (2012) mtDNA lineage analysis of mouse L-cell lines reveals the accumulation of multiple mtDNA mutants and intermolecular recombination. Genes & Development. 26, 384-394.
Fan, W., Waymire, K.G., Narula, N., Li, P., et al. (2008) A Mouse Model of Mitochondrial Disease Reveals Germline Selection Against Severe mtDNA Mutations. Science. 319, 958-962.
Foury, F., S & Vanderstraeten (1992) Yeast mitochondrial DNA mutators with deficient proofreading exonucleolytic activity. The EMBO journal. 11, 2717-2726.
Freyer, C., Cree, L., Mourier, A., Stewart, J., et al. (2012) Variation in germline mtDNA heteroplasmy is determined prenatally but modified during subsequent transmission. Nature genetics. 44, 1282-1285.
Fusté, J.M., Wanrooij, S., Jemt, E., Granycome, C.E., et al. (2010) Mitochondrial RNA Polymerase Is Needed for Activation of the Origin of Light-Strand DNA Replication. Molecular Cell. 37, 67-78.
Giles, R., Blanc, H. & Cann, H. (1980) Maternal inheritance of human mitochondrial DNA. Proceedings of the National Academy of Sciences. 77, 6715-6719.

Gray, M.W., Burger, G. & Lang, B.F. (1999) Mitochondrial evolution. Science (New York, N.Y.). 283, 1476-1481.
Graziewicz, M.A., Longley, M.J. & Copeland, W.C. (2006) DNA Polymerase ç in Mitochondrial DNA Replication and Repair. Chemical Reviews. 106, 383-405.
Greaves, L., Elson, J., Nooteboom, M. & Grady, J. (2012) Comparsion of mitochondrial mutation spectra in ageing human colonic epithelium and disease: Absence of evidence for purifying selection in somatic mitochondrial DNA point mutations. PLoS genetics. 8, 1-10 , e1003082.
Greaves, L.C., Beadle, N.E., Taylor, G.A., Commane, D., et al. (2009) Quantification of mitochondrial DNA mutation load. Aging cell. 8, 566-572.
Gyllensten, U., Wharton, D., Josefsson, A. & Wilson, A. (1991) Paternal inheritance of mitochondrial DNA in mice. Nature. 352, 255-257.
Hagelberg, E., Goldman, N., Lió, P., Whelan, S., et al. (1999) Evidence for mitochondrial DNA recombination in a human population of island Melanesia. Proceedings. Biological sciences / The Royal Society. 266, 485-492.
Harman, D. (1955) Aging: a theory based on free radical and radiation chemistry. Journal of Gerontology.11, 298-300
Holt, I., Harding, A. & Morgan-Hughes, J. (1988) Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. Nature. 331, 717-719.
Holt, I.J., Lorimer, H.E. & Jacobs, H.T. (2000) Coupled Leading- and Lagging-Strand Synthesis of Mammalian Mitochondrial DNA. Cell. 100, 515-524.
Inoue, K., Nakada, K., Ogura, A., Isobe, K., et al. (2000) Generation of mice with mitochondrial dysfunction by introducing mouse mtDNA carrying a deletion into zygotes. Nature genetics. 26, 176-181.
Jenuth, J., Peterson, A. & Shoubridge, E. (1997) Tissue-specific selection for different mtDNA genotypes in heteroplasmic mice. Nature genetics. 16, 93-95.
Jenuth, J.P., Peterson, A.C., Fu, K. & Shoubridge, E.A. (1996) Random genetic drift in the female germline explains the rapid segregation of mammalian mitochondrial DNA. Nature genetics. 14, 146-151.
Kaguni, L. & Olson, M. (1989) Mismatch-specific 3‘----5’exonuclease associated with the mitochondrial DNA polymerase from Drosophila embryos. In: Proceedings of the National Academy of Sciences. 86, 6469-6473
Kaneda, H., Hayashi, J. & Takahama, S. (1995) Elimination of paternal mitochondrial DNA in intraspecific crosses during early mouse embryogenesis. In*:* Proceedings of the National Academy of Sciences. 92, 4542-4546.
Kasiviswanathan, R., Collins, T.R.L., Copel, W.C., Collins, T.R., et al. (2012) The interface of transcription and DNA replication in the mitochondria. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms. 1819, 970-978.
Keohavong, P. & Thilly, W.G. (1989) Fidelity of DNA polymerases in DNA amplification. Proceedings of the National Academy of Sciences Academy of Sciences. 86, 9253-9257.
Kinde, I., Wu, J., Papadopoulos, N., Kinzler, K.W., et al. (2011) Detection and quantification of rare mutations with massively parallel sequencing. Proceedings of the National Academy of Sciences Academy of Sciences.108, 9530-9535.
Kirkwood, T. (2005) Understanding the odd science of aging. Cell. 120437-447.
Kivisild, T. & Villems, R. (2000) Questioning evidence for recombination in human mitochondrial DNA. Science. 288,1931a.

Kraytsberg, Y. & Khrapko, K. (2005) Single-molecule PCR: an artifact-free PCR approach for the analysis of somatic mutations. Expert review of molecular diagnostics. 5, 809-815.
Kraytsberg, Y., Schwartz, M., Brown, T.A., Ebralidse, K., et al. (2004) Recombination of human mitochondrial DNA. Science (New York, N.Y.). 304, 981.
Kujoth, G., Hiona, A., Pugh, T., Someya, S., 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., et al. (2011) Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA. Proceedings of the National Academy of Sciences Academy of Sciences. 108, 13534-13539.
Kuznetsov, A.V., Hermann, M., Saks, V., Hengster, P., et al. (2009) The cell-type specificity of mitochondrial dynamics. The International Journal of Biochemistry & Cell Biology. 41, 1928-1939.
Ladoukakis, E.D. & Eyre-Walker, A. (2004) Evolutionary genetics: Direct evidence of recombination in human mitochondrial DNA. Heredity. 93, 321-321.
Ladoukakis, E.D. & Zouros, E. (2001) Direct Evidence for Homologous Recombination in Mussel (Mytilus galloprovincialis) Mitochondrial DNA. Molecular Biology and Evolution. 18, 1168-1175.
Lane, N. (2007) Origin of mitochondria and hydrogenosomes. Spinger-Verlag Berlin Heidelberg.
Larsson, N. (2010) Somatic Mitochondrial DNA Mutations in Mammalian Aging. Annual Review of Biochemistry. 79, 683-706.
Li, M., Schönberg, A., Schaefer, M., Rol, et al. (2010) Detecting Heteroplasmy from High-Throughput Sequencing of Complete Human Mitochondrial DNA Genomes. The American Journal of Human Genetics. 87, 237-249.
Longley, M.J. (2001) The Fidelity of Human DNA Polymerase gamma with and without Exonucleolytic Proofreading and the p55 Accessory Subunit. Journal of Biological Chemistry. 276, 38555-38562.
McFarland, R., Taylor, R. & Turnbull, D. (2010) A neurological perspective on mitochondrial disease. The Lancet Neurology. 9, 829-840.
Metodiev, M.D., Lesko, N., Park, C.B., Yol, et al. (2009) Methylation of 12S rRNA Is Necessary for In Vivo Stability of the Small Subunit of the Mammalian Mitochondrial Ribosome. Cell Metabolism. 9, 386-397.
Murray, N. & Murray, K. (1974) Manipulation of restriction targets in phage lambda to form receptor chromosomes for DNA fragments. Nature. 251, 476-481.
Müller-Höcker, J. (1990) Cytochrome c oxidase deficient fibres in the limb muscle and diaphragm of man without muscular disease: An age-related alteration. Journal of the neurological sciences. 100, 14-21.
Müller-Höcker, J. (1989) Cytochrome-c-oxidase deficient cardiomyocytes in the human heart--an age-related phenomenon. A histochemical ultracytochemical study. The American journal of pathology. 134, 1167-1173.
Nabholz, B., Ellegren, H. & Wolf, J. (2013) High levels of gene expression explain the strong evolutionary constraint of mitochondrial protein-coding genes. Molecular Biology and Evolution. 30, 272-284.
Nachman, M. (1998) Deleterious mutations in animal mitochondrial DNA. Mutation and Evolution. 102, 10361-10369.

Nachman, M., Brown, W. & Stoneking, M. (1996) Nonneutral mitochondrial DNA variation in humans and chimpanzees. Genetics. 142, 953-963.
Nakada, K., Sato, A. & Hayashi, J. (2009) Mitochondrial functional complementation in mitochondrial DNA-based diseases. The International Journal of Biochemistry & Cell Biology. 41, 1907-1913.
Nunnari, J. & Suomalainen, A. (2012) Mitochondria: In Sickness and in Health. Cell. 148, 1145-1159.
Olivo, P., Walle, M., Laipis, P. & Hauswirth, W. (1983) Nucleotide sequence evidence for rapid genotypic shifts in the bovine mitochondrial DNA D-loop. 306, 400-402.
Pak, J., Herbst, A., Bua, E., Gokey, N., et al. (2003) Mitochondrial DNA mutations as a fundamental mechanism in physiological declines associated with aging. Aging cell. 2, 1-7.
Park, C.B. & Larsson, N.-. (2011) Mitochondrial DNA mutations in disease and aging. The Journal of Cell Biology. 193, 809-818.
Parsons, T., Muniec, D. & Sullivan, K. (1997) A high observed substitution rate in the human mitochondrial DNA control region. Nature. 59, 363-368.
Pavlov, A.R., Pavlova, N.V., Kozyavkin, S.A. & Slesarev, A.I. (2004) Recent developments in the optimization of thermostable DNA polymerases for efficient applications. Trends in Biotechnology. 22, 253-260.
Payne, B., Wilson, I., Yu-Wai-Man, P., Coxhead, J., et al. (2013) Universal heteroplasmy of human mitochondrial DNA. Human Molecular Genetics. 22, 384–390.
Piganeau, G. & Eyre-Walker, A. (2004) A reanalysis of the indirect evidence for recombination in human mitochondrial DNA. Heredity. 92, 282-288.
Pohjoismäki, J.L. & Goffart, S. (2011) Of circles, forks and humanity: Topological organisation and replication of mammalian mitochondrial DNA. BioEssays. 33, 290-299.
Pääbo, S., Irwin, D.M. & Wilson, A.C. (1990) DNA damage promotes jumping between templates during enzymatic amplification. The Journal of biological chemistry. 265, 4718-4721.
Rizzuto, R., de Stefani, D.D., Raffaello, A., Mammucari, C., et al. (2012) Mitochondria as sensors and regulators of calcium signalling. Nature Reviews Molecular Cell Biology.13, 566-578.
Rokas, A., Ladoukakis, E. & Zouros, E. (2003) Animal mitochondrial DNA recombination revisited. Trends in Ecology & Evolution. 18, 411-417.
Sacconi, S., Salviati, L., Nishigaki, Y., Walker, W.F., et al. (2008) A functionally dominant mitochondrial DNA mutation. Human Molecular Genetics. 17, 1814-1820.
Sanger, F., Coulson, A., Hong, G., Hill, D., et al. (1982) Nucleotide sequence of bacteriophage λ DNA. Journal of molecular Biology. 162, 729-773.
Sato, A., Nakada, K., Akimoto, M., Ono, T., et al. (2005) Rare creation of recombinant mtDNA haplotypes in mammalian tissues. In: Proceedings of the National Academy of Sciences Academy of Sciences. 102, 6057–6062.
Schaefer, A., McFarland, R., Blakely, el, He, L., et al. (2008) Prevalence of mitochondrial DNA disease in adults. Annals of neurology. 63, 35-39.
Schierup, M.H. & Hein, J. (2000) Consequences of recombination on traditional phylogenetic analysis. Genetics. 156, 879-891.
Schmitt, M.W., Kennedy, S.R., Salk, J.J., Fox, E.J., et al. (2012) Detection of ultra-rare mutations by next-generation sequencing. Proceedings of the National Academy of Sciences Academy of Sciences. 109, 14508-14513.

Seo, A.Y., Joseph, A., Dutta, D., Hwang, J.C., et al. (2010) New insights into the role of mitochondria in aging: mitochondrial dynamics and more. Journal of Cell Science. 123, 2533-2542.
Shoubridge, E. & Wai, T. (2007) Mitochondrial DNA and the mammalian oocyte. Current topics in developmental biology. 77, 87-111.
Shutt, T. & Gray, M. (2006) Bacteriophage origins of mitochondrial replication and transcription proteins. Trends in Genetics. 22, 90-95.
Smigrodzki, R., Parks, J. & Parker, W.D. (2004) High frequency of mitochondrial complex I mutations in Parkinson's disease and aging. Neurobiology of Aging. 25, 1273-1281.
Smith, D.J., Ng, H., Kluck, R.M. & Nagley, P. (2008) The mitochondrial gateway to cell death. IUBMB Life. 60, 383-389.
Stewart, J.B., Freyer, C., Elson, J.L., Wredenberg, A., et al. (2008) Strong Purifying Selection in Transmission of Mammalian Mitochondrial DNA. PLoS Biology. 6, e10..
Stoneking, M. (2000) Hypervariable Sites in the mtDNA Control Region Are Mutational Hotspots. The American Journal of Human Genetics. 67, 1029-1032..
Sutovsky, P., DMoreno, R., Ramalho-Santos, J., Dominko, T., et al. (2000) Ubiquitinated Sperm Mitochondria, Selective Proteolysis, and the Regulation of Mitochondrial Inheritance in Mammalian Embryos. Biology of reproduction. 63, 582-590.
Taylor, R., Barron, M., Borthwick, G., A, G., et al. (2003) Mitochondrial DNA mutations in human colonic crypt stem cells. Journal of Clinical Investigation. 112, 1351-1360.
Taylor, R.W. & Turnbull, D.M. (2005) Mitochondrial DNA mutations in human disease. Nature Reviews Genetics. 6, 389-402.
Thundathil, J., Filion, F. & Smith, L.C. (2005) Molecular control of mitochondrial function in preimplantation mouse embryos. Molecular Reproduction and Development. 71, 405-413.
Tolmachov, O. & Collombet, J.M. (2012) Direct mutation analysis by high-throughput sequencing: From germline to low-abundant, somatic variants. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 729, 1-15.
Trifunovic, A., Hansson, A., AT, R., E, D., et al. (2005) Somatic mtDNA mutations cause aging phenotypes without affecting reactive oxygen species production. Proceedings of the National Academy of Sciences. 102, 17993-17998.
Trifunovic, A., Wredenberg, A., Falkenberg, M., JN, S., et al. (2004) Premature ageing in mice expressing defective mitochondrial DNA polymerase. Nature. 429, 417-423.
Trounce, I., Byrne, E. & Marzuki, S. (1989) Decline in skeletal muscle mitochondrial respiratory chain function: possible factor in ageing. The Lancet. 25, 637-639.
Tully, L., Parsons, T., steighner, R., Holland, M., et al. (2000) A Sensitive Denaturing Gradient-Gel Electrophoresis Assay Reveals a High Frequency of Heteroplasmy in Hypervariable Region 1 of the Human mtDNA Control Region. The American Journal of Human Genetics. 67, 432-443.
Tyynismaa, H., Mjosund, K.P., Wanrooij, S., Lappalainen, I., et al. (2005) Mutant mitochondrial helicase Twinkle causes multiple mtDNA deletions and a late-onset mitochondrial disease in mice. Proceedings of the National Academy of Sciences Academy of Sciences .102, 17687-92.
Ujvari, B., Dowton, M. & Madsen, T. (2007) Mitochondrial DNA recombination in a free-ranging Australian lizard. Biology Letters. 3, 189-192.

Vartak, R., Porras, C.A. & Bai, Y. (2013) Respiratory supercomplexes: structure, function and assembly. Protein & Cell. 4, 582-590.
Vermulst, M., Wanagat, J., Kujoth, GC., Bielas, JH., et al., (2008) DNA deletions and clonal mutations drive premature aging in mitochondrial mutator mice. Nature Genetics. 40, 392-394.
Vermulst, M., Wanagat, J. & Loeb, L.A. (2009) On Mitochondria, Mutations, and Methodology. Cell Metabolism. 10, 437.
Wai, T., Teoli, D. & Shoubridge, E.A. (2008) The mitochondrial DNA genetic bottleneck results from replication of a subpopulation of genomes. Nature Genetics. 40, 1484-1488.
Wallace, D., Singh, G., Lott, M., Hodge, J., et al. (1988) Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. Science. 1427-1429.
Wanagat, J., Cao, Z., Pathare, P. & Aiken, J. (2001) Mitochondrial DNA deletion mutations colocalize with segmental electron transport system abnormalities, muscle fiber atrophy, fiber splitting, and oxidative damage in sarcopenia. The FASEB Journal. 15, 322-332.
Wanrooij, S., Fusté, J.M., Stewart, J.B., Wanrooij, P.H., et al. (2012) In vivo mutagenesis reveals that OriL is essential for mitochondrial DNA replication. EMBO reports. 13, 1130-1137.
Wu, R. & Taylor, E. (1971) Nucleotide sequence analysis of DNA: II. Complete nucleotide sequence of the cohesive ends of bacteriophage λ DNA. Journal of molecular Biology. 57, 491-511.
Yang, M.Y., Bowmaker, M., Reyes, A., Vergan, L., et al. (2002) Biased Incorporation of Ribonucleotides on the Mitochondrial L-Strand Accounts for Apparent Strand-Asymmetric DNA Replication. Cell. 1, 495-505.
Yonemura, I., Nakada, K., Sato, A., Hayashi, J., et al. (2007) Direct cloning of full-length mouse mitochondrial DNA using a Bacillus subtilis genome vector. Gene. 391, 171-177.
Zheng, L., Zhou, M., Guo, Z., Lu, H., et al. (2008) Human DNA2 is a mitochondrial nuclease/helicase for efficient processing of DNA replication and repair intermediates. Molecular Cell. 32, 325-336.
Zsurka, G., Kraytsberg, Y., Kudina, T., Kornblum, C., et al. (2005) Recombination of mitochondrial DNA in skeletal muscle of individuals with multiple mitochondrial DNA heteroplasmy. Nature Genetics. 37, 873-877.
de Souza-Pinto, N.C., Wilson, D.M., Stevnsner, T.V. & Bohr, V.A. (2008) Mitochondrial DNA, base excision repair and neurodegeneration. DNA Repair. 7, 1098-1109.