Mitochondrial DNA and genetic disease

Kim J. Krishnan and Doug M. Turnbull

Mitochondrial Research Group, Newcastle University Centre for Brain Ageing and Vitality, Institute for Ageing and Health, Newcastle University, Newcastle NE2 4HH, U.K.

Abstract

From their very beginning to the present day, mitochondria have evolved to become a crucial organelle within the cell. The mitochondrial genome encodes only 37 genes, but its compact structure and minimal redundancy results in mutations on the mitochondrial genome being an important cause of genetic disease. In the present chapter we describe the up-to-date knowledge about mitochondrial DNA structure and function, and describe some of the consequences of defective function including disease and aging.

Introduction

The human mitochondrial genome is a closed circular, double-stranded DNA molecule of size 16569 bp located within the mitochondrial matrix (Figure 1). The entire nucleotide sequence of human mtDNA (mitochondrial DNA) was first established in 1981 [1] and re-sequenced in 1999 [2]. The two strands that make up the circular genome are denoted heavy (H) and light (L) based on their different buoyant densities due to uneven nucleotide content, with the H-strand being guanine-rich, while the L-strand is guanine-poor. The entire mitochondrial genome encodes 37 genes which produce 22 tRNAs (transfer RNAs), 13 proteins involved in OXPHOS (oxidative phosphorylation) and two rRNAs (ribosomal RNAs). Most of these genes are encoded on the

1To whom correspondence should be addressed (email d.m.turnbull@ncl.ac.uk).
H-strand with only eight mitochondrial tRNAs and a single polypeptide encoded by the L-strand. All other proteins (~1000) required for mtDNA maintenance and expression are encoded by the nuclear genome translated in the cytosol and are imported into the mitochondrial matrix.

Unlike the nuclear genome, each cell in the human body contains multiple copies of the mitochondrial genome, typically 1000–5000 copies per somatic cell. The mitochondrial genome contains no introns, so it is a highly compact and organized molecule. There is a small: 16024–16576 [np (nucleotide position)] NCR (non-coding region) containing the major control site for mtDNA expression, which includes the origin of leading strand (H-strand) replication and the major promoters for transcription (Figure 2). In humans, part of the NCR (np:16106–16191) forms a triplex structure termed the displacement loop.
(D-loop), which is thought to have formed due to H-strand synthesis stalling shortly after initiation to create three strands [3].

It is often assumed that mtDNA molecules are free floating within the mitochondrial matrix, however, this is not the case. mtDNA are not naked molecules, but are associated with a number of proteins in what are now termed ‘nucleoids’. It is thought that each nucleoid contains between six and ten copies of mtDNA in cultured cells. The actual proteins associated with nucleoids remain uncertain, but they include proteins involved in replication and expression such as Twinkle (mitochondrial helicase), POLG (polymerase γ) and TFAM (transcription factor A) [4].

### Evolution of mtDNA

Mitochondria are thought to be remnants of ancient, bacterial intruders now surviving as highly derived endosymbionts within eukaryotic cells [5]. Over time, a large majority of the genes encoded by mtDNA have been transferred to the nuclear genome. However, some have been retained and it is still uncertain as to why this is, but there are some theories. One attractive theory proposes that it may be difficult to import some hydrophobic proteins into mitochondria and therefore it is easier for such proteins to be produced inside mitochondria [6]. Support for this theory comes from the fact that all of the genes encoded by the mitochondria are hydrophobic, particularly cytochrome b, which is the most hydrophobic of all proteins within the mitochondrion. Alternatively we proposed recently that mitochondria may act as long-term redox damage sensors to provide cumulative feedback on the bioenergetic and genomic status within cells [7]. This would provide a selection

---

**Figure 2. Schematic representation of the human NCR**

Displayed are the main elements and factors involved in transcription and replication initiation. H-strand transcription is initiated from either HSP1 or HSP2. Initiation from HSP1 generates a short transcript which terminates after 16S RNA within mitochondrial tRNA L (Term), this produces the 12S and 16S rRNAs. Initiation from HSP2 generates polycistronic transcripts of the entire H-strand. L-strand initiation starts at the promoter, LSP, and produces polycistronic transcripts. Transcription is tightly regulated with replication, and RNA precursors for H-strand replication from OH are also generated from LSP. Conserved sequence blocks (CSBs I–III) are conserved regions in human, mouse and rat that participate in the formation of RNA primers for replication.
mechanism to maintain the energetically ‘fittest’ cells and a means of removing cells that compromise survival of the organism.

**Mitochondrial DNA replication**

mtDNA replicates independently from nuclear DNA and the cell cycle. The actual mechanism of mtDNA replication is a highly debated topic. Since the 1970s, initially through electron microscopic analysis of replicative intermediates of mammalian mtDNA, it was suggested that mtDNA undergoes an unusual strand-displacement asynchronous replication [8]. This model proposed that there are two distinct origins of replication, the origin of H-strand replication (termed OH) which is located within the D-loop region, while the origin of L-strand replication (termed OL) is located two thirds of the genome distance away on the heavy strand (Figure 1). This mode of replication predicts that mtDNA synthesis starts at OH, displacing the H-strand and proceeds along the parental L-strand to replicate a new H-strand. When H-strand replication reaches OL, this initiates L-strand synthesis which proceeds in the opposite direction to that of H-strand replication. The two strands are produced asynchronously and continuously without the need for Okazaki fragments.

For quite some time the asynchronous method remained unchallenged. In 2000, the first study began to challenge this model, instead proposing a conventional, coupled leading and lagging strand replication [9]. The coupled-replication mechanism was based on the observation of replicative intermediates in two-dimensional gels. In the first of a series of studies, the data were interpreted to show that both modes of replication, asynchronous and synchronous, were occurring [9]. However, subsequent work has led to the suggestion that strand synchronous replication is the predominant mode of mtDNA replication in the systems analysed [10]. There is some suggestion that different modes of replication may exist in cultured cells compared with tissue, also with the incorporation of RNA preceding DNA in the lagging strand [RITOLS (RNA incorporation throughout the L-strand)], possibly to protect regions from being single-stranded [11]. It has been proposed that RNA loss during isolation and processing in the earlier studies could account for the single-stranded molecules observed in the asynchronous method. In addition, there is also much debate about whether mtDNA replication initiates uni- or bi-directionally from the D-loop or other initiation sites [11].

Proteins required for the replication of mtDNA include the mitochondrial specific DNA POLG. POLG is composed of two subunits, a large catalytic α subunit (POLG1) (125–140 kDa), which possesses both 5′ → 3′ polymerase and 3′ → 5′ exonuclease activity, and two smaller β subunits (POLG2) (30–54 kDa) which are thought to increase processivity [12]. Twinkle is a hexameric DNA helicase protein, required for unwinding stretches of DNA up to 25 bases. TFAM is required for mitochondrial transcription, but is also thought to have histone-like function by binding and coating the entire mitochondrial
An additional enzyme also considered essential for mtDNA replication is type I DNA topoisomerase.

**Mitochondrial DNA repair**

For a number of years it was widely believed that mitochondria lacked any repair mechanisms as it was shown that pyrimidine dimers, induced by UV-irradiation, were not repaired in mitochondria [14]. Although it still remains the case that mitochondria do not possess NER (nucleotide excision repair) which remove pyrimidine dimers, what is clear is that mitochondria can repair other lesions with mechanisms similar to those found in the nucleus. The most common adduct caused by oxidative stress is 8-oxoG (8-oxo-7,8-dihydroxyguanine) and, in mitochondria, BER (base excision repair) mechanisms efficiently repair these adducts [15]. Both short-patch and long-patch BER have now been shown to exist in mitochondria, as well as mismatch repair activity [16,17]. Our understanding of all repair processes within mitochondria is still not complete. The list of proteins involved in mitochondrial repair is still growing since many proteins are constantly being reported to have dual location in the nucleus and mitochondria, as well as dual function.

**Mitochondrial gene expression**

Transcription of the H- and L-strands occurs in the D-loop from separate promoters, HSP and LSP respectively (Figure 2). Despite the close proximity of HSP and LSP they have been shown to be functionally independent. The L-strand is transcribed as a single, polycistronic precursor RNA which encodes most of the L-strand genes. H-strand transcription is complicated by the presence of two initiation sites (H1 and H2). H1 is located 16 bp upstream of mitochondrial tRNA\(^{\text{Phe}}\), while H2 is located within the mitochondrial tRNA\(^{\text{Phe}}\) gene, close to the 5′ end of the 12S rRNA gene. Thus transcription of the H-strand produces two polycistronic RNA units, one which initiates at H1, incorporates the rRNA genes and terminates at the downstream end of the 16S rRNA gene, and one which initiates at H2 and transcribes almost the entire strand. Initiation of H-strand transcription is most frequent at H1 and this may account for the difference in the synthesis rates between the rRNAs and the mRNAs encoded on the H-strand [18].

A mitochondrial-specific RNApol (RNA polymerase) initiates transcription, but not on its own, it requires TFAM and the transcription factors TFB1M and TFB2M. It is thought that the binding of TFAM to the transcription initiation site will bend the DNA, which introduces partial unwinding, allowing TFB2 to bind to ssDNA (single-stranded DNA) and recruit RNApol to the promoter [3]. Also involved is the termination factor mTERF [19]. The exact role of TFB1M in transcription is unknown, and may not be involved at all, as previously thought [20,21].
All of the protein and rRNA genes encoded on mtDNA are flanked by at least one tRNA gene. Processing of the polycistronic RNA molecules is therefore believed to be initiated by folding of the mitochondrial tRNAs, suggesting that the mitochondrial tRNA secondary structure provides punctuation marks for mtRNA processing. The folded mitochondrial tRNAs are likely to provide a substrate for the mitochondrial RNase P, but the exact mechanism responsible for the RNA processing and other factors involved are not fully understood. Recently, the components of mitochondrial RNase P have been identified and it appears to be unique in structure compared with other RNase P enzymes, which raises important questions about its evolutionary origin [22].

The mRNA sequences of the mitochondrial transcripts are translated within the mitochondrial matrix by the mitoribosomes (mitochondrial ribosomes; generated from the two mitochondrial-encoded rRNAs) and approx. 85 ribosomal proteins encoded by the nucleus and imported from the cytosol [23]. The exact details of human mitochondrial translation are largely unknown, although it involves mitochondrial initiation factors (IF2 and IF3), four critical elongation factors (EFTu, EFTs, EFG1 and EFG2) and one identified termination release factor (mtRF1a) [23]. From what is known, it appears that human mitochondrial translation differs from that found in the cytoplasm of eukaryotes and bacteria. One of the most interesting differences is their deviation from the standard genetic code: for example, the stop codon TGA is in fact read as tryptophan in human mitochondria; mitochondrial stop codons include AGA and AGG, which conventionally encode arginine [23]. Other differences in mitochondrial translation include bicistronic transcripts and partial stop codons in some of the transcripts that are completed by polyadenylation [23].

**mtDNA mutations**

The mitochondrial genome is associated with the IMM (inner mitochondrial membrane) where a large amount of ROS (reactive oxygen species) are continually produced as a byproduct of OXPHOS. This high ROS production may be a large factor in mtDNA having a ~10–17-fold higher mutation rate than that of the nuclear genome, but may also be a result of the limited repair capacity and lack of protective histones. Most changes that occur on mtDNA result in neutral polymorphisms, which have proved useful in tracking human migrations [24]. It has been two decades since the first mutations on the mitochondrial genome were described, and since then more than 250 primary pathogenic mutations have been documented to date. Most mtDNA mutations are either single point mutations or large-scale deletions. Point mutations can occur within protein, mitochondrial tRNA or mitochondrial rRNA genes. However, more than half of the disease-related point mutations reported are located within mitochondrial tRNA genes, despite the fact that these genes represent only approx. 5% of the mitochondrial genome. Large-scale single mtDNA deletions can span several genes and vary in size from 1.3 to 8 kb. It is important to note that as mtDNA is present in multiple copies
within cells, any given mutation can exist among wild-type copies in a situation referred to as heteroplasmy; although, some mtDNA mutations can reach homoplasmic levels within a cell (i.e. 100%). A cell can tolerate an mtDNA mutation up to a certain threshold (typically above 85% for point mutations and 60% for deletions) through complementation of the remaining wild-type copies. Therefore the majority of mtDNA mutations are considered functionally recessive. Once the mutation reaches this threshold, the cell becomes respiratory-chain deficient and comprises the cell’s normal functions. Despite the small size of this genome, new defects are still being recognized. Very recently the first functionally dominant mitochondrial tRNA mutation, m.5545C>T in MTTW, has been identified, with levels of mutation at <25% in clinically affected tissues [25]. This has important implications for the diagnosis of mitochondrial disease, as mutations at low levels such as these may easily escape detection.

The incidence of mtDNA mutations was initially thought to be low, but the frequency of known pathogenic mutations in the general population has recently been investigated. These studies have shown that approx. 1 in 500 individuals carry the m.3243A>G mutation [26], which can cause severe neurological disease, with a similar figure observed for the m.1555A>G mutation, which causes aminoglycoside-induced deafness [27,28]. Another approach for studying the incidence of mtDNA disease is to document the number of clinically affected cases within a specific geographical region. This approach is limited because there is marked clinical variability, which means many patients go unrecognized, and because patients may not be referred to specialist centres. However, even these studies show there is a high disease burden with at least 1 in 10000 of the adult population suffering from mtDNA disease [29].

**Maternal transmission of mtDNA and bottleneck**

Numerous population-based genetic studies have shown that the transmission of mtDNA is purely maternal. Maternal inheritance is very important in the context of mitochondrial diseases. Large maternal pedigrees have been identified, although sometimes there are apparently unaffected family members. This may be due to variation in the expression of the disease or in the presence of low levels of heteroplasmic mutations. Heteroplasmic mutations can vary widely between offspring due to the presence of a mitochondrial bottleneck during development. The copy number of mtDNA molecules in primordial germ cells falls to very low levels [30–32]. This mitochondrial bottleneck has a crucial role in preventing the transmission of very deleterious mutations. This purifying selection seems to occur during the cell replication phase in primordial germ cells. The stage in the process in which differences in heteroplasmy between offspring occur is less certain, but it is a major feature in many families with heteroplasmic mtDNA disease.
**mtDNA disease**

Patients with mtDNA disease present with variable phenotypes and age-of-onset, even with identical mutations. This complicated presentation is in part due to varying heteroplasmic levels of the mutation within tissues and the threshold effect. A classical hallmark of heteroplasmic mtDNA disease is the presence of respiratory-chain-deficient cells. These can be observed in tissue sections using a dual histochemical stain which reveals the presence of respiratory-chain-deficient cells (blue) among normal cells (brown) (for an example see Figure 3).

Single, large-scale deletions are generally responsible for sporadic disorders, including PS (Pearson’s syndrome) and KSS (Kearns–Sayre syndrome) [33]. The most commonly reported deletion associated with mtDNA is a 4977 bp deletion often referred to as the ‘common deletion’. KSS is a multi-system disorder and patients often develop severe neurological complications,

![Figure 3. Respiratory-chain-deficient cells in human tissue sections](image)

Histochemical staining to observe the presence of respiratory-chain-deficient cells (blue) among respiratory chain normal cells (brown) in substantia nigra from an aged individual (**A**) and skeletal muscle from a patient with a mutation in POLG (**B**). Respiratory-chain-deficient cells are a characteristic hallmark of mtDNA disease.
including cerebellar ataxia, cognitive impairment and deafness, as well as non-neurological features of cardiomyopathy, complete heart block, deafness, short stature and dysphagia. The clinical course in children with PS can be severe, leading to early death; but in those that survive the blood disorder improves, however, they later develop the clinical features of KSS. As mitochondria are under the dual genetic control of both the mitochondrial and nuclear genome, mutations in nuclear-DNA-encoding genes involved in mtDNA maintenance and mitochondrial nucleotide metabolism can also cause mtDNA disease. For example, mutations in \textit{POLG}, \textit{PEO1} (encoding Twinkle) and \textit{SLC25A4} can cause multiple mtDNA deletions which are associated with the disease CPEO (chronic progressive external ophthalmoplegia) [33]. CPEO is characterized by a progressive paralysis of the eye muscles leading to impaired eye movement and ptosis (dropping of eyelids). In patients with CPEO there may be other features depending on the underlying genetic defect, but myopathy and fatigue are common in all patients.

Point mutations on the mitochondrial genome can cause both inherited and sporadic mtDNA disease. The most common of these are the m.3243A>G mutation associated with MELAS (mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes) and the m.8344A>G mutation associated with MERRF (myoclonic epilepsy with ragged red fibres). Patients with MELAS often present with stroke-like episodes and seizures. Other features include intermittent episodes of encephalopathy, vomiting, migraine-like headaches, ataxia and cognitive impairment. MERRF is a progressive, neurodegenerative disease, which often presents in childhood or early adulthood following normal development.

\textbf{Prevention of mtDNA disease transmission}

Since the outcome of specific pregnancies remains unpredictable due to the genetic bottleneck, accurate genetic counselling for families affected by mtDNA mutations is extremely challenging. Nonetheless, several approaches to help minimize the transmission of high mtDNA mutation loads are being considered. For example, prenatal genetic diagnosis from chorionic villus sampling or amniocentesis may be performed. PGD (preimplantation genetic diagnosis) of mtDNA using blastomeres removed from early embryos prior to implantation is an alternative approach [34]. Prevention of mtDNA transmission may also be possible by pronuclear transfer, the transfer of the pronucleus from an oocyte of a mother with mtDNA disease to an enucleated oocyte from a healthy donor female. Studies in heteroplasmic mice have suggested the feasibility of this approach [35], and very recently a related technique, spindle–chromosome complex transfer between unfertilized metaphase II oocytes, has been reported using non-human primate oocytes [36].
mtDNA mutations and aging

For over a decade, mtDNA mutations, identical with those found in mtDNA disease, have been shown to accumulate with age in a number of tissues [37]. Levels of the mutations were usually found at <0.1% and it was questioned as to what role they could actually play in the aging process. However, high levels of mtDNA mutations have since been found to occur focally in a small subset of cells which could compromise these cells and contribute to the age-related decline in tissues [38,39]. A boost to the field of mtDNA in aging was given by the creation of transgenic mice with knock-in mutations in POLG [40,41]. These mice developed features characteristic of premature aging and crucially had increased mtDNA mutations. This was the first direct link to show that mtDNA mutations could contribute to aging. More recently, we and an independent group found high levels of mtDNA deletions (50%) in neurons from aged individuals and patients with Parkinson’s disease [38,39]. Again this highlights the potential importance of investigating the role that mtDNA plays in aging and disease.

Conclusions

Although our understanding of mitochondrial genetics has increased considerably over the last couple of decades, there is still so much that remains unknown. For example, where exactly are mitochondria replicated within a cell? How often are mitochondria turned over? How are mutated mitochondria targeted for degradation? Answers to these and many more basic questions are fundamental to improving our knowledge of the functional implications of mtDNA mutations. There is as yet no effective treatment available for patients with mtDNA disease. This is in part due to mtDNA being multi-copy and also the variation in mutation types detected within patients. Future areas for investigation as potential treatments include exercise and gene therapy. Also there is much interest in developing techniques to prevent mtDNA disease transmission.

Summary

- Mitochondria produce ATP through OXPHOS, but are also key players in apoptosis, calcium storage and iron–sulfur cluster biogenesis.
- mtDNA is a 16.6 kb, double-stranded, circular genome distinct from nuclear DNA and is maternally inherited.
- Cells contain multiple copies of mtDNA.
- mtDNA replicates independently from the cell cycle, but the exact mechanism for mtDNA replication is still to be determined.
- mtDNA mutations are an important cause of human disease, with >250 known pathogenic mtDNA mutations to date.
• mtDNA diseases with onset in early infancy/childhood include: Leigh syndrome, KSS and PS.
• mtDNA diseases with onset in late childhood or adult life include: MELAS, CPEO, NARP (Neuropathy, Ataxia and Retinitis Pigmentosa), LHON (Leber’s hereditary optic neuropathy) and MERRF.
• mtDNA mutations accumulate with age in a number of tissues and cause respiratory-chain deficiency in those cells with a high mutation load.

We are grateful for support from the Alzheimer’s Research Trust, Marriott Foundation, Wellcome Trust and the Newcastle University Centre for Brain Ageing and Vitality supported by the BBSRC (Biotechnology and Biological Sciences Research Council), EPSRC (Engineering and Physical Sciences Research Council), ESRC (Economic and Social Research Council) and MRC (Medical Research Council) as part of the cross-council Lifelong Health and Wellbeing Initiative.

References


