SJOERD WANROOIJ

# MtDNA Maintenance by Twinkle and Polymerase Gamma in Health and Disease

■

UNIVERSITY OF TAMPERE

# Contents

**List of original publications**

This thesis is based on the following original communications. In addition, some unpublished data is included.

I.      Spelbrink JN, Li FY, Tiranti V, Nikali K, Yuan QP, Tariq M, **Wanrooij S**, Garrido N, Comi G, Morandi L, Santoro L, Toscano A, Fabrizi GM, Somer H, Croxen R, Beeson D, Poulton J, Suomalainen A, Jacobs HT, Zeviani M and Larsson C (2001): Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria. Nat Genet 28: 223-231.

II.     **Wanrooij S**, Luoma P, van Goethem G, van Broeckhoven C, Suomalainen A and Spelbrink JN (2004): Twinkle and POLG defects enhance age-dependent accumulation of mutations in the control region of mtDNA. Nucleic Acids Res 32: 3053-3064.

III     **Wanrooij S**, Goffart S, Pohjoismäki JLO, Yasukawa T and Spelbrink JN (2007): Expression of catalytic mutants of the mtDNA helicase and polymerase POLG causes distinct repication stalling phenotypes. Nucleic Acids Res, *in press* doi:10.1093/nar/gkm215

**Related publications**

Tyynismaa H, Mjosund KP, **Wanrooij S**, Lappalainen I, Ylikallio E, Jalanko A, Spelbrink JN, Paetau A and Suomalainen A (2005): Mutant mitochondrial helicase Twinkle causes multiple mtDNA deletions and a late-onset mitochondrial disease in mice. Proc Natl Acad Sci U S A 102: 17687-17692.

Pohjoismäki JL, **Wanrooij S**, Hyvärinen AK, Goffart S, Holt IJ, Spelbrink JN and Jacobs HT (2006): Alterations to the expression level of mitochondrial transcription factor A, TFAM, modify the mode of mitochondrial DNA replication in cultured human cells. Nucleic Acids Res 34: 5815-5828.

**Abbreviations**

| | |
|---|---|
| ad | autosomal dominant |
| ADP | adenosine diphosphate |
| AMP | adenosine monophosphate |
| ANT | adenine nucleotide transporter |
| ATP | adenosine triphosphate |
| APE | apurinic/apyrimidinic endonuclease |
| APP | amyloid precursor protein |
| ar | autosmal recessive |
| BER | base excision repair |
| bp | base pair |
| CoQ | coenzyme $Q_{10}$ or ubiquinone |
| COX | cytochrome c oxidase |
| CSB | conserved sequence block |
| CTP | cytidine triphosphate |
| CPEO | chronic progressive external ophthalmaplegia |
| Cyt *b* | cytochrome *b* |
| DC | doxyxycline |
| dCK | deoxycytidine kinase |
| dGK | deoxyguanosine kinase |
| DAPI | 4',6 diamidino-2-phenylindole |
| ddC | dideoxycytidine |
| D-loop | displacement loop |
| DMSO | dimethylsulfoxide |
| DNA | deoxyribonucleic acid |
| dNDP | deoxyribonucleoside diphosphate |
| dNMP | deoxyribonucleoside monophosphate |
| dNTP | deoxyribonucleoside triphosphate |
| dRP | deoxyribose phosphate |
| DSB | double-strand break |
| EST | expressed sequence taq |
| EtBr | ethidium bromide |
| FADH2 | reduced flavin adenine dinucleotide |
| GFP | green fluorescent protein |
| GMP | guanosine monophosphate |
| GTP | guanosine triphosphate |
| HEK | human embryonic kidney |
| HMG | high-mobility group |
| HSP | heavy strand promoter |
| H-strand | heavy strand |
| ICC | immunocytochemistry |
| kDA | kilodalton |
| KSS | Kearns-Sayre syndrome |
| LHON | Leber's hereditary optic neuropathy |
| LMPCR | ligation-mediated PCR |

| | |
|---|---|
| LSP | light strand promoter |
| L-strand | light strand |
| MDS | mitochondrial DNA depletion syndrome |
| MELAS | mitochondrial encephalopathy, lactic acidosis and stroke-like episodes |
| mETC | mitochondrial electron transport chain |
| MERFF | myoclonus epilepsy with ragged red fibres |
| MILS | maternally inherited Leigh's syndrome |
| MMR | mismatch repair |
| MNGIE | mitochondrial neurogastointestinal encephalomyopathy |
| mRNA | messenger ribonucleic acid |
| MRP | mitochondrial RNA processing |
| mt | mitochondrial |
| mtDNA | mitochondrial DNA |
| mtSSB | mitochondrial single-stranded DNA-binding protein |
| MTERF | mitochondrial transcription termination factor |
| $NAD^+$ | nicotinamide adenine dinucleotide |
| NADH | reduced nicotinamide adenine dinucleotide |
| NARP | neuropathy, ataxia, an retinitis pigmentosa |
| NCR | non coding region |
| nDNA | nuclear DNA |
| NDPK | nucleosidediphosphates kinase |
| NER | nucleotide excision repair |
| NHEJ | nonhomologous end-joining |
| NRTI | nucleotide reverse transcriptase inhibitors |
| Nts | nucleotides |
| $O_H$ | H-strand replication origin |
| $O_L$ | L-strand replication origin |
| ORF | open reading frame |
| OXPHOS | oxidative phosphorylation |
| PAGE | polyacrylamide gel electrophoresis |
| PBS | phosphate buffered saline |
| PCR | polymerase chain reaction |
| PEO | progressive external opthalmaplegia |
| POL $\gamma$ | polymerase gamma |
| POLG1 | polymerase gamma catalytic subunit |
| POLG2 | accessory subunit polymerase gamma |
| POLRMT | mitochondrial RNA polymerase |
| QPCR | quantitative polymerase chain reaction |
| RI | replication intermediate |
| RITOLS | ribonucleotide incorporation throughout the lagging strand |
| RNA | ribonucleic acid |
| RNR | ribonucleotide reductase |
| ROS | reactive oxygen species |
| rRNA | ribosomal RNA |
| RT | reverse transcriptase |
| SANDO | sensory ataxia with neuropathy dysarthria and opthalmoparesis |

| | |
|---|---|
| SDS | sodium dodecyl sulphate |
| SSA | single-strand annealing |
| ssDNA | single-stranded DNA |
| TAS | terminated associated sequence |
| TCA | tricarboxylic acid cycle |
| TFBM | mitochondrial transcription factor B |
| TFAM | mitochondrial transcription factor A |
| TIM | transporter of the inner membrane |
| TK | thymidine kinase |
| TOM | transporter of the outer membrane |
| TP | thymidine phosphorylase |
| tRNA | transfer RNA |
| TTP | thymidine triphosphate |
| Twinkle | T7 gp4-like protein with intramitochondrial nucleoid localization |
| UTP | uridine triphosphate |
| UTR | untranslated region |
| UV | ultraviolet |
| XRCC1 | x-ray repair cross-complementing gene 1 |
| 2DNAGE | two dimensional neutral/neutral agarose gel electophoresis |

**Abstract**

Faithful replication and maintenance of mitochondrial DNA (mtDNA) are essential to the oxidative phosphorylation (OXPHOS) system that produces of ATP, and therefore to life. Maintenance of mtDNA depends entirely on nuclear-encoded gene products. Many of these proteins are presumed not yet identified, and several are poorly characterized. Currently, there is no consensus on the mechanism by which mammalian mtDNA is replicating. One long-standing replication model suggests that mtDNA is replicated by a strand-asynchronous manner. More recent studies, suggest an alternative model, which includes two modes of mtDNA replication, a bidirectional coupled-leading and lagging-strand synthesis mode and a unidirectional mode in which the lagging-strand is initially laid-down as RNA by a still unknown mechanism (RITOLS mode). In search for novel factors involved in mtDNA maintenance we identified the Twinkle protein. Twinkle displays structural similarity to the phage T7 gene 4 primase/helicase and co-localized with the nucleoprotein structures within mitochondria referred to as nucleoids. Twinkle mutations co-segregated with autosomal dominant progressive external ophthalmaplegia (adPEO), a disease characterized by multiple mtDNA deletions. This finding indicated that Twinkle is a crucial protein for proper mtDNA maintenance. Besides Twinkle defects, adPEO can be caused by mutations in 3 other nuclear genes, encoding the adenine nucleotide translocator (ANT1), the mitochondrial polymerase γ (POLG1) and its accessory subunit (POLG2). By which mechanism these gene defects lead to the accumulation of multiple mtDNA deletion in post-mitotic tissue remains to be elucidated. One hypothesis suggested that a decrease in replication fidelity leads to frequent mutations throughout the mtDNA of adPEO patients, which in turn would promote slipped mispairing and formation of mtDNA deletions. Our investiagtions did not support this idea, since we could only observe a higher mutation level in the control region, but not in other parts of the mtDNA molecule. We found frequent deletion breakpoints at homopolymeric runs by deletion breakpoint analysis in adPEO muscle mtDNA. We propose that this is a feature of replication stalling, which could be the primary cause of mtDNA deletion formation. This conclusion is supported by overexpression of mutant POLG1 or Twinkle in human cell culture, which results in a decrease in mtDNA content as a consequence of extensive replication stalling. Interestingly mutant POLG1 and Twinkle expression resulted in distinct replication stalling phenotypes observed by two-dimensional neutral/neutral agarose gel-electrophoresis (2DNAGE). The different stalling phenotypes can best be understood by taking into consideration the proposed mechanisms of lagging-strand synthesis. Our study showed that while induction of a defective POLG1 still presents delayed lagging-strand synthesis, Twinkle-induced stalling resulted in maturated, essentially fully double-stranded DNA intermediates. These intermediates mimic conventional strand-coupled RIs, however initiation of leading-strand DNA synthesis mainly occurred at the non-coding region. Limited inihibition of POLG1 with dideoxycytidine restored the delay between leading- and lagging-strand synthesis, suggesting that the maturation of the lagging-strand involves POLG1. In conclusion the observed replication intermediates (RIs) by 2DNAGE analysis of human cell-lines expressing different variants of Twinkle and POLG1 are most compatible with the RITOLS mode of mtDNA replication. This study produced deeper comprehension of POLG1 and Twinkle functions in adPEO and

mtDNA maintenance in general and suggests a universal deletion formation mechanism in adPEO as a result of frequent replication stalling.

# 1. Introduction

Most of the energy needed for cellular function is produced by organelles called mitochondria. The total process of this energy production system located within the mitochondria is called oxidative phosphorylation (OXPHOS). Some of the genes that encode essential subunits of this OXPHOS system are located in the mitochondrial DNA (mtDNA), a small circular genome that is separated from the nuclear DNA. For this reason proper maintenance of mtDNA is crucial for life. To understand which factors and mechanisms are involved in mtDNA replication and repair is important since defects in mitochondrial DNA are a common cause of human neuromuscular genetic disease and are also implicated in aging and infertility. The maintenance and replication of mtDNA depends on nuclear encoded genes. A defect in one of the nuclear genes involved in mtDNA maintenance can result in the accumulation of mtDNA abnormalities, like depletion or multiple deletions. Autosomal dominant progressive external ophthalmaplegia (adPEO) with multiple mtDNA deletion is an example of such a disorder. In this disease the multiple mtDNA deletions accumulate progressively during life and cause gradual paralysis of eye movements, ptosis and exercise intolerance all as a consequence of a slow decrease in energy production. Although several distinct autosmal loci for this disorder have been identified, in most cases the underlying nuclear gene defect remains unclear.

In this study we tried to identify novel factors involved in mtDNA maintenance and adPEO disease. We sequenced DNA obtained from adPEO patient in order to find the disease-causing nuclear gene defect. Further we measured the *in vivo* mtDNA replication fidelity in adPEO patient and compared the findings to age matched control. In order to learn more about the adPEO disease mechanism we mapped multiple mtDNA deletions to investigate at what kind of DNA sequences breakpoints occur. Based on this study we proposed a hypothesis of the mechanism behind deletion formation in adPEO. During later studies we used an inducible expression system in human cultured cells to study the functions of two proteins, involved in mtDNA maintenance and associated with adPEO, the novel mitochondrial replication protein Twinkle and POLG1 polymerase. This study created better understanding of the disease process underlying adPEO and the functions of Twinkle and POLG1 in general mtDNA replication. This basic research, discussed in this thesis will contribute to the understanding of human diseases caused by mitochondrial dysfunction that could ultimately result to development of a therapy.

## 2. Review of the literature
## 2.1 Function and origins of mitochondria

Mitochondria are organelles that are situated in the cytoplasm of eukaryotic cells. The major responsibility of this organelle is to convert nutrients to ATP for the cell's energy demands. For this reason mitochondria are often referred to as the 'power houses' of the cell. It is then also not surprising that the amount of mitochondria mass per cell is proportional to the energy needs of a certain cell (Nelson et al, 2000). Mitochondria are surrounded by two membranes, the inner and outer membrane. The outer membrane encloses the entire organelle. The inner membrane is compartmentalized into a series of cristae, the compartments where ATP production takes place. ATP is generated with the help of five large enzyme complexes. The first four of these enzyme complexes together shape the respiratory chain. The energy that is harnessed from electron transport between the complexes is used to pump protons from the mitochondrial matrix to the mitochondrial inter-membrane space. During this process the four multiprotein complexes achieve reduction of 1 oxygen molecule to two molecules of water plus oxidation of NADH (or FADH2) that is produced in the Krebs cycle. The function of the ATP synthase, the fifth enzyme, is to use the energy of the proton ion gradient to form ATP from ADP and phosphate. The total procedure of this ATP production system is called oxidative phosphorylation (OXPHOS). It consists of the phosphorylation of ADP (by ATP synthase) and the oxidation of NADH and reduction of molecular oxygen (by the respiratory chain) (Wallace, 1999).

A second important metabolic process in mitochondria is the breakdown of fatty acids in the beta-oxidation. Glucose is broken down in the cytosol (via glycolysis), which requires no oxygen. However, the end product pyruvate is imported into mitochondria. Both these processes and the break down of amino acids, result in the formation of acetyl-CoA. Acetyl-CoA is oxidized in the Krebs cycle leading to reduction of $NAD^+$ into NADH. Re-oxidation of NADH by complex I of the electron transfer chain is essential to keep the Krebs cycle going and forms a connection between the OXPHOS system and the Krebs cycle (Figure 2.1). Mitochondria are also essential for the formation of cellular building blocks such as amino acids, nucleotides, heme, coenzyme Q and cardiolipin. Other important duties of the mitochondria involve cellular signaling events such as, apoptosis (Green and Reed, 1998) and $Ca^{2+}$ -homeostatis (Gunter et al, 2004).

According to the widely accepted endosymbiotic theory, mitochondria are eukaryotic organelles of bacterial origin. Phylogenetic data supporting this theory point to an ancient alpha-proteobacterium, establishing a symbiotic relationship inside a primitive eukaryotic cell circa 2 billion years ago (Anderson et al, 2003). This partnership was beneficial for both, it gave them the possibility to use the emerging oxygen in the atmosphere for energy, while oxygen was toxic to most other life forms. The absorbed cell evolved into mitochondria, which retained some of the DNA and ribosomes characteristic of their original ancestors. The best evidence supporting the endosymbiotic theory is based on the discovery that mitochondria contain their own separate DNA (Nass and Nass, 1963). The small amount of mitochondrial DNA in human codes for 13 proteins, the mitochondrial ribosomal (rRNAs) and transfer RNAs (tRNAs). All the 13 mRNAs are subunits of the

different OXPHOS complexes, most of them are the hydrophobic elements of the electron transport chain, which do not easily pass the outer membrane. The alpha-proteobacterium that became the mitochondrion could use many cellular functions of its host resulting in extensive overlap of genetic information. During evolution most of the genetic information of the alpha-proteobacterium was transferred to the nucleus and most of the hundreds of other mitochondrial proteins are now imported from the cytosol. However despite the low proportional contribution to the organelle's protein complement, mtDNA encoded proteins are essential for normal cell bioenergetics and physiology.
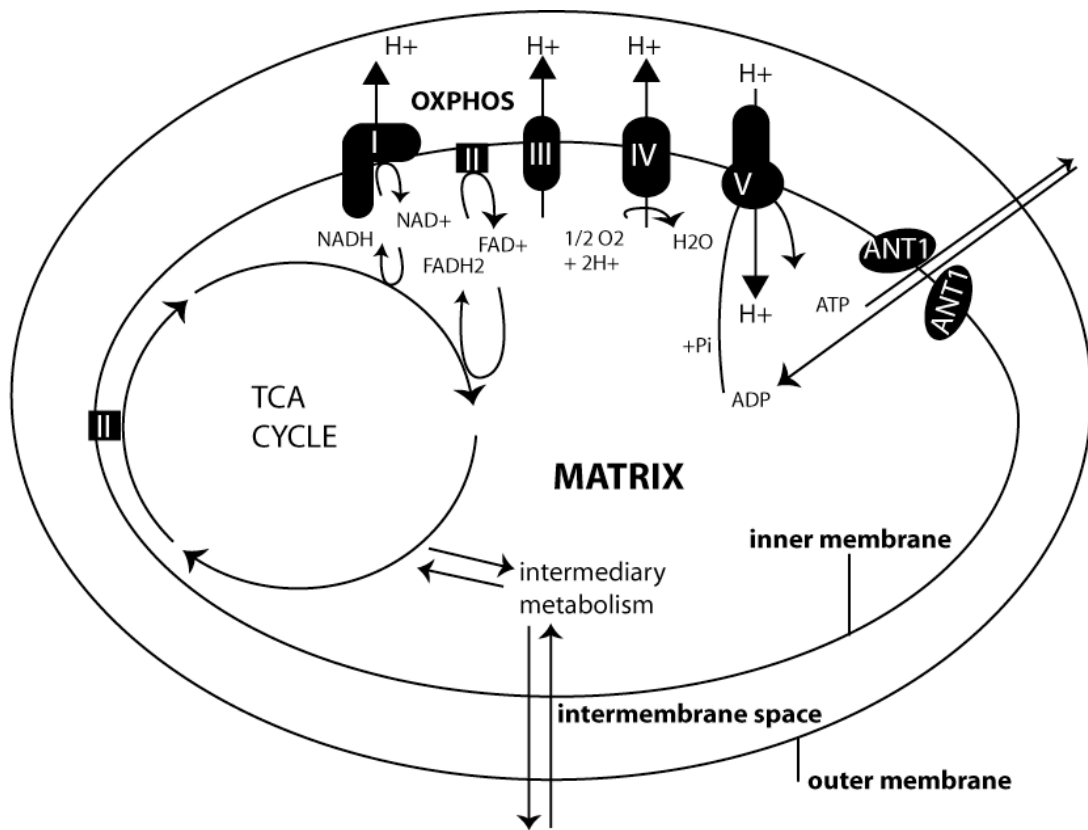


**Figure 2.1:** *Mitochondrial metabolism. Schematic presentation of a mitochondrion, with main metabolic activities (see text for more detail). For clarity, cristae are not shown. Electrons are transferred from complex I and II, requiring complex III, to complex IV which donates them to oxygen. Electron transfer from complex I, III and IV involves proton ($H^+$) translocation from the matrix into the inter membrane space. This results in an electrochemical proton gradient, which is used for ATP synthesis by complex V. Most of the produced ATP is exported to the cytosol in exchange for ADP by the adenine nucleotide translocator (ANT).*

## 2.2 Mitochondrial metabolism and structure
## 2.2.1 The dynamic structure and protein import of mitochondria

A distinctive feature of mitochondrial structure is the double membrane that separates the mitochondrial matrix from the cytoplasm of the cell. The mitochondrial matrix is the inner part of the mitochondrion that contains many enzymes involved in intermediary metabolism. It also contains multiple copies of the mitochondrial genome. The inter-membrane space is located between the outer and inner membranes. Whereas the smooth outer membrane of mitochondria is quite permeable, the inner membrane is impermeable, and specific transport systems are in place to allow movement of molecules in and out of mitochondria (Turcotte, 2003). The inner-membrane has many folds inwards into the mitochondrial matrix that are compartments referred to as cristae. These cristae expand the surface area of the inner-membrane drastically thus providing more surface for oxidative phosphorylation to occur (Manella, 2006).

Most notable of the mitochondrial composition is the variable appearance of the cristae. Depending on the cell type cistae are mostly seen as lamellar or as tubular shaped structures. The number and morphology of cristae is dynamic and depends on how the mitochondria react to the energy demands of a certain cells. Work in *Saccharomyces cerevisiae* has shown that when the respiration is repressed by growing the yeast on high glucose the morphology of the cristae changes dramatically, but not the number of mitochondria (Scheffler, 1999). This shows that the topology of the cristae can be rapidly adjusted and is a very dynamic process.

Mitochondria are dynamic organelles that continuously fuse and divide. Initially it was shown in *S. cerevisiae* that the morphology of mitochondria is regulated by fission and fusion of the mitochondrial membranes (Bleazard et al, 1999). The fission and fusion pathways, which are conserved in human, are two opposing forces that need to be in balance (Okamoto and Shaw, 2005). Disruption of fusion causes the tubular network of mitochondria to fragment into short rods due to ongoing fission events (Chen et al, 2003). On the other hand, disturbance of fission produces elongated, interconnected tubules (Smirnova et al, 2001). In yeast these morphology changes result in disruption of mitochondrial energy function and mtDNA stability. Besides the role of maintenance of morphology these mitochondrial dynamics have been suggested to regulate apoptosis. The current thought is that mitochondrial fission enhances apoptosis whereas mitochondrial fusion protects against this (Chen and Chan, 2005). The significance of these antagonistic pathways has been highlighted recently by the finding that two inherited human diseases are caused by mutations in two mitochondrial fusion genes (Olichon et al, 2006).

Mammalian mitochondria contain approximately 2000 different proteins. As mtDNA codes for only 13 of these, the contribution of the mitochondrial genome to the mitochondrial protein content is small. To properly execute their function mitochondria are dependent on recognition and import of these proteins from the cytosol. Protein translocases that are located in both inner and outer membrane are responsible for the extensive process of preprotein recognition and import. The translocase that facilitates the

transport of proteins from the cytosol across the outer mitochondrial membrane is called the TOM (translocase of the outer mitochondrial membrane) complex, it consists of cytosol-exposed receptors and a pore-forming center (Koehler, 2004). There are two translocases located in the inner membrane, which are involved in protein transport through the inner membrane. The TIM23 complex (a presequence translocase) forms a pore in the inner membrane and takes care of the transport across the inner membrane of presequence containing proteins. As soon as the TIM23 imported proteins enter the mitochondrial matrix the mitochondrial processing peptidase (MPP) cleaves off the presequence (Rehling et al, 2004). Another TIM complex, TIM22, uses the membrane potential to mediate the insertion of more complex proteins that have an internal targeting signal into the inner membrane.

## 2.2.2 Oxidative phosphorylation (OXPHOS)

Mitochondria generate cellular energy in the form of ATP by a process called oxidative phosphorylation (OXPHOS). Five large multiheteromeric complexes (I, II, III, IV, V) located inside the inner membrane shape this process (Zeviani and Di Donato, 2004)(Figure 2.1). The OXPHOS system is very complex, consisting of approximately 85 proteins encoded by two separate genomes (Smeitink et al, 2001). The 13 mtDNA and at least 72 nDNA encoded subunits that form this system are assembled together by a large group of chaperones and assembly factors. Coenzyme $Q_{10}$ (CoQ) and cytochrome c function as electron carriers respectivey linking complexes I-III and III-IV with each other (Wallace, 1999). These first four complexes and electron carriers together compose the mitochondrial electron transport chain (mETC) or respiratory chain. The mETC coordinates the transports of electrons and proton pumping across the inner membrane. This passage of electrons releases energy and is used by the last OXPHOS complex ATPase ($F_1F_0$- ATPase) to generate ATP from ADP and phosphate. Although some of the ATP produced is used by the mitochondria themselves most of it is exported by the adenine nucleotide transporter (ANT) and used for various cellular functions (Saraste, 1999).

Complex I of the mETC is NADH-ubiquinone oxireductase, the enzyme responsible for the oxidation of NADH originating from the oxidation of fatty acids, pyruvate and aminoacids (Carroll et al, 2003). This complex consists of 7 mtDNA and at least 36 nDNA encoded subunits (Janssen et al, 2006) and is sensitive to rotenone. In the process of oxidation, complex I transports electrons from NADH to the quinone acceptor CoQ, thus forming an ubiquinol, which can travel through the lipids in the membrane and transfer the electrons to complex III.

Part of the electrons derived from the Krebs cycle are transferred directly to CoQ, via complex II or succinate ubiquinone oxireductase. It is the smallest and simplest of all complexes. It is not involved in proton pumping, it simply transfers electrons from succinate to ubiquinone (CoQ) with the help of iron-sulfur redox cofactors (Cecchini, 2003). It is the only OXPHOS complex that does not have mtDNA encoded subunits, all 4 are nuclear encoded. CoQ delivers electrons to complex III, or ubiquinol-

ferricytochrome *c* oxidase. This antimycine sensitive complex contains one subunit, cytochrome *b*, derived from the mitochondrial genome and 10 subunits encoded by the nuclear genome. Complex III catalyses the reduction of cytochrome *c* by ubiquinol in association with proton pumping into the inter-membrane space. Cytochrome *c* is positioned in between complex III-IV and takes care of transfer of electrons from III to IV. The final electron acceptor complex is complex IV or cytochrome *c* oxidase. This cyanide sensitive complex holds 3 mtDNA and 10 nDNA encoded subunits and upon reoxidation of cytochrome *c* reduces one oxygen molecule to two water molecules (Saraste, 1999).

The function of the total mETC is to accepts electrons from NADH or FADH2 and transfer them through several successive oxidation/reduction reactions to oxygen to produce water. The electrons lose free energy at each step which is utilized and coupled to the transport of $H^+$ from the mitochondrial matrix to the inter-membrane space. This movement across the inner membrane results in an electrochemical gradient that provides the necessary energy for the second major step of oxidative phosphorylation to produce ATP by phosphorylation of ADP. This reaction is performed by ATP synthase ($F_1F_0$ ATPase or complex V), the last OXPHOS complex that includes 2 mtDNA and at least 12 nuclear encoded proteins (Saraste, 1999). The membrane potential generated by the mETC can be depolarized by transport of protons through a proton channel back into the mitochondrial matrix. Fo, the membrane component of the ATP synthase, forms the proton channel. This flow of protons drives the ATP production by the catalytic component $F_1$, located in the matrix side of the membrane (Birkenhäger et al, 1995). After production most of the ATP is then transported by the adenine nucleotide translocator (ANT) to the cytosol in exchange for ADP, thus ensuring a continuous supply of substrate in energy demanding situations (Figure 2.1).

Many different mutations in chaperones, assembly factors and mtDNA or nuclear DNA encoded subunits affect the OXPHOS system and results in a wide range of mitochondrial diseases. Defects in the OXPHOS system are among the most common genetic metabolic diseases (Thornburn, 2004), besides influencing energy production they also disturb other cellular processes such as ROS production or calcium homeostasis (Smeitink et al, 2001). Depending on the gene defect in the respiratory chain it leads to different cellular pathogenesis resulting in a great variability of often tissue specific clinical features.

**2.2.3 Other metabolic pathways in mitochondria**

Besides the oxidative phosphorylation mitochondria contain other metabolic pathways. Fuel molecule catabolism including glycolysis, carbohydrate, fatty acid and amino acid metabolism all produce acetyl-CoA. Acetyl-CoA is the initiator of the Kreb or citric acid cycle. The final accounting after the 9 enzymatic reactions including the key regulator enzyme pyruvate dehydrogenase is the conversion of pyruvate (C3) to 3 molecules of CO2. In this process four molecules of $NAD^+$ are reduces to NADH which can be used by complex I for the first oxidation reaction of the electron transport chain. The oxygen consumption, NADH oxidation and ATP synthesis during oxidative phosphorylation are

not only coupled to each other but also connected to the citric acids cycle. (Scheffler, 1999)

Most fatty acids are degraded in the beta oxidation. Fatty acids are activated in the cytosol to form acyl-CoA, after that they are transported to the mitochondria. The oxidation of the acyl-CoA occurs by a cycle of four enzymatic steps. In every cycle one molecule of acetyl-CoA and a two carbon molecule shorter acyl-CoA is formed. This acyl-CoA can then again be degraded by further cycles (Stryer, 1999).

Other important metabolic pathways localized in the mitochondria are: i) The urea cycle is a metabolic pathway that eliminates excess of aminoacids and nitrogen from the body (Nassogne et al, 2005). ii) Biosynthesis of the iron containing heme metalloproteins that have as main function retention of oxygen and delivering it for enzymatic reactions (Moraes et al, 2004). iii) Metabolism of lipids and cardiolipin, that are required for formation of the biological membranes (Mileykovskaya et al, 2005). iv) Biosynthesis of CoQ, this lipid is a member of the mitochondrial respiratory chain and beyond this it has several other functions of great importance for the cellular metabolism like control of cellular redox state and as antioxidant (Turunen et al, 2003; Crane, 2001).

## 2.3 The mitochondrial genome

The gene organization, size and expression mode of the mitochondrial genome is very different in diverse phyla, however the basic function and content of genes is exceptionally similar. The human mitochondrial genome was the first human genomic entity to be entirely sequenced in 1981, it was reported to be a circular double stranded molecule 16,569 base pairs in length (Anderson et al, 1981)(Figure 2.2). The mitochondrial DNA molecule consists of two strands, a guanine-rich heavy (H) strand and a cytosine-rich light (L) strand. Due to the presence of many copies per cell, it corresponds to about 1 % of the total DNA content in an average cell. The mitochondrial genome of mammals is organized extremely economically, showing a gene organization that is very compact without any introns (Fernandez-silva et al, 2003). In all vertebrates the DNA molecule holds one major non-coding region, part of this region contains the so called displacement loop (D-loop). The D-loop includes the promoters of transcription initiation and the suggested heavy strand replication origin (Shadel and Clayton, 1997). Even though the precise function of mitochondrial D-loops is not yet clear, D-loops are assumed to be involved in regulating mtDNA replication (Takamatsu et al, 2002). The name, D-loop originates from the three-stranded displacement (D) loop structure generated by a nascent short H-strand that displaces the parental H strand (ter Scheggert et al, 1971). The D-loop is the region with the most variable sequence among different species, even though it harbors some well conserved elements (Sbisa et al, 1997). The mitochondrial genome in mammals includes 37 different genes, of which 24 genes (22 tRNAs +2 rRNAs) encode the RNA components of the mitochondrial translation apparatus. The remaining 13 mRNAs are subunits of four of the OXPHOS complexes. Seven of those are subunits of complex I, one of complex III, three of complex IV and finally two subunits of complex V. The genetic information is divided asymmetrically on

the mtDNA molecule since the H-strand holds a majority of the genes (Anderson et al, 1981).



**Figure 2.2:** *The human mitochondrial genome. The human mitochondrial genome consists of 16568 base pairs (bp) and contains a heavy- and (H-strand) light-strand (L-strand). Complex I NADH dehydrogenase (ND) genes are shown in green; Complex III cytochrome b (Cytb) gene is shown in red; Complex IV cytochrome c oxidase (COX) genes are shown in light blue; Complex V ATP synthese (ATPase) genes are shown in gray. Transfer RNA genes in black and ribosomal RNA genes (rRNA) in yellow.*

The mitochondrial genome has some distinctive characteristics that are different when you compare it to the nuclear genome. One unique feature is the use of a differing genetic code, for this reason the cytosolic translation machinery is not able to translate mitochondrial mRNAs or vice versa. Secondly, the mitochondrial genome is normally inherited maternally; it has been proposed that the mitochondria from the sperm cell are actively eliminated by an ubiquitin-dependent mechanism (Sutovsky et al, 1999). Another difference of the mtDNA molecule is the higher mutation rate. MtDNA is positioned in the vicinity of the OXPHOS system, which is considered to be the major

producer of reactive oxygen species (ROS). This location might explain the increases mutation rate (Wallace, 2005). One more peculiarity of the mitochondrial genetic system is that cells are polyploid regarding mtDNA. While as a rule human cells hold two copies of the nuclear genome, mitochondria contain many more mtDNA copies. In normal tissue, all mtDNA molecules are the same (homoplasmy), however mutations can arise, be amplified to different levels and coexist with wild-type mtDNA, a state known as heteroplasmy. As a consequence of this it is common to find a 'threshold effect' in mtDNA related human diseases in which a minimum number of mutant mtDNA molecules is required to cause mitochondrial dysfunction in a particular tissue (Dimauro and Davidzon, 2005; Chinnery and Schon, 2003).
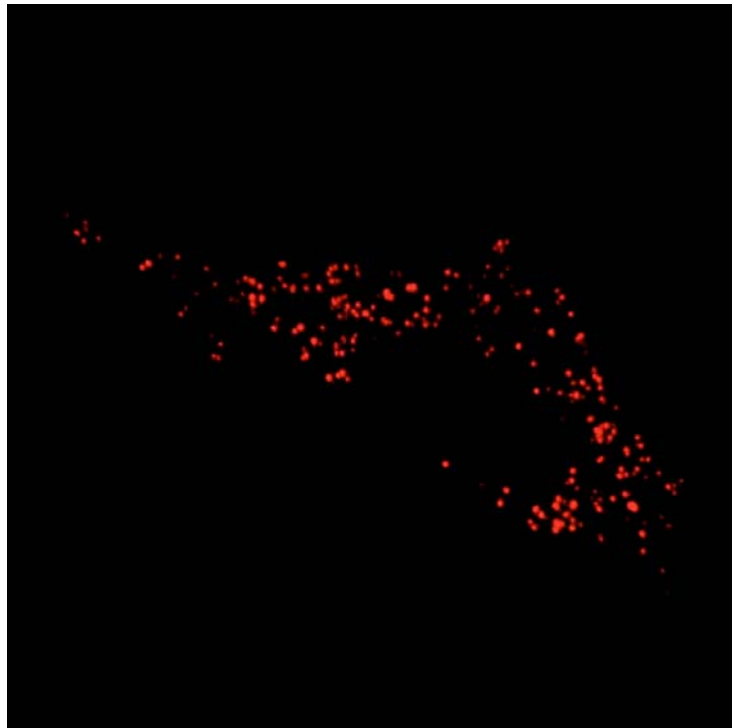


**Figure 2.3:** *Confocal microscopy picture of osteosarcoma cells that overexpress fluorescent mtSSB-DsRed, the mitochondrial ssDNA binding protein. The fluorescent signal of mtSSB shows punctate foci within the mitochondria, which is due to its co-localization with the mitochondrial nucleoids (kind gift from Dr. P. Martinsson).*

For some time now yeast mtDNA is known to be organized in a discrete number of nucleoprotein complexes, called nucleoids, that that can be visualized by fluorescence microscopy with DNA-binding dyes (Williamson and Fennell, 1975; Miyakawa et al, 1984). More recently similar protein-mtDNA complexes have been characterized in human mitochondria (Figure 2.3). These nucleoids each contain several mtDNA molecules and appear to be stable yet dynamic structures (Garrido et al, 2003; Legros et al, 2004; Iborra et al, 2004). Various strategies have been used to elucidate the composition of human and yeast mtDNA nucleoids (Miyakawa et al, 1987; Garrido et al, 2003; Chen and Butow, 2005; Wang and Bogenhagen, 2006).

## 2.4 Transcription and translation of the mitochondrial genome
## 2.4.1 Mitochondrial transcription

In human cells, each mtDNA strand includes at least one promoter for transcriptional initiation, the light strand promoter (LSP) and the heavy strand promoter (HSP) (Montoya et al, 1982). Transcription that initiates from these promoters produces precursor RNA encoding for all the genetic information in each of the strands. A later study suggested there is a second heavy strand initiation site of transcription (Montoya et al, 1983). It has been proposed that one initiation site operates much more frequently and is accountable for the synthesis of the two rRNA and two tRNAs. In this manner H-strand transcription of rRNA versus mRNA could be differentially regulated.

The transcription of mtDNA is performed by the mitochondrial RNA polymerase (POLRMT). This is a 120 Kda protein that is homologous with phage RNA polymerases (Titranti et al, 1997). For the initiation of transcription POLRMT needs the simultaneous interaction of mitochondrial transcription factor A (TFAM) and one of the transcription factors B1 (TFB1M) or B2 (TFB2M) (Falkenberg et al, 2002; Gaspari et al, 2004). The current hypothesis is that TFAM binds upstream of the promoters and introduces particular alterations in mtDNA which can stimulate transcription initiation by facilitating POLRMT/TFB2M binding to the initiation site (Shadel and Clayton, 1997; Gaspari et al, 2004). Both TFB1M and TFB2M stimulate transcription activity by direct contact with POLRMT forming a heterodimer, although showing very different levels of stimulation (Falkenberg et al, 2002). TFB1M has also been shown to contain RNA methyltransferase activity suggesting that it is a duel-function protein (McCulloch and Shadel, 2003). The human mitochondrial transcription termination factor (MTERF1) has been suggested to play an essential role in the regulation of H strand rDNA transcription by promoting initiation and termination, of the rDNA transcript (Martin et al, 2005). The recent finding that at least four MTERF family-members exist suggests that control of mitochondrial transcription is far more complex than previously expected (Linder et al, 2005).

## 2.4.2 Mitochondrial translation

Mitochondria contain their own separate synthesis apparatus for translation of the genetic information encoded by the mtDNA. In humans, the RNA components of this machinery are encoded by mtDNA, however all proteins necessary for mitochondrial translation have to be imported from the cytosol (Jacobs and Turnbull, 2005). Mitochondrial ribosomes are classified as a bacteria-type ribosome because of their antibiotic susceptibilities and sequence alignment with bacterial ribosomal proteins (Zhang et al, 2005)

The mitochondrial translation system possesses the distinct characteristic that RNAs are more simple then those of the cytosolic translation system. Whereas over 50 species of tRNA are crucial in cytoplasmic translation systems, only 22 tRNA species are enough for the translation of 60 codons in mitochondria. Mitochondrial genes are translated with a genetic code that is different from the universal code. In mammalian mitochondria,

UGA specifies tryptophan instead of a termination codon, AUA, AUC and AUU are used as initiation codons while AGA and AGG are termination codons (Attardi, 1988).

12S and 16S rRNAs are encoded by the mtDNA and combined with nuclear derived mitochondrial proteins, they shape the mitoribosomes. With support of the mitochondrial tRNAs the ribosomes are able to translate the mtDNA encoded mRNAs. The enzymes for mitochondrial rRNAs and tRNAs processing and base-modification are nuclear derived. Just like aminoacyl-tRNA synthetases, initiation, elongation and termination factors they need to be imported from the cytosol (Jacobs and Turnbull, 2005). The precise mechanism of mitochondrial translation has been extensively reviewed (Gillham et al, 1994).

**2.5 mtDNA replication mechanisms and proteins**
**2.5.1 mtDNA replication mechanisms**

For several decades mammalian mtDNA replication has been considered to occur through an asymmetric mechanism. In this strand-displacement model it was proposed that the two mtDNA strands replicate asymmetrically (Robberson et al, 1972)(Figure 2.4A). This model of mtDNA replication is based on interpretation of results obtained from several different experimental strategies such as electron microscopy, pulse- and pulse-chase labeling studies, strand isolation, and 5'end characterization (Clayton, 1982; Brown et al, 2005). The strand-displacement replication of mtDNA shares various similarities with ColE1 and some other plasmids. However, this model is not like eukaryote or eubacterial chromosomal DNA replication (Shadel and Clayton, 1997).

According to the strand-displacement model heavy-strand replication begins at the so-called H-strand replication origin ($O_H$), which is situated within the D-loop region. The RNA primer that is needed for the initiation of replication is suggested to be a processed L-strand promoter (LSP) transcript. This organization of initiation makes that mtDNA replication is coupled with and dependent on the mitochondrial transcription machinery. The formation and processing of the transcript requires the presence of three conserved sequence block (termed CBSI, CBSII, and CSBIII) (Xu and Clayton, 1996). The mechanism responsible for the regulation of full-length transcription versus primer formation has been suggested to involve cleavage of the primary LSP transcript by an endonuclease activity executed by RNase MRP (Lee and Clayton, 1997). However recent data obtained using a pure in vitro transcription system (Falkenberg et al, 2002) suggests a new RNase MRP independent mechanism significantly dependent on the CSBII sequence (Pham et al, 2006).
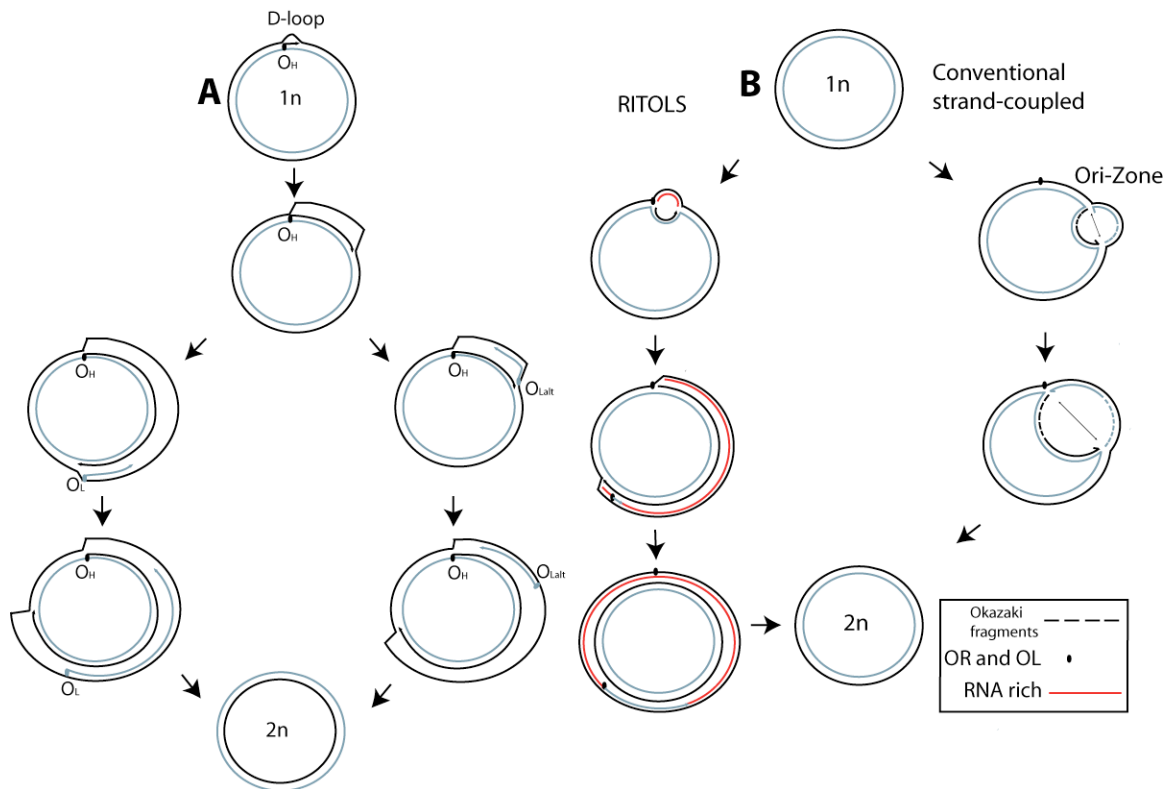
**Figure 2.4:** *Replication models of mammalian mtDNA. A) Strand-displacement replication model. Replication starts with leading strand H-strand (in black) initiation at $O_H$ producing a D-loop. Although often replication terminates at the termination associated sequence replication can proceed in this case the parental H-Strand is further displaced and can be used as template for L-strand synthesis (shown in gray). L-Strand synthesis normally initiates at $O_L$ (shown on the left) or at an alternative light-strand origin ($O_{LALT}$)(shown on the right of panel A). mtDNA replication by this model would result in extensive ssDNA RIs. This figure was adapted from (Brown et al, 2005). B) RITOLS- and strand coupled-model of mtDNA replication. RNA incorporation throughout the lagging strand (RITOLS) mode of replication is shown on the left in panel B. Replication initiates at a discrete origin $O_R$ in a unidirectional fashion. The newly synthesised lagging-strand is initially laid down as RNA (red) by a still unknown mechanism using the H-strand (black) as template. The slow step of this replication mode is the RNA conversion/maturation to DNA (gray), which uses $O_L$ as a major but not the only initiation site by an as yet unclear mechanism. Replication of mtDNA by this mode would result in prominent RNA/DNA hybrids RIs. The strand-coupled model is shown in the right pathway of panel B. During the strand-coupled mtDNA mode of replication, initiation occurs at an Ori-zone downstream of $O_H$ in a bidirectional manner. This is followed by progression of both forks until one fork arrest at $O_H$. A far majority of RIs produced by this mode will be dsDNA. This figure was adapted from (Yasukawa et al, 2006).*

The mitochondrial DNA polymerase γ (Pol γ) can start H-strand synthesis through the extension of an RNA primer (Ropp and Copeland, 1996). After H-strand replication has started there are two suggested possibilities. The first possibility is a replication arrests

around a 15bp conserved sequence that is positioned a short distance upstream from the 3' D-loop end, called the termination associated sequence (TAS) (Madsen et al, 1993). This is suggested to generate a more or less stable D-loop containing molecule. One suggested role for the D-loop strands is that they provide primers for re-initiation of replication, although no direct evidence has been presented yet. In the strand-asynchronous replication model, only a minority of the initiated nascent H-strand replication overcomes termination and proceeds over the entire length of the genome. The premature arrest could be one possible regulation point of mtDNA replication. Recent experimental data in support of this hypothesis show a decrease in TAS associated termination after an increased mtDNA replication rate. In both proliferating T-cells and HeLa cells recovering from mtDNA depletion, mtDNA replication is increased. It has been suggested this was a consequence of a decrease in termination (Kai et al, 1999; Brown and Clayton, 2002).

When in the strand-displacement model of mtDNA replication H-strand synthesis proceeds beyond the TAS sequence, Pol γ continues unidirectionally until two-third of the way around the mitochondrial genome, where it meets the second origin for light-strand replication ($O_L$). When the replication machinery passes this point, $O_L$ is exposed on the displaced heavy strand. This initiation for the L strand replication occurs at a small non-coding region (30bp) that is located in the WANCY tRNA coding region. It has been suggested that this particular DNA sequence is able to form a stable stem-loop structure after the newly synthesized H-strand passes the $O_L$ region and exposes the parental H-strand as a single-strand (Tapper and Clayton, 1981). In this step it is necessary that a small RNA molecule is generated by a primase in order to initiate L-strand synthesis. Pol γ then synthesizes the nascent L-strand in the opposite direction on the single-stranded H-strand template. Following this second initiation, L-strand replication proceeds over the whole length of the strand and ends after the synthesis of the nascent H-strand has terminated (Tapper and Clayton 1981). In this strand-asymmetric model of replication, DNA synthesis is continuous on both strands. When the synthesis is complete the new mtDNA molecules need to be separated, RNA primers removed and gaps ligated after which the replication cycle can be completed by the addition of superhelical turns and binding of diverse proteins. However, these steps are not yet well understood (Lecrenier and Foury, 2000)

Since 2000, several papers have appeared that re-exam this previously described model of mtDNA replication (Holt et al, 2000; Yang et al, 2002; Bowmaker et al, 2003; Reyes et al, 2005; Yasukawa et al, 2005; Yasukawa et al, 2006)(Figure 2.4B). These recent results obtained mainly by using two dimensional neutral/neutral agarose gel electrophoresis (2DNAGE) have lead to an intensive debate about the bona fide model of mammalian mtDNA replication (Bogenhagen and Clayton, 2003a; Holt and Jacobs, 2003; Bogenhagen and Clayton, 2003b). In the 2DNAGE technique (Brewer and Fangman, 1987; Brewer and Fangman, 1988), a widely used method to examine mechanisms of replication, endonuclease digested DNA fragments are separated in a first dimension agarose gel based on size and in a second dimension agarose gel based on size and strand configuration, by using ethidium bromide. In this system particular types of replication intermediates (RIs) migrate in specific and predictable ways. The 2DNAGE

data revealed that a significant fraction of the DNA formed duplex RIs that were seen as archetypal bubble and forks arcs in the experiments (Holt et al, 2000). Holt and co-workers concluded that these obtained RIs are derived from coupled leading and lagging-strand synthesis of mtDNA. In their opinion these RIs could not be explained with the asymmetric strand-displacement model of replication. Other RIs containing extensive regions of single-strandedness were initially suggested to represent the strand asymmetric mode of mtDNA replication. However, shortly after, new data showed that discontinuous ribonucleotide-rich synthesized lagging strands are more easily degraded due to an RNase H-like activity present in those preparations (Yang et al, 2002). This RNase H sensitivity showed that the strand-displacement RIs could be created by an artifact of preparation. It was thus suggested that mammalian mtDNA replication proceeds solely by coupled leading and lagging-strand synthesis. A recent study lead to the refinement of this replication model, the data obtained by 2DNAGE showed two types of RIs (Yasukawa et al, 2006), representing two distinct modes of mtDNA replication (Figure 2.4B). One class of RIs contain very few ribonucleotides and are fitting to strand-coupled DNA synthesis, initiating bidirectionally from a broad zone (Bowmaker et al, 2003; Reyes et al, 2005). The other class of RIs identified were ribonucleotide-rich and mapping of 5' ends showed this mtDNA synthesis initiated unidirectional strictly from the non-coding region (Yasukawa et al, 2005; Yasukawa et al, 2006). This mtDNA replication mode involves ribonucleotide incorporation throughout the lagging strand (RITOLS) across practically the complete genome. In RITOLS replication the lagging-strand is laid down initially as RNA before it is converted to DNA. The RNA might help to protect and stabilize the displaced ssDNA strand, it also could provide genetic informational back-up. Virtually nothing is known yet about the processing and maturation of these RNA-rich intermediates.

However, proponents of the strand-asynchronous model have put forward that the 2DNAGE method destabilizes many of the strand-displacement RIs resulting in an artifact that makes duplex RIs seem more prominent (Brown and Clayton, 2006). Further Clayton and co-workers believe that the resistance of some RIs to restriction-enzyme digestion is not a consequence of extensive incorporation of ribonucleotides but rather caused by mtDNA transcription intermediates.

**2.5.2 mtDNA Polymerase γ**

Human polymerase gamma (Pol γ) functions in a heterotrimer consisting of one catalytic subunit (p140) encoded by *POLG1* and two subunits of the accessory subunit (p55) encoded by *POLG2* (Yakubovskaya et al, 2006). Although 16 different polymerases have been identified in mammals, Pol γ is the only polymerase that is located and active within the mitochondria (Graziewicz et al, 2006). Thus Pol γ is indispensable for all replication and repair of mtDNA. POLG1 is nuclear encoded; the N-terminal region contains three conserved motifs that form the 3'→5'exonuclease domain (Exo1, Exo2 and Exo 3) while the C terminal region holds three conserved polymerase motifs (PolA, PolB and PolC) (Ropp and Copeland, 1996; Lecrenier et al, 1997)(Figure 2.5). The responsibility of the polymerase domain is the actual mtDNA synthesis whereas the exonuclease domain takes care of proofreading during replication. A region often referred to as "the linker region"

separates the polymerase and exonuclease domain. Based on the conserved sequence motifs Pol γ has been grouped in family A of DNA polymerases. This family includes other well-studied polymerases like T7 DNA polymerase and polymerase I of *E.coli*. While the Exo and Pol motifs of Pol γ are very conserved in polymerase family A, the length of the region that connects the exonucleolytic and polymerase catalytic domains is considerably increased in Pol γ (Lecrenier et al, 1997). Downstream of the N-terminal mitochondrial targeting signal the human *POLG1* gene includes a CAG tri-nucleotide repeat that normally codes for 10 glutamines. Although in cell culture deletion of this CAG repeat does not affect mtDNA maintenance (Spelbrink et al, 2000), alterations in the length of this repeat are associated with male infertility (Rovio et al, 2001).



**Figure 2.5:** *Schematic diagram of the 1239 amino acids (AA) long human POLG1 protein. Within the polymerase and exonuclease domain (shown in black) the conserved Exo (I-III) and Pol (A-C) active site motifs are indicated with white boxes. The POLG1 gene contains a N-terminal mitochondrial targeting signal (shown in gray).*

The first experiments that showed that Pol γ has a role in mtDNA maintenance were performed by inactivation of the yeast Pol γ gene MIP1 (Genga et al, 1986; Foury, 1989). Even though this first cloned mitochondrial polymerase MIP1 shares ~40 % identical residues in amino acids sequence with human POLG1 it differs in that it doesnot appear to require an accessory subunit (Lecrenier et al, 1997). The Pol γ holoenzymes from *Drosophila melanogaster*, *Xenopus laevis* and mice do posses two distinct subunits (Wernette and Kaguni, 1986; Insdorf and Bogenhagen, 1989; Carrodeguas et al, 2001). The accessory subunit of Pol γ as far as we know does not posses catalytic activities by itself, however it stimulates the activities of the catalytic subunit by increasing its affinity to DNA. The isolated human p140 catalytic subunit has modest processivity by itself, producing on average 100 nucleotide DNA product from a natural DNA template. However, when p55 is added, products as long as 7 kilo bases can be reached (Lim et al, 1999). A recent report identified the first human pathogenic mutation in the human p55 gene, POLG2 (Longley et al, 2006). Using purified proteins it was shown that this p55 variant is not capable of enhancing the affinity of p140 to DNA and thus affects Pol γ processivity resulting in accumulation of multiple mtDNA deletions in post mitotic tissues of this patient.

Purified Pol γ is able to replicate natural DNA templates and several homopolymers. Because Pol γ contains reverse transcriptase activity a Poly(rA).Oligo(dT) template is often used in enzymatic assays, since this gives the possibility to discriminate Pol γ activity from the other polymerases active in the cell (Longley et al, 1998b). Pol γ reverse transcriptase (RT) activity has been suggested to be of physiological significance, since its RT activity is slightly higher then HIV RT (Murakami et al, 2003). Nucleotide reverse transcriptase inhibitors (NRTI) are used as medicine for HIV infected patients as it slows down the disease progression by inhibiting HIV RT. However, a side effect is inhibition

of Pol γ activity leading to mtDNA depletion. It shows that inhibition of the reverse transcriptase ability of Pol γ causes loss of mitochondrial function in human (Dalakas et al, 1990)

Besides the very conserved exonuclease and polymerase motifs the *Drosophila*, *X. leavis* and human Pol γ also contains six moderate conserved sequences (γ1 till γ6) (Kaguni, 2004). Four of these elements (γ1-γ4) are located within the linker region, while the remaining two are located around the polymerase motifs. A Study on the linker region of T7 polymerase has shown that it is responsible for interaction with the accessory subunit (Doublie et al, 1998). Although the linker region of Pol γ has no known function a similar role has been suggested. Recent deletions studies on the drosophila Pol γ p140 support this idea. The authors showed that removing the moderate conserved sequence γ1 and γ4, resulted in a weaker interaction with the accessory subunit (Luo and Kaguni, 2005). The importance of this region is further emphasized by the observation that mutations in the POLG1 linker region are associated with several human diseases.

Currently no three-dimensional crystal structure of p140 is available, making it difficult to predict the functional effect of p140 mutant variants. However a structural model based on similarities of other family A polymerase members (e.g. T7 polymerase) has been successfully used to explain functions of many conserved amino acids (Graziewicz et al, 2004). The structural insight into less conserved regions of POLG1 will have to wait until a crystal structure is published.

The exonuclease motifs identified in all family A polymerase members is responsible for proofreading newly incorporated bases during ongoing DNA synthesis. Site directed mutagenesis of conserved amino acids located within the exonuclease motifs of MIP1 results in a strong increase in mtDNA mutation levels without affecting the actual replication rate (Foury and Vanderstraeten, 1992). In human cell-lines the expression of a proofreading deficient POLG1 with a single aminoacid change D198A in exonuclease motif 1 similarly resulted in accumulation of mtDNA mutations (Spelbrink et al, 2000). A similar point mutant in exonuclease motif 2 was recently produced in a homozygous knock-in mouse, giving rise to a 5 times higher mtDNA mutation level in heart, liver and brain (Trifunovic et al, 2004; Kujoth et al, 2005). The higher mutation levels in these mice coincided with a premature ageing phenotype and a reduced lifespan. These findings and the observation that all human disease mutation found in the exonuclease domain are not located in the conserved motifs demonstrate the importance of this proofreading mechanism for normal mitochondrial function. The collaboration of the polymerase domain and exonuclease domain gives Pol γ a high fidelity (Longley et al, 2001). However Pol γ is known to have difficulties with homopolymeric runs making it prone to create frame shift mutations (Longley et al, 2001).

Human Pol γ also contains 5'deoxyribose phosphate (dRP) lyase activity, which plays a role in base excision repair (BER)(Figure 2.6). It functions together with an AP endonuclease in the removal of the incorrect or damaged base that is recognized and marked by DNA glycosylases. After this removal Pol γ is also responsible for filling the

single-nucleotide gap (Pinz and Bogenhagen, 1998; Longley et al, 1998b). The accessory subunit increases the effectiveness of this reaction (Pinz and Bogenhagen 2006).

### 2.5.3 Other proteins involved in mtDNA replication

Using an *in vitro* reconstituted replisome Korhonen and co-workers demonstrated that addition of the mitochondrial single-stranded DNA-binding protein (mtSSB) stimulated the replication significantly (Korhonen et al, 2004). The cloning of human mtSSB showed that it is a functional homologue of the SSB of *Escherichia coli* (Tiranti et al, 1993). Similar mtSSB proteins have been identified from rat (Pavco and van Tuyle, 1985), *Xenopus laevis* (Tiranti et al, 1991), yeast (van Dyck et al, 1992), *Drosophila* (Thömmes et al, 1995) and mouse (Li and Williams, 1997). Many processes of DNA metabolism result in generation of single stranded DNA regions. SSBs, which have high specificity to ssDNA but no sequence specificity, stabilize these regions. MtSSB was shown to be essential for mtDNA maintenance in yeast, since deletion of the yeast gene RIM1 resulted in total loss of mtDNA (van Dyck et al, 1992). Later this was confirmed by RNAi studies in *drosophila* S2 cells, where reduction of mtSSB protein levels generated strong mtDNA depletion (Farr 2004). The crystal structure of the human mtSSB suggested that the protein wraps mtDNA around a homotetrameric protein consisting of ~16kDa subunits (Yang et al, 1997). MtSSB stimulates synthesis of mtDNA (Mignotte et al, 1988; Genuario and Wong, 1993) by increasing the primer recognition (Thömmes et al, 1995) and processivity (Farr et al, 1999) of pol γ. Further it functions in helix destabilization by preventing separated strands from renaturation. A single *in vitro* study suggests that human mtSSB might play an essential role in D-loop turnover (Takamatsu et al, 2002), however no *in vivo* evidence has reported the existence of this mechanism. MtSSB was shown to localize within nucleoids in vivo (Barat et al, 1985; Garrido et al, 2003).

For the progression of mtDNA replication supercoiled mtDNA must be made accessible for the replication machinery. DNA topoisomerases are enzymes that alter the tertiary structure of DNA without modifying its primary structure and are responsible for the disentangling of this intertwined DNA (Kellner et al, 2002). Isolation of a mitochondrial topoisomerase enzyme from rat showed that the organelle contained a distinct DNA topoisomerase activity (Fairfield et al, 1979) By now members of topoisomerase family IA and IB have been shown to be active in human mitochondria (Wang et al, 2002; Zhang et al, 2001). The TOP1mt gene, a family IB member, contains strong homology at the nucleotide level with the nuclear TOP1 gene. The resulting protein was confirmed to localize within the mitochondria after using GFP tagged reporter constructs (Zhang et al, 2001). A later study showed that an alternative transcription start site in the TOPIIIα gene results in addition of a mitochondrial targeting signal in this IA family member (Wang et al, 2002). The precise roles in mtDNA metabolism of these topoisomerases are still not elucidated. It is very likely that a topoisomerase II enzyme exists in human mitochondria since its activity was isolated from bovine heart (Low et al, 2003). Deletion mapping in patients carrying a single large-scale mtDNA deletion revealed a correlation of Topoisomerase II like cleavage sites with the breakpoints sequence. It was

hypothesized that this topoisomerase II activity was responsible for the production of this type of mtDNA deletions (Mita et al, 1990).

Another protein that is required for the end stages of both mtDNA replication and repair is a DNA ligase. For a long time there was just one report about the existence of such a mitochondrial ligase activity (Levin and Zimmerman, 1976). More then 2 decades after this isolation from rat mitochondria a DNA ligase III like protein was shown to be active in mitochondria of *X. laevis* (Pinz and Bogenhagen, 1998; Perez-jannotti et al, 2001). The human DNA ligase III gene, LIG3, encodes a nuclear and mitochondrial protein. They are translated from two different initiation codons. While the longer polypeptide is directed to the mitochondria, the shorter product goes to the nucleus (Lakshmipathy and Campbell, 1999). The human ligase III is essential in mtDNA maintenance, lowering the ligase III activity by transfecting cells with anti-sense DNA ligase III resulted in mtDNA depletion and increase of single-strand nicks (Lakshmipathy and Campbell, 2001). In the nuclei XRCC1 directs the ligase III to the single-strand nicks. Since this protein is absent in mitochondria it remains puzzling how ligase III finds the gaps in mtDNA (Lakshmipathy and Campbell, 2000). In *S. cerevisiae* a comparable translation system is effecting the CDC9 protein location (Willer et al, 1999). CDC9, a functional homologue of human DNA ligase I (Barnes et al, 1990) is the only ligase identified in yeast mitochondria where its inactivation leads to a decrease of mtDNA content (Donahue et al, 2001).

Mitochondrial transcription factor A (TFAM), a member of an HMGB subfamily in high-mobility group (HMG) protein families is in several ways implicated in mtDNA replication (Kang and Hamasaki, 2005). This multifunctional protein of DNA metabolism activates mitochondrial transcription, thus indirectly stimulating initiation of mtDNA replication (Parisi and Clayton, 1981). In organello footprint analysis demonstrated that TFAM binds conserved replication control regions within the D-loop (Ghivizzani et al, 1994). TFAM has also been proposed to regulate the D-loop structure by destabilization and in this way inhibiting the total mtDNA replication (Takamatsu et al, 2002). TFAMs affinity for holiday junctions (Ohno et al, 2000), the recombination activity of the yeast homologue Abf2p (Zelenaya-troitskaya et al, 1998) and the physical interaction with p53 (Yoshida et al, 2003) all suggest TFAM might play a role in mtDNA recombination as well. While heterozygous TFAM knockout mouse have reduced mtDNA levels, the homozygous knockout was embryonic lethal due to lack of mtDNA (Larson et al, 1998). It is not clear if this mtDNA loss is caused by affecting the mtDNA replication or DNA packaging histone like role of TFAM, or both (Fisher et al, 1992; Larsson et al, 1998). Like its yeast counter part Abf2, human TFAM is abundant enough to wrap the entire mtDNA molecule (Takamatsu et al, 2002) and it has been shown to be a core component of nucleoids (Garrido et al, 2003). Based on these observations it has been suggested that TFAM functions as main constitutive factor of nucleoid structure (Garrido et al, 2003; Alam et al, 2003). In a recent 2DNAGE based study, increased TFAM expression in human cells resulted in a decrease of RITOLS RIs. It was suggested that TFAM plays an important role in influencing the mtDNA replication mode (Pohjoismäki et al, 2006).

RNase H1 is another protein essential for mtDNA maintenance, loss of RNase H1 in mice caused failure of mtDNA replication resulting in mtDNA depletion and embryonic lethality (Cerritelli et al, 2003). In human and mice most RNase H1 is present in the nucleus, however the usage of an alternative AUG directs it to mitochondria. The precise function of RNase H1 in mtDNA metabolism is not yet known. In the asymmetric replication model (Clayton, 1982) it was proposed to remove the RNA primers for replication initiation. Based on the phenotype of the Rnase H1 -/- mouse (Cerritelli et al, 2003) it is however more likely that RNase H1 has a more significant participation in mtDNA metabolism. Another possibility is that during RITOLS mtDNA replication (Yang et al, 2002; Yasukawa et al, 2006) it functions during the initial step of maturation of mtDNA by RNA removal. The observation that RNase H1 leaves one or more ribonucleotide attached to the DNA (Rumbauch et al, 1999) is in agreement with the presence of RNA in non-replicating mtDNA molecules (Yang et al, 2002). A later publication showed that dimerization of eukaryotic RNase H1 gives processivity to the enzyme, suggesting a possible need to process long DNA/RNA hybrids during mtDNA replication (Gaidamakov et al, 2005).

## 2.6 Other processes of mtDNA maintenance
## 2.6.1 mtDNA Repair

MtDNA is estimated to have a higher mutation rate compared to DNA in the nuclei (Khrapko et al, 1997, Stuart and Brown, 2006). It has been proposed that mtDNAs association with the inner-membrane (Nass, 1969), which is the major site of ROS production, is the main contributor to this higher mutation rate (Hamilton et al, 2001). Another reasons might be the absence of a protective histone protein in mitochondria. However, the observation that TFAM is abundant enough to cover the whole genome suggests that such a protection might also be available inside the mitochondria (Takamatsu et al, 2002; Alam et al, 2003; Garrido et al, 2003). The vulnerability of mtDNA to DNA damage and the finding that mtDNA also replicates outside S-phase, when nDNA repair proteins are downregulated suggests that an extensive mitochondrial DNA repair system should exists to ensure proper mitochondrial function.

After the observation that UV induced mtDNA damage was not repaired in HeLa cells, it was suggested that there was no mtDNA repair (Clayton et al, 1974). Following this initial observation it took a long time to realize that certain kind of DNA damage could be repaired in mitochondrial extracts (Myers et al, 1988; Pettepher et al, 1991). These observations and the identification of several enzymes containing DNA repair activity in mitochondria proved that the organelle has DNA repair activity (Anderson and Friedberg, 1980; Thibodeau and Verlg, 1980; Tomkinson et al, 1988). The major DNA repair pathway working in mitochondria is base excision repair (BER)(Figure 2.6). All small non-bulky base damage is repaired using this pathway (LeDoux et al, 1992). The first step of the pathway is recognition of the modified bases by a damage specific DNA glycosylase, which leaves an apurinic site. In human at least 5 different mitochondrial enzymes are identified, uracil-DNA glycosylase (Slupphaug et al, 1993), adenine glycosylase (Slupska et al, 1999), 3-methyladenine glycosylase (Izumi et al, 1997), thymine glycol glycosylase (Takao et al, 1998) and 8-oxoguanine glycosylase (Nishioka

et al, 1999). The apurinic site generated by these enzymes is used by an AP endonuclease to catalyze incision of the DNA phosphate backbone creating a 5' abasic deoxyribose phosphate (dRP) and a free 3' hydroxyl group necessary for DNA synthesis. The main AP endonuclease in human is encoded by the gene APE1 and has been shown to localize to mitochondria (Tell et al, 2001). Another AP endonuclease, APE2, has been identified in mitochondria (Tsuchimoto et al, 2001) but this seems much less active then APE1. Its significance in mtDNA metabolism is unknown (Hadi et al, 2000). The dRP is removed by the dRP-lyase activity while the 3'OH site is used for the insertion of a new nucleotide. An *in vitro* study showed that both functions can be performed by Pol γ (Longley et al, 1998a). However these Pol γ activities have not been demonstrated *in vivo* yet. This is followed by ligation of the broken strand by mitochondrial targeted DNA ligase III (Lakshmipathy and Campbell, 1999). Based on studies on calorie-restricted mice, it has been proposed that mtBER activity is regulated by ROS production (Stuart et al, 2004). Calorie restricted mice have a lower ROS production and this is accompanied with a lower mtBER activity.
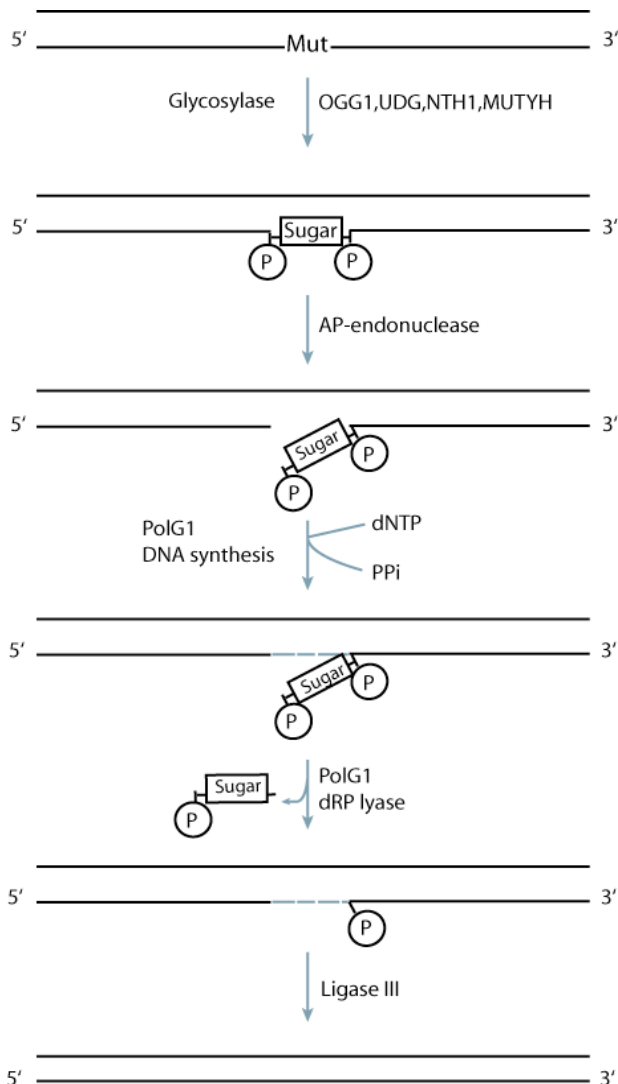


**Figure 2.6:** Enzymes and enzymatic steps of mitochondrial base excision repair. For more detail see text of section 2.6.1. Figure adapted from (Graziewicz et al, 2006).

30

Nucleotide excision repair NER is a pathway that is responsible for removal of bulky lesions that cause structural deformation of the DNA helix. NER activity has not been demonstrated inside mitochondria. However some adducts normally removed by NER in nuclear DNA are removed from mtDNA (Shen et al, 1995). Removal of nuclear DNA base mismatches is performed by mismatch repair (MMR). Although MMR activity has been identified in mitochondrial extracts from rat liver (Mason et al, 2003), no components of the MMR family were found within mammalian mitochondria. These findings suggest MMR activity is different from the corresponding activity in the nucleus (Mason et al, 2003). Although recombinational repair in the mitochondria in yeast is known for years (Foury and Lahaye, 1987), in mammalian cells it is much less clear. In a recent study two human cell lines were fused with similar nuclear background, each containing a different mtDNA mutation that abolishes oxidative phosphorylation. By growing the resulting hybrid cells without pyruvate and uridine they could select for hybrids where oxidative phosphorylation was rescued after recombination between the two genomes (D'Aurelio et al, 2004). 2DNAGE analyses showed that recombination-like intermediates are prominent in human heart muscle. These X-spike replication intermediates were resolved with the bacterial holiday junction resolvase RuvC and were suggested to represent mtDNA recombination repair intermediates (Kajander et al, 2001). MtDNA recombination has also been reported after analyzing distribution of allelic combinations in skeletal muscle of ten individuals with multiple mtDNA heteroplasmy (Zsurka et al, 2005). However the same study indicated mtDNA recombination is of low frequency in human cells.

**2.6.2 Mitochondrial nucleotide metabolism**

Precise regulation of the four deoxyribonucleoside triphosphates (dNTPs) is essential for proper DNA maintenance, since unbalance in the nucleotide pool can lead to many forms of mutagenesis (Kunz, 1988). The mitochondrial nucleotide pool is distinct from the cyctosolic nucleotide pool and is separated by the impermeable inner membrane (Figure 2.7). All cellular dNTPs can be produced by *de novo* synthesis. This pathway consists of the successive reaction of NDPs with ribonucleotide reductase (RNR) and dNDP kinases producing dNTPs. RNR activity is high during S-phase when nuclear DNA is replicated. In terminally differentiated and quiescent cells resting in $G_o$, RNR activity is strongly downregulated (Eriksson et al, 1984). In contrast to nuclear DNA replication, mtDNA replication is continuous and also replicates outside the S-phase (Bogenhagen and Clayton 1977; Song et al, 2003). In differentiated cells and outside the S-phase mtDNA replication can not be sufficiently supplied with dNTPs produced by the *de novo* pathway, but is dependent on the salvage synthesis of dNTPs. The salvage pathway makes use of deoxynucleosides derived from nutrients or degraded DNA to produce the building blocks of DNA synthesis. Mitochondria have their own salvage pathway in which nucleosides are imported by a still unknown mechanism from the cytosol and phosphorylated inside the mitochondria to dNMPs by deoxyribonucleoside kinases (Berk and Clayton, 1973). After phosphorylation dNMPs cannot leave the mitochondrial compartment due to their negative charge, making this the key regulatory step (Arner and Eriksson, 1995). There are 4 deoxyribonuleoside kinases in human identified; thymidine kinase 1 (TK1) (Bradshaw, 1983) and deoxycytidine kinase (dCK) (Chottiner et al, 1991)

are localized in the cytosol whereas thymidine kinase 2 (TK2) (Johansson and Karlsson, 1997) and deoxyguanosine kinase (dGK) (Johansson and Karlsson, 1996) are situated in the mitochondria. TK2 performs the initial step in dTTP, dUTP and dCTP production while dGK fulfils this role for dGTP and dATP synthesis.
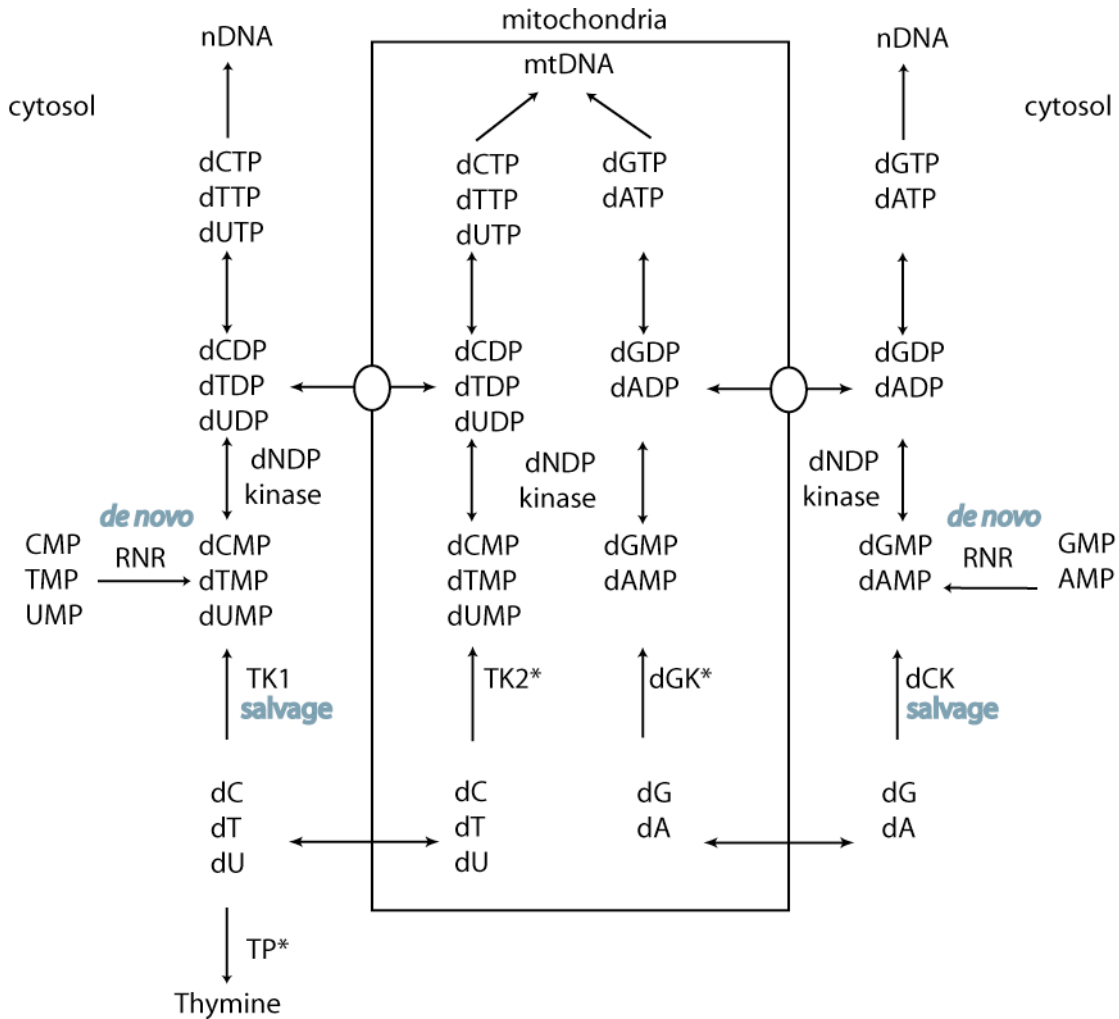


**Figure 2.7:** *Pathways of de novo and salvage deoxynucleotides (dNTP) synthesis in cytosol and mitochondria. The (\*) indication shows the genes that are associated with mitochondrial DNA disorders. For more clarity see text section 2.6.2. Figure adapted from (Marti 2002).*

Mutations in TK2 and dGK are both associated with mitochondrial DNA depletion syndrome (MDS) (Saada et al, 2001; Mandel et al, 2001). The defects in these two enzymes results in severe imbalance of the mitochondrial dNTP pool (Saada et al, 2003). As a result of this the patients have tissue specific mtDNA depletion that results in OXPHOS defects and clinically heterogeneous symptoms (Wang et al, 2005). Unlike control cells, fibroblast derived from a patient with dGK defects only synthesized mtDNA during the S-phase (Taanman et al, 2003). In optimal cell growth conditions when cells are exponentially growing this S-phase synthesis was enough to maintain

mtDNA content. However during resting state achieved by serum withdrawal the patient derived fibroblast quickly lost mtDNA. This depletion could be prevented when dGMP and dAMP were added to the cell growth medium (Taanman et al, 2003).

Mitochondrial neurogastointestinal encephalomyopathy (MNGIE) is an autosomal recessive disease associated with multiple deletions and partial depletion of mtDNA (Hirano et al, 1994). The defect in nucleotide metabolism is caused by mutations in thymidine phosphorylase (TP) (Nishino et al, 1999). TP degrades thymidine to thymine in the cytosol and a defect in this enzymatic reaction results in thymidine accumulation (Spinazzola et al, 2002). Elevated levels of thymidine cause an overproduction of dTTP by the salvage pathways, leading to nucleotide imbalances, particularly inside the mitochondria. Sequence analyses showed a high level T to C transition mutations in patient mtDNA probably as a result of increased dTTP levels (Nishigaki et al, 2003). The observation that addition of extra cellular thymidine to HeLa cells causes mitochondrial nucleotide imbalance that can lead to mtDNA deletion formation supports this disease mechanism hypothesis (Song et al, 2003). The reason why mutations in this cytosolic enzyme results in defective mtDNA metabolism could be due to the absence of nDNA replication in terminally differentiated cells, giving mitochondrial TK2 an abundance of substrate to produce dTTP (Ferraro et al, 2005).

A study of mitochondrial nucleotide pools in rat heart and muscle tissues demonstrated they are significantly different from whole cell extract nucleotide pools. *In vitro* fidelity assays showed that under these physiological conditions present in heart and muscle, mitochondria Pol γ makes 3 fold more replication errors (Song et al, 2005). This study suggests that normal mitochondrial nucleotide pool concentrations could contribute to the elevated mitochondrial mutation rate compared to the nuclear genome.

**2.6.3 Mitochondrial nucleoids**

Like chromosomal DNA in bacteria (Stonington and Pettijohn, 1971), DNA inside the mitochondria is organized in protein-DNA structures called nucleoids. Similar mitochondrial protein-DNA structures have been isolated from various organisms, *Physarum polycephalum* (Kuroiwa et al, 1977; Suzuki et al, 1982), yeast (Miyakawa et al, 1987; Kaufman et al, 2000; Chen et al, 2005), *X.laevis* (Barat et al, 1985; Bogenhagen et al, 2003) and human (Albring et al, 1977; Garrido et al, 2003; Wang and Bogenhagen, 2006). Nucleoids have fusion and fission capability and show dynamic movements within the cell which suggest that they are the mitochondrial units of inheritance (Garrido et al, 2003). While in yeast and slime molds nucleoids can easily be visualized with 4',6 diamidino-2-phenylindole (DAPI) staining, in human cells other methods have been developed to visualized nucleoids properly. Nucleoid studies in human cells have used other DNA binding dyes (Ashley et al, 2005), green fluorescent protein (GFP) tagged nucleoid proteins (Spelbrink et al, 2001) or incorporation of a thymidine analogue in mtDNA (Garrido et al, 2003; Magnusson et al, 2003). These studies showed punctate nucleoids that are consistently scattered inside the mitochondrial network in both yeast and man. The yeast nucleoid is much larger (~0,5 $\mu$m) compared to the human nucleoid (~0,065 $\mu$m) (Iborra et al, 2004). Also the number of mtDNA molecules shaping the

nucleoids and the amount of nucleoids per cell is significant different between yeast (~1-2 mtDNA/nucleoid, ~40-60 nucleoids/cells) and man (~2-15 mtDNA/nucleoid, ~300-800 nucleoids/cell) (Legros et al, 2004).

Attempts to purify the nucleoid protein components from yeast, *X.Laevis* and human (Bogenhagen et al, 2003; Garrido et al, 2003; Chen et al, 2005; Wang and Bogenhagen, 2006) have to date revealed only two corresponding proteins. However this might be due to use of diverse purification strategies and the dynamic characteristics of nucleoids. It is thought that TFAM and its yeast and *X.leavis* homologues are the core-packaging elements of the nucleoids. MtSSB was, as can be expected from a protein with strong single stranded DNA binding affinity, found to be part of the nucleoids. In cultured human cells, antibodies against mtSSB and TFAM showed a strong co-localization of these proteins with other nucleoid markers  (Garrido et al, 2003).

Studies on budding yeast have been most successful in revealing the nucleoids components. Most nucleoid proteins were identified using *in organello* formaldehyde cross-linking to mtDNA followed by density gradient centrifugation (Kaufman et al, 2000; Chen et al, 2005). Some of the proteins found in these studies have a known function in mtDNA metabolism like the yeast mtDNA polymerase gamma Mip1p and RNA polymerase Rpo41p. The other group of proteins found to localize in the yeast nucleoids were directly or indirectly involved in the citric acid cycle or amino acid metabolism.  Functional studies of some of these proteins show they contain a dual function and they couple mtDNA maintenance to metabolic regulation. (Chen et al, 2005; Chen and Butow, 2005). Two extensively studied bi-functional nucleoid proteins are Ilv5p, functional in amino acid biosynthesis and aconitase (AcoIp) a krebs-cycle enzyme. Both enzymes are known to be metabolically regulated, Ilv5p expression is enhanced during amino-acid starvation and AcoIp expression goes up when a cell is shifting to respiratory conditions. An hypothesis is that increased levels of Ilv5p and AcoIp causes alterations in mtDNA packaging resulting in a more metabolically favorable conformation of the nucleoid (Chen and Butow, 2005). Although not identified in nucleoids isolations mitochondrial membrane protein (Mmm1p) is essential for mtDNA stability and the fusion protein Mmm1p-GFP localizes to dot-like structures that are flanking the nucleoids (Hobbs et al, 2001). Mmm1p forms a complex with the outer membrane proteins Mdm10p and Mdm12p. It was shown that Mmm1p complex connects the mitochondria to the outer membrane and cytoskeleton (Boldogh et al, 1998, Meeusen and Nunnari, 2003). The connection of this complex with nucleoids seems to involve cooperation of the inner membrane proteins Mdm31p and Mdm32p (Dimmer et al, 2005). Deletion of any of these membrane proteins results in nucleoid instability.

Much less is known about the composition and dynamics of the human nucleoids. Many of the yeast nucleoid proteins previously discussed donot have homologues in human, suggesting human nucleoids are somewhat different. Enrichment of human nucleoids resulted in co-purification of both subunits of Pol γ (Garrido et al, 2003). However, Pol γA-GFP did not show a puncate mitochondrial localization (Spelbrink et al, 2000), perhaps suggesting the mitochondrial polymerase is only recruited to the nucleoids when needed for mtDNA replication or repair. Twinkle-GFP however seems to be a structural

nucleoid protein, showing co-localization with mt-SSB, TFAM and mtDNA. Human nucleoids from HeLa cells can be separated in two fractions based on their interaction with the cytoskeleton. Immunoaffinity purification using antibodies against TFAM and SSB with the fraction that didnot contain cytoskeletal protein contamination resulted in identification of ~20 human nucleoid proteins. Several of these proteins were proteins potentially involved in DNA metabolism , which included two helicases Suv3 and DHX30. Similar to yeast, other identified proteins consisted of chaperones and metabolic enzymes. Although these findings are similar to the observations in yeast almost all dual-function proteins seem to differ in yeast and human nucleoids (Wang and Bogenhagen, 2006).


**2.6.4 mtDNA copy-number regulation**

MtDNA copy-number is regulated under different physiological and developmental conditions. The mechanisms of this regulation in addition appears tissue and cell specific (Veltri et al, 1990), but remains poorly understood. It has been suggested that the high copy number of mtDNA is a protection mechanism preventing accumulation of mutations to critical threshold levels (Moraes, 2001). Cytoplasmic hybrids and patient cell lines containing mtDNA molecules with large scale deletions have a higher mtDNA copy number then their full length wild-type counter part (Spelbrink et al, 1997; Tang et al, 2000). It was suggested that the total mtDNA mass is regulated rather then amount of mtDNA genomes (Tang et al, 2000).

One of the factors that seems to be involved in the regulation of mammalian mtDNA copy-number is TFAM. Elimination of the yeast TFAM homologue (Abf2p), or RNAi inhibition of the *Drosophila* homologue in Schneider cells results in a decrease of ~50% in mtDNA content (Zelenaya-Troitskaya et al, 1998; Goto et al, 2001). This mtDNA depletion was achieved without affecting the mitochondrial *in vivo* transcription activities or steady state levels of mitochondrial mRNAs, suggesting that in these organisms TFAM does not function directly in transcription-mediated initiation of replication. A 3-fold increase of Abf2 expression resulted in a 2 fold increase in mtDNA copy number (Zelenaya-Troitskaya et al, 1998), clearly demonstrating Abf2's role in yeast mtDNA copy number regulation. However, this increase was not unlimited since a 10-fold Abf2 overexpression resulted in mtDNA depletion. These observations can be explained by TFAM's mtDNA packaging functions. Whereas a 2-fold increase stabilizes mtDNA genomes, more TFAM binding makes mtDNA inaccessible for the replication machinery (Zelenaya-Troitskaya et al, 1998). Similarly, TFAM heterozygous knockout mice have ~about 40% copy number reduction (Larsson et al, 1998), while heterologous overexpression of human TFAM results in mtDNA copy number increase in mice (Ekstrand et al, 2004). The authors chose to overexpress the human variant of this protein because it poorly stimulates mouse transcription initiation *in vitro* without loosing its strong mtDNA binding activity. They suggest that copy number goes up because mtDNA molecules are stabilized by human TFAM binding. It canot however be excluded that the increase of mtDNA levels is a result of stimulation of replication initiation *in vivo*. In *drosophila* cultured cells a similar effect was observed, after stimulation of transcription

initiation by overexpressing the TFB2M homologue the mtDNA copy increased (Matsushima et al, 2004). Transient transfection of TFAM in HeLa and HEK cells was shown to result in a moderate overexpression of TFAM without change in mtDNA levels (Maniura-Weber et al, 2004), suggesting that other factors are required to increase mtDNA copy number in human cells.

Another suggested mechanism is that different modes of mtDNA replication could regulate copy- number (Moraes, 2001). The recently proposed non-asynchronous replication model of mtDNA suggests two modes of replication based on analysis of RIs by 2DNAGE (Yang et al, 2002; Bowmaker et al, 2003; Yasukawa et al, 2005). Some RIs are very sensitive to RNase H and a detailed mapping has recently shown that these RIs represent RNA:DNA hybrids essentially along the entire length of mtDNA (Yasukawa et al, 2006). This RNA-rich replication mode is in fact very alike the asynchronous model with the distinction that the lagging-strand is mostly protected by RNA and not by mtSSB. The RIs that are not degraded by RNase H treatment have been suggested to arise by a conventional strand-coupled replication mode. The suggestion that cells recovering from mtDNA depletion induced by EtBr mainly use the RNA-poor replication mode shows that mtDNA replication rate might be regulated by a switch in replication mode. It was suggested based on these results that the RNA-poor mode represents a faster mode. A recent publication showed that in cultured human cells TFAM inhibition with RNAi results in somewhat increased RNA-rich RIs, whereas TFAM overexpressing caused accumulation of RNA poor RIs. It was suggested that TFAM levels might regulate the rate of mtDNA replication (Pohjoismäki et al, 2006). However, a moderate overexpression of TFAM did not result in mtDNA copy-number increase.

Transgenic mice that overexpress the wild-type mtDNA helicase Twinkle show an mtDNA copy-number increase in heart and muscle (Tyynismaa et al, 2004). In contrast, RNAi inhibition of Twinkle reduced mtDNA copy number in cultured human cells. These experiments suggest that Twinkle may be a key regulator of mtDNA copy-number in mammals, it is possible that Twinkle is rate limiting for replication initiation.

In yeast the mechanisms controlling mtDNA copy-number are starting to be unraveled. Mounting evidence points to the nucleotide pools as regulator of mtDNA copy number. Overexpression of the large subunit of ribonucleotide reductase (RNR) in *S.cerevisiae* decreases the amount of rho- mutants observed in diploids that contain a single copy of the yeast pol γ MIP1 gene (Lecrenier and Foury, 1995). This result suggested that nucleotide pool concentration is rate limiting for mtDNA replication in yeast. Another way to induced RNR and with that *de novo* dNTP production is by a transduction pathway mediated by MEC1 and the protein kinase RAD53 (Huang and Elledge, 1997). The deletion of a DNA helicase RRM3 causes replication fork stalling that is followed by phosphorlyation activation of RAD53 (Ivessa et al, 2003). This activation results in increased dNTP synthesis by RNR and a higher mtDNA copy-number (Taylor et al, 2005). Similarly, deletion of the RNR inhibitory protein Sml1 and over-expression of RNR-1 subunit caused an increase in *de novo* synthesis of dNTPs (Zhao et al, 2001; O'rourke et al, 2005), which was accompanied by an mtDNA copy-number increase (Taylor et al, 2005). Taken together these results suggest that mtDNA copy number in

yeast is regulated by the activity of the RAD53/Mec1 pathway, which in turn regulates RNR activity. Deletion of Abf2 did not alter the RNR induced copy-number change showing this is an Abf2 independent regulation of mtDNA copy number (Taylor et al, 2005). The RAD53/Mec1 pathway is very conserved and it might be that the human homologous protein kinases will affect mtDNA copy-number in a similar fashion, however no experimental evidence has been reported for this yet.

## 2.7 Mitochondrial DNA related diseases

Mitochondrial DNA related diseases are genetically heterogeneous disorders with a very broad phenotypic range. In most of these disorders there is involvement of the central nervous system resulting in neurological presentations such as stroke-like episodes, seizures or movement disorders. Often there is a combination of neurological and skeletal muscle defects and therefore these are often referred to as encephalomyopathies. In some cases other organs are affected such as eyes, heart and kidney (Larsson and Clayton, 1995). The cause of these diseases is generally believed to be caused primarily by a defective OXPHOS system. This can be a result of mutations in genes involved in mitochondrial protein synthesis or complex assembly or more directly by mutations in mtDNA encoded subunits of the OXPHOS system. Since the coding potential of mtDNA is very limited, many disorders that result in an OXPHOS defect are primarily caused by mutations in nuclear genes. One special group are those nuclear genes involved in mtDNA maintenance which can secondarily cause mtDNA mutations.

Several aspects of mtDNA genetics that are very important in understanding the phenotypic presentation of those disorders that directly or indirectly involve mtDNA are related to the high mtDNA copy-number per cell. Patients with mtDNA mutations often have two mtDNA populations, one mutated and the other wild-type (heteroplasmy). It has been shown that the distribution of mtDNA molecules to daughter cells during cell division can be explained by a stochastic process (Chinnery and Samuels, 1999). The result of this so-called segregation is that the proportion of mutant versus wild-type mtDNA molecules may be changed in the daughter cells. Given enough cell divisions segregation can result in cells with purely mutant and purely wild type mtDNA, assuming no selection of any kind takes place. Tissues with a high proportion of wild-type molecules are less likely to suffer from OXPHOS defects than those carrying high proportion of mutant mtDNA molecules (Chinnery et al, 1997). The proportion of mutated mtDNA molecules that can be tolerated is tissue specific and this threshold is usually lower in tissue with a high energy demand. The threshold is reached when the wild-type mtDNA molecules can no longer compensate for the mutant mtDNA. The threshold level is usually between 70-90%, depending on the mutation, tissue and nuclear background of the patient (Hammans et al, 1993). Many of the mtDNA related diseases are progressive probably caused by a steady increase of the proportion of mutant mtDNA in post-mitotic tissue and/or by a slow decline of OXPHOS capacity associated with normal ageing (Trounce et al, 1989; Larsson et al, 1990; Cooper et al, 1992).

**2.7.1 mtDNA point mutations**

A far majority of pathogenic point mutations in mtDNA encoded OXPHOS subunits are linked with four diseases. Leber's hereditary optic neuropathy (LHON), maternally inherited Leigh's syndrome (MILS), mitochondrial encephalopathy lactic acidosis and stroke-like episodes (MELAS) and neuropathy, ataxia, and retinitis pigmentosa (NARP) (Wallace et al, 1988; Santorelli et al, 1993; Goto et al, 1990; Holt et al, 1990).

LHON is a disease characterized with acute loss of central vision that presents itself in early adulthood and was the first disease identified to associate with mtDNA point mutations (Wallace et al, 1988). Most mtDNA point mutations are found in the genes ND1, ND4 and ND6, which all code for subunits of complex I (Wallace et al, 1988; Huopenen et al, 1991; Johns et al, 1992). Although in most LHON patients symptoms are restricted to the optic system, the most severe mutations can lead to additional neurological features like movement disorders and dystonia (Nikoskelainen et al, 1995). Enzymological analysis on lymphoblast mitochondria extracted from LHON patients revealed that several patient mutations affect complex I activity. The demonstration that these LHON mutations result in a complex I that is more sensitive to complex I inhibitors that interact with CoQ10, suggests that LHON is a consequence of defective interaction between CoQ10 and complex I, which might lead to increased ROS production (Brown et al, 2000). The finding that neuronal cell lines harboring LHON mutations show a neuron-specific ROS increase supports the hypothesis that increased ROS generation results in damage to the optic nerve (Wong et al, 2002). Haplogroup J shows an increase in LHON disease penetrance (Howell et al, 2003) and since a founder event was excluded these findings imply a true mtDNA background effect (Carelli et al, 2006).

In NARP, the most common clinical manifestations are damage to the peripheral nervous system, ataxia, and abnormalities in the pigment epithelium of the retina. A base modification at position 8993 of the mitochondrial genome in the ATP6 gene is the most frequent observed mutation in NARP patients, although other rarer mutations also cause the disease (Holt et al, 1990). NARP patients normally contain 70%-90% of mutant mtDNA (Makela-Bengs et al, 1995). When the mutant load is even higher (>95%) this ATP6 mutations may cause the fatale infantile encephalopathy MILS (Santorelli et al, 1993). The 8993 mutation impairs the $F_o$ portion ATPase resulting in a defective ATP production (Tatuch and Robinson, 1993). Based on blue native electrophoresis studies others have suggested that the ATP synthesis defects in patients carrying the 8993 mutation are a result of incorrect assembly of complex V (Nijtmans et al, 2001). NARP patients skin fibroblasts were found to contain high superoxide dismutase activity, suggesting higher ROS production (Geromel et al, 2001). The demonstration that antioxidants partially recover ATP synthesis in cells harboring the NARP/MILS mutation suggests that the ROS increase plays an important pathogenic role (Mattiazzi et al, 2004).

Point mutations in tRNA genes that affect protein synthesis result most often in two distinct syndromes, mitochondrial encephalopathy lactic acidosis and stroke-like episodes (MELAS) and myoclonus epilepsy with ragged red fibres (MERRF). MELAS is a

progressive neurodegenerative disease, the most common features are included in the name of this disorder. The most common A to G mutation is located in the tRNA$^{Leu(UUR)}$ gene at position 3243 (A3243G) (Goto et al, 1990) and is found in approximately 80% of MELAS cases (Goto et al, 1992). A low mutation load of 5-30% in blood and 40% in muscle is associated with relatively mild features like diabetes, and deafness (Ouwenland van den et al, 1995; Goto, 1995), whereas a high proportion of >85% mutant mtDNA results in MELAS. A high proportion of A3243G containing mtDNA molecules results in defects in complex I and IV. This MELAS mutation has been shown to affect the post-transcriptional modification of the tRNA $^{Leu(UUR)}$ (Helm et al, 1999) and reduces the stability of mitochondrial encoded proteins (Janssen et al, 1999). These observations make it plausible that faulty incorporation of amino acids at leucine codons is the primary defect of this disease. The most common mutation found in MERRF, the A8344G nucleotide substitution, results in similar tRNA lysine defects as previously described in MELAS disease concerning tRNA leucine (Shoffner et al, 1990; Enriquez et al, 1995). Although in both syndromes the biochemical defect is a tRNA modification it results in different clinical phenotypes in the patients. The reports that the primary affect of the MELAS mutations is on complex I activity and MERFF mutations normally show strongest reduction in complex IV activity might explain these different clinical features (Morgan-Hughes et al, 1995; Enriquez et al, 1995). A study of rho$^o$ cell transformants carrying the MERFF (8344) mutation revealed a defect in the specific aminoacylation capacity of the affected tRNA$^{lys}$. This resulted in accumulation of abnormal complex I pepetides by premature termination of translation, specific at lysine residues (Enriquez et al, 1995).


## 2.7.2 Large-scale mtDNA rearrangements

Sporadic mtDNA deletions were the first mtDNA mutations identified to be associated with a group of disorders with diverse clinical symptoms (Holt et al, 1988). In order of increasing severity, single mtDNA deletions gives rise to pure mitochondrial myopathies including chronic progressive external ophthalmaplegia (CPEO) (Holt et al, 1988; Moraes et al, 1989), Kearns-Sayre syndrome (KSS) (Zeviani et al, 1988) and Pearson's syndrome (Rotig et al, 1989). The more severe symptoms in Pearson's syndrome and KSS might be explained by the observation that mutant mtDNA molecules are found in many different tissues while in CPEO they are confined to muscle. In CPEO eye and skeletal muscle are affected while KSS includes also neurological features like ataxia. Pearson's syndrome has additional defects in bone marrow and pancreas. These sporadic deletions are heteroplasmic and the size of the deletion can vary. However, the most common 4977 base pair deletion is found in 30% of the cases (McFarland et al, 2002). The proportion of deleted mtDNA molecules can change with time and can vary between different tissues. Studies on cybrid cell lines harboring an mtDNA deletion suggests that a mutation load of 60% is required to cause OXPHOS dysfunction (Hayashi et al, 1991). However, this study was done in Hela cells and other studies have shown that this threshold is different in other cell-lines or tissues (Bourgeron et al, 1993; Spelbrink et al, 1994). The general thought is that a sporadic deletion arises in the oocyte and the resulting embryo will have the deleted mtDNA (Shapira, 2006). It has also been reported

that mtDNA duplications might co-exist in a patient with a single mtDNA deletion, although one report suggested an clinical association between duplication and KSS (Poulton et al, 1994), in other reports they did not appear pathogenically significant and were suggested to represent an intermediate step in the formation of mtDNA deletions (Cormier et al, 1990; Rotig et al, 1992; Brockington et al, 1995). Further studies are needed to confirm a phenotype association of mtDNA duplications. Due to the dense gene content of mtDNA, deletions will generally result in the loss of several tRNA genes as well as OXPHOS protein coding genes. This will often result in a combined OXPHOS deficiency i.e. affecting several of the OXPHOS enzymes (Shapira, 2006; Laforet et al, 1995)

### 2.7.3 Somatic mtDNA mutations

The mitochondrial theory of aging suggests that mtDNA alterations accumulate during ageing resulting in OXPHOS defects. This is suggested to increase ROS production leading to more mtDNA mutations followed by an even more pronounced OXPHOS defect and ROS increase (Trounce et al, 1989). This vicious cycle then ultimately results in loss of tissue function through apoptosis and energy insufficiency. Somatic mtDNA mutations and deletions have been reported by many studies to accumulate in human aging (Cortopassi et al, 1992), neurodegeneration (Corraldebrinski et al, 1994) and cancer (Polyak et al, 1998) in numerous tissues. However, there is some doubt about the significance of these mtDNA point-mutations and deletions since mostly they do not reach a high heteroplasmy level. The maximum proportion of the common mtDNA 4977 base pair deletion was found in ageing muscle at ~0,1%, much lower than the high levels generally associated with mtDNA disorders (see previous sections). Studies on single muscle fibers show that clonal expansion of one mutation during life can occur, most often these are fibers without cytochrome c oxidase (COX) activity (Del Bo et al, 2002; Coller et al, 2002; Nekhaeva et al, 2002). It seems logical that accumulation of mtDNA point-mutations or deletions will lead to OXPHOS defects in individual cells however it is unclear to what extent this will effect the whole organism. The mechanism of mutant mtDNA accumulation is unclear. One hypothesis is based on the observation that smaller genomes containing an mtDNA deletion repopulated faster than the full length genomes in cell lines recovering from EtBr-induced mtDNA depletion (Diaz et al, 2002). However, mathematical modeling has shown that the clonal expansion can be achieved by random genetic drift (Elson et al, 2001).

The mitochondrial theory of aging is in part based on the knowledge that ROS is a byproduct of normal OXPHOS energy production. Moreover, studies done in cybrid cell lines harboring a MERFF (Mattiazzi et al, 2004) or LHON (Wong et al, 2002) mutation have shown that some OXPHOS defects result in increase of ROS production. This supports at least one step of the vicious cycle, showing that OXPHOS defects are in some cases able to increase ROS production.

To start testing the mitochondrial aging hypothesis, stable HEK cell lines expressing a proofreading deficient Pol γ A were generated (Spelbrink et al, 2000). Like expected the lack of proofreading during mtDNA replication resulted in an increase of random point

mutations in genes encoding for the respiratory chain subunits, however these mutations did not affect the oxygen consumption of the cells. The reason for this might be that long-term cell culture leads to selection against the most deleterious mutations. The study showed proof of the principle that an exonuclease deficient Pol γ A would result in an accumulation of mtDNA point mutations. Subsequently, in knock-in mice expressing a similar proofreading deficient mouse Pol γ A, a similar accumulation of point mutations was observed in several tissues of several month old mice. In this case, a decrease in OXPHOS activity was observed (Trifunovic et al, 2004). This result showed there is a direct link between increased somatic mutation levels and a decrease in OXPHOS activity at least in mice. In accordance with the mitochondrial theory of ageing this increased mtDNA mutation level resulted in a premature aging phenotype in the mice. However, unlike what this hypothesis would predict the OXPHOS defects in these mice didnot lead to an increase of ROS (Kujoth er al, 2005; Trifunovic et al, 2005). The level of somatic mtDNA mutations in these 'mutator' mice increases linearly over time and not exponentially as expected when ROS production increases over time. The authors of this work suggest that ROS is not involved in the premature aging induced by increased levels of mtDNA somatic mutation, suggesting no role for oxidative stress in this aging model. Their idea is that aging is caused directly by a decrease in OXPHOS resulting in energy insufficiency and apoptosis. This is in agreement with a majority of studies on OXPHOS activity in different age groups, which show an age-related decrease in OXPHOS activity (Cottrell and Turnbull, 2000). In contrast to these results, others have shown that catalase enzyme targeted to mouse mitochondria reduces ROS by decomposition of $H_2O_2$ and increases the mice life span (Schriner et al, 2005). This affect was absent when the enzyme was directed to the peroxisome or nucleus suggesting that mitochondrial ROS can play a role in determining longevity in mammals. More studies are required to conclusively determine the precise role of ROS in aging.

Specific studies on the common 4977 base pair mtDNA deletion revealed that areas of the brain with high dopamine metabolism, especially in the substantia nigra, have a high prevalence of this mtDNA deletion (Soong et al, 1992). However, in this study the whole nigral tissue was taken and only one type of mtDNA was analyzed. Recently two reports have come out that use different PCR methods to measure all mtDNA deletions in individual dopaminergic neuron (Bender et al, 2006; Kraytsberg et al, 2006). High levels of mtDNA deleted molecules were detected and showed a correlation with age. In some neurons obtained from older individuals the deletions exceeded 60%, which has been shown to be the threshold for impaired OXPHOS activity (Kraytsberg et al, 2006; Hayashi et al, 1991). Long-range PCR detected many different mtDNA deletions, however each individual neuron showed only one type, confirming that these are somatic mtDNA deletions (Bender et al, 2006). The cause of the specific accumulation of mtDNA deletions in dopamingeric neurons has been suggested to involve dopamine breakdown, a process in which $H_2O_2$ is formed that might cause double-strand break formation (Manfredi, 2006). In mice double-strand breaks can result in mtDNA deletions, since skeletal muscle expression of a mitochondrially targeted restriction endonuclease that cuts mtDNA twice resulted in the accumulation of mtDNA deletions, which mostly involved one of the restriction site (Srivastava and Moraes, 2005). The free ends

generated by the endonuclease were often recombined with a region at the 3'end of the D-loop, deleting part of the mitochondrial genome (Srivastava and Moraes, 2005).

Somatic mtDNA point mutations, of which some homoplasmic, are increased in some cancers (Polyak et al, 1998; Fliss et al, 2000; He et al, 2003). The mechanisms by which these mutations accumulate and if they have any influence on cancer development is not known. This accumulation of mutations and the high copy-number of mtDNA have lead to the suggestion it could function as an easy detection for cancer in diagnostics (Fliss et al, 2000)

During the aging of a human, mtDNA mutation load increases with the strongest accumulation of base changes occurring in the control region (Jazin et al, 1996). The same conclusions came out of a later study, investigating control region mutation levels in skin fibroblasts obtained from individuals ranging from 20-week fetal to 101 years old. A sensitive method was used that could pick up single point mutations at low heteroplasmy levels (Michikawa et al, 1999). Besides a relatively low proportion of mutations at random positions, it revealed a clear accumulation of mutations at specific positions of the mtDNA control region. The most striking mutation that was specific to skin fibroblast of older individuals (>65 years) was the T to G transversion at position 414, which is positioned within the L-strand promoter. The accumulation of this specific mutation was absent in skeletal muscle, heart (Wang et al, 2001), brain (Murdock et al, 2000) and leukocytes (Zhang et al, 2003) of controls. Interestingly one report showed that a majority of Alzheimer's disease brain tissue harbors the T414G mutation, which goes along with a reduced L-strand transcription (Coskun et al, 2004). In skeletal muscle two other tissue specific control region mutations accumulate with age, A189G and T408A, which are critical sites for either mtDNA replication or transcription (Wang et al, 2001). The C150T transition mutation that is near a proposed origin of H-strand replication (Robberson et al, 1972) was shown to be significant higher in leukocytes of Italian centenarians (Zhang et al, 2003). 5'strand analysis of nascent heavy-strand replication initiation in fibroblast showed that this C150T mutation results in change of position of mtDNA replication initiation. It was suggested that this change in replication initiation gives a survival advantage, which would explain its high prevalence in centenarians. The accumulation of tissue and site specific control region mutations is interesting indeed, however all mechanistic explanations for this accumulation till so far are purely based on speculation (Coskun et al, 2003)

### 2.7.4 mtDNA maintenance disorders

Accumulation of multiple mtDNA deletions and mtDNA depletion results in several disorders but are primarily associated with nuclear gene mutations. Even though this gives rise to heterogeneous clinical features it involves neurological syndromes that have quite some common characteristics, symptoms such as PEO, myopathy, neuropathy, epilepsy, ataxia and hepatopathy.

Mitochondrial depletion syndrome (MDS) was first identified in two related patients with fatal mitochondrial disease. The affected tissues, muscle or liver contained low mtDNA content (Moraes et al, 1991). Most patients suffering from these severe syndromes, which are phenotypically heterogeneous, die early in childhood and onset of the disease is mostly in the first year of life. Later studies revealed that mutations in at least 6 different nuclear encoded genes cause MDS. The different gene defects result in mtDNA depletion in different tissues, which lead to different tissue specific disease features. Three genes found to be associated with MDS, encode enzymes involved in nucleotide metabolism, thymidine phosphorylase (TP) (Nishino et al, 1999), thymidine kinase 2 (TK2) (Saada et al, 2001) and deoxyguanosine kinase (dGK) (Mandel et al, 2001) and have been discussed previously in relation to their importance in understanding the fundamental mechanisms of mitochondrial nucleotide metabolism.

Alper's syndrome is a MDS associated with Pol γ defects that has brain and liver involvement (Naviaux et al, 1999). Most patients with Alper's syndrome have been found to be compound heterozygotes, carrying two pathogenic mutations one on each Pol γ allele. The A467T mutation seems the most common mutation in this group of patients (Nguyen et al, 2005). *In vitro* studies showed that this linker region mutation result in 96 % reduction of polymerase activity and a weak interaction with the accessory subunit (Chan et al, 2005a; Luoma et al, 2005). Mono-allelic expression of A467T caused by a premature stop-codon on the other allele leads to an earlier disease onset then di-allelic expression of a homozygous A467T patient, suggesting this mutant protein is still capable at least to some extent to replicate mtDNA (Chan et al, 2005b).

The SUCLA2 gene encodes for the β subunit of the TCA cycle enzyme succinyl-CoA synthetase. Although it is probably not directly involved in DNA metabolism, mutations in this gene can lead to MDS (Elpeleg et al, 2005). Succinyl-CoA synthetase functions as heterodimer. The yeast homologue of the α-subunit is part of the mtDNA nucleoid (Chen and Butow, 2005). Immunoprecipitation studies with antiserum directed against succinyl-CoA synthetase showed that succinyl-CoA interacts with nucleoside diphosphate kinase (NDPK), which catalyzes the transformation from nucleoside diphosphates to nucleoside triphosphates (Kowluru et al, 2004). Although it seems likely that the mtDNA depletion is a result of a secondary NDPK dysfunction the actual mechanism is currently unknown. MDS can also be the result of a protein defect in the inner mitochondrial membrane protein MPV17 (Spinazzola et al, 2006; Karadimas et al, 2006). MPV17 knockout mice have decreased mtDNA copy number in liver, muscle and brain, which is similar to the patients suffering from this gene defect (Weiher et al, 1990; Spinazzola et al, 2006). The precise role of MPV17 in mtDNA maintenance still needs to be elucidated.

Gene defects that result in the accumulation of multiple mtDNA deletions without a decrease in mtDNA content cause autosomal dominant (ad) or recessive (ar) progressive external ophthalmaplegia (PEO). The most common feature of this late onset disease (20-40 years) is progressive weakness of the extraocular muscle. Besides general muscle weakness additional symptoms vary and can include ataxia, hearing loss, ptosis, parkinsonism and psychiatric conditions (Zeviani et al, 1989; Melberg et al, 1996; Suomalainen et al, 1997). Symptoms are progressive, probably due to continuous

increase of deleted mtDNA molecules in post-mitotic tissue. Muscle biopsies from patients often show ragged red fibers and a decrease of OXPHOS activity (Suomalainen et al, 1997).

The gene for the muscle and heart specific isoform of the adenine nucleotide tranlocator (ANT1) was the first of four identified nuclear genes that associates with ad/arPEO (Kaukonen et al, 2000). Mutation in ANT1, a protein responsible for the export of ATP out of the mitochondrial matrix in exchange for ADP (Li et al, 1989), results in a relatively mild form of adPEO (Kaukonen et al, 2000). The connection between ANT1 and mtDNA maintenance is still unclear. A possibility is that the adPEO mutations in ANT1 increases ROS production, which might lead to mtDNA instability. More likely however is that ANT1 regulates ATP concentration inside the mitochondria, and that too high or low levels of ATP might distrubt mtDNA replication maybe even lead to an imbalanced intramitochondrial dNTP pool. This is supported by the demonstration that mutations in thymidine phosphorylase in MNGIE result not only in mtDNA depletion but also in mtDNA deletion accumulation (Nishino et al, 1999), showing that nucleotide imbalance can indeed lead to multiple mtDNA deletion accumulation. Studies in yeast have shown that the corresponding adPEO mutations in the yeast homologue of ANT1 (Aac2) result in a similar mtDNA deletion accumulation and OXPHOS defect (Fontanesi et al, 2004).

Recessive PolG1 mutations are the most common in sporadic cases with multiple mtDNA deletions (Agostino et al, 2003). Dominant PEO POLG1 mutations often give additional symptoms like movement, voice and swallowing disorders (Luoma et al, 2004). All reported dominant mutations that cause PEO are localized in the polymerase domain, while recessive mutation can be found throughout the gene sequence (Graziewicz et al, 2006). Biochemical characterization of some dominant mutations *in vitro*, showed a strong decrease in polymerase activity (>70%) and lower nucleotide selectivity and affinity (Graziewicz et al, 2004). This all might leads to loss of processive synthesis by pol γ *in vivo*. The same study showed that some of these mutant pol γ variants bind DNA more efficiently than the wild-type variant. This might explain why these mutant variants compete in a dominant negative fashion. ArPEO patients with POLG1 mutations are often compound heterozygous and have two mutant alleles (Graziewicz et al, 2006). Some of these mutations that cause arPEO are also found in other syndromes like Alper's (Nguyen et al, 2005) and Sensory ataxia with neuropathy dysarthria and ophthalmoparesis (SANDO) (Gago et al, 2006). One single mutation in POLG2, the accessory subunit of pol γ, has been found in a PEO patient (Longley et al, 2006). In this report it was shown that the mutant subunit is unable to stimulate the processivity of the catalytic subunit.

The precise mechanism by which these four nuclear gene defects result in the accumulation of multiple mtDNA deletions is still not clear. One study showed that in a single COX negative muscle fibers one single mtDNA deletion could be detected (Moslemi et al, 1996) They suggested a two-hit mechanism in which a mutant nuclear gene predisposes to deletion formation followed by clonal expension in post-mitotic tissue. The multiple deletions might accumulate because the nuclear gene defects cause a

strong increase in mtDNA rearrangement formation, which also accumulate in normal conditions during ageing but at very low levels. Another possibility is that the PEO gene mutations result in a defect in the mechanism to lose these mtDNA rearrangements. Alternatively there is a selective advantage for the mtDNA deleted molecules in post-mitotic tissue of PEO patients (Spinazzola and Zevianni, 2005). More study is needed to elucidate the precise mechanism of multiple deletion formation in adPEO patients for this purpose an *in vivo* and *in vitro* model system is needed.

## 3.      Aims of the study

AdPEO is a mitochondrial disease associated with multiple mtDNA deletions, which manifests itself in its weakest form as an isolated ophthalmoparesis and ptosis (Suomalainen et al, 1995).  Previous studies reported at least 3 distinct loci for this disease (Suomalainen et al, 1995; Kaukonen et al, 1996; Kaukonen et al, 1999). The 4q34-linked adPEO disease gene was identified as the heart and skeletal muscle specific isoform of the adenine nucleotide translocator (ANT1) (Kaukonen et al, 2000). Other genes associated with adPEO remained unidentified at the onset of this PhD. For lack of knowledge of the disease genes, nothing was known about the mechanisms behind the accumulation of multiple mtDNA deletions.

The overall aims of the work were:


1   To identify and characterize novel proteins involved in adPEO and mtDNA maintenance.

2   To learn more about the manner in which nuclear gene defects in adPEO patients lead to the accumulation of multiple deletion in the mitochondrial genome.

3   To characterize the general functions of mtDNA replication factors in mtDNA maintenance.


Fulfilling these aims was expected to lead to a further basic understanding of the general function and composition of the mtDNA replication machinery. It was furthermore anticipated to provide insight into the mechanisms of mtDNA maintenance in both health and disease.

# 4. Materials and methods

## 4.1 Patient samples

We obtained DNA and tissue samples from affected individuals and healthy family members with informed consent following the ethical guidelines of the medical center in question.

For PCR amplification of the five exons of the Twinkle gene a salting out procedure was used to extract DNA from blood lymphocytes (Davis et al, 1986). The patient material used to measure the mtDNA mutation load and for the mapping of mtDNA deletions consisted of mtDNA isolated from six POLG1 (1-6) and five Twinkle (6-11) affected PEO patients. In the case of patient 7 mtDNA was extracted from extraocular eye muscle tissue however for all the other patients and controls skeletal muscle tissue was used for DNA extracting. Additional mtDNA samples were extracted from blood lymphocytes (patient 1 and 2) and frontal cortex tissue (patient 8). Controls Δ1-5 are sporadic patients carrying a single mtDNA deletion resulting in PEO (Δ1-4) or mild KSS (Δ5). Patient Δ5 was 5 years of age at time of biopsy, while the others (Δ1-4) were adults. Healthy controls 1-6 are unaffected relatives of PEO patients 3-5, control 6 being a distant relative and controls 1-5 being brothers and sisters, thus sharing the same mitochondrial background. Healthy controls, 7-10 are direct relatives of patients 7-11 all sharing mtDNA transmitted through the same maternal line. In the statistical analysis comparing healthy controls with PEO patients, external eye muscle from patient 6 was counted as skeletal muscle.

## 4.2 PCR amplification and sequencing

PCR of Twinkle genomic or cDNA was performed with routine protocols. Oligonucleotides were as follows (see Table 4.1): for exon 1, Tw-5'UTR2 and TwIn1R; for exon 2, Tw-In1F and Tw-In2R; for exons 3 and 4, Tw-In2F and Tw-In4R; for exon 5, Tw-In4F and Tw-3' UTR1. All PCR-products were column purified (Qiaquick PCR purification kit, Qiagen) before genomic sequencing. Cycle sequencing was performed with the same oligonucleotides as for PCR amplification plus the following additional exon 1 oligonucleotides: Tw-Ex1-1, Tw-Ex1-2, Tw-Ex1-3 and Tw-Ex1-4. For sequencing we used BigDye-terminator chemistry (Applied Biosystems). Samples were run on ABI Prism 310 and 3100 instruments, with kit reagents supplied by the manufacturer.

The mtDNA control region and cytochrome *b* gene (cyt *b*) from PEO patients and controls were PCR amplified using the primers described in table 4.1, giving rise to a 576 bp (control region 35L-611H) or a 624 bp (cyt b 14682L-15516H) fragment. The reaction conditions in a total volume of 50 $\mu$L were as follows: 100 ηg of total genomic DNA, 1,5 U of Pfu DNA polymerase (Promega), 250 $\mu$M dNTPs, 10 $\mu$M of each primer and the buffer supplied by the manufacturer. The PCR amplification conditions consisted of an initial denaturation step of 2 min at 95°C, followed by 21 cycles of 30s at 95°C, 1 min at 57°C and 2 min at 72°C. Finally an extension step of 10 min at 72°C. For the detection of multiple mtDNA deletions, mtDNA was amplified using the primer FR31 in combination with 611H-long, 15996H, 15516H or 14338H, resulting in multiple PCR fragments according to sizes of the deletions. PCR reaction conditions were as described above but

with 5% dimethylsulfoxide (DMSO) and 20 $\mu$M of each primer. The amplification conditions were: initial denaturation of 2 min at 93°C, 29 cycles of 30s at 93°C, 1min at 57°C and 7 min at 72°C, with a final extension step of 15 min at 72°C. Deletion containing fragments, control region or cyt *b* fragments were cloned using the TOPO Zero-Blunt PCR cloning kit (Invitrogen). Plasmids from individual bacterial colonies were isolated with the Macherey-Nagel Nucleospin Robot-96 plasmid kit on a Tecan multipipeting robot. The presence of an insert was confirmed by restriction digestion and agarose gel electrophoresis. After this the insert was sequenced on an ABI Prism 3100 DNA sequencer, using BigDye terminator chemistry (Applied Biosystems) and M13 forward and reverse oligonucleotides or mtDNA specific primers. Sequences were analyzed using the Seqman™II software (DNASTAR). Alterations of the Cambridge reference sequence present in all clones were excluded as polymorphisms. The control region homopolymeric D310 tract (nt 303-315) was discarded from our analysis since it is known to be highly polymorphic. Full length and deleted mtDNA sequence characteristics like base composition, nucleotide word counts, palindrome and direct repeat searches were carried out with the Java interface (J)EMBOSS package (Rice et al, 2000) at http//www.hgmp.mrc.ac.uk/Software/EMBOSS/Jemboss/.

| Name | Sequence 5' to 3' | notes |
|---|---|---|
| Tw-5'UTR2 | GTTGGTCTAGTGAAGGCACG | Exon1Forward |
| Tw-In1R | CCCACTTGCTTTTGTCACCTG | Exon1Reverse |
| Tw-In1F | GTCTTGGTTTCAAGGGTAGG | Exon2 Forward |
| Tw-In2R | GATATGTCTGGGAAAGCAAGG | Exon2 Reverse |
| Tw-In2F | GGTGGTCTAGAGACAACTTG | Exon3-4 Forward |
| Tw-In4R | GGACAGTCAAGACGATTAAGG | Exon3-4 Reverse |
| Tw-In4F | CTTTCTGCTTTGCTCATGTCC | Exon5 Forward |
| Tw-3' UTR1 | CCTTGCAGAGTTTTATGCTCC | Exon5 Reverse |
| Tw-Ex1-1 | AGTGGATGGGTCGGAGGGGC | Exon1 69Forward |
| Tw-Ex1-2 | GCAGAGAAAGTGGCCTGTGG | Exon1344Reverse |
| Tw-Ex1-3 | CCAGAATTTGAGGACAGCGAG | Exon1 440Forward |
| Tw-Ex1-4 | CTTACCCCCTGCCTTACTCC | Exon1 871Forward |
| 35L | GGAGCTCTCCATGCATTTGG | Control region |
| 611H | CAGTGTATTGCTTTGAGGAGG | Control region |
| 14682L | CACGGACTACAACCACGACC | Cytochrome b |
| 15516H | GTATAATTGTCTGGGTCGCCTAGG | Cytochrome b |
| FR31 | CTTCCCACAACACTTTCTCGGCCTA | COXI |
| 611H/Long | CAGTGTATTGCTTTGAGGAGGTAAGCTACATAA | Control Region |
| 15996H | GCTTTGGGTGCTAATGGTGG | tRNA$^{pro}$ |
| 14338H | GGTGGTTGTGGTAAACTTTAA | ND6 |
| EcoRI/Tw1F | CCGGAATTCTAGGAATGTGGGTCCTCCTCC | EcoRI site |
| SacII/Tw-1R | AGCCGCGGCTTTGAACGCTTGGAGGTGTC | SacII site |
| HindIII/Tw-1R | CCCAAGCTTCTTTGAACGCTTGGAGGTGTCTG | HindIII site |
| Tw-dup1F | ACGATGCTTCTTGGTGCGACC | |
| Tw-dup1R | CGAAGCTGCCGGAAAGATACG | |
| Tw-ex2-1F | CTCCTCACCCAGGTCTGTTC | |
| Tw-ex2-1R | CCCTGCCCTCTCATTCTTTG | |
| Diagnostic primer | CTGGAGGCCCTGACGGCTTCAATCTT*GG*TC | |

Table 4.1: Primers used in PCR and sequencing. For details see text.

Total cellular DNA was extracted to measure mtDNA mutation load from various transgenic 293-Flp-In™T-REx™ cells expressing POLG1 variants after 0 and 60 days of doxycycline induction (3 ng/ml) by Proteinase K digestion and subsequent phenol/chloroform extraction and ethanol precipitation. Point mutation levels in the control and cyt *b* region of mtDNA were measured as described above.


### 4.3. Cloning of expression constructs

The full Twinkle coding-region cDNA was amplified with oligonucleotides EcoRI/Tw1F and SacII/Tw-1R or HindIII/Tw-1R (Table 4.1). The newly created restrictions sites were used to clone the cDNA by standard procedures in the vectors pEGFP-N1 (Clontech), pCMV-Tag4 (Stratagene) or pcDNA3.1(-)/MycHisA (Invitrogen). The Twinkle mutant variants were generated, using site directed mutatgensis (Spelbrink et al, 2000).

Similar to the Twinkle variants, the full length cDNA of POLG1 variants were originally cloned in the pcDNA3.1(-)/Myc-His A, using site directed mutagenesis as previously described (Spelbrink et al, 2000).

The full coding-region cDNA from mtSSB was amplified and cloned into the pcDNA 3.1(-)/Myc-His A. POLG2-HA was constructed in pcDNA3.1-hygro (+) by recloning the full-length cDNA from a non-commercial vector. All Constructs were re-cloned in the pcDNA5/FrT/TO vector taking advantage of the two *Pme*I restriction sites in the multiple cloning sites of the pcDNA3 vectors. The resulting fusion proteins contains the sequence of the protein precursor followed by either the Myc-His tag or the HA tag. These pcDNA5/FrT/TO constructs were used for stable transfection of 293 Flp-In™T-REx™ (Invitrogen) cell-lines. Complete DNA sequencing of the insert and flanking sequences verified all plasmid-constructs.


### 4.4 Genotyping

We found all Twinkle disease mutations by direct sequencing of DNA samples on both strands. To detect the disease causing Twinkle 39-bp duplication of the linker region in family members we amplified the genomic DNA with two primers Tw-dup1F and Tw-dup1R (Table 4.1), with Dynazyme II DNA polymerase (Finnzymes, Finland) following the manufacturer's directions. The PCR products were analyzed on a 1,5 % agarose gel, yielding a 39 bp longer (177 bp) fragment in the PEO family with this specific linker region duplication. To exclude a common polymorphism, we analyzed 40 pooled DNA samples of 10 control subjects each which yielded a 138bp fragment. The A475P Twinkle patient mutation is a result of a G-to-C change which gives rise to a new recognition site for *Hae*III at position 1423 of the Twinkle cDNA. We amplified a exon 2 PCR fragment with the primers Tw-Ex2-1F and Tw-Ex2-1R (Table 4.1). Digestion or this 413 bp long PCR product with *Hae*III resulted in two fragment (35bp and 378 bp), however in the heterozygous (A475P) mutant the original 378 bp fragment was digested to give two bands of 201 bp and 177 bp. The R354P Twinkle patient mutation is a result of G-to-C change at position 1061 which creates an *Ava*II site by using a diagnostic primer and Tw-In1R. *Ava*II digestion of this 283 bp PCR product resulted in two fragments. Similar procedures were used to identify other Twinkle mutations in family members of index patients.

**4.5 Cell culture and transfections**

A549 adenocarcinoma, HEK293T and 143Bosteosarcoma cells were cultured as described (Spelbrink et al, 2000). Cells were seeded in six-well plates at about 50% density 1 or 2 days prior to transfection. For the transfections of HEK293T and A549 cells we used 10 µl of lipofectamine (Gibco) diluted into 1 ml of Opti-MEM (Gibco) following the manufacturer's instructions. We added 2 ml of fresh medium five hours after transfection and refreshed this medium 19 hours later. FuGENE 6 (Roche) was used according the manufacturer's protocols for the transfection of 143B cells.

The Flp-In™ T-REx™ 293 host cell line (Invitrogen), an HEK-293 variant that includes a Flp-In recombination site at a transcriptionally active locus, was grown in D-MEM medium (Cambrex Bioscience) with 2 mM L-glutamine (Cambrex Bioscience), 10% FCS (Euroclone,) and 50 µg/ml uridine (Sigma) supplemented with 100 µg/ml zeocin (Invivogen, San Diego, USA) and 15 µg/ml blasticidin (Invivogen) in a 37 °C incubator with 8,5% $CO_2$. 48 hours prior to transfection cells were split into 10 cm plates and grown to ~80% confluence on the day of transfection in medium lacking antibiotics. Cells were co-transfected with the appropriate pcDNA5/FrT/TO construct and pOG44 (invitrogen), a plasmid encoding the Flp-recombinase enzyme necessary for targeted stable intergration (Dymecki, 1996). The TransFectin™ (BioRAD) reagent was used to transfect the cells according to the manufacturer's recommended protocol. Briefly, 3.6 µg pOG44 and 0.4 µg of a pcDNA5/FrT/TO construct were diluted to a 1,5 ml in OptiMEM-I serum-free medium (Invitrogen); similarly, 30 $\mu$l TransFectin™ was diluted to a volume of 1,5 ml. The diluted DNA and TransFectin™ were combined and incubated 20 minutes at room temperature and added to the cell culture medium. Six hours after transfection, the medium was replaced with medium lacking antibiotics. One day after transfection this medium was replaced by medium supplemented with the selective antibiotics hygromycin (150 µg/ml) (Invivogen) and blasticidin (150 µg/ml). Selective medium was replaced every 2 days for cell maintenance.

Inducible cell lines expressing variants of POLG1, Twinkle, POLG2 and mtSSB were created according this method. The obtained cell lines were verified for inducible expression of the transgenic protein by Western blotting. To induce expression the indicated amount of doxycycline (Sigma) was added to the growth medium.

**4.6. Monitoring GFP reporter gene expression and immunocytochemistry**

Fluorescence microscopy, cell preparation and Mitotracker Red staining were done as described (Spelbrink et al, 2000). The *in vivo* staining of mtDNA with EtBr was done essentially as described (Hayashi et al, 1994; Davis and Clayton 1996), however using a longer incubation. Briefly, cells were grown on coverslips and then incubated for 2 hours at 37 °C with EtBr (1 µg/ml) in cell culture medium, followed by 2 washes of medium and 2 washes of phosphate buffered saline (PBS). The mounting was performed in PBS containing 5% Sucrose. For confocal laser microscopy, we used a Perkin Elmer/Wallace Ultra *View* LCI system.

Immunofluorescent detection was done as previously described (Garrido et al, 2003). Monoclonal anti-DNA antibody AC-30-10 (PROGEN) was used to detect mtDNA and

was done as described (Legros 2004). Secondary antibodies were anti-mouse IgM-Alex Fluor®568 (DNA) and IgG-Alex Fluor®488 (Invitrogen).

## 4.7 Cell lysis, PAGE and western blotting

Cell lysates were prepared and analyzed to verify the protein expression by immunoblotting after SDS-PAGE (Spelbrink et al, 2000). Primary monoclonal antibodies against the c-myc (Roche) or HA tag (BAbCO) were used. The peroxidase-coupled secondary antibody (horse-anti-mouse) was obtained from Vector Laboratories. Chemiluminescence detection was performed as previously described (Spelbrink et al, 2000).

Minimal denaturing SDS gels (6% polyacrylamide) contained 0,02% SDS instead of 0,1 %. About 40 μg of protein was mixed 1:1 before loading with a 1:10 dilution of standard SDS sample buffer (but with 12% glycerol) and was not heat denaturated. The western blotting was performed according standard protocols (Spelbrink et al, 20000). The native PAGE markers were obtained from Amersham/Pharmacia.

## 4.8 Twinkle isolation and DNA helicase activity

*In vitro* helicase activity was measured in highly enriched Twinkle preparations derived from 293 Flp-In™T-REx™ cells. The cells treated for 2 days with 50 ng/ml doxycycline (Sigma) were harvested and mitochondria were isolated by hypotonic lysis and differential centrifugation (Garrido et al, 2003). The mitochondrial pellet was lysed in high salt buffer (50 mM KxPO4 pH 7,0, 1M NaCl, 1% Triton X-100, 1x complete Protease inhibitors EDTA-free, Roche) and sonicatd on ice (Sonics Vibra-cell, 1 min 40% amplitude, 1s pulses with 2s break). The insoluble DNA pelleted for 10 min at 12.000g (4°C). The supernatant was incubated for 1-2 h (4°C) with Talon metal-affinity resin (Clontech) to allow binding of the His-tagged proteins. We washed the resin twice with high salt buffer and twice with low salt buffer (25 mM TrisHCl pH 7,6, 40 mM NaCl, 4,5 mM MgCl$_2$, 10% glycerol, 100 mM L-Arginine) containing 20 mM Imidazole. Elution was performed with low salt buffer containing 500 mM Imidazole. The supernatant of this step was shock-frozen in liquid nitrogen and stored at –80°C.

In the helicase assay the standard substrate used was a radioactively end-labeled 60 bp oligonucleotide hybridized to M13 ssDNA (5'ACATGATAA-GATACATGGATGAGTTTGGACAAACCACAACGTAAAACGACGGCCAGTGCC3 '), generating a 20 bp double-stranded stretch with a 40bp 5' overhang. Other substrates were prepared by annealing two oligonucleotides of 25 bp, generating 20-30 bp double-stranded stretches. The helicase assay was performed by incubating 1 ng Twinkle protein in 40 μl helicase buffer (25 mM Tris pH 7,6, 40 mM NaCl, 4,5 mM MgCl$_2$, 100 mM L-Arginine-HCl pH 7,6, 10% glycerol, 3 mM UTP, 1mM DTT and 5 μM unspecific oligonucleotide) with 2 amol substrate for 30 min at 37°C. The reaction was terminated by addition of 10 μl loading buffer (90 mM EDTA, 6% SDS, 30% glycerol, 0.25% bromophenol blue). The resulting 20 μl reaction mix was separated over a 15% acrylamide gel in 1x TBE.

## 4.9 Real time PCR (QPCR)

The mtDNA content per cell was determined by real time PCR of cytochrome b (cyt b) with the gene encoding the amyloid precursor protein *APP* as a nuclear standard as described (Tyynismaa et al, 2004). DNA extracts were obtained from cells by lysis, proteinase K treatment and subsequent isopropanol precipitation. The copy numbers of cyt *b* and *APP* were determined in a duplex Taqman PCR on an Abiprism 7000 (Applied Biosciences) using pCR® 2.1-TOPO® (Invitrogen) containing the cyt *b* and *APP* amplicon as standards.

## 4.10 Brewer-Fangman 2D neutral/neutral Agarose electrophoresis

Mitochondria were purified using cytochalasine B (Sigma-Aldrich) and nucleic acids were purified essentially as described (Yasukawa et al, 2005). The obtained mtDNA was digested with either *Hinc*II, *Bcl*I or *Acc*I (Fermentas) and were in some experiments further treated with RNaseH or S1 nuclease (Fermentas). The fragments were separated using neutral 2D-AGE after which the gels were blotted and hybridized with 32P-labelled DNA probes (Yasukawa et al, 2005). The migration of the *Hinc*II digest fragment (mtDNA nts 13636 – 1006) was studied with a probe containing the mtDNA nts 14846 – 15357. The migration of the *Bcl*I digest fragments nts 3658 – 7657 and nts 8591 – 11921 were studied using probes containing mtDNA nts 4480 – 4984 and 9275 – 9774 respectively. The migration of *Acc*I digest fragments nts 15255 – 1504 and 1504 – 6286 were studied with probes of mtDNA nts 652 – 1156 and 3314– 3845 respectively.

# 5.    Results

## 5.1 Identification of a T7 primase/helicase homologue Twinkle

AdPEO is an inherited adult-onset disorder with large-scale deletions of mtDNA in post-mitotic tissue resulting in exercise intolerance, muscle weakness, peripheral neuropathy, deafness, ataxia, cataracts, and hypogonadism (Suomalainen et al, 1995). This mtDNA maintenance disorder has been shown to associate with mutations in the adenine nucleotide translocator (ANT1) at locus 4q34-35 and POLG1 at locus 15q25 (Kaukonen et al, 2000, Van Goethem et al, 2001). Previous studies involving several adPEO families also identified an association between the disease and a region on human chromosome 10q24 (Suomalainen et al, 1995; Li et al, 1999). This region was screened for open reading frames (ORFs) and a novel gene similar to bacteriophage T7 primase/helicase gp4 (T7gp4) was identified, C10orf2 (later renamed PEO1). A single ORF of about 2,1 kb, with an ATG start codon was identified by cDNA sequencing, exon and intron prediction programs and EST database searches. cDNA amplification followed by direct sequencing revealed two different transcripts, one produces the full-length protein Twinkle and a second transcript yielded a shorter splice variant named Twinky that included a 4 amino acids extended exon 4 followed by a stop codon but lacked exon 5. The resulting protein is a predicted 102 amino acids shorter product. The full-length protein Twinkle has a molecular mass of 77 kD and consists of 684 amino acids. Further, similar proteins were identified in *Caenorhabditis elegans*, *Drosophila melanogaster* and mouse. The homology of these proteins with T7gp4 was strongest in de C-terminal helicase domain. Mouse Twinkle shares a high level of sequence identity with human Twinkle (89%) and like in human the use of a downstream exon 4 splice-donor site is predicted to yield a splice variant Twinky (Tyynismaa et al, 2004). A human multitissue Nothern blot hybridised with a full-length PEO1 cDNA probe showed that human Twinkle expression was highest in skeletal muscle and pancreas.

## 5.2 Twinkle co-localizes with mitochondrial nucleoids and includes helicase activity

After the identification of this PEO1 gene we generated several expression constructs for the characterization of the protein products. MitoProt II predicted that the PEO1 gene contains a putative amino-terminal mitochondrial targeting sequence thus suggesting that Twinkle is a mitochondrial protein. Human cells (HEK293T,143B and A549) transfected with reporter constructs of full length Twinkle containing a Green Fluorescent Protein (GFP) tag showed punctate fluorescence restricted to multiple discrete foci (Figure 5.1). Mitotracker red staining and confocal microscopy showed that this fluorescence is strictly mitochondrial. Transient transfection of a Twinkle-GFP fusion protein with a 21 N-terminal amino acids deletion predicted to contain part of the putative mitochondrial targeting signal, resulted in uniform cytoplasmic fluorescence. MtDNA stained with Etbr co-localized with the punctate green fluorescence obtained after Twinkle-GFP transfection, showing that Twinkle localizes to mtDNA. Similar nucleoprotein structures have earlier been identified in fungal, plant and yeast mitochondria and are referred to as nucleoids. The clear night-sky like pattern observed after transfection of the Twinkle GFP fusion protein, its homology to T7gp4 and the mitochondrial nucleoid localization characteristic of the gene product of PEO1 gave the name to the full-length protein

Twinkle (for T7gp4-like protein with intramitochondrial nucleoid localization). Twinky-GFP did not give a punctate green fluorescence suggesting that the C-terminal region of Twinkle is at least partially responsible for its nucleoid localization. Twinkle-GFP transfection of different cell types showed there is variation in amount of nucleoids between different human cell-lines, 143B cell contained the highest number of foci whereas A549 contained fewer of this nucleoprotein structures compared to 293T cells.

**A**          **B**          **C**
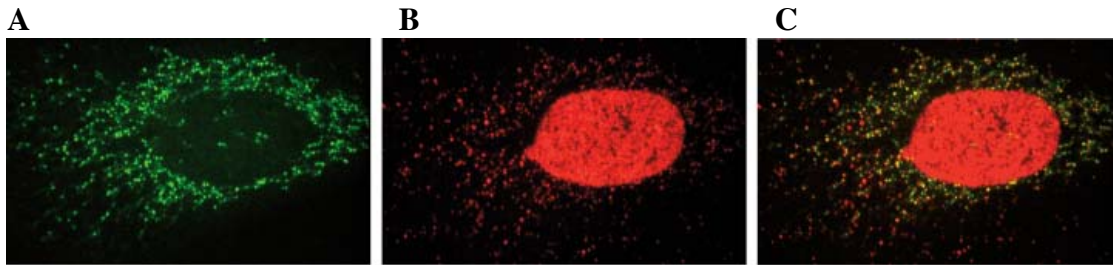


**Figure 5.1:** *Twinkle co-localizes with mtDNA. Human U2OS were transfected with a Twinkle GFP expression construct. Cells were fixed and mtDNA detected using a DNA-specific antibody. a) TwinkleGFP; b) anti-DNA; c) overlay of panels a and b showing extensive overlap of TwinkleGFP and DNA fluorescent signals. (Images by JN Spelbrink).*

The Twinkle protein belongs to a class of hexameric helicases. Most of these hexameric helicases are involved in DNA replication and are needed to unwind the duplex DNA ahead of the DNA polymerase (Patel and Picha, 2000). Our results showed that mitochondrial extracts isolated from HEK293T cells were able to unwind an M13 oligonucleotide substrate. This *in vitro* DNA helicase activity was enhanced by 50% in mitochondrial extracts obtained from cells that overexpress Twinkle after transient transfection. The splice variant Twinky did not enhance DNA helicase activity in crude mitochondrial extracts. The stimulation of this activity was most pronounced with ATP.

**5.3 Twinkle mutations associate with adPEO disease and multiple mtDNA deletions**
Two large adPEO families had previously been shown to have linkage to chromosome 10q24 (Suomalainen et al, 1995; Li et al, 1999). The Twinkle gene was located within this critical region and the protein product showed mitochondrial localization and furthermore suggested a direct involvement in DNA metabolism making it a very strong PEO disease gene candidate. For these reasons we analysed the 5'- and 3'UTR, exons and splice junctional sequences of this gene. In one of the two families originally used for the linkage analysis, of Pakistani origin, we found by direct sequencing that the disease phenotype within this family segregated with an amino acid change A475P in the gene for Twinkle, which was originally called *C10orf2* but has been renamed *PEO1* . Analysis of ~600 control individuals from which 88 from Pakistani/Indian origin did not reveal this mutation. In the second 'linkage' family, of Finnish origin, direct sequencing revealed a heterozygous 39 bp duplication, resulting in an in frame duplication of aminoacids 352 till 364. PCR amplification and agarose gel electrophoresis analysis showed an additional 39bp longer PCR fragment in all affected family members. After this we performed direct sequencing of about ~70 families diagnosed with definite or

possible adPEO. In 10 of these families, 9 additional different Twinkle mutations were identified (Figure 5.2) that were not detected in appropriate control groups of 100 individuals. In one family the disease segregated with an amino acid A359T mutation, resulting in a relatively mild phenotype in heterozygous individuals but an earlier disease onset and more severe symptoms in the homozygous patients.
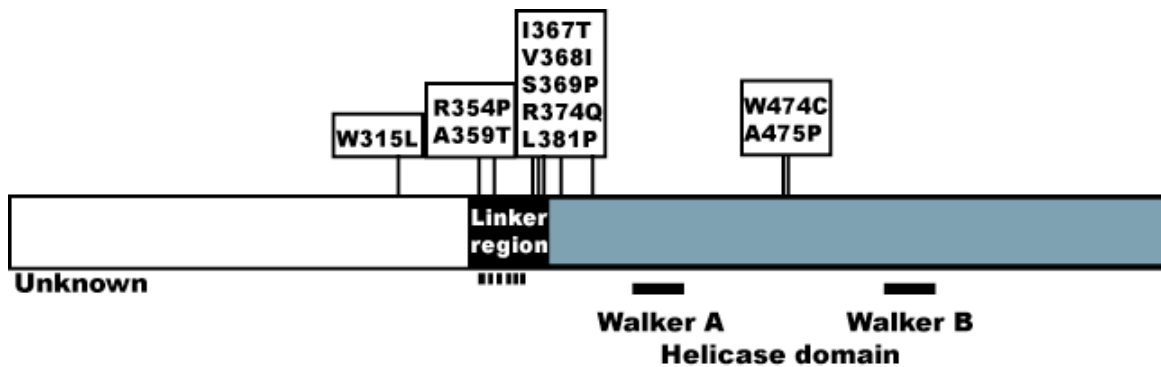


**Figure 5.2:** *Schematic presentation of the human PEO1 gene and indication of the positions of the mutations that we mapped in our sequence analysis of different adPEO families. The sequence duplicated in a finnish adPEO family is indicated with a dotted bar. The V368I mutation was later shown not to segregate with adPEO (Arenas 2003).*

### 5.4 Multimerization and localization of mutant Twinkle

After we found that Twinkle mutations are associated with adPEO, we continued investigating if these gene mutations obstruct Twinkle's nucleoid localization. For this purpose Twinkle-EGFP fusion proteins were constructed containing adPEO patient mutations. Transient transfection of Twinkle W474C and the 13 amino acids linker region duplication (AA 352-364) showed a similar punctate fluorescence as the wild-type Twinkle-EGFP. In later studies using the same experimental approach similar punctate nucleoid localization was obtained for the 9 other identified adPEO Twinkle mutations (unpublished data). Suggesting that the adPEO Twinkle mutations do not affect mitochondrial and nucleoid localization.

Most of the Twinkle mutations we identified (8 out of 11) are confined to a small region of 30 amino acids referred to as the linker region. The T7gp4 helicase/primase functions as a hexamer and also includes such a linker region. Various experiments including native gel electrophoretic analysis have shown this region is essential for oligomerization and function of the protein (Guo et al, 1999). Native gel electrophoretic analysis was used to examine Twinkle's ability to form hexamers. For this purpose 293T cells were transfected with Twinkle-FLAG constructs and protein lysates were run on a minimal denaturing SDS-PAGE gel. After western blotting and antibody detection it was noted that most Twinkle-FLAG migrated as multimers, consisting of 2 to 6 subunits. Two adPEO (W474C and duplication 352-364) mutations did not affect this multimerization

significantly. Interestingly, Twinky-FLAG fusion protein showed very poor mulitmerization.

## 5.5 High mutation load in the mtDNA control region of PEO patients
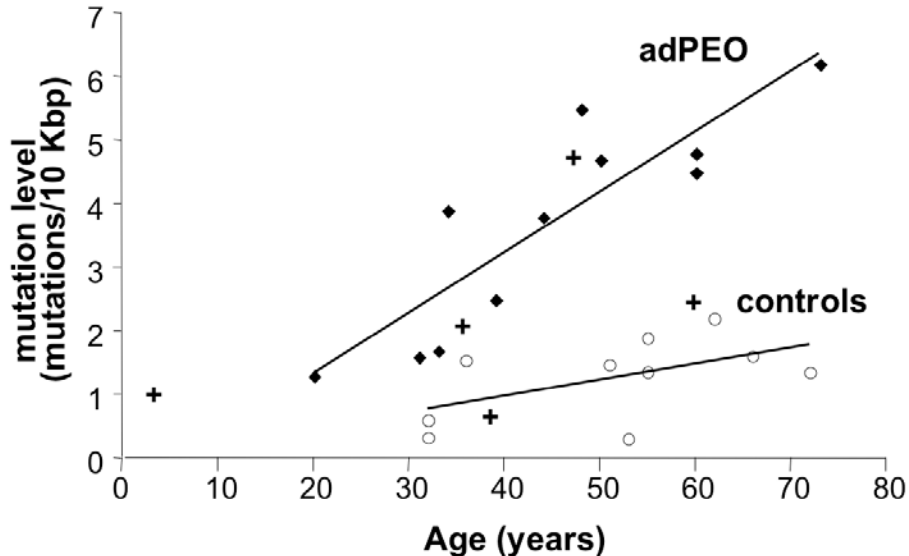


**Figure 5.3:** *Mitochondrial control region mutation levels in PEO patients (filled diamonds), healthy controls (open circles) and single deletion controls (+), related to age at the time of the biopsy. Lines for the PEO patients and healthy controls indicate trend lines (Wanrooij et al, 2004). Reprinted from original article (II), copyright (2004), by permission of Oxford University Press.*

Although several adPEO disease genes were now identified it remained unclear by which mechanism mutations in PEO1, POLG1 or ANT1 result in the accumulation of multiple mtDNA deletions. One hypothesis suggested that mtDNA replication infidelity could stimulate slipped mispairing, by which it would promote deletion formation (Ponamarev et al, 2002). This suggestion was based on the observation that the Y955C POLG1 mutation showed a decrease in DNA synthesis accuracy *in vitro*. In order to asses the fidelity of POLG1 in PEO *in vivo* we measured mtDNA point mutation levels in ad/arPEO patients with multiple deletions by high fidelity mtDNA PCR amplification, followed by cloning and sequencing of individual plasmids. Results were compared to the levels in age-matched individuals. We analysed the point mutation levels in two regions of the mtDNA genome, most of the control region (nts 35-611) and a fragment of the cyt *b* gene (nts 14682-15516) by determining the amount of mutations found per 10 Kbp DNA sequence. We found comparable low mutation levels (0,2-1,4 mut/10kbp) in the cyt b region of ad/arPEO samples and controls. Partial analyses of two other regions of mtDNA, the Hyper Variable Region 1 (16000-16569) and COXI (7203-7445) gave similar low mutation levels. Control region mutation levels in skeletal muscle mtDNA

obtained from control individuals revealed a weak tendency to increase with age (Figure 5.3). However this age related increase in control region mutation levels was much more pronounced in ar/adPEO patients. From all investigated patients the youngest ar/adPEO patient analysed (aged 20 years) showed the lowest control region mutation level (1,6mut/10 Kbp), which was comparable to the levels found in elderly (>60 years) control subjects. The older ar/adPEO patients (>35 years) had an average mutation level of 4,4mut/10kbp (ranged from 2,4 to 6,3 mut/10kbp), which was significantly (p < 0,0002) higher then the 1,5mut/10kbp found in the control samples (all lower then 2,2mut/10kbp) (Figure 5.4). This increased control region mutation load seemed confined to post-mitotic tissue since leukocytes mtDNA isolated from 2 different ad/arPEO patients showed low control region mutation levels. Further we determined the mutation levels in 5 patients with a sporadic single mtDNA deletion and found elevated control region mutation levels (4.3, 2.1 and 1.7mut/10kbp) in 3 of this group of patients.
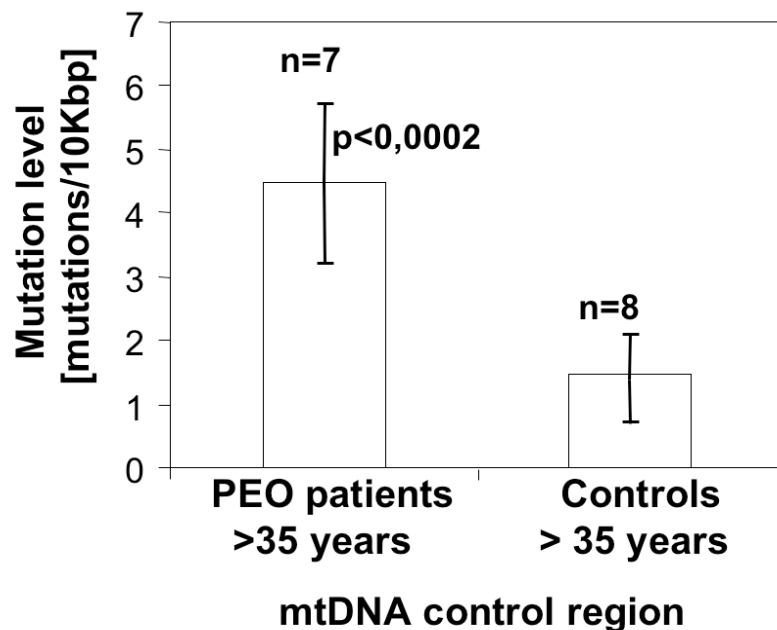


**Figure 5.4:** *Number of different control region mutations in muscle mtDNA of PEO patients >35 years old grouped together, and in the >35-year-old controls. The probability (P) of both groups being identical was calculated using the unpaired Student's t-test. The indicated results show that mutation level differences between both groups are highly significant (Wanrooij et al, 2004). Reprinted from original article (II), copyright (2004), by permission of Oxford University Press.*

**5.6 Distribution of the control region mutations in ar/adPEO skeletal muscle mtDNA**

We observed a non-random distribution of the control region mutations that accumulate in ad/ar PEO patients. The majority (81%) of control region mutations we found in the six POLG1/PEO patients that we analysed were located in a region that contains the heavy strand replication origin ($O_H$) and the conserved sequence block I (CSBI) (nts 150 to 250) (Figure 5.5). The most remarkable mutations were found in patient 1 (compound

heterozygous R3P and A467T) and patient 2 (Y955C). In Patient 1 we identified several duplications varying from 2-32 bp partly overlapping with CSBI. The mtDNA of patient 2 contained an A to G mutational hotspot with heteroplasmy of about 20% at positions close to the $O_H$, 178, 183, 185 and 189. Patient 3 is unrelated to patient 2 and carrying the same Y955C POLG1 mutation, but no such A to G mutational hotspot was found in patient 3. However the A183G mutation was found at high heteroplasmy levels. This mutation was also found in 15% or higher heteroplasmy in patients 2,3,4,5 and 6, which are all individuals carrying a POLG1 mutations. In all controls and Twinkle/PEO patients this mutation was not detected at all.
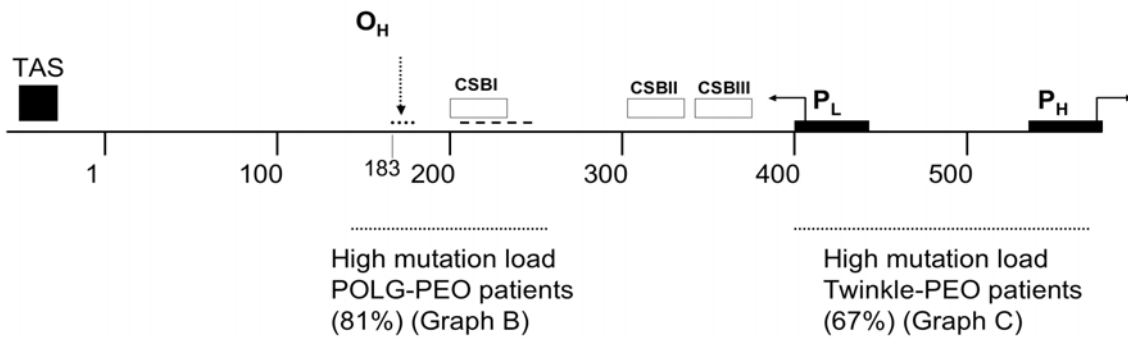


**Figure 5.5:** *Non-random distribution of control region mutations. Diagram of the mtDNA control region and mutational hotspots. Abbreviations and symbols; OH, heavy strand replication origin; PH and PL, H and L strand promoters; TAS, termination associated sequence; CSB I, II, and III, conserved sequence blocks I, II and III; bold dotted lines, the regions where most mutation were found in POLG1/PEO and Twinkle/PEO patients' muscle mtDNA; percentage in parentheses, the percentage of mutated fragments with mutations in the indicated region; small dotted line adjacent to OH, A to G conversionmutational hotspot region of patient 2 (Y955C POLG1); dashed line, position of the duplication hotspot in patient 1 (compound heterozygous R3P and A467T POLG1) (Wanrooij et al, 2004). Reprinted from original article (II), copyright (2004), by permission of Oxford University Press.*

In contrast to the POLG1 affected individuals the Twinkle/PEO patients accumulated most of their control region mutations (67%) in another region stretching from nucleotide 385 till 570 (Figure 5.5). In this region that contains the heavy and light strand promoters we found many different mutations that were generally present at low levels. The G564A mutation was found in all Twinkle patients and had high heteroplasmy levels in two patients but absent in controls. The sporadic patients carrying a single mtDNA deletion, showed a similar control region mutation distribution as the Twinkle/PEO patients.

Many of the control region mutations in PEO patients that we identified were located in conserved regions suggested to be replication control elements. We aimed to study if these control region mutations could predispose the DNA molecules to deletion formation. For this reason we investigated if control region mutation preferentially

accumulate in deleted molecules, we specifically amplified deleted mtDNA molecules with long range Pfu PCR. We compared the mutation levels of deleted molecules in six different ar/adPEO patients with the mutation level found earlier in the total population of mtDNA molecules (Figure 5.6) and found no significant difference in control region mutation level of deleted molecules.
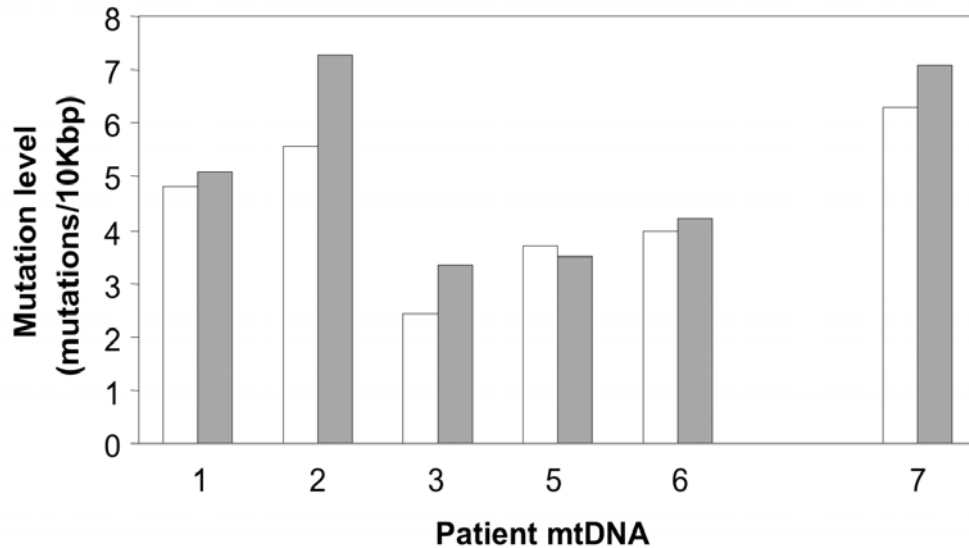


**Figure 5.6:** *Comparison of control region mutation levels between deleted and the total mtDNA population. The number of point mutations in the control region detected in deleted mtDNA molecules (gray bars) and in the total population of mtDNA molecules (white bars) of patients 1-3 and 5-7 (Wanrooij et al 2004). Reprinted from original article (II), copyright (2004), by permission of Oxford University Press.*

**5.7 Mapping of multiple mtDNA deletions in PEO patients**

Our mtDNA mutation analysis revealed that there is no fidelity decrease of POLG1 in PEO *in vivo*. This observation makes the slipped mispairing hypothesis of deletion formation in adPEO patients unlikely. To learn more about the formation mechanism of multiple mtDNA deletions in ad/arPEO patients we mapped the mtDNA deletion breakpoints. For this purpose we sequenced over the deletion breakpoints of cloned fragments obtained after specific amplification of deleted mtDNA molecules. We used several different primer sets (see material and methods) to avoid that certain deletions would be over-represented. It is likely to see an over-representation of the largest mtDNA deletions by using a single set of primers since those yield a shorter PCR fragment, which might be preferred in the ligation procedure. In both Twinkle and POLG1 PEO patients we identified a major deletion breakpoint at the 'termination associated sequence' (around bp 16070). Apart from this, the DNA sequence of the deletion breakpoints had some distinctive features; we found a high prevalence of homopolymeric runs and several micro-satellite-like sequences at the deletion boundaries. About 20% of the breakpoints were within or precisely flanked by a 4-nucleotide or more homopolymeric run. This was significantly higher than the predicted frequency (11%) if deletions would be randomly distributed . Moreover when we excluded the deletion hotspot around 16070 this

frequency was even higher (33%). We did not observe any difference in deletion boundaries when we compared deletions that occurred at different regions of the mitochondrial genome. A combination of deletion mapping and control region mutations analyses in the same mtDNA molecule revealed there was no obvious association between a particular control region mutation with a specific deletion. Finally, we did not identify any mutations in the boundary sequences of the deletion breakpoints.

**5.8 Establishing stable inducible cell lines expressing mtDNA replication factors**
In order to obtain more insight in the mechanisms by which mtDNA is maintained we established stable inducible HEK293 Flp-In™ T-REx™ cell-lines that express different variants of POLG1, Twinkle, POLG2 and mtSSB (Table 5.1).

| Variant | Size (kDa) |
| --- | --- |
| Twinkle (wt), K421A, G575D | 74 |
| Twinkle ΔAA346-376 | 71 |
| Twinky | 64 |
| Twinkle ΔAA70-343 | 41 |
| POLG1(wt), D198A, D890N, D1135A | 142 |
| POLG1 ΔCAG | 141 |
| POLG2 | 55 |
| MtSSB | 15 |

**Table 5.1:** *Established 293 Flp-In™ T-REx™ cell-lines.*

We previously demonstrated this inducible system is a useful tool to study different properties of another protein (TFAM) involved in mtDNA maintenance (Pohjoismäki et al, 2006). This cell culture system allowed us to study protein function and mtDNA dynamics *in vivo*. Immunoblotting showed clear expression of the exogenous protein in all cell-lines after 72 hours induction with 3 ng/ml doxycycline (DC) and even higher expression levels when 10 ng/ml DC was used. (Figure 5.7). Some protein variants showed reproducible lower expression levels (Twinky and POLG1 D1135A), which might be due to lower protein stability. We confirmed that all expressed proteins were mitochondrial localized by immunocytochemistry (ICC) (data not shown). Previous studies have shown that constitutive expression of an exonuclease deficient POLG1 (D198A) variant in human cells resulted in accumulation of mtDNA point mutations (Spelbrink et al, 2000). In order to validate our inducible expressing system we established a cell-line expressing this same variant and determined the mtDNA mutation levels after 60 days of induced expression. As expected, overexpression of this D198A

variant resulted in a 4 times higher mutation level in both regions that were investigated (cyt *b* and control region) (Figure 5.8).
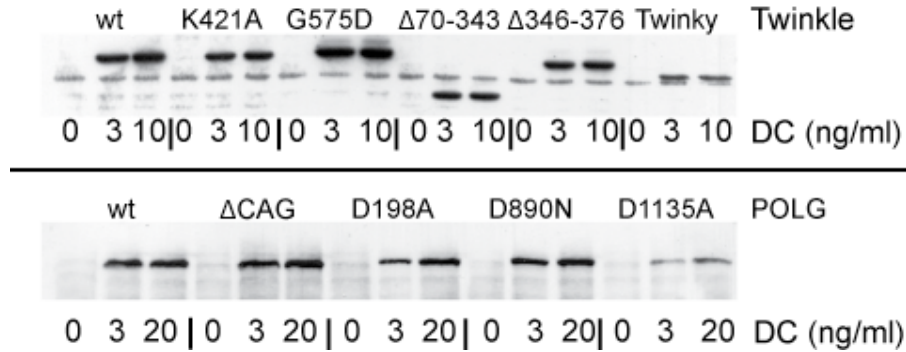


**Figure 5.7:** *Expression of mtDNA maintenance proteins in HEK293 Flp-In™ T-REx™ cells. Expression of all proteins used in this study was confirmed by Western blot analysis comparing no induction with 3ng/ml DC and full (10 or 20 ng/ml DC) induction for 3 days. Note that full induction with the POLG1 cell-lines was chosen as 20 ng/ml, only in order to make sure maximum expression was reached as expression of POLG1 was generally much weaker when compared with the expression of Twinkle variants (Wanrooij et al, 2007). Reprinted from original article (III), copyright (2007), by permission of Oxford University Press.*
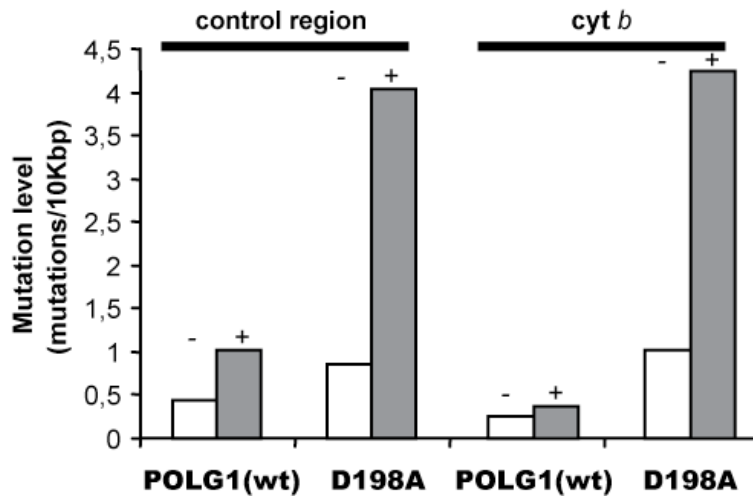


**Figure 5.8:** *POLG1 wild-type (wt) and POLG1 D198A expressing cells were grown for 2 months either without (-) or with (+) 3ng/ml DC. MtDNA mutation levels were determined for part of the cytochrome b gene (cyt b) and part of the non-coding control region using a PCR based approach (see Material and Methods). Results show that the D198A accumulated an approximately 5-10 fold increased level of mutations compared to non-induced D198A and induced POLG1 wt (Wanrooij et al, 2007). The mutation pattern was similar as reported earlier for this POLG1 variant (Spelbrink et al, 2000). Reprinted from original article (III), copyright (2007), by permission of Oxford University Press.*

Cells that were grown without DC showed nearly no expression; only on overexposed western blots some expression was detected (data not shown). However the observations that uninduced cells did not give any fluorescence signal above the background with ICC (data not shown) and no rise in mutation levels was detected in the non-induced D198A cell-line (Figure 5.8) showed that expression levels were very low.

## 5.9 MtDNA content after expression of mtDNA replication proteins

The established 293 Flp-In™ T-REx™ cell-lines were used to examine the influence of different protein variants on mtDNA content. Comparative Q-PCR of mtDNA and the single copy nuclear gene (*APP*) was used to measure mtDNA copy number.
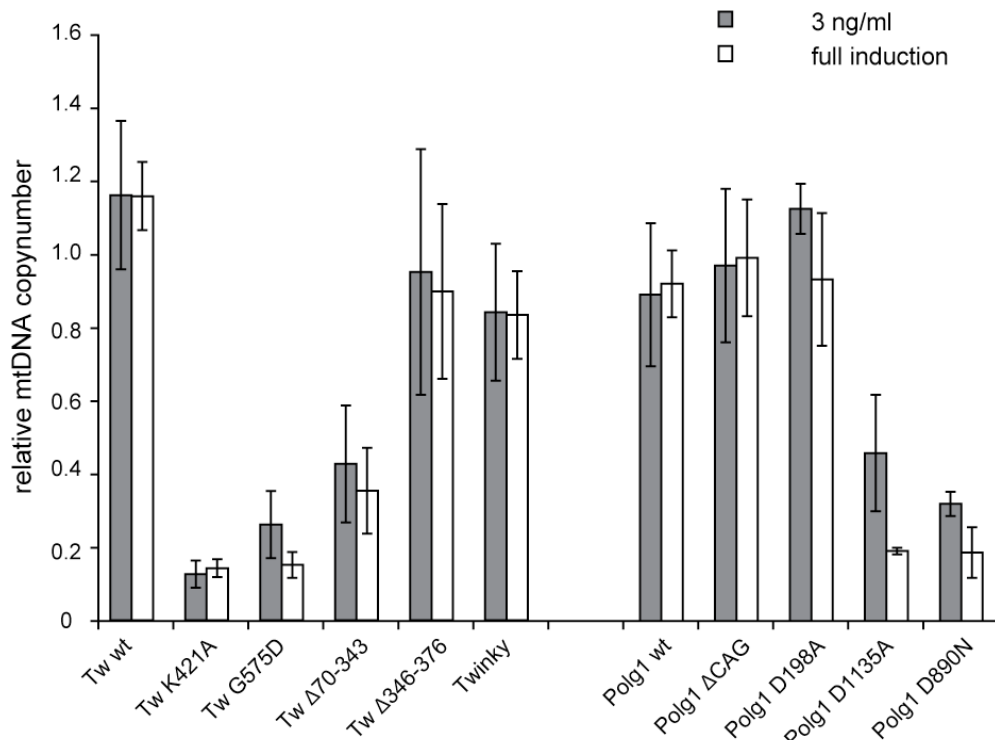


**Figure 5.9:** *POLG1 and Twinkle mutants can cause mitochondrial DNA copy-number depletion. MtDNA copy-number was determined using duplex Taqman PCR. Shown are the copy-numbers for each cell-line relative to their copy-number in non-induced cells following either 3 ng/ml or full induction (10ng/ml for Twinkle variants; 20 ng/ml for POLG1 variants). Note that although Twinkle K421A and G575D cell-lines had a reduced steady-state copy-number in non-induced cells, induction resulted in a further strong drop in copy-number. The same general trends as observed here were obtained by Southern analysis (not shown). Copy-number decline was highly significant in Twinkle K421A, G575D, and Δ70-343 lines as well as in the POLG1 D890N and D1135A mutant lines (Wanrooij et al, 2007). Reprinted from original article (III), copyright (2007), by permission of Oxford University Press.*

62

Expression induced by DC of wild-type POLG1, POLG2, Twinkle and mtSSB did not alter the content of mtDNA (Figure 5.9 and unpublished data). Similarly, some POLG1 (D198A and ΔCAG) and Twinkle (Twinky and Δlinker) variants did not change copy number. However, expression of POLG1 mutants D890N or D1135A and Twinkle mutants K421A, G575D and Δ70-343 lead to a strong depletion of mtDNA. Depletion caused by the expression of these mutants was dose dependent, lower expression levels show a slower decrease in mtDNA content. Even though the transgene expression was very low in the K421A-Twinkle and G575D-Twinkle cell-lines when grown without DC, we measured a 60% decrease in mtDNA level when cells were grown under this condition. This suggests that the K421A and G575D mutant Twinkle variants are strongly dominant. Both the Twinkle and POLG1 mtDNA depleting variants resulted after 3 days of full DC induction in a 70%-80% decrease in mtDNA levels compared to non-induced cells, suggesting no new complete mitochondrial DNA molecules were synthesized in this period.

**5.10 Twinkle mutations and their influence on nucleoid localization and helicase activity**

Twinkle is a DNA helicase with 5' to 3' directionality and distinct substrate requirements (Korhonen et al, 2003). Together with POLG1, POLG2 and mtSSB, Twinkle forms a minimal replisome showing that Twinkle is the mtDNA replicative helicase (Korhonen et al, 2004). In agreement with this our results showed that his affinity purified Twinkle has clear unwinding activity in an *in vitro* assay (Figure 5.10). In contrast, the various Twinkle variants (K421A, G575D, Twinky and Δlinker) did not possess any helicase activity or a significantly reduced activity (Δ70-343).
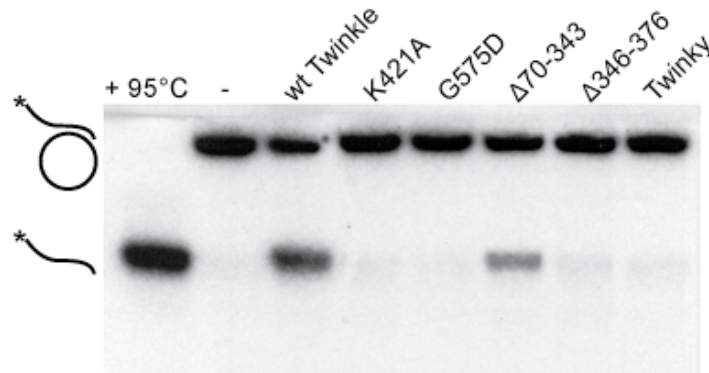


**Figure 5.10:** *Helicase activities of purified Twinkle protein fractions. An example of a helicase activity assay using the various Twinkle variants purified to near homogeneity. As indicated on the left hand side, the substrate consisted of M13 plasmid with an annealed 60 nt 5'end-labeled oligo with a long (40 nt) single-stranded 5' overhang. (-) indicates the purified substrate whereas the denatured substrate (95ºC) indicates the released labelled product, which is also released when a purified protein contains helicase activity (Wanrooij et al, 2007). Reprinted from original article (III), copyright (2007), by permission of Oxford University Press.*

Our initial characterization of the Twinkle protein revealed that Twinkle co-localizes with protein-DNA complexes referred to as mitochondrial nucleoids. With the use of the established 293 Flp-In™ T-REx™ cell-lines we studied intra-mitochondrial localization

of different Twinkle variants by ICC. As expected from the nucleoid localization from Twinkle, the wild-type protein showed a punctuated pattern that co-localized with an anti-DNA antibody. Twinky and Δlinker-Twinkle proteins showed a diffuse mitochondrial staining. Although the G575D variant showed to some extend a punctuate pattern, it was clearly more diffuse compared to the wild-type protein. The K421A mutant showed enlarged foci, suggesting abnormal shaped nucleoids in these cells. Anti-DNA antibodies showed that the mutant Twinkle variants K421A and G575D were still part of nucleoids, however their expression resulted in a reduced number of nucleoids.

## 5.11 Expression of different Twinkle variants and effect on mtDNA replication

We used two-dimensional neutral/neutral agarose gel electrophoresis (2DNAGE) followed by southern blotting to study replication intermediates (RI) in 293 Flp-In™ T-REx™ cells. Previous studies have shown this method is a valuable tool to visualize different types of RIs in cultured human cells (Yasukawa et al, 2005). In a recent publication 2DNAGE was used to show that the most dominant form of replication in cultured human cells initiates unidirectional in the non-coding region and RNA is deposited on the displaced H-strand, thus forming RITOLS (ribonucleotide incorporation throughout the lagging strand) (Yasukawa et al, 2006). Although we analysed different regions of the mitochondrial genome, most experiments were done on a 3,9 kb mtDNA fragment (mtDNA between bp 13636 and bp 1006) that includes the whole non-coding region. If not stated otherwise all the following results involve this particular region of the mitochondrial genome. (Figure 5.11) shows a schematic diagram that indicates some RIs that can be visualised by use of the 2DNAGE method. More explanation about the different types of RIs can be found in the legends of this figure. Two levels of transgene expression were used in our study, moderate (3 ng/ml DC) and full expression level (10-20 ng/ml DC).
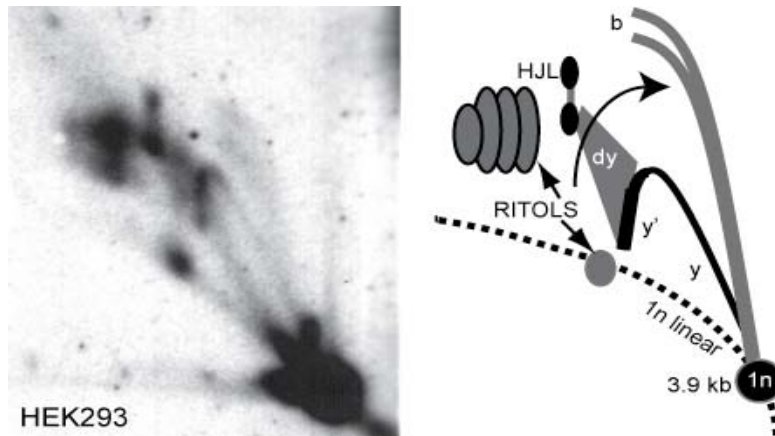


**Figure 5.11:** *Figures show the RIs of HEK293 cells and the interpretation based primarily on earlier 2DNAGE analysis of mtDNA RIs (Yasukawa et al, 2005). Abbreviations: 1n, 3.9 kb non-replicating HincII fragment. (b) bubble arcs. MtDNA bubble arcs are usually very sensitive to RNase H due to the presence of patches of RNA-DNA especially on the lagging strand; these therefore also fall in the category of RITOLS as do various other RIs. y and y' indicate ascending and descending parts of the y arc. (dy) indicates double-Y structures. These will eventually form resolution intermediates resembling Holliday junctions (HJL-Holliday junction like molecules) (Wanrooij et al 2007). Reprinted from original article (III), copyright (2007), by permission of Oxford University Press.*

Moderate expression levels of wild-type Twinkle did not influence RIs. Although also high expression (10 ng/ml) shows no noticeable affect on bubble- and y arc there was a clear decrease in resolution intermediated compared to non-induced 293 Flp-In™ T-REx™ -Twinkle cells (Figure 5.12A). Expression of the K421A and G575D Twinkle variants resulted in a strong increase in bubble- and y arc RIs (Figure 5.12B). Integration of this data with the earlier QPCR results suggested that severe replication stalling caused mtDNA depletion in these 293 Flp-In™ T-REx™ cells. Most RIs were found in the bubble arc, which are replication initiation intermediates, suggesting replication stalls at an early stage of replication. It was further apparent that this Twinkle induced replication stalling coincided with a strong decrease in RITOLS RIs. The limited effect of RNaseH and S1 nuclease on the RIs obtained by Twinkle stalling confirmed the RNA poor and double stranded nature of these RIs. Expression of mtSSB or Twinkle variants Twinky and Δlinker did not show any obvious change in RIs.
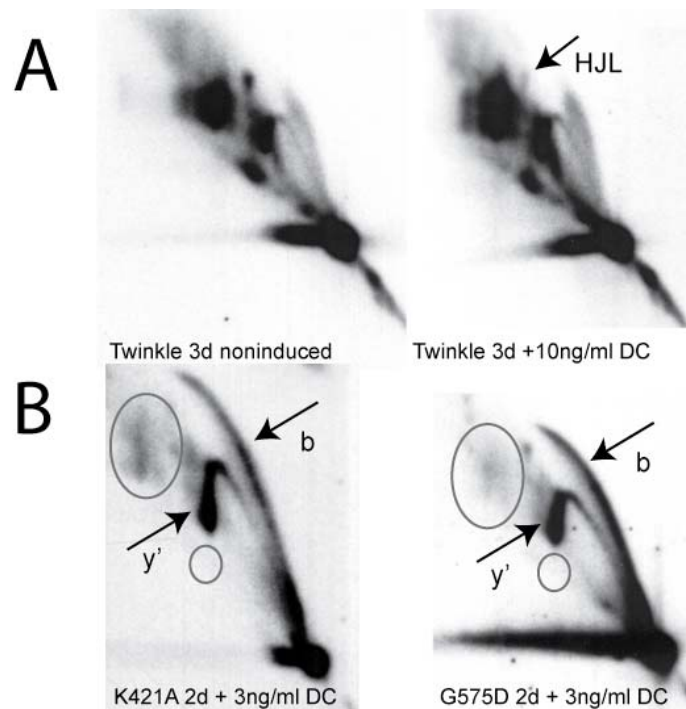


**Figure 5.12:** *Some Twinkle mutants cause replication stalling. 2DNAGE samples for all panels consisted of purified mtDNA digested with HincII and probed with a radiolabeled cytochrome b gene fragment (nt 14,846–15,357). The detected fragment includes the non-coding region of mtDNA also including the cytochrome b, ND6, part of the ND5 gene and intervening tRNA genes (nt 13,636–1,006). (**A**) A comparison of RIs of non-induced (left) and fully induced (right) cells expressing Twinkle wt. The only notable difference in this case is a reproducible reduction in one of the HJL RIs as indicated. (**B**) K421A and G575D show similar patterns of replication stalling, with increased bubble (b) and descending Y-arc (y') intensities, a sharpening and lengthening of the bubble arc and loss of RITOLS (ovals) (Wanrooij et al, 2007). Reprinted from original article (III), copyright (2007), by permission of Oxford University Press.*

## 5.12 Expression of different POLG1 variants and effect on mtDNA replication intermediates

Moderate or full expression of wild type POLG1, ΔCAG or POLG2 does not affect RIs. In contrast even a moderate expression (3ng/ml DC) of the D890N and D1135A variants increased bubble- and y arc RIs significantly (Figure 5.13), probably due to replication stalling. Unlike the Twinkle stalling variants, the POLG1 mutant induced stalling showed abundant RITOLS RIs. In the cell-line expressing the exonuclease deficient PolG1 mutant (D198A) we observed a mild phenotype. Full induction of D198A (20ng/ml) showed enhanced bubble and y-arc RIs, however this increase was clearly less than observed with D890N/D1135A at low induction.
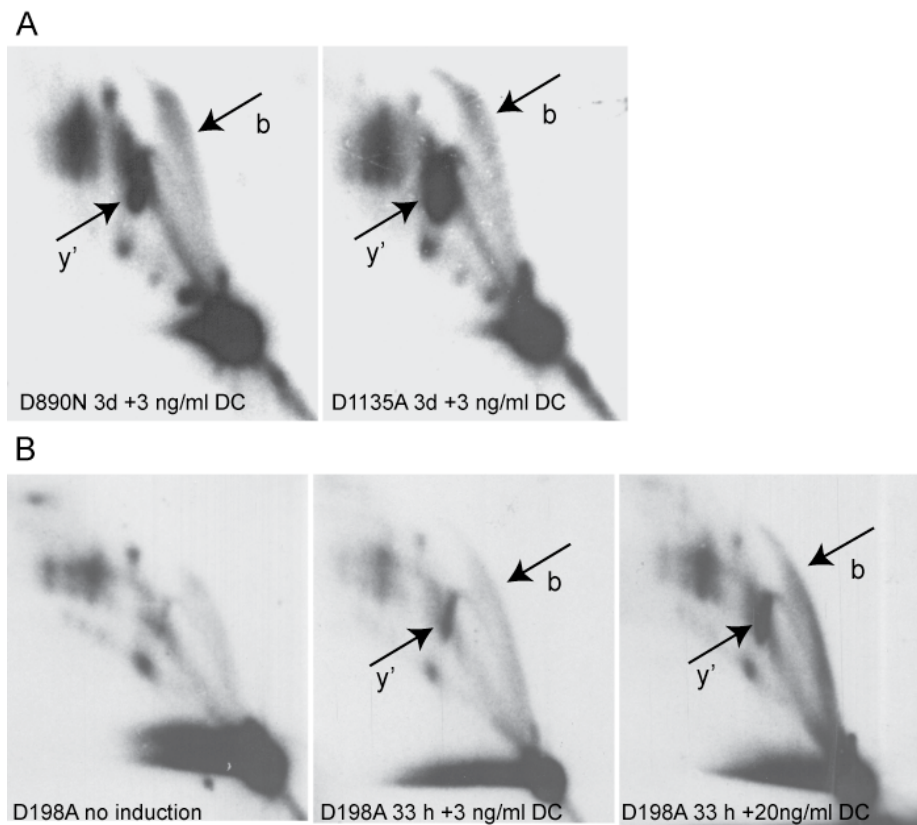


**Figure 5.13:** *Replication stalling by POLG1 mutants. (**A**) Induction of POLG1 mutants D890N and D1135A results in a similar pattern of replication stalling with an increased intensity both of bubble (b) and descending y (y') RIs. RITOLS are not affected in sharp contrast with Twinkle-mutant induced stalling (Figure 5.12). (**B**) Fully induced expression of POLG1 D198A results in a modest change in RIs similar to but less severe than low level induction of D890N and D1135A. This did however not result in copy-number depletion (Figure 5.9). All panels in **B** are cropped images from a single autoradiograph (Wanrooij et al, 2007). Reprinted from original article (III), copyright (2007), by permission of Oxford University Press.*

## 5.13 Different POLG1 and Twinkle induced replication stalling phenotypes

The low abundance of RITOLS RIs in the stalling Twinkle cell-lines and the clear presence of these RIs after mutant PolG1 induced stalling, suggested that POLG1 is involved in their maturation. To test this hypothesis, that mtDNA depletion in the stalling POLG1 cell-lines is caused by a combination of replication stalling and loss of lagging strand maturation, we inhibited POLG1 in the Twinkle K421A and G575D cell-lines by dideoxycytidine (ddC) treatment. In the absence of DC the G575D cell-line already expressed its transgene at very low levels. Consequently, cells grown without DC already showed a moderate stalling phenotype, including loss of RITOLS (Figure 5.14). RITOLS re-emerged after short POLG1 inhibition by 4 hours ddC (200 µM) treatment. Interestingly, treatment with ddC after strong stalling by full induction of G575D (10ng/ml) failed to cause re-appearance of RITOLS RIs (data not shown). This suggested that after full G575D induction replication was so severely impaired that virtually all replication events stall in the beginning of the replication process and do not reach the later stages of replication that involves these RITOLS intermediates.
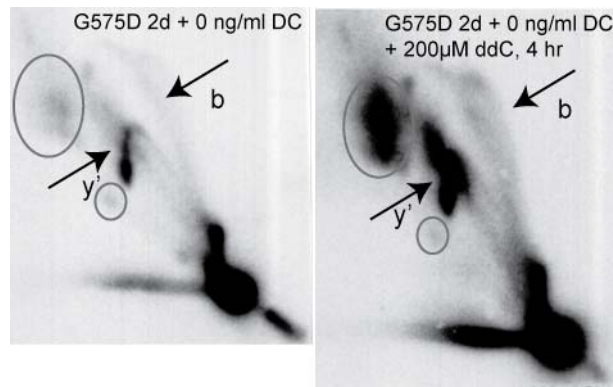


**Figure 5.14:** Inhibition of POLG1 by ddC results in the reappearance of RITOLS replication intermediates following stalling induced by Twinkle G575D. Non-induced Twinkle G575D cells, already showing signs of replication stalling due to leaky expression, were treated for four hrs with 200 $\mu$M ddC prior to mtDNA isolation. Note that exposures of these two samples are identical, as both samples were run on a single gel and blot. Markings are similar as in figure 5.12 (Wanrooij et al, 2007). Reprinted from the original article (III), copyright (2007), by permission of Oxford University Press.

## 6. Discussion

### 6.1 Twinkle, a mitochondrial nucleoid-localized replicative mtDNA helicase with striking sequence similarity to the bacteriophage T7 gene 4 protein (gp4)

In 2000 the mammalian mtDNA replication machinery remained poorly defined. Although the mammalian mitochondrial DNA polymerase gamma and the mitochondrial SSB protein were cloned and characterized (Ropp et al, 1996; Tiranti et al, 1993), other factors remained unknown. In particular a mtDNA replicative helicase unwinding dsDNA in front of polymerase gamma was on the list of required but unidentified proteins. Further a DNA primase, an essential enzyme in most systems of DNA replication that synthesizes an RNA template for the DNA polymerase to initiate replication, was not identified within mammalian mitochondria.

The gene coding for the human T7 gp4-like protein with intramitochondrial nucleoid localization or Twinkle was localized in a region on chromosome 10q24 previously shown to be linked to adPEO (Li et al, 1999). As its name implies the 77-KD protein shows primary sequence similarity to the bacteriophage T7 gp4 primase/helicase. The strongest similarity is found in the C-terminal helicase domain. Similar to the phage T7 protein it contains conserved sequence motifs in this region, which are typical for DNA helicases of the RecA/DnaB superfamily. This suggested that Twinkle is a DNA helicase and indeed initial experiments showed that overexpressing Twinkle in human cells resulted in an increase of mtDNA helicase activity, however this was a very modest increase. In later experiments purified recombinant Twinkle was confirmed to contain DNA helicase activity and this activity could be specifically stimulated by the mitochondrial single-stranded DNA-binding protein (mtSSB) (Korhonen et al, 2003). The demonstration that Twinkle and the pol γ holoenzyme together are able to processively synthesize single-stranded DNA from a double-stranded DNA template suggests that Twinkle is indeed the helicase at the mtDNA replication fork (Korhonen et al, 2004). The protein was shown to catalyzes the ATP-dependent unwinding of a dsDNA (20bp) in a 5' to 3' direction. Further Twinkle needs specific substrates with a single-stranded 5'DNA loading site and a short 3'-tail to initiate unwinding. This substrate is similar to the DNA at the replication fork and to the substrate used by the T7 gp4 helicase/primase (Korhonen et al, 2004; Debyser et al, 1994). During mtDNA replication Twinkle is positioned on the lagging-strand template and separates dsDNA to make the ssDNA needed for mtDNA replication by POLG1 on the leading-strand template. Further experiments using the mitochondrial *in vitro* replication system showed that Twinkle dependent stimulation of POLG1 is specific. Twinkle is not able to stimulate T7 DNA polymerase in rolling circle replication (Korhonen et al, 2004)

The N-terminus containing the primase domain in the phage T7 protein shows less sequence similarity with human Twinkle. Based on much weaker overall sequence similarity in this region and on the absence of a zinc-binding domain crucial for primase activity it was initially suggested that Twinkle does not include primase activity. However, the recent finding that one Twinkle homologue (Pfprex) in the malaria parasite *Plasmodium faciparum* contains primase activity questions this (Seow et al, 2005). A

more recent sequence alignment study of many Twinkle homologues in eukaryotes showed conserved primase motifs including the zinc finger motif in almost all Twinkle sequences, however not in Metazoa. Thus it is most likely that Twinkle functions as the primase for mtDNA replication in most eukaryotes, with the exception of Metazoa (Shutt and Gray, 2006). In accordance with this observation human Twinkle seems to lack *in vitro* primase activity (Personal communication, Farge and Falkenberg).

The N-terminus and helicase domain in Twinkle are separated by a conserved sequence often referred to as the linker region. In the phage T7 gp4 protein, that needs to form hexamers in order to be active as a helicase, this linker region is required for the actual multimerization (Guo et al, 1999). We have proposed that the linker region in Twinkle would fulfil a similar role in hexamerization. Indeed, native-PAGE experiments with low SDS percentage have shown that human Twinkle is capable of forming multimers. Nevertheless, in our study a variant of Twinkle that lacks the linker-region showed an abundant multimeric form that could correspond with a hexa- and/or heptamer after glutaraldehyde crosslinking.

Identification of the mitochondrial DNA and RNA polymerase showed that they are closely related to their counterparts from T-odd bacteriophages (Ropp et al, 1996; Tiranti et al, 1997). Twinkle's similarity to the T7 primase/helicase thus confirmed that the core of the mtDNA replication machinery is closely related to the T7 phage replisome. This suggests that much can be learned about molecular mechanisms and mode of mtDNA replication, from the intensively studied T-odd phage replication systems (Richardson, 1983). Moreover this raises a question about a role for a T-odd-like phage in the evolution of mitochondria (Shutt and Gray, 2006).

Another interesting feature of Twinkle is that it co-localizes with mitochondrial nucleoids. Although you would expect this localization from a protein implicated in mtDNA metabolism, other proteins involved in this process like POLG1 and POLG2 are more uniformly distributed (Garrido et al, 2003). This suggests that some proteins involved in mtDNA replication are only recruited to the nucleoids when needed for replication. Twinkle however seems to have a more permanent position within the nucleoids, a characteristic that is shared with TFAM and mtSSB. Further it has been proposed that Twinkle has a role in mtDNA copy number control. This is based on the observations that transgenic mice that overexpress Twinkle have increased mtDNA content, whereas reduction of Twinkle expression in human cells by RNAi results in mtDNA depletion (Tyynismaa et al, 2004). This increase of mtDNA content caused by overexpression of Twinkle in mouse could be a result of increased mtDNA replication, or alternatively Twinkle as a permanent constituent of the nucleoids stabilizes mtDNA molecules. However in contrast to the findings in mice, we found that overexpression of human Twinkle in human cell-lines did not increase the level of mtDNA. The mechanism behind this difference needs to be addressed in future experiments.

The splice variant of Twinkle, a 66kDA protein named Twinky that is missing a portion of the C-terminus, does not localize to nucleoids suggesting that this region might be in part responsible for the nucleoid localization. A function for Twinky has not been

reported yet, it could be an inactive form of Twinkle, making it possible to regulate by splicing the Twinkle protein expression levels. This splice variant seems to be conserved in mouse, however mRNA coding for Twinky has not yet been detected in mouse tissue (Tyynismaa et al, 2004). Nevertheless, minisequencing data showed that ≈20% of the mRNAs transcribed from the human PEO1 gene codes for Twinky hinting towards a more significant role for this splice variant (Nikali et al, 2005). Overexpression of Twinky in HEK 239 Flp-In™T-Rex™ cells did not affect RIs. This was expected since our study showed that Twinky does not form multimers and thus by analogy to the T7 helicase/primase is not only expected not to contain helicase activity (Egelman et al, 1995), but also not to interfere with mtDNA replication. In agreement with this we showed that purified Twinky did not posses helicase activity.

Overexpression of wild-type Twinkle in HEK 239 Flp-In™T-Rex™ cells did not show apparent changes in y- and bubble arc RIs, nevertheless a strong decrease in an alleged resolution RI was observed. This shows that Twinkle might be involved in resolving these termination junctions. In heart tissue four way (Holliday) junction types of RIs are prominent and have been proposed to represent recombination intermediates (Kajander et al, 2001). A role for Twinkle in resolving termination junctions and in recombination is not unlikely since the helicase domain shows the same conserved organization as RecA. The Rec A gene was the first gene discovered to mediate homologous recombination (Clark and Margulies, 1965). Nonetheless experimental evidence is needed to confirm such a recombination role for Twinkle.

## 6.2 Mutations in Twinkle underlie adPEO, a disorder associated with multiple mtDNA deletions

AdPEO is a mitochondrial disease associated with multiple mtDNA deletions, which generally manifests itself as exercise intolerance and muscle weakness, most harshly in the ocular muscles (Suomalainen et al, 1995). Previous studies have reported at least 3 distinct loci for this disease (Suomalainen et al, 1995; Kaukonen et al, 1996; Kaukonen et al, 1999). The 4q34-linked adPEO disease gene has been identified as the heart and skeletal muscle specific isoform of the adenine nucleotide translocator (ANT1) (Kaukonen et al, 2000). Twinkle is localized in a region previously identified as being linked to adPEO on chromosome 10 (Suomalainen et al, 1995; Li et al, 1999). Our extensive sequence analysis of the PEO1 gene in adPEO families revealed 11 different mutations. The absence of these gene variants in ethnically matched controls confirmed they are not common polymorphisms. The co-segregation of the Twinkle mutations within all adPEO families and the phylogentically conserved regions in which they occurred provides strong evidence that the mutations are pathogenic significant in this disease. However, the V368I change in Twinkle we initially reported to be pathogenic was later shown to be a nonpathogenic polymorphism (Arenas et al 2003). Based on the T7 gp4 crystal structure of the helicase domain none of the Twinkle mutations associated with adPEO are suggested to be critical for DNA helicase activity (Sawaya et al, 1999). Two mutations, W474C and A475P were identified in less essential residues within the helicase domain and a third mutation F485L was identified more recently (Kiechl et al, 2004)(Figure 6.1). The amino acids corresponding to W474 and A475 in human Twinkle

70

are located in the helical loop that contacts the linker region of a neighbouring subunit in the multameric complex (Kiechl et al, 2004). Most of the patient mutations identified in adPEO patients are located within this linker region and might be predicted to interfere with the multimerization of the protein. However, initial experiments with 2 adPEO-Twinkle variants containing a 39 bp linker region duplication (aa 352-364) or W474C did not show any effect on Twinkle in vitro mulitmerization. This suggested that at least these mutations might have other consequences on the function of Twinkle. Perhaps the multimerization defect is subtle and does not show in these *in vitro* experiments, but does interfere with the helicase processivity *in vivo*. In order to investigate the function of the linker-region we established a cell-line expressing a variant that lacks the whole linker-region (Δ346-376). Induced expression did not lead to change in RIs analyzed by 2DNAGE or mtDNA copy number. This could be explained by our observation that this protein did not show a distinct nucleoid localization and is inactive in helicase assays. This mutant is thus unlikely to form hetero-multimers with the endogenous enzyme.

Our study and subsequently others identified several N-terminal Twinkle mutations (Lewis et al, 2002)(Figure 6.1). The observation that Twinkle lacks primase activity (personal communication, Farge and Falkenberg) suggests that the N-terminal Twinkle adPEO causing mutations must therefore affect other essential functions of the protein. One Twinkle mutation Y508C has been shown to cause a severe autosomal recessive disorder, infantile onset spinocerebellar ataxia (IOSCA) (Nikali et al, 2005). In several Finnish patients there was no indication of mitochondrial dysfunction or mtDNA instability in muscle tissue. The pure central nervous system phenotype in IOSCA suggests there might be another cellular role for Twinkle and/or Twinky. This idea is supported by our unpublished findings that overexpression of the Y508C mutation in human cell culture does not change mtDNA copy number or replication intermediates observed by 2DNAGE (Goffart et al, unpublished data). The mutation might however influence mtDNA maintenance in a cell-type specific manner for example by interaction with a cell-type specific unknown partner.
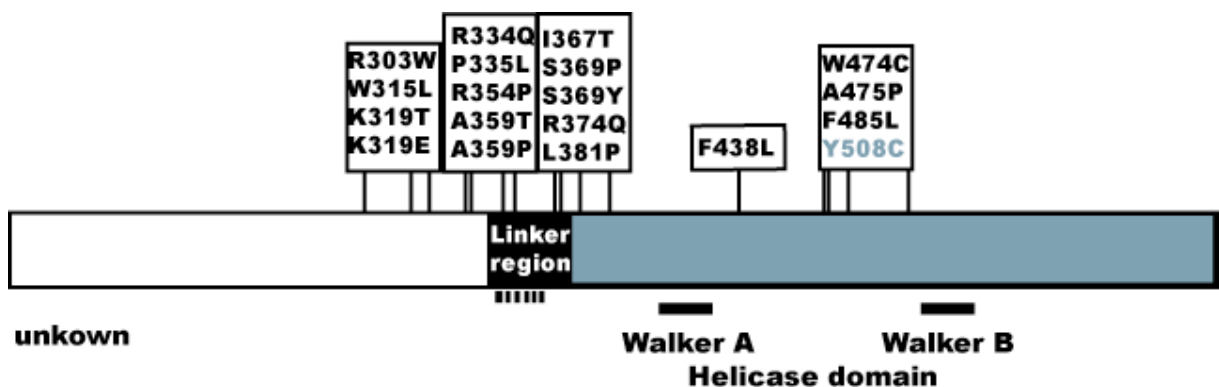


**Figure 6.1:** *Schematic presentation of the human PEO1 gene and visual overview of all the mutations that are mapped by us and others following sequence analysis of different adPEO family (Spelbrink et al, 2001; Lewis et al, 2002; Agostino et al, 2003; Kiechl et al, 2004; Hudson et al, 2005). The Y508C mutation has been associated infantile onset spinocerebellar ataxia (Nikali et al, 2005). The dotted line indicates the 13AA duplication in a Finnish adPEO family.*

71

Recently, a study showed that expression of mouse Twinkle with a corresponding PEO1 patient mutation (duplication aa 352-364) resulted in accumulation of multiple mtDNA deletions and similar histological features as seen in the PEO patients. Surprisingly the physical performance of this 'deletor' mouse was not compromised, unlike in PEO patients (Tyynismaa et al, 2005). The mild affect on the mouse might be useful however for evaluation of therapeutic strategies for mitochondrial myopathies. In the deletor mice, the first histological signs of mitochondrial myopathy, COX negative muscle fibers, were identified at the age of 12 months. When you take into account the expected lifespan this mimics the relative age of onset in human disease. Similar to adPEO patients the mice accumulate mtDNA deletions in muscle and brain, however the amount of deleted mtDNA in the mice was quite low. Interestingly the 3'deletion breakpoints were found to overlap with the human base pair 16070 which we and others identified as a breakpoint hotspot in PEO-Twinkle patients (see below). This suggests that both in the deletor mouse and adPEO patients the mtDNA deletions accumulate by a similar mechanism.

## 6.3 Does the mtDNA control region in ad/arPEO patient muscle show evidence of premature ageing?

At the same time as our report of Twinkle heterozygous mutations in adPEO families another group showed that mutations in the catalytic subunit of DNA polymerase γ can lead to adPEO (Van Goethem et al, 2001). *In vitro* studies on the most common POLG1 mutation (Y955C) associated with adPEO showed a 45-fold reduced binding affinity for the incoming nucleoside triphosphate (Ponamarev et al, 2002). This suggested a higher misinsertion rate during mtDNA replication for this adPEO-POLG1 variant, but in the same study a functional exonuclease domain was shown to counteract most of these misinsertions. Based on these observations the authors nevertheless presented a model in which an enhanced base substitution rate, such as induced by the Y955C mutation, could provoke a slippage event between direct repeats. A missinsertion could transform an imperfect repeat into a perfect repeat; this could promote slipped mispairing resulting in the formation of a mtDNA deletion. In order to assess the *in vivo* replication fidelity in PEO patients we measured mtDNA mutation levels in affected tissues from PEO patients. From a previous cell culture and mtDNA mutator mouse model we know that reduced POLG1 fidelity results in misinsertion mutations randomly distributed throughout the mitochondrial genome (Spelbrink et al, 2000, Trifunovic et al, 2004). Our data showed that PEO patients with a defect in POLG1 or Twinkle indeed have a higher mutation level in muscle mtDNA, but that this mutation level increase was specific to the control region and was not detected in coding regions like cyt *b*. These findings showed that there is no replication infidelity in ad/arPEO patients making the strand-slippage model for deletion formation unlikely. Furthermore, contrary to our findings in muscle mtDNA of PEO patients, mtDNA mutation levels in the mutator mice are significant higher in coding regions, compared to the control region (Rovio, 2006). Moreover control region mutation levels increase during normal human aging whereas no such increase has been reported for the coding regions (Michikawa et al, 1999). A recent re-analysis of mtDNA mutation levels in heterozygous mutator mice showed that mice can sustain a 500-fold increased mtDNA mutation burden compared to normal mice without causing premature

aging (Vermulst et al, 2007). Unfortunately, in this study the non-coding region was not included in the analysis. The control region contains many conserved sequences elements that regulate mitochondrial transcription and replication; mutations in these elements might disrupt mitochondrial transcription and or replication. This interference with replication or transcription might be much more deleterious to functional oxidative phosphorylation than low heteroplasmy level coding-region point-mutations. If this is true adPEO might be seen as a tissue specific (muscle and brain) pre-mature aging syndrome. However, more experimental work is needed to assess the importance of control region mutation in aging, muscle mtDNA of PEO patients.

One possibility that remained was that POLG1 and Twinkle mutations result in a replication infidelity that is specific for the control region. However, the non-random distribution of control region mutations refutes this possibility. Several of the observed mtDNA mutations like A189G and T408A have previously been described to accumulate during aging (Michikawa et al, 1999; Wang et al, 2001). Interestingly in agreement with these studies we find a very mild increase of mtDNA control-region mutation level in skeletal muscle from healthy individuals. Skeletal muscle mtDNA from PEO patient showed a much stronger age dependent increase of control-region mutation levels, suggesting that those mutations accumulate by a mechanism similar to that seen in normal aging. The majority of the mutations identified in PEO skeletal muscle mtDNA were however not previously reported to accumulate during normal aging. Although we see that the control region mutations progressively accumulate during aging this different mutation accumulation suggests they arise by a different mechanism. Moreover in contrast to earlier findings we find the A189A and T408A mutations at very low heteroplasmy levels in muscle mtDNA even in the oldest individuals (Wang et al, 2001). However in our study only 4 individuals were older than 60, significant less then than the 26 in the study perform by Wang et al, suggesting the difference could be a sampling effect.

We also observed enhanced control region mutation levels in 3 out of 5 patients with a single sporadic mtDNA deletion. This implies that the control region mutations are probably the indirect result of the presence of an mtDNA deletion or multiple deletions and are unlikely a primary consequence of the POLG1 or Twinkle gene defects. Deletion PCR analysis indicated that mtDNA deletions are already evident in younger Twinkle/PEO patients, while no increase in control region point-mutations were detected. This again supports the idea that the deletion formation is an earlier event.

It remains formally possible that PEO patients do have enhanced mutations accumulation randomly throughout the mitochondrial genome; a higher level of repair in the coding regions compared to the control region might explain why we only detect an increase of mutation levels in the non-coding region. Indeed in one study in which mitochondrial DNA was exposed to electrophile and oxidant damage, the D-loop was found to be more susceptible to mutations (Mambo et al, 2003). Although it was suggested this could be due to the proximity of the D-loop to the mitochondrial inner membrane that may hinder access of repair proteins, an enhanced vulnerability to damage or inefficient repair within this region could not be distinguished. From our data it seems more likely that the control region mutations accumulate as a result of mtDNA deletions i.e. as a result of an

enhanced vulnerability to DNA damage. This is supported by our ability to detect enhanced control region mutation levels in some of the sporadic single deletion patients. MtDNA deletions often result in a combined OXPHOS deficiency i.e. affecting several of the OXPHOS enzymes. OXPHOS deficiency can lead to enhanced ROS production (Esposito et al, 1999). The unique triple-stranded D-loop structure might predispose this region to more excessive DNA damage by ROS (Figure 6.2).

Our finding that control region mutations accumulate during aging without any detectable mutation accumulation in the coding region has direct bearing on the question of the mitochondrial theory of aging as proposed by Harman more then 50 years ago (Harman, 1956). Many today still support this theory, however dissenting opinions about it also exist (Sanz et al, 2006; Magalhaes, 2005; Vermulst et al, 2007). According to this theory, ROS are endogenously produced by mitochondria as a byproduct of the OXPHOS system under normal physiological conditions. They are produced continuously during life and have deleterious effects on DNA. Somatic mutation would especially accumulate on mtDNA during aging because of its location close to the respiratory chain, leading to a progressively reduced efficiency of the OXPHOS system. To complete the circle of this theory this would result in even more ROS production, which would then again produce more mtDNA damage. In disagreement with this theory we could not detect an increase of coding region mtDNA mutations during aging in the muscle mtDNA samples of controls and adPEO patients. Moreover from our data it seems that if mtDNA alterations can lead to increased ROS production this mainly affects the control region, which contains the triple stranded D-loop structure. Indeed ssDNA has been found to be more susceptible to ROS than dsDNA (Frederico et al, 1990). Perhaps control region mutations are the actual mutations that contribute to limited natural life span in human and mice.

Most control region mutations (67%) in the muscle mtDNA of Twinkle/PEO patients were identified between bp 400-560, an area containing the L and major H-strand promoter for transcription initiation, whereas POLG1/PEO patients accumulated far more mutations (81%) in a region between bp 150-250, that includes $O_H$ and conserved sequence block I (CSBI). Our mtDNA deletion mapping indicated that both patient groups accumulated similar mtDNA deletions suggesting that the differences in control region mutation distribution are caused by the specific gene defect. Mutant Twinkle or POLG1 may affect the susceptibility of these mtDNA sequences to oxidative damage or alternatively provide a replicative advantage to the mutant molecules.
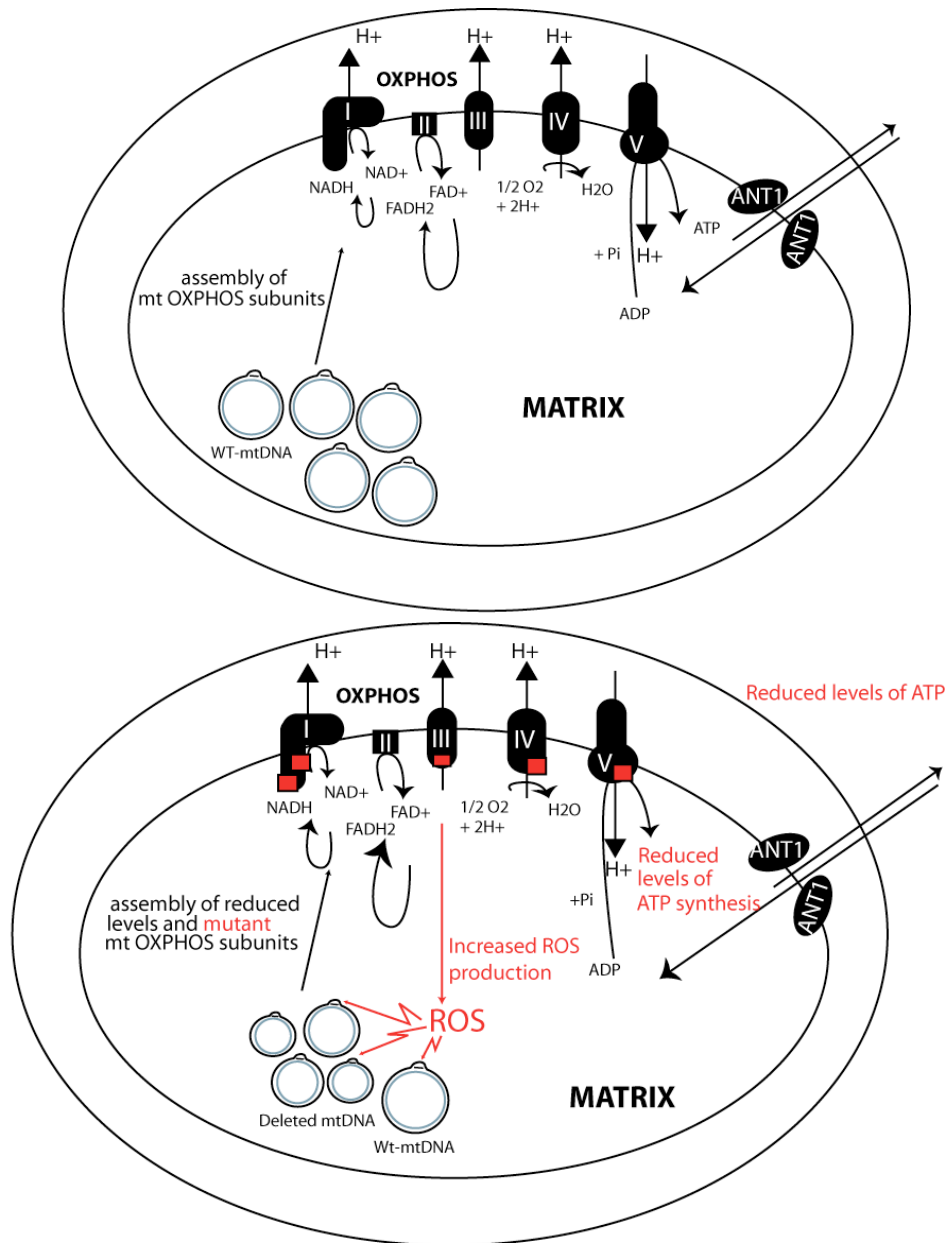
**Figure 6.2:** *Schematic presentation of a mitochondrion, with mtDNA molecules and OXPHOS system leading to ATP production. A) Situation in a healthy individual: wild-type mtDNA encodes for several OXPHOS subunits which are assembled into complex I, III, IV and V after transcription and translation. The functional OXPHOS system produces ATP (for more detail see literature review section 2.2); most of the ATP is exported by the adenine nucleotide transporter (ANT) and used for various cellular functions. B) Situation in post-mitotic tissue of adPEO patients: deleted mtDNA molecules accumulate during aging, as a consequence reduced levels and mutant OXPHOS subunits are assembled into complex I, III, IV and V resulting in a deficient OXPHOS system producing low levels of ATP and the disease symptoms. Another consequence of improper OXPHOS complexes could be an increase of reactive oxygen species (ROS) as a by-product of the oxygen consumption. The increased ROS might damage mtDNA, in particular the control region because of it triple strand structure and its close proximity to the mitochondrial membrane.*

## 6.4 Twinkle and POLG1 PEO mutations result in the accumulation of mtDNA deletions at difficult sites for the mtDNA replication machinery

The mutation analysis that we performed of skeletal muscle mtDNA samples obtained from PEO patients, excluded a decrease in replication fidelity. This is strong evidence against the strand-slippage mechanism of multiple deletion formation. Also our extensive deletion break analysis supports a different mechanism. We did not detect long direct repeats surrounding the deletion breakpoints. Secondly we did not find misinsertion mutations in the vicinity of the deletion breakpoint as would be predicted by the strand-slippage mechanism. Most direct repeats that were identified were short, often imperfect, and hardly ever flanked the deletions precisely. This is similar to the findings of earlier less extensive mtDNA deletion mapping studies (e.g. Zeviani et al, 1989). As general rule we found the deletions breakpoints occurring at sites that are difficult for the replication machinery. First, we frequently detected homopolymeric runs, which are known to be difficult sequences for the replication machinery to replicate (Bebenk et al, 1989; Kunkel et al, 1994; Tran et al, 1997). During the replication of a repetitive sequence the concentration of the inserted dNTP might be locally very low. A mutation in one of the key enzymes of the replication machinery might make it very sensitive to single nucleotide depletion. It is plausible that such a situation in which for example a deficient polymerase is not capable of incorporating the subsequent nucleotide might result in stalling or pausing of the replication fork. This might also explain the adenine nucleotide translocator (ANT1) gene mutations found in adPEO patients since it is directly involved in nucleotide metabolism (Kaukonen et al, 2000). A frequent stalling or pausing mechanism is supported by the observation that overexpression of the Y955C POLG1 variant in cell culture results in the accumulation of replication intermediates which suggests stalling resulting in mtDNA depletion (unpublished data). A recent study showed that Y955C POLG1 targeted transgenically to the mouse heart also leads to mtDNA depletion in heart tissue (Lewis et al 2007). In adPEO the Y995C heterozygous mutation results in accumulation of multiple mtDNA deletions in post-mitotic tissue without a clear reduction of mtDNA copy number. Expression of POLG1 Y955C in HEK293 Flp-In™ T-Rex™ cells or in the transgenic mice leads nonetheless to strong mtDNA depletion. In the transgenic mice however Y995C POLG1 was 10 to 20 fold the endogenous expression level, while in adPEO patients you would expect from a heterozygous mutant a comparable expression of wild-type and mutant POLG1. The cell-culture system we used consists of proliferating HEK293 cells that need to replicate mtDNA continuously in order to maintain constant mtDNA levels in daughter cells. This is quite different from post-mitotic tissues, which are mostly non-dividing cells and thus do not require a high rate of mtDNA replication to maintain constant levels.

Similarly, nucleotide pool involvement in frequent replication stalling/pausing might also explain the tissue specific accumulation of mtDNA deleted molecules seen in adPEO patients. Post-mitotic tissues consist mostly of differentiated non-dividing cells that consequently do not replicate nDNA. These cells do not go through S-phase and outside the S-phase *de novo* dNTPs synthesis in the cytosol is strongly downregulated (Eriksson et al, 1984). Under such conditions mtDNA synthesis relies mostly on the salvage synthesis of dNTPs (Bogenhagen et al, 1977, Song et al, 2003). Although mitochondria

produce their own dNTPs, the nucleotide concentrations might get critically low when little cytosolic dNTPs can be imported. It is plausible that a defective replication machinery resulting from a POLG1, Twinkle or POLG2 mutation is not able to deal with such conditions leading to uncontrolled and frequent replication pausing at for example homopolymeric runs. In support the above hypothesis we also found several mtDNA deletions that were flanked with microsatellite types of repeats, which are known sites for replication stalling and genome instability (Gordenin and Resnick, 1998). Finally others (Zeviani et al, 1989) and we found a frequent replication breakpoint at the replication pause site 16070. This observation is in agreement with a replication stalling model of mtDNA deletion formation, replication might slow down or stall at this pause site and may be a favoured substrate for error-prone recombination (Bowmaker et al, 2003).

In yeast and bacteria it has been well documented that replication forks frequently stall because of mutations in replication proteins especially in sequence repeats or at natural replication termination pause sites (Samadashwily et al, 1997; McGlynn et al, 2001; Cha and Kleckner, 2002). A common transitional state in restoring replication forks after stalling are double strand breaks (DSBs), which are either repaired by recombination or non-homologues end-joining (NHEJ). Enhancement of these repair pathways might produce genome instability (Michel et al, 1997; Weitao et al, 2003). Integration of this data with our deletion mapping makes it plausible that multiple mtDNA deletions in adPEO patients are the result of frequent replication stalling caused by a mutation in one of the replication proteins (Twinkle, POLG1, POLG2), followed by DSBs formation and error-prone repair by recombination or non-homologues end joining (NHEJ) (Figure 6.3). Recombination and double strand brake rejoining activity has indeed been reported in mammalian mitochondrial extracts (Thyagarajan et al, 1996; Lakashmipathy and Campbell, 1999). In mice expression of a mitochondrially targeted restriction endonuclease (*Pst*I) resulted in mtDNA DSBs formation, since mtDNA harbors two *Pst*I specific restriction sites (Srivastava and Moraes, 2005). Interestingly the mtDNA DSBs in these mice resulted in the formation of multiple mtDNA deletions, with molecular characteristics that are especially similar to the multiple deletions that we found in muscle of adPEO patients. In my opinion this strongly supports our proposed hypothesis that DSBs lead to the formation of multiple mtDNA deletions in the adPEO disorder.

Twinkle and POLG1 defects seemed to result in very similar mtDNA deletions, thus suggesting that for both gene defects the pathogenetic pathways resulting in deletion formation are nevertheless closely related.
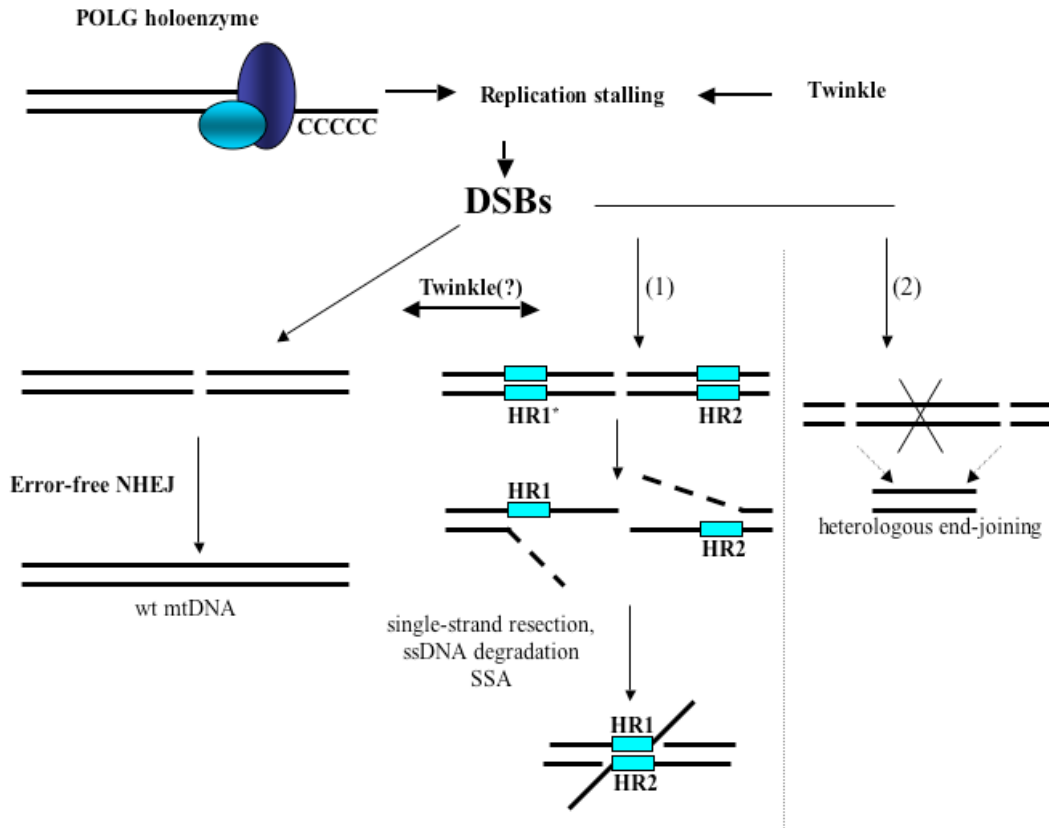
**Figure 6.3:** *Proposed model for multiple deletion formation by error prone NHEJ. Based on the data presented in this thesis we propose that multiple deletion formation are induced by replication stalling, for example during replication of homopolymeric runs, followed by replication stalling and DSB formation. While the vast majority of DSB are repaired by precise endjoining, a minor proportion is repaired via error-prone mechanisms: (1) via a single strand annealing type of mechanism (SSA) following a single DSB, or (2) via intramolecular heterologous end-joining in molecules with more than 1 DSB. \* HR: homology region, can be very short and imperfect, from 2-7 nucleotides.*

## 6.5 The HEK293 Flp-In™ T-Rex™ inducible expression system to study mtDNA maintenance

The use of cultured cells as experimental approach has several advantages compared to the use of an isolated preparation or intact animals. Cultured cells provide a more controlled and easily manipulated cellular environment (Adler, 2006). In order to study the mtDNA maintenance machinery we established cell-lines that express different variants of proteins involved in mtDNA maintenance using HEK 239 Flp-In™T-Rex™ cells. We established cell-lines to study expression of transcription factor A (TFAM), mitochondrial termination factor 1, MTERF1, and other mtDNA maintenance related proteins (Pohjoismäki et al, 2006, Hyvärinen et al, unpublished data). The amount of expression of the exogenous protein could be readily manipulated by titration of DC to

the cell culture growth medium resulting in an increase of expression between 1 and 10 ng DC per ml growth medium. By use of immuno cytochemistry (ICC) we showed that all expressed protein variants were targeted to the mitochondrial compartment. With the exception of some of the Twinkle mutants, using the HEK 239 Flp-In™T-Rex™ cells we did not observe loss of expression of the exogenous protein even after long-term cell-culture growth. In earlier studies performed with the Hela tet-on and 293T tet-on inducible system long-term cell culture growth resulted in a mixed cell-culture population, in which a majority of cells had lost exogenous protein expression (unpublished results). This loss of expression was probably due to a relatively high exogenous protein expression when cells were grown without inducer, presumably caused by so-called position effects. This so called 'leaky expression' was very low in HEK 239 Flp-In™T-Rex™ cells since this system uses site-specific recombination at a single genomic locus to integrate the recombinant expression plasmid. Leaky expression could only be detected in highly overexposed Western blots (data not shown). In conclusion, the cell-line system we established is a useful tool to determine enzymatic and structural roles of mtDNA maintenance related proteins.

## 6.6 The HEK293 Flp-In™ T-Rex™ inducible expression of different POLG1 variants

Maintenance of mtDNA is critically dependent on Pol γ, since it is the only polymerase known to be active in the mitochondrial compartment. The catalytic subunit of Pol γ is a family A polymerase encoded by the POLG1 gene, which is implicated in a number of human disorders (Graziewicz et al, 2006). To study mtDNA replication and more generally the function of POLG1 we established several inducible HEK 239 Flp-In™T-Rex™ cells which upon DC induction express different POLG1 variants. Overexpression of exogenous wild-type POLG1 did not affect mtDNA copy number and we did not observed obvious changes in replication intermediates (RIs), showing that in cell culture POLG1 alone is not rate limiting for mtDNA replication. This might be explained by its interaction with the processivity factor POLG2. Pol γ has been shown by surface plasmon resonance to function as heterotrimer in the same study they confirmed these interactions by analytical gel filtration and isothermal titration calorimetry (Yakubovskay et al, 2006). Based on *in vitro* studies POLG1 without its subunit (POLG2) is not expected to be able to synthesise more than 100 nucleotides before it falls of the mtDNA strand (Lim et al, 1999). However 293T cells that constitutively overexpress both POLG1 and POLG2 also do not increase mtDNA copy number (Spelbrink, unpublished data).

Downstream of the N-terminal mitochondrial targeting signal the human POLG1 gene includes a CAG tri-nucleotide repeat that normally codes for 10 glutamines (Ropp & Copeland, 1996; Lecrenier et al, 1997). It remains unclear why alterations in the POLG1 CAG repeat can lead to male infertility (Rovio et al, 2001). After overexpression of a POLG1 variant that lacks this CAG repeat (ΔCAG) we did not observe any obvious effect on RIs or mtDNA copy number. Sperm however is a very specialized cell type in the testis, a tissue that undergoes many rounds of cell cycles and thus susceptible for problems in energy production (Diez-Sanchez et al, 2003). Glutamine tracts are regarded as regions of protein-protein interactions; therefore the CAG repeat might be important in

a sperm-specific interaction. Alterations in the CAG repeat could hinder such an interaction leading to tissue specific mtDNA alterations. MtDNA deletions in sperm have indeed been described and suggested to result in reduced fertility (Reynier and Malthiery, 1995; Reynier et al, 1998; Kao et al, 1998). Another possibility is that alterations in the CAG repeat lead to a very subtle defect in POLG1 activity, not enough to cause trouble in other tissues, but problematic during a strong increase of mtDNA copy number per cell volume which has been suggested to happen during spermatogenesis (Diez-Sanchez et al, 2003). One study did not find any increase in single nucleotides substitutions in sperm mtDNA of infertile males carrying a mutant POLG1 allele, however with the method used only heteroplasmy levels as high as 12% could be detected (Harris et al, 2006). Another study measured mutation level by high fidelity mtDNA PCR amplification, followed by cloning and sequencing of individual plasmids. It showed that alterations in the POLG1 CAG repeat lead to a sperm specific accumulation of mtDNA mutation (Rovio, 2006). This higher mtDNA mutation level was confined to the control region and the mutations were only found at very low heteroplasmy. This questions the pathogenic relevance of these control region mutations (Rovio, 2006).

The exonuclease activity identified in POLG1 is responsible for proofreading newly incorporated bases during ongoing mtDNA synthesis (Kunkel and Mosbaugh, 1989). In cell culture constitutive expression of a POLG1 variant (D198A) that is deficient in proofreading results in accumulation of randomly distributed mtDNA mutations (Spelbrink et al, 2000). To validate our HEK 239 Flp-In$^{TM}$ T-Rex™ cell culture system we determined the mtDNA mutation level after 60 days of D198A induced expression. In agreement with earlier studies we found elevated levels of mtDNA mutations in at least two regions (cyt b and control region). Earlier studies showed that D198A is not compromised in its polymerase activity (Spelbrink et al, 2000). Modest induction (3ng/ml DC) of D198A expression showed little change in RIs. Interestingly, after full expression (20ng/ml DC) we did observe a more enhanced y- and bubble arc, suggesting replication stalling. However this phenotype was clearly less severe than that after low D1135A and D890N expression and did not result in a decreased mtDNA level. The results might be explained by slowing down of mtDNA replication after overexpression of D198A, however not slow enough to result in mtDNA depletion. This slowing down could be a result of the D198A POLG1 variant repeatedly trying in vain to switch from exonuclease to polymerase active site but further investigation is needed to explain this observed stalling.

## 6.7 Is replication stalling the first step towards multiple deletion formation in adPEO?

Based on our mutation analysis and deletion mapping in adPEO patients we hypothesized that the mutations in POLG1 or Twinkle lead to frequent replication stalling or pausing resulting in the accumulation of multiple deletions. In order to test this hypothesis we constructed HEK 239 Flp-In$^{TM}$T-Rex$^{TM}$ cells expressing Twinkle or POLG1 variants carrying non-disease catalytic (Wanrooij et al, 2007) as well as adPEO associated mutations (Goffart et al, manuscript in preparation).

The non-disease Twinkle variants that we constructed were a K421A and G575D substitution, besides another Twinkle variant was used containing a 13 AA duplication (aa 352-364) within the linker region that we previously identified in a Finnish adPEO family. The K421A substitution was selected because the amino acid is analogous to the T7 primase/helicase lysine 318 substitution by alanine that leads to a 200-fold decrease of dTTPase activity resulting in complete loss of helicase activity (Patel et al, 1994). Furthermore, it is a completely conserved residue in the Walker A motif of the superfamily of ATPases to which Twinkle belongs and that also include the mitochondrial $F_1$ATPase. In agreement with the T7 primase/helicase mutant variant, our mutant did not contain helicase activity. G575D is a Twinkle variant that contains a mutation in the H4 helicase motif. Previous studies on the analogous mutation in T7 primase/helicase (G488D) demonstrated that this residue is important for DNA binding and is critical for DNA unwinding (Washington et al, 1996). We also constructed a HEK 293 Flp-In™T-Rex™ cell-line expressing the Y955C POLG1 variant, a dominant mutation resulting in PEO, two non-disease catalytic POLG1 variants which we studied were D1135A and D890N. Transient expression of human POLG1 variants containing mutations at the corresponding aspartate 890 and 1135 residues leads to mtDNA depletion in cell culture (Spelbrink et al, 2000).

We have shown that expression of these POLG1 (D890N,Y955C, D1135A) and Twinkle (duplication aa 352-364, K421A, G575D) variants resulted in a strong decrease of mtDNA copy number (Wanrooij et al, 2007 and Goffart et al, manuscript in preparation). The observed reduction of mtDNA (more then 80%) after 3 days of overexpression suggested a complete halt of newly synthesized mtDNA molecules. This rapid decrease of mtDNA content coincided with a substantial increase in y- and bubble arc RIs detected by 2DNAGE. These observations indicate that that expression of these mutants results in severe replication stalling/pausing. The expression of these Twinkle variants seemed to result in a stronger stalling phenotype compared to the POLG1 variants since most intermediates were found within the bubble arc, suggesting most replication stalls during the early stage of replication. However, this difference might be explained by expression level differences, 3ng/ml DC induction results in a recombinant Twinkle 8-10 fold the endogenous Twinkle level, whereas the same cell growth conditions resulted in comparable POLG1 recombinant and endogenous POLG1 protein levels. These experiments showed that expression of a polymerase deficient POLG1 variant or helicase deficient Twinkle variant can lead to replication stalling and supports the idea that POLG1 or Twinkle mutation in adPEO patients lead to a frequent stalling of mtDNA replication. In agreement with this, overexpression of the most common POLG1 adPEO mutation (Y955C) and a Twinkle adPEO variant (duplication aa 352-364) in our established cell culture system resulted in a similar stalling phenotype (unpublished data).

**6.8. The different mtDNA replication stalling phenotypes in mutant Twinkle and POLG1 expressing cell-lines can be explained by the RITOLS model of mtDNA replication**

Although for long mtDNA was believed to be replicated via continuous, but asynchronous synthesis of each strand (Robberson et al, 1972). 2DNAGE analysis of mtDNA revealed RIs with properties of conventional strand-coupled DNA synthesis (Holt et al, 2000). These strand-coupled RIs co-existed with another type of RIs, which were initially believed to account for asynchronous DNA synthesis. Subsequently these RIs were shown to posses different features as predicted by the asynchronous replication model (Yang et al, 2002). The RIs were shown to be RNA-rich and it was suggested that mtDNA replication often involves RNA incorporation throughout the lagging strand (RITOLS) (Yasukawa et al, 2006).

The stalling phenotype as a consequence of induced expression of G575D and K421A resulted in reduced fork movement throughout the genome. We used 3 different probes for fragments that cover half of the genome and stalling RIs were visible by use of all. When a probe was used against a fragment that contains $O_H$ we obtained intense bubble arcs and relatively weak ascending y arcs, suggesting initiation mainly occurred in the non-coding region. The initiation site near $O_H$ suggests intermediates represent RITOLS RIs or strand-displacement mtDNA replication (Clayton, 1982; Yasukawa et al, 2006). However the lack of ssDNA RIs as expected by the strand-displacement model of replication and the RNA-poor intermediates show that the RIs mainly consist of dsDNA resembling strand-coupled RIs. The dsDNA nature of these RIs was confirmed by RIs purification followed by digestion with dsDNA and ssDNA specific restriction endonucleases. However the predominant initiation close to $O_H$ shows they canot represent conventional coupled leading- and lagging-strand RIs, since one property of coupled leading- and lagging-strand mtDNA replication is its initiation in a broad zone referred to as OriZ, which does not include the non-coding region (Bowmaker et al, 2003). Thus the dsDNA RIs detected after mutant Twinkle proteins induce replication stalling can only be explained by the RITOLS model of replication in which an increased rate of lagging-strand initiation and/or maturation relative to rate of fork progression would lead to accumulation of dsDNA molecules.

The replication mode of mtDNA remains controversial (Bogenhagen and Clayton, 2003a; Holt and Jacobs, 2003; Bogenhagen and Clayton, 2003b). Proponents of the strand-displacement model of replication have suggested that the RNA-rich intermediates do not represent bonafide replication intermediates. Instead they suggest the RNA is produced by the transcription machinery and cannot be considered RIs. In disagreement with this is that the origin of the RITOLS products have been mapped around $O_H$ (Yasukawa et al, 2006) and not close to the transcription promoters, which would be expected from mitochondrial transcription products. Indeed on the dsDNA mtDNA molecule RNA polymerase is only able to start transcription at the promoters (Asin-Cayuela and Gustafsson, 2007), however on ssDNA mammalian RNA polymerase can start RNA production unspecifically (Tsurumi and Lehman, 1990). Perhaps a stronger argument in favour of the existence of the RITOLS replication model is the observation that the RNA

intermediates are specific to the lagging strand (Yasukawa et al, 2006), if they are products from the transcription machinery it seems hard to explain why such RNA-rich fragments are not found in the leading strand. Moreover our experiments showed that POLG1 is involved in the maturation of the RITOLS intermediates. This suggests that RITOLS are RIs, since till date POLG1 has only been reported to be involved in different forms of mtDNA metabolism. However, little is known yet about the precise RITOLS mechanism and therefore it does not present a complete model of mtDNA replication. Future experiments are needed to show which enzymes incorporate the initial RNA on the lagging-strand and what is the mechanism behind RNA maturation to dsDNA. Although we clearly demonstrated that POLG1 is involved in the RITOLS maturation it is likely that more enzymes are involved in this process. Both the strand-coupled and strand-displacement replication models need one or more short RNA primers to start lagging-strand DNA synthesis, but to date no primase has been identified. In conclusion the mtDNA replication mechanism needs to be investigated further to assess the actual replication mode.

**Summary**

In this study we showed the identification of a gene, *PEO1,* encoding for Twinkle, a new mitochondrial protein with striking similarity to the phage T7 primase/helicase. It was shown that Twinkle colocalizes with mtDNA in nucleoprotein complexes referred to as mitochondrial nucleoids. The co-segregation of the Twinkle mutations within all adPEO families, the lack of these mutations in controls and the strong conservation of mutated residues provided strong evidence that the mutations are of pathogenic significance in this disease. With this we showed evidence that Twinkle mutations cause a disorder of mtDNA maintenance, indicating that the activity of Twinkle is needed for maintaining mtDNA integrity.

In order to test one of the proposed mechanisms (slipped mispairing) by which POLG1 and Twinkle mutations result in the accumulation of mtDNA deletions in adPEO we measured mtDNA mutation levels in affected tissues from PEO patients. We found enhanced control region mutation levels in muscle mtDNA of PEO patients but no evidence of a general decrease in replication fidelity. The demonstration of a similar high control region mutation burden in sporadic single deletion patients and the late age at which these mutations start to accumulate suggested they were a secondary consequence of mtDNA deletions that obstruct the OXPHOS system. Specific amplification of mtDNA molecules containing deletions showed that breakpoints were generally found at regions of difficulty for the replication machinery such as homopolymeric runs and microsatelite-type of repeats and suggested frequent replication stalling/pausing as a primary cause of mtDNA deletion formation. It was indeed demonstrated that expression of Twinkle and POLG1 mutants in human cell culture decreased mitochondrial DNA copy-number by severe replication stalling thus supporting our hypothesis and suggesting that replication stalling could be the first step towards multiple deletion formation in adPEO.

# 7.    Conclusions and future prospects

## 7.1 Conclusions

Correct maintenance of mtDNA is critical for life. Mechanisms of mtDNA maintenance have recently acquired wide interest due to their role in inherited diseases as well as in aging. In our study we identified a human helicase Twinkle, a novel key player in maintenance of the mitochondrial genome. Twinkle defects result in accumulation of mtDNA abnormalities in the form of multiple mtDNA deletions leading to the adPEO disorder. The mechanism by which mutations in different nDNA encoded genes results in mtDNA deletion formation is of essential importance in understanding the disease pathology. Our results strongly suggest that mtDNA deletions in adPEO accumulate as a result of frequent replication stalling.  This study has generated deeper insight in the adPEO disease mechanism and the proteins involved in this disorder. This has in addition lead to a further basic understanding of the mechanisms of mtDNA maintenance. My main conclusions are the following:

1)     We identified human gene PEO1, encoding for a protein similar to the phage T7 primase/helicase gp4.  The protein products of this gene are the full-length 77kD protein Twinkle and a shorter 66kD splice variant Twinky.
2)     Twinkle is a protein localized to mitochondrial nucleoids.
3)     Overexpression of Twinkle in mitochondrial extracts result in a modest increase of mtDNA helicase activity.
4)     Mutations in Twinkle co-segregate with the adPEO disorder. We identified 11 different mutations in 12 adPEO families of various ethnic origins.
5)     Twinkle and POLG1 adPEO mutations do not result in a general decrease of replication fidelity.
6)     Twinkle and POLG1 adPEO defects enhance age-dependent accumulation of mostly point-mutations in the control region of mtDNA.
7)     mtDNA deletions in adPEO post-mitotic tissues occur at sites that are difficult for the mitochondrial replication machinery, such as homopolymeric runs and microsatelite type of repeats.
8)     The HEK 239 Flp-In™T-Rex™ cells that we have established provides an excellent tool to study the function of proteins in mtDNA maintenance.
9)     Expression of several Twinkle and POLG1 mutants results in mtDNA copy-number depletion. Overexpression of wild-type variants of these proteins does not alter mtDNA content in cultured human cells.
10)    Expression of mutant Twinkle and POLG1 variant result in different replication stalling phenotypes. The stalling induced by deficient Twinkle results in RIs that mimic conventional strand-coupled RIs and suggest that initiation of lagging-strand DNA synthesis or maturation occurs at multiple sites across the genome. We furthermore propose that this maturation involves POLG1.

## 7.2 Future prospects

Initial studies on the Twinkle gene variants associated with adPEO, showed that the mutations do not affect nucleoids localization or multimerization. This study however

contained only two different adPEO mutations (duplication AA352-364 and W474C) and does not exclude that other Twinkle adPEO mutations affect these protein characteristics. It is possible however, that some Twinkle mutations found in adPEO patients would affect the multimerization of the protein, or nucleoid co-localization and with that the helicase activity. The established cell culture expression system (HEK 239 Flp-In™T-Rex™) could be used to study such *in vivo* effects of mutations in Twinkle and POLG1 associated with human disorders of mtDNA maintenance.

Under normal cell culture condition, cell-lines expressing Twinkle and POLG1 adPEO variants do not appear to accumulate mtDNA deletions at high levels (data not shown). Probably, under optimal cell culture conditions the late onset disease feature of PEO (slow accumulation of mtDNA deletions) could be a limiting factor. Serum starvation and perhaps nucleotide pool depletion could result in a higher rate of DSBs and deletion accumulation. Cell-lines that would accumulate multiple mtDNA deletions would be a valuable tool in studying adPEO disease mechanism.

Our results strongly suggest that POLG1 is involved in lagging-strand maturation in RITOLS mtDNA replication. Little is known about the lagging strand maturation and processing of RNA-DNA hybrids to form dsDNA. It would be interesting to assess the precise role of POLG1 in this process and moreover what additional proteins interacting with POLG1 are involved in this process. Another interesting point concerning the RITOLS replication mechanism that needs to be clarified is which enzyme initially lays down RNA on the lagging-strand.

Although our results suggest that the higher control region mutation level found in adPEO is an indirect consequence of the mtDNA deletions, the non-random distribution of the accumulated mutations remains intriguing. Especially the observation that in all our POLG1 adPEO patients the fast majority (81%) of the mutations accumulate in a small region responsible for replication control could be addressed in future experiments. To study the significance of the A183G mutation found in high heteroplasmy in nearly all POLG1 adPEO patients might also provide valuable information concerning mtDNA maintenance.

Overexpression of the exonuclease deficient POLG1 D198A did not seriously affect mtDNA copy-number or the abundance of RNA-rich versus RNA-poor RIs. Nevertheless, moderate changes in replication intermediates at high D198A expression levels were observed by 2DNAGE and warrant further investigation.

These and other future experiments will provide deeper comprehension of mtDNA replication and cross-talk between mitochondria and the nucleus. Furthermore, this basic research and the resulting understanding of the mechanisms that lead to multiple deletion formation will contribute to the development of potential therapies of adPEO and possibly other human diseases associated with mitochondrial dysfunction.

## Acknowledgements

I will remember my years as a PhD student in Tampere as being very eventful; although I had my bad moments, all in all I consider my stay in Tampere as a happy time. During these years a lot of people have helped me with work-related and personal issues and finally I like to thank them

First of all I would like to thank the person who made the biggest contribution to this thesis, my supervisor Hans Spelbrink. I thank you for teaching me many laboratory techniques and anything else that is important in the world of science from basic things like attending seminars to the actual theoretical biology. Thank you for taking me as PhD student and organizing my funding, something I never had to worry about. Furthermore, thanks for good games of badminton, tennis and pool and together with Herma Renkema for pleasant nights out and lifts from and to Lentävänniemi during my first year in Tampere.

I am grateful to Howy Jacobs for giving me the job as technician when I came from Amsterdam to Tampere and for many relevant and stimulating questions during times I presented my research. I enjoyed 4 years of Murikka and they were always hard work but very fun and stimulating, thanks for that! I would like to acknowledge Anu Wartiovaara for providing patient samples, being essential for my research and as a "follow-up-group" member of my thesis for showing interest in my results.

I like to especially acknowledge Professor Carlos Moraes for his rapid review of my manuscript. His comments were very valuable. Professor Ronald Wanders I thank for help in order to get the position in Tampere.

I would like to thank Steffi Goffart for introducing me to the world in 2 dimensions, for her very hard work on our common projects and especially for carrying out the last experiments that were required for the final article of this thesis when I had left Tampere.

Merja Jokela for excellent cell-culture help, many western blots and lots of other things. It has always been nice to work with you and besides that you are just very good in your job, whether you believe it or not. Together with Outi Kurronen I also thank you for keeping the lab running, making it so much easier to work for all other people in the lab, including me.

I am grateful to Maria Falkenberg and Claes Gustafsson for giving me the time to finalise my thesis, for taking time to review it and for being very pleasant people to work with! To the people of their labs, thanks for receiving me so warmly as newcomer to the lab. To Zsolt thanks for keeping me positive about the progress of my thesis. Nicola Solaroli thanks for introducing me to FC Internovum.

~~Pim!~~Gertjan Hakkaart, the most misunderstood person I have ever met ("Why do you say like that?!"). Thank you for all the fun we had together in Tampere. At work or out in town, it does not matter - when you are around things are just never boring. Moreover for

keeping me focused on my thesis and on other things that are important and keeping a cool head despite other disturbing things. Your advice has been crucial for finishing this PhD ("all by myself!"). Further for your friendship, great humour and for always telling the truth, something very few people dare to do ("excuse me!").

Peter M(elberg)artinsson. Although I will be living in Sweden now for a while, I am pretty sure I will never meet anybody who is more Swedish than you. You are a warm person and good friend. Thanks for endless discussions about football and even more important things in life. I apologize for complaining to you too much about the proceedings of my thesis the last year, however you always gave me good support. Tack så mycket!

I thank my friend Emilia for extreem good cooking, for the great times when we were out and for being a very nice office/lab companion. Dakujem! Hope to see you soon in Stockholm!

Ilkka (Junthla) Junttila, ystäväni, kiitos for running the Helsinki city marathon together. Moreover for endless talks about gossip, sports, work, women, Michael Jackson ☺ and religion. In particular I want to thank you for being a very big support for me when my world was falling apart…I will not forget that.

Jaakko Pohjoismäki, my Murikka cell-mate, lounaskaveri and co-PhD student. For nice conversations during lunch throughout many years and for fun in the lab and Murikka. Your knowledge about science, insects, politics and many other things have impressed me many times, but moreover you are just a very nice straightforward guy. All the best to you! Nuria Garrido for being a great positive spirit in the lab and for introducing me to good música Español. ¡muchas kiitos, querida Nuria! I also thank Sanna Horttanainen for important help in the lab and for the fun we had together at work and during nights out in Manse.

Kaisu Pekoselle suuret kiitokset monista pienistä mutta tärkeistä asioista, ja airon Tampereem murteen opettamisesta! Also Ilkka Mäkelä for nice conversations during the last year. Anne Hyvärinen for pleasant long talks in the cell-culture room. Katariina Juhola for practical help during the years. Janne Toivonen for fun evenings. Eric Dufour and Alberto Sanz for bringing a positive new spirit to the labs in Tampere. Esko Kemppainen for organizing football matches and downloading music. Marina Toompuu for giving me a good example of how a good PhD thesis should be written. Katja Peltonen for being very fun. Rhoda for teaching me to dance like a Greek. I thank other previous Hans/Howy-lab members for various things, Henna Tyynismaa, Kaisa Luoto, Helen Lindsay, John Walker, Colin McLeod, Satu Kärkkäinen and Olli Kajander. To other people I am thankful for pleasant conversations on and around the IMT campus: Taisto "What's happening?" Heinonen, Keijo Viiri, Iivari Kleino, Jukka Pitkänen, Heimo Koskinen, Anne Koskela and the nice neighbour Professor Markku Kulomaa. I thank Eino Palin and Nuno Raimundo for fun at FinMit and Murikka.

Now I would like to thank the relatives and friends that have been important for me during the years I spent in Tampere. Philipp Beierer, Gerbrand Bijmolt, Rickard Farnemyhr and Frederik Vermaete I thank for lots of sports, for fun evenings out in town or playing Risk, for brewing the best beer in town (okay, I confess the banana beer was not a success) and most of all for their friendship. Further I would like to thank the "Hollantilaiset Hullut" for relaxing Thursday evenings over the years in Salhojankadunpubbi.

I would like to thank the whole family Jokinen for many pleasant Sundays I spent at your home in Pirkkala in particular thanks to Anne for many things.

My group of friends from back in Amsterdam, already going strong since high school. Marcel Hasper for too much to put into words, Martijn Heuckelbach for always having time to listen to my problems, Remco Keijser for friendship and together with Ronald van Olffen for great help in surviving the terror of the HLO; Bas van der Horst for staying in touch and your fun visit to Finland.

I would like to thank Timo, Hely, Tomi and Henri ("mun kaveli Henli") Matikainen for many pleasant days I spent in your warm home in Lempäälä, and for help with many things, from moving to pakoputki- and other car problems.

My aunt and uncle Anni and Mar Wanrooij, thank you for always showing interest in my life and studies, as well for keeping me up to date with the latest sport results. I would like to thank my Father and Frances for their love and support during these years. Especially many thanks for your help during difficult times. Jeroen and Minna for being very good friends and the best family one could wish for. Ester for being such a positive, lively nice "iso tyttö" and of course Lukas. I am grateful to my Mother for her many visits to Finland and for her continuous unconditional support throughout my 28 years.

Most of all I would like to thank my love Paulina for her endless patience with me. I am happy and proud that you are my woman. Thanks for joining me to Stockholm, Minä Rakastan Sinua!

Stockholm May 2007

Sjoerd Wanrooij

# References

Adler EM (2006): Teaching resources. Cell culture as a model system for teaching: using PC12 cells. Sci STKE 334: tr5.

Agostino A, Valletta L, Chinnery PF, Ferrari G, Carrara F, Taylor RW, Schaefer AM, Turnbull DM, Tiranti V and Zeviani M (2003): Mutations of ANT1, Twinkle and POLG1 in sporadic progressive external ophthalmoplegia (PEO). Neurology 60: 1354-1356.

Albring M, Griffith J and Attardi G (1977): Association of a protein structure of probable membrane derivation with HeLa cell mitochondrial DNA near its origin of replication. Proc Natl Acad Sci USA74: 1348-1352.

Alam TI, Kanki T, Muta T, Ukaji K, Abe Y, Nakayama H, Takio K, Hamasaki N and Kang D (2003): Human mitochondrial DNA is packaged with TFAM. Nucleic Acids Res 31: 1640-1645.

Anderson CT and Friedberg EC (1980): The presence of nuclear and mitochondrial uracil-DNA glycosylase in extracts of hum KB cells. Nucleic Acids Res 8: 875-888.

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R and Young IG (1981): Sequence and organization of the human mitochondrial genome. Nature 290: 457-465.

Andersson GE, Karlberg O, Canbäck B and Kurland CG (2003): On the origin of mitochondria: a genomics perspective. Philos Trans R Soc Lond B Biol Sci 358: 165-177.

Ashley N, Harris D and Poulton J (2005): Detection of mitochondrial DNA depletion in living human cells using PicoGreen staining. Exp Cell Res 303: 432-436.

Asin-Cayuela J and Gustafsson CM (2007): Mitochondrial transcription and its regulation in mammalian cells. Trends Biochem Sci 32: 111-117.

Attardi G and Schatz G (1988): Biogenesis of mitochondria. Annu Rev Cell Biol 4: 289-333.

Arenas J, Briem E, Dahl H, Hutchison W, Lewis S, Martin MA, Spelbrink H, Tiranti V, Jacobs H and Zeviani M (2003): The V368i mutation in Twinkle does not segregate with adPEO. Ann Neurol 53: 278.

Arner ESJ and Eriksson S (1995): Mammalian deoxyribonucleoside kinases. Pharmac Ther 67: 155-186.

Barat M, Rickwood D, Dufresne C and Mounolou JC (1985): Characterization of DNA-protein complexes from the mitochondria of Xenopus laevis oocytes. Exp Cell Res 157: 207-217.

Barnes DE, Johnston LH, Kodama K, Tomkinson AE, Lasko DD and Lindahl T (1990): Human DNA ligase I cDNA: cloning and functional expression in Saccharomyces cerevisiae. Proc Natl Acad Sci USA 87: 6679-6683.

Bebenek K, Abbotts J, Roberts JD, Wilson SH and Kunkel TA (1989): Specificity and mechanism of error-prone replication by human immunodeficiency virus-1 reverse transcriptase. J Biol Chem 264: 16948-16956.

Bender A, Krishnan KJ, Morris CM, Taylor GA, Reeve AK, Perry RH, Jaros E, Hersheson JS, Betts J, Klopstock T, Taylor RW and Turnbull DM (2006): High levels of mitochondrial DNA deletions in substantia nigra neurons in aging and Parkinson disease. Nat Genet 38: 515-517.

Berk AJ and Clayton DA (1973): A genetically distinct thymidine kinase in mammalian mitochondria. Exclusive labeling of mitochondrial deoxyribonucleic acid. J Biol Chem 248: 2722-2729.

Birkenhager R, Hoppert M, Deckers-Hebestreit G, Mayer F and Altendorf K (1995): The F0 complex of the Escherichia coli ATP synthase. Investigation by electron spectroscopic imaging and immunoelectron microscopy. Eur J Biochem 230: 58-67.

Bleazard W, McCaffery JM, King EJ, Bale S, Mozdy A, Quinton T, Nunnari J and Shaw JM (1999): The dynamin-related GTPase Dnm1 regulates mitochondrial fission in yeast. Nat Cell Bio 1: 298-304.

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

Bogenhagen DF and Clayton DA (2003a): The mitochondrial DNA replication bubble has not burst. Trends Biochem Sci 28: 357-360.

Bogenhagen DF and Clayton DA (2003b): Concluding remarks: the mitochondrial DNA replication bubble has not burst. Trends Biochem Sci 28: 404-405.

Bogenhagen DF, Wang Y, Shen EL and Kobayashi R (2003): Protein components of mitochondrial DNA nucleoids in higher eukaryotes. Mol Cell Proteomics 2: 1205-1216.

Boldogh I, Vojtov N, Karmon S and Pon LA (1998): Interaction between mitochondria and the actin cytoskeleton in budding yeast requires two integral mitochondrial outer membrane proteins, Mmm1p and Mdm10p. J Cell Biol 141: 1371-1381.

Bourgeron T, Chretien D, Amati P, Rotig A, Munnich A and Rustin P (1993): Expression of respiratory chain deficiencies in human cultured cells. Neuromuscul Disord 3: 605-608.

Bowmaker M, Yang MY, Yasukawa T, Reyes A, Jacobs HT, Huberman JA and Holt IJ (2003): Mammalian mitochondrial DNA replicates bidirectionally from an initiation zone. J Biol Chem 278: 50961-50969.

Bradshaw HD jr (1983): Molecular cloning and cell cycle-specific regulation of a functional human thymidine kinase gene. Proc Natl Acad Sci USA 80: 5588-5591.

Brewer BJ and Fangman WL (1987): The localization of replication origins on ARS plasmids in S. cerevisiae. Cell 51: 463-471.

Brewer BJ and Fangman WL (1988): A replication fork barrier at the 3' end of yeast ribosomal RNA genes. Cell 55: 637-643.

Brockington M, Alsanjari N, Sweeney MG, Morgan-Hughes JA, Scaravilli F and Harding AE (1995): Kearns-Sayre syndrome associated with mitochondrial DNA deletion or duplication: a molecular genetic and pathological study. J Neurol Sci 131: 78-87.

Brown MD, Trounce IA, Jun AS, Allen JC and Wallace DC (2000): Functional analysis of lymphoblast and cybrid mitochondria containing the 3460, 11778 or 14484 Leber's hereditary optic neuropathy mitochondrial DNA mutation. J Biol Chem 275: 39831-39836.

Brown TA and Clayton DA (2002): Release of replication termination controls mitochondrial DNA copy number after depletion with 2',3'-dideoxycytidine. Nucleic Acids Res 30: 2004-2010.

Brown TA, Cecconi C, Tkachuk AN, Bustamante C and Clayton DA (2005): Replication of mitochondrial DNA occurs by strand displacement with alternative light-strand origins, not via a strand-coupled mechanism. Genes Dev 19: 2466-2476.

Brown TA and Clayton DA (2006): Origins and migrations in asymmetrically replicating mitochondrial DNA. Cell Cycle 5: 917-921.

Carelli V, Achilli A, Valentino ML, Rengo C, Semino O, Pala M, Olivieri A, Mattiazzi M, Pallotti F, Carrara F, Zeviani M, Leuzzi V, Carducci C, Valle G, Sadun AA and Torroni A (2006): Haplogroup effects and recombination of mitochondrial DNA: novel clues from the analysis of Leber hereditary optic neuropathy pedigrees. Am J Hum Genet 78: 564-574.

Carrodeguas JA, Theis K, Bogenhagen DF and Kisker C (2001): Crystal structure and deletion analysis show that the accessory subunit of mammalian DNA polymerase γ, PolγB, functions as a homodimer. Mol Cell 7: 43-54.

Carroll J, Fearnley IM, Shannon RJ, Hirst J and Walker JE (2003): Analysis of the subunit composition of complex I from bovine heart mitochondria. Mol Cell Proteomics 2: 117-126.

Cecchini G (2003): Function and structure of complex II of the respiratory chain. Annu Rev Biochem 72: 77-109.

Cerritelli SM, Frolova EG, Feng C, Grinberg A, Love PE and Crouch RJ (2003): Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. Mol Cell 11: 807-815.

Cha RS and Kleckner N (2002): ATR homolog Mec1 promotes fork progression, thus averting breaks in replication slow zones. Science 297: 602-606.

Chan SS, Longley MJ, Copeland WC (2005a): The common A467T mutation in the human mitochondrial DNA polymerase (POLG) compromises catalytic efficiency and interaction with the accessory subunit. J Biol Chem 280: 31341-31346.

Chan SSL, Longley MJ, Naviaux RK and Copeland WC (2005b): Mono-allelic POLG expression resulting from a nonsense-mediated decay and alternative splicing in a patient with Alpers syndrome. DNA Repair 4: 1381-1389.

Chen H, Detmer SA, Ewald AJ, Griffin EE, Fraser SE and Chan DC (2003): Mitofusins Mfn1 and Mfn2 coordinately regulate mitochondrial fusion and are essential for embryonic development. J Cell Biol 160: 189-200.

Chen H and Chan DC (2005): Emerging functions of mammalian mitochondrial fusion and fission. Hum Mol Genet 14: R283-R289.

Chen XJ and Butow RA (2005): The organization and inheritance of the mitochondrial genome. Nat Rev Genet 6: 815-825.

Chen XJ, Wang X, Kaufman BA and Butow RA (2005): Aconitase couples metabolic regulation to mitochondrial DNA maintenance. Science 307: 714-717.

Chinnery PF, Howell N, Lightowlers RN and Turnbull DM (1997): Molecular pathology of MELAS and MERFF. The relationship between mutation load and clinical phenotypes. Brain 120: 1713-1721.

Chinnery PF and Samuels DC (1999): Relaxed replication of mtDNA: A model with implications for the expression of disease. Am J Hum Genet 64: 1158-1165.

Chinnery PF and Schon EA (2003): Mitochondria. J Neurol Neurosurg Psychiatry 74: 1188-1199.

Chottiner EG, Shewach DS, Datta NS, Ashcraft E, Gribbin D, Ginsburg D, Fox IH and Mitchell BS (1991): Cloning and expression of human deoxycytidine kinase. Proc Natl Acad Sci USA 88: 1531-1535.

Clark AJ and Margulies AD (1965): Isolation and characterization of recombination-deficient mutants of Escherichia coli K12. Proc Natl Acad Sci USA 53: 451-459.

Clayton DA, Doda JN and Freidberg EC (1974): The absence of a pyrimidine dimer repair mechanism in mammalian mitochondria. Proc Natl Acad Sci USA 71: 2777-2781.

Clayton DA (1982): Replication of animal mitochondrial DNA. Cell 28: 693-705.

Coller HA, Bodyak ND and Khrapko K (2002): Frequent intracellular clonal expansions of somatic mtDNA mutations: significance and mechanisms. Ann N Y Acad Sci 959: 434-447.

Cooper JM, Mann VM and Schapira AH (1992): Analyses of mitochondrial respiratory chain function and mitochondrial DNA deletions in human skeletal muscle: effect of aging. J Neurol Sci 113: 91-98.

Cormier V, Rotig A, Quartino AR, Forni GL, Cerone R, Maier M, Saudubray JM and Munnich A (1990): widespread multi-tissue deletions of the mitochondrial genome in the Pearson marrow-pancreas syndrome. J Pediatr 117: 599-602.

Cortopassi GA, Shibata D, Soong NM and Arnheim N (1992): A pattern of accumulation of a somatic deletion of mitochondrial DNA in aging human tissues. Proc Natl Acad Sci USA 89: 7370-7374.

Corral-Debrinski M, Horton T, Lott MT, Shoffner JM, McKee AC, Beal MF, Graham BH and Wallace DC (1994): Marked changes in mitochondrial DNA deletion levels in Alzheimer brains. Genomics 23: 471-476.

Coskun PE, Ruiz-Pesini E and Wallace DC (2003): Control region mtDNA variants: Longevity, climatic adaptation, and a forensic conundrum. Proc Natl Acad Sci USA 100: 2174-2176.

Coskun PE, Beal MF and Wallace DC (2004): Alzheimer's brains harbor somatic mtDNA control-region mutations that suppress mitochondrial transcription and replication. Proc Natl Acad Sci USA 101: 10726-10731.

Cottrell DA and Trunbull DM (2000): Mitochondria and ageing. Curr Opin Clin Nutr Metab Care 3: 473-478.

Crane FL (2001): Biochemical functions of coenzyme $Q_{10}$. J Am Coll Nutr 20: 591-598.

Dalakas MC, Illa I, Pezeshkpour GH, Laukaitis JP, Cohen B and Griffin JL (1990): Mitochondrial myopathy caused by long-term zidovudine therapy. N Engl J Med 322: 1098-1105.

D'Aurelio M, Gajewski CD, Lin MT, Mauck WM, Shao LZ, Lenaz G, Moraes CT and Manfredi G (2004): Heterologous mitochondrial DNA recombination in human cells. Hum Mol Genet 13: 3171-3179.

Davis AF and Clayton DA (1996): In situ localization of mitochondrial DNA replication in intact mammalian cells. J Cell Biol 135: 883-893.

Davis LG, Dibner MF and Battey JF (1986): Basic methods in molecular biology, Elsevier Science.

Debyser Z, Tabor S and Richardson CC (1994): Coordination of leading and lagging strand DNA synthesis at the replication fork of bacteriophage T7. Cell 77: 157-166.

Del Bo R, Crimi M, Sciacco M, Malferrari G, Bordoni A, Napoli L, Prelle A, Biunno I, Moggio M, Bresolin N, Scarlato G and Pietro-Comi G (2003): High mutational burden in the mtDNA control region from aged muscles: a single-fiber study. Neurobiol Aging 24: 829-838.

Diaz F, Bayona-Bafaluy MP, Rana M, Mora M, Hao H and Moraes CT (2002): Human mitochondrial DNA with large deletions repopulates organelles faster than full-length genomes under relaxed copy number control. Nucleic Acids Res 30: 4626-4633.

Diez-Sanchez C, Ruiz-Pesini E, Lapena AC, Montoya J, Perez-Martos A, Enriquez JA and Lopez-Perez MJ (2003): Mitochondrial DNA content of human spermatozoa. Biol Reprod 68: 180-185.

Dimauro S and Davidzon G (2005): Mitochondrial DNA and disease. Ann Med 37: 222-232.

Dimmer KS, Jakobs S, Vogel F, Altmann K and Westermann B (2005): Mdm31 and Mdm32 are inner membrane proteins required for maintenance of mitochondrial shape and stability of mitochondrial DNA nucleoids in yeast. J Cell Biol 168: 103-115.

Donahue SL, Corner BE, Bordone L and Campbell C (2001): Mitochondrial DNA ligase function in Saccharomyces cerevisiae. Nucleic Acids Res 29: 1582-1589.

Doublie S, Tabor S, Long AM, Richardson CC and Ellenberger T (1998): Crystal structure of a bacteriophage T7 replication complex at 2.2 Å resolution. Nature 391: 251-258.

Dymecki SM (1996): Flp recombinase promotes site-specific DNA recombination in embryonic stem cells. Proc Natl Acad Sci USA 93: 6191-6196.

Egelman EH, Yu X, Wild R, Hingorani MM and Patel SS (1995): Bacteriophage T7 helicase/primase proteins form rings around single-stranded DNA that suggest a general structure for hexameric helicases. Proc Natl Acad Sci USA 92: 3869-3873.

Ekstrand MI, Falkenberg M, Rantanen A, Park CB, Gaspari M, Hultenby K, Rustin P, Gustafsson CM and Larsson NG (2004): Mitochondrial transcription factor A regulates mtDNA copy number in mammals. Hum Mol Genet 13: 935-944.

Elpeleg O, Miller C, Hershkovitz E, Bitner-Glindzicz M, Bondi-Rubinstein G, Rahman S, Pagnamenta A, Eshhar S and Saada A (2005): Deficiency of the ADP-forming succinyl-CoA synthase activity is associated with encephalomyopathy and mitochondrial DNA depletion. Am J Hum Genet 76: 1081-1066.

Elson JL, Samuels DC, Turnbull DM and Chinnery PF (2001): Random intracellular drift explains the clonal expansion of mitochondrial DNA mutations with age. Am J Hum Genet 68: 802-806.

Enriquez JA, Chomyn A and Attardi G (1995): MtDNA mutation in MERFF syndrome causes defective aminoacylation of tRNA(lys) and premature translation termination. Nat Genet 10: 47-55.

Eriksson S, Graslund A, Skog S, Thelander L and Tribukait B (1984): Cell cycle-dependent regulation of mammalian ribonucleotide reductase. The S phase-correlated increase in subunit M2 is regulated by de novo protein synthesis. J Biol Chem 259: 11695-11700.

Esposito LA, Melov S, Panov A, Cottrell BA and Wallace DC (1999): Mitochondrial disease in mouse results in increased oxidative stress. Proc Natl Acad Sci USA 96: 4820-4825.

Fairfield FR, Bauer WR and Simpson MV (1979): Mitochondria contain a distinct DNA topoisomerase. J Biol Chem 254: 9352-9354.

Falkenberg M, Gaspari M, Rantanen A, Trifunovic A, Larsson NG and Gustafsson CM (2002): Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA. Nat Genet 31: 289-294.

Farr CL, Wang Y and Kaguni LS (1999): Functional interactions of mitochondrial DNA polymerase and single-stranded DNA-binding protein. J Biol Chem 274: 14779-14785.

Farr CL, Matsushima Y, Lagina AT, Luo N and Kaguni LS (2004): Physiological and biochemical defects in functional interactions of mitochondrial DNA polymerase and DNA-binding mutants of single-stranded DNA-binding protein. J Biol Chem 279: 17047-17053.

Fernandez-silva P, Enriquez JA and Montoya J (2003): Replication and transcription of mammalian mitochondrial DNA. Exp Physiol 88: 41-56.

Ferraro P, Pontarin G, Crocco L, Fabris S, Reichard P and Bianchi V (2005): Mitochondrial deoxynucleotide pools in quiescent fibroblasts. J Biol Chem 280: 24472-24480.

Fisher RP, Lisowsky T, Parisi MA and Clayton DA (1992): DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein. J Biol Chem 267: 3358-3367.

Fliss MS, Usadel H, Caballero OL, Wu L, Buta MR, Eleff SM, Jen J and Sidranshy D (2000): Facile detection of mitochondrial DNA mutations in tumors and bodily fluids. Science 287: 2017-2019.

Fontanesi F, Palmieri L, Scarcia P, Lodi T, Donnini C, Limongelli A, Tiranti V, Zeviani M, Ferrero I and Viola AM (2004): Mutations in AAC2, equivalent to human adPEO-associated ANT1 mutations, lead to defective oxidative phosphorylation in Saccharomyces cerevisiae and affect mitochondrial DNA stability. Hum Mol Genet 13: 923-934.

Foury F and Lahaye A (1987): Cloning and sequencing of the PIF gene involved in repair and recombination of yeast mitochondrial DNA. EMBO J 6: 1441-1449.

Foury F (1989): Cloning and sequencing of the nuclear gene MIP1 encoding the catalytic subunit of the yeast mitochondrial DNA polymerase. J Biol Chem 264: 20552-20560.

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

Gago MF, Rosas MJ, Guimaraes J, Ferreira M, Vilarinho L, Castro L and Carpenter S (2006): SANDO: Two novel mutations in POLG1 gene. Neuromuscul Disord 16: 507-509.

Gaidamakov SA, Gorshkova II, Schuck P, Steinbach PJ, Yamada H, Crouch RJ and Cerritelli SM (2005): Eukaryotic RNases H1 act processively by interactions through the duplex RNA-binding domain. Nucleic Acids Res 33: 2166-2175.

Garrido N, Griparic L, Jokitalo E, Wartiovaara J, van der Bliek AM and Spelbrink JN (2003): Composition and dynamics of human mitochondrial nucleoids. Mol Bio Cell 14: 1583-1596.

Gaspari M, Larsson NG and Gustafsson CM (2004): The transcription machinery in mammalian mitochondria. Biochim Biophys Acta 1659: 148-152.

Genga A, Bianchi L and Foury F (1986): A nuclear mutant of *saccharomyces cerevisiae* deficient in mitochondrial DNA replication and polymerase activity. J Biol Chem 261: 9328-9332.

Genuario R and Wong TW (1993): Stimulation of DNA polymerase gamma by mitochondrial single-strand DNA binding protein. Cell Mol Biol Res 39: 625-634.

Geromel V, Kadhom N, Cebalos-Picot I, Ouari O, Polidori A, Munnich A, Rotig A and Rustin P (2001): Superoxide-induced massive apoptosis in cultured skin fibroblasts harboring the neurogenic ataxia retinitis pigmentosa (NARP) mutation in the ATPase-6 gene of the mitochondrial DNA. Hum Mol Genet 10: 1221-1228.

Ghivizzani SC, Madsen CS, Nelen MR, Ammini CV and Hauswirth WW (1994): In organello footprint analysis of human mitochondrial DNA: human mitochondrial transcription factor A interactions at the origin of replication. Mol Cell Biol 14: 7717-7730.

Gillham NW, Boynton JE and Hauser CR (1994): Translational regulation of gene expression in chloroplasts and mitochondria. Annu Rev Genet 28: 71-93.

Gordenin DA and Resnick MA (1998): Yeast ARMs (DNA at-risk moifs) can reveal sources of genome instability. Mutat Res 400: 45-58.

Goto A, Matsushima Y, Kadowaki T and Kitagawa Y (2001): Drosophila mitochondrial transcription factor A (d-TFAM) is dispensable for the transcription of mitochondrial DNA in Kc167 cells. Biochem J 354: 243-248.

Goto Y, Nonaka I and Horai S (1990): A mutation in the tRNA(Leu)(UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. Nature 348: 651-653.

Goto Y, Horai S, Matsuoka T, Koga Y, Nihei K, Kobayashi M and Nonaka I (1992): Mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes (MELAS): a correlative study of the clinical features and mitochondrial DNA mutation. Neurology 42: 545-550.

Goto Y (1995): Clinical features of MELAS and mitochondrial DNA mutations. Muscle Nerve 3: S107-S112.

Graziewicz MA, Longley MJ Bienstock RJ, Zeviani M and Copeland WC (2004): Structure-function defects of human mitochondrial DNA polymerase in autosomal dominant progressive external ophthalmoplegia. Nat Struct Mol Biol 11: 770-776.

Graziewicz MA, Longley MJ and Copeland WC (2006): DNA polymerase γ in mitochondrial DNA replication and repair. Chem Rev 106: 383-405.

Green DR and Reed JC (1998): Mitochondria and apoptosis. Science 281: 1309-1312.

Gunter TE, Yule DI, Gunter KK, Eliseev RA and Salter JD (2004): Calcium and mitochondria. FEBS Lett 567: 96-102.

Guo S, Tabor S and Richardson CC (1999): The linker region between the helicase and primase domains of the bacteriophage T7 gene 4 protein is critical for hexamer formation. J Biol Chem 274: 30303-30309.

Hadi MZ, Coleman MA, Fidelis K, Mohrenweiser HW and Wilson DM 3rd (2000): Functional characterization of Ape1 variants identified in the human population. Nucleic Acids Res 28: 3871-3879.

Hamilton ML, Van Remmen H, Drake JA, Yang H, Guo ZM, Kewitt K, Walter CA and Richardson A (2001): Does oxidative damage to DNA increase with age? Proc Natl Acad Sci USA 98: 10469-10474.

Hammans SR, Sweeney MG, Brockington M, Lenox GG, Lawton NF, Kennedy CR, Morgan-Hughes JA and Harding AE (1993): The mitochondrial DNA transfer RNA(Lys)A→G(8344) mutation and the syndrome of myoclonic epilepsy with ragged red fibres (MERFF). Relationship of clinical phenotype to proportion of mutant mitochondrial DNA. Brain 116: 617-632.

Harman D (1956): Aging: a theory based on free radical and radiation chemistry. J Gerontol 11: 298-300.

Harris TP, Gomas KP, Weir F, Holyoake AJ, McHugh P, Wu M, Sin Y, Sin IL and Sin FY (2006): Molecular analysis of polymerase gamma gene and mitochondrial polymorphism in fertile and subfertile men. Int J Androl 29: 421-433.

Hayashi J, Ohta S, Kikuchi A, Takemitsu M, Goto Y and Nonaka I (1991): Introduction of disease-related mitochondrial DNA deletions into HeLa cells lacking mitochondrial DNA results in mitochondrial dysfunction. Proc Natl Acad Sci USA 88: 10614-10618.

Hayashi J, Takemitsu M, Goto Y and Nonaka I (1994): Human mitochondria and mitochondrial genome function as a single dynamic cellular unit. J Cell Biol 125: 43-50.

He L, Luo L, Proctor SJ, Middleton PG, Blakely EL, Taylor RW and Turnbull RW (2003): Somatic mitochondrial DNA mutations in adult-onset leukaemia. Leukemia 17: 2487-2491.

Helm M, Florentz C, Chomyn A and Attardi G (1999): Search for differences in post-transcriptional modification patterns of mitochondrial DNA-encoded wild-type and mutant human tRNA$^{lys}$ and tRNA $^{Leu(UUR)}$. Nucleic Acids Res 27: 756-763.

Hirano M, Silvestri G, Blake DM, Lombes A, Minetti C, Bonilla E, Hays AP, Lovelace RE, Butler I and Bertorini TE (1994): Mitochondrial neurogastrointestinal encephalomyopathy (MNGIE): clinical, biochemical, and genetic features of an autosomal recessive mitochondrial disorder. Neurology 44: 721-727.

Hobbs AE, Srinivasan M, McCaffery JM and Jensen RE (2001): Mmm1p, a mitochondrial outer membrane protein, is connected to mitochondrial DNA (mtDNA) nucleoids and required for mtDNA stability. J Cell Biol 152: 401-410.

Holt IJ, Harding AE and Morgan-Hughes JA (1988): Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. Nature 331: 717-719.

Holt IJ, Harding AE, Petty RK and Morgan-Hughes JA (1990): A new mitochondrial disease associated with mitochondrial DNA heteroplasmy. Am J Hum Genet 46: 428-433.

Holt IJ, Lorimer HE and Jacobs HT (2000): Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA. Cell 100: 515-524.

Holt IJ and Jacobs HT (2003): Response: the mitochondrial DNA replication bubble has not burst. Trends Biochem Sci 28: 355-356.

Howell N, Oostra RJ, Bolhuis PA, Spruijt L, Clarke LA, Mackey DA, Preston G and Herrnstadt C (2003): Sequence analysis of the mitochondrial genomes from dutch pedigrees with Leber hereditary optic neuropathy. Am J Hum Genet 72: 1460-1469.

Huang M and Elledge SJ (1997): identification of RNR4, encoding a second essential small subunit of ribonucleotide reductase in Saccharomyces cerevisiae. Mol Cell Biol 17: 6105-6113.

Hudson G, Deschauer M, Busse K, Zierz S and Chinnery PF (2005): Sensory ataxic neuropathy due to a novel C10Orf2 mutation with probable germline mosaicism. Neurology 64: 371-373.

Huoponen K, Vilkki J, Aula P, Nikoskelainen EK and Savontaus ML (1991): A new DNA mutation associated with Leber hereditary optic neuropathy. Am J Hum Genet 48: 1147-1153.

Iborra FJ, Kimura H and Cook PR (2004): The functional organization of mitochondrial genomes in human cells. BMC Biol 2:9.

Insdorf NF and Bogenhagen DF (1989): DNA polymerse γ from *xenopus laevis*. J Biol Chem 264: 21491-21497.

Ivessa AS, Lenzmeier BA, Bessler JB, Goudsouzian LK, Schnakenberg SL and Zakian VA (2003): The Saccharomyces cerevisiae helicase Rrm3p facilitates replication past nonhistone protein-DNA complexes. Mol Cell: 1525-1536.

Izumi T, Tatsuka M, Tano K, Asano M and Mitra S (1997): Molecular cloning and characterization of the promoter of the human N-methylpurine-DNA glycosylase (MPG) gene. Carcinogenesis 18: 1837-1839.

Jacobs HT and Turnbull DM (2005): Nuclear genes and mitochondrial translation: a new class of genetic disease. Trends Genet 21: 312-314.

Janssen GMC, Maassen JA and van den Ouweland JMW (1999): The diabetes-associated 3243 mutation in the mitochondrial tRNA[Leu(UUR)] gene causes severe mitochondrial dysfunction without a strong decrease in protein synthesis rate. J Biol Chem 274: 29744-29748.

Janssen RJRJ, Nijtmans LG, van den Heuvel LP and Smeitink JAM (2006): Mitochondrial complex I: structure, function and pathology. J Inherit Metab Dis 29: 499-515.

Jazin EE, Cavelier L, Eriksson I, Oreland L and Gyllensten U (1996): Human brain contains high levels of heteroplasmy in the noncoding regions of mitochondrial DNA. Proc Natl Acad Sci USA 93: 12382-12387.

Johansson M and Karlsson A (1996): Cloning and expression of human deoxyguanosine kinase cDNA. Proc Natl Acad Sci USA 93: 7258-7262.

Johansson M and Karlsson A (1997): Cloning of the cDNA and chromosome localization of the gene for human thymidine kinase 2. J Biol Chem 272: 8454-8458.

Johns DR, Neufeld MJ and Park RD (1992): An ND-6 mitochondrial DNA mutation associated with Leber hereditary optic neuropathy. Biochem Biophys Res Commun 187: 1551-1557.

Kaguni LS (2004): DNA polymerase γ, the mitochondrial replicase. Annu Rev Biochem 73: 293-320.

Kai Y, Miyako K, Muta T, Umeda S, Irie T, Hamasaki N, Takeshige K and Kang D (1999): Mitochondrial DNA replication in human T lymphocytes is regulated primarily at the H-strand termination site. Biochem Biophys Acta 1446: 126-134.

Kajander OA, Karhunen PJ, Holt IJ and Jacobs HT (2001): Prominent mitochondrial DNA recombination intermediates in human heart muscle. EMBO Rep 2: 1007-1012.

Kang D and Hamasaki N (2005): Mitochondrial transcription factor A in the maintenance of mitochondrial DNA. Ann N Y Acad Sci 1042: 101-108.

Kao SH, Chao HT and Wei YH (1998): Multiple deletions of mitochondrial DNA are associated with the decline of motility and fertility of human spermatozoa. Mol Hum Reprod 4: 657-666.

Karadimas CL, Vu TH, Holve SA, Chronopoulou P, Quinzii C, Johnsen SD, Kurth J, Eggers E, Palenzuela L, Tanji K, Bonilla E, De Vivo DC, DiMauro S and Hirano M (2006): Navajo neurohepatopathy is caused by a mutation in the MPV17 gene. Am J Hum Genet 79: 544-548.

Kaufman BA, Newman SM, Hallberg RL, Slaughter CA, Perlman PS and Butow RA (2000): In organello formaldehyde crosslinking of proteins to mtDNA: identification of bifunctional proteins. Proc Natl Acad Sci USA 97: 7772-7777.

Kaukonen JA, Amati P, Suomalainen A, Rotig A, Piscaglia MG, Salvi F, Weissenbach J, Fratta G, Comi G, Peltonen L and Zeviani M (1996): An autosomal locus predisposing to multiple deletions of mtDNA on chromosome 3p. Am J Hum Genet 58: 763-769.

Kaukonen JA, Zeviani M, Comi P, Piscaglia MG, Peltonen L and Suomalainen A (1999): A third locus predisposing to multiple deletions of mtDNA in autosomal dominant progressive external ophthalmoplegia. Am J Hum Genet 65: 256-261.

Kaukonen JA, Juselius JK, Tiranti V, Kyttala A, Zeviani M, Comi GP, Keranen S, Peltonen L and Suomalainen A (2000): Role of adenine nucleotide translocator 1 in mtDNA maintenance. Science 289: 782-785.

Kellner U, Sehested M, Jensen PB, Gieseler F and Rudolph P (2002): Culprit and victim – DNA topoisomerase II. Lancet Oncol 3: 235-243.

Khrapko K, Coller HA, Andre PC, Li XC, Hanekamp JS and Thilly WG (1997): Mitochondrial mutational spectra in human cells and tissues. Proc Natl Acad Sci USA 94: 13798-13803.

Kiechl S, Horvath R, Luoma P, Kiechl-Kohlendorfer U, Wallacher-Scholz, Stucka R, Thaler C, Wanschitz J, Suomalainen A, Jaksch M and Willeit J (2004): Two families with autosomal dominant progressive external ophthalmoplegia. J Neurol Neurosurg Psychiatry 75: 1125-1128.

Koehler CM (2004): New developments in mitochondrial assembly. Annu Rev Cell Dev Biol 20: 309-335.

Korhonen JA, Gaspari M and Falkenberg M (2003): TWINKLE has 5' → 3' DNA helicase activity and is specifically stimulated by mitochondrial single-stranded binding protein. J Biol Chem 278: 48627-48632.

Korhonen JA, Pham XH, Pellegrini M and Falkenberg M (2004): Reconstruction of a minimal mtDNA replisome *in vitro*. EMBO J 23: 2423-2429.

Kowluru A, Tannous M and Chen HQ (2002): Localization and characterization of the mitochondrial isoform of the nucleoside diphosphate kinase in the pancreatic beta cell: evidence for its complexation with mitochondrial succinyl-CoA synthetase. Arch Biochem Biophys 398: 160-169.

Kraytsberg Y, Kudryavtseva E, McKee AC, Geula C, Kowall NW and Khrapko K (2006): Mitochondrial DNA deletions are abundant and cause functional impairment in aged human substantia nigra neurons. Nat Genet 38: 518-520.

Kujoth GC, Hiona A, Pugh TD, Someya S, Panzer K, Wohlgemuth SE, Hofer T, Seo AY, Sullivan R, Jobling WA, Morrow JD, Van Remmen H, Sedivy JM, Yamasoba T, Tanokura M, Weindruch R, Leeuwenburgh C and Prolla TA (2005): Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. Science 309: 481-484.

Kunkel TA and Mosbaugh DW (1989): Exonucleolytic proofreading by a mammalian DNA polymerase. Biochemistry 28: 988-995.

Kunkel TA, Patel SS and Johnson KA (1994): Error-prone replication of repeated DNA sequences by T7 DNA polymerase in the absence of its processivity subunit. Proc Natl Acad Sci USA 91: 6830-6834.

Kunz BA (1988): Mutagenesis and deoxyribonucleotide pool imbalance. Mutat Res 200: 133-147.

Kuroiwa T, Kawano S and Hizume M (1977): Studies on mitochondrial structure and function in Physarum polycephalum. V. Behaviour of mitochondrial nucleoids throughout mitochondrial division cycle. J Cell Biol 72: 687-694.

Laforet P, Lombes A, Eymard B, Danan C, Chevallay M, Rouche A, Frachon P and Fardeau M (1995): Chronic progressive external ophthalmoplegia with ragged-red fibres: clinical, morphological and genetic investigations in 43 patients. Neuromuscul Disord 5: 399-413.

Lakshmipathy U and Campbell C (1999a): Double strand break rejoining by mammalian mitochondrial extracts. Nucleic Acids Res 27: 1198-1204.

Lakshmipathy U and Campbell C (1999b): The human DNA ligase III gene encodes nuclear and mitochondrial proteins. Mol Cell Biol 19: 3869-3876.

Lakshmipathy U and Campbell C (2000): Mitochondrial DNA ligase III function is independent of Xrcc1. Nucleic Acids Res 28: 3880-3886.

Lakshmipathy U and Campbell C (2001): Antisense-mediated decrease in DNA ligase III expression results in reduced mitochondrial DNA integrity. Nucleic Acids Res 29: 668-676.

Larsson NG, Holme E, Kristiansson B, Oldfors A and Tulinius M (1990): Progressive increase of the mutated mitochondrial DNA fraction in Kearns-Sayre Syndrome. Pediatr Res 28: 131-136.

Larsson NG and Clayton DA (1995): Molecular genetic aspects of human mitochondrial disorders. Annu Rev Genet 29: 151-178.

Larsson NG, Wang J, Wilhelmsson H, Oldfros A, Rustin P, Lewandoski M, Barsh GS and Clayton DA (1998): Mitochondrial transcription factor A is necessary for mtDNA maintenance and embryogenesis in mice. Nat Genet 18: 231-236.

Lecrenier N and Foury F (1995): Overexpression of the RNR1 gene rescues *Saccharomyces cerevisiae* mutants in the mitochondrial DNA polymerase-encoding MIP1 gene. Mol Gen Genet 249: 1-7.

Lecrenier N, van de Bruggen P and Foury F (1997): Mitochondrial DNA polymerases from yeast to man: a new family of polymerases. Gene 185: 147-152.

Lecrenier N and Foury F (2000): New features of mitochondrial DNA replication system in yeast and man. Gene 246: 37-48.

LeDoux SP, Wilson GL, Beecham EJ, Stevnsner T, Wasserman K and Bohr VA (1992): Repair of mitochondrial DNA after various types of DNA damage in Chinese hamster ovary cells. Carcinogenesis 13: 1967-1973.

Lee DY and Clayton DA (1997): RNase mitochondrial RNA processing correctly cleaves a novel R loop at the mitochondrial DNA leading-strand origin of replication. Genes Dev 11: 582-592.

Legros F, Malka F, Frachon P, Lombes A and Rojo M (2004): Organization and dynamics of human mitochondrial DNA. J Cell Sci 117: 2653-2662.

Levin CJ and Zimmerman SB (1976): A DNA ligase from mitochondria of rat liver. Biochem Biophys Res Commun 69: 514-520.

Lewis S, Hutchison W, Thyagarajan D and Dahl HH (2002): Clinical and molecular features of adPEO due to mutations in the Twinkle gene. J Neurol Sci 201: 39-44.

Lewis W, Day BJ, Kohler JJ, Hosseini SH, Chan SS, Green EC, Haase CP, Keebaugh ES, Long R, Ludaway T, Russ R, Steltzer J, Tioleco N, Santoianni R and Copeland WC (2007): Decreased mtDNA, oxidative stress, cadiomyopathy, and death from transgenic cardiac targeted human mutant polymerase gamma. Lab Invest 87: 326-335.

Li FY, Tariq M, Croxen R, Morten K, Squier W, Newson-Davis J, Beeson D and Larsson C (1999): Mapping of autosomal dominant preogressive external ophthalmoplegia to a 7-cM critical region on 10q24. Neurology 53: 1265-1271.

Li K, Warner CK, Hodge JA, Minoshima S, Kudoh J, Fukuyama R, Maekawa M, Shimizu Y, Shimizu N and Wallace DC (1989): A human muscle adenine nucleotide translocator gene has four exons, is located on chromosome 4, and is differentially expressed. J Biol Chem 264: 13998-14004.

Li K and Williams RS (1997): Tetramerization and single-stranded DNA binding properties of native and mutated forms of murine mitochondrial single-stranded DNA-binding proteins. J Biol Chem 272: 8686-8694.

Lim SE, Longley MJ and Copeland WC (1999): The mitochondrial p55 accessory subunit of human DNA polymerase γ enhances DNA binding, promotes processive DNA synthesis, and confers N-ethylmaleimide resistance. J Biol Chem 274: 38197-38203.

Linder T, Park CB, Asin-Cayuela J, Pellegrini M, Larsson NG, Falkenberg M, Samuelsson T and Gustafsson CM (2005): A family of putative transcription termination factors shared amongst metazoans and plants. Curr Genet 48: 265-269.

Longley MJ, Prasad R, Srivastava DK, Wilson SH and Copeland WC (1998a): Identification of 5' deoxyribose phosphate lyase activity in human DNA polymerase γ and its role in mitochondrial base excision repair *in vitro*. Proc Natl Acad Sci USA 95: 12244-12248.

Longley MJ, Ropp PA, Lim SE and Copeland WC (1998b): Characterization of the native and recombinant catalytic subunit of human DNA polymerase γ: identification of residues critical for exonuclease activity and dideoxynucleotide sensitivity. Biochemistry 37: 10529-10539.

Longley MJ, Nguyen D, Kunkel TA and Copeland WC (2001): The fidelity of human DNA polymerase γ with and without exonucleolytic proofreading and the p55 accessory subunit. J Biol Chem 276: 38555-38562.

Longley MJ, Clark S, Yu Wai Man C, Hudson G, Durham SE, Taylor RW, Nightingale S, Turnbull DM, Copeland WC and Chinnery PF (2006): Mutant POLG2 disrupts DNA polymerase γ subunits and causes progressive external ophthalmoplegia. Am J Hum Genet 78: 1026-1034.

Low RL, Orton S and Friedman DB (2003): A truncated form of DNA topoisomerase IIβ associates with the mtDNA genome in mammalian mitochondria. Eur J Biochem 270: 4173-4186.

Luo N and Kaguni LS (2005): Mutations in the spacer region of *drosophila* mitochondrial DNA polymerase affect DNA binding, processivity, and the balance between pol and exo function. J Biol Chem 280: 2491-2497.

Luoma P, Melberg A, Rinne JO, Kaukonen JA, Nupponen NN, Chalmers RM, Oldfors A, Rautakorpi I, Peltonen L, Majamaa K, Somer H and Suomalainen A (2004): Parkinsonism, premature menopause, and mitochondrial DNA polymerase gamma mutations: clinical and molecular genetic study. Lancet 364: 875-882.

Luoma P, Luo N, Loscher WN, Farr CL, Horvath R, Wanschitz J, Kiechl S, Kaguni LS and Suomalainen A (2005): Functional defects due to spacer-region mutations of human mitochondrial DNA polymerase in a family with an ataxia-myopathy syndrome. Hum Mol Genet 14: 1907-1920.

Madsen CS, Ghivizzani SC and Hauswirth WW (1993): Protein binding to a single termination-associated sequence in the mitochondrial DNA D-loop region. Mol Cell Biol 13: 2162-2171.

de Magalhaes JP (2005): Open-minded scepticism: inferring the causal mechanisms of human ageing from genetic perturbations. Ageing Res Rev 4: 1-22.

Magnusson J, Orth M, Lestienne P and Taanman JW (2003): Replication of mitochondrial DNA occurs throughout the mitochondria of cultured human cells. Exp Cell Res 289: 133-142.

Makela-Bengs P, Suomalainen A, Majander A, Rapola J, Kalimo H, Nuutil A and Pihko H (1995): Correlation between the clinical symptoms and the proportion pf mitochondrial DNA carrying the 8893 point mutation in the NARP syndrome. Pediatr Res 37: 634-639.

Mambo E, Gao X, Cohen Y, Guo Z, Talalay P and Sidransky D (2003): Electrophile and oxidant damage of mitochondrial DNA leading to rapid evolution of homoplasmic mutations. Proc Natl Acad Sci USA 100: 1838-1843.

Mandel H, Szargel R, Labay V, Elpeleg O, Saada A, Shalata A, Anbinder Y, Berkowitz D, Hartman C, Barak M, Eriksson S and Cohen N (2001): The deoxyguanosine kinase gene is mutated in individuals with depleted hepatocerebral mitochondrial DNA. Nat Genet 29: 337-341.

Manfredi G (2006): MtDNA clock runs out for dopaminergic neurons. Nat Genet 38: 507-508.

Maniura-Weber K, Goffart S, Garstka HL, Montoya J and Wiesner RJ (2004): Transient overexpression of mitochondrial transcription factor A (TFAM) is sufficient to stimulate mitochondrial DNA transcription but not sufficient to increase mtDNA copy number in cultured cells. Nucleic Acids Res 32: 6015-6027.

Mannella CA (2006): The relevance of mitochondrial membrane topology to mitochondrial function. Biochim Biophys Acta 1762: 140-147.

Marti R, Spinazzola A, Nishino I, Andreu AL, Naini A, Tadesse S, Olivier JA and Hirano M (2002): Mitochondrial neurogastrointestinal encephalomyopathy and thymidine metabolism: results and hypotheses. Mitochondria 2: 143-147.

Martin M, Cho J, Cesare AJ, Griffith JD and Attardi G (2005): Termination factor-mediated DNA loop between termination and initiation sites drives mitochondrial rRNA synthesis. Cell 123: 1127-1240.

Mason PA, Matheson EC, Hall AG and Lightowlers RN (2003): Mismatch repair activity in mammalian mitochondria. Nucleic Acids Res 31: 1052-1058.

Matsushima Y, Garesse R and Kaguni LS (2004): Drosophila mitochondrial transcription factor B2 regulates mitochondrial DNA copy number and transcription in schneider cells. J Biol Chem 279: 26900-26905.

Mattiazzi M, Vijayvergiya C, Gajewksi CD, DeVivo DC, Lenaz G, Wiedmann M and Manfredi G (2004): The mtDNA T8993G (NARP) mutation results in an impairment of oxidative phosphorylation that can be improved by antioxidants. Hum Mol Genet 13: 869879.

McCulloch V and Shadel GS (2003): Human mitochondrial transcription factor B1 interacts with the C-terminal activation region of h-mtTFA and stimulates transcription independently of its RNA methyltransferase activity. Mol Cell Biol 23: 5816-5824.

Mcfarland R, Taylor RW and Turnbull DM (2002): The neurology of mitochondrial DNA disease. Lancet Neurol 1: 343-351.

McGlynn P, Lloyd RG and Marians KJ (2001): Formation of Holliday junctions by regression of nascent DNA in intermediates containing stalled replication forks: RecG stimulates regression even when the DNA is negatively supercoiled. Proc Natl Acad Sci USA 98: 8235-8240.

Meeusen S and Nunnari J (2003): Evidence for a two membrane-spanning autonomous mitochondrial DNA replisome. J Cell Biol 163: 503-510.

Melberg A, Arnell H, Dahl N, Stalberg E, Raininko R, Oldfors A, Bakall B, Lundberg PO and Holme E (1996): Anticipation of autosomal dominant progressive external ophthalmoplegia with hypogonadism. Muscle Nerve 19: 1561-1569.

Michel B, Ehrlich SD and Uzest M (1997): DNA double-strand breaks caused by replication arrest. EMBO J 16: 430-438.

Michikawa Y, Mazzucchelli F, Bresolin N, Scarlato G and Attardi G (1999): Aging-dependent large accumulation of point mutations in the human mtDNA control region for replication. Science 286: 774-779.

Mignotte B, Marsault J and Barat-Gueride M (1988): Effects of the Xenopus Laevis mitochondrial single-stranded DNA-binding protein on the activity of DNA polymerase gamma. Eur J Biochem 174: 479-484.

Mileykovskaya E, Zhang M and Dowhan W (2005): Cardiolipin in energy transducing membranes. Biochemistry 70: 154-158.

Mita S, Rizzuto R, Moraes CT, Shanske S, Arnaudo E, Fabrizi GM, Koga Y, DiMauro S and Schon EA (1990): Recombination via flanking direct repeats is a major cause of large-scale deletions of human mitochondrial DNA. Nucleic Acids Res 18: 561-567.

Miyakawa I, Aoi H, Sando N and Kuroiwa T (1984): Fluorescence microscopic studies of mitochondrial nucleoids during meiosis and sporulation in the yeast *saccharomyces cerevisiae*. J Cell Sci 66: 21-38.

Miyakawa I, Sando N, Kawano S, Nakamura S and Kuroiwa T (1987): Isolation of morphologically intact mitochondrial nucleoids from the yeast, *saccharomyces cerevisia*. J Cell Sci 88: 431-439.

Montoya J, Christianson T, Levens D, Rabinowitz M and Attardi G (1982): Identification of initiation site for heavy-strand and light-strand transcription in human mitochondrial DNA. Proc Natl Acad Sci USA 79: 7195-7199.

Montoya J, Gaines GL and Attardi G (1983): The pattern of transcription of the human mitochondrial rRNA genes reveals two overlapping transcription units. Cell 34: 151-159.

Moraes CT, DiMauro S, Zeviani M, Lombes A, Shanske S, Miranda AF, Nakase H, Bonilla E, Werneck LC and Servidei S (1989): Mitochondrial DNA deletions in progressive external opthalmoplegia and Kearns-Sayre syndrome. N Engl J Med 320: 1293-1299.

Moraes CT, Shanske S, Trischler HJ, Aprille JR, Andreetta F, Bonilla E, Schon EA and DiMauro S (1991): MtDNA depletion with variable tissue expression: a novel genetic abnormality in mitochondrial diseases. Am J Hum Genet 48: 492-501.

Moraes CT (2001): What regulates mitochondrial DNA copy numbr in animal cells? Trends Genet 17: 199-205.

Moraes CT, Diaz F and Barrientos A (2004): Defects in the biosynthesis of mitochondrial heme *c* and heme *a* in yeast and mammals. Biochim Biophys Acta 1659: 153-159.

Morgan-Hughes JA, Sweeney MG, Cooper JM, Hammans SR, Brockington M, Schapira AH, Harding AE and Clark JB (1995): Mitochondrial DNA (mtDNA) diseases: correlation of genotype to phenotype. Biochim Biophys Acta 1271: 135-140.

Moslemi AR, Melberg A, Holme E and Oldfors A (1996): Clonal expansion of mitochondrial DNA with multiple deletions in autosomal dominant progressive external ophthalmoplegia. Ann Neurol 40: 707-713.

Murakami E, Feng JY, Lee H, Hanes J, Johnson KA and Anderson KS (2003): Characterization of novel reverse transcriptase and other RNA-associated catalytic activities by human DNA polymerase γ. J Biol Chem 278: 36403-36409.

Murdock DG, Christacos NC and Wallace DC (2000): The age-related accumulation of a mitochondrial DNA control region mutation in muscle, but not brain, detected by a sensitive PNA-directed PCR clamping based method. Nucleic Acids Res 28: 4350-4355.

Myers KA, Saffhill R and O'Connor PJ (1988): Repair of alkylated purines in the hepatic DNA of mitochondria and nuclei in the rat. Carcinogenesis 9: 283-292.

Nass MM and Nass S (1963): Intramitochondrial fibers with DNA characteristics. I. Fixation and electron staining reactions. J Cell Biol 19: 593-611.

Nass MM (1969): Mitochondrial DNA.I. Intramitochondrial distribution and structural relations of single- and double-length circular DNA. J Mol Biol 42: 521-528.

Nassogne MC, Heron B, Touati G, Rabier D and Saudubray JM (2005): Urea cycle defects: management and outcome. J Inherit Metab Dis 28: 401-414.

Naviaux RK, Nyhan WL, Barshop BA, Poulton J, Markusic D, Karpinski NC and Haas RH (1999): Mitochondrial DNA polymerase gamma deficiency and mtDNA depletion in a child with Alpers' syndrome. Ann Neurol 45: 54-58.

Nekhaeva E, Bodyak ND, Kraytsberg Y, McGrath SB, Van Orsouw NJ, Pluzhnikov A, Wei JY, Vijg J and Khrapko K (2002): Clonally expanded mtDNA point mutations are abundant in individual cells of human tissues. Proc Natl Acad Sci USA 99: 5521-5526.

Nelson DL, Cox MM and Lehninger (2000): Lehninger principle of biochemistry 3rd edition, Worth Publishers Inc.

Nguyen KV, Ostergaard E, Ravn SH, Balslev T, Danielsen RV, Vardag A, McKiernan PJ, Gray G and Naviaux RK (2005): PolG mutations in Alpers syndrome. Neurology 65: 1493-1495.

Nijtmans LG, Henderson NS, Attardi G and Holt IJ (2001): Impaired ATP synthase assembly associated with a mutation in the human ATP synthase subunit 6 gene. J Biol Chem 276: 6755-6762.

Nikali K, Suomalainen A, Saharinen J, Kuokkanen M, Spelbrink JN, Lonnqvist and Peltonen L (2005): Infantile onset spinocerebellar ataxia is caused by recessive mutations in mitochondrial proteins Twinkle and Twinky. Hum Mol Genet 14: 2981-2990.

Nikoskelainen EK, Marttila RJ, Huoponen K, Juvonen V, Lamminen T, Sonninen P and Savontaus ML (1995): Leber's "plus": neurological abnormalities in patients with Leber's hereditary optic neuropathy. J Neurol Neurosurg Psychiatry 59: 160-164.

Nishigaki Y, Marti R, Copeland WC and Hirano M (2003): Site-specific somatic mitochondrial DNA point mutations in patients with thymidine phosphorylase deficiency. J Clin Invest 111: 1913-1921.

Nishino I, Spinazzola A and Hirano M (1999): Thymidine phosphorylase gene mutations in MNGIE, a human mitochondrial disorder. Science 283: 689-692.

Nishioka K, Ohtsubo T, Oda H, Fujiwara T, Kang D, Sugimachi K and Nakabeppu Y (1999): Expression and differential intercellular localization of two major forms of human 8-oxoguanine DNA glycosylase encoded by alternatively spliced OGG1 mRNAs. Mol Biol Cell 10: 1637-1652.

Ohno T, Umeda S, Hamasaki N and Kang D (2000): Binding of human mitochondrial transcription factor A, an HMG box protein, to a four-way DNA junction. Biochem Biophys Res Commun 271: 492-498.

Okamoto K and Shaw JM (2005): Mitochondrial morphology and dynamics in yeast and multicellular eukaryotes. Annu Rev Genet 39: 503-536.

Olichon A, Guillou E, Delettre C, Landes T, Arnaune-Pelloquin L, Emorine LJ, Mils V, Daloyau M, Hamel C, Amati-Bonneau P, Bonneau D, Reynier P, Lenaers G and Belenguer P (2006): Mitochondrial dynamics and disease, OPA1. Biochim Biophys Acta 1763: 500-509.

O'Rourke TW, Doudican NA, Zhang H, Eaton JS, Doetsch PW and Shadel GS (2005): Differential involvement of the related DNA helicases Pif1p and Rrm3p in mtDNA point mutagenesis and stability. Gene 354: 86-92.

van den Ouweland JM, Lemkes HH, Gerbitz KD and Maassen JA (1995): Maternally inherited diabetes and deafness (MIDD): a distinct subtype of diabetes with a mitochondrial tRNA(Leu)(UUR) gene point mutation. Muscle Nerve 3: S142-S130.

Parisi MA and Clayton DA (1991): Similarity of human mitochondrial transcription factor 1 to high mobility group proteins. Science 252: 965-969.

Patel SS, Hingorani MM and Ng WM (1994): The K318A mutant of bateriophage T7 DNA primase-helicase protein is deficient in helicase but not primase activity and inhibits primase-helicase protein wild-type activities by heterooligomer formation. Biochemistry 33: 7857-7868.

Patel SS and Picha KM (2000): Structure and function of hexameric helicases. Annu Rev Biochem 69: 651-697.

Pavco PA and van Tuyle GC (1985): Purification and general properties of the DNA-binding protein (p16) from rat liver mitochondria. J Cell Biol 100: 258-264.

Perez-Jannotti RM, Klein SM and Bogenhagen DF (2001): Two forms of mitochondrial DNA ligase III are produced in *Xenopus laevis* oocytes. J Biol Chem 276: 48978-48987.

Pettepher CC, LeDoux SP, Bohr VA and Wilson GL (1991): Repair of alkali-labile sites within the mitochondrial DNA of RINr 38 cells after exposure to the nitrosourea streptozotocin. J Biol Chem 266: 3113-3117.

Pham XH, Farge G, Shi Y, Gaspari M, Gustafsson CM and Falkenberg M (2006): Conserved sequence box II directs transcription termination and primer formation in mitochondria. J Biol Chem 281: 24647-24652.

Pinz KG and Bogenhagen DF (1998): Efficient repair of abasic sites in DNA by mitochondrial enzymes. Mol Cell Biol 18: 1257-1265.

Pinz KG and Bogenhagen DF (2006): The influence of the DNA polymerase γ accessory subunit on base excision repair by the catalytic subunit. DNA Repair 5: 121-128.

Pohjoismäki JL, Wanrooij S, Hyvärinen AK, Goffart S, Holt IJ, Spelbrink JN and Jacobs HT (2006): Alterations to the expression level of mitochondrial transcription factor A, TFAM, modify the mode of mitochondrial DNA replication in cultured human cells. Nucleic Acids Res 34: 5815-5828.

Polyak K, Li Y, Zhu H, Lengauer C, Willson JK, Markowitz SD, Trush MA, Kinzler KW and Vogelstein B (1998): Somatic mutations of the mitochondrial genome in human colorectal tumours. Nat Genet 20: 291-293.

Ponamarev MV, Longley MJ, Nguyen D, Kunkel TA and Copeland WC (2002): Active site mutation in DNA polymerase gamma associated with progressive external ophthalmoplegia causes error-prone DNA synthesis. J Biol Chem 277: 15225-15228.

Poulton J, Morten KJ, Weber K, Brown GK and Bindoff L (1994): Are duplications of mitochondrial DNA characteristic of Kearns-Sayre syndrome? Hum Mol Genet 3: 947-951.

Rehling P, Brandner K and Pfanner N (2004): mitochondrial import and twin-pore translocase. Nat Rev Mol Cell Biol 5: 519-530.

Reyes A, Yang MY, Bowmaker M and Holt IJ (2005): Bidirectional replication initiates at sites throughout the mitochondrial genome of birds. J Biol Chem 280: 3242-3250.

Reynier P and Malthiery Y (1995): Accumulation of deletions in mtDNA during tissue aging: analysis by long PCR. Biochem Biophys Res Commun 217: 59-67.

Reynier P, Chretien MF, Savagner F, Larcher G, Rohmer V, Barriere P and Malthiery Y (1998): Long PCR analysis of human gamete mtDNA suggests defective mitochondrial maintenance in spermatozoa and supports the bottleneck theory for oocytes. Biochem Biophys Res Commun 252: 373-377.

Rice P, Longden I and Bleasby A (2000): EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet 16: 276-277.

Richardson CC (1983): Bacteriophage T7: minimal requirements for the replication of a duplex DNA molecule. Cell 33: 315-317.

Robberson DL, Kasamatsu H and Vinograd J (1972): Replication of mitochondrial DNA. Circular replicative intermediates in mouse L cells. Proc Natl Acad Sci USA 69: 737-741.

Ropp PA and Copeland WC (1996): Cloning and characterization of the human mitochondrial DNA polymerase, DNA polymerase γ. Genomics 36: 449-458.

Rotig A, Colonna M, Bonnefont JP, Blanche S, Fischer A, Saudubray JM and Munnich A (1989): Mitochondrial DNA deletion in Pearson's marrow/pancrease syndrome. Lancet 1: 902-903.

Rotig A, Bessis JL, Romero N, Cormier V, Saudubray JM, Narcy P, Lenoir G, Rustin P and Munnich A (1992): Maternally inherited duplication of the mitochondrial genome in a syndrome of proximal tubulopathy, diabetes mellitus, and cerebellar ataxia. Am J Hum Genet 50: 364-370.

Rovio AT, Marchington DR, Donat S, Schuppe HC, Abel J, Fritsche E, Elliott DJ, Laippala P, Ahola AL, McNay D, Harrison RF, Hughes B, Barrett T, Bailey DMD, Mehmet D, Jequier AM, Hargreave TB, Kao SH, Cummins JM, Barton DE, Cooke HJ, Wei YH, Wichmann L, Poulton J and Jacobs HT (2001): Mutations at the mitochondrial DNA polymerase (POLG) locus associated with male infertility. Nat Genet 29: 261-262.

Rovio A (2006): DNA polymerase gamma mutations in male infertility and ageing. PhD thesis, University of Tampere, Finland

Rumbaugh JA, Henricksen LA, DeMott MS and Bambara RA (1999): Cleavage of substrates with mismatched nucleotides by Flap endonuclease-1. Implications for mammalian Okazaki fragment processing. J Biol Chem 274: 14602-14608.

Saada A, Shaag A, Mandel H, Nevo Y, Eriksson S and Elpeleg O (2001): Mutant mitochondrial thymidine kinase in mitochondrial DNA depletion myopathy. Nat Gent 29: 342-344.

Saada A, Ben-Shalom E, Zyslin R, Miller C, Mandel H and Elpeleg O (2003): Mitochondrial deoxyribonucleoside triphosphate pools in thymidine kinase 2 defieciency. Biochem Biophys Res Commun 310: 963-966.

Samadashwily GM, Raca G and Mirkin SM (1997): Trinucleotide repeats affect DNA replication in vivo. Nat Genet 17: 298-304.

Santorelli FM, Shanske S, Macaya A, DeVivo DC and DiMauro S (1993): The mutation at nt 8993 of mitochondrial DNA is a common cause of Leigh's syndrome. Ann Neurol 34: 827-834.

Sanz A, Pamplona R and Barja G (2006): Is the mitochondrial free radical theory of aging intact? Antioxid Redox Signal 8: 582-599.

Saraste M (1999): Oxidative phosphorylation at the *fin de siecle*. Science 283: 1488-1493.

Sawaya MR, Guo S, Tarbor S, Richardsson CC and Ellenberger T (1999): Crystal structure of the helicase domain from the replicative helicase-primase of bacteriophage T7. Cell 99: 167-177.

Sbisa E, Tanzariello F, Reyes A, Pesole G and Saccone C (1997): Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. Gene 205: 125-140.

Schapira AH (2006): Mitochondrial disease. Lancet 368: 70-82.

Scheffler IE (1999): Mitochondria, Wisley-Liss Inc.

ter Scheggert J, Flavell RA and Borst P (1971): DNA synthesis by isolated mitochondria. 3. characterization of D-loop DNA, a novel intermediate in mtDNA synthesis. Biochem Biophys Acta 254: 1-14.

Schriner SE, Lindford NJ, Martin GM, Treuting P, Ogburn CE, Emond M, Coskun PE, Ladiges W, Wolf N, Van Remmen H, Wallace DC and Rabinovitch PS (2005): Extension of murine life span by overexpression of catalase targeted to mitochondria. Science 308: 1909-1911.

Seow F, Sato S, Janssen CS, Riehle MO, Mukhopadhyay A, Phillips RS, Wilson RJ and Barrett MP (2005): The plastidic DNA replication enzyme complex of Plasmodium falciparum. Mol Biochem Parasitol 141: 145-153.

Shadel GS and Clayton DA (1997): Mitochondrial DNA maintenance in vertrbrates. Annu Rev Biochem 66: 409-435.

Shen CC, Wertelecki W, Driggers WJ, LeDoux SP and Wilson GL (1995): Repair of mitochondrial DNA damage induced by bleomycin in human cells. Mutat Res 337: 19-23.

Shoffner JM, Lott MY, Lezza AM, Seibel P, Ballinger SW and Wallace DC (1990): Myoclonic epilepsy and ragged-red fiber disease (MERFF) is associated with a mitochondrial DNA tRNA(Lys) mutation. Cell 61: 931-937.

Shutt TE and Gray MW (2006): Twinkle, the mitochondrial replicative DNA helicase, is widespread in the eukaryotic radiation and may also be the mitochondrial DNA primase in most eukaryotes. J Mol Evol 62: 588-599.

Slupphaug G, Markussen FH, Olsen LC, Aasland R, Aarsaether N, Bakke O, Krokan HE and Helland DE (1993): Nuclear and mitochondrial forms of human uracil-DNA glycosylase are encoded by the same gene. Nucleic Acids Res 21: 2579-2584.

Slupska MM, Luther WM, Chiang JH, Yang H and Miller JH (1999): Functional expression of hMYH, a human homolog of Escherichia coli MutY protein. J Bacteriol 181: 6210-6213.

Smeitink J, van den Heuvel L, DiMauro S (2001): The genetics and pathology of oxidative phosphorylation. Nat Rev Genet 2: 342-452.

Smirnova E, Griparic L, Shurland DL and van der Bliek A (2001): Dynamin-related protein Drp1 is required for mitochondrial division in mammalian cells. Mol Biol Cell 12: 2245-2256.

Song S, Wheeler LJ and Mathews CK (2003): Deoxyribonucleotide pool imbalance stimulates deletions in HeLa cell mitochondrial DNA. J Biol Chem 278: 43893-43896.

Song S, Pursell ZF, Copeland WC, Longley MJ, Kunkel TA and Mathews CK (2005): DNA precursor asymmetries in mammalian tissue mitochondria and possible contribution to mutagenesis through reduced replication fidelity. Proc Natl Acad Sci USA 102: 4990-4995.

Soong NW, Hinton DR, Cortopassi G and Arnheim N (1992): Mosaicism for a specific somatic mitochondrial DNA mutation in adult human brain. Nat Genet 2: 318-323.

Spelbrink JN, Van Oost BA, Van den Bogert C (1994): The relationship between mitochondrial genotype and mitochondrial phenotype in lymphoblasts with a heteroplasmic mtDNA deletion. Hum Mol Genet 3: 1989-1997.

Spelbrink JN, Zwart R, Van Galen MJ, Van den Bogert C (1997): Preferential amplification and phenotypic selection in a population of deleted and wild-type mitochondrial DNA in cultured cells. Curr Genet 32: 115-124.

Spelbrink JN, Toivonen JM, Hakkaart GA, kurkela JM, Cooper HM, Lehtinen SK, Lecrenier N, Back JW, Speijer D, Foury F and Jacobs HT (2000): *In vivo* functional analysis of the human mitochondrial DNA polymerase POLG expressed in cultured human cells. J Biol Chem 275: 24818-24828.

Spelbrink JN, Li FY, Tiranti V, Nikali K, Yuan QP, Tariq M, Wanrooij S, Garrido N, Comi G, Morandi L, Santoro L, Toscano A, Fabrizi GM, Croxen R, Beeson D, Poulton J, Suomalainen A, Jacobs HT, Zeviani M and Larsson C (2001): Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria. Nat Genet 28: 223-231.

Spinazzola A, Marti R, Nishino I, Andreu AL, Naini A, Tadesse S, Pela I, Zammarchi E, Donati MA, Oliver JA and Hirano M (2002): Altered thymidine metabolism due to defects of thymidine phosphorylase. J Biol Chem 277: 4128-4133.

Spinazzola A and Zeviani M (2005): Disorders of nuclear-mitochondrial intergenomic signaling. Gene 354: 162-168.

Spinazzola A, Viscomi C, Fernandez-Vizarra E, Carrara F, D'Adamo P, Calvo S, Marsano RM, Donnini C, Weiher H, Strisciuglio P, Parini R, Sarzi E, Chan A, DiMauro S, Rötig A, Gasparini P, Ferrero I, Mootha VK, Tiranti V and Zeviani M (2006): MPV17 encodes an inner mitochondrial membrane protein and is mutated in infantile hepatic mitochondrial DNA depletion. Nat Genet 38: 570-575.

Stonington OG and Pettijohn DE (1971): The folded genome of Escherichia coli isolated protein-DNA-RNA complex. Proc Natl Acad Sci 68: USA 6-9.

Srivastava S and Moraes CT (2005): Double-strand breaks of mouse muscle mtDNA promote large deletions similar to multiple mtDNA deletions in humans. Hum Mol Genet 14: 893-902.

Stryer L (1999): Biochemistry 4th edition, W.H. Freeman company.

Stuart JA, Karahalil B, Hogue BA, Souza-Pinto NC and Bohr VA (2004): Mitochondrial and nuclear DNA base excision repair are affected differently by caloric restriction. FASEB J 18: 595-597.

Stuart JA, Brown MF (2006): Mitochondrial DNA maintenance and bioenergetics. Biochim Biophys Acta 1757: 79-89.

Suomalainen A, Kaukonen J, Amati P, Timonen R, Haltia M, Weissebach J, Zeviani M Somer H and Peltonen L (1995): An autosomal locus predisposing to deletions of mitochondrial DNA. Nat Genet 9: 146-151.

Suomalainen A, Majander A, Wallin M, Setala K, Kontula K, Leinonen H, Salmi T, Paetau A, Haltia M, Valanne L, Lonnqvist J, Peltonen L and Somer H (1997): Autosomal dominant progressive external ophthalmoplegia with multiple deletions of mtDNA: clinical, biochemical and molecular genetic features of the 10q-linked disease. Neurology 48: 1244-1253.

Sutovsky P, Moreno RD, Ramalho-Santos J, Dominko T, Simerly C and Schatten G (1999): Ubiquitin tag for sperm mitochondria. Nature 402: 371.

Suzuki T, Kawano S and Kuroiwa T (1982): Structure of three-dimensionally rod-shaped mitochondrial nucleoids isolated from the slime mould Physarum polycephalum. J Cell Sci 58: 241-261.

Taanman JW, Muddle JR and Muntau AC (2003): Mitochondrial DNA depletion can be prevented by dGMP and dAMP suppemation in a resting culture of deoxyguanosine kinase-deficient fibroblast. Hum Mol Genet 12: 1839-1845.

Takamatsu C, Umeda S, Ohsato T, Ohno T, Abe Y, Fukuoh A, Shinagawa H, Hamasaki N and Kang D (2002): Regulation of mitochondrial D-loops by transcription factor A and single-stranded DNA-binding protein. EMBO rep 3: 451-456.

Takao M, Aburatani H, Kobayashi K and Yasui A (1998): Mitochondrial targeting of human DNA glycosylases for repair of oxidative DNA damage. Nucleic Acids Res 26: 2917-2922.

Tang Y, Schon EA, Wilichowski E, Vazquez-Memije ME, Davidson E and King MP (2000): Rearrangements of human mitochondrial DNA (mtDNA): New insights into the regulation of mtDNA copy number and gene expression. Mol Biol Cell 11: 1471-1485.

Tapper DP and Clayton DA (1981): Mechanism of replication of human mitochondrial DNA. Localization of the 5' ends of nascent daughter strands. J Biol Chem 256: 5109-5115.

Tatuch Y and Robinson BH (1993): The mitochondrial DNA mutation at 8993 associated with NARP slows the rate of ATP synthesis in isolated lymphoblast mitochondria. Biochem Biophys Res Commun 192: 124-128.

Taylor SD, Zhang H, Eaton JS, Rodeheffer MS, Lebedeva MA, O'Rourke TW, Siede W and Shadel GS (2005): The conserved Mec1/Rad53 nuclear checkpoint pathway regulates mitochondrial DNA copy number in *Saccharomyces cerevisiae*. Mol Cell Biol 16: 3010-3018.

Tell G, Crivellato E, Pines A, Paron I, Pucillo C, Manzini G, Bandiera A, Kelley MR, Di Loreto C and Damante G (2001): Mitochondrial localization of APE/Ref-1 in thyroid cells. Mutat Res 485: 143-152.

Thibodeau L and Verly WG (1980): Cellular localization of the apurinic/apyrimidinic endodeoxyribonucleases in rat liver. Eur J Biochem 107: 555-563.

Thömmes P, Farr CL, Marton RF, Kaguni LS and Cotterill S (1995): Mitochondrial single-stranded DNA-binding protein from *drosophila* embryos. J Biol Chem 270: 21137-21143.

Thorburn DR (2004): Mitochondrial disorders: prevalence, myths and advances. J Inherit Metab Dis 27: 349-362.

Thyagarajan B, Padua RA and Campbell C (1996): Mammalian mitochondria possess homologous DNA recombination activity. J Biol Chem 271: 27536-27543.

Tiranti V, Barat-Gueride M, Bijl J, DiDonato S and Zeviani M (1991): A full-length cDNA encoding a mitochondrial DNA-specific single-stranded DNA binding protein from *Xenopus laevis*. Nucleic Acids Res 19: 4291.

Tiranti V, Rocchi M, DiDonato S and Zeviani M (1993): Cloning of human and rat cDNAs encoding the mitochondrial single-stranded DNA-binding protein (SSB). Gene 126: 219-225.

Tiranti V, Savoia A, Forti F, D'Apolito MF, Centra M, Rocchi M and Zeviani M (1997). Identification of the gene encoding the human mitochondrial RNA polymerase (h-mtRPOL) by cyberscreening of the expressed sequence tags database. Hum Mol Genet 6: 615-625.

Tomkinson AE, Bonk RT and Linn S (1988): Mitochondrial endonuclease activities specific for apurinic/apyrimidine sites in DNA from mouse cells. J Biol Chem 263: 12532-12537.

Tran HT, Keen JD, Kricker M, Resnick MA and Gordenin DA (1997): Hypermutability of homonucleotide runs in mismatch repair and DNA polymerase proofreading yeast mutants. Mol Cell Biol 17: 2859-2865.

Trifunovic A, Wredenberg A, Falkenberg M, Spelbrink JN, Rovio AT, Bruder CE, Bohlooly-Y M, Gidlöf S, Oldfors A, Wibom R, Törnell J, Jacobs HT and Larsson NG (2004): Premature ageing in ice expressing defective mitochondrial DNA polymerase. Nature 429: 417-423.

Trifunovic A, Hansson A, Wredenberg A, Rovio AT, Dufour E, Khvorostov I, Spelbrink JN, Wibom R, Jacobs HT and Larsson NG (2005): Somatic mtDNA mutations cause aging phenotypes without affecting reactive oxygen species production. Proc Natl Acad Sci USA 102: 17993-17998.

Trounce I, Byrne E and Marzuki S (1989): Decline in skeletal muscle mitochondrial respiratory chain function: possible factor in ageing. Lancet 25: 637-639.

Tsuchimoto D, Sakai Y, Sakumi K, Nishioka K, Sasaki M, Fujiwara T and Nakabeppu Y (2001): Human APE2 protein is mostly localized in the nuclei and to some extent in the mitochondria, while nuclear APE2 is partly associated with proliferating cell nuclear antigen. Nucleic Acids Res 29: 2349-2360.

Tsurumi T and Lehman IR (1990): Release of RNA polymerase from vero cell mitochondria after herpes simplex virus type 1 infection. J Virol 64: 450-452.

Turcotte LP (2003): Mitochondria: biogenesis, structure, and function. Med Sci Sport Exerc 35: 82-85.

Turunen M, Olsson J and Dallner G (2004): Metabolism and function of coenzyme Q. Biochem Biophys Acta 1660: 171-199.

Tyynismaa H, Sembongi H, Bokori-Brown M, Granycome C, Ashley N, Poulton J, Jalanko A, Spelbrink JN and Suomalainen A (2004): Twinkle helicase is essential for mtDNA maintenance and regulates mtDNA copy number. Hum Mol Genet 13: 3219-3227.

Tyynismaa H, Mjosund KP, Wanrooij S, Lappalainen I, Ylikallio E, Jalanko A, Spelbrink JN, Paetau A and Suomlainen A (2005): Mutant mitochondrial helicase Twinkle causes multiple mtDNA deletions and a late-onset mitochondrial disease in mice. Proc Natl Acad Sci USA 102: 17687-17692.

Van Goethem G, Dermaut B, Lofgren A, Martin JJ and Van Broeckhoven C (2001): Mutation of POLG is associated with progressive external ophthalmoplegia characterized by mtDNA deletions. Nat Genet 28: 211-212.

Van Dyck E, Foury F, Stillman B and Brill SJ (1992): A single-stranded DNA binding protein required for mitochondrial DNA replication in S.cerevisiae is homologous to E.coli SSB. EMBO J 11: 3421-3430.

Veltri KL, Espiritu M and Singh G (1990): Distinct genomic copy number in mitochondria of different mammalian organs. J Cell Physiol 143: 160-164.

Vermulst M, Bielas JH, Kujoth GC, Ladiges WC, Rabinovitch PS, Prolla TA and Loeb LA (2007): Mitochondrial point mutations do not limit the natural lifespan of mice. Nat Genet 39: 540-543.

Wallace DC, Singh G, Lott MT, Hodge JA, Schurr TG, Lezza AM, Elsas LJ 2nd and Nikoskelainen EK (1988): Mitochondrial DNA mutations associated with Leber's hereditary optic neuropathy. Science 242: 1427-1430.

Wallace DC (1999): Mitochondrial disease in man and mouse. Science 283: 1482-1488.

Wallace DC (2005): A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: a dawn for evolutionary medicine. Annu Rev Genet 39: 359-407.

Wang L, Limongelli A, Vila MR, Carrara F, Zeviani M and Eriksson S (2005): Molecular insight into mitochondrial DNA depletion syndrome in two patients with novel mutations in the deoxyguanosine kinase and thymidine kinase 2 genes. Mol Genet Metab 84: 75-82.

Wang Y, Michikawa Y, Mallidis C, Bai Y, Woodhouse L, Yarasheski KE, Miller CA, Askanas V, Engel WK, Bhasin S and Attardi G (2001): Muscle-specific mutations accumulate with aging in critical human mtDNA control sites for replication. Proc Natl Acad Sci USA 98: 4022-4027.

Wang Y, Lyu YL and Wang JC (2002): Dual localization of human DNA topoisomerase IIIα to mitochondria and nucleus. Proc Natl Acad Sci USA 99: 12114-12119.

Wang Y and Bogenhagen DF (2006): Human mtDNA nucleoids are linked to protein folding machinery and metabolic enzymes at the mitochondrial inner membrane. J Biol Chem 281: 25791-25802.

Washington MT, Rosenberg AH, Griffin K, Studier FW and Patel SS (1996): Biochemical analysis of mutant T7 primase/helicase proteins defective in DNA binding, nucleotide hydrolysis, and the coupling of hydrolysis with DNA unwinding. J Biol Chem 271: 26825-26834.

Weiher H, Noda T, Gray DA, Sharpe AH and Jaenisch R (1990): Transgenic mouse model of kidney disease: insertional inactivation of ubiquitously expressed gene leads to nephrotic syndrome. Cell 62: 425-434.

Weitao T, Budd M, Hoopes LL and Campbell JL (2003): Dna2 helicase/nuclease causes replicative fork stalling and double-strand breaks in the ribosomal DNA of Saccharomyces cerevisiae. J Biol Chem 278: 22513-22522.

Wernette CM and Kaguni LS (1986): A mitochondrial DNA polymerase from embryos *drosophila melanogaster*. J Biol Chem 261: 14764-14770.

Willer M, Rainey M, Pullen T and Stirling CJ (1999): The yeast CDC9 gene encodes both a nuclear and a mitochondrial form of DNA ligase I. Curr Biol 9: 1085-1094.

Williamson DH and Fennell DJ (1975): The use of fluorescent DNA-binding agent for detecting and separating yeast mitochondrial DNA. Methods Cell Biol 12: 335-351.

Wong A, Cavelier L, Collins-Schramm HE, Seldin MF, McGrogan M, Savontaus ML and Cortopassi GA (2002): Differentiation-specific effects of LHON mutations introduced into neuronal NT2 cells. Hum Mol Genet 11: 431-438.

Xu B and Clayton DA (1996): RNA-DNA hybrid formation at the human mitochondrial heavy-strand origin ceases at replication start sites: an implication for RNA-DNA hybrids serving as primers. EMBO J 15: 3135-3143.

Yakubovskaya E, Chen Z, Carrodeguas JA, Kisker C and Bogenhagen DF (2006): Functional human DNA polymerase γ forms a heterotrimer. J Biol Chem 281: 374-382.

Yang C, Curth U, Urbanke V and Kang C (1997): Crystal structure of human mitochondrial single-stranded DNA binding protein at 2.4 Å resolution. Nat Struct Biol 4: 153-157.

Yang MY, Bowmaker M, Reyes A, Vergani L, Angeli P, Gringeri E, Jacobs HT and Holt IJ (2002): Biased incorporation of ribonucleotides on the mitochondrial L-strand accounts for apparent strand-asymmetric DNA replication. Cell 111: 495-505.

Yasukawa T, Yang MY, Jacobs HT and Holt IJ (2005): a bidirectional origin of replication maps to the major noncoding region of human mitochondrial DNA. Mol Cell 18: 651-662.

Yasukawa T, Reyes A, Cluett TJ, Yang MY, Bowmaker M, Jacobs HT and Holt IJ (2006): Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand. EMBO J 25: 5358-5371.

Yoshida Y, Izumi H, Torigoe T, Ishiguchi H, Itoh H, Kang D and Kohno K (2003): P53 physically interacts with mitochondrial transcription factor A and differentially regulates binding to damaged DNA. Cancer Res 63: 3729-3734.

Zelenaya-Troitskaya O, Newman SM, Okamoto K, Perlman PS and Butow RA (1998): Functions of the high mobility group protein, Abf2p, in mitochondrial DNA segregation, recombination and copy number in Saccharomyces cerevisiae. Genetics 148: 1763-1776.

Zeviani M, Moraes CT, DiMauro S, Nakase H, Bonilla E, Schon EA and Rowland LP (1988): Deletions of mitochondrial DNA in Kearns-Sayre syndrome. Neurology 51: 1525-1533.

Zeviani M, Servidei S, Gellera C, Bertini E, DiMauro S and DiDonato S (1989): An autosomal dominant disorder with multiple deletions of mitochondrial DNA starting at the D-loop region. Nature 339: 309-311.

Zeviani M and Di Donato S (2004): Mitochondrial disorders. Brain 127: 2153-2172.

Zhang H, Barcelo JM, Lee B, Kohlhagen G, Zimonkic DB, Popescu NC and Pommier Y (2001): Human mitochondrial topoisomerase I. Proc Natl Acad Sci USA 98: 10608-10613.

Zhang J, Asin-Cayuela J, Fish J, Michikawa Y, Bonafe M, Olivieri F, Passarino G, De Benedictis G, Franceschi C and Attardi G (2003): Strikingly higher frequency in centenarians and twins of mtDNA mutation causing remodeling of replication origin in leukocytes. Proc Natl Acad Sci USA 100: 1116-1121.

Zhang L, Ging NC, Komoda T, Hanada T, Suzuki T and Watanabe K (2005): Antibiotic susceptibility of mammalian mitochondrial translation. FEBS Lett 579: 6423-6427.

Zhao X, Chabes A, Domkin V, Thelander L and Rothstein R (2001): The ribonucleotide reductase inhibitor Sml1 is a new target of the Mec1/Rad53 kinase cascade during growth and in response to DNA damage. EMBO J 20: 3544-3553.

Zsurka G, Kraytsberg Y, Kudina T, Kornblum C, Elger CE, Khrapko K and Kunz WS (2005): Recombination of mitochondrial DNA in skeletal muscle of individuals with multiple mitochondrial DNA heteroplasmy. Nat Genet 37: 873-877.

# Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria

Johannes N. Spelbrink[1]*, Fang-Yuan Li[2]*, Valeria Tiranti[3], Kaisu Nikali[4], Qiu-Ping Yuan[2], Muhammed Tariq[5], Sjoerd Wanrooij[1], Nuria Garrido[1], Giacomo Comi[6], Lucia Morandi[7], Lucio Santoro[8], Antonio Toscano[9], Gian-Maria Fabrizi[10], Hannu Somer[11], Rebecca Croxen[12], David Beeson[12], Joanna Poulton[13], Anu Suomalainen[4,14], Howard T Jacobs[1,15], Massimo Zeviani[3] & Catharina Larsson[2]

*These authors contributed equally to this work.

**The gene products involved in mammalian mitochondrial DNA (mtDNA) maintenance and organization remain largely unknown. We report here a novel mitochondrial protein, Twinkle, with structural similarity to phage T7 gene 4 primase/helicase and other hexameric ring helicases. Twinkle colocalizes with mtDNA in mitochondrial nucleoids. Screening of the gene encoding Twinkle in individuals with autosomal dominant progressive external ophthalmoplegia (adPEO), associated with multiple mtDNA deletions, identified 11 different coding-region mutations co-segregating with the disorder in 12 adPEO pedigrees of various ethnic origins. The mutations cluster in a region of the protein proposed to be involved in subunit interactions. The function of Twinkle is inferred to be critical for lifetime maintenance of human mtDNA integrity.**

## Introduction

Maintenance and expression of mtDNA depend on many nuclear-encoded gene products. A minority of these proteins have been identified and characterized. In particular, the mammalian mtDNA replication machinery remains poorly defined. The only DNA polymerase that has been identified in mammalian mitochondria, DNA polymerase γ (POLG), is believed to be the replicative mtDNA polymerase. The orthologous enzymes in yeast (Mip1p) and *Drosophila* (encoded by the gene *tamas*) are required for mtDNA maintenance[1,2]. Processivity is ensured by the accessory subunit of POLG[3], but the helicase(s) involved in mtDNA replication remains unidentified. The asynchronous model of mtDNA replication involves prolonged extension of the leading strand until a dedicated lagging-strand origin is encountered[4]. The priming mechanism on the lagging strand is obscure, but this replication mode requires involvement of the mitochondrial single-stranded DNA binding protein SSBP. Under certain conditions, mtDNA in mammalian cells replicates by a more conventional, strand-synchronous mechanism[5], which implies

the presence of a primase and a second molecule of the polymerase in the replication complex. The proteins involved in primer removal, closure and ligation, as well as the topoisomerase(s) involved in decatenation, also remain uncertain.

The closest relatives of mitochondrial POLG are members of the PolA superfamily from bacteria (for example, *Escherichia coli* PolI) and T-odd bacteriophages[6]. Phage T7 DNA replication involves only a small set of proteins: T7 DNA polymerase, T7 RNA polymerase, SSBP, DNA ligase, the multifunctional primase/helicase encoded by phage gene 4 and host-encoded thioredoxin[7]. The relatedness of POLG to T7 DNA polymerase and also of the mitochondrial RNA polymerase to T7 RNA polymerase[8] raises the question of whether other components of the mtDNA maintenance machinery might be closely related to their counterparts from T-odd bacteriophages.

PEO associated with multiple mtDNA deletions is one of the few examples of a human disorder in which the mtDNA maintenance machinery is defective. Typical symptoms are ophthalmoparesis and exercise intolerance. Some people also may
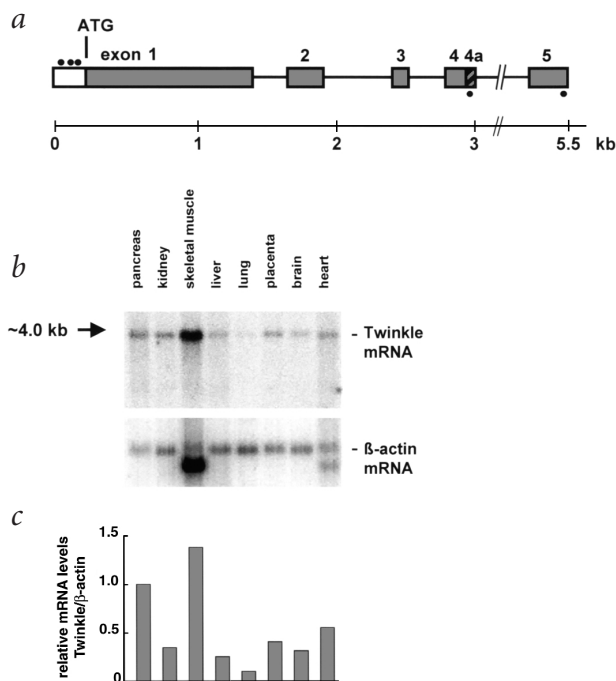
*article*

**Fig. 1** *C10orf2* gene structure and mRNA expression in human tissues. *a*, *C10orf2* exon-intron structure, based on cDNA and genomic sequence. Stop codons are indicated by filled circles; the 43-bp exon 4 extension, exon 4a, is indicated by a hatched box. Sequence is numbered from the first bp known to be present in the mRNA, np 70863 of *C10orf2* and ends at the stop codon (np 76356). *b*, An 8-lane human multitissue northern blot, hybridized with a full-length *C10orf2* cDNA probe (upper) and rehybridized with a β-actin probe (lower). *c*, Twinkle mRNA levels relative to β-actin control mRNA levels.

develop cardiomyopathy, cataracts, ataxia, peripheral neuropathy, hypogonadism or major depression[9–13]. The diagnosis depends on the demonstration, by Southern blotting, of multiple deletions of mtDNA in muscle biopsy specimens[9]. Such deletions also can be detected in brain and heart[11]. Accumulated deletions can be clonal, which suggests that deletion events are rare[14]. Familial PEO is genetically heterogeneous, occurring in an autosomal dominant form (adPEO (ref. 9)) or as an autosomal recessive disorder (arPEO (ref. 15); A. Suomalainen *et al.* unpublished data). Multiple mtDNA deletions with recessive inheritance can also be found in a more severe disorder involving infantile cardiomyopathy[16]. adPEO involves at least three different disease gene loci. Two of these have been mapped to chromosome regions 10q24 and 4q34 (refs. 13,17,18). The 4q-linked adPEO gene was recently identified as *ANT1*, which encodes the heart-muscle-brain isoform of the adenine nucleotide translocator[19]. Another autosomal recessive disorder associated with mtDNA deletions and depletion, MNGIE (mitochondrial neurogastrointestinal encephalomyopathy), was recently found to be caused by mutations in the thymidine phosphorylase gene[20].

Localization of the 10q-linked adPEO gene to a 7-cM critical interval[13] prompted us to search this region for plausible candidate genes likely to be involved in mtDNA replication or nucleotide metabolism. We report here the identification of the chromosome 10q-linked adPEO gene, which encodes a novel mitochondrial protein with striking similarity to the phage T7 gene 4 primase/helicase, that colocalizes with mtDNA in structures designated as mitochondrial nucleoids.

## Results

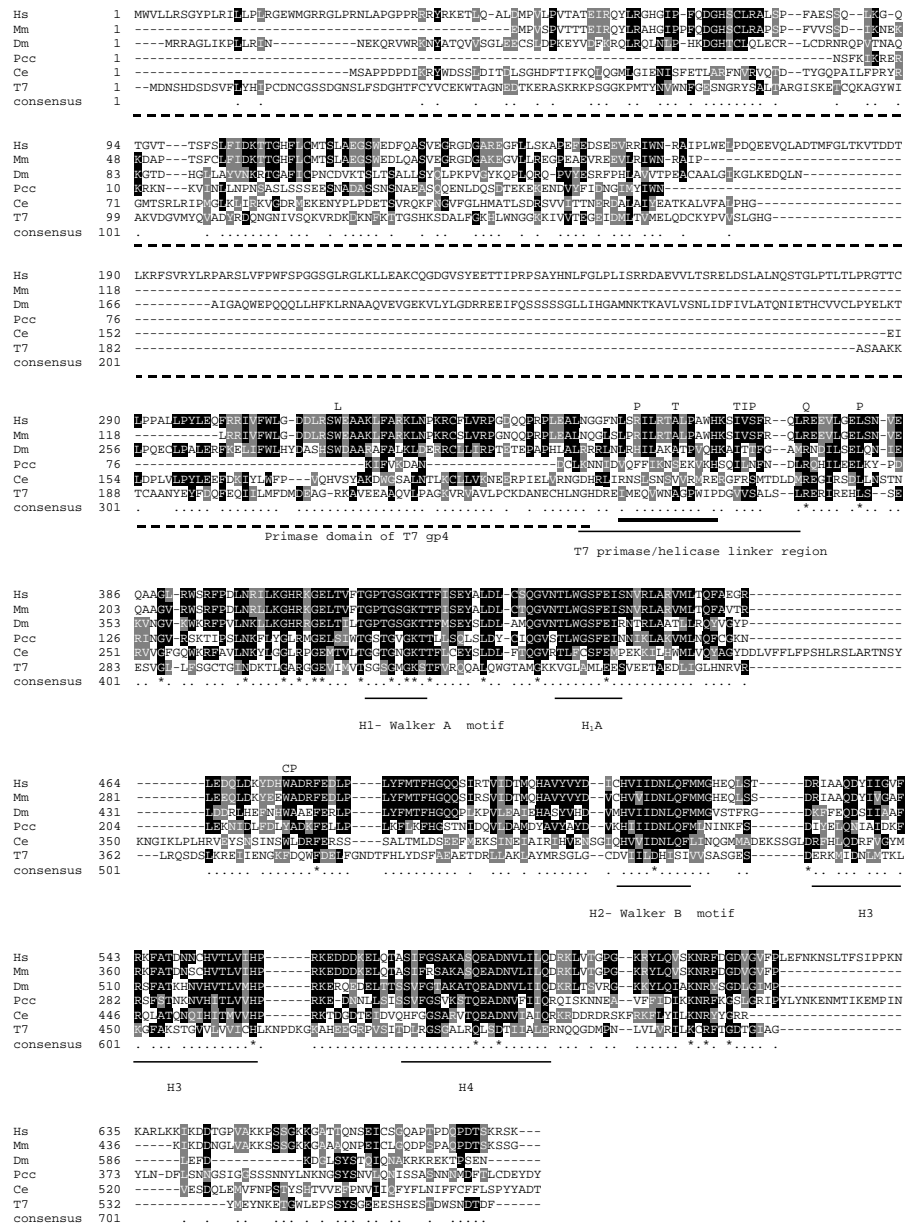### A protein similar to T7 primase/helicase is encoded by a human gene on chromosome 10

We identified human gene *C10orf2*, encoding a protein similar to T7 primase/helicase (gene 4 protein, gp4), by searching for open reading frames (ORFs) in a region previously identified as being linked to adPEO on chromosome 10q24 (refs. 13,17).

Based on a partial sequence, the carboxy-terminal portion of C10orf2 was previously suggested to be a possible helicase[21]. A combination of cDNA sequencing, resequencing of IMAGE cDNA clones 2413374, 33362 and 1113546 (ref. 22), intron-exon prediction programs and EST database searches revealed a single ORF of 2.1 kb, with a plausible start codon (Fig. 1*a*). We detected a transcript corresponding to this ORF by cDNA amplification and characterized it by direct sequencing. Several in-frame stop codons are found in the immediate 5′ untranslated region (UTR), which indicates that the ATG at position 186 is the initiation codon. There is almost complete coverage of the full-length mRNA with human as well as mouse ESTs, except for a short region at the 5′ end. The full-length ORF identified here is likely to be mitochondrial because of the presence of a putative amino-terminal mitochondrial targeting sequence predicted by MitoProt II (http://www.mips.biochem.mpg.de/proj/medgen/mitop/).

The sequence of an additional cDNA amplification product reveals the presence, in human cells, of a splice variant resulting from the use of a downstream exon 4 splice-donor site and leading to a 43-base-pair (bp) insertion between the regular exon 4–exon 5 sequence. This splice variant is also found in the human and mouse EST databases.

We hybridized multitissue northern blots with oligonucleotide and cloned cDNA fragment probes for the *C10orf2* mRNAs. On the basis of two independent blots, the full-length mRNA is approximately 4 kb, significantly larger than the length of the ORF. The expression pattern of the full-length mRNA (Fig. 1*b,c*), based on phosphorimaging analysis of two blots, showed high relative levels in skeletal muscle and pancreas but, perhaps surprisingly, a lower level of expression in the heart. The relative level of Twinkle expression in skeletal muscle is probably underestimated because of the strong cross-reacting α-actin mRNA species just below the β-actin mRNA. Hybridization with an oligonucleotide probe specifically detecting the splice variant mRNA revealed a similar expression pattern, as did an oligonucleotide detecting both splice forms (data not shown).

The predicted full-length protein is 684 amino acids, with a molecular mass of 77 kD, whereas the splice variant mRNA encodes a 66-kD product of 582 amino acids, lacking residues 579–684 and terminating with four unique amino acids. Both cDNA variants, when expressed transiently in HEK293T cells, direct the synthesis of proteins close to the expected molecular mass (see below). *Plasmodium chabaudi chabaudi*, *Caenorhabditis elegans* and *Drosophila melanogaster* have similar proteins, also related to T7 gp4. The fission yeast *Schizosaccharomyces pombe* has a few homologs with low similarity, but we could not identify a related sequence in *Saccharomyces cerevisiae*. Multiple sequence alignment of these various proteins shows greatest similarity in the domain responsible for helicase activity in the phage T7 protein, which is located in its C-terminal half[23] (Fig. 2). The degree of identity between the human polypeptide and a partial mouse sequence assembled from ESTs is approximately 85%, whereas between the mammalian and invertebrate sequences it is in the range 35–40%.

## *C10orf2* encodes a protein localized to mitochondrial nucleoids

To test mitochondrial targeting of the *C10orf2* gene products, we generated expression constructs for various reporter fusion proteins. The full-length protein of 684 amino acids, fused to Enhanced Green Fluorescent Protein (EGFP), when expressed in a variety of cell lines (HEK293T, 143B osteosarcoma and A549 adenocarcinoma) shows a punctuate fluorescence that, as shown by Mitotracker counterstaining and confocal microscopy, is clearly localized within mitochondria (Fig. 3). An EGFP fusion protein starting from the second methionine, 21 amino acids downstream of the first, gives uniform cytoplasmic fluorescence, demonstrating that the N-terminal region is necessary for mitochondrial targeting (Fig. 3g-i). To visualize mtDNA, we counterstained transfected and untransfected 143B cells with ethidium bromide (EtBr), a DNA intercalating reagent[24,25]. In both the absence and presence of the EGFP fusion protein, we observed bright, rapidly bleaching, punctate cytoplasmic fluorescence against a fainter background of more diffuse cytoplasmic fluorescence. The long incubation times with ethidium bromide (EtBr) allow it to intercalate into mtDNA and inhibit mitochondrial RNA synthesis. The resulting punctate staining resembles previous findings with bromodeoxyuridine and 4′,6-diamidino-2-phenylindole (DAPI) labeling[25,26], which indicates specific mtDNA staining. In transfected cells, punctate green fluorescence clearly colocalizes with the bright EtBr staining (Fig. 4a-c), which indicates nucleoprotein structures that contain both the *C10orf2* product and mtDNA. By analogy with such complexes previously inferred

to exist in fungal, plant and protozoan mitochondria, we designate them mitochondrial nucleoids. A549 cells consistently contain fewer dots than 293T cells, whereas 143B cells contain more. The number of foci observed is typically between 50 and several hundred per cell and agrees well with the number of bright fluorescent EtBr spots in 143B cells.

We also observed punctate mitochondrial fluorescence for the EGFP reporter in ρ° cells, indicating that Twinkle localization is not strictly dependent on DNA binding (Fig. 4d–f), although the number of localization sites per cell is lower than in the parental ρ+ cell line and some cells show uniform mitochondrial fluorescence, coincident with Mitotracker staining. Because of the unusual localization pattern, reminiscent of twinkling stars, we designate the full-length protein Twinkle (for T7 gp4-like protein with intramitochondrial nucleoid localization). An EGFP reporter fusion for the polypeptide encoded by the C-terminal splice variant, lacking amino acids 584–684, shows uniform mitochondrial fluorescence (Fig. 3d–f), as observed in human cells with other mitochondrially targeted proteins such as

**Fig. 2** Multiple sequence alignment of C10orf2 protein with T7 gp4 and several eukaryotic homologs. Full-length proteins from human (Hs), *P. chabaudi chabaudi* (Pcc, POM1), *C. elegans* (Ce), *D. melanogaster* (Dm), phage T7 gp4 (T7) and partial protein based on EST data from *Mus musculus* (Mm) were aligned using a ClustalW server (http://www.ebi.ac.uk/clustalw/) with default settings. For clarity the C-terminal part of POM1 is omitted. Asterisks below the alignment indicate amino acid identity in all species. Dots indicate similarity or identity in at least half the sequences. Identity is indicated by black shading and similarity is indicated by gray shading. Underlined regions indicate conserved motifs among ring helicases as well as the T7 gp4 linker-like region. A more thorough alignment of the helicase domain of C10orf2 protein with various phage and eubacterial RecA and DnaB-like proteins has been published[21]. The primase domain of T7 gp4 is indicated by a dashed line. Amino acids mutated in the various adPEO pedigrees are indicated in boldface above the aligned sequence. The sequence duplicated in the Finnish adPEO family is indicated with thick underlining.
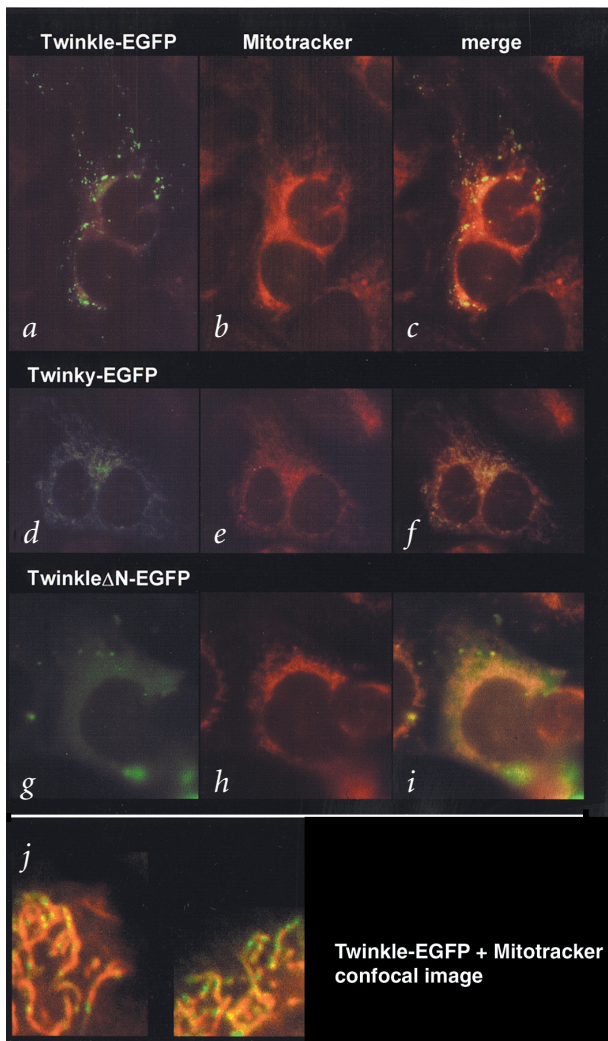
**Fig. 3** Subcellular localization of C10orf2-EGFP fusion proteins. *a–c*, Subcellular localization of full-length C10orf2-EGFP fusion protein (Twinkle-EGFP, *a* and *c*) compared with Mitotracker Red (*b* and *c*) in 143B osteosarcoma cells. Punctate green fluorescence always coincides with mitochondria, as confirmed by confocal microscopy (*j*). *d–f*, Colocalization of the C10orf2 splice-variant protein fused to EGFP (Twiny-EGFP, *d* and *f*), with Mitotracker Red (*e* and *f*) in 143B cells. Loss of the C-terminal tail results in uniform mitochondrial fluorescence. *g–i*, Uniform cytoplasmic localization of a C10orf2-EGFP fusion lacking the natural N terminus (TwinkleΔN-EGFP, commencing at Met-22). The N-terminal 21 amino acids are thus essential for mitochondrial localization. (*j*) Details from a confocal scanning laser microscopy image combining green fluorescence of Twinkle-EGFP with red fluorescence of Mitotracker Red. Twinkle fluorescence is clearly restricted to discrete foci located within the thread-like mitochondria stained by Mitotracker. Note that there are typically multiple Twinkle foci within each mitochondrion.

POLG[27] and TFAM (J.N.S., unpublished data). We designate the splice variant-encoded protein that lacks nucleoid localization Twinky. A further EGFP reporter fusion protein with a larger C-terminal truncation (amino acids 454–684) also shows uniform mitochondrial fluorescence, which indicates that the C terminus is necessary for nucleoid localization (data not shown).

## Cells overexpressing Twinkle have a modestly increased mtDNA helicase activity

Crude mitochondrial extracts from HEK293 cells show clear ATP- or dATP-stimulated DNA helicase activity on an M13 oligonucleotide substrate (Fig. 5). Overexpression of FLAG-tagged or untagged Twinkle (but not Twinky) consistently gives a modest increase (51% ± 10%, n=4) in this activity. Addition of 10 mM EDTA to the assay mix completely abolishes helicase activity. A second, forked M13 substrate with a longer stretch of double-stranded DNA, as used to detect T7 gp4 helicase[28], is not unwound in the mitochondrial extracts. These properties are similar to the DNA helicase activity described in purified bovine mitochondria[29]. Affinity-purified Twinkle with a C-terminal His$_6$ tag shows no detectable helicase activity (data not shown).

## Mutant Twinkle in individuals with chromosome 10q24-linked adPEO

So far, two large adPEO families, one Finnish and one Pakistani, show linkage to chromosome 10q24 (refs. 13,17, respectively). Many genes mapping within this interval have already been excluded, based on SSCP or sequencing of DNA from these families[30] (A. Suomalainen *et al.* and F.Y. Li *et al.*, unpublished data). Because *C10orf2* maps within the critical interval, we analyzed the 5′ UTR, coding regions, splice junctional sequences and 3′ UTR of the gene in affected and unaffected individuals from these families. The Finnish proband reveals a heterozygous 39-bp duplication nt 1053–1092 counting from the ATG start codon), resulting in duplication of amino acids 352–364 of Twinkle. Polymerase chain reaction (PCR) analysis of DNA from other family members shows complete segregation of the duplication with the disease phenotype (Fig. 6*a*). The Pakistani proband (IV:19; Fig 6*b*) shows a single, heterozygous, missense G→C transversion at bp 1423, resulting in an amino acid change of Ala-475 to Pro (A475P). This point mutation creates a *Hae*III restriction site, and analysis of restriction digests of amplicons from the family shows all affected family members to be heterozygous for this mutation. Analysis of over 400 Finnish controls for the duplication and 194 mixed controls including 88 of Pakistani/Indian origin for the A475P-associated *Hae*III site failed to find either mutation.

In addition to these two families, we sequenced *C10orf2* exons and intron boundaries in ~70 adPEO families where no significant linkage data had been derived. Most of these families were of Italian origin; a few were of British, Hungarian and American origin. We identified nine additional heterozygous mutations: eight in Italian families and one in a British family. We found one of the mutations twice in supposedly unrelated families, but it originated from the same geographic region of Italy (Web Table A). Genotype analysis showed co-segregation in all these families (Fig. 6*c,d* shows two of the larger Italian families). The large family 3 (Web Table A), with the A359T mutation, shows multiple cases of consanguinity, with two homozygous individuals having a more severe phenotype with earlier onset than their heterozygous relatives. We did not detect any of the mutations in appropriate control groups of 100 subjects. (Fig. 2 and Web Table A summarize all mutations.) Remarkably, 8 of the 11 mutations cluster in a small region of about 30 amino acids, corresponding to the linker region of the phage T7 gp4 between the primase and helicase domains. The exact function of this region is unknown, although it shows 100% sequence identity between human and mouse. However, mutations in this region might affect subunit interactions, because the corresponding segment of the T7 protein is required for hexamerization[31]. Two of the three other mutations, W474C and A475P, are in a moderately conserved region between Walker motifs A and B of the helicase domain. In the T7 gp4 they lie in an α-helix also implicated in subunit interactions[32]. Mutation W315L is also in a moderately conserved segment upstream of the linker region.

## *C10orf2* mutations do not affect Twinkle localization or multimerization

To test whether mutations in Twinkle could interfere with mitochondrial (or nucleoid) targeting, we prepared two Twinkle-EGFP fusion proteins equivalent to the 352–364 duplication and

**Fig. 4** Colocalization of Twinkle-EGFP with EtBr-stained mtDNA. ***a–c***, Cells transfected with Twinkle-EGFP were treated with EtBr to stain mtDNA. Some examples of clear colocalization of the Twinkle-EGFP fusion protein with EtBr are indicated by arrows. ***d–f***, Colocalization of Twinkle-EGFP with Mitotracker Red in (mtDNA-less) ρ° cells. Punctate green mitochondrial fluorescence can still be observed even in the absence of mtDNA.
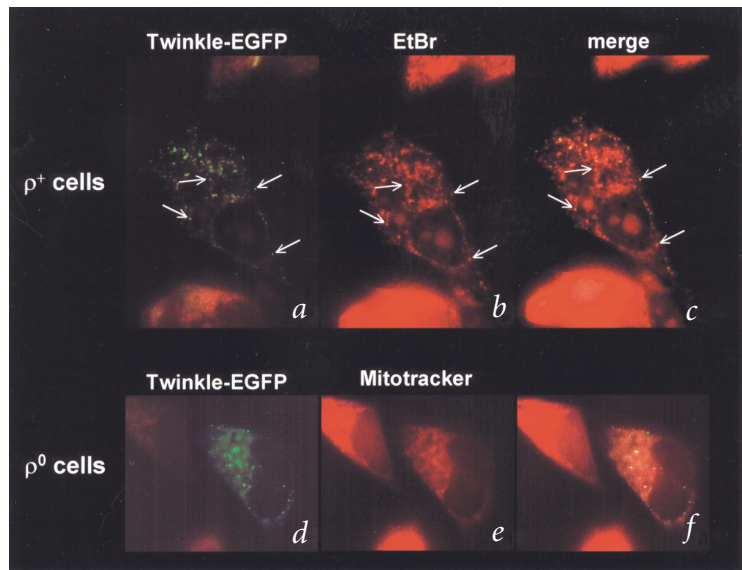


W474C mutants. Both proteins show the same punctate mitochondrial fluorescence as the wildtype protein fused to EGFP (data not shown).

To investigate whether adPEO-associated mutations could interfere with multimerization of Twinkle, we transfected FLAG epitope-tagged wildtype Twinkle, Twinky, Twinkle$^{dup352–364}$ and Twinkle$^{W474C}$ mutant variants into HEK293T cells. Ring helicases to which T7 gp4 belongs form hexamers *in vivo* when bound to DNA in the presence of Mg$^{2+}$ and a nucleotide—for example, dTTP (refs. 33,34). At high protein concentration, dimers and higher order multimers form spontaneously *in vitro*[23,33]. Western blot analysis (Fig. 7a) reveals wildtype Twinkle, Twinkle$^{dup352–364}$ and Twinkle$^{W474C}$ as monomers with the expected molecular mass, plus dimers, trimers and higher order multimers. Twinky presents as a monomer. Under less denaturing conditions (Fig. 7b), wildtype Twinkle, like T7 gp4, mainly forms multimers in such extracts, most prominently dimers, trimers/tetramers and hexamers. Again, Twinky presents mainly as a monomer. Cotransfection with Twinkle-myc does not significantly alter the multimerization pattern of Twinky-FLAG, which indicates that the two *C10orf2* products do not form heterooligomers. Neither the duplication nor the W474C adPEO mutation affects Twinkle multimerization.

The tissue specificity and late onset of the disorder indicate that Twinkle mutations are likely to have only minimal, short-term effects on mtDNA integrity in cell culture. We tested this by Southern blotting DNA of cell lines from affected individuals.

mtDNA deletions were undetectable, even in a lymphoblastoid cell line homozygous for the G1075A mutation (data not shown). Deletions were also undetectable in HEK293 cells transiently transfected with either wildtype or mutant (dup352–364) Twinkle (data not shown).

## Discussion

We show here that individuals with chromosome 10q24-linked adPEO have heterozygous mutations altering phylogenetically conserved amino acids of Twinkle, a new mitochondrial protein with striking similarity to phage T7 gp4. Twinkle colocalizes with mtDNA in nucleoprotein complexes designated mitochondrial nucleoids. Strict co-segregation with the disease identifies the adPEO mutations as causal, although they do not affect Twinkle nucleoid localization or multimerization. Most mutations cluster in a segment of the protein corresponding to a region of T7 gp4 implicated in subunit interactions.

### Structure and function of Twinkle

Twinkle has considerable similarity with T7 gp4, mostly in the helicase domain, which supports the hypothesis that it could function as a mtDNA helicase. Multiple sequence alignments show that the putative helicase domain of Twinkle has the same conserved organization as other members of the RecA/DnaB superfamily, and invariant amino acids of this superfamily are strictly conserved in Twinkle (see Fig. 2; ref. 21). These include G420/K421 of the Walker A motif, which forms part of the NTP binding and hydrolysis site, E445 of the H$_1$A motif, D516 of the Walker B motif and various other residues believed to be important for helicase function. Twinkle forms multimers *in vitro*, including hexamers, similar to those found for T7 primase/helicase and other ring helicases, with the hexamer being the functional form of the enzyme. Inside mitochondria, Twinkle colocalizes with mtDNA, as might be expected for a protein
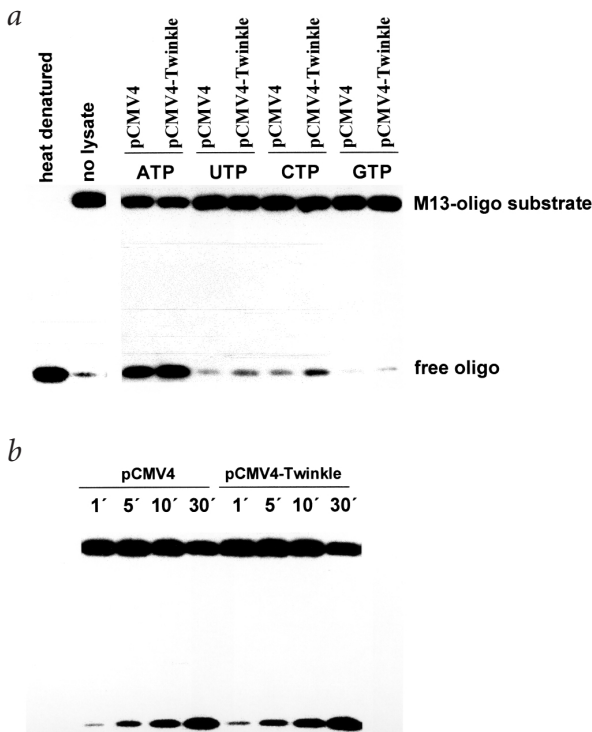


**Fig. 5** Enhancement of mitochondrial DNA helicase activity by transient expression of Twinkle tagged (or untagged). Twinkle was overexpressed in HEK293 cells by transient transfection. DNA helicase activity was measured in crude mitochondrial lysates and compared with activity from mock-transfected cells. ***a***, Comparison of helicase activity (30 min) from mock-transfected (pCMV4) and Twinkle-transfected (pCMV4-Twinkle) cells in the presence of various ribonucleotides. ATP (also dATP; data not shown) gave the most marked stimulation. Twinkle-transfected cells showed a modest, but reproducible, increase in activity. ***b***, Time course of the ATP-stimulated DNA helicase activity comparing mock-transfected and Twinkle-transfected cells.

involved in mtDNA metabolism. In addition, cells overexpressing Twinkle show a modestly increased ATP/dATP-stimulated mtDNA helicase activity, indicating that Twinkle is an adenine nucleotide-dependent DNA helicase similar to that previously purified from bovine mitochondria[29]. Because transfection efficiency in these cells is high, the weak enhancement of DNA helicase activity in transient transfection experiments probably indicates that Twinkle requires other, limiting factors for full activity. Alternatively, post-translational modification might regulate its activity but is not carried out efficiently in the overexpressed protein.

A role for Twinkle as a mitochondrial primase is less likely, because primase domain motifs found in the N-terminal region of phage T7 gp4 and bacterial and phage DnaG-like primases[21,35] are not identifiable in Twinkle. Furthermore, very little overall identity exists between the N-terminal regions of the various Twinkle-like proteins in eukaryotes.

### Localization of Twinkle in mitochondrial nucleoids

In fungi, plants and protozoa, large protein-mtDNA complexes called nucleoids[36-38] exist, which are probably also the functional units of mtDNA segregation. DAPI staining of human cells shows compact mtDNA-containing structures[26]. Based on the results presented above, Twinkle also localizes to such structures, which are visible by EtBr staining in untransfected cells. By analogy with other eukaryotes, we designate these structures mitochondrial nucleoids,

although their exact function(s) in mtDNA maintenance or partition remains to be elucidated. The fact that, based on similar reporter assays, some other mitochondrial proteins involved in DNA metabolism, such as POLG, are uniformly distributed within mitochondria indicates that the latter may need to be recruited to mtDNA by low-abundance proteins, perhaps in a temporally regulated fashion. The number of fluorescent Twinkle foci per cell is approximately constant for a given cell line but systematically different between cell lines. It is also significantly smaller than estimates of mammalian mtDNA copy number[39,40], which implies that mtDNA nucleoids are multicopy with respect to mtDNA.

The variant polypeptide designated Twinky, which is not nucleoid localized, lacks part of the helicase H4 motif plus the polar C-terminal tail and does not efficiently form multimers *in vitro*. Although the H4 motif has been implicated in DNA binding[28], the fact that punctate mitochondrial fluorescence can still be observed when full-length Twinkle is expressed in ρ° cells suggests that these punctate structures are determined by multimerization or the C-terminal domain instead of just DNA binding.

### Pathogenic mechanism of *C10orf2* mutations in adPEO

We have identified 11 different *C10orf2* mutations in 12 adPEO families. The following evidence supports the idea that the mutations are of pathogenic significance in adPEO, thus specifying Twinkle as the 10q24 adPEO gene: their absence from ethnically matched controls, their occurrence within phylogenetically
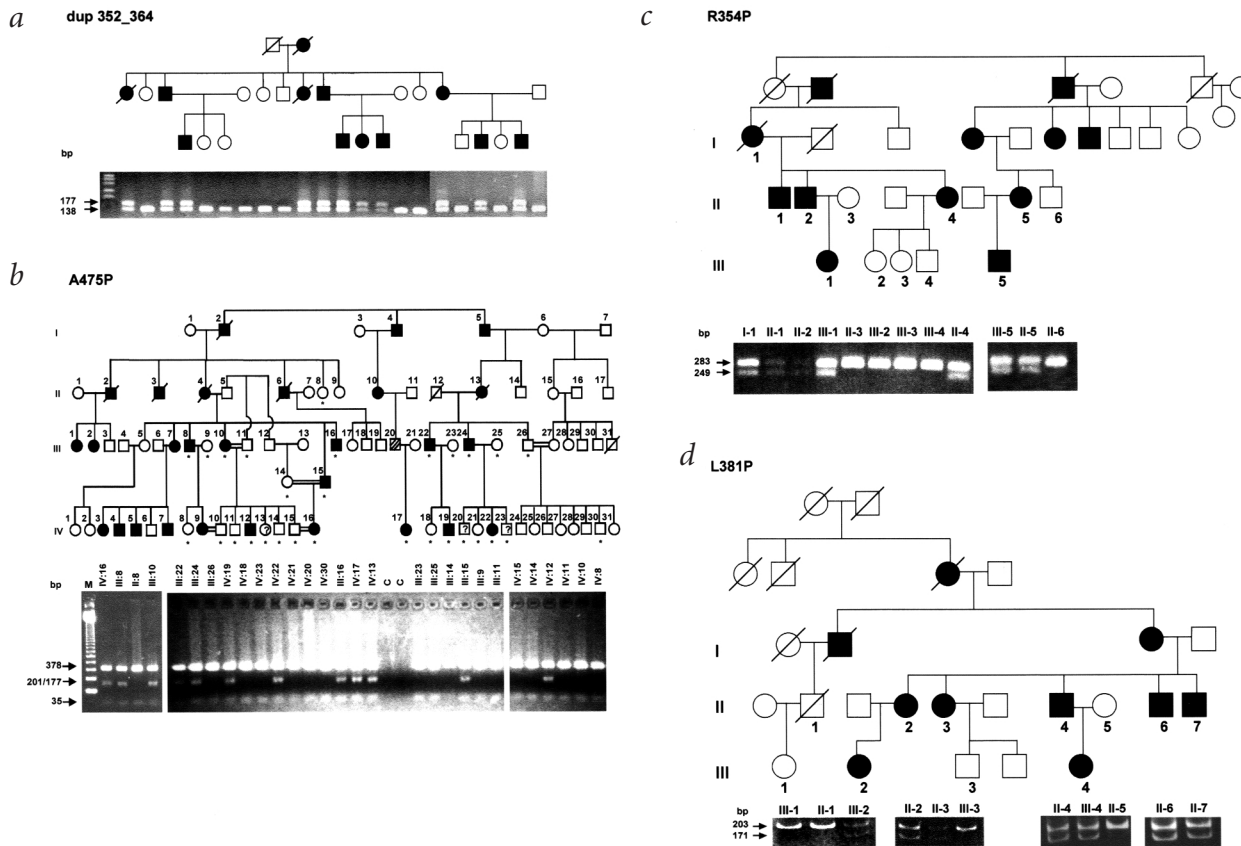


**Fig. 6** Twinkle mutations and segregation analysis in large adPEO pedigrees In each pedigree mutations were initially identified by direct sequencing of genomic DNA from at least one individual. Other individuals were screened either by PCR for the duplication (Finnish pedigree, *a*) or by restriction fragment length polymorphism PCR for the other mutations (*b–d*). *a*, Finnish family. All second- and third-generation members indicated were analyzed by PCR with primers flanking the duplication site. All affected individuals were heterozygous, whereas the larger allele could not be amplified from healthy family members or 400 controls. *b*, Pakistani family. The np G1423C (A475P) mutation creates a *Hae*III site, giving two additional products in mutant heterozygotes. C, control PCRs. Family members analyzed are indicated by asterisks. Note that family tree and analysis results are not aligned. For other details of the pedigree refer to Li *et al*[13]. *c,d*, Italian adPEO families 2, G1061C (R354P) and T1142C (L381P), were genotyped by restriction fragment length polymorphism PCR.

**Fig. 7** Multimerization of Twinkle and its variants. *a*, Western blots (SDS–10% PAGE) of FLAG-tagged Twinkle, Twinky and Twinkle variants carrying the W474C and dup352-364 adPEO mutations. Wildtype or mutant Twinkle forms dimers, trimers and higher order multimers under these conditions, whereas Twinky is almost exclusively in the monomeric form. *b*, Western blots (low percent SDS–6% PAGE) of the same variants. Twinkle multimers predominate, and their formation is essentially unaffected by adPEO mutations. Twinky multimerization is weaker, even in the presence of coexpressed Twinkle-myc.
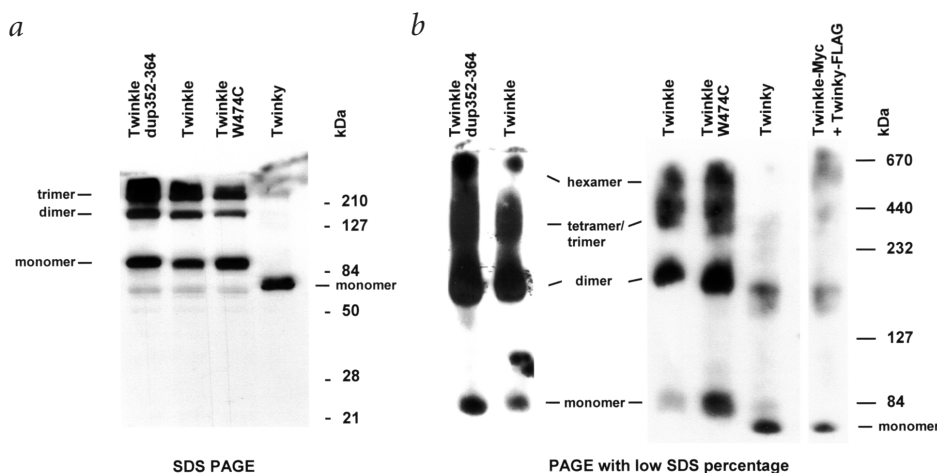
conserved regions at least among all metazoans and frequently also in T7, their largely nonconservative nature and their strict co-segregation with disease in all the families analyzed. The pattern of largely private mutations is as expected for a rare, dominantly inherited disorder.

None of the adPEO-associated mutations affects amino acids corresponding to those known to be catalytically essential for helicase activity in T7 gp4, although two are present within the putative helicase domain. All the disease mutations lie in a region implicated in subunit interactions in T7 gp4 (refs. 31,32,41,42), although neither of the disease mutations tested affects multimerization *in vitro*. In fact, the A257T mutation in T7 gp4, which is analogous to Twinkle A359T, appears to stabilize multimers *in vitro*[28,31]. Assuming that Twinkle functions *in vivo* as a multimer, heterozygosity for a mutation affecting subunit interactions, even in a subtle manner, could have dominant-negative effects on the catalytic properties of the enzyme. Based on the observation that the T7 gp4 A257T mutant exhibits enhanced dNTPase activity, most marked with dTTP in the absence of substrate DNA[28], we prefer a gain-of-function hypothesis. If A359T or other adPEO mutations in Twinkle also causes enhanced dNTP breakdown, an attractive parallel would exist with the two other genetic defects so far identified in multiple mtDNA deletion disorders. Both of these affect proteins involved in nucleotide metabolism: thymidine phosphorylase in MNGIE, an enzyme of the dTTP salvage pathway[20], and ANT1 in chromosome 4q-linked adPEO, which could influence intramitochondrial deoxynucleotide levels[19]. A functional deficiency of dNTPs could impair the processivity of mtDNA replication, creating a higher concentration of 3′ ends that would be substrates for illegitimate recombination. Because mtDNA deletions in tissues can be clonal[14] and cell lines from affected individuals or cells that transiently overexpress mutant Twinkle do not have detectable levels of these molecules, deletion events should be rare and probably limited to postmitotic tissues.

Many other possibilities remain. Healthy individuals also harbor the types of mtDNA rearrangement found in adPEO at low levels, and we previously suggested that their accumulation in pathological states could result from an increased rate of their generation, a reduced rate of their clearance, or an alteration in their selective value[43]. Repeated unequal partition combined with the slight replication advantage of rearranged mtDNAs or even just a defect in mtDNA repair, promoting enhanced mtDNA turnover, could favor the progressive accumulation of deleted molecules in a mosaic fashion, especially within postmitotic tissues such as muscle. Either a mtDNA repair defect or a replication-stalling mechanism is also suggested by the association of a POLG mutation in a highly conserved residue with adPEO in a

Belgian pedigree[44]. There are precedents for involvement of RecQ-type helicases in disorders of nuclear DNA repair[45], and mutants in the yeast RecQ-like helicase *SGS1* accumulate extrachromosomal ribosomal DNA minicircles[46].

**Clinical implications**

Within the largely outbred Italian population, we found *C10orf2* mutations in approximately 15% of cases of familial PEO showing multiple mtDNA deletions. Mutations in *ANT1* account for an even smaller proportion of Italian adPEO families[19]. This strongly implies additional genetic causes for the disorder.

The individuals of the adPEO family with the most drastic mutation, giving the 13 amino acid duplication, had severe retarded depression and avoidant personality features in addition to PEO. Three autopsies from family members showed the highest amounts of deleted mtDNA in the cerebral cortex and basal ganglia (~60% of total mtDNA), followed by skeletal muscle and heart, with only small amounts in kidney[47]. However, this distribution does not correspond exactly with the pattern of Twinkle expression, arguing for variable sensitivity of different tissues to Twinkle dysfunction.

The evidence that Twinkle mutations can cause adPEO indicates that the activity of the protein is required for maintaining mtDNA integrity during life. With appropriate model systems, in which we systematically manipulate *ANT1* and *C10orf2* expression, it should be possible to establish the roles of these genes in mtDNA replication and maintenance, leading to an understanding of adPEO pathogenesis.

**Methods**

**Monitoring GFP reporter gene expression.** Mitotracker Red staining, cell preparation and fluorescence microscopy were done as described[27]. We adapted *in vivo* staining of mtDNA with EtBr from published protocols[24,25] but used longer incubation. Cells grown on coverslips were incubated for 2 h with EtBr at 1 µg/ml in culture medium at 37 °C, washed twice with medium, allowed to stand for 10 min, washed twice with phosphate-buffered saline (PBS), mounted in PBS containing 5% sucrose and examined by fluorescence microscopy. For confocal scanning laser microscopy, we used a Perkin Elmer/Wallac Ultra*View* LCI system.

**PAGE and western blotting.** Standard SDS–PAGE used 7.5–12% polyacrylamide gels[49] as described previously[27]. Minimally denaturing SDS gels (6% polyacrylamide) used 0.02% SDS instead of the usual 0.1%. For these gels, we mixed 40 µg of protein 1:1 just before loading with a 1:10 dilution of the standard SDS sample buffer in 12% glycerol that was not heat denatured. Native PAGE markers were from Amersham/Pharmacia. We used standard methods for western blotting[27]. Primary antibodies used were mouse anti-myc monoclonal antibody 9E10 (Boehringer Mannheim), 1:15,000 dilution of a 5 mg/ml stock, and mouse anti-FLAG M2 monoclonal antibody (Stratagene), 1:5,000 dilution of a 2 mg/ml stock.

**Cell lysis.** Cell pellets for PAGE were diluted in PBS, lysed by the addition of 1 vol of digitonin (4 mg/ml), vortexed and kept on ice for 15 min. Six volumes of PBS were added and samples were centrifuged for 10 min at 16,000$g_{max}$ at 4 °C. Pellets were washed once with PBS, further lysed by adding PBS containing 1.5% (w/v) lauryl maltoside and 2.5 mM phenylmethylsulfonyl fluoride, vortexed, kept at 4 °C for 15 min and centrifuged at 16,000 $g_{max}$ for 1 min at 4 °C. Supernatants were stored at –80 °C or used immediately.

**DNA helicase activity measurements.** Mitochondria were isolated and lysed as described[27]. We measured DNA helicase activity at 37 °C (ref. 29) in 25-µl reaction mixtures containing 0.5–1 nM M13 substrate and 10 µg of mitochondrial protein. The substrate comprised 80 nM end-labeled M13 universal oligonucleotide (17 nucleotides) and 50 nM M13mp18(+) single-stranded DNA in 50 mM NaCl, 40 mM Tris-HCl (pH 7.5), heated at 75 °C for 5 min, cooled slowly to 30–37 °C and purified on a Centricon 100 column. We stopped reactions by adding 6 µl of agarose gel buffer containing 100 mM EDTA and 0.3% SDS. We separated the M13 substrate and unwound oligonucleotides with native 10% PAGE, analyzed them by phosphorimaging and calculated activity as the ratio (unwound oligonucleotide)/(total substrate).

**Miscellaneous DNA and RNA manipulations.** Total cellular RNA isolation used TRIzol reagent (Life Technologies) according to the manufacturer's recommendations. Poly(A) mRNA isolation used the Oligotex mRNA isolation kit (Qiagen). We synthesized cDNA with random hexamers in combination with M-MLV reverse transcriptase (GibcoBRL) or with the first-strand cDNA synthesis kit (Amersham Pharmacia) according to the manufacturer's recommendations. We synthesized northern hybridization probes by random-primed labeling of full-length cloned cDNA in the presence of [α-$^{32}$P]dCTP (Amersham; 3,000–6,000 Ci/mmol) and hybridized them to the human eight-lane MTN blot (Clontech) in ExpressHyb solution according to the manufacturer's instructions. The manufacturer provided the β-actin probe. After hybridization, we autoradiographed membranes at –70 °C overnight. PCR amplification of Twinkle genomic or cDNA used routine protocols. Genomic oligonucleotides were as follows: for exon 1, Tw-5′ UTR2 (GTTTG-GTCTAGTGAAGGCACG) and Tw-In1R (CCCACTTGCTTTTGTCAC CTG); for exon 2, Tw-In1F (GTCTTGGTTTCAAGGGTAGG) and Tw-In2R (GATATGTCTGGGAAAGCAAGG); for exons 3+4, Tw-In2F (GGTGGTC-TAGAGACAACTTG) and Tw-In4R (GGACAGTCAAGACGATTAAGG); for exon 5, Tw-In4F (CTTTCTGCTTTGCTCATGTCC) and Tw-3′ UTR1 (CCTTGCAGAGTTTTATGCTCC). We column-purified PCR products for genomic sequencing (Qiagen PCR purification kits) and used cycle sequencing with the same oligonucleotides as for PCR plus the following additional exon 1-specific oligonucleotides: Tw-Ex1-1 (AGTGGATGGGTCG-GAGGGGC), Tw-Ex1-2 (GCAGAGAAAGTGGCCTGTGG), Tw-Ex1-3 (CCAGAATTTGAGGACAGCGAG) and Tw-Ex1-4 (CTTACCCCCTGCCT-TACTCC). Sequencing used dye-terminator chemistry on Perkin Elmer ABI 310, 377 or 3100 instruments, with kit reagents supplied by the manufacturer. We amplified full coding-region cDNA with oligonucleotides *Eco*RI/Tw-1F (CCGGAATTCTAGGAATGTGGGTCCTCCTCC) and *Hin*dIII/Tw-1R (CCC AAGCTTCTTTGAACGCTTGGAGGTGTCTG) or *Sac*II/Tw-1R (AGCCGC

GGCTTTGAACGCTTGGAGGTGTC). We cloned cDNAs with the appropriate restriction enzymes in vector pEGFP-N1 (Clontech), pcDNA3.1(–)/MycHisA (Invitrogen) or pCMV-Tag4 (Stratagene) by standard procedures. Complete sequencing of the insert and flanking sequences verified all constructs.

**Genotyping.** We identified all disease mutations by direct DNA sequencing of DNA samples on both strands. To detect the 39-bp duplication in the Finnish family, the 3′ end of exon 1, we amplified from the genomic DNA with primers Tw-duplF (ACGATGCTTCTTGGTGCGACC) and Tw-duplR (CGAAGCTGCCGGAAAGATACG), using Dynazyme II DNA polymerase (Finnzymes, Finland) according to the manufacturer's instructions. The duplication altered the product size, visualized on 1.5% agarose gels, from 138 to 177 bp. To exclude it as a common polymorphism, we analyzed 40 pooled DNA samples of 10 control subjects each. The G-to-C mutation at position 1423 creates a new recognition site for *Hae*III. We amplified a 413-bp exon 2 PCR fragment with oligonucleotides Tw-ex2-1F (CTCCTCACC CAGGTCTGTTC) and Tw-ex2-1R (CCCTGCCCTCTCATTCTTTG). Digestion of the wildtype product of 413 bp gave two fragments of 378 and 35 bp and mutant heterozygotes gave extra fragments of 201 and 177 bp. G1061C creates an *Ava*II site by using the diagnostic oligonucleotide (CTGGAGGCCCTGACGGAGGCTTCAATCTT*GG*TC; deviations from wild type are in underlined italics) and Tw-In1R (see above). *Ava*II digestion of the resulting 283-bp PCR product gives two fragments of 249 and 34 bp only for the mutant allele. T1142C creates an *Ava*II site by using the diagnostic oligonucleotide (TTCCGGCAGCTTCGGGAGGAGGTGCTAG GAG*G*AC) and Tw-In1R. *Ava*II digestion of the resulting 203-bp PCR fragment gives two fragments of 171 and 32 bp only for the mutant allele. We used similar procedures to detect other mutations. Agarose or polyacrylamide gel electrophoresis visualized all products.

**GenBank accession numbers.** Twinkle, AF292004; Twinky, AF292005; *C10orf2*, AL133215. *P. chabaudi chabaudi* POM1, PDB AAA84746, *C. elegans* hypothetical protein F46G11.1, PDB T16375; *D. melanogaster* CG5924 gene product, PDB AAF52820. A possible Twinkle homolog can also be found in the fission yeast *S. pombe* (see Sanger Centre database entry SPCC830.03).

1. Foury, F. Cloning and sequencing of the nuclear gene *MIP 1* encoding the catalytic subunit of the yeast mitochondrial DNA polymerase. *J. Biol. Chem.* **264**, 20552–20560 (1989).
2. Iyengar, B., Roote, J. & Campos, A.R. The tamas gene, identified as a mutation that disrupts larval behavior in *Drosophila melanogaster*, codes for the mitochondrial DNA polymerase catalytic subunit (DNApol-gamma125). *Genetics* **153**, 1809–1824 (1999).
3. Lim, S.E., Longley, M.J. & Copeland, W.C. The mitochondrial p55 accessory subunit of human DNA polymerase gamma enhances DNA binding, promotes processive DNA synthesis, and confers N-ethylmaleimide resistance. *J. Biol. Chem.* **274**, 38197–38203 (1999).
4. Clayton, D.A. Replication of animal mitochondrial DNA. *Cell* **28**, 693–705 (1982).
5. Holt, I.J., Lorimer, H.E. & Jacobs, H.T. Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA. *Cell* **100**, 515–524 (2000).
6. Lecrenier, N., Van Der Bruggen, P. & Foury, F. Mitochondrial DNA polymerases from yeast to man: a new family of polymerases. *Gene* **185**, 147–152 (1997).
7. Kornberg, A. & Baker, T.A. *DNA Replication* (W.H. Freeman and Company, New York, 1992).
8. Tiranti, V. *et al.* Identification of the gene encoding the human mitochondrial RNA polymerase (h-mtRPOL) by cyberscreening of the Expressed Sequence Tags database. *Hum. Mol. Genet.* **6**, 615–625 (1997).
9. Zeviani, M. *et al.* An autosomal dominant disorder with multiple deletions of mitochondrial DNA starting at the D-loop region. *Nature* **339**, 309–311 (1989).

10. Zeviani, M. *et al.* Nucleus-driven multiple large-scale deletions of the human mitochondrial genome: A new autosomal dominant disease. *Am. J. Hum. Genet.* **47**, 904–914 (1990).

11. Suomalainen, A. *et al.* Multiple deletions of mitochondrial DNA in several tissues of a patient with severe retarded depression and familial progressive external ophthalmoplegia. *J. Clin. Invest.* **90**, 61–66 (1992).

12. Melberg, A., Lundberg, P.O., Henriksson, K.G., Olsson, Y. & Stalberg, E. Muscle-nerve involvement in autosomal dominant progressive external ophthalmoplegia with hypogonadism. *Muscle Nerve* **19**, 751–757 (1996).

13. Li, F.Y. *et al.* Mapping of autosomal dominant progressive external ophthalmoplegia to a 7-cM critical region on 10q24. *Neurology* **53**, 1265–1271 (1999).

14. Moslemi, A.-R., Melberg, A., Holme, E. & Oldfors, A. Clonal expansion of mitochondrial DNA with multiple deletions in autosomal dominant progressive external ophthalmoplegia. *Ann. Neurol.* **40**, 707–713 (1996).

15. Santorelli, F.M. *et al.* Multiple mtDNA deletions: clinical and molecular correlations. *J. Inherit. Metab. Dis.* **23**, 155–161 (2000).

16. Bohlega, S. *et al.* Multiple mitochondrial DNA deletions associated with autosomal recessive ophthalmoplegia and severe cardiomyopathy. *Neurology* **46**, 1329–1334 (1996).

17. Suomalainen, A. *et al.* An autosomal locus predisposing to deletions of mitochondrial DNA. *Nature Genet.* **9**, 146–151 (1995).

18. Kaukonen, J. *et al.* A third locus predisposing to multiple deletions of mtDNA in autosomal dominant progressive external ophthalmoplegia *Am. J. Hum. Genet.* **65**, 256–261 (1999).

19. Kaukonen, J. *et al.* Role of adenine nucleotide translocator 1 in mtDNA maintenance. *Science* **289**, 782–785 (2000).

20. Nishino, I., Spinazzola, A. & Hirano, M. Thymidine phosphorylase gene mutations in MNGIE, a human mitochondrial disorder. *Science* **283**, 689–692 (1999).

21. Leipe, D.D., Aravind, L., Grishin, N.V. & Koonin, E.V. The bacterial replicative helicase DnaB evolved from a RecA duplication. *Genome Res.* **10**, 5–16 (2000).

22. Lennon, G., Auffray, C., Polymeropoulos, M. & Soares, M.B. The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression. *Genomics* **33**, 151–152 (1996).

23. Bird, L.E., Hakansson, K., Pan, H. & Wigley, D.B. Characterization and crystallization of the helicase domain of bacteriophage T7 gene 4 protein. *Nucleic Acids Res.* **25**, 2620–2626 (1997).

24. Hayashi, J.-I., Takemitsu, M., Goto, Y.-I. & Nonaka, I. Human mitochondria and mitochondrial genome function as a single dynamic cellular unit. *J. Cell Biol.* **125**, 43–50 (1994).

25. Davis, A.F. & Clayton, D.A. In situ localization of mitochondrial DNA replication in intact mammalian cells. *J. Cell Biol.* **135**, 883–893 (1996).

26. Satoh, M. & Kuroiwa, T. Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell. *Exp. Cell Res.* **196**, 137–140 (1991).

27. Spelbrink, J.N. *et al.* In vivo functional analysis of the human mitochondrial DNA polymerase POLG expressed in cultured human cells. *J. Biol. Chem.* **275**, 24818–24828 (2000).

28. Washington, M.T., Rosenberg, A.H., Griffin, K., Studier, F.W. & Patel, S.S. Biochemical analysis of mutant T7 primase/helicase proteins defective in DNA binding, nucleotide hydrolysis, and the coupling of hydrolysis with DNA unwinding. *J. Biol. Chem.* **271**, 26825–26834 (1996).

29. Hehman, G.L. & Hauswirth, W.W. DNA helicase from mammalian mitochondria. *Proc. Natl. Acad. Sci. USA* **89**, 8562–8566 (1992).

30. Li, F. *et al.* Characterization of a novel human putative mitochondrial transporter homologous to the yeast mitochondrial RNA splicing proteins 3 and 4. *FEBS Lett.* **494**, 79–84 (2001).

31. Guo, S., Tabor, S. & Richardson, C.C. The linker region between the helicase and primase domains of the bacteriophage T7 gene 4 protein is critical for hexamer formation. *J. Biol. Chem.* **274**, 30303–30309 (1999).

32. Singleton, M.R., Sawaya, M.R., Ellenberger, T. & Wigley, D.B. Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell* **101**, 589–600 (2000).

33. Patel, S.S. & Hingorani, M.M. Oligomeric structure of bacteriophage T7 DNA primase/helicase proteins. *J. Biol. Chem.* **268**, 10668–10675 (1993).

34. Hingorani, M.M. & Patel, S.S. Cooperative interactions of nucleotide ligands are linked to oligomerization and DNA binding in bacteriophage T7 gene 4 helicases. *Biochemistry* **35**, 2218–2228 (1996).

35. Ilyina, T.V., Gorbalenya, A.E. & Koonin, E.V. Organization and evolution of bacterial and bacteriophage primase- helicase systems. *J. Mol. Evol.* **34**, 351–357 (1992).

36. Miyakawa, I., Sando, N., Kawano, S., Nakamura, S. & Kuroiwa, T. Isolation of morphologically intact mitochondrial nucleoids from the yeast, *Saccharomyces cerevisiae. J. Cell Sci.* **88**, 431–439 (1987).

37. Kuroiwa, T. *et al.* The division apparatus of plastids and mitochondria. *Int. Rev. Cytol.* **181**, 1–41 (1998).

38. Kaufman, B.A. *et al.* In organelle formaldehyde crosslinking of proteins to mtDNA: identification of bifunctional proteins. *Proc. Natl. Acad. Sci. USA* **97**, 7772–7777 (2000).

39. Robin, E.D. & Wong, R. Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. *J. Cell. Physiol.* **136**, 507–513 (1988).

40. Zhang, H. *et al.* Quantitation of mitochondrial DNA in human lymphoblasts by a competitive polymerase chain reaction method: application to the study of inhibitors of mitochondrial DNA content. *Mol. Pharmacol.* **46**, 1063–1069 (1994).

41. Doublie, S., Tabor, S., Long, A.M., Richardson, C.C. & Ellenberger, T. Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 A resolution *Nature* **391**, 251–258 (1998).

42. Sawaya, M.R., Guo, S., Tabor, S., Richardson, C.C. & Ellenberger, T. Crystal structure of the helicase domain from the replicative helicase-primase of bacteriophage T7. *Cell* **99**, 167–177 (1999).

43. Kajander, O.A. *et al.* Human mtDNA sublimons resemble rearranged mitochondrial genomes found in pathological states. *Hum. Mol. Genet.* **9**, 2821–2835 (2000).

44. Van Goethem, G., Löfgren, A., Dermaut, B., Martin, J-J. & Van Broeckhoven, C. Mutation of *POLG* is associated with autosomal dominant progressive external opthalmoplegia with multiple mtDNA deletions. *Nature Genet.* **28**, 211–212 (2001).

45. Yu, C.E. *et al.* Positional cloning of the Werner's syndrome gene *Science* **272**, 258–262 (1996).

46. Sinclair, D.A. & Guarente, L. Extrachromosomal rDNA circles—a cause of aging in yeast. *Cell* **91**, 1033–1042 (1997).

47. Suomalainen, A. *et al.* Autosomal dominant progressive external ophthalmoplegia with multiple deletions of mtDNA: clinical, biochemical, and molecular genetic features of the 10q-linked disease. *Neurology* **48**, 1244–1253 (1997).

48. Davis, L.G., Dibner, M.D. & Battey, J.F. *Basic Methods in Molecular Biology* (Elsevier Science, New York, 1986).

49. Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685 (1970).

# Twinkle and POLG defects enhance age-dependent accumulation of mutations in the control region of mtDNA

**Sjoerd Wanrooij, Petri Luoma[1], Gert van Goethem[2], Christine van Broeckhoven[2], Anu Suomalainen[1] and Johannes N. Spelbrink***

Institute of Medical Technology and Tampere University Hospital, Tampere, Finland, [1]Programme of Neurosciences and Department of Neurology, Biomedicum-Helsinki, Helsinki University, Helsinki, Finland and [2]Department of Molecular Genetics (VIB8), Flanders Interuniversity Institute for Biotechnology and Born-Bunge Foundation, University of Antwerp, Antwerpen, Belgium

## ABSTRACT

**Autosomal dominant and/or recessive progressive external ophthalmoplegia (ad/arPEO) is associated with mtDNA mutagenesis. It can be caused by mutations in three nuclear genes, encoding the adenine nucleotide translocator 1, the mitochondrial helicase Twinkle or DNA polymerase γ (POLG). How mutations in these genes result in progressive accumulation of multiple mtDNA deletions in post-mitotic tissues is still unclear. A recent hypothesis suggested that mtDNA replication infidelity could promote slipped mispairing, thereby stimulating deletion formation. This hypothesis predicts that mtDNA of ad/arPEO patients will contain frequent mutations throughout; in fact, our analysis of muscle from ad/arPEO patients revealed an age-dependent, enhanced accumulation of point muta-tions in addition to deletions, but specifically in the mtDNA control region. Both deleted and non-deleted mtDNA molecules showed increased point mutation levels, as did mtDNAs of patients with a single mtDNA deletion, suggesting that point muta-tions do not cause multiple deletions. Deletion breakpoint analysis showed frequent breakpoints around homopolymeric runs, which could be a sig-nature of replication stalling. Therefore, we propose replication stalling as the principal cause of deletion formation.**

## INTRODUCTION

Human mitochondrial DNA (mtDNA) is maintained in multiple copies in each mitochondrion, and is present in thousands of copies per cell in most human tissues. mtDNA is organized in protein–DNA complexes often referred to as nucleoids (1,2). Sporadic germline and maternally inherited mutations of mtDNA, including single deletions/duplications and point mutations, are a frequent cause of human disease (3).

A unique group of disorders is associated with mtDNA depletion (4,5) or multiple deletions of mtDNA in somatic cells, but is primarily caused by defects in nuclear genes. The nuclear background of these syndromes, autosomal dominant or recessive progessive external ophthalmoplegia (ad/arPEO) with multiple mtDNA deletions (6,7), and mitochondrial neurogastrointestinal encephalomyopathy (MNGIE) (8,9), is starting to unravel. PEO in its mildest form presents as external eye muscle weakness, eye lid ptosis and skeletal muscle weakness. To date, mutations in three nuclear genes have been linked to this disorder. These include the muscle-, brain- and heart-specific isoform of the adenine nucleotide translocator (*ANT1*) (10), *C10orf2* encoding the mitochondrial helicase Twinkle (11,12), and the mitochondrial DNA polymerase γ gene, *POLG* (13). ArPEO has been associated exclusively with compound heterozygous mutations in *POLG* (13,14). In addition, recessive *POLG* mutations are associated with more severe or complex disorders (15,16).

The most frequently reported mutation in POLG, Y955C, occurs, at the active site of its polymerase domain. *In vitro* analysis of the enzymatic properties of this mutant protein has led to a suggestion that reduced POLG fidelity could precipitate adPEO (17,18). However, POLG proofreading was also shown to be capable of counteracting infidelity resulting from the Y955C mutation in the polymerase motif B. In addition, the Y955C mutation resulted in a 45-fold decreased affinity for dNTPs (17), suggesting that under conditions of nucleotide limitation, replication stalling could occur. This effect might be enhanced at replication of short homopolymeric runs, frequently occurring in mtDNA.

MNGIE is caused by mutations in the gene *ECGF1*, encoding cytoplasmic thymidine phosphorylase (19). This protein is involved in the nucleoside/nucleotide salvage pathway, which is important for nucleotide homeostasis in post-mitotic cells and dividing cells outside the S phase. *ECGF1* mutations in MNGIE are generally loss-of-function mutations resulting in enhanced levels of thymidine and deoxyuridine in, for example, plasma (20,21). Interestingly, in

**Table 1.** Patient description and materials

| Patient | Mutation | Tissue (mtDNA) | Age | Reference |
|---|---|---|---|---|
| 1 | POLG R3P/A467T | Skeletal muscle | 60 | (13) |
| 2 | POLG Y955C | Skeletal muscle | 48 | (13) |
| 3 | POLG Y955C | Skeletal muscle | 39 | (24) |
| 4[a] | POLG N468D/A1105T | Skeletal muscle | 50 | (24) |
| 5[a] | POLG N468D/A1105T | Skeletal muscle | 44 | |
| 6[a] | POLG N468D/A1105T | Skeletal muscle | 34 | |
| 7[a] | 352–364 AA duplication Twinkle | Extraocular eye muscle | 73 | (23) |
| 8[a] | 352–364 AA duplication Twinkle | Skeletal muscle | 60 | |
| 8[a] | 352–364 AA duplication Twinkle | Frontal cortex | 60 | |
| 9[a] | 352–364 AA duplication Twinkle | Skeletal muscle | 33 | |
| 10[a] | 352–364 AA duplication Twinkle | Skeletal muscle | 31 | |
| 11[a] | 352–364 AA duplication Twinkle | Skeletal muscle | 20 | |
| 1L | POLG R3P/A467T | Blood lymphocytes | 60 | (13) |
| 2L | POLG Y955C | Blood lymphocytes | 48 | (13) |
| Controls | | | | |
| Δ1 | Single mtDNA deletion | Skeletal muscle | 60 | |
| Δ2 | Single mtDNA deletion | Skeletal muscle | 48 | |
| Δ3 | Single mtDNA deletion | Skeletal muscle | 35 | |
| Δ4 | Single mtDNA deletion | Skeletal muscle | 40 | |
| Δ5 | Single mtDNA deletion | Skeletal muscle | 5 | |
| 1[a] | None | Skeletal muscle | 32 | |
| 2[a] | None | Skeletal muscle | 36 | |
| 3[a] | None | Skeletal muscle | 51 | |
| 4[a] | None | Skeletal muscle | 53 | |
| 5[a] | None | Skeletal muscle | 55 | |
| 6[b] | None | Skeletal muscle | 72 | |
| 7[a] | None | Skeletal muscle | 32 | |
| 8[a] | None | Skeletal muscle | 55 | |
| 9[a] | None | Skeletal muscle | 62 | |
| 10[a] | None | Skeletal muscle | 66 | |

Age of patients refers to the age at the time of the biopsy.
[a]Patients and controls indicated in Figure 1.
[b]Control 6 is a distant paternal relative of the POLG N468D/A1105T family members. Patients 1–3 and all single deletion cases are unrelated.

cultured skin fibroblasts, this appears to result in site-specific point mutations in mtDNA mostly affecting residues preceding runs of Ts in the *de novo* synthesized strand (21). Most mutations could be explained by a next nucleotide effect and a slippage mechanism in homopolymeric A–T runs. mtDNA deletions are usually not observed in cultured skin fibroblasts, but are prominent in skeletal muscle of MNGIE and PEO patients. Conversely, site-specific point mutations were demonstrated at very low levels in skeletal muscle of the MNGIE patients, while being prominent in fibroblasts (21). Recent deletion mapping of mtDNA of a MNGIE patient showed both perfect and imperfect direct repeats flanking some of the more frequent deletions (22). Microdeletions within the longer imperfect direct repeats were also observed. The data were considered to be most compatible with a homologous recombination mechanism, because that could possibly explain features, such as the microdeletions, via branch migration.

In order to study how multiple mtDNA deletions are provoked in ad/arPEO, we sequenced regions of mtDNA of PEO patients with mutations in either *POLG* or Twinkle, and related the findings to the pathogenic mechanism. The results do not support a generalized decrease in mtDNA replication fidelity in either POLG or Twinkle patients. Rather, deletion breakpoint mapping suggests stalling at regions of difficulty

for DNA polymerases, such as homopolymeric runs and microsatellite-type repeats, and the additional involvement of the replication fork barrier at bp 16 070, as a primary cause of mtDNA deletion formation.

## MATERIALS AND METHODS

### Patient samples

Table 1 describes the patient material used in this study. PEO-affected individuals 1, 2, 6 and 7 and sample preparation were described previously (13,23). Patients 3–6 were described elsewhere (24). Controls Δ1–5 are sporadic patients affected by PEO (Δ1–4) or mild Kearns Sayre syndrome (KSS) (Δ5) but carrying a single mtDNA deletion; all patients were adults at the time of biopsy except for patient Δ5 who was 5 years of age. Healthy controls 1–6 are unaffected relatives of PEO patients 3–5 (Fig. 1A), controls 1–5 being brothers and sisters, thus sharing the same mitochondrial background, and control 6 being a distant paternal relative. Healthy controls, 7–10 are all direct relatives of patients 7–11 all sharing mtDNA transmitted through the same maternal line (Fig. 1B). External eye muscle from patient 6 has been treated as skeletal muscle in the statistical analysis comparing the healthy controls with PEO patients. All the samples have been taken in accordance with the Declaration of Helsinki and with informed consent.
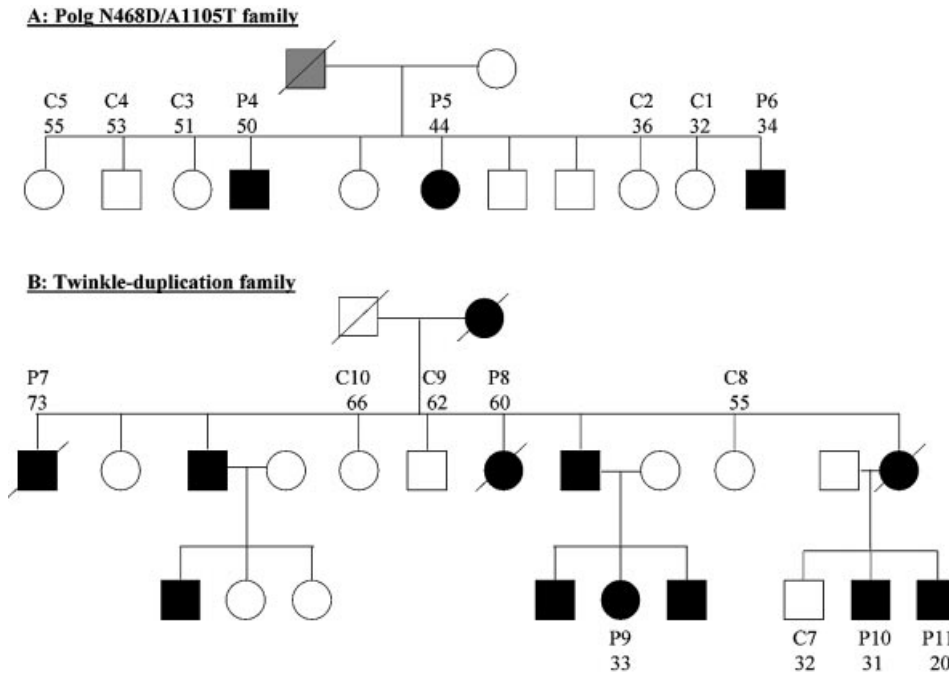
**Figure 1.** Pedigrees of PEO families. (**A**) POLG N468D/A1105T pedigree including patients 3–5 and controls 1–5 (see also Table 1). Black symbols indicate affected family members, white symbols unaffected ones, and the gray symbol indicates unknown disease status of a subject. A slash over a symbol denotes deceased individuals. (**B**) Twinkle dupAA352–364 pedigree including patients 7–11 and controls 7–10. Numbers below patient numbers indicate age at the time of biopsy.

**Table 2.** PCR primers for mtDNA amplification

|  | Sequence 5′→3′ | 5′ start | Location |
|---|---|---|---|
| (a) Short oligonucleotide primers for fragment amplification of the mtDNA control region and cyt b used for sequencing | | | |
| 35L | 5′-GGAGCTCTCCATGCATTTGG-3′ | 35 | Control region D-loop |
| 611H | 5′-CAGTGTATTGCTTTGAGGAGG-3′ | 611 | Control region D-loop |
| 14682L | 5′-CACGGACTACAACCACGACC-3′ | 14 682 | Cyt b |
| 15516H | 5′-GTATAATTGTCTGGGTCGCCTAGG-3′ | 15 516 | Cyt b |
| (b) Primers used for amplification of multiple mtDNA deletions | | | |
| FR31 | 5′-CTTCCCACAACACTTTCTCGGCCTA-3′ | 7177 | COX I |
| 611H/l | 5′-CAGTGTATTGCTTTGAGGAGGTAAGCTACATAA-3′ | 611 | Control region D-loop |
| 15996H | 5′-GCTTTGGGTGCTAATGGTGG-3′ | 15 996 | tRNA^Pro |
| 15516H | 5′-GTATAATTGTCTGGGTCGCCTAGG-3′ | 15 516 | Cyt b |
| 14338H | 5′-GGTGGTTGTGGTAAACTTTAA-3′ | 14 338 | ND 6 |

Numbers refer to the Cambridge reference sequence (25). L and H refer to the strand of the sequence, not to the complementary strand.

## PCR amplification, cloning and sequencing

Parts of the mtDNA control region and cytochrome *b* gene (cyt b) were PCR amplified using the primers shown in Table 2a, yielding a 576 bp (control region) or a 654 bp (cyt b) fragment. Reactions were performed in a 50 µl volume containing 100 ng of total genomic DNA template, 1.5 U of Pfu DNA polymerase (Promega), 250 µM dNTPs, 10 µM of each primer, and the buffer supplied by the manufacturer. The amplification conditions were as follows: initial denaturation of 2 min at 95°C; 21 cycles of 30 s at 95°C, 1 min at 57°C and 2 min at 72°C; final extension step of 10 min at 72°C. mtDNA deletions were PCR amplified using various primer pairs shown in Table 2b, yielding multiple fragments according to the sizes of the deletions. Reactions were as above but with 5%

dimethylsulfoxide, 20 µM of each primer, and the buffer supplied by the manufacturer. The amplification conditions were: initial denaturation of 2 min at 93°C; 29 cycles of 30 s at 93°C, 1 min at 57°C and 7 min at 72°C; final extension step of 15 min at 72°C. Following PCR amplification, control region and cyt b fragments as well as the deletion-containing PCR products were cloned using the TOPO Zero-Blunt PCR cloning kit (Invitrogen). Plasmids from individual bacterial colonies were purified using the Macherey–Nagel Nucleospin Robot-96 plasmid kit on a Tecan multipipetting robot, and verified for the presence of inserted fragments by restriction digestion and agarose gel electrophoresis. The plasmids carrying the inserts were sequenced on an ABI Prism 3100 DNA sequencer, using BigDye terminator chemistry (Applied Biosystems) and M13 forward and reverse oligonucleotides as

well as mtDNA-specific oligonucleotides. Sequences were analyzed using the SeqMan™II program of DNASTAR. Differences from the Cambridge reference sequence (25) present in all clones, and for patients 1 and 2 present in leukocyte mtDNA, were discarded as polymorphisms. In addition, the control region homopolymeric D310 tract was deliberately excluded from our analysis since it is highly polymorphic (see http://www.mitomap.org/). Rare 1–2 bp insertions or deletions are included with point mutation counting, but the small duplications identified in patients 1, 3 and 4 were not. Repeated PCR cloning and sequencing of one of the samples showed good reproducibility. Full-length and deleted mtDNA sequence features such as base composition, nucleotide word counts, palindrome and direct repeat searches were performed with the Java interface (J)EMBOSS package (26) at http://www.hgmp.mrc.ac.uk/Software/EMBOSS/Jemboss/.

## RESULTS

### High mutation levels in the control region of muscle mtDNA from PEO patients

To determine the point mutation levels of mtDNA from ad/arPEO patients with multiple deletions, we sequenced mtDNA from PEO patients' tissues, healthy controls and patients with single sporadic deletions, followed by cloning and sequencing of individual plasmids, and comparison with the Cambridge reference sequence to score mutations. It is important to note here that all of the healthy controls except control 6 were close relatives of the patients 3–5 and 7–11 (see Fig. 1 and Table 1 for family pedigrees and additional patient information). Therefore, their maternally inherited mtDNA was similar, and the possible differences in mtDNA seen between controls and patients are due to the inherited nuclear factors.

We analyzed two regions of mtDNA, a fragment of the control region that includes part of the D-loop [nucleotides 35–611 of the Cambridge reference sequence (25)] and a fragment of the cyt b gene (nucleotide 14 682–15 516). The mutation level was defined as the number of different mutations per 10 kb of sequenced cloned fragments of the control region. This is a measure of sequence variability, but not necessarily a strict measure of mutation burden since repeated clones with an identical mutation are not counted separately (see also below). The Twinkle/PEO patient samples included several relatively young individuals, one of whom, patient 11 (Fig. 1), had not yet presented disease symptoms at the age of biopsy. The biopsy DNA sample had, however, shown weak multiple mtDNA deletions by Southern blot and clear deletions by long-range PCR (23). The youngest patients had the lowest levels of control region point mutation of the patient samples, with levels being similar to those of elderly controls (Fig. 1). With increasing age ($\geqslant$35 years), control region mtDNA of PEO patients showed a significantly higher mutation level ($P < 0.0002$) than that of the control samples (>35 years) (Fig. 2A, B and G). The mutation level in patients ranged from 1.6 (patient 11) to 6.3 (patient 7) mutations/10 kb, whereas in the 10 healthy control individuals of varying ages, the level was lower than 2.2 mutations/10 kb (controls 1–10), closer to the PCR-induced mutation background for Pfu polymerase (27). A weak tendency to higher mutation levels

with increasing age was observed in the controls (Fig. 2H), while the ad/arPEO patients showed a strong increase of control region mutations during aging (Fig. 2H). Three additional controls, from patients with a sporadic single mtDNA deletion, showed mutation levels of 4.3, 2.1 and 1.7 mutations/10 kb, which is higher than the levels in controls of a similar age (Fig. 2H). Point mutation levels in leukocyte DNA of two PEO patients (1L and 2L) were much lower than in muscle DNA, comparable with controls.

Analysis of the sequence data revealed that mutations found in the control region were not randomly distributed and recurred in several clones from the same patient, suggesting that a considerable proportion of mtDNAs carried these mutations (Fig. 3). The amount of mutant A189G was verified by Genescan analysis and found to be in good agreement with the levels based on the frequency of the mutation in the cloned fragments (data not shown). The total mutation load was calculated by counting the total number of all clones of a sample with one or more mutations. This is a direct measure of the percentage of the total mtDNA population with one or more point mutations (Fig. 2C and D). This analysis of total mutant load showed that in the tissues of PEO patients, 5% (patient 11) to 96% (patient 2) of mtDNA carried control region mutations (Fig. 2C), whereas in the controls this load was <7% for most samples (Fig. 2D), except for three of the oldest individuals who showed a mutant load of between 18 and 25%. Similar to the mutation level analysis, total mutant load in PEO patients was always higher than that of controls of a similar age. Two of five single deletion patients also showed a higher percentage of mutant fragments than did controls of a similar age. mtDNA from leukocytes of POLG-PEO patients 1 and 2 showed similar values to control muscle mtDNA (2L, 3%; 1L, 6%).

### Comparable low mutation levels in the cyt b region in muscle mtDNA from PEO patients and controls

Contrary to the mtDNA control region, no statistically significant differences were observed in mutation levels between patients and controls in the cyt b region (Fig. 2E and F). The mutation levels in patients ranged from 0.2 (patient 3) to 1.4 (patient 8) mutation/10 kb. In skeletal muscle from the 10 healthy individuals and lymphocyte mtDNA from patient 1, the mutation levels were also within this range. Total mutation loads revealed no difference between PEO patients and controls (data not shown). Partial analyses of other mtDNA regions, COX I (7203–7445) and HVR1 (16 000–16 569), showed levels in PEO patients comparable with those observed for the cyt b region (not shown). Similar results to those in skeletal muscle were obtained in mtDNA control and cyt b regions in brain and extraocular muscle of one of the Twinkle/PEO patients (see also Fig. 2). However, no appropriate controls were available for these two tissues.

### Distribution of heteroplasmic control region mutations

The heteroplasmic control region mutations found in the patients were non-randomly distributed (Fig. 3). In POLG/PEO patients, the majority of mtDNA mutations (86%) accumulated in the region from nucleotide 150 to 250, particularly close to the presumptive heavy strand replication origin $O_H$ and in conserved sequence block I (CSB I). Specific mutation hotspots at the area could be identified (Fig. 3B).
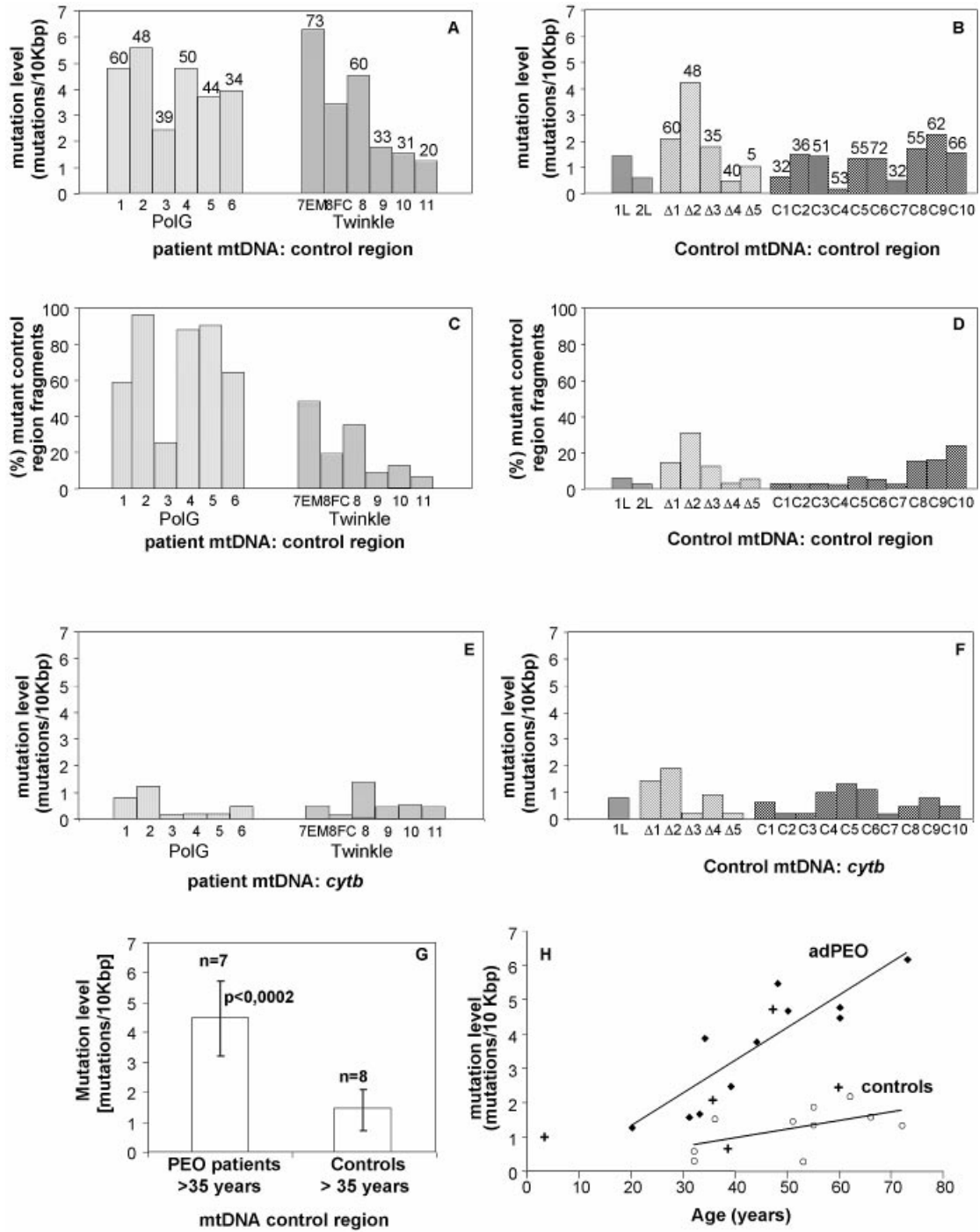
**Figure 2.** Point mutation levels in the control and cyt b region of PEO patient muscle mtDNA. (**A** and **B**) Number of different mutations detected per 10 kb of control region mtDNA in PEO patients (A) and in different types of controls (B). Numbers above the bars indicate the age at the time of biopsy. (**C** and **D**) Number of cloned mtDNA control region fragments carrying at least one point mutation in PEO patients (C) and in controls (D), expressed as a percentage of the total number of clones that were sequenced. (**E** and **F**) Number of different mutations detected per 10 kb of cyt b gene sequence in PEO patients (E) and in controls (F). Patients and abbreviations: 1–6, POLG/PEO patients, skeletal muscle; 7EM, Twinkle/PEO patient, extraocular muscle; 8FC, Twinkle/PEO patient, frontal cortex; 1L, 2L, POLG/PEO patients 1 and 2, leukocytes; Δ1–Δ5, patients with sporadic single mtDNA deletions, skeletal muscle; C1–C10, healthy control samples, skeletal muscle. (**G**) Number of different control region mutations in muscle mtDNAs of PEO patients >35 years old grouped together, and in the >35-year-old controls. The probability (*P*) of both groups being identical was calculated using the unpaired Student's *t*-test. The indicated results show that mutation level differences between both groups are highly significant. (**H**) Mitochondrial control region mutation levels in PEO patients (filled diamonds), healthy controls (open circles) and single deletion controls (+), related to the age at the time of biopsy, based on the values also indicated in (A) and (B). Lines for the PEO patients and healthy controls indicate trend lines.
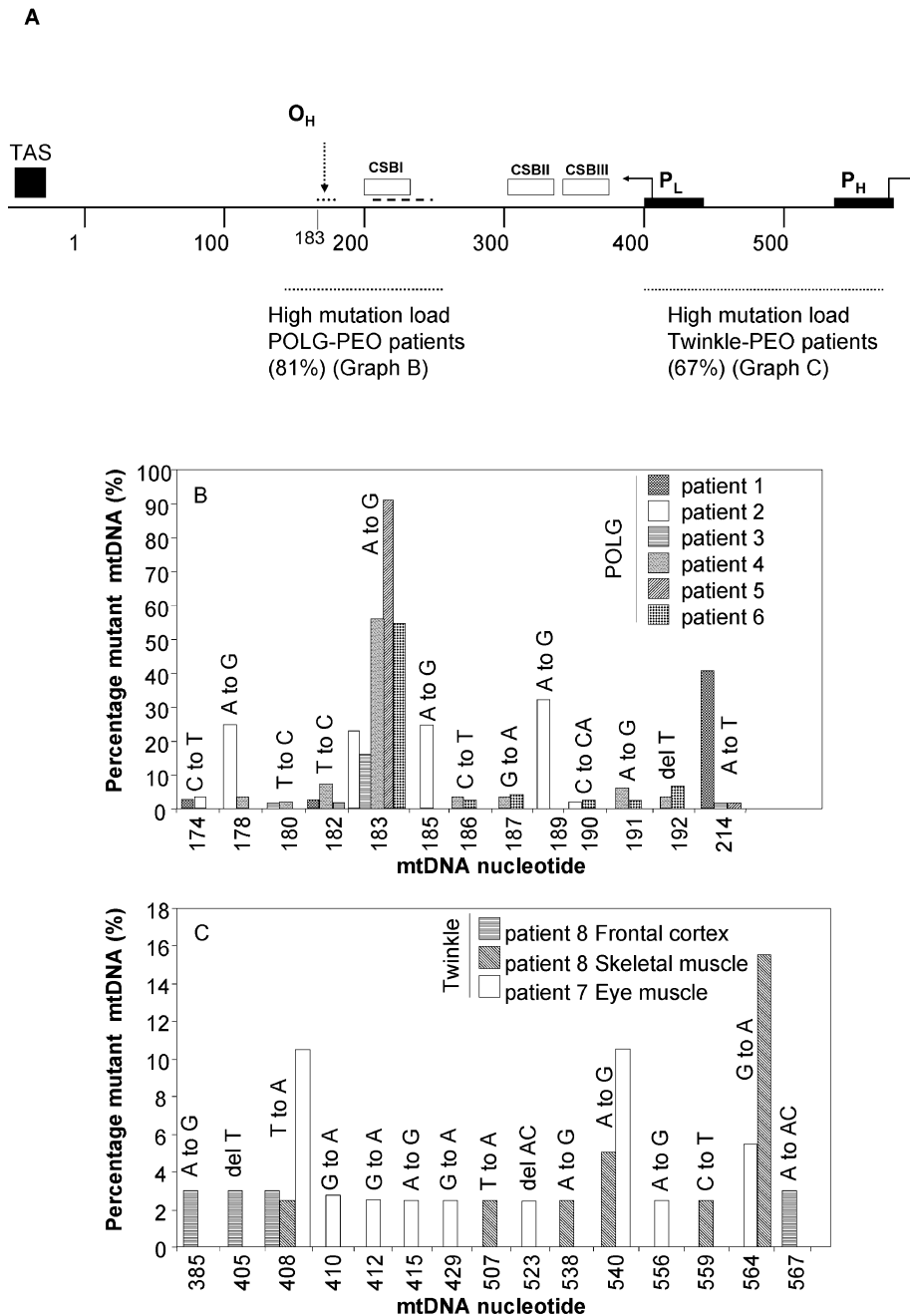
**A**





**Figure 3.** Non-random distribution and heteroplasmy levels of control region mutations. (**A**) Diagram of the mtDNA control region and mutational hotspots. Abbreviations and symbols: $O_H$, heavy strand replication origin; $P_H$ and $P_L$, H and L strand promoters; TAS, termination-associated sequence; CSB I, II and III, conserved sequence blocks I, II and III; bold dotted lines, the regions where most mutations were found in POLG/PEO and Twinkle/PEO patients' muscle mtDNA; percentages in parentheses, the percentage of mutated fragments with mutations in the indicated region; small dotted line adjacent to $O_H$, A to G conversion mutational hotspot region of patient 2; dashed line, position of the duplication hotspot in patient 1. (**B**) Pattern of mutations in the mtDNA replication control region of POLG/PEO patients between nucleotides 174 and 214. (**C**) Mutation pattern in the mtDNA transcription control region of Twinkle patients 7 and 8. The percentage of mutant mtDNA in (B) and (C) indicates the number of clones with the indicated mutation in the indicated patients relative to the total number of clones sequenced for each patient.

Most striking were the mutations found in PEO patients 1 and 2, carrying the recessive POLG R3P and A467T mutations in the compound heterozygous state, and the dominant Y955C POLG mutation, respectively. Patient 1 had several duplications (2–32 bp, Supplementary fig. 1B available at NAR Online), partly overlapping with CSB I. Patients 3 and 4 each also showed one clone with a small duplication

(Supplementary fig. 1B). Patient 2 presented an A to G transition 'hotspot' close to $O_H$ at nucleotides 178, 183 and 189 (~20% heteroplasmy). In addition, based on the leukocyte mtDNA sequence, patient 2 had the less common 185A polymorphism. This polymorphic nucleotide 185 also appeared as an A to G transition hotspot in this patient. Although several clones showed two of these transitions, most

clones contained only one, sometimes in combination with an additional point mutation. The unrelated patient 3, with the same Y955C mutation in POLG, only carried the mtDNA A183G mutation at a high percentage. This change was also detected at levels >15% of total mtDNA in POLG/PEO patients 3, 4, 5 and 6, but was not observed in Twinkle/PEO patients' tissues or controls, nor is it a recognized polymorphism or disease-associated mutation (http://www.mitomap.org/). An aging-associated nucleotide change, A189G (28), was detected at high levels in POLG patient 2 and in one of about 50 clones from each of the two elderly Twinkle patients' samples, but not in any other POLG/PEO patients or the younger Twinkle/PEO patients. This mutation was also frequently found in the control samples, where it represented 53% of all mutated clones of these samples. The T408A mutation represented 16% of all mutated clones in the controls.

The distribution of control region point mutations in the Twinkle/PEO patients (patients 7–11) differed from POLG/PEO patients, since the majority of the Twinkle-associated mutations (67%) were located between nucleotides 385 and 570. In the two elderly Twinkle patients 7 and 8, who showed the highest point mutation levels, this value was 87%. For samples 7EM and 8, of the total number of clones sequenced, 42 and 31% had mutations between 385 and 570. For the POLG patients, these values were typically around 5%, with the exception of patient 1 who had 24% of all clones with a mutation between 385 and 570. Individual point mutations were generally present at low levels (note the difference in scale between Fig. 3B and C), but the number of different mutations, at least for the oldest Twinkle patients, was nevertheless comparable with POLG/PEO patient mutation levels (compare Fig. 2A and C). The 564 G to A transition mutation, which was seen in the highest percentage of clones in patients 7 and 8, was detected in all Twinkle/PEO patients, including patients 9–11, but only in the POLG patient 1 and one of the elderly controls. The above data are not skewed by large differences in the number of sequenced clones or differences in total number of bases that were sequenced: from each sample for each region, approximately 50 clones were sequenced resulting in ~30 kb of sequence, except for a young Twinkle/PEO patient 9, of whom we sequenced ~150 kb. Because point mutations in the samples from the young Twinkle patients are rare and mutation levels are only marginally higher than those of controls of a similar age, the analysis of regional variation could not be done with confidence. Nevertheless, the analysis of patient 9 showed that she had already had several of the same mutations observed in her older relatives with PEO. Since her maternally inherited mtDNA came from her unaffected mother, this suggests that the occurrence of these mutations at young age are associated with the nuclear disease allele. The sporadic patients with a single mtDNA deletion (∆1–5) showed a distribution of control region point mutations similar to the Twinkle/PEO patients.

## Comparison of mutation load in deleted and non-deleted mtDNA molecules

The low mutation load in the cyt b gene suggested close to normal mutant POLG fidelity (see Discussion). To study whether control region point mutations predispose the

molecules to deletion formation, we specifically amplified deleted mtDNA molecules including the control region. We compared control region mutation levels in these deleted molecules with the levels in the total population of mtDNA molecules obtained by the amplification and sequencing of just the control region, as already shown in Figure 2. The outermost PCR primers used (Table 2b) create an ~10 kb fragment using wild-type mtDNA and the appropriate PCR conditions (not shown). However, a short extension time was used in the PCR to promote preferential amplification of deleted mtDNA molecules. In all samples, a set of products generally shorter than 4 kb was amplified (Fig. 4A), cloned and the control region sequenced. Figure 4B and C shows the mutation levels and total mutational loads in deleted molecules compared with the values for the total population of mtDNA molecules (as shown in Fig. 2) for patients 1–3, 5, 6 (POLG/PEO) and 7 (Twinkle/PEO). No significant POLG- or Twinkle-specific differences were observed.

## Mapping of multiple mtDNA deletions

Another prediction of strand slippage, resulting from decreased replication fidelity, is the occurrence of direct repeats flanking the deletions. In addition, with the current knowledge of the patients' nuclear gene mutations, we asked whether different mutations in one gene such as *POLG* or mutations in different genes, such as *POLG* and *C10orf2* (Twinkle), result in systematic differences in deletion breakpoints, possibly pointing to different mechanism(s) of deletion formation. Finally, we questioned the hypothetical relationship between particular control region point mutations and specific mtDNA deletions. For these purposes, we mapped the breakpoints of mtDNA deletions from the cloned fragments of deleted molecules that were also used to determine the mutation levels in the control region of the deleted molecules (see Table 3 and Supplementary fig. 1). Many breakpoints were located in the mitochondrial genes from cyt b to COX I. Using the outermost primer pair (611H/l-FR31), which also amplified the control region, a major 3′ (L-strand orientation) deletion breakpoint was identified around bp 16 070 at or near the 'termination-associated sequence' (TAS). This breakpoint was identified in the first patients identified with adPEO (6), most of whom carried the POLG Y955C mutation (14). Our results indicate that bp 16 070 is a persistent breakpoint in both POLG and Twinkle/PEO patients. The breakpoint is rarely observed in sporadic cases with single mtDNA deletions, but is frequently encountered in so-called 'sublimons' from control tissues (29). However, with the PCR primers used here, these molecules are probably over-represented since they are short and preferentially amplified in PCR. In addition, the ligation–cloning procedure used here strongly favors the shortest PCR products, thus enriching for the largest mtDNA deletions in the derived bacterial clones. In fact, the shortest fragments visible in our deletion PCR (Fig. 4A) from ~1200–1300 bp very probably represent those molecules with breakpoints at around 16 070 at the 3′ end and close to 7177 at the 5′ end. To avoid the strong prevalence of the 16 070 breakpoint, we used several primer pairs for more extensive mapping of deletions from patient 1. In this patient, breakpoints of deletions in different regions of the mitochondrial genome showed the same general features. We detected short or imperfect direct repeats at most breakpoints (Table 3 and
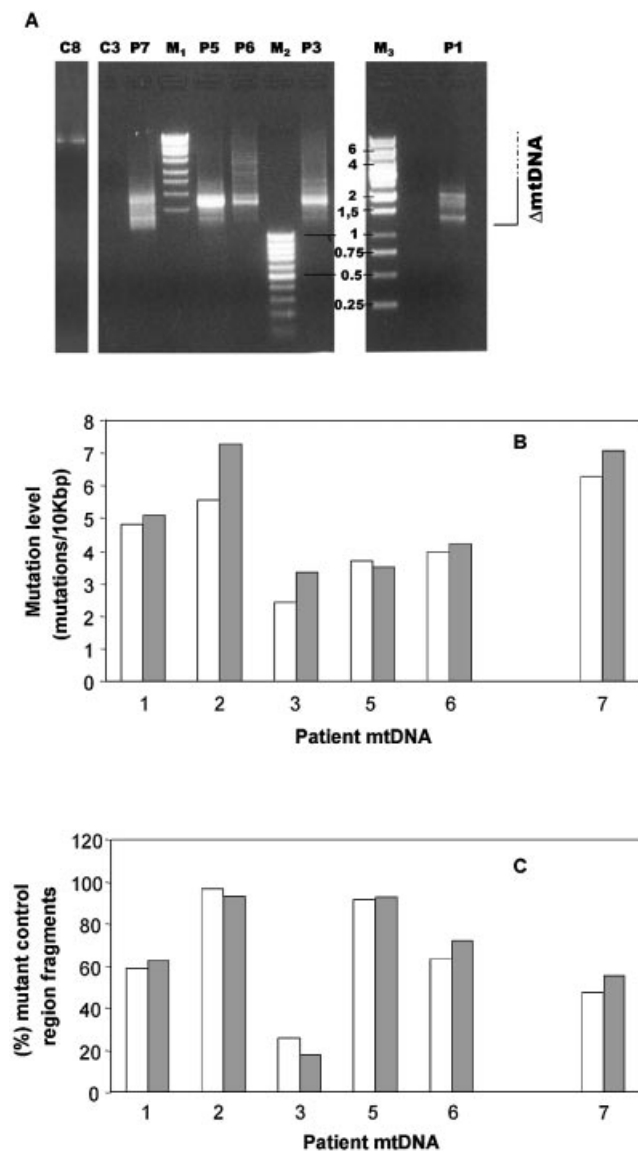
**Figure 4.** Detection of multiple mtDNA deletions in PEO patients by long-range Pfu PCR and comparison of control region mutation levels between the deleted and the total mtDNA population. (**A**) DNA fragments amplified from muscle mtDNA of patients 1, 3, 5, 6 and 7 and control muscle mtDNA were separated on a 1% agarose gel and stained with ethidium bromide. Abbreviations: M1, M2 and M3, DNA size markers indicated in kb; C3 and C8, controls 3 and 8. Deletion PCR was performed as described in Materials and Methods using primers FR31 and 611H/l (Table 2b), amplifying the mtDNA region between nt 7177 and 611. Note that the PCR from C8 did show a faint band corresponding to the full-length mtDNA fragment. (**B**) The number of point mutations in the control region detected in deleted mtDNA molecules (gray bars) and in the total population of mtDNA molecules as determined in Figure 1 (white bars) of patients 1–3 and 5–7. (**C**) The number of mtDNA fragments carrying a mutation in skeletal muscle mtDNA of patients 1–3 and 5–7.

Supplementary fig. 1A) and occasional imperfect palindromic sequences (data not shown). Most notably, frequent short homopolymeric runs of four or more nucleotides were seen close to the breakpoint. Several breakpoints showed microsatellite-like sequences (Supplementary fig. 1A). Allowing one different nucleotide in a run of six nucleotides, 111 out of a total of 142 deletions had runs within 10 nucleotides of the

breakpoint at one or both ends. We analyzed the occurrence of homopolymeric runs at the 'deletion' region (bp 7177–16 100), as defined by the region we amplified with the outermost PCR primer pair and assuming no deletion breakpoints occur beyond the 16 070 region. Over the entire length of the 7177–16 100 region, 816 bp are in randomly distributed perfect homopolymeric runs of four or more nucleotides. Based on this, the predicted frequency of randomly occurring deletions within or at the precise boundary of a homopolymeric run is 1:9.0, assuming that a run of four nucleotides has five positions at which the break can occur. Analysis of the deletion boundaries of the patient samples showed that 60 of 284 3′ or 5′ breakpoints are at the exact end or within a perfect homopolymeric run of four or more nucleotides, indicating a frequency of 1:4.8. When excluding all the boundaries located at the PEO deletion hotspot area around mtDNA bp 16 070 from this analysis, 60 out of 186 breakpoints are within or precisely flanked by a homopolymeric run, giving a frequency of 1:3.1, which is 3-fold higher than predicted by random deletion occurrence. The bias towards deletions in the vicinity of homopolymeric runs would have been even more pronounced had we included imperfect but long homopolymeric runs, such as the *cgccctccc* at nt 7815 (28 of 142 5′ breakpoints). The most frequently encountered 5′ breakpoints around nt 7400 (30 of 142 5′ breakpoints) included in the calculation occurred within a sequence stretch *cccccaccc*. This is a perfect 6 nt homopolymeric run and a 9 out of 10 nt imperfect run. No association of particular control region point mutations with any specific mtDNA deletions were detected (data not shown). No obvious differences were observed in types of deletion breakpoints between Twinkle/PEO and POLG/PEO patients (Supplementary fig. 1). Finally, no point mutations were observed in the deletion boundaries of any of the 142 deletions we sequenced (data not shown).

## DISCUSSION

Here, we catalog mtDNA point mutations accumulated in PEO patients' post-mitotic tissues because of POLG and Twinkle defects. Our results strongly suggest that the mutation mechanism in the two pathological conditions is closely related but is not caused by generalized replication infidelity as previously suggested. Instead, we present evidence for the hypothesis that multiple deletion formation in PEO is initiated by frequent replication stalling.

A recent study (17) showed that *in vitro*, the Y955C POLG mutation affected the accuracy of mtDNA synthesis, resulting in enhanced point mutation levels caused by misinsertion. By analogy with mutations in the synonymous amino acid in the Klenow fragment of *Escherichia coli* DNA polymerase I, POLG Y955 is an important residue for base discrimination and binding [(18) and references therein]. Misinsertions following direct repeats were suggested to promote slipped mispairing and hence deletion formation. On the other hand, the same study showed that the enhanced misinsertion rate was counteracted by the proofreading capabilities of the POLG exonuclease domain. In order to assess the fidelity of POLG in PEO *in vivo*, we established the point mutation levels in the tissues of the patients carrying either POLG or Twinkle mutations. In addition, we mapped the breakpoints of 142 ad/arPEO-associated mtDNA deletions in order to analyze: (i) the

**Table 3.** Statistics of deletion mapping

| Deletion PCR | Mapped deletions | Flanking repeat length | | | Frequent breakpoint elements | | |
|---|---|---|---|---|---|---|---|
| | | >5 bp | >6 bp | >7 bp | (3′) 16 070 | (5′) 7400 cccccacc | (5′) 7815 cgccctccc |
| FR31–611H/l (7177–611) | 105 | 29 | 0 | 0 | 98 | 20 | 28 |
| FR31–15 996H (7177–15 996) | 10 | 5 | 0 | 0 | NA | 0 | 0 |
| FR31–15 516H (7177–15 516) | 14 | 2 | 2 | 0 | NA | 4 | 0 |
| FR31–14 338H (7177–14 338) | 13 | 1 | 1 | 0 | NA | 6 | 0 |
| Total | 142 | 37 | 3 | 0 | | 30 | 28 |

Multiple mtDNA deletions were amplified using the indicated primer pairs, yielding PCR products of varying length, as shown for example in Figure 4A. The PCRs were subsequently used for direct bacterial cloning, with random clones picked for sequence analysis (see Materials and Methods for details). Primer pair FR31–611H/l was used on DNA from patients 1–3 and 5–7, while the other primer pairs were only used on DNA from patient 1. Detailed mapping of the deletions is shown in Supplementary figure 1. The table shows the absence of long flanking direct repeats in this series of deletions, while the most frequent 5′ or 3′ breakpoint elements are indicated (see Results and Discussion for more details).

occurrence and types of direct repeats; (ii) a possible correlation between point mutation and deletion occurrence; and (iii) a possible correlation between disease genotype and mtDNA genotypes. The results lead to the following conclusions: (i) POLG and Twinkle PEO patients, as well as three out of five patients with single sporadic mtDNA deletions, show increased levels of point mutations in the mtDNA control region compared with controls of a similar age, but not in the cyt b-encoding gene region, strongly suggesting that the PEO patients do not have a generalized decrease in replication/ repair accuracy. (ii) The number and types of point mutations in deleted molecules do not differ from the number and type in the total population of mtDNA molecules, arguing against the idea that point mutations predispose to deletions. (iii) Deletion breakpoint mapping in skeletal muscle showed no obvious differences between POLG/PEO and Twinkle/PEO patients, suggesting a similar mechanism of deletion formation. (iv) Analysis of deletion-flanking sequences does not show prominent features suggestive of a strand slippage mechanism, such as precisely flanked long (8–13 nt) direct repeats. Instead, the occurrence of frequent deletion breakpoints at known problem sites for DNA polymerases, most notably homo-polymeric runs, suggests that deletions in ad/arPEO might arise following replication stalling.

## Increased mutational load in the mtDNA control region but not in the cyt b region of ad/arPEO skeletal muscle

Dominant mutations in POLG cluster in and around the polymerase motif B, while the recessive mutations are mostly not in recognized functional domains, although they are generally closer to the exonuclease domains (13,14,30,31). *In vitro* studies on the consequences of the polymerase B domain Y955C mutation suggest that POLG mutations in PEO generally cause decreased accuracy of the POLG protein during replication and/or repair, resulting in an increased nucleotide misinsertion rate leading to point mutations (17). Furthermore, our previous cell culture model, which over-expressed a mutant POLG lacking proofreading activity [(27); S. Horttanainen, S. Wanrooij, J. Kurkela, M. Jokela, H. T. Jacobs and J. N. Spelbrink, unpublished data), predicts that reduced POLG fidelity causes misinsertion mutations ran-domly throughout the mitochondrial genome. This is now also confirmed for a physiologically more relevant model system, the (mtDNA mutator) mouse (32). However, our present data show that in the muscle of the POLG/PEO patients, this is not the case: point mutations accumulate in the control region, but

not in the coding region. No evidence on other pathological mutations of mtDNA support selection against the mutant mtDNA in post-mitotic tissues, such as muscle, whereas in leukocytes that may be the case (33). Post-mitotic tissues are usually quite resistant to mutant mtDNA, and become affected only when high proportions of mutant mtDNA are reached (3). mtDNA mutator mice indeed do not show any evidence of rapid selection against coding region mutations in the post-mitotic tissues brain, heart and liver (32). On the contrary, point mutation levels at 2 or 6 months of age are ~2-fold higher in the cytochrome *b* gene region compared with the control region. Nucleotide pool imbalance has been implied to predispose to replication errors in the MNGIE syndrome. In this case, point mutations in the coding regions were shown to occur in a wide variety of tissues (21). Thus, the present findings of enhanced point mutation levels in the mtDNA control region, but not in the cyt b gene or other mtDNA regions, refute the idea of increased replication infidelity in ad/arPEO.

The increased level of control region mutations begs the question, cause or effect? If control region point mutations did predispose deletion formation, then they should occur more frequently in deleted molecules. This was not the case; all mtDNA, deleted and full length, showed an increased control region mutation burden. That enhanced levels of control region mutations were also detected in patients with a sporadic single deletion of mtDNA suggests that the mere presence of deleted mtDNA in a cell can induce mutagenesis, without the need for a nuclear gene mutation. The observed low point mutation levels in the control region of young Twinkle/PEO patients, who already had multiple mtDNA deletions and showed disease symptoms, further supports this hypothesis.

A further argument against replication infidelity as a cause of increased point mutation levels is the non-random distri-bution of mutations within the control region, and the observation that several of the mutations we have detected have also been described to accumulate with increasing age [see, for example, Wang *et al.* (28), Michikawa *et al.* (34) and Chomyn and Attardi (35)]. The POLG/PEO patients accumu-late mutations around the heavy strand replication origin ($O_H$) and CSB I (bp 150–250), whereas Twinkle/PEO patients showed more mutation accumulation in the region containing the L-strand and major H-strand promoter, which are import-ant for transcription initiation (36). However, since the Twinkle/PEO patients studied here were related, other Twinkle/PEO families should be studied to confirm this

difference in the mutation distribution pattern from POLG/PEO patients.

$O_H$ and CSB I are important replication initiation sites in the strand-asynchronous model of mtDNA replication. Although this model is now hotly debated (37,38), both $O_H$ and CSB I are conserved features of mammalian mtDNA, suggesting that they do play a regulatory role. $O_H$ and CSB I are also sites of DNA–protein interactions (39). The occurrence and frequency of point mutations in this region suggest that some of these mutated molecules could have a replicative advantage, as previously proposed for the C150T mutation and possibly other mutations that accumulate during aging (40), or are mutational hotspots (41). This might explain the high proportions of some of these mutations.

Two POLG/PEO patients showed either an A to G transition hotspot (patient 2, Y955C) or a duplication mutation hotspot (patient 1, R3P/A467T). The absence of such clear mutation hotspots in the other samples could suggest that the hotspots are related to the specific mutations found in POLG. This was, however, not supported by the fact that two unrelated patients with the Y955C mutation did not share a hotspot. A lower number of individual mutations in patient 3, as well as a considerably lower point mutation load could reflect a difference in lifestyle, genetic background or age.

The A189G mutation was previously shown to accumulate with increasing age (28). In our material, it was usually present at low levels, with the exception of patient 2, who showed multiple A to G transitions near $O_H$ including A189G. In all our other cases, the percentage of mtDNA with A189G was not particularly high, comparable with age-related reported percentages (28). This finding is somewhat different from a report in which patients with multiple mtDNA deletions showed a higher risk for the A189G transition at ages <53 years (42). The novel A183G mutation was recently also identified in three out of eight POLG patients, two of them with a high percentage of mtDNA showing this change (43). In our samples, it was present at 15% or more of total mtDNA in all but one POLG/PEO patient. The absence of this mutation in all controls and in Twinkle/PEO patients suggests that this nucleotide is normally not mutation prone. Although close to $O_H$, the function of the A183 nucleotide has, to our knowledge, not been established.

In a recent paper, Del Bo and co-workers also used a sequencing strategy for mutation analysis in PEO patients (43). In sharp contrast to our results, they failed to detect significant differences between PEO patients carrying mutations in the POLG polymerase motifs or in Twinkle, and control samples. However, in their study, the mutation load in controls with an average age of 45 years was 30%. In our study, it was just 8.2% in controls with an average age of 51 years, three elderly controls contributing significantly to this percentage. The difference probably reflects their use of an error-prone Taq polymerase and high PCR cycle number, which probably has masked this finding [see also Chinnery (44)]. Although our data oppose the idea of replication infidelity as a major problem in the PEO patients studied here, we do not exclude that some PEO patients would show such a defect as a result of a clearly deficient POLG proofreading activity caused by one of more mutations in the conserved exonuclease motifs. This has indeed been observed by Del Bo *et al.* (43), but their generalization that this shows the

predisposing effect of high mtDNA point mutation levels to multiple deletion formation is supported neither by their own nor by our data.

Enhanced control region point mutation levels in some of the sporadic single deletion patients shows that the presence of deleted mtDNA alone is enough to provoke this effect. The most obvious explanation for this is an enhanced production of reactive oxygen species (ROS), caused by a progressive OXPHOS deficiency. Perhaps the best evidence to suggest that mitochondrial dysfunction can result in enhanced ROS generation in mitochondria, which subsequently may result in mtDNA damage, comes from mouse genetic studies (45,46). Why particularly the mtDNA control region would be prone to ROS remains an open question, but perhaps has something to do with the mtDNA topology in that region. For example, single-stranded DNA is possibly as much as a 1000-fold more prone to DNA damage by ROS than double-stranded DNA (47). The POLG and Twinkle PEO mutations could perhaps have an effect on mtDNA topology by stalling at inappropriate sites, in the replication or the transcription control region, respectively, increasing the sensitivity of the region to DNA damage.

### Deletion breakpoint analysis suggests replication stalling as a major event leading to deletions

Single, large mtDNA rearrangements are usually associated with strictly sporadic mitochondrial disorders such as KSS and isolated PEO, and often flanked by relatively long, direct or inverted repeats. Deletions in these sporadic cases previously have been put into two classes. Class I represents deletions with perfect direct repeats that flank the deletion boundaries precisely; class II contains deletions without flanking repeats, with imperfect repeats, with perfect repeats that do not flank the deletions precisely and also with some palindromic sequences (48,49). Notably, the majority of sporadic deletions (~70%) belong to the first category. Patients with single deletions usually also show evidence of partial duplications (50) whereas, to our knowledge, in PEO cases with multiple deletions, duplications have not been observed. Some studies have excluded duplications in multiple deletion cases (23,51), which was taken to suggest that single and multiple deletions might arise by different mechanisms. Our analysis of deletion breakpoints in ad/arPEO and published deletion breakpoints of adPEO patients [e.g. Zeviani *et al.* (6)] suggest that the majority of multiple deletions do not correspond to class I of single sporadic deletions. Most direct repeats we observed were short, often imperfect, and hardly ever flanked the deletions precisely.

By which mechanism could POLG and Twinkle dysfunction cause the late-onset multiple deletion pattern of ad/arPEO? Our data and that from published studies suggest that deletion formation might be induced by replication stalling, which in turn has been shown to induce double-strand breaks (DSBs) (52,53). Replication stalling is supported, first of all, by the 2- to 3-fold over-representation of homopolymeric runs at deletion breakpoints. This can be understood for the POLG Y955C mutation by the reported 45-fold reduction in affinity for the incoming nucleotide (17). Stalling at homopolymeric runs could be explained by frequent reiteration of a single nucleotide, which might result in local depletion of that nucleotide and make it difficult for a deficient polymerase to

incorporate the subsequent nucleotide. It was shown that a mutation of the corresponding Y844 in adenovirus 5 DNA polymerase also results in a dramatic decrease in polymerase activity and DNA-binding capacity (54), supporting a propensity to stalling. Several other dominant POLG mutations affect the polymerase B domain and probably result in similar defects. Such expectations are supported by extensive mutation studies of the orthologous Klenow and T7 polymerase, as well as their crystal structures [for a recent review, see Copeland *et al.* (18)]. In addition to homopolymeric runs, several deletions were flanked by microsatellite types of repeats, which are known sites of replication stalling and genome instability [see, for example, Gordenin and Resnick (55)]. A further finding in favor of the idea of increased DSBs in the case of POLG dysfunction is the higher proportion of restriction enzyme-undigested but 'spontaneously' linearized full-length muscle mtDNA molecules in POLG/PEO patients (including patient 2 of this study) when compared with control muscle mtDNA (56). Finally, the nt 16 070 breakpoint boundary is compatible with a replication stalling model since it is conceived as a fork barrier as it the defines the boundary of the TAS element and could thus be a frequent replication stalling site, as previously suggested (6). This is consistent with it being a conserved protein-binding site, as shown by footprinting analysis (57,58). It is also in line with the recent suggestion that the mtDNA D-loop is a fork arrest region (59). In that model, the TAS region could be the 5′ boundary of this region.

Our data on Twinkle patients' mtDNA suggest that the pathogenetic pathways leading to deletion formation are closely linked. The role of Twinkle as a 5′–3′ helicase was recently suggested (12) and this, together with its high homology to replication helicases, supports its role in mtDNA replication. Since the precise role(s) for Twinkle in mtDNA replication and/or repair is still to be established, the mechanism of deletion formation by defective Twinkle remains speculative. Apart from a direct role in replication fork movement, it could be directly involved in DSB repair, by analogy with, for example, the Werner syndrome helicase (60), or it could play a role in re-initiation of stalled replication forks.

## SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

## ACKNOWLEDGEMENTS

## REFERENCES

1. Satoh,M. and Kuroiwa,T. (1991) Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell. *Exp. Cell Res.*, **196**, 137–140.
2. Garrido,N., Griparic,L., Jokitalo,E., Wartiovaara,J., Van Der Bliek,A.M. and Spelbrink,J.N. (2003) Composition and dynamics of human mitochondrial nucleoids. *Mol. Biol. Cell.*, **14**, 1583–1596.
3. Grossman,L.I. and Shoubridge,E.A. (1996) Mitochondrial genetics and human disease. *Bioessays*, **18**, 983–991.
4. Mandel,H., Szargel,R., Labay,V., Elpeleg,O., Saada,A., Shalata,A., Anbinder,Y., Berkowitz,D., Hartman,C., Barak,M. *et al.* (2001) The deoxyguanosine kinase gene is mutated in individuals with depleted hepatocerebral mitochondrial DNA. *Nature Genet.*, **29**, 337–341.
5. Saada,A., Shaag,A., Mandel,H., Nevo,Y., Eriksson,S. and Elpeleg,O. (2001) Mutant mitochondrial thymidine kinase in mitochondrial DNA depletion myopathy. *Nature Genet.*, **29**, 342–344.
6. Zeviani,M., Servidei,S., Gellera,C., Bertini,E., DiMauro,S. and DiDonato,S. (1989) An autosomal dominant disorder with multiple deletions of mitochondrial DNA starting at the D-loop region. *Nature*, **339**, 309–311.
7. Zeviani,M., Bresolin,N., Gellera,C., Bordoni,A., Pannacci,M., Amati,P., Moggio,M., Servidei,S., Scarlato,G. and DiDonato,S. (1990) Nucleus-driven multiple large-scale deletions of the human mitochondrial genome: a new autosomal dominant disease. *Am. J. Hum. Genet.*, **47**, 904–914.
8. Bardosi,A., Creutzfeldt,W., DiMauro,S., Felgenhauer,K., Friede,R.L., Goebel,H.H., Kohlschutter,A., Mayer,G., Rahlf,G., Servidei,S. *et al.* (1987) Myo-, neuro-, gastrointestinal encephalopathy (MNGIE syndrome) due to partial deficiency of cytochrome-c-oxidase. A new mitochondrial multisystem disorder. *Acta Neuropathol. (Berl.)*, **74**, 248–258.
9. Hirano,M., Silvestri,G., Blake,D.M., Lombes,A., Minetti,C., Bonilla,E., Hays,A.P., Lovelace,R.E., Butler,I., Bertorini,T.E. *et al.* (1994) Mitochondrial neurogastrointestinal encephalomyopathy (MNGIE): clinical, biochemical and genetic features of an autosomal recessive mitochondrial disorder. *Neurology*, **44**, 721–727.
10. Kaukonen,J., Juselius,J.K., Tiranti,V., Kyttala,A., Zeviani,M., Comi,G.P., Keranen,S., Peltonen,L. and Suomalainen,A. (2000) Role of adenine nucleotide translocator 1 in mtDNA maintenance. *Science*, **289**, 782–785.
11. Spelbrink,J.N., Li,F.Y., Tiranti,V., Nikali,K., Yuan,Q.P., Tariq,M., Wanrooij,S., Garrido,N., Comi,G., Morandi,L. *et al.* (2001) Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria. *Nature Genet.*, **28**, 223–231.
12. Korhonen,J.A., Gaspari,M. and Falkenberg,M. (2003) Twinkle has 5′ to 3′ DNA helicase activity and is specifically stimulated by mtSSB. *J. Biol. Chem.*, **278**, 48627–48632.
13. Van Goethem,G., Dermaut,B., Lofgren,A., Martin,J.J. and Van Broeckhoven,C. (2001) Mutation of POLG is associated with progressive external ophthalmoplegia characterized by mtDNA deletions. *Nature Genet.*, **28**, 211–212.
14. Lamantea,E., Tiranti,V., Bordoni,A., Toscano,A., Bono,F., Servidei,S., Papadimitriou,A., Spelbrink,H., Silvestri,L., Casari,G. *et al.* (2002) Mutations of mitochondrial DNA polymerase gammaA are a frequent cause of autosomal dominant or recessive progressive external ophthalmoplegia. *Ann. Neurol.*, **52**, 211–219.
15. VanGoethem,G., Schwartz,M., Lofgren,A., Dermaut,B., Van Broeckhoven,C. and Vissing,J. (2003) Novel POLG mutations in progressive external ophthalmoplegia mimicking mitochondrial neurogastrointestinal encephalomyopathy. *Eur. J. Hum. Genet.*, **11**, 547–549.
16. VanGoethem,G., Martin,J.J., Dermaut,B., Lofgren,A., Wibail,A., Ververken,D., Tack,P., Dehaene,I., Van Zandijcke,M., Moonen,M. *et al.* (2003) Recessive POLG mutations presenting with sensory and ataxic neuropathy in compound heterozygote patients with progressive external ophthalmoplegia. *Neuromusc. Disord.*, **13**, 133–142.
17. Ponamarev,M.V., Longley,M.J., Nguyen,D., Kunkel,T.A. and Copeland,W.C. (2002) Active site mutation in DNA polymerase gamma associated with progressive external ophthalmoplegia causes error-prone DNA synthesis. *J. Biol. Chem.*, **277**, 15225–15228.
18. Copeland,W.C., Ponamarev,M.V., Nguyen,D., Kunkel,T.A. and Longley,M.J. (2003) Mutations in DNA polymerase gamma cause error

prone DNA synthesis in human mitochondrial disorders. *Acta Biochim. Pol.*, **50**, 155–167.

19. Nishino,I., Spinazzola,A. and Hirano,M. (1999) Thymidine phosphorylase gene mutations in MNGIE, a human mitochondrial disorder. *Science*, **283**, 689–692.

20. Spinazzola,A., Marti,R., Nishino,I., Andreu,A.L., Naini,A., Tadesse,S., Pela,I., Zammarchi,E., Donati,M.A., Oliver,J.A. *et al.* (2002) Altered thymidine metabolism due to defects of thymidine phosphorylase. *J. Biol. Chem.*, **277**, 4128–4133.

21. Nishigaki,Y., Marti,R., Copeland,W.C. and Hirano,M. (2003) Site-specific somatic mitochondrial DNA point mutations in patients with thymidine phosphorylase deficiency. *J. Clin. Invest.*, **111**, 1913–1921.

22. Nishigaki,Y., Marti,R. and Hirano,M. (2004) ND5 is a hot-spot for multiple atypical mitochondrial DNA deletions in mitochondrial neurogastrointestinal encephalomyopathy. *Hum. Mol. Genet.*, **13**, 91–101.

23. Suomalainen,A., Majander,A., Wallin,M., Setala,K., Kontula,K., Leinonen,H., Salmi,T., Paetau,A., Haltia,M., Valanne,L. *et al.* (1997) Autosomal dominant progressive external ophthalmoplegia with multiple deletions of mtDNA: clinical, biochemical and molecular genetic features of the 10q-linked disease. *Neurology*, **48**, 1244–1253.

24. Luoma,P., Melberg,M., Rinne,J.O., Kaukonen,J.A., Nupponen,N.N., Chalmers,R.M., Oldfors,A., Rautakorpi,I., Peltonen,L., Majamaa,K. *et al.* (2004) Parkinsonism, premature menopause and mitochondrial DNA polymerase gamma mutations. *Lancet*, in press.

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

26. Rice,P., Longden,I. and Bleasby,A. (2000) EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.*, **16**, 276–277.

27. Spelbrink,J.N., Toivonen,J.M., Hakkaart,G.A., Kurkela,J.M., Cooper,H.M., Lehtinen,S.K., Lecrenier,N., Back,J.W., Speijer,D., Foury,F. *et al.* (2000) *In vivo* functional analysis of the human mitochondrial DNA polymerase POLG expressed in cultured human cells. *J. Biol. Chem.*, **275**, 24818–24828.

28. Wang,Y., Michikawa,Y., Mallidis,C., Bai,Y., Woodhouse,L., Yarasheski,K.E., Miller,C.A., Askanas,V., Engel,W.K., Bhasin,S. *et al.* (2001) Muscle-specific mutations accumulate with aging in critical human mtDNA control sites for replication. *Proc. Natl Acad. Sci. USA*, **98**, 4022–4027.

29. Kajander,O.A., Rovio,A.T., Majamaa,K., Poulton,J., Spelbrink,J.N., Holt,I.J., Karhunen,P.J. and Jacobs,H.T. (2000) Human mtDNA sublimons resemble rearranged mitochondrial genomes found in pathological states. *Hum. Mol. Genet.*, **9**, 2821–2835.

30. VanGoethem,G., Martin,J.J. and Van Broeckhoven,C. (2003) Progressive external ophthalmoplegia characterized by multiple deletions of mitochondrial DNA: unraveling the pathogenesis of human mitochondrial DNA instability and the initiation of a genetic classification. *Neuromol. Med.*, **3**, 129–146.

31. Agostino,A., Valletta,L., Chinnery,P.F., Ferrari,G., Carrara,F., Taylor,R.W., Schaefer,A.M., Turnbull,D.M., Tiranti,V. and Zeviani,M. (2003) Mutations of ANT1, Twinkle and POLG1 in sporadic progressive external ophthalmoplegia (PEO). *Neurology*, **60**, 1354–1356.

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

33. Rahman,S., Poulton,J., Marchington,D. and Suomalainen,A. (2001) Decrease of 3243 A→G mtDNA mutation from blood in MELAS syndrome: a longitudinal study. *Am. J. Hum. Genet.*, **68**, 238–240.

34. Michikawa,Y., Mazzucchelli,F., Bresolin,N., Scarlato,G. and Attardi,G. (1999) Aging-dependent large accumulation of point mutations in the human mtDNA control region for replication. *Science*, **286**, 774–779.

35. Chomyn,A. and Attardi,G. (2003) MtDNA mutations in aging and apoptosis. *Biochem. Biophys. Res. Commun.*, **304**, 519–529.

36. Clayton,D.A. (1992) Transcription and replication of animal mitochondrial DNAs. *Int. Rev. Cytol.*, **141**, 217–232.

37. Bogenhagen,D.F. and Clayton,D.A. (2003) The mitochondrial DNA replication bubble has not burst. *Trends Biochem. Sci.*, **28**, 357–360.

38. Holt,I.J. and Jacobs,H.T. (2003) Response: the mitochondrial DNA replication bubble has not burst. *Trends Biochem. Sci.*, **28**, 355–356.

39. Ghivizzani,S.C., Madsen,C.S., Nelen,M.R., Ammini,C.V. and Hauswirth,W.W. (1994) *In organello* footprint analysis of human mitochondrial DNA: human mitochondrial transcription factor A interactions at the origin of replication. *Mol. Cell. Biol.*, **14**, 7717–7730.

40. Zhang,J., Asin-Cayuela,J., Fish,J., Michikawa,Y., Bonafe,M., Olivieri,F., Passarino,G., De Benedictis,G., Franceschi,C. and Attardi,G. (2003) Strikingly higher frequency in centenarians and twins of mtDNA mutation causing remodeling of replication origin in leukocytes. *Proc. Natl Acad. Sci. USA*, **100**, 1116–1121.

41. Stoneking,M. (2000) Hypervariable sites in the mtDNA control region are mutational hotspots. *Am. J. Hum. Genet.*, **67**, 1029–1032.

42. Del Bo,R., Bordoni,A., Martinelli Boneschi,F., Crimi,M., Sciacco,M., Bresolin,N., Scarlato,G. and Comi,G.P. (2002) Evidence and age-related distribution of mtDNA D-loop point mutations in skeletal muscle from healthy subjects and mitochondrial patients. *J. Neurol. Sci.*, **202**, 85–91.

43. Del Bo,R., Bordoni,A., Sciacco,M., Di Fonzo,A., Galbiati,S., Crimi,M., Bresolin,N. and Comi,G.P. (2003) Remarkable infidelity of polymerase gammaA associated with mutations in POLG1 exonuclease domain. *Neurology*, **61**, 903–908.

44. Chinnery,P.F. (2003) Mitochondrial disorders come full circle. *Neurology*, **61**, 878–880.

45. Esposito,L.A., Melov,S., Panov,A., Cottrell,B.A. and Wallace,D.C. (1999) Mitochondrial disease in mouse results in increased oxidative stress. *Proc. Natl Acad. Sci. USA*, **96**, 4820–4825.

46. Melov,S., Coskun,P., Patel,M., Tuinstra,R., Cottrell,B., Jun,A.S., Zastawny,T.H., Dizdaroglu,M., Goodman,S.I., Huang,T.T. *et al.* (1999) Mitochondrial disease in superoxide dismutase 2 mutant mice. *Proc. Natl Acad. Sci. USA*, **96**, 846–851.

47. Frederico,L.A., Kunkel,T.A. and Shaw,B.R. (1990) A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry*, **29**, 2532–2537.

48. Mita,S., Spadari,S., Moraes,C.T., Shanske,S., Arnaudo,E., Fabrizi,G.M., Koga,Y., DiMauro,S. and Schon,E.A. (1990) Recombination via flanking direct repeats is a major cause of large scale deletions of human mitochondrial DNA. *Nucleic Acids Res.*, **18**, 561–567.

49. Solano,A., Gamez,J., Carod,F.J., Pineda,M., Playan,A., Lopez-Gallardo,E., Andreu,A.L. and Montoya,J. (2003) Characterisation of repeat and palindrome elements in patients harbouring single deletions of mitochondrial DNA. *J. Med. Genet.*, **40**, e86.

50. Schon,E.A. (2000) Mitochondrial genetics and disease. *Trends Biochem. Sci.*, **25**, 555–560.

51. Carrozzo,R., Hirano,M., Fromenty,B., Casali,C., Santorelli,F.M., Bonilla,E., DiMauro,S., Schon,E.A. and Miranda,A.F. (1998) Multiple mtDNA deletion features in autosomal dominant and recessive diseases suggest distinct pathogeneses. *Neurology*, **50**, 99–106.

52. Michel,B., Ehrlich,S.D. and Uzest,M. (1997) DNA double-strand breaks caused by replication arrest. *EMBO J.*, **16**, 430–438.

53. Lewis,L.K. and Resnick,M.A. (2000) Tying up loose ends: nonhomologous end-joining in *Saccharomyces cerevisiae*. *Mutat. Res.*, **451**, 71–89.

54. Liu,H., Naismith,J.H. and Hay,R.T. (2000) Identification of conserved residues contributing to the activities of adenovirus DNA polymerase. *J. Virol.*, **74**, 11681–11689.

55. Gordenin,D.A. and Resnick,M.A. (1998) Yeast ARMs (DNA at-risk motifs) can reveal sources of genome instability. *Mutat. Res.*, **400**, 45–58.

56. Van Goethem,G., Martin,J.-J., Löfgren,A., Dehaene,I., Tack,P., Van Zandijcke,M., Ververken,D., Ceuterick,C. and Van Broeckhoven,C. (1997) Unusual presentation and clinical variability in Belgian pedigrees with progressive external ophthalmoplegia and multiple deletions of mitochondrial DNA. *Eur. J. Neurol.*, **4**, 476–484.

57. Saccone,C., Attimonelli,M. and Sbisa,E. (1987) Structural elements highly preserved during the evolution of the D-loop-containing region in vertebrate mitochondrial DNA. *J. Mol. Evol.*, **26**, 205–211.

58. Madsen,C.S., Ghivizzani,S.C. and Hauswirth,W.W. (1993) Protein binding to a single termination-associated sequence in the mitochondrial DNA D-loop region. *Mol. Cell. Biol.*, **13**, 2162–2171.

59. Bowmaker,M., Yang,M.Y., Yasukawa,T., Reyes,A., Jacobs,H.T., Huberman,J.A. and Holt,I.J. (2003) Mammalian mitochondrial DNA replicates bidirectionally from an initiation zone. *J. Biol. Chem.*, **278**, 50961–50969.

60. Oshima,J., Huang,S., Pae,C., Campisi,J. and Schiestl,R.H. (2002) Lack of WRN results in extensive deletion at nonhomologous joining ends. *Cancer Res.*, **62**, 547–551.

# Expression of catalytic mutants of the mtDNA helicase Twinkle and polymerase POLG causes distinct replication stalling phenotypes

**Sjoerd Wanrooij[1], Steffi Goffart[1], Jaakko L.O. Pohjoismäki[1], Takehiro Yasukawa[2] and Johannes N. Spelbrink[1],***

[1]Institute of Medical Technology and Tampere University Hospital, Tampere, Finland and [2]MRC-Dunn Human Nutrition Unit, Wellcome Trust-MRC Building, Cambridge, UK

## ABSTRACT

**The mechanism of mitochondrial DNA replication is a subject of intense debate. One model proposes a strand-asynchronous replication in which both strands of the circular genome are replicated semi-independently while the other model proposes both a bidirectional coupled leading- and lagging-strand synthesis mode and a unidirectional mode in which the lagging-strand is initially laid-down as RNA by an unknown mechanism (RITOLS mode). Both the strand-asynchronous and RITOLS model have in common a delayed synthesis of the DNA-lagging strand. Mitochondrial DNA is replicated by a limited set of proteins including DNA polymerase gamma (POLG) and the helicase Twinkle. Here, we report the effects of expression of various catalytically deficient mutants of POLG1 and Twinkle in human cell culture. Both groups of mutants reduced mitochondrial DNA copy number by severe replication stalling. However, the analysis showed that while induction of POLG1 mutants still displayed delayed lagging-strand synthesis, Twinkle-induced stalling resulted in maturated, essentially fully double-stranded DNA intermediates. In the latter case, limited inhibition of POLG with dideoxycytidine restored the delay between leading- and lagging-strand synthesis. The observed cause-effect relationship suggests that Twinkle-induced stalling increases lagging-strand initiation events and/or maturation mimicking conventional strand-coupled replication.**

## INTRODUCTION

Human mitochondrial DNA (mtDNA) is a closed circular molecule of $\sim$16.5 kb and was sequenced 25 years ago (1,2). The two strands of mtDNA are denoted as the Heavy-(H)-strand and the Light-(L)-strand on the basis of their mobility in a denaturing caesium chloride gradient.

The strand-asynchronous or strand-displacement model for mammalian mitochondrial DNA replication was first proposed in the early 70s (3). In this model, synthesis of the nascent H-strand starts at a fixed point in the major non-coding region (NCR) of mtDNA denoted $O_H$. $O_H$ was originally defined by mapping the 5′ ends of the so called D-loop and is located on the L-strand upstream of three conserved sequence blocks. Leading-strand (nascent H-strand) synthesis proceeds two-thirds of the way around the molecule, displacing the parental H-strand in the process with mitochondrial single-stranded DNA-binding protein (mtSSB) suggested to provide protection against the action of nucleases and other insults such as reactive oxygen species. Following exposure of the lagging-strand initiation site ($O_L$) synthesis of the nascent L-strand begins (4,5).

More recently, Holt and co-workers proposed two models of mtDNA replication, one a more conventional strand-synchronous theta mode (6–9) where mtDNA replication initiates bidirectionally at various sites across an initiation zone (OriZ). In this case termination occurs at or near $O_H$. The other mode of replication is similar to the strand-asynchronous mode of replication so that the nascent L-strand DNA was suggested also to be synthesized with a considerable delay. Initiation is essentially unidirectional and occurs in the NCR, importantly however RNA is deposited on the displaced H-strand rather than mtSSB, thus forming ribonucleotide incorporation throughout the lagging strand (RITOLS) intermediates, which is a crucial difference from the strand-asynchronous model (10). Although the high levels

of mtSSB (11) could be seen as supporting the strand displacement model, also for example *Escherichia coli* is estimated to have several thousands of molecules of SSB (12) even though it contains a single copy genome and replicates via conventional theta replication. SSB is nevertheless essential, as it would be in mammalian mitochondria, not only at the replication fork but also in repair, recombination and other DNA maintenance processes. Given the various essential functions of SSB, the high levels might simply reflect a cell's precaution to ensure it is readily available.

The RITOLS model requires that the RNA is replaced by DNA to produce a dsDNA lagging-strand. It was shown that the RITOLS replication intermediates (RIs) are prone to RNaseH degradation during mtDNA purification leaving a single-stranded parental H-strand (7), thus producing RIs originally predicted by the strand-asynchronous model. Strand-asynchronous RIs are therefore considered purification/degradation artefacts. In rodent and chick liver and cultured human cells under normal culture conditions RITOLS intermediates are the predominant class (6,9,10). However, in cultured human cells recovering from mtDNA depletion, the majority of the replication intermediates are essentially double-stranded DNA suggesting a switch from the RITOLS replication mode to more conventional theta replication (9). Alternatively, initiation of lagging-strand DNA synthesis occurs more frequently resulting in an increased rate of conversion of RITOLS RIs to dsDNA RIs.

All proteins responsible for mammalian mtDNA maintenance are encoded in the nucleus, translated by cytosolic ribosomes and imported into the mitochondrial compartment. So far, a limited number of proteins has been identified. These include the mitochondrial DNA polymerase gamma (POLG1) and its accessory subunit (POLG2) [see, e.g. (13)], the mitochondrial DNA helicase Twinkle (14,15), mitochondrial single-stranded DNA-binding protein (mtSSB) (16) and various proteins with a more general role in mtDNA maintenance. The POLG holoenzyme, Twinkle and mtSSB can form a minimal mitochondrial replisome capable of genome length DNA synthesis on an artificial template (17). Some of the components of the mitochondrial replisome and transcription machinery show similarity to their counterparts in T-odd bacteriophages suggesting that a T-odd phage ancestor contributed to the early 'mitochondrial' endosymbiosis event (18). For example, Twinkle shows striking similarity to the T7 phage primase/helicase protein gp4 (T7 gp4) (14). The Metazoan primase domain of Twinkle has diverged from the ones of more primitive Eukaryotes and T-odd phages suggesting it has lost its primase function (19).

The genes for Twinkle, POLG1 and more recently POLG2 (20) have been implicated in human diseases. Autosomal dominant (ad) mutations in Twinkle are associated with Progressive External Ophthalmoplegia (adPEO)(14), while a single recessive mutation is associated with infantile onset spinocerebellar ataxia or IOSCA (21). Mutations in POLG1 are associated with a variety of disorders, including dominant and recessive PEO, various types of ataxia, Parkinsonism and the severe

mtDNA depletion Alpers syndrome (see, e.g. (22) and references therein, and http://dir-apps.niehs.nih.gov/polg/).

The catalytic subunit of polymerase gamma, POLG1, is well-characterized biochemically (see (13), bearing similarity with prokaryotic A-type DNA polymerases such as *E. coli* DNA polymerase I and T7 DNA polymerase. Conserved regions include a C-terminal domain responsible for polymerase activity and an N-terminal 3′-5′ exonuclease domain involved in proofreading. Several disease associated POLG1 mutations have been studied using purified recombinant enzyme. These include the common autosomal dominant Y955C and other mutations, which result in a moderate to severe decrease in polymerase activity, reduced nucleotide selectivity or reduced processivity (23).

*In vivo*, the properties of POLG1 have also been partly characterized in yeast and in cultured human cells (24–27). In both cases, expression of a mutant form of the protein deficient in 3′-5′ exonuclease activity results in the accumulation of mtDNA mutations. An exonuclease deficient variant in mouse also results in a mutator phenotype and shows a whole-organism phenotype of reduced lifespan with a variety of tissue specific ageing associated defects (28,29). Expression of an adPEO associated Twinkle mutation in transgenic mice has shown a late onset phenotype with striking similarities to late onset PEO (30).

Although there is a need for further biochemical characterization of POLG1 and Twinkle mutants, and animal models can provide a wealth of information on disease aetiology and pathogenesis, both approaches have their limitations. We therefore chose an alternative approach of inducible expression of wild-type and mutant POLG1 and Twinkle in cultured human cells, allowing us to study protein function and mtDNA replication dynamics *in vivo*. Using this inducible system in combination with two-dimensional neutral/neutral agarose gel-electrophoresis (2DNAGE), we show here that the induced expression of either Twinkle or POLG1 mutants results in distinct replication stalling phenotypes suggesting defined roles for both proteins in mtDNA replication and in particular the frequency of initiation of lagging-strand maturation/synthesis.

## MATERIALS AND METHODS

### Cloning of expression constructs

The full-length cDNA of POLG1 and Twinkle variants were originally cloned in the pcDNA3.1(−)/Myc-His A (Invitrogen, Carlsbad, CA, USA), as previously described (14,27). All constructs were re-cloned in the pcDNA5/FrT/TO vector (Invitrogen) taking advantage of two PmeI restriction sites flanking the multiple cloning sites of the original pcDNA3 vectors and target vector. The resulting fusion proteins contained the sequence of the respective proteins followed by the Myc-His. All resulting plasmid-constructs were confirmed by DNA sequencing.

## Creation and maintenance of stable transfected inducible cell lines

The Flp-In$^{TM}$ T-REx$^{TM}$ 293 host cell-line (Invitrogen), a HEK293 variant containing a Flip recombination site at a transcriptionally active locus, was grown in DMEM medium (Cambrex Bioscience, Walkersville, MD, USA) with 2 mM L-glutamine (Cambrex Bioscience), 10% FCS (Euroclone, Milan, Italy) and 50 µg/ml uridine (Sigma, St. Louis, MO, USA) supplemented with 100 µg/ml Zeocin (Invivogen) and 15 µg/ml Blasticidin (Invivogen) in a 37°C incubator at 8.5% CO$_2$. Two-day prior to transfection cells were split to 10 cm plates and grown to ~80% confluence in medium lacking antibiotics. Cells were co-transfected with TransFectin (Bio-Rad, Hercules, CA, USA) according the manufacturer's protocol with the appropriate pcDNA5/FrT/TO construct (0.4 µg) and pOG44 (Invitrogen; 3.6 µg), a plasmid encoding the Flp-recombinase necessary for targeted stable integration. Six hour later, transfection medium was replaced with regular fresh medium lacking antibiotics. Twenty-four hour after transfection the selective antibiotics Hygromycin (150 µg/ml) (Invivogen) and Blasticidin (15 µg/ml) were added. Selective medium was replaced every 2 days for cell maintenance. All inducible cell lines were created according this method. To induce expression the indicated amount of doxycycline (Sigma) was added to the growth medium, and cells were processed for further analyses. With longer than 2 days induction medium was refreshed every 2 days.

## Western blot analysis

Cell lysates were prepared and analyzed for protein expression by immunoblotting after SDS-PAGE (27). A primary monoclonal c-myc (Roche Molecular Biochemicals, Nutley, NJ, USA) antibody was used for detection of recombinant proteins. Peroxidase-coupled secondary antibody horse-anti-mouse was obtained from Vector Laboratories, Burlingame, KS, USA. Enhanced Chemiluminescence detection was done essentially as described (27).

## Quantitative PCR

The copy number of mitochondrial DNA per cell was determined by real time PCR of *cytochrome b* using the gene for amyloid precursor protein *APP* as a nuclear standard as described (31). Briefly, crude nucleic acid extracts were obtained from cells by lysis, proteinase K digest and subsequent isopropanol precipitation, and copy numbers of *cytochrome b* and *APP* were determined in a duplex Taqman PCR on an Abiprism 7000 (Applied Biosciences, Foster City, CA, USA) using pCR 2.1-TOPO (Invitrogen) vectors containing the *cytochrome b* and *APP* amplicon as standards.

## Immunocytochemistry

Immunofluorescent detection was done essentially as described previously (32). For the detection of mtDNA we used a monoclonal anti-DNA antibody AC-30-10 (PROGEN, Shingle Springs, CA, USA) as described

previously (33). Secondary antibodies were anti-mouse IgG-Alexa Fluor®488 (Invitrogen; Myc) and anti-mouse IgM-Alexa Fluor®568 (DNA). Image acquisition using confocal microscopy was done as described (32).

## Protein isolation and helicase assays

In vitro assays for determination of helicase activities were performed with highly enriched Twinkle preparations derived from 293 Flp-In$^{TM}$ T-REx$^{TM}$ cells. The cells were induced with 50 ng/ml doxycycline (Sigma) for 2 days, harvested and mitochondria isolated by hypotonic lysis and differential centrifugation (32). The mitochondrial pellet obtained was lysed in high salt buffer (50 mM KH$_2$PO$_4$ pH 7.0, 1 M NaCl, 1% Triton X-100, 1 × complete Protease inhibitors EDTA-free, Roche) and sonicated on ice (Sonics Vibra-cell, 1 min 40% amplitude, 1 s pulses with 2 s break). The insoluble DNA fraction was pelleted for 10 min at 12 000 g and 4°C. Supernatant was incubated with Talon metal-affinity resin (Clontech, Palo Alto, CA, USA) for 1–2 h at 4°C to allow binding of His-tagged proteins. Resin was washed twice with high salt buffer and twice with low salt buffer (25 mM Tris–HCl pH 7.6, 40 mM NaCl, 4.5 mM MgCl$_2$, 10% glycerol, 100 mM L-Arginine) containing 20 mM Imidazole. Elution was carried out with low salt buffer containing 500 mM Imidazole. The supernatant of this step was shock-frozen in liquid nitrogen and stored at −80°C.

As standard substrate for helicase assays a radioactively end-labeled 60 nt oligonucleotide hybridized to M13 ssDNA was used (5′ACATGATAAGATACATGGATG AGTTTGGACAAACCACAACGTAAAACGACGGCC AGTGCC 3′), forming a 20 nt double-stranded stretch with a 40 nts 5′ overhang.

The assay was performed by incubating 1 ng Twinkle protein in 40 µl helicase buffer (25 mM Tris pH 7.6, 40 mM NaCl, 4.5 mM MgCl$_2$, 100 mM L-Arginine-HCl pH 7.6, 10% glycerol, 3 mM UTP, 1 mM DTT, 5 µM unspecific oligonucleotide) with 2 amol substrate for 30 min at 37°C. The reaction was stopped by adding 10 µl loading buffer (90 mM EDTA, 6% SDS, 30% glycerol, 0.25% bromo-phenol blue). Twenty microliter reaction mixes were separated on a 15% acrylamide gel in 1 × TBE, dried on a vacuum gel drier and exposed to X-ray film or quantified by phosphoimager.

## Mutation sequencing

Point mutation levels in the NCR and *cytochrome b* region of mtDNA of POLG1 cell lines were measured as previously described (34).

## Brewer-Fangman 2D neutral/neutral Agarose electrophoresis

Mitochondrial nucleic acids were extracted using cytocha-lasine (Sigma-Aldrich) as described (9). Purified mtDNA was digested with HincII and where mentioned further treated with RNase H or S1 nuclease (Fermentas, Hanover, MD, USA) with the indicated amounts and time. The fragments were separated by 2DNAGE as described (35,36) and the gels were blotted and hybridized with a $^{32}$P-labeled DNA probe for human mtDNA nts 14 846–15 357 (9).

**Table 1.** Expressed proteins with their predicted sizes

| Variant | Size (kDa) |
| --- | --- |
| Twinkle[a] wt | 74 |
| K421A, G575D | |
| Twinkle[a] △AA346–376 (△346–376) | 71 |
| Twinky[a] | 64 |
| Twinkle[a] △AA70–343 (△70–343) | 41 |
| POLG1 wt | 142 |
| D198A, D890N, D1135A | |
| POLG1 △CAG | 141 |

All expressed proteins had the MycHis epitope tag and this tag was included in the calculation of size in kiloDaltons (kDa).
[a]Size based on predicted processing at AA42.

## RESULTS

### Establishing stable inducible cell lines expressing mtDNA replication factors

In order to study the mtDNA maintenance machinery in cultured human cells, inducible cell lines expressing wild-type and mutant variants of Twinkle, the Twinkle splice variant Twinky (14) or POLG1 were established using HEK293 Flp-In™ T-REx™ cells (see Table 1 for a list of all variants). Twinkle mutants included a lysine mutation (K421A) in the highly conserved Walker A motif implicated in ATP binding and hydrolysis; a mutation (G575D) in helicase motif H4 and implicated in DNA binding (37); a deletion of 31 amino acids (△346–376) of the region that shows similarity with the T7 gp4 linker region that has been implicated in multimer formation [see e.g. (38)]; a large deletion (△70–343) in the region of the protein that shows similarity with the T7 gp4 primase domain (19). POLG1 mutants included two polymerase deficient mutants (D890N and D1135A), one exonuclease deficient mutant D198A and a non-deleterious deletion mutant (ΔCAG) of 10 consecutive glutamines in the N-terminus, all as previously described (27). All cell lines and >99% of all cells expressed the recombinant proteins upon doxycycline (DC) induction and expressed proteins were all targeted to the mitochondrial compartment (Figure 3 and not shown).

Since DC is a mitochondrial protein synthesis inhibitor at μg/ml concentrations, we first determined the lowest possible levels of DC to achieve full induction. Figure 1A shows an increase of protein expression of wild-type (wt) Twinkle in cells with increasing DC concentrations (0–1000 ng/ml.) Both after one or three days, maximum induction levels were reached at low ng/ml concentrations, but at slightly lower concentrations after three days induction.

Figure 1B shows the expression of all proteins used in this study confirmed by immunoblotting, following 72 h of treatment with 0, 3 and 10 or 20 ng/ml DC. All proteins were detected using their respective epitope tag and gave bands of the expected size (Table 1) upon induction. In the absence of DC we could detect leaky expression of most Twinkle variants, but only when films were overexposed considerably (not shown). However,
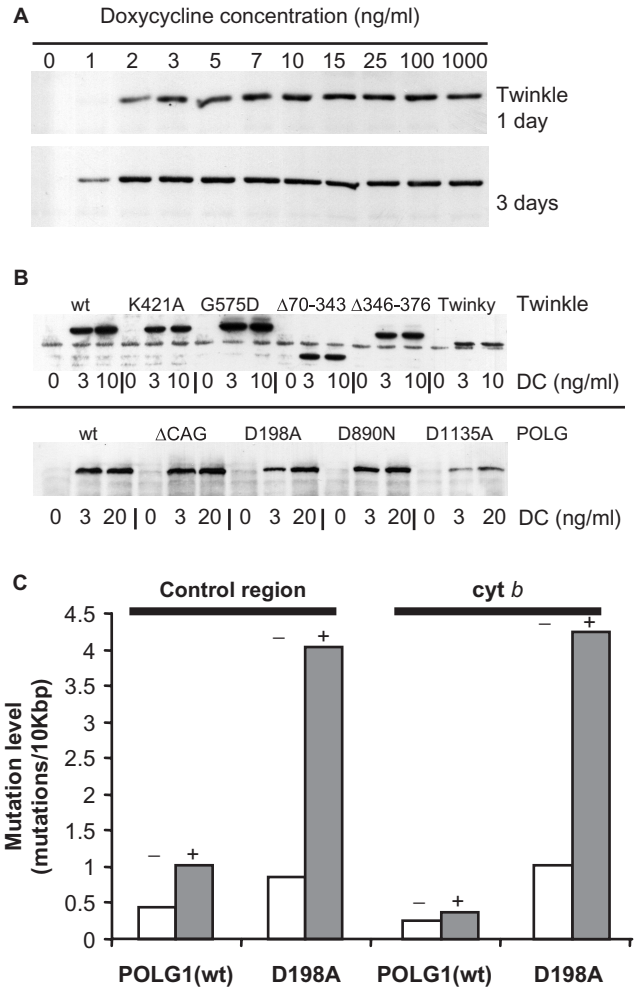


**Figure 1.** Expression of mtDNA maintenance proteins in HEK293 Flp-In™ T-REx™ cells. (**A**) Twinkle wild-type (wt) expressing cells were induced with various concentrations of doxycycline (DC) and analyzed by western blot analysis after 1 or 3 days of induction. Maximum expression levels were reached within the low ng/ml range. Similar results were obtained for all other cell lines. (**B**) Expression of all proteins used in this study was confirmed by western blot analysis comparing no induction with 3 ng/ml DC and full (10 or 20 ng/ml DC) induction for 3 days. Note that full induction with the POLG1 cell lines was chosen as 20 ng/ml, only in order to make sure maximum expression was reached as expression of POLG1 was generally much weaker when compared with the expression of Twinkle variants. (**C**) POLG1 wild type (wt) and POLG1 D198A expressing cells were grown for 2 months either without (−) or with (+) 3 ng/ml DC. MtDNA mutation levels were determined for part of the *cytochrome b* gene (*cyt b*) and part of the non-coding control region using a PCR-based approach (see Materials and Methods section). Results show that the D198A accumulated ~5–10-fold increased level of mutations compared to non-induced D198A and induced POLG1 wt. The mutation pattern was similar as reported earlier for this POLG1 variant (27).

immunofluorescence in the absence of DC induction did not result in any mitochondrial signal above background fluorescence suggesting the expression levels were very low (see also below and Figure 3A). Some of the analyzed Twinkle variants, such as Twinky, showed reproducibly lower protein levels with full induction of expression, indicating differences in protein (or mRNA)
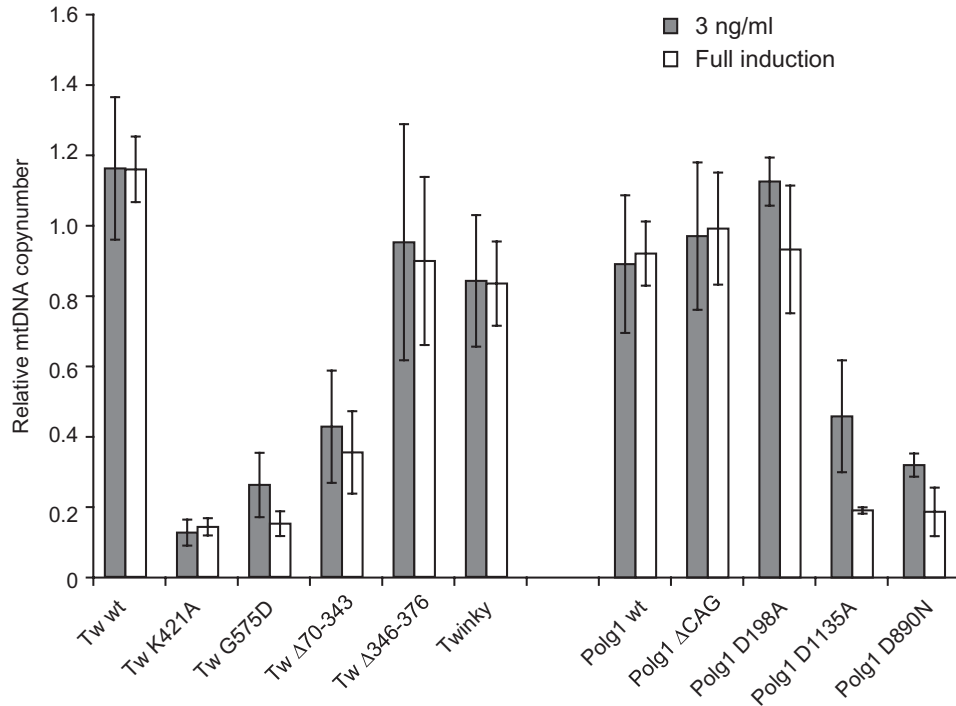
**Figure 2.** POLG1 and Twinkle mutants can cause mitochondrial DNA copy-number depletion. MtDNA copy number was determined using duplex Taqman PCR. Shown are the copy numbers for each cell line relative to their copy number in non-induced cells following either 3 ng/ml or full induction (10 ng/ml for Twinkle variants; 20 ng/ml for POLG1 variants). Note that although Twinkle K421A and G575D cell lines had a reduced steady-state copy number in non-induced cells, induction resulted in a further strong drop in copy number (see also main text). The same general trends as observed here were obtained by Southern analysis (not shown). Copy number decline was highly significant in Twinkle K421A, G575D and △70–343 lines as well as in the POLG1 D890N and D1135A mutant lines.

stability of these variants. Similarly, the POLG1 D1135A mutant showed lower expression levels than the other POLG1 variants, suggesting that the mutant protein is less stable.

As a final test for the inducible expression system, we created a cell-line expressing a POLG1 variant (D198A) in which exonuclease activity is abolished. We have shown previously that constitutive expression of D198A in cultured human cells results in the accumulation of point mutations in mtDNA (27). To validate the obtained inducible cell-line we determined the point mutation levels in two regions of mtDNA (Figure 1C). After 60 days of induction, both the *cytochrome b* and control region showed elevated mutation levels in the D198A cell-line while non-induced D198A cells showed low mutation levels similar to cells expressing POLG1 wt.

**Expression of several Twinkle and POLG1 mutants results in mtDNA copy-number depletion**

The relative mtDNA copy number in the various inducible cell lines was compared by quantitative real-time PCR (QPCR) using the nuclear amyloid precursor protein (APP) gene as a standard (31) (Figure 2). The absolute copy number determined by us for the HEK293 Flp-In™ T-REx™ and the majority of non-induced transgenic cells was ∼3000 copies/cell (2798 ± 450 (*n* = 4) for the non-induced parental cell line). Induced overexpression of

POLG1 wt or Twinkle wt did not significantly change mtDNA copy number per cell, indicating that abundance of these proteins is not rate-determining for mtDNA replication at least in cell culture. More importantly it also showed that overexpression per se did not otherwise interfere with mtDNA replication. Similarly, overexpression of Twinky, Twinkle △346–376, POLG1-D198A and POLG1-△CAG did not influence steady-state mtDNA levels. In contrast even low-level expression of Twinkle mutants K421A or G575D and POLG1 mutants D890N or D1135A lead to a dramatic decrease of mtDNA levels within a few days. The Twinkle K421A and G575D cell lines showed a significant steady-state reduction in mtDNA copy number of ∼60% even prior to induction, presumably caused by the slightly leaky expression of the Twinkle variants. This suggests these mutants are strongly dominant in nature. In contrast, the POLG1 D890N and D1135A mutants did not show a significant copy-number reduction without induction (not shown). Notwithstanding the minor leakiness, depletion upon induction was dose-dependent, as higher expression levels lead to a faster depletion than low-level expression (data not shown). For the Twinkle K421A and G575D as well as the POLG1 D890N and D1135A mutants, the mtDNA levels after three days of full DC induction were ∼20–30% compared to non-induced cells, indicating a complete abolishment of successful replication and dilution of mtDNA by cell division.
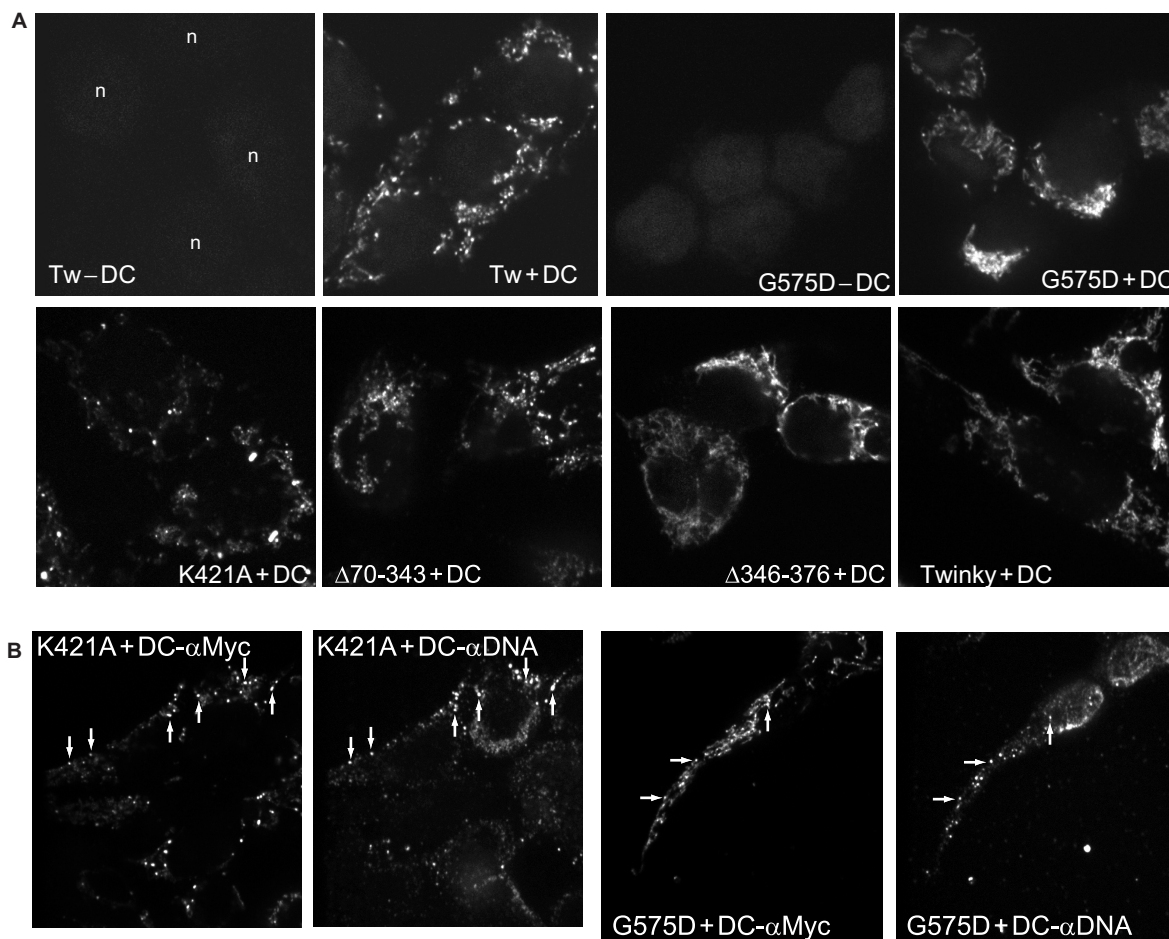
**Figure 3.** Subcellular localization of Twinkle variants. (**A**) Localization of all Twinkle variants was determined by immunocytochemistry. Two examples compare the obtained signals without (−DC) induction with induction (+DC) at 3 ng/ml DC. For this comparison images were taken with exactly the same settings for exposure, laser intensity, etc. Images were in addition processed for brightness and contrast in the same manner. The nucleus in most images in Figure A is also faintly visible due to the inclusion of DAPI in the mounting medium. In the first panel (*n*) indicates the nuclei. (**B**) Two examples of co-localization of mutant Twinkle variants visualized using anti-Myc tag antibody (−αMyc) with mtDNA using anti-DNA antibody (−αDNA) are shown. Arrows indicate a selection of foci that co-localize but by now means show all co-localizing foci as most protein foci did co-localize with more or less clear foci stained for mtDNA.

## Several Twinkle mutants show altered nucleoid localization

The localization of Twinkle variants was analyzed by immunofluorescence using the myc-tag of the recombinant proteins (Figure 3A). Overexpressed Twinkle wt showed the typical punctuated pattern within mitochondria, indicating the normal localization in mtDNA nucleoids. Twinky and △346–376 both showed diffuse mitochondrial staining. The N-terminal deletion variant △70–343 showed almost normal punctuate nucleoid-like localization. The variant G575D could be detected in punctate foci, but in addition showed enhanced diffuse staining in mitochondria. The K421A variant having a mutation in the WalkerA motif was detected in few enlarged spots, indicating either abnormal nucleoid segregation or protein aggregation. To differentiate between these possibilities we used an anti-DNA antibody to see to what extent this and other Twinkle mutants co-localized with mtDNA. As previously shown (32), Twinkle wt showed excellent co-localization with DNA as did △70–343 (not shown).

The K421A variant, despite its abnormal appearance of enlarged foci, did co-localize with mtDNA (Figure 3B). The number of DNA foci, however, was severely reduced compared to non-expressing or Twinkle wt expressing cells. For Twinkle G575D, the number of nucleoids was again severely reduced (Figure 3B) but most of the protein-foci co-localized with mtDNA. Both Twinky and the △346–376 variant showed normal nucleoid numbers judging from detection with the anti-DNA antibody.

## Helicase activity of Twinkle mutants is reduced or absent

The helicase activity of Twinkle variants was compared using His-affinity purified protein in an *in vitro* helicase assay. Twinkle wt showed the expected helicase activity and was able to unwind a DNA substrate with a 5′ overhang of >20 bases, as long as the double-stranded part was less than 25 base pairs (Figure 4 and data not shown). No such helicase activity could be detected with △346–376 or Twinky. Similarly, proteins bearing the
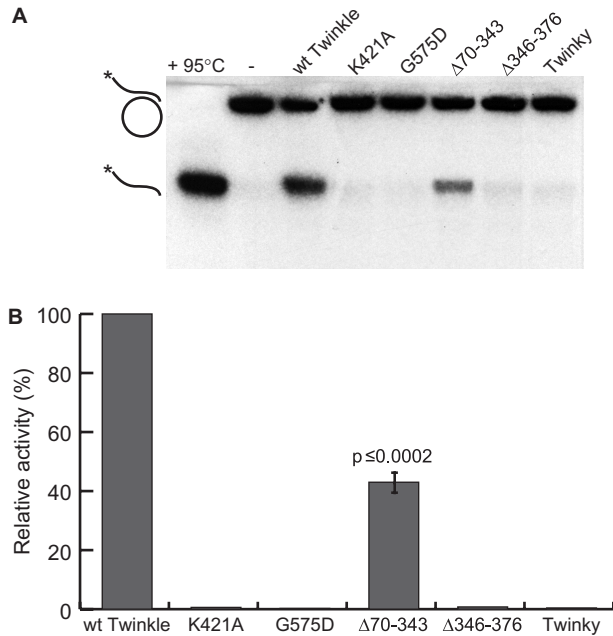
**A**



**B**



**Figure 4.** Helicase activities of purified Twinkle protein fractions. (**A**) An example of a helicase activity assay using the various Twinkle variants purified to near homogeneity. As indicated on the left-hand site, the substrate consisted of M13 plasmid with an annealed 60-nt 5′end-labeled oligo with a long (40 nt) single-stranded 5′ overhang. (−) indicates the purified substrate whereas the denatured substrate (95°C) indicates the released labeled product, which is also released when a purified protein contains helicase activity. (**B**) This panel shows the averaged results of six independent measurements, using several independent protein preparations. Except for the △70–343 mutant, which still had activity though significantly (Student's *t*-test) reduced, all other mutants had essentially lost all helicase activity.

mutation K421A in the Walker A motif or the G575D mutation in the helicase motif H4 also had less than 5% residual unwinding activity. △70–343 had ∼40% helicase activity of the wild-type protein.

**Twinkle and POLG1 mutants show different replication stalling phenotypes**

The effect of overexpression of the various mutant proteins on replication was studied using two-dimensional neutral/neutral agarose gel electrophoresis (2DNAGE) and Southern blotting, as first established by Brewer and Fangman (36,39). This method allows visualization of RIs, as DNA fragments are separated both by size and shape.

When applied to the analysis of mtDNA isolated from cultured cells, Holt and co-workers showed that various types of RIs can be detected [see e.g. (9)]. Figure 5 shows a schematic figure of human mtDNA, indicating the appropriate restriction fragments and probes used for detection of these fragments on 2DNAGE gel blots. The first 2 panels of Figure 6A show the example of a HincII digest of purified HEK293 mtDNA run on a 2DNAGE gel and probed for the 3.9 kb mtDNA fragment (nucleotide number [nt] 13 636–1006) that includes the whole NCR (for a detailed explanation of the various RIs
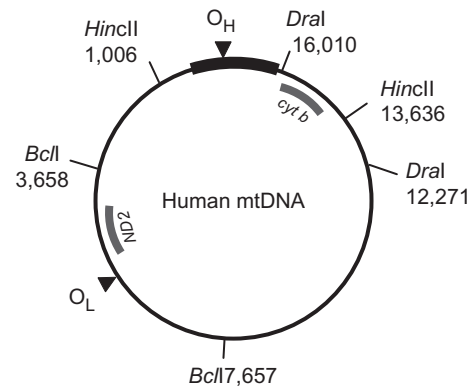


**Figure 5.** Digests and probes for human mtDNA. A schematic map of human mtDNA is shown with the positions of the fragments detected by 2DNAGE as in Figures 6–8 and Supplementary Figures 2 and 3 (HincII digests-fragment nt 13 636–1006 detected with a *cytochrome b* probe (*cyt b*, nt 14 846–15 357)), Supplementary Figure 1 (BclI digest-fragment nt 3658–7657 detected with an *ND2* probe, nt 4480–4984) and Supplementary Figure 5 (DraI digest-fragment nt 12 271–16 010 also detected with the *cyt b* probe). Probes are indicated with names and a thick gray line. The NCR is indicated with a thick black line. Please note that for clarity only the restriction sites of both ends of the probed fragments are shown here, even though all three enzymes cut at multiple sites on human mtDNA.

see Supplementary Data). All the analyses shown in the subsequent figures consider this HincII fragment. Analysis of a second region of the genome was also performed (Supplementary Figure 1).

We applied the 2DNAGE methodology to analyze the effects of overexpression of POLG1 and Twinkle variants on mtDNA RIs and in particular to test the hypothesis that the variants depleting mtDNA do so by causing non-site specific, general replication stalling or pausing. First we examined the effect of overexpression of wild-type Twinkle by comparing mtDNA RIs with and without induction (Figure 6A, right two panels). Very little change in quality or quantity of RIs was observed, with the exception of a presumed resolution intermediate which markedly decreased after induction with >3 ng/ml DC. Similar results were observed with induced expression of untagged wild-type Twinkle (not shown). Increased expression of △346–376 or Twinky showed no effect on overall replication (Figure 6B) or RIs.

When overexpressing the Twinkle mutant K421A or G575D, a considerable increase in y- and bubble RIs was observed (Figure 6C). Since at the same time the total amount of mtDNA decreased strongly, these results indicate a severe, non-site specific, slowing down of replication fork movement. In parallel there was a concomitant decrease in RNA containing RIs, RITOLS or partially single-stranded RIs. Thus, the bubble arc was not only sharper and longer than normal (Figure 6C) but was substantially resistant to RNase H treatment (Supplementary Figure 2). In K421A and especially G575D expressing cells the majority of mtDNA molecules were found on the bubble arc, indicating stalling occurred in the early phase of replication as replicating molecules forming bubble arcs on 2DNAGE harbor initiation site(s) in the fragment. The fraction of molecules in a replicative
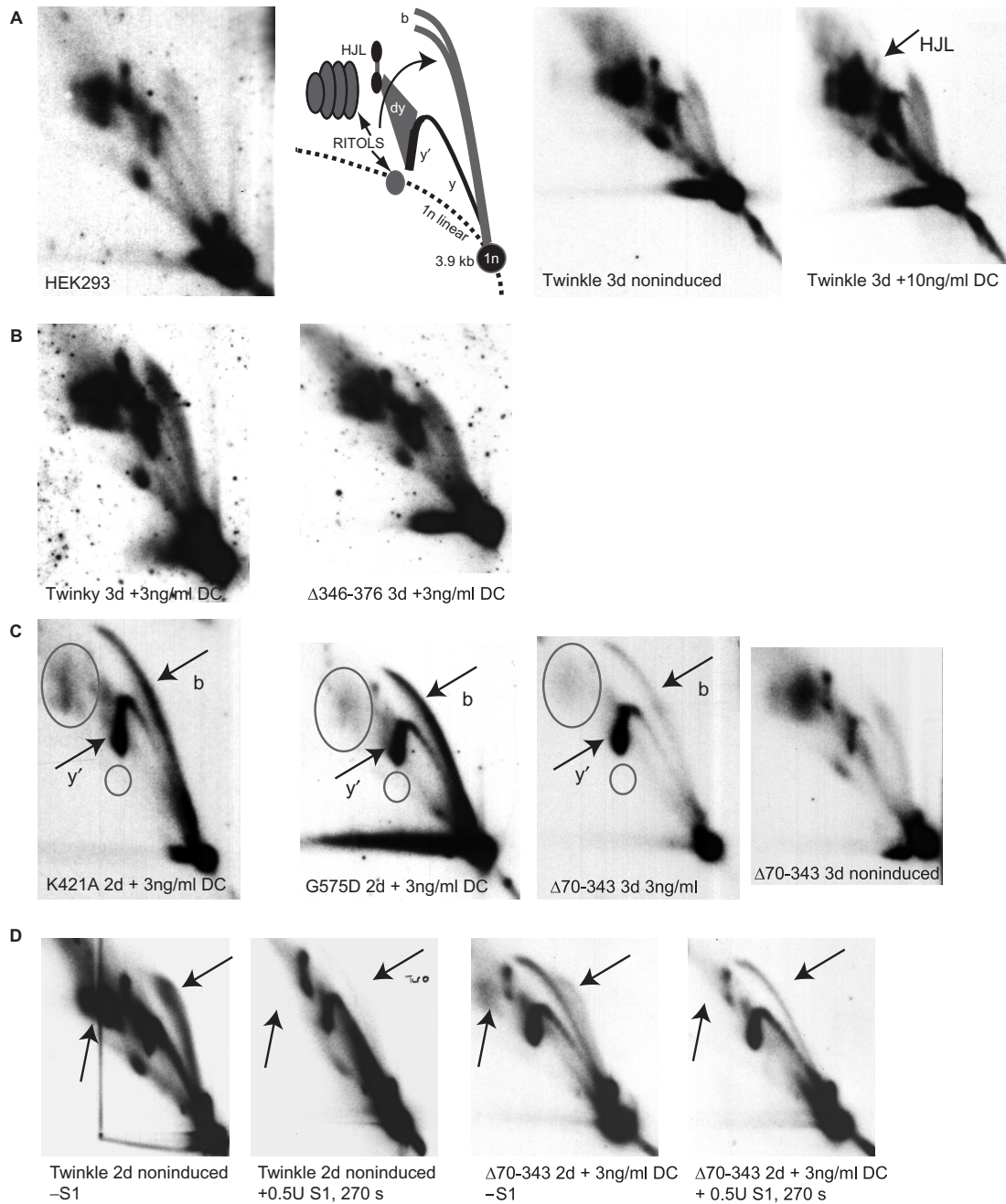
**Figure 6.** Some Twinkle mutants cause replication stalling. 2DNAGE samples for all panels consisted of purified mtDNA digested with HincII and probed with a radiolabeled *cytochrome b* gene fragment (nt 14 846–15 357). The detected fragment includes the non-coding region of mtDNA also including the *cytochrome b*, *ND6*, part of the *ND5* gene and intervening tRNA genes (nt 13 636–1006). (**A**) The first two panels on the left show the RIs of HEK293 cells and the interpretation based primarily on earlier 2DNAGE analysis of mtDNA RIs (7,9). Abbreviations: 1n, 3.9 kb non-replicating HincII fragment. (b) bubble arcs. MtDNA bubble arcs are usually very sensitive to RNase H due to the presence of patches of RNA–DNA especially on the lagging strand; these therefore also fall in the category of RITOLS as do various other RIs. Here, y and y′ indicate ascending and descending parts of the y arc and (dy) indicates double-Y structures. These will eventually form resolution intermediates resembling Holliday junctions (HJL-Holliday junction like molecules). The two panels on the right show a comparison of RIs of non-induced and fully induced cells expressing Twinkle wt. The only notable difference in this case is a reproducible reduction in one of the HJL RIs as indicated. (**B**) A normal pattern of RIs similar to non-expressing cells was observed in cells expressing Twinky or the △Linker variant. (**C**) K421A, G575D and △70–343 show similar patterns of replication stalling, with increased bubble (b) and descending Y-arc (y′) intensities, a sharpening and lengthening of the bubble arc and loss of RITOLS (ovals). The right-most panel shows the same exposure 2D gel pattern of the non-induced △70–343 line showing the typical HincII fragment pattern, including abundant RITOLS. (**D**) A limited S1 digestion illustrates that stalled RIs observed in panel C are S1 insensitive (right two panels). The effectiveness of the S1 treatment is illustrated by the left two panels, showing the effect on Twinkle non-induced cells. Similar to the S1 treatment RNase H treatment shows that the stalling RIs observed with Twinkle mutants are largely insensitive to this enzyme (Supplementary Figure 2), showing that the observed stalled RIs in panel C are essentially dsDNA. Although the intensities in subfigures A–D cannot be directly compared due to differences in exposure time, each panel contains appropriate controls of similar exposure. For example, the exposures of the left two panels in C have been chosen to be similar in comparison to the right-most panels to properly illustrate the severity of the stalling phenotype.

state was ∼30% in the 3.9 kb HincII fragment, compared to <5% in control cells. In contrast, the △70–343 mutant showed a somewhat milder stalling phenotype (Figure 6C), with the majority of RIs in the upper part of the bubble arc and on the descending area of the y-arc suggesting that replication is not aborted so frequently in the very early stages as was the case for the K421A and G575D mutants. Nevertheless, in this case also RITOLS and single-stranded RIs were considerably reduced. Finally, S1 nuclease treatment for various lengths of time with a fixed amount of enzyme resulted only in very minor changes in the abundance of the stalled RIs in Twinkle stalling mutants (Figure 6D shows the treatment of △70–343 isolated mtDNA). Only with the longest treatment was there some reduction in the bubble arc intensity with a slight concomitant increase in the ascending part of the y-arc, suggesting some single-strandedness in the bubbles resulting in broken bubbles upon S1 nuclease treatment. This is not unexpected as some single-strandedness is always expected close to the junction point of the bubble structure of replication intermediates. The △70–343 stalling in this case still showed some RITOLS intermediates and these were efficiently removed by the S1 treatment (compare regions marked by arrows in the right two panels of Figure 6D). Nevertheless, most RIs were insensitive to the S1 nuclease showing that they are essentially dsDNA (this was further supported by the analyses shown in Supplementary Figures 5 and 6). In contrast, simultaneous S1 treatment in samples with abundant RITOLS showed a strong reduction in RITOLS RIs, illustrating the effectiveness of the S1 treatment (compare regions marked by arrows in the left two panels in Figure 6D).

Overexpression of POLG1 wt (Figure 7A) and POLG1 △CAG (not shown) did not result in any obvious effect on the RIs whereas even modest expression of the mutant variants D890N and D1135A lead to a clear increase in y- and bubble arcs, suggesting these POLG1 mutants also caused replication stalling (Figure 7B). However, unlike the stalling Twinkle variants, RITOLS and single-stranded RIs persisted in the case of D890N and D1135A POLG1 mutants (Figure 7B), even after full induction of the mutant proteins for three days (not shown).

The proofreading deficient D198A POLG1 showed little change in the appearance of RIs. Three ng/ml DC slightly enhanced both the Y- and bubble arc (Figure 7C). Only after full induction did we observe a phenotype suggestive of stalling, similar to but clearly less severe than that observed with D890N and D1135A overexpression at low induction levels.

Comparison of Twinkle induced replication stalling with POLG induced replication stalling showed a considerable quantitative difference in RITOLS and single-stranded RIs (see also Supplementary Figures 3 and 5). We hypothesized (see also Discussion) that this difference could be explained by involvement of POLG1 in initiation of lagging-strand DNA synthesis or maturation. In this scenario, expression of POLG1 mutants would not only result in stalling of the leading-strand but also in delayed lagging-strand DNA synthesis. To test whether

inhibition of POLG1 function could delay lagging-strand DNA synthesis in cells overexpressing either Twinkle wt or the stalling mutations △70–343, G575D and K421A, cells were treated with the cytidine analogue dideoxycytidine (ddC), a competitive inhibitor of POLG. Incorporation instead of deoxycytidine can result in chain termination. Surprisingly, already a short 3–4 h treatment with a high dose of ddC (200 μM) in cells showing modest stalling due to Twinkle mutants like G575D, resulted in the reappearance of considerable levels of RITOLS, while accumulated y- and bubble-arc RIs were as prominent as in untreated cells (Figure 8). The appearance of RIs under the applied conditions was highly similar to the appearance observed with POLG1 stalling (compare lower panels of Figure 8 with those in Figure 7B). Due to the slightly leaky nature of the Flp-In$^{TM}$ T-REx$^{TM}$ system, the strongest Twinkle stalling mutants G575D and K421A already showed signs of stalling without induction. In particular, the G575D mutant had already lost most RITOLS in the absence of DC. Treatment of these cells with ddC showed a similar reappearance of RITOLS (Figure 8). The same regime of ddC treatment did not result in obvious changes in RIs in cells expressing Twinkle wt or in non-induced Twinkle wt cells (not shown). Finally, ddC treatment of cells showing a strong stalling phenotype caused by higher level expression of Twinkle mutants did not result in a clear increase in RITOLS (not shown), suggesting that replication was completely abolished under these conditions (see Discussion section).

## DISCUSSION

Because mtDNA encodes some of the central components required for cellular energy metabolism, its maintenance is essential for development and overall cell function. While human and mouse mtDNA were sequenced 25 years ago, for a long time the knowledge of the replication and repair machinery was lacking. This made it very difficult to test predictions of replication models for example by reconstituting the various components needed for replication or by manipulating the individual enzymes.

Using inducible overexpression of wild type or mutant variants of Twinkle and POLG1 in cell culture, we show here that replication stalling results in changed patterns of RIs that can best be understood by considering the proposed mechanisms of lagging-strand synthesis. Most notably our results show that stalling induced by deficient Twinkle results in RIs that mimic conventional strand-coupled RIs and suggest that initiation of lagging-strand DNA synthesis or maturation occurs at multiple sites across the genome. We furthermore propose that this maturation involves POLG1.

### Catalytically deficient POLG1 and Twinkle mutants

In this study we have used the HEK293 Flp-In$^{TM}$ T-Rex$^{TM}$ inducible expression system to study mutants of POLG1 and Twinkle expected to result in severe catalytic defects of these enzymes.
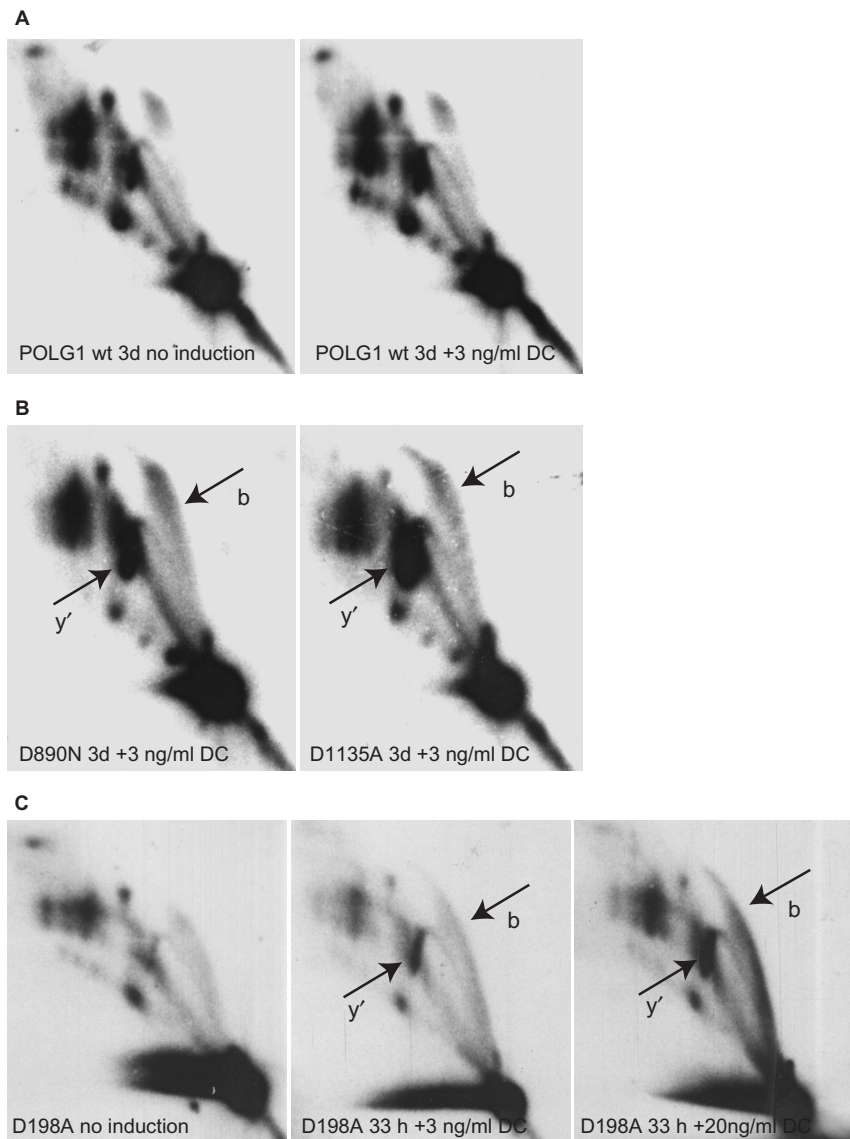
**Figure 7.** Replication stalling by POLG1 mutants. (**A**) Induction of wt POLG1 does not modify RIs compared to non-induced cells. (**B**) Induction of POLG1 mutants D890N and D1135A results in a similar pattern of replication stalling with an increased intensity both of bubble (b) and descending y (y′) RIs. RITOLS are not affected in sharp contrast with Twinkle-mutant induced stalling (Figure 6). All panels in **A** and **B** are cropped images from a single autoradiograph. (**C**) Fully induced expression of POLG1 D198A results in a modest change in RIs similar to but less severe than low-level induction of D890N and D1135A. This did, however, not result in copy number depletion (Figure 2). All panels in C are cropped images from a single autoradiograph.

As the Flp-In™ T-Rex™ system is a transgenic system, endogenous copies of the genes under study are still being expressed. Thus, in order to discuss here the severity of the mutations studied and to better value the use of the inducible Flp-In™ T-Rex™ system we tried to estimate relative levels of endogenous and recombinant proteins. The results (presented in Supplementary Figures 7 and 8) showed that at 3 ng/ml DC following 2 days induction, POLG1 recombinant and endogenous protein levels were comparable, suggesting this regime recreates the effects of expression of a 'heterozygous' mutant. The situation was different for Twinkle. Both northern blot analysis and western blot analysis suggested that at 1 ng/ml DC for 2 days, the wt

recombinant Twinkle protein level is 4–5-fold the level of endogenous Twinkle, whereas at 3 ng/ml for 2 days recombinant Twinkle is ∼8–10-fold the endogenous level.

We have previously shown that the polymerase mutants D890N and D1135A of POLG1 are deficient in a commonly used reverse transcriptase assay. Transient expression in cell culture showed modest mtDNA deple-tion (27). The corresponding aspartate 890 and 1135 residues in *E. coli* polymerase I (Asp[705] and Asp[882]) have both been shown to be essential for catalysis (40,41). A model of POLG1 based on the family A polymerase structures including *E. coli* polymerase I places both these residues at critical sites for catalysis (23), which is corroborated by the fact that cell lines for stable
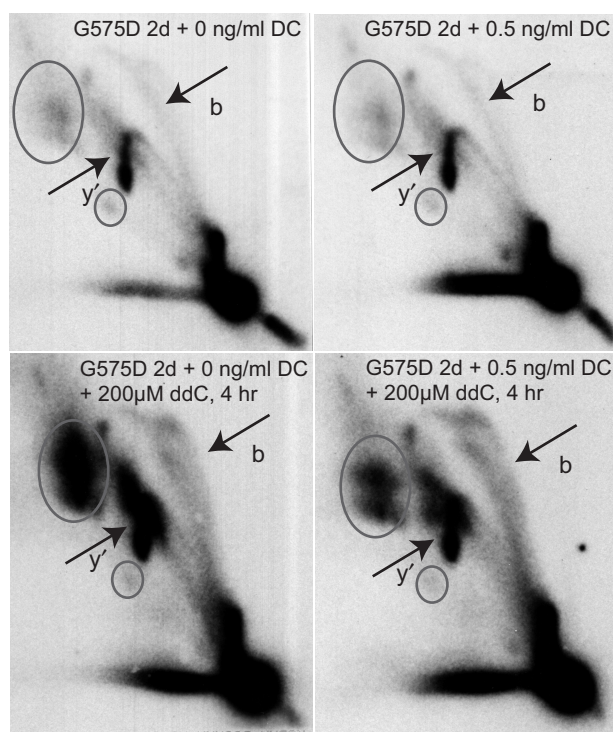
**Figure 8.** Inhibition of POLG1 by ddC results in the reappearance of RITOLS RIs following stalling induced by Twinkle G575D. Non-induced and induced (0.5 ng/ml DC) Twinkle G575D cells were treated for 4 h prior to mtDNA isolation with 200 μM ddC. Note that exposures of these four samples are identical as all four samples were run on a single gel and blot. Markings are similar as in Figure 6.

constitutive expression of these variants could not be established. More recently, HEK293 T-Rex cells were established expressing the D1135A mutation (42). The authors showed, as we repeated here, that induced expression of this variant resulted in fast mtDNA copy-number depletion. Similarly we now show that the D890N mutation resulted in similar rapid mtDNA depletion upon induction. For both variants we observed a reduction in copy number that suggests that upon full induction mtDNA is diluted by half with each cell division and thus can no longer be replicated. The generally low expression levels of recombinant POLG1 and the 1:1 ratio to endogenous PolG1 at 3 ng/ml DC imply that both the D890N and D1135A mutation behave as dominant and furthermore suggest that the results are unlikely due to the capture of the entire pool of the POLG1 accessory subunit POLG2 by excessive expression of catalytically inactive mutants.

Like the POLG1 polymerase domain mutations, the Twinkle K421A mutation is predicted to abolish catalytic activity. This mutation is an invariable residue of the Walker A motif that based on the T7 helicase crystal structures is essential for nucleotide binding and hydrolysis (43,44). Indeed, using partially purified Twinkle K421A, we showed that this mutation abolished helicase activity of Twinkle whereas wild-type purified protein showed robust helicase activity and similar specificity (not shown) as published (15). The splice variant Twinky

showed no evidence of being able to unwind the helicase substrate, even though most of the core helicase motifs except for the last few amino acids of motif H4 are left intact. The absence of helicase activity is nevertheless not unexpected as we showed previously that this variant is mostly monomeric (14), whereas the helicase activity is expected to require hexamer formation. In contrast to full-length Twinkle, overexpressed Twinky does not co-localize with mtDNA in nucleoids within the mitochondrial network. Likewise, the 31 amino acid deletion of the region, that by analogy with the region in the T7 primase/helicase we tentatively call the linker region, did not show nucleoid localization and was deficient in the unwinding assay. However, using glutaraldehyde cross-linking (Supplementary Figure 4) this mutant was still able to form multimers, unlike Twinky. The analogous mutation to G575D has previously been studied in the T7 primase/helicase (G488D) and was shown to be severely defective in its unwinding as well as its primase activity (37), although the remaining primase activity was still sufficient to allow for phage growth. It was furthermore shown to have weak DNA-binding capacity, in part explaining its weak enzymatic activities. In our hands, the Twinkle G575D mutant did not show helicase activity and its nucleoid localization was partially compromised consistent with weak DNA binding. Of the variants tested here, G575D and K421A behaved as dominant mutations since they reduced mtDNA copy number even at very low levels of expression, which occurred without induction by DC caused by leakiness of protein expression. Based on the comparison above and the difficulties to detect the protein without induction we are confident that under these conditions all recombinant Twinkle mutants are present at levels well below that of the endogenous protein. Twinky and △346–376 did not reduce copy number and are thus suggested not to compromise the activity of the endogenous Twinkle still expressed in these cell lines. Based on recent mechanistic insight in the T7 primase/helicase (45) and the fact that the purified △346–376 enzymes is inactive, this mutant is unlikely to form hetero-multimers with the endogenous enzyme. Finally, the △70–343 mutant reduced mtDNA copy number and helicase activity. Although this showed the importance of this region for Twinkle function, these results and the analysis by 2DNAGE have not yet established a specific function for this part of the protein such as a primase activity.

### Replication stalling phenotypes are compatible with delayed lagging-strand synthesis

One of the powers of 2DNAGE is that it allows accurate predictions of results on the basis of a given replication model. The method is therefore very useful to generate and test hypotheses as demonstrated by the adaptation of the method in a computer model (46). One of the unusual features of the strand-displacement model (4) is the prediction of extensive single-stranded RIs. Although these RIs have been observed by electron microscopy and more recently by AFM after mtDNA coating by SSB (47), the RITOLS replication model considers them

artefacts of degradation of the RNA patches of the initial lagging-strand. Indeed, single-stranded RIs predicted depending on the origin of replication and the restriction enzymes used (7,48) are not readily observed by 2DNAGE. It has been suggested however that the failure to observe these RIs is due to extensive branch-migration (47), thus arguing that the 2DNAGE technique is flawed and that the majority of RIs predicted by the strand-displacement model are not detected as such by this technique. This however does not provide an explanation for the observation of abundant RITOLS RIs.

Our observations of mtDNA copy-number depletion and a concomitant overall increase in RIs observed by 2DNAGE are evidence of extensive replication stalling or slowing down of replication fork movement in a non-site-specific manner. Replication stalling induced by the Twinkle K421A and G575D mutations has direct bearing on the question of delayed lagging-strand DNA synthesis as proposed by both the strand-displacement and RITOLS models. Using 2DNAGE we have shown here that accumulated RIs were visible using 3 different restriction enzymes (HincII in Figures 6–8, BclI in Supplementary Figure 1 and DraI in Supplementary Figure 5A) and probes for fragments that cover half of the genome. The results show that reduced fork progression occurs throughout mtDNA and not at a few specific sites since we did not observe appearance of discrete spots on replication arcs but a general increase of the intensities of arcs of RIs. Probing for the 3.9 kb HincII fragment from mtDNA nt 13 636–1006, which includes $O_H$ as well as most of the initiation zone (OriZ) covering the cytochrome b, ND6, part of the ND5 gene region (8), showed both an intense bubble and descending y-arc typical of a fragment containing multiple initiation sites within that fragment. However, since the ascending y-arc signal in this fragment was generally weak even when the bubble arc signal was very strong, the results nevertheless suggest that most initiation occurred in the NCR region, possibly at or near $O_H$ and not in OriZ, which would result in a more uniform y-arc (8). 2DNAGE analysis of a *Dra*I fragment spanning most or all of OriZ but not including $O_H$ further supports this hypothesis as it showed a strong y-arc signal in Twinkle stalling mutants but only a very weak bubble arc (Supplementary Figure 5A). These results thus suggest that despite the apparent lack of RITOLS on 2DNAGE gels in the Twinkle stalling mutants, the predominant initiation occurs at $O_H$ as was proposed for both the RITOLS (10) and the strand-displacement model (4). Since the RIs observed with the Twinkle stalling mutants were RNaseH and S1 nuclease insensitive they were most likely double-stranded DNA and thus resemble strand-coupled RIs. The double-stranded nature of the RIs in question was further confirmed by comparing RIs obtained by digestion with DraI and analysis of a fragment close to but not including the NCR origin (Supplementary Figure 5A and B). As explained in detail in Supplementary Figure 5C the prediction is that double-stranded RIs would yield a conventional Y-arc, whereas RITOLS and single-stranded intermediates being non-digestible on the nascent lagging-strand would yield a retarded arc resembling a Y-arc

(SMY or slow-moving Y). Also by comparison with Supplementary Figure 5A, the results show that the vast majority of RIs resulting from Twinkle stalling mutants are on a conventional Y-arc whereas by comparison RIs from cells expressing Twinkle wt or expressing the POLG1 D1135A mutant are on the predicted SMY (7,10). As an alternative we purified replication intermediates from preparative Twinkle G575D 2DNAGE gels and analyzed the purified RIs by restriction digestion using two dsDNA-specific enzymes (XhoI and DraI) and AccI that is capable of digesting ds and ssDNA. As demonstrated and illustrated in Supplementary Figure 6, these digests also fully support the conclusion that RIs observed with the Twinkle stalling mutants are essentially double-stranded DNA. As it is highly unlikely that leading- and lagging- strand synthesis have become coupled due to fork stalling, the observed double-stranded RIs can only be reasonably explained by an increased rate of lagging-strand initiation and/or maturation relative to the rate of fork movement. The results furthermore imply this to occur at sites other than the proposed lagging-strand initiation site(s) because the 3.9 kb HincII is unlikely to contain any, based on the mapping of the mouse second preferred maturation start-site proposed to be between nt 12 966 and 12 671 (10). Although not excluded at the moment, in the strand displacement model this would require novel priming events by unknown proteins and mechanisms. In the case of RITOLS intermediates, processing of incorporated RNA can be more easily envisaged to yield fully dsDNA intermediates.

The conclusions above are further strengthened by the ddC experiments. In our study overexpression of POLG1 mutants induced replication stalling but maintained RITOLS RIs. In contrast, Twinkle stalling mutants invariably resulted in an extensive loss of RITOLS RIs on 2DNAGE and in an associated increase in dsDNA RIs. The suggested cause-effect that dsDNA RIs increase as RITOLS decrease is further reinforced by the demonstration that a short inhibition of POLG by ddC led to a re-emergence of the RITOLS RIs under mild stalling conditions induced by Twinkle mutants and demonstrates that RITOLS intermediates precede dsDNA replication intermediates. Importantly, higher expression levels of, e.g. G575D for two or three days followed by ddC treatment did not result in the reappearance of RITOLS. We interpret this as a sign of a complete halt in ongoing replication and initiation whilst visible stalled RIs have already been maturated. Because under these conditions no replication (re)initiation occurs, also no new RITOLS intermediates are observed upon inhibition of POLG1 by ddC. In contrast, low G575D expression levels allow for a continued fork progression albeit at a reduced rate. Accordingly, very low (leaky) G575D expression resulted in a reduced mtDNA copy number, but a steady-state level was maintained over many cell generations (not shown). As the only plausible target for ddC treatment is POLG, the results seem to imply that the lagging-strand POLG is more sensitive to ddC treatment than the leading-strand polymerase, suggesting a different composition of leading and lagging-strand polymerase holoenzymes. Although the proposal of RITOLS replication (10)

leaves open many questions regarding the synthesis of the initial RNA-rich lagging-strand, our data suggest that POLG is involved in lagging-strand maturation. Furthermore, as proposed previously (7) our data confirm that this is one of the slower steps in mitochondrial replication. Although various RNA-associated activities of POLG1 have recently been characterized (49), to our knowledge a displacement synthesis where POLG1 would displace annealed RNA from a template strand has not yet been demonstrated. Alternatively, POLG1 exonuclease activity might be involved in removing nascent RNA before initiating DNA synthesis. Our data however seem to exclude this idea, as overexpression of the exonuclease deficient POLG1 D198A did not seriously affect mtDNA copy number or the abundance of RNA-rich versus RNA-poor RIs. Nevertheless, moderate changes in replication intermediates at high D198A expression levels were observed by 2DNAGE and warrant further investigation.

The system we have established here provides an excellent tool to study the function of other proteins in mtDNA maintenance, to establish their enzymatic or structural roles in, e.g. replication or repair and to address questions concerned with the mechanisms of DNA replication. It will also be invaluable to determine the *in vivo* effects of mutations in Twinkle and POLG1 associated with human disorders of mtDNA maintenance. Low-level expression of the severe Twinkle mutations K421A and G575D due to the leakiness of the expression system also in the absence of DC caused reduced mtDNA steady state levels of ~60%. Medium to high induction of the same mutations ultimately lead to the loss of protein expression, probably due to growth arrest or death of cells that lost their mtDNA completely and outgrowth of a few remaining cells that either had lost protein expression or did not have expression to start with, very similar to our earlier observations with constitutive expression of the POLG1 D198A mutant in 293T cells (27). In contrast expression of the null mutations did not have any effect on mtDNA levels also after 7 days and no negative selection occurred. These considerations also illustrate that the system is best used for transient expression of mutant proteins, that care must be taken when long-term effects are studied and that long-term maintenance of master plates without DC is not without hazard.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

*Conflict of interest statement*. None declared.

## REFERENCES

1. Anderson,S., Bankier,A.T., De Bruijn,M.H.L., Coulson,A.R., Drouin,J., Eperon,I.C., Nierlich,D.P., Roe,B.A., Sanger,F. *et al.* (1981) Sequence and organization of the human mitochondrial genome. *Nature*, **290**, 457–465.
2. Andrews,R.M., Kubacka,I., Chinnery,P.F., Lightowlers,R.N., Turnbull,D.M. and Howell,N. (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.*, **23**, 147.
3. Robberson,D.L., Kasamatsu,H. and Vinograd,J. (1972) Replication of mitochondrial DNA. Circular replicative intermediates in mouse L cells. *Proc. Natl Acad. Sci. USA*, **69**, 737–741.
4. Clayton,D.A. (1982) Replication of animal mitochondrial DNA. *Cell*, **28**, 693–705.
5. Spelbrink,J.N. (2003).Replication, repair and recombination of mitochondrial DNA. In Holt,I.J. (ed), *Genetics of Mitochondrial Diseases*. Oxford University Press, Oxford.
6. Holt,I.J., Lorimer,H.E. and Jacobs,H.T. (2000) Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA. *Cell*, **100**, 515–524.
7. Yang,M.Y., Bowmaker,M., Reyes,A., Vergani,L., Angeli,P., Gringeri,E., Jacobs,H.T. and Holt,I.J. (2002) Biased incorporation of ribonucleotides on the mitochondrial L-strand accounts for apparent strand-asymmetric DNA replication. *Cell*, **111**, 495–505.
8. Bowmaker,M., Yang,M.Y., Yasukawa,T., Reyes,A., Jacobs,H.T., Huberman,J.A. and Holt,I.J. (2003) Mammalian mitochondrial DNA replicates bidirectionally from an initiation zone. *J. Biol. Chem.*, **278**, 50961–50969.
9. Yasukawa,T., Yang,M.Y., Jacobs,H.T. and Holt,I.J. (2005) A bidirectional origin of replication maps to the major noncoding region of human mitochondrial DNA. *Mol. Cell*, **18**, 651–662.
10. Yasukawa,T., Reyes,A., Cluett,T.J., Yang,M.Y., Bowmaker,M., Jacobs,H.T. and Holt,I.J. (2006) Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand. *EMBO J.*, **25**, 5358–5371.
11. Takamatsu,C., Umeda,S., Ohsato,T., Ohno,T., Abe,Y., Fukuoh,A., Shinagawa,H., Hamasaki,N. and Kang,D. (2002) Regulation of mitochondrial D-loops by transcription factor A and single-stranded DNA-binding protein. *EMBO Rep.*, **18**, 18.
12. Meyer,R.R. and Laine,P.S. (1990) The single-stranded DNA-binding protein of *Escherichia coli*. *Microbiol. Rev.*, **54**, 342–380.
13. Kaguni,L.S. (2004) DNA polymerase gamma, the mitochondrial replicase. *Annu. Rev. Biochem.*, **73**, 293–320.
14. Spelbrink,J.N., Li,F.Y., Tiranti,V., Nikali,K., Yuan,Q.P., Tariq,M., Wanrooij,S., Garrido,N., Comi,G. *et al.* (2001) Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria. *Nat. Genet.*, **28**, 223–231.
15. Korhonen,J.A., Gaspari,M. and Falkenberg,M. (2003) Twinkle has 5′ to 3′ DNA helicase activity and is specifically stimulated by mtSSB. *J. Biol. Chem.*, **278**, 48627–48632.
16. Tiranti,V., Rocchi,M., DiDonato,S. and Zeviani,M. (1993) Cloning of human and rat cDNAs encoding the mitochondrial single- stranded DNA-binding protein (SSB). *Gene*, **126**, 219–225.

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

18. Shutt,T.E. and Gray,M.W. (2006) Bacteriophage origins of mitochondrial replication and transcription proteins. *Trends Genet.*, **22**, 90–95.

19. Shutt,T.E. and Gray,M.W. (2006) Twinkle, the Mitochondrial Replicative DNA Helicase, Is Widespread in the Eukaryotic Radiation and May Also Be the Mitochondrial DNA Primase in Most Eukaryotes. *J. Mol. Evol.*, **62**, 588–599.

20. Longley,M.J., Clark,S., Yu Wai Man,C., Hudson,G., Durham,S.E., Taylor,R.W., Nightingale,S., Turnbull,D.M., Copeland,W.C. *et al.* (2006) Mutant POLG2 disrupts DNA polymerase gamma subunits and causes progressive external ophthalmoplegia. *Am. J. Hum. Genet.*, **78**, 1026–1034.

21. Nikali,K., Suomalainen,A., Saharinen,J., Kuokkanen,M., Spelbrink,J.N., Lonnqvist,T. and Peltonen,L. (2005) Infantile onset spinocerebellar ataxia is caused by recessive mutations in mitochondrial proteins Twinkle and Twinky. *Hum. Mol. Genet.*, **14**, 2981–2990.

22. Longley,M.J., Graziewicz,M.A., Bienstock,R.J. and Copeland,W.C. (2005) Consequences of mutations in human DNA polymerase gamma. *Gene*, **354**, 125–131.

23. Graziewicz,M.A., Longley,M.J., Bienstock,R.J., Zeviani,M. and Copeland,W.C. (2004) Structure-function defects of human mitochondrial DNA polymerase in autosomal dominant progressive external ophthalmoplegia. *Nat. Struct. Mol. Biol.*, **11**, 770–776.

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

25. Hu,J., Vanderstraeten,S. and Foury,F. (1995) Isolation and characterization of ten mutator alleles of the mitochondrial DNA polymerase- encoding MIP1 gene from *Saccharomyces cerevisiae*. *Gene*, **160**, 105–110.

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

27. Spelbrink,J.N., Toivonen,J.M., Hakkaart,G.A., Kurkela,J.M., Cooper,H.M., Lehtinen,S.K., Lecrenier,N., Back,J.W., Speijer,D. *et al.* (2000) In vivo functional analysis of the human mitochondrial DNA polymerase POLG expressed in cultured human cells. *J. Biol. Chem.*, **275**, 24818–24828.

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

29. Kujoth,G.C., Hiona,A., Pugh,T.D., Someya,S., Panzer,K., Wohlgemuth,S.E., Hofer,T., Seo,A.Y., Sullivan,R. *et al.* (2005) Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. *Science*, **309**, 481–484.

30. Tyynismaa,H., Mjosund,K.P., Wanrooij,S., Lappalainen,I., Ylikallio,E., Jalanko,A., Spelbrink,J.N., Paetau,A. and Suomalainen,A. (2005) Mutant mitochondrial helicase Twinkle causes multiple mtDNA deletions and a late-onset mitochondrial disease in mice. *Proc. Natl Acad. Sci. USA*, **102**, 17687–17692.

31. Tyynismaa,H., Sembongi,H., Bokori-Brown,M., Granycome,C., Ashley,N., Poulton,J., Jalanko,A., Spelbrink,J.N., Holt,I.J. *et al.* (2004) Twinkle helicase is essential for mtDNA maintenance and regulates mtDNA copy number. *Hum. Mol. Genet.*, **13**, 3219–3227.

32. Garrido,N., Griparic,L., Jokitalo,E., Wartiovaara,J., Van Der Bliek,A.M. and Spelbrink,J.N. (2003) Composition and dynamics of human mitochondrial nucleoids. *Mol. Biol. Cell*, **14**, 1583–1596.

33. Legros,F., Malka,F., Frachon,P., Lombes,A. and Rojo,M. (2004) Organization and dynamics of human mitochondrial DNA. *J. Cell Sci.*, **117**, 2653–2662.

34. Wanrooij,S., Luoma,P., van Goethem,G., van Broeckhoven,C., Suomalainen,A. and Spelbrink,J.N. (2004) Twinkle and POLG defects enhance age-dependent accumulation of mutations in the control region of mtDNA. *Nucleic Acids Res.*, **32**, 3053–3064.

35. Friedman,K.L. and Brewer,B.J. (1995) Analysis of replication intermediates by two-dimensional agarose gel electrophoresis. *Methods Enzymol.*, **262**, 613–627.

36. Brewer,B.J. and Fangman,W.L. (1987) The localization of replication origins on ARS plasmids in S. cerevisiae. *Cell*, **51**, 463–471.

37. Washington,M.T., Rosenberg,A.H., Griffin,K., Studier,F.W. and Patel,S.S. (1996) Biochemical analysis of mutant T7 primase/ helicase proteins defective in DNA binding, nucleotide hydrolysis, and the coupling of hydrolysis with DNA unwinding. *J. Biol. Chem.*, **271**, 26825–26834.

38. Guo,S., Tabor,S. and Richardson,C.C. (1999) The linker region between the helicase and primase domains of the bacteriophage T7 gene 4 protein is critical for hexamer formation. *J. Biol. Chem.*, **274**, 30303–30309.

39. Brewer,B.J. and Fangman,W.L. (1988) A replication fork barrier at the 3′ end of yeast ribosomal RNA genes. *Cell*, **55**, 637–643.

40. Polesky,A.H., Steitz,T.A., Grindley,N.D.F. and Joyce,C.M. (1990) Identification of residues critical for the polymerase activity of the Klenow fragment of DNA polymerase I from *Escherichia coli*. *J. Biol. Chem.*, **265**, 14579–14591.

41. Polesky,A.H., Dahlberg,M.E., Benkovic,S.J., Grindley,N.D.F. and Joyce,C.M. (1992) Side chains involved in catalysis of the polymerase reaction of DNA polymerase I from *Escherichia coli*. *J. Biol. Chem.*, **267**, 8417–8428.

42. Jazayeri,M., Andreyev,A., Will,Y., Ward,M., Anderson,C.M. and Clevenger,W. (2003) Inducible Expression of a Dominant Negative DNA Polymerase-gamma Depletes Mitochondrial DNA and Produces a rho 0 Phenotype. *J. Biol. Chem.*, **278**, 9823–9830.

43. Sawaya,M.R., Guo,S., Tabor,S., Richardson,C.C. and Ellenberger,T. (1999) Crystal structure of the helicase domain from the replicative helicase- primase of bacteriophage T7. *Cell*, **99**, 167–177.

44. Singleton,M.R., Sawaya,M.R., Ellenberger,T. and Wigley,D.B. (2000) Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell*, **101**, 589–600.

45. Crampton,D.J., Mukherjee,S. and Richardson,C.C. (2006) DNA-induced switch from independent to sequential dTTP hydrolysis in the bacteriophage T7 DNA helicase. *Mol. Cell*, **21**, 165–174.

46. Viguera,E., Rodriguez,A., Hernandez,P., Krimer,D.B., Trellez,O. and Schvartzman,J.B. (1998) A computer model for the analysis of DNA replication intermediates by two-dimensional agarose gel electrophoresis. *Gene*, **217**, 41–49.

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

48. Belanger,K.G., Mirzayan,C., Kreuzer,H.E., Alberts,B.M. and Kreuzer,K.N. (1996) Two-dimensional gel analysis of rolling circle replication in the presence and absence of bacteriophage T4 primase. *Nucleic Acids Res.*, **24**, 2166–2175.

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