

Ultra-Deep Sequencing of Mouse Mitochondrial DNA: Mutational Patterns and Their Origins

Adam Ameur¹*, James B. Stewart²*, Christoph Freyer², Erik Hagström³, Max Ingman¹, Nils-Göran Larsson^{2,3}, Ulf Gyllensten^{1*}

1 Department of Immunology, Genetics, and Pathology, SciLifeLab Uppsala, Rudbeck Laboratory, Uppsala University, Uppsala, Sweden, **2** Max Planck Institute for Biology of Ageing, Cologne, Germany, **3** Division of Metabolic Diseases, Department of Laboratory Medicine, Karolinska Institutet, Stockholm, Sweden

Abstract

Somatic mutations of mtDNA are implicated in the aging process, but there is no universally accepted method for their accurate quantification. We have used ultra-deep sequencing to study genome-wide mtDNA mutation load in the liver of normally- and prematurely-aging mice. Mice that are homozygous for an allele expressing a proof-reading-deficient mtDNA polymerase (mtDNA mutator mice) have 10-times-higher point mutation loads than their wildtype siblings. In addition, the mtDNA mutator mice have increased levels of a truncated linear mtDNA molecule, resulting in decreased sequence coverage in the deleted region. In contrast, circular mtDNA molecules with large deletions occur at extremely low frequencies in mtDNA mutator mice and can therefore not drive the premature aging phenotype. Sequence analysis shows that the main proportion of the mutation load in heterozygous mtDNA mutator mice and their wildtype siblings is inherited from their heterozygous mothers consistent with germline transmission. We found no increase in levels of point mutations or deletions in wildtype C57Bl/6N mice with increasing age, thus questioning the causative role of these changes in aging. In addition, there was no increased frequency of transversion mutations with time in any of the studied genotypes, arguing against oxidative damage as a major cause of mtDNA mutations. Our results from studies of mice thus indicate that most somatic mtDNA mutations occur as replication errors during development and do not result from damage accumulation in adult life.

Citation: Ameur A, Stewart JB, Freyer C, Hagström E, Ingman M, et al. (2011) Ultra-Deep Sequencing of Mouse Mitochondrial DNA: Mutational Patterns and Their Origins. *PLoS Genet* 7(3): e1002028. doi:10.1371/journal.pgen.1002028

Editor: Gregory S. Barsh, Stanford University School of Medicine, United States of America

Received: November 10, 2010; **Accepted:** February 1, 2011; **Published:** March 24, 2011

Copyright: © 2011 Ameur et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was supported by Swedish Research Council (VR) grants (to UG and N-GL) and the Knut and Alice Wallenberg Foundation, through the Center for Metagenomic Sequence Analysis (CMS), as well as an ERC Advanced Investigator grant to N-GL. The funders had no role in study design, data collection, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: ulf.gyllensten@igp.uu.se

† These authors contributed equally to this work.

Introduction

A decline of mitochondrial function has been observed in a variety of aging mammalian tissues and is implicated as a driving force behind the aging process [1–2]. A somatic mammalian cell carries thousands of copies of the mitochondrial DNA (mtDNA) chromosome, which encodes essential subunits of the respiratory chain protein complexes as well as rRNAs and tRNAs needed for mitochondrial translation. Expression of mtDNA is required for maintenance of oxidative phosphorylation and accumulation of somatic mtDNA mutations has been suggested as a cause of the observed decrease in respiratory chain function during aging [3–4]. A variety of low levels of point mutations and rearrangements of mtDNA are found in aging mammals. Rare mutational events tend to undergo clonal expansion as exemplified by human aging where clonal expansion of somatic mtDNA mutations cause a mosaic respiratory chain deficiency in tissues such as brain, heart, skeletal muscle and large intestine [4]. Point mutations as well as insertions and deletions (indels) of mtDNA are observed in tissues of aging humans [5–9], primates [10] and rodents [11], but the relative contribution of each of these different types of mutations to the aging process is unknown.

The mtDNA mutator mice (genotype PolgA^{mut}/PolgA^{mut}) express a proof-reading-deficient mtDNA polymerase (PolgA^{D257A}) and have provided experimental support that accumulation of mtDNA mutations can lead to a premature aging syndrome [12–14]. These mice contain high levels of point mutations in their mtDNA and high levels of several species of large linear deleted mtDNA molecules. Although the linear deleted molecules are abundant (~25–30% of total mtDNA in liver) the corresponding reduction in levels of full-length mtDNA molecules is on its own not sufficient to cause an impairment of respiratory chain function [13]. A detailed molecular characterization of the mtDNA mutator mice has shown that the high levels of point mutations are the likely explanation for the respiratory chain deficiency and the premature aging syndrome [15]. One report claims that the presence of a third type of mtDNA mutation, i.e. circular mtDNA molecules with large deletions, may be of critical importance in driving the premature aging phenotype of mtDNA mutator mice [16–17]. However, this finding has been refuted by several other reports showing that the levels of such deleted molecules are extremely low [15,18–21]. In addition, the biochemical phenotype in mtDNA mutator mice can be fully explained by the finding that high levels of point mutations in

Author Summary

Mitochondria represent the powerhouses of cells and have their own DNA. Mutations in the mitochondrial genome are associated with a range of human diseases and have also been implicated as a driving force behind the aging process. We have used ultra-deep sequencing to study the genome-wide mutation load in the mitochondrial DNA (mtDNA) of liver from normal inbred mice and mice that express a proof-reading-deficient mtDNA polymerase (mtDNA mutator mice) that cause premature aging. The mtDNA mutator mice show a dramatic increase of point mutations with age and have 10-times-higher point mutation levels than wildtype siblings or normal C57Bl/6N mice. Circular mtDNA molecules with large deletions occur at very low frequencies in mtDNA mutator mice and are therefore unlikely to contribute to the premature aging phenotype. We found no increase in levels of point mutations or deletions in normal mice with increasing age, arguing against the accumulation of mtDNA mutations as contributing to aging. Our results indicate that most somatic mtDNA mutations occur as replication errors during the rapid amplification of mtDNA during embryogenesis and do not result from damage accumulation in adult life.

mtDNA leads to the synthesis of respiratory chain subunits with abundant amino acid replacements, which, in turn, cause instability of the respiratory chain complexes [15].

Oxidative damage has for more than 50 years been proposed as a central mechanism in aging, but the supporting evidence is mainly correlative. Interestingly, the mtDNA mutator mice have no or minor increase in levels of reactive oxygen species (ROS) production and oxidative damage despite a severe decline in oxidative phosphorylation capacity. This finding refutes the popular notion of a vicious cycle whereby somatic mtDNA mutations lead to increased ROS production, which, in turn, creates additional mtDNA mutations that further increase ROS production [12,14].

Genome-wide estimates of intra-individual mtDNA variability are needed to determine the importance of the various types of mtDNA mutations in aging. Recent studies using next-generation sequencing of the human mitochondrial genome have identified a small number of sites, in which normal individuals carry both the wildtype copy and a high-frequency mutant allele [22–23]. In their efforts to eliminate false positive calls, these studies limited their analyses to high-frequency mutations and excluded rare variants.

We hypothesised that a DNA sequencing technology that minimize the levels of false positive calls caused by technical errors could be applied to assess the intra-individual mtDNA variability caused by low frequency mutations. We investigated this possibility by using the ABI SOLiD technology to sequence mtDNA of normal aging mice and prematurely aging mtDNA mutator mice. To control for technical errors we sequenced a complete mtDNA genome inserted in a lambda clone.

Results

We used the SOLiD sequencing platform to sequence mtDNA purified from liver mitochondria of normal C57Bl/6N mice (henceforth denoted wt^{B6}) at 30, 40 and 84 weeks of age. In addition, we analysed liver mtDNA from mtDNA mutator mice ($PolgA^{mut}/PolgA^{mut}$; henceforth denoted mutators), heterozygous mtDNA mutator mice ($+/PolgA^{mut}$; henceforth denoted heterozygotes) and their wildtype siblings (henceforth denoted wt^{mut}) at 30 and 40 weeks of age. An mtDNA molecule cloned in the lambda

phage (henceforth referred to as λ^{mtDNA}) was also sequenced and used as a control and for correction of errors introduced by the SOLiD sequencing technique. Samples and sequence read information is shown in Table S1.

Sequence coverage

The sequence coverage for 99% of the mtDNA bases was at least 1800x in each sample and the coverage for 70% of the mtDNA bases was at least 10,000x (Figure S1). The wt^{B6} , heterozygotes and wt^{mut} mice showed a similar sequencing coverage between samples. The coverage profiles were quite similar to that obtained by sequencing λ^{mtDNA} (Figure 1). By contrast, in the mutator samples there was a pronounced decrease in sequence coverage for approximately one third of the genome (Figure 1). This under-represented region corresponds to the small arc between the two origins of replication for the mtDNA molecule. The mtDNA mutator mice have been reported to contain ~25–30% of truncated, linear double-stranded mtDNA molecules, ranging from the origin of heavy strand replication (O_H) to the origin of light strand replication (O_L) [13,19]. The unequal sequence coverage curve thus likely reflects the presence of these linear mtDNA molecules with a large deletion in the mutator.

Point mutation loads

The point mutation frequency was estimated as the median number of mutations per nucleotide site and varied dramatically in mice of different genotypes (Figure 2, Figure 3). The wt^{B6} mice showed a median mutation frequency of $1.3\text{--}1.8 \times 10^{-4}$ per site, while the mutators had frequencies of about 12×10^{-4} (Table 1). The heterozygotes also had an elevated point mutation frequency both at 30 and 40 weeks of age, as compared with their wildtype siblings. The wt^{mut} and wt^{B6} mice showed similar point mutation frequencies ($1.3\text{--}1.8 \times 10^{-4}$, Table 1). We also determined the number of high frequency point mutation sites, defined as the number of sites with single nucleotide variant (SNV) frequencies $>0.5\%$. Interestingly, there are approximately the same number of such sites in heterozygotes and wt^{mut} (Figure 3), arguing that this mutation load is inherited from their common mother which is heterozygous for the mtDNA mutator allele. In contrast, the number of high frequency point mutation sites in wt^{B6} mice was only half of the value of wt^{mut} mice (Figure 3). We observed no difference in mutation frequencies (Table 1) or in the number of high frequency point mutation sites in wt^{B6} mice between the ages of 30 and 84 weeks (Figure 3).

Genomic distribution of point mutations

In the mutators the protein coding genes, tRNA genes and rRNA genes showed similar mutation loads, whereas the mutation load was 59–66% lower in the control region (also referred to as the major non-coding region or D-loop region) (Table 1; Figure 4). A particularly conserved part of the control region, the conserved sequence blocks (CSB), had an almost 80% reduction in the mutation load compared with the coding regions. There was also a modest reduction (34–42% decrease) in mutation load in the control region of heterozygotes, whereas no difference was observed in wt^{B6} or wt^{mut} mice (Table 1).

Absence of a signature of oxidative damage

The obtained mutational spectrum allowed us to investigate whether oxidative damage is a source of mtDNA damage during aging, since oxidative damage is expected to increase the number of observed transversion mutations, as exemplified by the G/C to T/A transversions caused by the oxidative adduct 8-oxo-guanine [4]. The

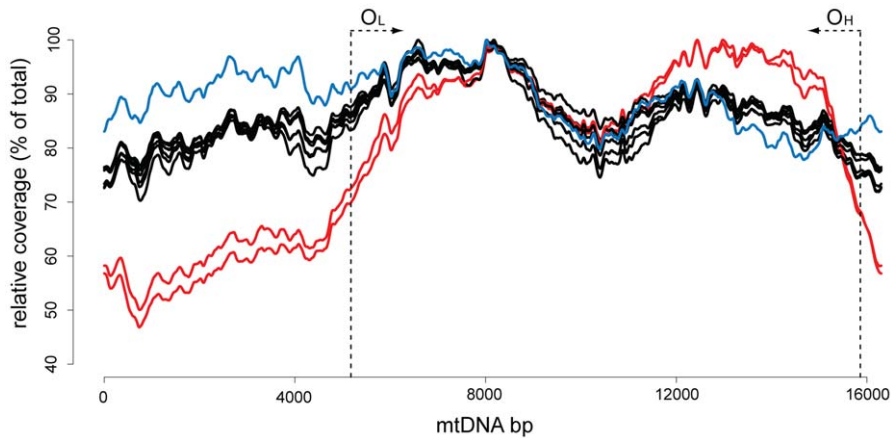


Figure 1. Sequence coverage for the mouse mtDNA samples. Mitochondrial genome position (x-axis) versus sequence coverage divided by maximum coverage for each sample. The coverage was calculated using a 2 kb sliding window average. The red lines correspond to mtDNA mutator samples. All other mouse samples are in black. The blue line represents the sequence coverage of the λ^{mtDNA} control. The approximate locations of the origins of light-strand (O_L) and heavy-strand (O_H) replication are indicated by dotted lines with arrows.
doi:10.1371/journal.pgen.1002028.g001

number of transitions versus transversions did not change as a function of age in any of the samples, implying that mtDNA polymerase errors are responsible for the majority of the observed point mutations (Table S2). The heterozygotes and the mutators showed increased relative levels of transitions in the two samples, as expected under conditions of excess polymerase errors. These observations argue against oxidative damage as main contributor to the observed mutation pattern.

Shared mutations among litter mates

In order to identify mutations that are common to littermates, we calculated the number of high frequency point mutation sites that are shared among offspring (mutators, heterozygotes and wt^{mut} mice) obtained by mating heterozygous mtDNA mutator mice (Figure 5). We found more than 800 high frequency point mutation sites in all siblings within litters of 30 and 40-weeks of age. Approximately 85% of the high frequency variable point mutation sites were present in single animals and most of these were mutators (56–59%). Approximately 5% of the sites ($n = 44$ – 42) were shared between the littermates. Interestingly, the wt^{B6} mice obtained from independent matings shared a substantial number of these high frequency point mutation sites ($n = 35$), suggesting the occurrence of mutational hotspots.

Mutational hotspots

The point mutation frequency varied between nucleotide sites, with some sites experiencing 100–1,000 times higher mutation

frequencies than an average variable site. A number of regions also showed several neighbouring, but not necessarily adjacent, positions with high mutation frequencies (Table S3). Seven of the twelve identified hotspot regions in this study corresponded to the regions with increased levels of inherited mtDNA mutations we previously have described in wildtype mouse strains derived from female mtDNA mutator mice [24]. Similar hotspot regions have been reported in human mtDNA [22,25,26] and such mutational hotspots may play an important role in the generation of the common disease alleles reported for mtDNA [27].

Insertion/deletion mutations

Although we did not utilize paired-end sequencing, it was still possible to detect structural re-arrangements, such as indel mutations, by using the *SplitSeek* method [28]. This method splits the sequence reads and aligns the two parts independently to the reference sequence. We found that indels had a median frequency of 10^{-3} – 10^{-4} (Table 1). The mutators had 4–6.6 times more indels than their wt^{mut} siblings and 2–10.5 times more indels than wt^{B6} mice (Table 1). Most indels were small (Table S4) and only five deletions involved more than 1 kb of DNA. Four of these large deletions were found in a single mutator and the remaining deletion was present in a heterozygote (Table S5). The number of indels did not increase with age in mice of any of the studied genotypes (Table 1).

The indels showed an uneven distribution across the mtDNA genome with a clustering of sites in two regions around genome

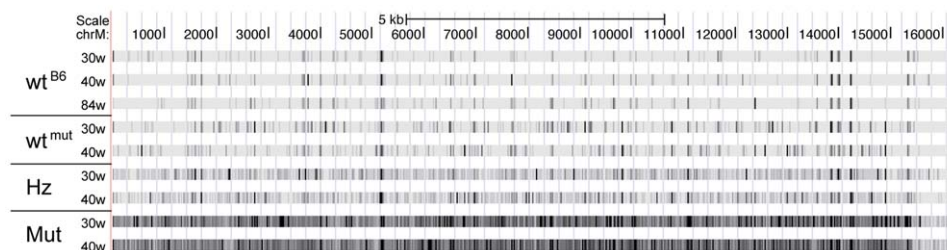


Figure 2. Corrected mutational frequencies viewed in the UCSC genome browser. The scale ranges from low mutation frequencies (light colors) to high mutation frequencies (black). (Abbreviations: wt^{B6} = C57Bl/6N animal, wt^{mut} = wildtype sibling of the mtDNA mutator mouse, Hz = heterozygous sibling of the mtDNA mutator mouse, Mut = mtDNA mutator mouse, 30w & 40w = 30 and 40 weeks of age, respectively).
doi:10.1371/journal.pgen.1002028.g002

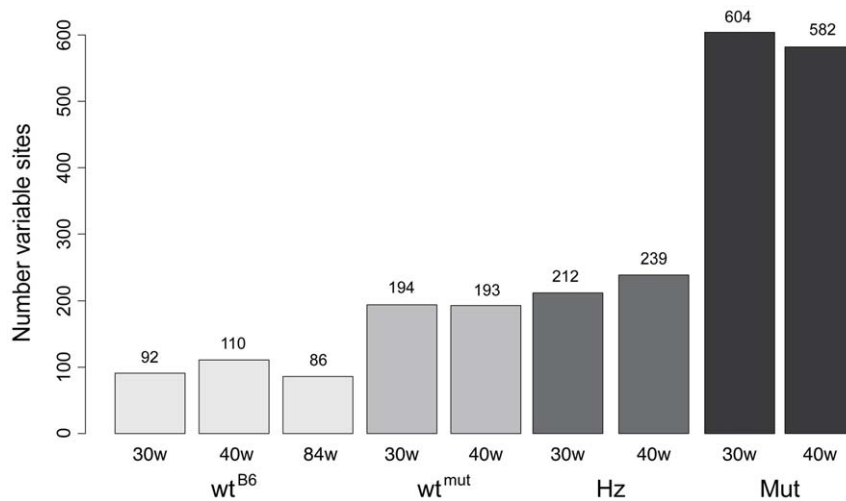


Figure 3. Number of mutations with a frequency of at least 0.5% in each sample. (Abbreviations: wt^{B6} = C57Bl/6N animal, wt^{mut} = wildtype sibling of the mtDNA mutator mouse, Hz = heterozygous sibling of the mtDNA mutator mouse, Mut = mtDNA mutator mouse, 30w & 40w = 30 weeks of age and 40 weeks of age, respectively). doi:10.1371/journal.pgen.1002028.g003

positions 1,000–5,000 and 12,000–15,000 (Figure S2). Small indels found in the mutator mouse sibling sets were often observed in proximity to mononucleotide stretches. These small indels were present in mutators, heterozygotes and wt^{mut}, but were very infrequent in wt^{B6} mice and essentially absent in λ^{mtDNA} (Table S4). These data demonstrate that these small indel events are induced by the *PolgA^{mut}* allele. The presence of indels in wt^{mut} animals, who lack the *PolgA^{mut}* allele, can be explained by inheritance of these indels from their heterozygous mothers.

Functional consequence of mutations

Adult homozygous mtDNA mutator mice are predicted to encode about 7 amino acid substitutions per mtDNA molecule, compared with 2 substitutions in their wt^{mut} siblings and 1 or 2 substitutions in wt^{B6} mice (Table 1). The presence of many amino acid substitutions in the mutators has been demonstrated to impair respiratory chain function due to destabilization of the respiratory

chain complexes [15]. Our results provide additional support for the conclusion that the premature aging syndrome in mtDNA mutator mice is due to accumulation of point mutations in mtDNA, which, in turn, cause amino acid substitutions that impair mitochondrial function.

Discussion

Our analysis has revealed a number of novel aspects of the accumulation of mtDNA mutations in wildtype and mtDNA mutator mice. The mutators show highly elevated point mutation frequencies as compared to their wt^{mut} siblings, consistent with previous results [12–14,16,21,29–30]. The SOLiD sequencing estimates of mutation loads presented here are similar to the mutation load estimates that we and others previously have obtained by Sanger sequencing of cloned PCR fragments [12–14]. In a recent study using a different next-generation DNA sequencing technology, He et al. [22] reported

Table 1. Distribution of mutations in the mtDNA samples.

Sample	Mutation Load (median per site frequency (value $\times 10^{-4}$))						Estimated coding changes per mtDNA molecule				Indels		
	All mtDNA	Protein coding	tRNA	rRNA	Control Region		Syn.	Non-Synonymous		STOP codon	Number of indels	Avg freq ($\times 10^{-4}$)	Max Freq ($\times 10^{-4}$)
					All	CSB		All	AA change				
Wt ^{B6} 30w	1.8	1.8	2.1	1.8	2.0	1.8	0.20	1.38	1.21	0.17	1	4	4
Wt ^{B6} 40w	1.3	1.3	1.5	1.4	1.6	1.5	0.17	1.18	1.04	0.15	6	3	5
Wt ^{B6} 84w	1.6	1.6	1.9	1.6	1.5	1.8	0.19	1.37	1.20	0.17	4	6	6
Wt ^{mut} 30w	1.4	1.4	1.7	1.3	1.4	1.1	0.31	2.34	2.11	0.23	12	9	26
Wt ^{mut} 40w	1.5	1.5	1.9	1.5	1.7	1.5	0.23	2.42	2.06	0.36	7	5	5
Hz 30w	3.8	3.9	3.9	3.7	2.5	1.7	0.50	3.32	2.87	0.45	15	11	20
Hz 40w	3.3	3.4	3.8	3.1	1.9	1.2	0.42	3.48	2.95	0.52	31	7	30
Mut 30w	11.7	12.5	11.5	10.7	4.8	2.8	1.26	8.98	7.56	1.42	53	12	56
Mut 40w	12.5	13.6	12.4	11.2	4.2	2.7	1.21	7.94	6.90	1.04	41	19	204

doi:10.1371/journal.pgen.1002028.t001

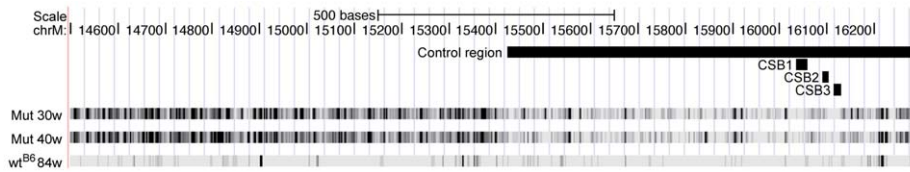


Figure 4. Mutation frequencies for the two mtDNA mutator mice (Mut 30w and Mut 40w) and the C57Bl/6N 84-week-old mouse (wt^{B6}) in a region including the control region. The locations of the control region and the conserved sequence blocks (CSB) 1–3 are indicated by horizontal black lines.

doi:10.1371/journal.pgen.1002028.g004

eight sites with a mutation frequency $>1.6\%$ in a human mtDNA sample. If we apply their detection threshold to our data, we identify five such sites in the wt^{B6} mice. Techniques that enable identification of low frequency variants, such as those used in our study, are likely to uncover additional variability and provide a more complete understanding of the mitochondrial mutation load.

We hypothesized that our analysis criteria should also uncover gradual increase in the mutation loads with natural aging, consistent with published results obtained by other mutation detection methods (reviewed in [4]). However, we neither detected a difference in the mtDNA mutation load in liver mtDNA from wt^{B6} mice at different ages (30–84 weeks), nor did we see a shift in the mutational spectrum consistent with oxidative damage causing

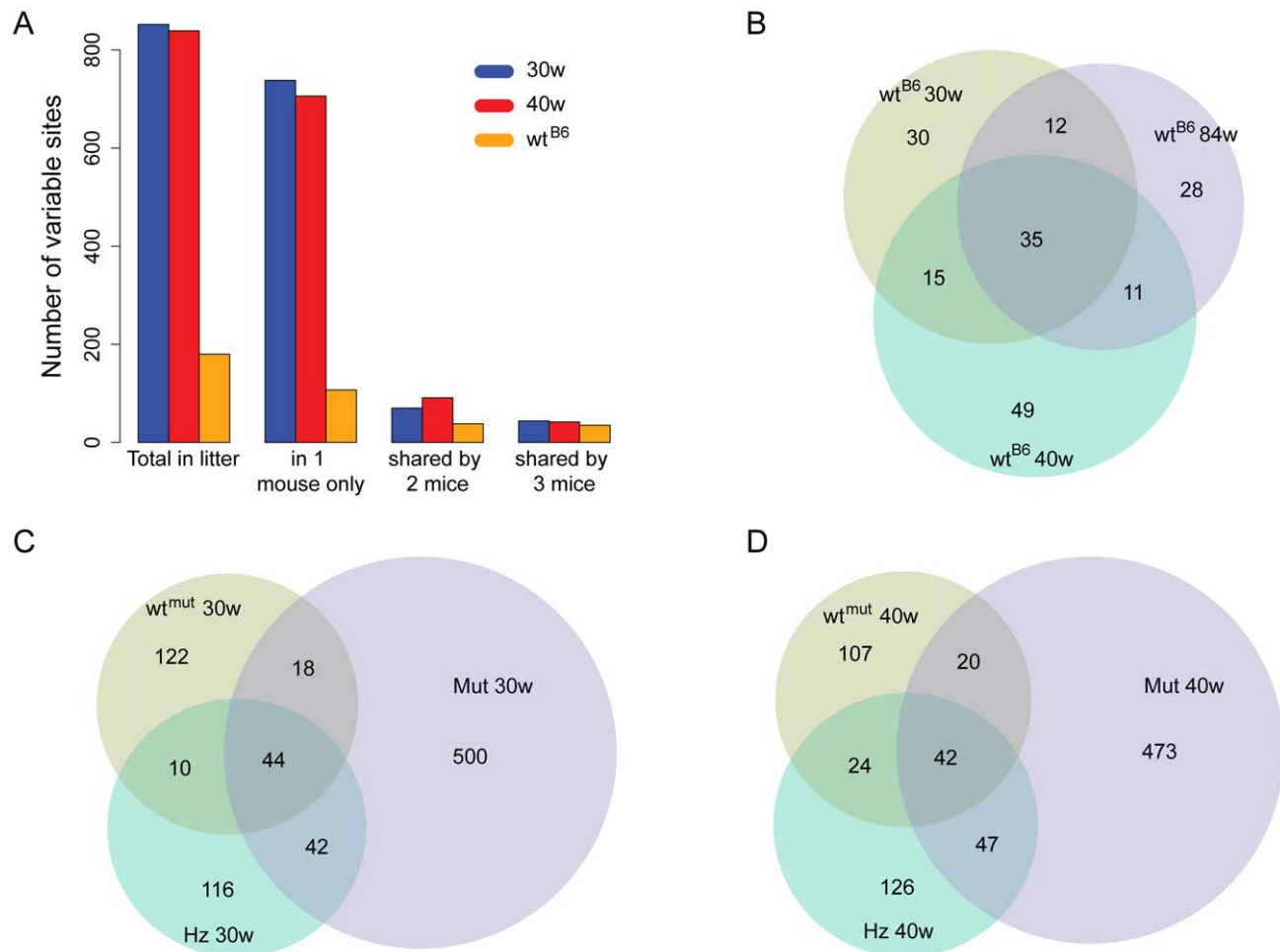


Figure 5. Number of shared high-frequency mtDNA mutation in the C57Bl/6N and mtDNA mutator sibling sets. A) Number of high frequency mutations (per site frequency $>0.5\%$) in the 30-week-old (30w; blue) and 40-week-old (40w; red) mutator siblings and C57Bl/6N mice (orange). The columns refer to the total number of such positions (“Total in litter”), positions found in only one of the three mice in each group (“in 1 mouse only”), in two of the three mice in each group (“shared by 2 mice”) and finally those shared between all three mice within each of the groups (“shared by 3 mice”). Venn Diagrams showing the distribution of high frequency mutations that are unique mutations and those shared between, B) C57Bl/6N samples, C) the 30-week-old mtDNA mutator sibling set and D) the 40-week-old mtDNA mutator sibling set. (Abbreviations: wt^{B6} = C57Bl/6N animal, wt^{mut} = wildtype sibling of the mtDNA mutator mouse, Hz = heterozygous sibling of the mtDNA mutator mouse, Mut = mtDNA mutator mouse, 30w and 40w = 30 weeks of age and 40 weeks of age, respectively).

doi:10.1371/journal.pgen.1002028.g005

mtDNA mutations in mice of the different genotypes. Together, the data we present here suggest that most mtDNA mutations are due to mtDNA replication errors and that oxidative damage of mtDNA does not drive the aging process in liver.

Our estimates of the mutation load in the wt^{B6} and wt^{mut} mice are in good agreement with estimates based on sequencing of cloned PCR fragments [12–14], but they are higher than estimates obtained by a restriction enzyme digestion-based assay [30] (summarized in Table S6). Also, we observed no increase in mutation load in wt^{B6} mice with age, as reported elsewhere [11,30]. A possible explanation for this difference from results in the literature is that the next-generation sequencing method has an inherent limitation in detecting extremely low levels of mutations. It remains possible that there is a slight increase of mutation load with age in wt^{B6} mice, as reported in other studies [11], and the true mutation levels may be below the detection threshold of the SOLiD sequencing method. But in any case, the mutation levels seen in the aging wt^{B6} mice were very modest in comparison to those of age-matched mtDNA mutator mice. Also, we chose to only analyze mtDNA from mouse liver as this tissue made it possible for us to obtain sufficient quantities of pure mtDNA for direct sequencing, without excess DNA amplification. A continuously dividing tissue like liver may show a different pattern of accumulation of mtDNA point mutations with age in comparison with a postmitotic tissue such as brain or heart.

The different coding regions showed similar mutation frequencies, while the control region had a much lower mutation frequency. A reduction in the overall mutation load has previously been observed in the mitochondrial control region of mtDNA mutator mice [12–13,29]. The control region contains crucial sequence elements controlling replication and transcription of mtDNA [31]. It is therefore likely that mutations that inhibit mtDNA maintenance or expression could undergo strong selection and be eliminated from the mtDNA pool.

Some of the mtDNA point mutations observed in siblings to mtDNA mutator mice are likely to have been inherited via the maternal gamete of their heterozygous mother. The wt^{mut} mice from this cross carry approximately twice the number of high frequency point mutations (>0.5%) in comparison with wt^{B6} mice. In addition, an elevated number of small indel mutations are observed in the wt^{mut} mice, but not in wt^{B6} mice (Table S4). We cannot exclude that some of the shared point mutations represent extreme mutational hotspots, however, a more likely explanation is that these shared mutations are maternally inherited. A similar phenomenon has been observed in humans, where the variability at several positions was shown to be inherited from the common maternal cytoplasm instead of representing repeated *de novo* mutational events [22].

In mtDNA mutator mice, approximately 30% of the mtDNA molecules are non-replicating, linear mtDNA molecules with large deletions [13,19]. It could be speculated that these molecules contain an elevated mutation load and that most of the mutation load is sequestered in these molecules. By ultra-deep sequencing, we were able to determine that the mutation load in the region covering the linear molecule did not vary from the global mtDNA mutation load in the mutators. A recent publication suggests the PolgA^{mut} polymerase may pause at the control and OriL regions during mtDNA replication, which may explain the generation of the linear molecules with large deletions [19]. Our results are consistent with a hypothesis that altered processivity of the mutator polymerase and not the point mutations *per se*, are responsible for the creation of the linear molecules. The physiological consequences of the linear molecules, and their contribution to the premature aging in the mtDNA mutator mice, remain unclear and worthy of further investigation.

Circular mtDNA molecules with deletions have been suggested to be the driving force of the aging phenotype in the mutator mice, and are reliably detected in human tissues during aging. High levels of these circular mtDNA molecules with deletions lead to mitochondrial dysfunction in human patients [32] or mice engineered to contain these mutations [33–34]. However, we found that the circular mtDNA molecules with deletions are exceedingly rare in mtDNA mutator mice, with only 4 breakpoints being detected in the millions of reads of two mutator samples. Very low levels of this type of deleted molecules have also been reported by studies using different PCR based analyses [15,20]. Recently, an independent next-generation sequencing analysis detected exceedingly low levels of this type of mutation in brain and heart of mtDNA mutator mice [21]. Large deletion of mtDNA are known to impair mitochondrial translation due to lack of one or more tRNA genes [35]. However, mtDNA mutator mice do not display impaired mitochondrial translation in heart or liver [15] and the levels of deleted mtDNA are much lower than the levels observed in respiratory chain deficient mouse strains with single [33] and multiple [34] deletions of mtDNA. Together, these observations provide strong evidence that the circular deleted mtDNA molecules are not the causative factor in the aging phenotype of the mtDNA mutator mice. A recent study made a remarkable observation that similar low levels of circular deletions were accumulating in a mouse model with a tissue-specific disruption of the mitochondrial fusion process [36]. The presence of very low levels of circular mtDNA molecules with deletions in two very different models of mitochondrial dysfunction suggests these rare events are being generated as a secondary consequence of mitochondrial dysfunction. Another possibility is that these molecules are continuously generated at low frequency during normal mtDNA replication and that mitochondrial dysfunction limit their clearance.

The large numbers of point mutations in adult mtDNA mutator mice result in production of highly mutated mtDNA-encoded respiratory chain subunits, causing the experimentally observed instability of the respiratory chain complexes [15,37]. There is likely a threshold for the tolerance of point mutations, where eventually the combined effect of the many amino acid changes in mtDNA mutator mice cause destabilization of respiratory chain complexes and lead to mitochondrial dysfunction. Our results support the assertion that the accumulation of point mutations has an adversary effect on mitochondrial function and cause the premature aging syndrome in the mtDNA mutator mice.

Materials and Methods

Preparation of purified mitochondria and mitochondrial DNA

This animal study was approved by the animal welfare ethics committee and performed in compliance with Swedish law. Three C57Bl/6N males mice were obtained from the animal unit's breeding colony. Two sibling sets of three males were also obtained, each containing the three genotypes expected from a PolgA^{mut}/PolgA^{wt} intercross. These two litters were not from the same heterozygous mother.

Mitochondria were isolated using standard protocols. Briefly, whole livers were homogenised under ice-cold conditions and cell debris pelleted by low speed centrifugation (600 g, 4°C for ten minutes). The supernatant was transferred and the mitochondria pelleted by centrifugation at 5000 g for 10 min. Resuspended mitochondria were isolated by centrifugation in a 1.0M/1.5M two-phase sucrose gradient. Isolated mitochondria were lysed in 1% Sarkosyl and DNA purified by organic extraction, followed by

salt precipitation. The DNA preparation was treated with RNase and the DNA precipitated prior to use.

Isolated mtDNA from a liver preparation from a C57Bl/6N mouse was digested in BglII and cloned into lambda using the Lambda FIX II/Xho I and Gigapack III Gold Packaging kits (Stratagene). A single clone was expanded and the DNA was extracted by standard phage DNA extraction protocols.

Sequencing library preparation and DNA sequencing using SOLiD

Sequencing libraries were prepared from 1 µg of purified mtDNA following manufacturer's instructions (ABI). Emulsion-PCR was performed according to the manufacturer's instructions (ABI), and then applied to standard slide and sequenced with 50 base pair read length, using an ABI SOLiD 3 sequencing system. The reads were aligned to the C57Bl/6J mouse mtDNA reference sequence (NC_005089.1), using the corona lite mapping algorithm (Applied Biosystems) with default settings. The first 49 bases of the mtDNA sequence were appended to the end of the reference to avoid that reads fail to align due to the circularity of the mitochondrial genome. This alignment procedure attempts to map each read at full length to the reference sequence, allowing for at most 6 mismatches for each 50 bp read.

Calculating point mutation frequencies

In the SOLiD sequencing technology, a SNV is represented by two valid adjacent mismatches in an aligned read. We used the valid adjacent mismatch calls to calculate the mutation frequencies for all samples at every position of the mtDNA molecule by the following method. At each position the number of nucleotide substitutions were calculated for each of the three alternative bases. By dividing these numbers with the total read coverage we obtained SNV frequencies for each of the three possible mutant alleles, with their sum representing the total mutation frequency at the specific position. This method may include some false positive SNV calls, which implies that our mutation frequencies will be over-estimated. The error frequencies may be dependent both on the sequence context and the sequencing technology used, and is likely to vary substantially at different positions of the mtDNA molecule. We therefore used a λ^{mtDNA} control sample as a means to correct the mutation rates.

Correcting mutation frequencies using cloned mtDNA

The mutation frequencies were calculated as described above for all samples, including the λ^{mtDNA} control. For each of the mtDNA samples, the SNV frequencies for the λ^{mtDNA} clone were subtracted to obtain corrected frequencies of all nucleotide changes at each position. In the cases where the λ^{mtDNA} showed a higher rate of some nucleotide compared to the mtDNA sample, the corrected value was set to zero. Several sites in the λ^{mtDNA} control sample showed elevated rates of mutations. The mutations found in λ^{mtDNA} can partly be explained by technical errors in the SOLiD sequencing, but there may also be a set of true polymorphisms that were incorporated during the replication of the λ^{mtDNA} clone. Since we cannot distinguish between these two sources of variability we have taken a conservative approach and subtracted the entire λ^{mtDNA} mutation signal from the mouse mtDNA samples. As a consequence there may be sites where the λ^{mtDNA} mutation rate is higher than in the native mtDNA.

Estimating number of protein coding nucleotide changes per mtDNA molecule

We used our SNV frequencies as an estimate of the average number of mutations per mtDNA molecule. All SNVs within

protein coding genes were extracted and their frequencies were calculated. These SNVs were then grouped into three different categories, 'synonymous', 'amino acid change' or 'stop codon', depending on the effect on the protein sequence. For each of the categories the number of changes per mtDNA molecule was calculated as the sum of all SNV frequencies belonging to that group. To remove the effect of extreme mutational hotspots, all mutations with an observed per site frequency of >0.5% were excluded from this analysis. If such sites are not removed there is a risk that a few sites with extremely high frequency will have too large an effect on the estimate since their frequencies can sometimes be 1,000 times higher than at other sites. By removing the highest frequency sites we instead focus on the combined effect of the lower frequency mutations on the proteome. These estimates are therefore likely to be conservative.

Indel analysis

Reads containing insertions and deletions will not be aligned to the reference sequence with the corona lite program. The unmapped reads were analyzed for indels using the SplitSeek strategy [28]. This strategy was originally developed for junction detection in RNA-seq data, but it can also be used for detecting insertions and deletions. The unmapped reads were aligned using version 1.1 of the AB WT pipeline (<http://solidssoftwaretools.com/gf/project/transcriptome/>), a software that performs split read alignment of SOLiD data, using the same settings as in the RNA-seq study [28] and the alignment results were used as input to the SplitSeek program. We required each indel to be supported by at least 5 reads with unique starting points, and by reads on both strands. For each indel we calculated a frequency for its occurrence in the mtDNA samples by the ratio $r_i/(r_i+r_c)$. In this formula r_i is the number of reads that supports an indel, while r_c is the total coverage over the indel calculated from the initial full-length mapping of reads.

Identifying mutational hotspots

Mutational hotspots were defined as regions with elevated mutation frequency over at least 20 bases. The regional hotspots were identified by first calculating the median SNV frequency in a 20 bp window around each base in each sample to obtain a smoothed signal of the mutation rates. For each sample, we then selected those positions with a median window frequency above the 90th percentile of all SNV frequencies in the entire mtDNA genome. In this way we select only those positions with a substantially elevated mutation rate over a number of neighbouring bases. Furthermore, we required the same region to be identified in at least three of the nine mice. This analysis resulted in the 12 hotspot regions presented in Table S2.

Identifying shared high frequency mutation sites

We used the same cut-off to detect high frequency SNVs as was used for filtering out positions with high signal in the negative control sample (i.e. a per site frequency of 0.005 or 0.5%). For each of the samples we extracted all positions with mutation frequencies of at least 0.5%.

Mitochondrial fragments in the nuclear genome

Parts of the mitochondrial genome can exist as nuclear-inserted mitochondrial pseudogenes (NucMts) [38]. In mice, these NucMts normally differ substantially from the mtDNA, though a large insert with few variations from the C57Bl/6J mtDNA sequence is known [39]. Since the purified mitochondrial fraction will be contaminated with small amounts of nuclear DNA, sequences that

appear to be mtDNA but are derived from NucMts could be present in the sequenced reads and affect the mutation rate. However, mtDNA represents about 1% of the DNA in a cell. Assuming a ratio of 90:10 between mitochondrial and nuclear DNA in the sample, a sequencing reaction that generates a total of 1×10^9 bases will then include 9×10^8 bases of mtDNA from the mitochondria and the remaining from the nuclear genome. If there are 100 mtDNA fragments inserted in each nuclear genome, the 1×10^8 bases that derives from the nuclear genome represent roughly 4% of a nuclear mouse genome, or about 4 mtDNA fragments. These sequences should be compared to the 9×10^8 bases of mtDNA sequence from the mitochondria, which correspond to about 5×10^4 mtDNA genomes. Given that only a fraction of the nuclear insert sequences have a nucleotide that deviates from the consensus, nuclear inserts are likely to contribute less than one deviating read out of 5×10^4 reads at each position. Thus, the effect of nuclear inserts on the estimate is likely to be very small.

Data accession

The SOLiD data is available from the EMBL-EBI Sequence Read Archive. Study accession number: ERP000469 (<http://www.ebi.ac.uk/ena/data/view/ERP000469>).

Supporting Information

Figure S1 Sequence coverage overview. The lines represent the coverage for the different samples. (DOC)

Figure S2 Distribution of sites for insertions and deletions in the mtDNA genome. (DOC)

References

- Harman D (1972) The biologic clock: the mitochondria? *J Am Geriatr Soc* 20: 145–147.
- Miquel J, Economos AC, Fleming J, Johnson JE, Jr. (1980) Mitochondrial role in cell aging. *Exp Gerontol* 15: 575–591.
- Krishnan KJ, Greaves LC, Reeve AK, Turnbull D (2007) The ageing mitochondrial genome. *Nucleic Acids Res* 35: 7399–7405.
- Larsson NG (2010) Somatic mitochondrial DNA mutations in mammalian aging. *Annu Rev Biochem* 79: 683–706.
- Cortopassi GA, Arnheim N (1990) Detection of a specific mitochondrial DNA deletion in tissues of older humans. *Nucleic Acids Res* 18: 6927–6933.
- Ikebe S, Tanaka M, Ohno K, Sato W, Hattori K, et al. (1990) Increase of deleted mitochondrial DNA in the striatum in Parkinson's disease and senescence. *Biochem Biophys Res Commun* 170: 1044–1048.
- Katayama M, Tanaka M, Yamamoto H, Ohbayashi T, Nimura Y, et al. (1991) Deleted mitochondrial DNA in the skeletal muscle of aged individuals. *Biochem Int* 25: 47–56.
- Yen TC, Su JH, King KL, Wei YH (1991) Ageing-associated 5 kb deletion in human liver mitochondrial DNA. *Biochem Biophys Res Commun* 178: 124–131.
- Hattori K, Tanaka M, Sugiyama S, Obayashi T, Ito T, et al. (1991) Age-dependent increase in deleted mitochondrial DNA in the human heart: possible contributory factor to presbycardia. *Am Heart J* 121: 1735–1742.
- Schwarze SR, Lee CM, Chung SS, Roecker EB, Weindruch R, et al. (1995) High levels of mitochondrial DNA deletions in skeletal muscle of old rhesus monkeys. *Mech Ageing Dev* 83: 91–101.
- Piko L (1992) Accumulation of mtDNA defects and changes in mtDNA content in mouse and rat tissues with aging. *Ann N Y Acad Sci* 663: 450–452.
- Trifunovic A, Hansson A, Wredenberg A, Rovio AT, Dufour E, et al. (2005) Somatic mtDNA mutations cause aging phenotypes without affecting reactive oxygen species production. *Proc Natl Acad Sci U S A* 102: 17993–17998.
- Trifunovic A, Wredenberg A, Falkenberg M, Spelbrink JN, Rovio AT, et al. (2004) Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature* 429: 417–423.
- Kujoth GC, Hiona A, Pugh TD, Someya S, Panzer K, et al. (2005) Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. *Science* 309: 481–484.
- Edgar D, Shabalina I, Camara Y, Wredenberg A, Calvaruso MA, et al. (2009) Random point mutations with major effects on protein-coding genes are the driving force behind premature aging in mtDNA mutator mice. *Cell Metab* 10: 131–138.
- Vermulst M, Wanagat J, Kujoth GC, Bielas JH, Rabinovitch PS, et al. (2008) DNA deletions and clonal mutations drive premature aging in mitochondrial mutator mice. *Nat Genet* 40: 392–394.
- Vermulst M, Wanagat J, Loeb LA (2009) On mitochondria, mutations, and methodology. *Cell Metab* 10: 437.
- Edgar D, Larsson NG, Trifunovic A (2010) Response: Point Mutations Are Causing Progeroid Phenotypes in the mtDNA Mutator Mouse. *Cell Metab* 11: 93.
- Bailey LJ, Cluett TJ, Reyes A, Prolla TA, Poulton J, et al. (2009) Mice expressing an error-prone DNA polymerase in mitochondria display elevated replication pausing and chromosomal breakage at fragile sites of mitochondrial DNA. *Nucleic Acids Res* 37: 2327–2335.
- Kraytsberg Y, Simon DK, Turnbull DM, Khrapko K (2009) Do mtDNA deletions drive premature aging in mtDNA mutator mice? *Aging Cell* 8: 502–506.
- Williams SL, Huang J, Edwards YJ, Ulloa RH, Dillon LM, et al. (2010) The mtDNA mutation spectrum of the progeroid Polg mutator mouse includes abundant control region multimers. *Cell Metab* 12: 675–682.
- He Y, Wu J, Dressman DC, Iacobuzio-Donahue C, Markowitz SD, et al. (2010) Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. *Nature* 464: 610–614.
- Li M, Schonberg A, Schaefer M, Schroeder R, Nasidze I, et al. (2010) Detecting heteroplasmy from high-throughput sequencing of complete human mitochondrial DNA genomes. *Am J Hum Genet* 87: 237–249.
- Stewart JB, Freyer C, Elson JL, Wredenberg A, Cansu Z, et al. (2008) Strong Purifying Selection in Transmission of Mammalian Mitochondrial DNA. *PLoS Biol* 6: e10.1371/journal.pbio.0060010.
- Khrapko K, Coller HA, Andre PC, Li XC, Hanekamp JS, et al. (1997) Mitochondrial mutational spectra in human cells and tissues. *Proc Natl Acad Sci U S A* 94: 13798–13803.
- Zheng W, Khrapko K, Coller HA, Thilly WG, Copeland WC (2006) Origins of human mitochondrial point mutations as DNA polymerase gamma-mediated errors. *Mutat Res* 599: 11–20.
- Elliott HR, Samuels DC, Eden JA, Relton CL, Chinnery PF (2008) Pathogenic mitochondrial DNA mutations are common in the general population. *Am J Hum Genet* 83: 254–260.

Table S1 Samples and sequence read information. (DOC)

Table S2 Transition and transversion ratios of observed mutations. (DOC)

Table S3 Mutational hotspot regions. (DOC)

Table S4 Small indels (<10 bp) observed in the samples by the SplitSeek method. (DOC)

Table S5 Large deletions (>10 bp) observed in the samples by the SplitSeek method. (DOC)

Table S6 Comparison of obtained mutation loads to literature values. Values reported as the per-site mutation $\times 10^{-4}$. (DOC)

Acknowledgments

The SOLiD sequencing was performed at Swedish National Infrastructure for Large-scale DNA sequencing (SNISS) and SciLifeLab Genomics Platform in Uppsala.

Author Contributions

Conceived and designed the experiments: UG N-GL AA JBS. Performed the experiments: JBS CF EH. Analyzed the data: AA MI JBS. Wrote the paper: AA JBS UG N-GL.

28. Ameur A, Wetterbom A, Feuk L, Gyllenstein U (2010) Global and unbiased detection of splice junctions from RNA-seq data. *Genome Biol* 11: R34.
29. Rovio AT (2006) DNA Polymerase Gamma Mutations in Male Infertility and Ageing. Tampere: University of Tampere. 137 p.
30. Vermulst M, Bielas JH, Kujoth GC, Ladiges WC, Rabinovitch PS, et al. (2007) Mitochondrial point mutations do not limit the natural lifespan of mice. *Nat Genet* 39: 540–543.
31. Falkenberg M, Larsson NG, Gustafsson CM (2007) DNA replication and transcription in mammalian mitochondria. *Annu Rev Biochem* 76: 679–699.
32. Holt IJ, Harding AE, Morgan-Hughes JA (1988) Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. *Nature* 331: 717–719.
33. Inoue K, Nakada K, Ogura A, Isobe K, Goto Y, et al. (2000) Generation of mice with mitochondrial dysfunction by introducing mouse mtDNA carrying a deletion into zygotes. *Nat Genet* 26: 176–181.
34. Tyynismaa H, Mjosund KP, Wanrooij S, Lappalainen I, Ylikallio E, et al. (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.
35. Nakase H, Moraes CT, Rizzuto R, Lombes A, DiMauro S, et al. (1990) Transcription and translation of deleted mitochondrial genomes in Kearns-Sayre syndrome: implications for pathogenesis. *Am J Hum Genet* 46: 418–427.
36. Chen H, Vermulst M, Wang YE, Chomyn A, Prolla TA, et al. (2010) Mitochondrial fusion is required for mtDNA stability in skeletal muscle and tolerance of mtDNA mutations. *Cell* 141: 280–289.
37. Kukat A, Trifunovic A (2009) Somatic mtDNA mutations and aging—facts and fancies. *Exp Gerontol* 44: 101–105.
38. Jensen-Seaman MI, Wildschutte JH, Soto-Calderon ID, Anthony NM (2009) A comparative approach shows differences in patterns of numt insertion during hominoid evolution. *J Mol Evol* 68: 688–699.
39. Bayona-Bafaluy MP, Acin-Perez R, Mullikin JC, Park JS, Moreno-Loshuertos R, et al. (2003) Revisiting the mouse mitochondrial DNA sequence. *Nucleic Acids Res* 31: 5349–5355.