

# Ultra-Sensitive Sequencing Reveals an Age-Related Increase in Somatic Mitochondrial Mutations That Are Inconsistent with Oxidative Damage

Scott R. Kennedy<sup>1</sup>, Jesse J. Salk<sup>1,2</sup>, Michael W. Schmitt<sup>1,2</sup>, Lawrence A. Loeb<sup>1,3\*</sup>

**1** Department of Pathology, University of Washington, Seattle, Washington, United States of America, **2** Department of Medicine, University of Washington, Seattle, Washington, United States of America, **3** Department of Biochemistry, University of Washington, Seattle, Washington, United States of America

## Abstract

Mitochondrial DNA (mtDNA) is believed to be highly vulnerable to age-associated damage and mutagenesis by reactive oxygen species (ROS). However, somatic mtDNA mutations have historically been difficult to study because of technical limitations in accurately quantifying rare mtDNA mutations. We have applied the highly sensitive Duplex Sequencing methodology, which can detect a single mutation among  $>10^7$  wild type molecules, to sequence mtDNA purified from human brain tissue from both young and old individuals with unprecedented accuracy. We find that the frequency of point mutations increases  $\sim 5$ -fold over the course of 80 years of life. Overall, the mutation spectra of both groups are comprised predominantly of transition mutations, consistent with misincorporation by DNA polymerase  $\gamma$  or deamination of cytidine and adenosine as the primary mutagenic events in mtDNA. Surprisingly, G $\rightarrow$ T mutations, considered the hallmark of oxidative damage to DNA, do not significantly increase with age. We observe a non-uniform, age-independent distribution of mutations in mtDNA, with the D-loop exhibiting a significantly higher mutation frequency than the rest of the genome. The coding regions, but not the D-loop, exhibit a pronounced asymmetric accumulation of mutations between the two strands, with G $\rightarrow$ A and T $\rightarrow$ C mutations occurring more often on the light strand than the heavy strand. The patterns and biases we observe in our data closely mirror the mutational spectrum which has been reported in studies of human populations and closely related species. Overall our results argue against oxidative damage being a major driver of aging and suggest that replication errors by DNA polymerase  $\gamma$  and/or spontaneous base hydrolysis are responsible for the bulk of accumulating point mutations in mtDNA.

**Citation:** Kennedy SR, Salk JJ, Schmitt MW, Loeb LA (2013) Ultra-Sensitive Sequencing Reveals an Age-Related Increase in Somatic Mitochondrial Mutations That Are Inconsistent with Oxidative Damage. *PLoS Genet* 9(9): e1003794. doi:10.1371/journal.pgen.1003794

**Editor:** Bennett Van Houten, University of Pittsburgh, United States of America

**Received:** April 22, 2013; **Accepted:** July 29, 2013; **Published:** September 26, 2013

**Copyright:** © 2013 Kennedy 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:** Support for this research was provided by the following grants: NIA P01-AG01751, NCI P01-CA77852, and NCI R01-CA102029 to LAL. SRK was further supported by the Genetic Approaches to Aging Training Grant (NIA T32-AG000057). Tissues were collected by the University of Washington Alzheimer's Disease Research Center (P50AG05136). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

\* E-mail: laloeb@u.washington.edu

## Introduction

Mitochondria are the primary source of energy for cells. Owing to their evolutionary history, these organelles harbor a small, independently replicated genome (mtDNA). Human mtDNA encodes two rRNA genes, 13 protein coding genes that are essential components of the electron transport chain (ETC), and a full complement of 22 tRNAs used in translation of the ETC peptides. The escape of electrons from the ETC can lead to the formation of reactive oxygen species (ROS), which are capable of damaging a variety of cellular components, including DNA. Due to its proximity to the ETC, absence of protective histones, and a lack of nucleotide excision or mismatch repair, mtDNA is thought to be especially vulnerable to ROS-mediated damage and the generation of mutations. Failure to faithfully transmit the encoded information during mtDNA replication leads to the production of dysfunctional ETC proteins, leading to the release of more free electrons and ROS in what has been termed 'the vicious cycle' [1,2]. Thus, it is not surprising that mutations in mtDNA have been associated with a decline in energy production, a loss of organismal fitness, an increased

propensity for a number of pathological conditions, and aging (reviewed in [3,4]).

Numerous lines of evidence have suggested mtDNA mutations are involved in the aging process. In particular, ETC activity declines with age [5,6], and this decrease is coincident with accumulation of mitochondria with large deletions in their mtDNA [7,8,9,10]. Large, kilobase-sized deletions in mtDNA become more prevalent with age in a variety of tissues, including brain [11], heart [12], and skeletal muscle [7]. Furthermore, these large deletions have been shown to increase in frequency in a number of neurodegenerative conditions, including Parkinson's disease [13,14] and Alzheimer's disease [15]. In addition, DNA damage, predominantly in the form of 8-hydroxy-2'-deoxyguanosine (8-oxo-dG) [16], increases with age in both nuclear and mitochondrial DNA [17,18,19,20]. While the role of mtDNA deletions in aging is well established, the role of point mutations remains controversial [21,22].

Several previous studies have examined the accumulation of point mutations in human aging and disease [23,24,25,26]. Until very recently, hypotheses that required the observation of rare mutations in mtDNA have been extremely difficult to experimen-

## Author Summary

Owing to their evolutionary history, mitochondria harbor independently replicating genomes. Failure to faithfully transmit the genetic information of mtDNA during replication can lead to the production of dysfunctional electron transport proteins and a subsequent decline in energy production. Cellularly-derived reactive oxygen species (ROS) and environmental agents preferentially damage mtDNA compared to nuclear DNA. However, little is known about the consequences of mtDNA damage for mutagenesis. This lack of knowledge stems, in part, from an absence of methods capable of accurately detecting these mutations throughout the mitochondrial genome. Using a new, highly sensitive DNA sequencing strategy, we find that the frequency of point mutations is 10–100-fold lower than what has been previously reported using less precise means. Moreover, the frequency increases 5-fold over an 80 year lifespan. We also find that it is predominantly transition mutations, rather than mutations commonly associated with oxidative damage to mtDNA, that increase with age. This finding is inconsistent with free radical theories of aging. Finally, the mutagenic patterns and biases we observe in our data are similar to what is seen in population studies of mitochondrial polymorphisms and suggest a common mechanism by which somatic and germline mtDNA mutations arise.

tally validate due to: 1) the lack of genetic tools for introducing reporters or selectable markers into mtDNA; 2) the high background error rate of most DNA sequencing methods [27,28]; and 3) the sampling limitations of the few available high-sensitivity mutation assays that screen only a tiny subset of the genome [29]. The mitochondrial genome is 16,569 bp, and individual human cells frequently contain hundreds to thousands of molecules of mtDNA; thus, a single human cell typically contains millions of nucleotides of mtDNA sequence. The rate of accumulation of mtDNA mutations has previously been estimated as  $6 \times 10^{-8}$  mutations per base pair per year [30]. Therefore, reliable study of spontaneous mtDNA mutations requires methodologies that can accurately detect a single mutation among  $>10^6$  wild-type base-pairs. However, most prior studies of mtDNA mutations and aging have relied upon methods with background error frequencies of  $10^{-3}$  to  $10^{-4}$ ; hence the many reported differences likely reflect changes in mutation clonality or technical artifacts (e.g. due to increases in DNA damage with age) rather than true spontaneous mutations.

Massively parallel sequencing technologies allow mtDNA to be subjected to ‘deep sequencing’ in order to detect rare/sub-clonal mutations on a genome-wide level. However, these new sequencing methods are highly error prone, with artifactual error rates of approximately one spurious mutation per 100 to 1,000 nucleotides sequenced. These high error rates have precluded the study of spontaneous mutations in mtDNA [31]. To circumvent this limitation, we recently developed a new, highly accurate sequencing methodology termed Duplex Sequencing (DS), which has the unique ability to detect a single mutation among  $>10^7$  sequenced bases [32].

In the study herein, we determined the effect of aging on mtDNA mutation burden by using DS to compare human mtDNA purified from brain tissue of five young individuals (ages  $<1$ ) and five aged individuals (ages 75–99 years) obtained via rapid autopsy (Table S1). As brain is among the most metabolically active tissues in the human body, we reasoned it

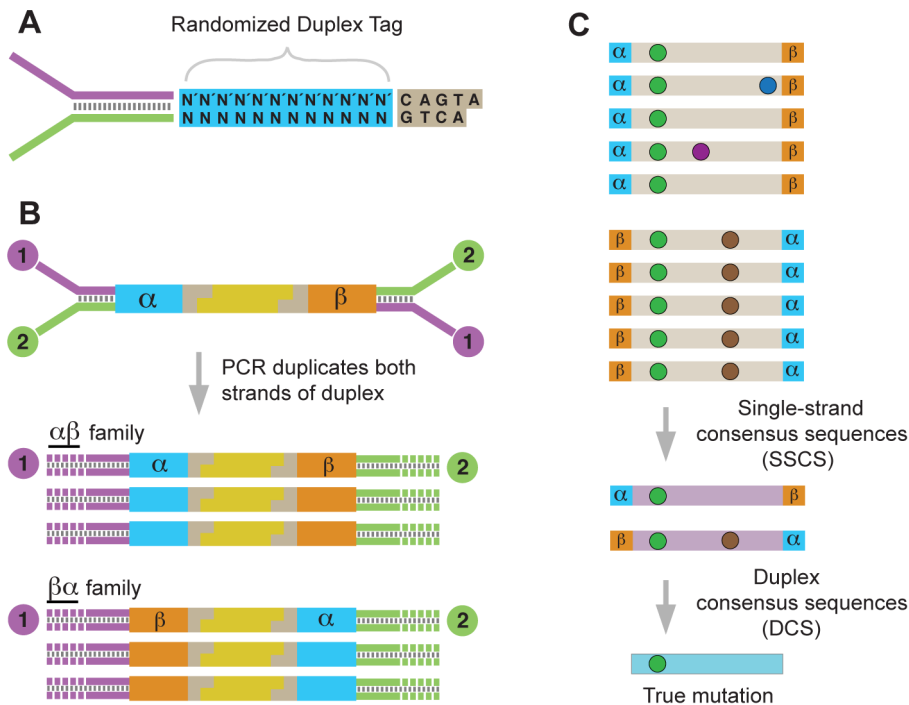
to be particularly prone to damage from ROS, and thus, an optimal tissue for comparison between age groups. We assessed the relative frequency, spectrum, and distribution of mtDNA mutations in the two groups. We find that point mutations increase with age, but do so in a non-uniform manner. Furthermore, we find that mutations show a bias in their occurrence with respect to both genome location and strand orientation. The types of mutations we detect are inconsistent with oxidative damage being a major driver of mtDNA mutagenesis.

## Results

Duplex Sequencing relies on the concept of molecular tagging and the fact that the two strands of DNA contain complementary information. Fragmented duplex DNA is tagged with adapters bearing a random, yet complementary, double-stranded nucleotide sequence (Fig. 1A). Following ligation, the individually labeled strands are PCR amplified, thus creating sequence “families” that share a common tag sequence, derived from each of the two single parental strands (Fig. 1B). After sequencing, members of each duplicate family are grouped by tag, and a consensus sequence is calculated for each family, creating a single strand consensus sequence (SSCS) (Fig. 1C). This step eliminates random sequencing or PCR errors that occur during library amplification; however, the single-stranded consensus process does not filter out artifactual mutations that are the consequence of first round PCR errors, such as those commonly caused by DNA damage. To remove this latter type of error, the complementary SSCS families derived from the two single-stranded halves of the original DNA duplex are compared to each other (Fig. 1C). The base identity at each position in a read is kept in the final consensus only if the two strands match perfectly at that position. Upon remapping of these duplex consensus sequence (DCS) reads back to the reference genome, any deviations from the reference sequence are considered true mutations. The frequency of mutations in a sampled population of mtDNA is calculated as the number of DCS mutant molecules divided by the number of DCS wild-type molecules observed at any given genomic position.

### Point mutations accumulate with aging in human mtDNA

Point mutations in mtDNA could be the result of maternal inheritance or a *de novo* mutation event. Maternally inherited mutations or mutations arising during early embryonic development are more likely to be clonal (i.e. the same mutation being present at the same location in most or all mtDNA molecules). Therefore, in order to quantify the frequency of *de novo* events, we used a clonality cutoff that excluded any positions with variants occurring at a frequency of  $>1\%$ , and scored each type of mutation only once at each position of the genome. Based on these criteria, the mtDNA from aged individuals show a highly significant  $\sim 5$ -fold increase in mutational frequency, relative to those obtained from young individuals (Young:  $3.7 \pm 0.9 \times 10^{-6}$  vs. Aged:  $1.9 \pm 0.2 \times 10^{-5}$ ,  $p < 10^{-4}$ , two-sample t-test) (Fig. 2A). These mutation frequencies are between one and two orders of magnitude lower than the previously reported values for both young and old individuals using PCR-based methods or conventional next-generation sequencing [24,33,34]. This discordance likely stems from artifactual scoring of mutations by these latter methods due to misinsertion of incorrect bases at sites of damage in template DNA during the PCR steps. Duplex Sequencing, in contrast, is unaffected by DNA damage [32].



**Figure 1. Overview of the Duplex Sequencing methodology.** (A) Adapter design with random double-stranded tag sequence and invariant spacer sequence. (B) Ligation of adapters to fragmented DNA generates unique 12 bp tags on each end ( $\alpha$  and  $\beta$ ). PCR amplification of the two strands produces two related, but distinct products. (C) Sequence reads sharing unique  $\alpha$  and  $\beta$  tags are grouped into families of  $\alpha$ - $\beta$  or  $\beta$ - $\alpha$  orientation. Mutations are of three different types: sequencing mistakes or late arising PCR errors (blue or purple spots); first round PCR errors (brown spots); true mutations (green spots). Comparing SSCSs from the paired families generates a DCS, which eliminates all but true mutations. doi:10.1371/journal.pgen.1003794.g001

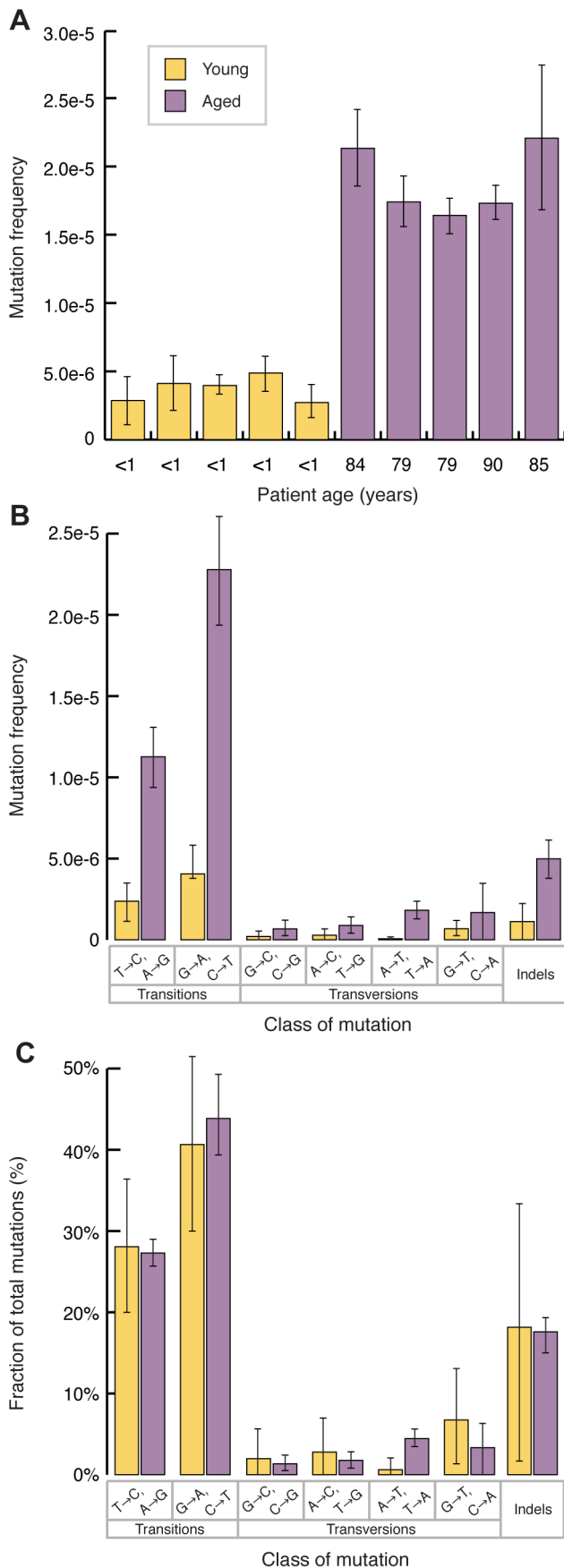
Inspection of the mutation spectra for both the young and old samples reveals that all samples are significantly biased towards transitions (Fig. 2B). Specifically, the most common mutation type,  $G \rightarrow A/C \rightarrow T$ , is consistent with either misincorporation by DNA polymerase  $\gamma$  or deamination of cytosine to form uracil, as being the largest mutagenic drivers in mtDNA [35,36]. The second most common mutation type,  $T \rightarrow C/A \rightarrow G$ , is consistent with either deamination of adenosine to inosine or a T-dGTP mispairing, the primary base misinsertion mistake made by DNA polymerase  $\gamma$  [37,38,39,40]. Plotting the frequency of each type of mutation as a proportion of total mutations (Fig. 2C) reveals that the relative abundance of each mutation type is the same in young and old, suggesting that the mutagenic pressures that result in the observed spectra are constant throughout the human lifespan.

Surprisingly, comparison of the mutation spectra of the young and old samples reveals a notable absence of the mutational signature of oxidative damage. A number of studies have shown that oxidative damage to DNA accumulates in both the nuclear and mitochondrial genomes as a function of age, as well as several age-associated pathologies [17,18,19,20,41]. The most frequent alteration produced by oxidative damage is 8-oxo-dG, which, when copied during replication or repair, results in dA substitutions, yielding  $G \rightarrow T/C \rightarrow A$  transversions [42]. A number of theories of aging invoke ROS-mediated damage to mtDNA as being a major driver of the aging phenotype (reviewed in [43] and [44]). A key prediction for these theories is that the frequency of  $G \rightarrow T/C \rightarrow A$  mutations would be expected to increase with time. We failed to find either a preponderance of  $G \rightarrow T/C \rightarrow A$  substitutions or a proportionally greater increase with age in this type of mutation relative to other types, despite a span of >80 years between our sequenced sample groups (Fig. 2C).

### Deleterious mutations increase with age

Our data indicate that point mutations increase with age and that these mutations are inconsistent with oxidative damage being a primary driver of mutagenesis; we next assessed whether these mutations lead to alterations in the protein coding sequence. We find that in the aged samples, 78.3% of mutations are non-synonymous. The incidence of non-synonymous mutations is close to the expected value of 75.7% for mtDNA that would occur if non-synonymous and synonymous mutations occur randomly. In contrast, only 62.9% of mutations are non-synonymous in the young samples. The reduced mutation load observed in the young samples is consistent with that negative intergenerational selection against such mutations and that this selection is relieved during development and could play a role in the aging phenotype.

However, the existence of a high load of non-synonymous mutations does not necessarily mean that the coding changes lead to functional protein alteration. To examine this possibility, we compared the predicted “pathogenicity” of all non-synonymous mutations in both the young and aged samples using MutPred [45], a software package that calculates the likelihood of a mutation being deleterious based on a number of criteria, including protein structure, the presence of functional protein motifs, evolutionary conservation, and amino acid composition bias. A score between zero and one is assigned to each mutation, with a higher score denoting a higher likelihood of being deleterious. Based on this analysis (Fig. S1), the predicted pathogenicity of mutations, indeed, increases with age ( $p < 0.02$ , Wilcoxon Rank Sum analysis), suggesting that mutations acquired during aging may have functional consequences for the electron transport chain. A similar increase in predicted deleterious mutations was also observed using the SWIFT software package



**Figure 2. Mitochondrial point mutations increase with age and are biased to transitions.** (A) mtDNA point mutations burden is higher in older individuals (purple) than young individuals (yellow) ( $p < 10^{-4}$ , two-tailed t-test). Error bars represent the 95% confidence interval for each sample (Wilson Score interval) (B) The mutation spectra of both young (yellow) and aged (purple) individuals shows an excess of transitions, relative to transversions. Frequencies were calculated by dividing the number of mutations of each type by the number of times the wild-type base of each mutation type was sequenced. Indels were calculated independently as events per total number of bases sequenced. Error bars represent one standard deviation. (C) The mutation spectra, reported as the relative proportion of the different mutation types, do not change with age. Error bars represent one standard deviation. Significance was tested using the two-tailed t-test. doi:10.1371/journal.pgen.1003794.g002

(data not shown). The increase in predicted pathogenicity is consistent with mutations causing coding changes occurring randomly and argues against a mechanism by which point mutations are selected against by the cell. Similar finding in clonally expanded mutations were recently reported in colon tissue show a similar increase in predicted pathogenic mutation in mtDNA [46].

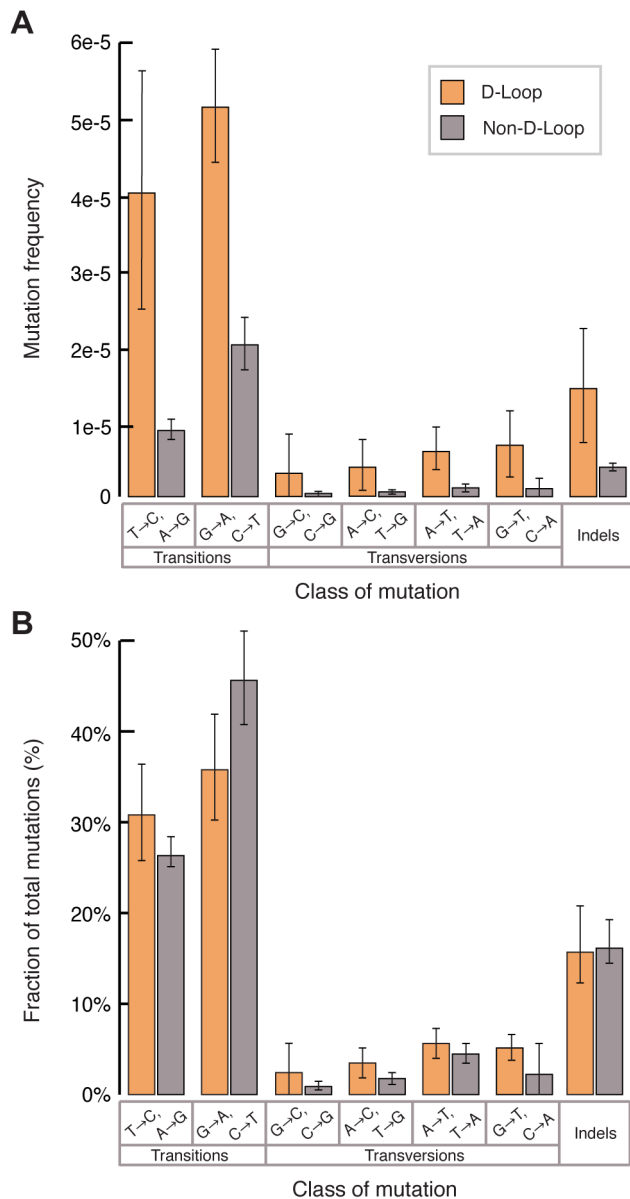
### The D-loop of mtDNA exhibits an elevated mutation burden but is not a mutagenic 'hotspot'

The mitochondrial genome can be divided into three different regions: 1) protein coding genes, 2) RNA coding genes (consisting of both rRNA and tRNA), and 3) non-coding/regulatory regions including the origin of replication known as the D-loop. Phylogenetic analysis of both human and other mammalian lineages has shown that population level single nucleotide variants (SNVs) tend to cluster in a number of 'hotspots' in the mitochondrial genome, most notably in Hypervariable Regions I and II of the D-loop [47,48,49]. We sought to determine if the distribution of non-clonal mutations within the mtDNA of individuals exhibited a uniform distribution or if certain regions of the genome similarly show variations in mutation frequency. Comparison of the mutation frequencies of the RNA coding genes to the protein coding genes yielded no significant differences in either the young or old samples ( $p = 0.15$ , two-tailed t-test).

In contrast, we observed a significant increase in mutation frequency of the D-loop (bp 16024-576) relative to the coding regions (bp 577-16023) in both young (D-loop:  $1.5 \pm 0.6 \times 10^{-5}$  vs. coding region:  $2.9 \pm 0.7 \times 10^{-6}$ ,  $p < 0.01$ , two-tailed t-test) and aged (D-loop:  $5.7 \pm 1.5 \times 10^{-5}$  vs. coding region:  $1.65 \pm 0.2 \times 10^{-5}$ ,  $p < 0.01$ , two-tailed t-test) samples, suggesting that the D-loop is a mutagenic hotspot. However comparing the relative increase in the mutation frequency of the D-loop between the young and old sample groups ( $3.8 \pm 1.6$ -fold increase) to the relative increase seen between the two sample groups in the non-D-loop regions ( $5.6 \pm 2.0$  fold increase) shows no difference. This finding is inconsistent with the idea that the D-loop accumulates significantly more mutations during aging than the rest of the mitochondrial genome. Spectrum analysis shows a similar predominance of transition mutations in both the D-loop and coding regions of the genome (Fig. 3A), with no significant difference in the relative abundance of the different mutation types (Fig. 3B). Taken together, our data suggest that the mutagenic processes of mtDNA are largely uniform across the genome.

### Mutations accumulate asymmetrically on the two strands of mtDNA

The human mitochondrial genome has a significant bias in the cytosine/guanine composition between the two strands. Specifi-



**Figure 3. The D-loop has an elevated mutation burden but its mutation spectrum is similar to the remainder of the mitochondrial genome.** (A) The D-loop (orange) exhibits a higher aggregate mutation burden than the rest of the genome (grey). Frequencies were calculated by dividing the number of mutations of each type by the number of times the wild-type base of each mutation type was sequenced. Indels were calculated independently as events per total number of bases sequenced. Error bars represent one standard deviation. (B) The relative fraction of mutations exhibits no difference between the D-loop (orange) and non-D-loop (grey) portions of the genome, suggesting a similar underlying mutagenic process. Error bars represent one standard deviation. doi:10.1371/journal.pgen.1003794.g003

cally, the light strand (L-strand), which is the coding strand for only nine genes, contains about three-fold more cytosine than guanine, whereas the heavy strand (H-strand) codes for the remaining 28 genes and has the opposite composition bias. Human population studies, as well as the comparative analysis of evolutionarily related species, have shown a bias towards the occurrence of G→A and T→C SNPs of the L-strand [50,51,52,53]. These population-level compositional biases are

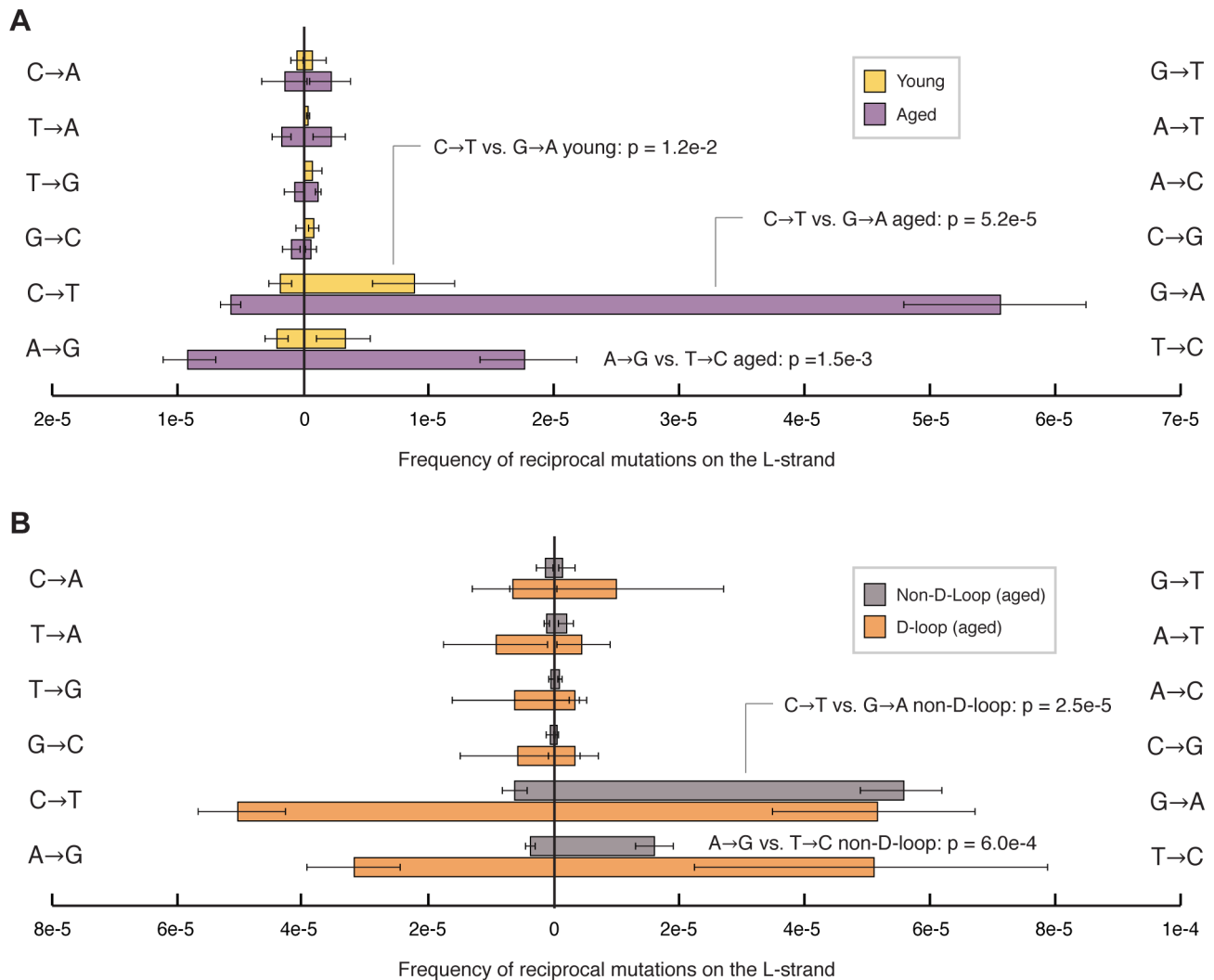
hypothesized to be due to an asymmetric accumulation of mutations between the two strands of mtDNA in the germline; however, to date, the biases have not been observed at the sub-clonal/random level within individuals. To examine this, we compared the frequency of reciprocal mutations occurring on the L-strand (i.e. G→A on the L-strand vs. C→T mutations on the L-strand). By definition, mutations cause complementary sequence changes on both strands of a DNA molecule. Therefore, if a bias does not exist in the orientation of specific mutations towards a particular strand, then the frequency of reciprocal mutations on the same strand would be expected to be equal. Alternatively, the presence of a strand orientation bias would manifest itself in the form of a particular type of mutation occurring more frequently than its reciprocal mutation.

We find that the majority of the human mitochondrial genome shows a significant strand orientation bias in the occurrence of transitions, whereas transversions show no apparent asymmetry (Fig. 4A). Specifically, in young samples, G→A/C→T mutations are more likely to occur when the dG base is present on the L-strand and the dC base is in the H-strand, respectively. This pattern is even more pronounced in aged individuals, consistent with this bias being due to ongoing mutagenic process and not the result of maternal inheritance. In addition to the G→A/C→T bias, the aged samples also exhibit a strand orientation bias in the occurrence of T→C/A→G, where dT is more likely to be mutated to a dC when it is located on the L-strand than on the H-strand. Interestingly, this bias, which appears uniformly throughout most of the mtDNA, is uniquely absent in the D-loop region (Fig. 4B). Thus, both the spectrum and strand orientation asymmetry of somatic mtDNA mutation accumulation recapitulates what has been previously recognized in population studies.

## Discussion

The accumulation of somatic mutations in mtDNA has frequently been hypothesized to drive the aging process and its associated pathologies, including neurodegeneration, cancer, and atrophy (reviewed in [54]). The underlying mechanisms by which these mutations occur and accumulate have been the subject of intense study, but remain incompletely defined. One of the major limitations has been the lack of methodologies with sufficient sensitivity to detect rare mutations among a much larger population of wild-type molecules. We recently developed a robust next-generation sequencing methodology, termed Duplex Sequencing, which is able to detect a single point mutation among  $>10^7$  sequenced bases [32] and has now enabled us to precisely characterize the genome-wide frequency, spectrum, and distribution of somatic mtDNA mutations in aging human brain with unprecedented accuracy.

Our data show a significant increase in the load of point mutations as a function of human age, with absolute frequencies 10–100 fold lower than what has been typically reported in the literature using less sensitive assays. Recent work using the Random Mutation Capture assay has reported an age associated increase in mtDNA point mutation frequencies in mice and *Drosophila* that are on par with the values that we have determined here; however, these studies were limited to only a very small region of the genome [22] (Leo Pallanck-submitted). Of particular interest, despite a ~1000-fold difference in lifespan, the increase in mutation load with age appears to be highly consistent among multiple species. This surprising finding suggests that the underlying mechanisms behind the age-dependent accumulation of point mutations in mtDNA are conserved between humans, flies, and mice and merit more detailed comparison.



**Figure 4. Transition mutations show a pronounced strand bias in their occurrence.** (A) Comparison of frequencies of reciprocal mutation types on the L-strand for both young (yellow) and aged (purple) sample groups shows that young individuals exhibit a strand bias in the occurrence of G→A/C→T mutations. Aged samples exhibit an even stronger bias in both G→A/C→T and T→C/A→G mutations. No strand bias was observed for transversion mutations. (B) Comparison of frequencies of reciprocal mutation types on the L-strand from aged individuals shows no significant strand bias in the D-loop (bp 16,025–576) region (grey) of mtDNA, while the non-D-loop portion (bp 577–16,024) (orange) of the genome from the same aged samples exhibits a strong strand bias. Frequencies were calculated by dividing the number of mutations of each type by the number times the wild-type base of each mutation type was sequenced. Error bars represent one standard deviation. Significance was tested using the two-tailed t-test. doi:10.1371/journal.pgen.1003794.g004

Oxidative damage to DNA, most notably in the form of 8-oxo-dG, has long been believed to be a primary driver of mutagenesis in both nuclear and mitochondrial DNA [42,55,56]. However, our results do not support this hypothesis. In our data, the relative proportion of G→T/C→A mutations is quite low in the young samples examined and, importantly, does not show a disproportionate increase with age relative to other types of mutations. Other recent reports, which used less sensitive methods to detect intermediate frequency sub-clonal mutations, have similarly failed to detect this classic signature of oxidative damage to DNA. For instance, one conventional deep sequencing analysis of aged mice reported no significant burden of G→T/C→A transversions [57]. Even more surprising is the observation that a transgenic mouse strain deficient for both MutY and OGG1, which are the two primary enzymes responsible for repairing 8-oxo-dG, do not exhibit an increase in mtDNA mutations [58].

Comparison of the spectrum of our reported data (Fig. 2B) to that of the clonal SNV's in our data (i.e. mutations present at >90%), as well as those reported in the Mitomap database [59], reveals an identical bias towards transitions with a minimal number of G→T/C→A transversions (Fig. S2). Indeed, a similar propensity towards transitions has been noted in numerous animal phylogenies [60]. This consistency in mutational pattern suggests that the mutagenic processes that cause the accumulation of mutations in somatic tissue are also responsible for clonal population variants arising in the maternal germline.

In addition to 8-oxo-dG, ROS can also cause a number of other mutagenic lesions, including thymine glycol and deamination of cytidine and adenosine, all of which can induce transition mutations [37,61,62]. Our data clearly show an excess of transitions relative to transversions, which could be consistent with oxidatively induced deamination events becoming fixed as

mutations. It is well established that ROS production and oxidative damage increase with age [17,18,19,20]. Yet, if oxidative damage were the main driver for deamination events, this model would predict that the relative proportion of transitions should be disproportionately higher in aged individuals, which is not the case with our data. While we cannot conclusively rule out a role for ROS in inducing transition mutations, the excess of transitions could additionally be explained by spontaneous deamination of either cytidine or adenosine, especially in single-stranded replication intermediates, or base misincorporation events by DNA polymerase  $\gamma$  during genome replication, which has a known propensity for transition mutations [36,39,40]. Overall, the absence of key mutagenic signatures of oxidative damage argues against ROS being the major driver of mutagenesis in mtDNA in normal aging brain.

Several explanations may account for why we observed few 8-oxo-dG associated mutations in mtDNA despite an extensive literature showing that 8-oxo-dG levels increase with age. First, rapid DNA repair may remove 8-oxo-dG prior to genome replication and this repair capacity may increase with age [63], thereby keeping 8-oxo-dG mutagenesis to a minimum. Secondly, mitochondrial quality control pathways may simply eliminate mitochondria with damaged mtDNA. Consistent with this idea is the observation that oxidatively damaged mtDNA rapidly disappears from cells treated with  $H_2O_2$  [64]. In addition, the cellular levels of parkin, a major component of the quality control pathway involved in mitochondrial turnover, increase under conditions of high oxidative stress [65]. Finally, DNA polymerase  $\gamma$  itself may actively discriminate against incorporating 8-oxo-dG [66]. Regardless of the mechanism, our data suggest that cells have evolved one or more strategies to effectively deal with the challenge of replicating mtDNA in the highly damaging environment of the mitochondria.

We find a striking excess of mutations in the D-loop in young individuals; however, D-loop mutations accumulate during aging at the same rate as other parts of the mitochondrial genome, consistent with the D-loop not being inherently error prone. Instead, our data suggest that young individuals are born with a higher aggregate mutation burden in the D-loop relative to the rest of the genome, thus preserving the D-loop's disproportionate mutation load as mutations accumulate during life. The higher proportion of low-level/random mutations in the D-loop also offers a likely explanation as to why population level SNVs tend to cluster in a number of 'hotspots' in the D-loop.

The disproportionate mutation load of the D-loop at birth provides a likely explanation as to why the D-loop has long been considered prone to mutagenesis. Due to the D-loop's higher aggregate burden, the previously used assays were only sensitive enough to detect mutations located in the D-loop. Thus, even though point mutations increase uniformly throughout mtDNA with age, the mutation load in the non-D-loop portions of the genome would likely be below the sensitivity of the previously available assays, thus giving the appearance that the D-loop's mutations increased with age.

Human population studies have previously identified a bias in the occurrence of G $\rightarrow$ A and T $\rightarrow$ C SNPs of the L-strand, as has comparison of human mtDNA sequences with those of evolutionary related species [50,51,52]. This imbalance has been hypothesized to be the result of preferential deamination of cytidine and adenosine on the H-strand; however, because this strand orientation bias has only been previously observed at the clonal population scale, the influence of natural selection could not be discounted [50,52]. Our data demonstrate that this bias originates from mutagenesis of mtDNA through a process that is continuous

throughout life; however, the source of this bias is unclear. One possibility is that this bias arises during mtDNA replication. Replication of mtDNA is thought to occur in an asymmetric manner [67,68]. In this model, replication of the H-strand (G-rich strand of mtDNA) begins at the H-strand origin, using the L-strand (C-rich strand of mtDNA) as its template. During this time, the parental H-strand remains in a single-stranded state until the synthesis of the L-strand is initiated at the L-strand origin. Several studies previously documenting a strand orientation bias of clonal SNVs at the population level hypothesized that the bias is due to an increased rate of spontaneous hydrolytic deamination of cytidine and adenosine on the H-strand spending a greater amount of time in an unprotected single-stranded state during replication [50,51,52,69]. Indeed, there is *in vitro* evidence that replication past a deaminated dC (i.e. dU) by DNA polymerase  $\gamma$  is responsible for the observed asymmetry. Zheng *et al.* used DNA polymerase  $\gamma$  to amplify small regions of homoduplex mtDNA via PCR and observed the emergence of an excess of G $\rightarrow$ A/C $\rightarrow$ T mutations in these reactions [40]. Their experimental conditions involved extended heating of the DNA, which is known to increase the rate of deamination of single-stranded DNA. While our results do not, on their own, provide direct proof that an asymmetry in deamination of single-stranded replication intermediates is responsible for the observed inter-strand mutational skew, our data are consistent the findings and conclusions in Zheng *et al.*, thus providing strong support for this model [40]. Furthermore, the observation that a similar strand bias exists in the absence of intergenerational selection argues that it is due to an ongoing mutagenic process within the cell and that the bias observed at the population level is likely due to the same process.

In summary, Duplex Sequencing is a powerful technology for high-sensitivity study of mtDNA mutagenesis that has enabled us to uncover a number of mutagenic patterns and biases in human brain tissue that have not been previously observed at the somatic level. Taken together, these mutation patterns argue against oxidative DNA damage being a major driver of aging and suggest that replication errors by DNA polymerase  $\gamma$  and/or spontaneous base hydrolysis are responsible for the bulk of point mutations that accumulate in mtDNA. In light of this finding, the central role for oxidative damage to mtDNA in theories of human aging and disease merits re-evaluation.

## Materials and Methods

### Tissues and DNA isolation

Mitochondria were isolated from the pre-frontal cortex of either young (<1 yr) or old (>75 yr) brain tissue with no known brain pathologies (Table S1). Approximately 500–800 mg of brain tissue was sub-divided into smaller 100–150 mg pieces, with the mitochondria from each piece purified independently from the other pieces of the same tissue sample. Each tissue fraction was dounced in 5 mL of homogenization buffer (0.32 M sucrose, 1 mM EDTA, 10 mM Tris-HCl, pH 7.8) with no more than five strokes, as excessive douncing increases the contamination of nuclear DNA. The homogenate was centrifuged at 1000 g for 20 min at 4°C. The supernatant was transferred to a new centrifuge tube and centrifuged at 1000 g for 10 min at 4°C. Samples that exhibited a pellet at the bottom of the tube after the second centrifugation step were centrifuged a third time using the same condition. The supernatant was removed and centrifuged at 13,000 g for 30 min at 4°C to give a crude mitochondrial pellet. Each crude pellet was resuspended in Mito-DNase buffer (0.3 M sucrose, 10 mM MgCl<sub>2</sub>, 0.15% BSA (w/v), 20 mM Tris-HCl, pH 7.5) and DNase added to a concentration of 0.01 mg/mL and

incubated at 37°C for 1.5 hrs. This step helps remove contaminating nuclear DNA while not affecting mtDNA protected within intact mitochondria. Mitochondria were then re-pelleted at 13,000 g for 30 min and the supernatant discarded. The mitochondrial pellets were washed two times by resuspending the pellets in 1 mL of Mito-DNase buffer followed by centrifugation at 13,000 g for 30 min. Finally, the enriched mitochondrial pellets were resuspended in 500  $\mu$ L Lysis Buffer (150 mM NaCl, 20 mM EDTA, 1% SDS (w/v), 10 mM Tris-HCl, pH 7.8, 0.2 mg/mL Proteinase K, 0.01 mg/mL RNase) and incubated at 56°C for 1 hr. The mtDNA was extracted using a standard phenol:chloroform approach. The relative purity of each mtDNA prep from nuclear DNA was determined by qPCR using the following primer sets: nDNA: *fvd* 5'- TTGCCAGACCATGGGATTGTCTCA, *rev* 5'-TTCCTACCGAACGAGGACTC-CAAA; mtDNA: *fvd* 5'-ACAGTTTATGTAGCTTACCTC, *rev* 5'-TTGCTGCGTGCTTGATGCTTG. DNA fractions from the same tissue samples exhibiting a  $\Delta C(t) \geq 17.5$  cycles (corresponding to ~50% mtDNA by mass) were pooled and used for library preparation and sequencing. Typical yields were between 100–500 ng of highly pure mtDNA.

### Sequencing

We carried out synthesis and preparation of the Duplex Tag-labeled adapters and sequencing library preparation as previously described with the following minor mtDNA specific modifications [32]: 1) 100–500 ng of mtDNA was sheared using the Covaris AFA system with a duty cycle of 10%, intensity of 5, cycles/burst 100, time 20 seconds $\times$ 5, temperature of 4°C; 2) prior to adapter ligation, the DNA was quantified and a 40:1 molar excess of adapters was used for the ligation step; 3) after adapter ligation and clean up, the library was re-quantified and ~37 fmol of product was amplified by PCR for a total of 20 cycles. The resulting libraries were subjected to sequencing on an Illumina HiSeq 2000/2500 platform using 101 bp, paired-end reads.

**Data analysis.** The overall sequencing pipeline has been previously described [32]. Briefly, all raw sequencing reads are filtered based on the location of an expected fixed sequence within each read, as well as the fact that each position of the 12 nucleotide random Duplex Tag sequence can only contain one of the four canonical bases. Any reads not conforming to these criteria are discarded. The random tag sequences from each read pair are computationally concatenated into a 24-nt random sequence and appended to the header of each read-pair. In order to remove potential artifactual mutations arising from the end-repair and ligation reactions, each read was trimmed by four additional bases. The reads were then aligned against build 19 of the human genome using the Burrows-Wheeler Aligner (BWA) [70]. Reads not aligning to the human mitochondrial genome were filtered out. A consensus sequence for each tag family (e.g. reads sharing identical tag sequences) was computationally determined using software written in-house. The consensus for any position in a read was considered undefined if the position was represented by fewer than three instances in the family or if less than 70% of the sequences at that position in the read were in agreement. The SSCS reads were then realigned to the revised mitochondrial Cambridge Reference Sequence using BWA. After filtering for unmapped reads, DCS reads were constructed by pairing SSCS with their respective strand-mates by grouping the 24 nucleotide tag in read 1 with the appropriate 24 nucleotide tag in read 2 and the sequence identity at each position of the two reads were compared to one another. The sequence information was kept only if the base identity of both reads were identical. Next, to remove alignment artifacts common at the ends of reads, five

nucleotides from each end of the duplex reads were soft-clipped using the Genome Analysis Toolkit. Finally, all reads that mapped to the same genomic coordinates and contained tags that differed in their sequence by less than three nucleotides were removed from the data set. Reads found to harbor a mutation were then realigned against the entire human genome using BLAST. Reads that had a lower e-score for a region of the nuclear genome than the mtDNA were removed in order to eliminate any potential contribution from nuclear mitochondrial pseudogenes. Notably, on post-analysis review, no correlation existed between sample mutation frequency and degree of nuclear genome contamination, which would be expected if mutation data were confounded by contaminating clonal mutations in nuclear pseudogenes (Fig. S3). Scripts to calculate mutation frequencies and locations were written in-house and are provided as supplemental files. Due to the high redundancy of the sequencing data in Duplex Sequencing, relatively few “genome equivalents” are sequenced (~5,000 “genomes”/sample), thus the ability to detect sub-clonal, multi-kilobase, deletions and rearrangements is limited and was, therefore, not evaluated in this study.

### Ethics statement

This study involved the use of human brain tissue obtained from autopsy. The tissues were collected at University of Washington under the direction of the Alzheimer’s Research Center and Seattle Children’s hospital under their approved IRB protocols. Our research did not fall under human subjects requirements due to the samples being anonymous.

### Supporting Information

**Figure S1** Non-synonymous mutations are predicted to be more deleterious with age. All non-synonymous, non-clonal (i.e. occurring at <1% clonality) mutations were scored for pathogenicity using MutPred. Young (*gray*) samples had a significantly lower pathogenicity score than the aged samples (*light blue*) ( $p < 0.02$ , Wilcoxon Rank Sum analysis). (TIF)

**Figure S2** Population level SNPs, homoplasmic variants, and non-clonal mutations all show a similar mutation spectra. Mitomap.org data (*light blue*) were downloaded from <http://www.mitomap.org/bin/view.pl/MITOMAP/> PolymorphismsCoding and <http://www.mitomap.org/bin/view.pl/MITOMAP/PolymorphismsControl>. Insertion/deletion mutations were ignored and only point mutations were tabulated. Homoplasmic variants from our data set (*dark blue*) are defined as variants occurring in >90% of mapped reads. We detected a total of 134 unique homoplasmic variants among all young and aged samples. Data for the Non-clonal variants (*green*) are derived from the aged sample data in Figure 2B. (TIF)

**Figure S3** Nuclear DNA contamination does not affect observed mtDNA mutation frequency. Analysis was performed on reads only mapping to either the nuclear or mitochondrial genome. Percentage of mtDNA was calculated by dividing the number of reads mapping to the mtDNA by the total number of mapped reads. Linear regression is non-significant ( $R^2 = 0.15$ ). (TIF)

**Table S1** Sample information and sequencing statistics. (DOCX)

**Text S1** Scripts used for data processing and analysis. (ZIP)



## Acknowledgments

We would like to acknowledge and thank both Raj J. Kapur (Seattle Children's Hospital) and the University of Washington Alzheimer's Disease Research Center for the tissue samples used in these experiments. We would also like to thank Alan J. Herr for critical reading of the manuscript and his work on Figure 1. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## References

- Harman D (1956) Aging: A theory based on free radical and radiation chemistry. *J Gerontol* 11: 298–300.
- Harman D (1972) The biological clock: The mitochondria? *J Am Geriatr Soc* 20: 145–147.
- Wallace DC (1999) Mitochondrial diseases in man and mouse. *Science* 283: 1482–1488.
- Taylor RW, Turnbull DM (2005) Mitochondrial DNA mutations in human disease. *Nat Rev Genet* 6: 389–402.
- Short KR, Bigelow ML, Kahl J, Singh R, Coenen-Schimke J, et al. (2005) Decline in skeletal muscle mitochondrial function with aging in humans. *Proc Nat Acad Sci U S A* 102: 5618–5623.
- Greaves LC, Barron MJ, Campbell-Shiel G, Kirkwood TBL, Turnbull DM (2011) Differences in the accumulation of mitochondrial defects with age in mice and humans. *Mech Ageing Dev* 132: 588–591.
- Cao Z, Wanagat J, McKiernan SH, Aiken JM (2001) Mitochondrial DNA deletion mutations are concomitant with ragged red regions of individual, aged muscle fibers: analysis by laser-capture microdissection. *Nuc Acid Res* 29: 4502–4508.
- Hsieh R, Hou J, Hsu H, Wei W (1994) Age-dependent respiratory function decline and DNA deletions in human muscle mitochondria. *Biochem Mol Biol Int* 32: 1009–1022.
- Kraytsberg Y, Kudryavtseva E, McKee AC, Geula C, Kowall NW, et al. (2006) Mitochondrial DNA deletions are abundant and cause functional impairment in aged human substantia nigra neurons. *Nat Genetics* 38: 518–520.
- Lezza A, Boffoli D, Cantatore P, Gadaleta M (1994) Correlation between mitochondrial DNA 4977-bp deletion and respiratory chain enzyme activities in aging human skeletal muscles. *Biochem Bioph Res Co* 205: 772–779.
- Corral-Debrinski M, Horton T, Lott MT, Shoffner JM, Beal MF, et al. (1992) Mitochondrial DNA deletions in human brain: Regional variability and increase with advanced age. *Nat Genetics* 2: 324–329.
- Mohamed SA, Hanke T, Erasmi AW, Bechtel MJF, Scharfshwerdt M, et al. (2006) Mitochondrial DNA deletions and the aging heart. *Exp Gerontol* 41: 508–517.
- Bender A, Krishnan KJ, Morris CM, Taylor GA, Reeve AK, et al. (2006) High levels of mitochondrial DNA deletions in substantia nigra neurons in aging and Parkinson disease. *Nat Genetics* 38: 515–517.
- Gu G, Reyes PF, Golden GT, Woltjer RL, Hulette C, et al. (2002) Mitochondrial DNA deletions/rearrangements in parkinson disease and related neurodegenerative disorders. *J Neuropath Exp Neurol* 61: 634–639.
- Corral-Debrinski M, Horton T, Lott MT, Shoffner JM, McKee AC, et al. (1994) Marked changes in mitochondrial DNA deletion levels in Alzheimer brains. *Genomics* 23: 471–476.
- Kasai H, Nishimura S (1984) Hydroxylation of deoxyguanosine at the C-8 position by ascorbic acid and other reducing agents. *Nuc Acid Res* 12: 2137–2145.
- Fraga CG, Shigenaga MK, Park J-W, Degan P, Ames BN (1990) Oxidative damage to DNA during aging: 8-Hydroxy-2'-deoxyguanosine in rat organ DNA and urine. *Proc Nat Acad Sci U S A* 87: 4533–4537.
- Mecocci P, MacGarvey U, Kaufman AE, Koontz D, Shoffner JM, et al. (1993) Oxidative damage to mitochondrial DNA shows marked age-dependent increases in human brain. *Ann Neurol* 34: 609–616.
- Asuncion JGdl, Millan A, Pla R, Bruseghini L, Esteras A, et al. (1996) Mitochondrial glutathione oxidation correlates with age-associated oxidative damage to mitochondrial DNA. *FASEB J* 10: 333–338.
- Barja G, Herrero A (2000) Oxidative damage to mitochondrial DNA is inversely related to maximum life span in the heart and brain of mammals. *FASEB J* 14: 312–318.
- 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, Bielas JH, Kujoth GC, Ladiges WC, Rabinovitch PS, et al. (2007) Mitochondrial point mutations do not limit the natural lifespan of mice. *Nat Genetics* 39: 540–543.
- Lin MT, Cantuti-Castelvetri I, Zheng K, Jackson KE, Tan YB, et al. (2012) Somatic mitochondrial DNA mutations in early Parkinson and incidental Lewy Body Disease. *Ann Neurol* 71: 850–854.
- Lin MT, Simon DK, Ahn CH, Kim LM, Beal MF (2002) High aggregate burden of somatic mtDNA point mutations in aging and Alzheimer's disease brain. *Hum Mol Genet* 11: 133–145.

## Author Contributions

Conceived and designed the experiments: SRK LAL. Performed the experiments: SRK. Analyzed the data: SRK JJS MWS LAL. Contributed reagents/materials/analysis tools: SRK MWS. Wrote the paper: SRK JJS MWS LAL.

- Nekhaeva E, Bodyak ND, Kraytsberg Y, McGrath SB, Orsouw NJV, et al. (2002) Clonally expanded mtDNA point mutations are abundant in individual cells of human tissues. *Proc Nat Acad Sci U S A* 99: 5521–5526.
- Michikawa Y, Mazzucchelli F, Bresolin N, Scarlato G, Attardi G (1999) Mutations in the human mtDNA control region for replication. *Science* 286: 774–779.
- Loeb LA (2011) Human cancers express mutator phenotypes: Origin, consequences and targeting. *Nat Rev Cancer* 11: 450–457.
- Salk JJ, Fox EJ, Loeb LA (2010) Mutational heterogeneity in human cancers: Origin and consequences. *Annu Rev Pathol-Mech* 5: 51–75.
- Vermulst M, Bielas JH, Loeb LA (2008) Quantification of random mutations in the mitochondrial genome. *Methods* 46: 263–268.
- Marcelino LA, Thilly WG (1999) Mitochondrial mutagenesis in human cells and tissues. *Mutat Res* 434: 177–203.
- Harismendy O, Ng PC, Strausberg RL, Wang X, Stockwell TB, et al. (2009) Evaluation of next generation sequencing platforms for population targeted sequencing studies. *Genome Biol* 10: R32.
- Schmitt MW, Kennedy SR, Salk JJ, Fox EJ, Hiatt JB, et al. (2012) Detection of ultra-rare mutations by next-generation sequencing. *Proc Nat Acad Sci U S A* 109: 14508–14513.
- Marinho ANdR, Moraes MRd, Santos S, Ribeiro-dos-Santos Á (2011) Human aging and somatic point mutations in mtDNA: A comparative study of generational differences (grandparents and grandchildren). *Genet Mol Biol* 34: 31–34.
- 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.
- Duncan BK, Miller JH (1980) Mutagenic deamination of cytosine residues in DNA. *Nature* 287: 560–561.
- Spelbrink JN, Toivonen JM, Hakkaart GAJ, Kurkela JM, Cooper HM, 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.
- Nordmann PL, Makris JC, Reznikoff WS (1988) Inosine induced mutations. *Mol Gen Genet* 214: 62–67.
- Song S, Pursell ZF, Copeland WC, Longley MJ, Kunkel TA, et al. (2005) DNA precursor asymmetries in mammalian tissue mitochondria and possible contribution to mutagenesis through reduced replication fidelity. *Proc Nat Acad Sci U S A* 102: 4990–4995.
- Longley MJ, Nguyen D, Kunkel TA, Copeland WC (2001) The fidelity of human DNA polymerase  $\gamma$  with and without exonucleolytic proofreading and the p55 accessory subunit. *J Biol Chem* 276: 38555–38562.
- Zheng W, Khrapko K, Collier HA, Thilly WG, Copeland WC (2006) Origins of human mitochondrial point mutations as DNA polymerase  $\gamma$ -mediated errors. *Mutat Res-Fundamental and Molecular Mechanisms of Mutagenesis* 599: 11–20.
- Mecocci P, MacGarvey U, Beal MF (1994) Oxidative damage to mitochondrial DNA is increased in Alzheimer's disease. *Ann Neurol* 36: 747–751.
- Cheng KC, Cahill DS, Kasais H, Nishimura S, Loeb LA (1992) 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G→T and A→C substitutions. *J Biol Chem* 267: 166–172.
- Beckman KB, Ames BN (1998) The free radical theory of aging matures. *Physiol Rev* 78: 547–581.
- Kirkwood TBL, Kowald A (2012) The free-radical theory of ageing-older, wiser and still alive. *Bioessays* 34: 692–700.
- Li B, Krishnan VG, Mort ME, Xin F, Kamati KK, et al. (2009) Automated inference of molecular mechanisms of disease from amino acid substitutions. *Bioinformatics* 25: 2744–2750.
- Greaves LC, Elson JL, Nootboom M, Grady JP, Taylor GA, et al. (2012) Comparison of mitochondrial mutation spectra in ageing human colonic epithelium and disease: Absence of evidence for purifying selection in somatic mitochondrial DNA point mutations. *Plos Genet* 8: e10003082.
- Galtier N, Enard D, Radondy Y, Bazin E, Belkhir K (2006) Mutation hot spots in mammalian mitochondrial DNA. *Genome Res* 16: 215–222.
- Meyer S, Weiss G, Haeseler Av (1999) Pattern of nucleotide substitution and rate heterogeneity in the hypervariable regions I and II of human mtDNA. *Genetics* 152: 1103–1110.
- Wakeley J (1993) Substitution rate variation among sites in hypervariable region 1 of human mitochondrial DNA. *J Mol Evol* 37: 613–623.

50. Reyes A, Gissi C, Pesole G, Saccone C (1998) Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. *Mol Biol Evol* 15: 957–966.
51. Belle EMS, Piganeau G, Gardner M, Eyre-Walker A (2005) An investigation of the variation in the transition bias among various animal mitochondrial DNA. *Gene* 355: 58–66.
52. Tanaka M, Ozawa T (1994) Strand asymmetry in human mitochondrial DNA mutations. *Genomics* 22: 327–335.
53. Xia X (2012) DNA replication and strand asymmetry in prokaryotic and mitochondrial genomes. *Curr Genomics* 13: 16–27.
54. Kennedy SR, Loeb LA, Herr AJ (2012) Somatic mutations in aging, cancer and neurodegeneration. *Mech Ageing Dev* 133: 118–126.
55. Cooke MS, Evans MD, Dizdaroglu M, Lunec J (2003) Oxidative DNA damage: mechanisms, mutation, and disease. *FaSEB J* 17: 1195–1214.
56. Marnett LJ (2000) Oxyradicals and DNA damage. *Carcinogenesis* 21: 361–370.
57. Ameer 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: e1002028.
58. Halsne R, Esbensen Y, Wang W, Scheffler K, Suganthan R, et al. (2012) Lack of the DNA glycosylases MYH and OGG1 in the cancer prone double mutant mouse does not increase mitochondrial DNA mutagenesis. *DNA Repair* 11: 278–285.
59. Ruiz-Pesini E, Lott MT, Procaccio V, Poole JC, Brandon MC, et al. (2007) An enhanced MITOMAP with a global mtDNA mutational phylogeny. *Nuc Acids Res* 35: D823–D828.
60. Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, et al. (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc Nat Acad Sci U S A* 86: 6196–6200.
61. Basu AK, Loechler EL, Leadon SA, Essigmann JM (1989) Genetic effects of thymine glycol: Site-specific mutagenesis and molecular modeling studies. *Proc Nat Acad Sci U S A* 86: 7677–7681.
62. Kreuzer DA, Essigmann JM (1998) Oxidized, deaminated cytosines are a source of C→T transitions *in vivo*. *Proc Nat Acad Sci U S A* 95: 3578–3582.
63. Souz-Pinto NC, Croteau DL, Hudson EK, Hansford RG, Bohr VA (1999) Age-associated increase in 8-oxo-deoxyguanosine glycosylase/AP lyase activity in rat mitochondria. *Nuc Acid Res* 27: 1935–1942.
64. Furda AM, Marrangoni AM, Lokshin A, Houten BV (2012) Oxidants and not alkylating agents induce rapid mtDNA loss and mitochondrial dysfunction. *DNA Repair* 11: 684–692.
65. Yang YX, Muqit MMK, Latchman DS (2006) Induction of parkin expression in the presence of oxidative stress. *Eur J Neurosci* 24: 1366–1372.
66. Hanes JW, Thal DM, Johnson KA (2006) Incorporation and replication of 8-oxo-deoxyguanosine by the human mitochondrial DNA polymerase. *J Biol Chem* 281: 36241–36248.
67. Brown TA, Ceccconi C, Tkachuk AN, Bustamante C, 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.
68. Clayton DA (1982) Replication of animal mitochondrial DNA. *Cell* 28: 693–705.
69. Frederico LA, Kunkel TA, Shaw BR (1990) A sensitive genetic assay for the detection of cytosine deamination: Determination of rate constants and the activation energy. *Biochemistry* 29: 2532–2537.
70. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25: 1754–1760.