OPEN ACCESS Freely available online

PLOS | GENETICS

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

Scott R. Kennedy[1], Jesse J. Salk[1,2], Michael W. Schmitt[1,2], Lawrence A. Loeb[1,3]*

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 ~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 γ or deamination of cytidine and adenosine as the primary mutagenic events in mtDNA. Surprisingly, G→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→A and T→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 γ 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 PO1-CA77852, and NCI RO1-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-

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

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 6×10^{-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 ~5-fold increase in mutational frequency, relative to those obtained from young individuals (Young: $3.7\pm0.9\times10^{-6}$ vs. Aged: $1.9\pm0.2\times10^{-5}$, $p<10^{-4}$, 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].

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

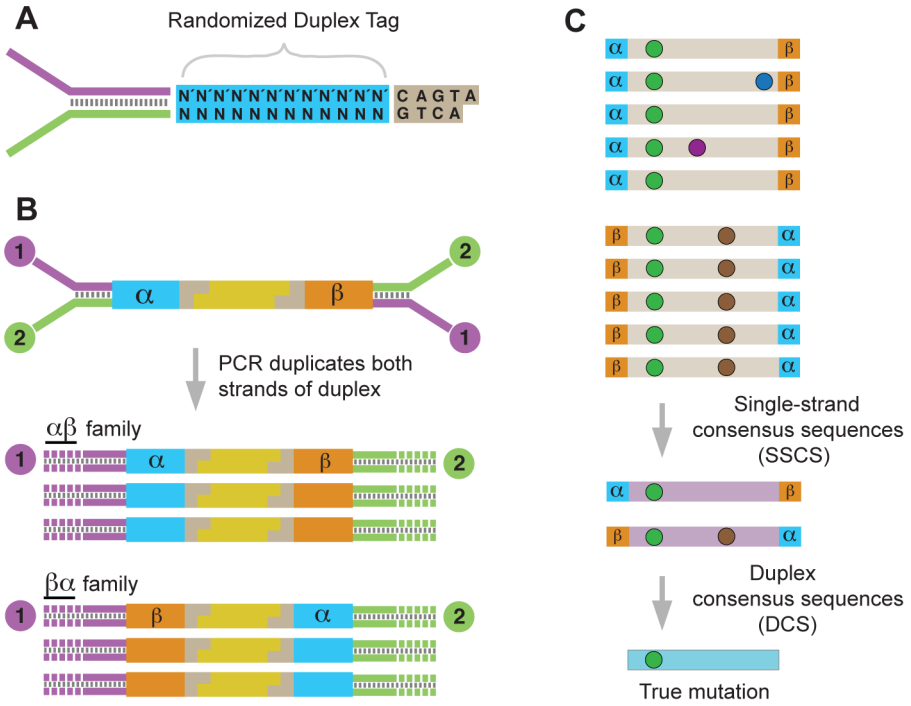

Figure 1. Overview of the Duplex Sequencing methodology. (A) Adapter design with random double-stranded tag sequence and invariant spacer sequence. (B) Ligation of adapters to fragmented DNA generates unique 12 bp tags on each end (α and β). PCR amplification of the two strands produces two related, but distinct products. (C) Sequence reads sharing unique α and β tags are grouped into families of α-β or β-α 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→A/C→T, is consistent with either misincorporation by DNA polymerase γ or deamination of cytosine to form uracil, as being the largest mutagenic drivers in mtDNA [35,36]. The second most common mutation type, T→C/A→G, is consistent with either deamination of adenosine to inosine or a T-dGTP mispairing, the primary base misinsertion mistake made by DNA polymerase γ [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→T/C→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→T/C→A mutations would be expected to increase with time. We failed to find either a preponderance of G→T/C→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

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

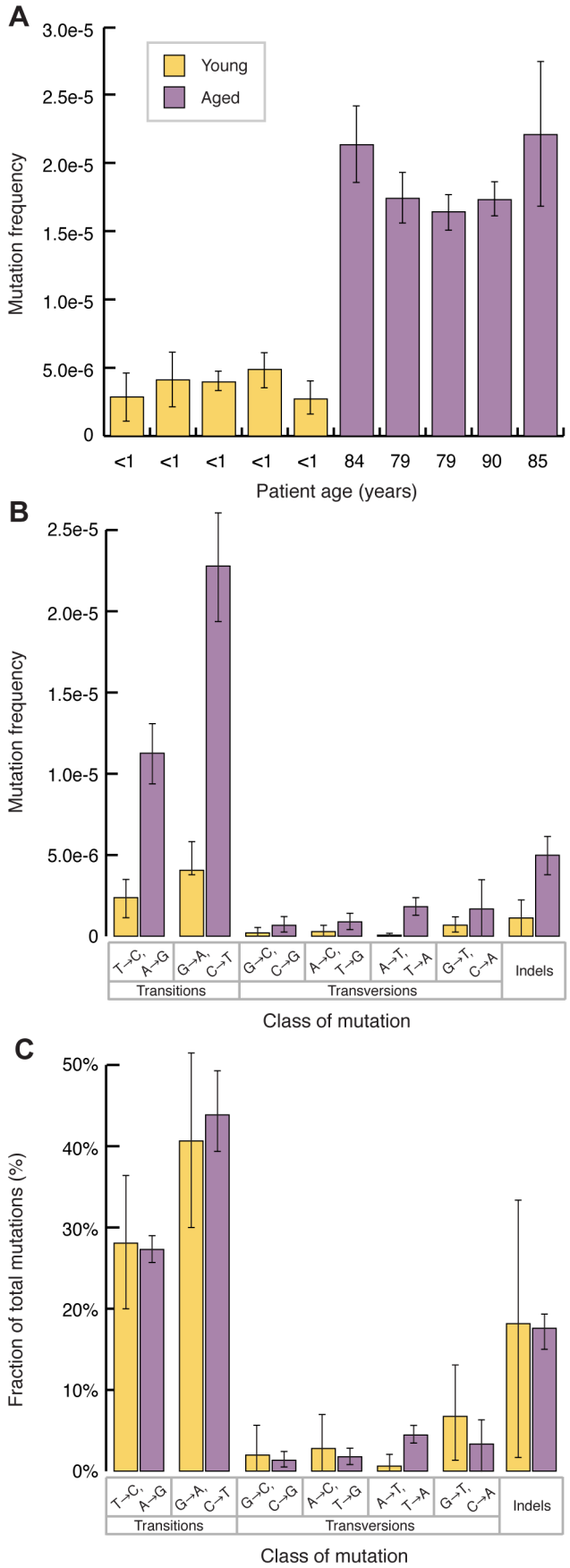

Figure 2. Mitochondrial point mutations increase with age and are biased to transitions. (A) mtDNA point mutations burden is higher in older individuals (*purple*) than young individuals (*yellow*) ($p<10^{-4}$, two-tailed t-test). Error bars represent the 95% confidence interval for each sample (Wilson 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: $1.5\pm0.6\times10^{-5}$ vs. coding region: $2.9\pm0.7\times10^{-6}$, $p<0.01$, two-tailed t-test) and aged (D-loop: $5.7\pm1.5\times10^{-5}$ vs. coding region: $1.65\pm0.2\times10^{-5}$, $p<0.01$, 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 ± 1.6-fold increase) to the relative increase seen between the two sample groups in the non-D-loop regions (5.6 ± 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-

DOCKET ALARM
Find authenticated court documents without watermarks at docketalarm.com.

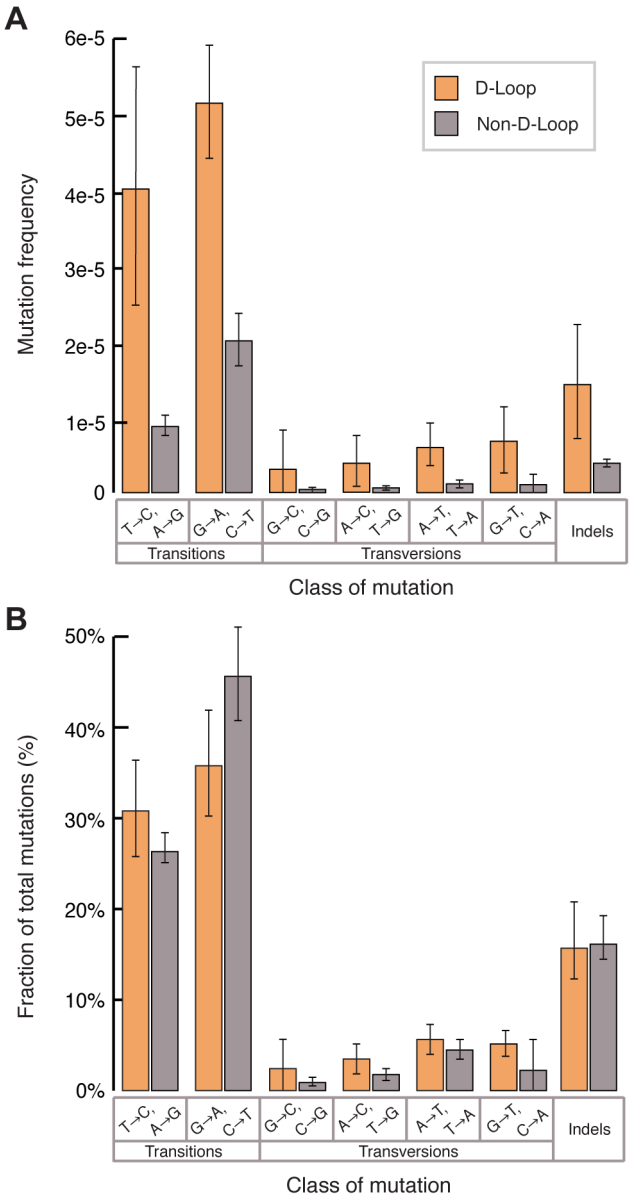

Figure 3. The D-loop has an elevated mutation burden but its mutation spectrum is similar to the remainder of the mitochondrial genome. (A) The D-loop (*orange*) exhibits a higher aggregate mutation burden than the rest of the genome (*grey*). Frequencies were calculated by dividing the number of mutations 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 subclonal/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.

DOCKET ALARM
Find authenticated court documents without watermarks at docketalarm.com.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.