PNAS RESEARCH ARTICLE | GENETICS OPEN ACCESS

High-frequency and functional mitochondrial DNA mutations at the single-cell level

Xiaoxian Guo[a,b,1], Weilin Xu[a,1], Weiping Zhang[a], Cuiping Pan[c], Anna E. Thalacker-Mercer[a], Hongxiang Zheng[d], and Zhenglong Gu[a,c,d,2]

Edited by Douglas Wallace, University of Pennsylvania, Philadelphia, PA; received January 27, 2022; accepted November 10, 2022

Decline in mitochondrial function underlies aging and age-related diseases, but the role of mitochondrial DNA (mtDNA) mutations in these processes remains elusive. To investigate patterns of mtDNA mutations, it is particularly important to quantify mtDNA mutations and their associated pathogenic effects at the single-cell level. However, existing single-cell mtDNA sequencing approaches remain inefficient due to high cost and low mtDNA on-target rates. In this study, we developed a cost-effective mtDNA targeted-sequencing protocol called single-cell sequencing by targeted amplification of multiplex probes (scSTAMP) and experimentally validated its reliability. We then applied our method to assess single-cell mtDNA mutations in 768 B lymphocytes and 768 monocytes from a 76-y-old female. Across 632 B lymphocyte and 617 monocytes with medium mtDNA coverage over >100×, our results indicated that over 50% of cells carried at least one mtDNA mutation with variant allele frequencies (VAFs) over 20%, and that cells carried an average of 0.658 and 0.712 such mutation for B lymphocytes and monocytes, respectively. Surprisingly, more than 20% of the observed mutations had VAFs of over 90% in either cell population. In addition, over 60% of the mutations were in protein-coding genes, of which over 70% were nonsynonymous, and more than 50% of the nonsynonymous mutations were predicted to be highly pathogenic. Interestingly, about 80% of the observed mutations were singletons in the respective cell populations. Our results revealed mtDNA mutations with functional significance might be prevalent at advanced age, calling further investigation on age-related mtDNA mutation dynamics at the single-cell level.

aging | single cell | mitochondrial DNA

Significance

Mitochondrial oxidative phosphorylation (OXPHOS) system is encoded partially by mitochondrial DNA (mtDNA). Decline of mitochondrial function may underlie the pathogenesis of many age-related diseases, but the mechanisms underlying mitochondrial functional decline with advancing age remain controversial. Our results in this study showed the prevalence of high-frequency pathogenic mutations at the single-cell level. Given the high percentage of cells harboring these mutations, our results demonstrate the diversity of single-cell mtDNA mutations that were previously underappreciated and highlight the importance of investigating single-cell mtDNA mutations and their roles in human aging.

Mitochondria are double-membraned organelles in cytoplasm of almost all eukaryotic cells. They play important roles in energy production, immune response, epigenetic regulation, apoptosis, and many other metabolic and signaling processes (1–5). Due to their symbiotic origin during evolution, mitochondria are the only organelle in animal cells that carry their own genomes (6). In humans, the mitochondrial DNA (mtDNA) is an approximately 16.6 kb double-stranded molecule that encodes for 13 proteins of the oxidative phosphorylation (OXPHOS) complexes, in addition to 22 tRNAs and 2 rRNAs. Each human cell contains many copies of mtDNA, and mutations can affect various proportions of the mtDNA molecules (6).

Mutations in mtDNA occur at a much higher rate than in nuclear DNA (7). Because approximately 95% of the entire mtDNA genome are coding regions, mtDNA mutations could lead to mitochondrial functional decline, alterations in cellular activities and even diseases (8–11). In addition to well-known mitochondrial diseases such as mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes (MELAS), myoclonic epilepsy with ragged-red fibers (MERRF), or neurogenic weakness with ataxia and retinitis pigmentosa (NARP), mtDNA mutations might also contribute to age-related complex disorders, such as type-2 diabetes, neurodegenerative diseases, and multiple cancers (8–10, –14).

The majority of reported mtDNA mutations implicated in human diseases are heteroplasmic, meaning that mutant and wild-type mtDNAs coexist in a cell or tissue (15). While research on heteroplasmy has recently expanded (16–30), more effort is needed to investigate the pattern of mtDNA heteroplasmy at the single-cell level (31–36). Theoretical simulations demonstrated that different mtDNA mutations can drift to high frequencies in cell subpopulations as a result of cell division or internal mitochondrial turnover over a person's lifetime (37–39). Sequencing a population of cells with heterogeneous mtDNA mutations can potentially mask distinct mtDNA mutations in individual cells. It is therefore essential to quantify mtDNA mutations and their associated pathogenic effects at the single-cell level to understand the real distribution of mtDNA mutations.

Multiple sequencing protocols have been developed and utilized for assessment of human mtDNA mutations (21–23, 26, 27, 404142–43). However, the majority of these protocols are limited to cell population-based samples. In recent years, single-cell approaches to

Author contributions: Z.G. designed research; X.G., W.X., and Z.G. performed research; A.E.T.-M. contributed new reagents/analytic tools; X.G., W.X., W.Z., H.Z., and Z.G. analyzed data; and X.G., W.X., W.Z., C.P., A.E.T.-M., H.Z., and Z.G. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]X.G. and W.X. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: zg27@cornell.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2201518120/-/DCSupplemental.

Published December 28, 2022.

sequencing mtDNA have emerged to uncover the rich diversity of mtDNA and resolve clonal dynamics (34–, –46). Single cells assessment of mtDNA mutations was also performed with scRNA-seq data to provide insights into the clonal properties in various model systems (47, 48). Several protocols based on single-cell assay for transposase-accessible chromatin with sequencing (scATAC-seq) have been developed to quantify mtDNA mutations in single cells and have become increasingly popular (35, 36, 44). However, the intrinsic drawback of scATAC-seq associated with allelic dropout has been widely ignored in determining the mutational profile of mitochondrial genome (49). Meanwhile, it was estimated that only about 20% of the total scATAC-seq reads from an individual cell were mtDNA fragments (44), resulting in high cost of the protocol, insufficient depth of mtDNA coverage, and subsequently unreliable mutation calling. Rolling circle amplification (RCA) preferentially amplifies circular mtDNA over linear nuclear genome, and it has been adapted by scMito-seq and MitoSV-seq to generate enough replicates of mtDNAs from single-cell lysate. In conjugate with tagmentation-based library preparation, these techniques have enabled the characterization of mtDNA genomes of single cells. Although the preamplification step using RCA can mitigate allelic dropout in comparison with direct tagmentation of initial mtDNA, the amplification products are inevitably contaminated with nuclear genome and eventually lead to low mtDNA coverage. While scRNA-seq data can help identify mtDNA mutations for lineage tracing, it cannot reveal overall single-cell mtDNA mutation load. This raises the need for cost-effective and targeted-sequencing approaches for single-cell mtDNA.

We recently developed a novel, cost-effective human mtDNA sequencing method called sequencing by targeted amplification of multiplex probes (STAMP) for assessing mtDNA mutations in cell population samples (25). We demonstrated that STAMP has both high on-target rates and high sensitivity in detecting single-nucleotide mutations (25, 30, 50). Due to its flexibility, STAMP can be readily adapted to single-cell mtDNA sequencing. In this study, we established a novel method, single-cell STAMP (scSTAMP), for cost-effective and sensitive assessment of mtDNA mutations in single cells. The method was applied to 768 B lymphocytes and 768 monocytes for a human adult aged at 76. We systematically evaluated the patterns and dynamics of mtDNA mutations at the single-cell level and demonstrated prevalence of unique, high-frequency, and functional mutations at the single-cell level. Our results indicate that mtDNA mutations, varying among different cells in the same person, could potentially be an important source of mitochondrial functional decline at the tissue and/or organism level with advancing age.

Results

Accurate Detection of mtDNA Mutations at the Single-Cell Level. STAMP is a cost-effective targeted-sequencing approach developed for assessing mtDNA contents and mutations in bulk cell population samples (25). STAMP uses 46 pairs of single-stranded oligonucleotide probes for capturing the entire human mtDNA genome with an extension-ligation (EL) reaction, the product of which are then subjected to PCR amplification, library purification, and massively parallel sequencing. To extend STAMP for assessing mitochondrial genomes in individual cells, we incorporated a PCR amplification step before the EL-capture reaction to generate sufficient replicates of mtDNA molecules for target capturing. Specifically, the whole mtDNA genomes were amplified as two overlapping fragments directly from single-cell lysates in a single reaction. The resulting amplicons were processed by STAMP to generate sequencing libraries (Fig. 1*A*).

Given the multiplicity of mtDNA molecules in single cells and most mtDNA mutations are heteroplasmic, we conducted a validation experiment to assess the technical variation of scSTAMP. We isolated, from a donor of ~60-y-old, 24 B lymphocytes and 24 monocytes representing the lymphoid lineage and the myeloid lineage, respectively (51, 52). We split each cell lysate evenly into two aliquots and independently applied scSTAMP to both. To stringently control for false positive mutation calls, we decided to focus on single-cell mtDNA mutations with over 20% of variant allele frequency (VAF). Mathematically, it has been shown that high-VAF mutations were unlikely to originate from PCR errors (34). Biologically, it was recently reported that transcriptional reprogramming starts to occur for some mutations with VAFs as low as 20%, and a defect in OXPHOS is usually detectable after the corresponding mutations reach high-frequency levels in cells (53–56). We additionally required that both aliquots derived from the same cells have over 100× median depth of mtDNA coverage, which resulted in 22 B lymphocytes and 24 monocytes with median coverage ranging from 180.0× to 4489.0×. Across 46 probe regions, the distributions of high-quality consensus reads were highly correlated between two aliquots of the same cells. In total, we identified 27 mutations that were shared by both aliquots of the same cells with both VAFs over 20% (Fig. 1*B*). The correlation between VAFs of shared mutations in both aliquots of the same cells was 0.88 (p-value = 1.407×10^{-9}). Furthermore, 21 out of the 27 (77.78%) mutations that appeared in both aliquots with VAFs over 20% had an absolute VAF difference within 10% (Fig. 1*B*).

We further investigated mutations that appeared in only one aliquot, which totaled nine mutations. Among these, five mutations were found in both aliquots but were filtered out in the further analysis of one of the two aliquots due to stringent quality control. The remaining four mutations were present in only one aliquot (*SI Appendix*, Table S1). The uneven split of mitochondrial contents during library preparation and/or biased amplification of certain mtDNA molecules could potentially contribute to the differences in VAFs of mtDNA mutations in two aliquots of the same cells. Indeed, three out of the four mutations that were only detected in one aliquot had VAFs under 30%, making them susceptible to these random events. Our results indicated that scSTAMP can reliably identify mutations at medium to high VAFs in single cells.

mtDNA Mutations are Prevalent in Individual Cells. mtDNA is an important target of age-related accumulation of mutations (21, 29, 57–59). Our validation results suggested diverse mtDNA mutations at the single-cell level (Fig. 1*B*). To further investigate this at the single-cell level, we used scSTAMP to assess mitochondrial genomes of individual B lymphocytes and monocytes from peripheral blood that was obtained from a second donor, a healthy 76-y-old female. We constructed 768 single-cell mtDNA libraries for each of the two cell types. High-throughput sequencing yielded median depths of mtDNA coverage of 744.03× and 640.16× for the B lymphocytes and the monocytes, respectively (*SI Appendix*, Fig. S1*A*). Among these, 632 (82.3%) B lymphocytes and 617 (80.3%) monocytes had a median depth of coverage over 100×, which were retained for further analysis (*SI Appendix*, Fig. S1*B*). The distribution of high-quality consensus reads across 46 probes was highly correlated between the two cell types (Pearson's correlation = 0.91, *SI Appendix*, Fig. S2). The uniformly high mtDNA coverage enabled us to reliably identify and compare mtDNA mutations in these two cell types (*SI Appendix*, Figs. S3 and S4).

We identified 473 mutations with VAF over 20% at 416 mtDNA sites in the B lymphocytes, and 505 such mutations at

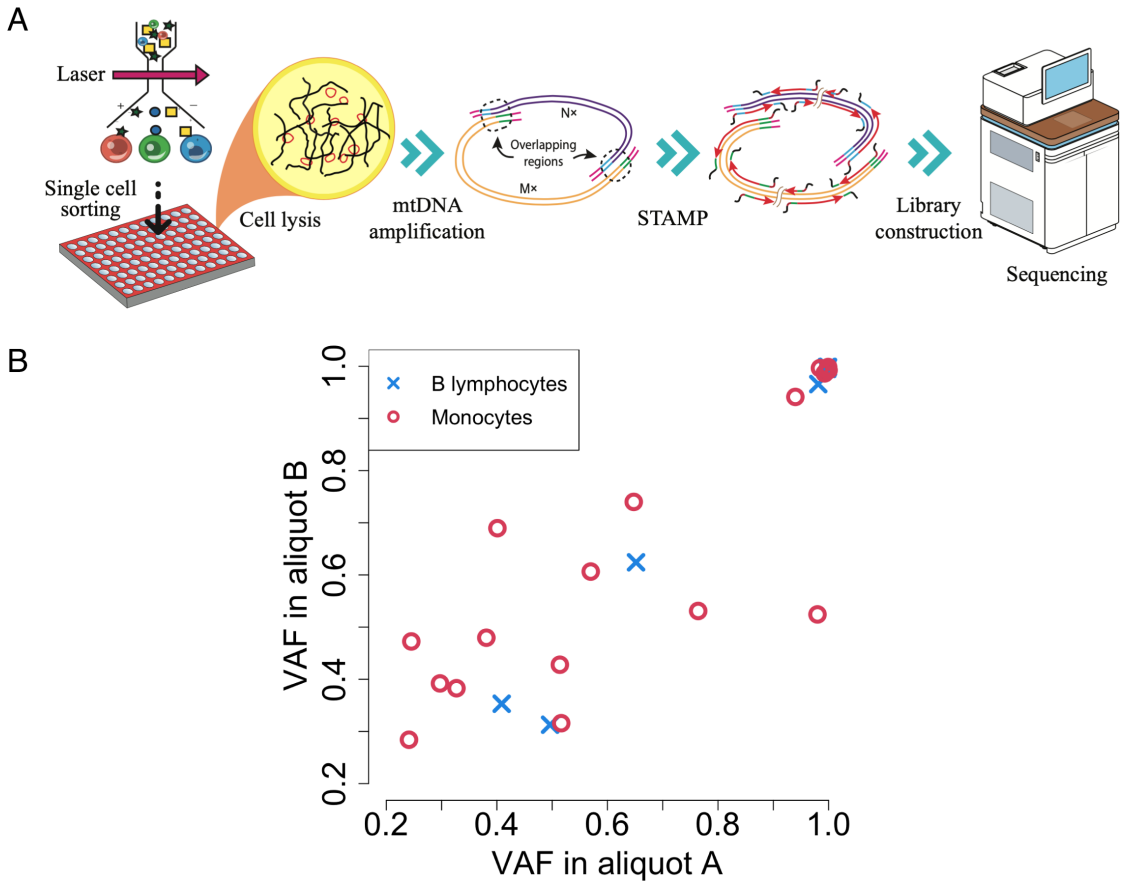

Fig. 1. Experimental workflow for scSTAMP and validation. (*A*) Schematics of scSTAMP. After single-cell sorting and cell lysis, the whole mtDNA genomes are amplified as two overlapping fragments directly from single-cell lysates in a single reaction. The resulted amplicons were processed by STAMP to generate sequencing libraries (25). (*B*) mtDNA mutations with VAFs > 20% in both aliquots derived from the same cells had highly correlated VAFs (Pearson's correlation = 0.88, *p*-value = 1.407×10^{-9}). We constructed 48 libraries each (24 cells × 2 aliquots) for B lymphocytes and monocytes. After sequencing, 2 B lymphocytes were excluded because both cells had sequencing reads returned for only one aliquot. The remaining 22 B lymphocytes and 24 monocytes had median depth of mtDNA coverage of over 100×.

439 sites in the monocytes (Fig. 2*A*). Each B lymphocyte and monocyte carried an average of 0.658 and 0.712 mutation sites, respectively. We noted that there were no significant differences in the number of mutations among plates (*SI Appendix*, Fig. S5*A* and Table S3) or significant correlations between number of mutations and median depth of coverage in a cell (*SI Appendix*, Fig. S5*B* and Table S3). Among all sites, 97.37% and 96.83% were transitions, respectively, leading to transition/transversion ratios of 37.0 and 30.5 in the B lymphocytes and the monocytes populations. Transitions G > A and T > C were the two major types of mutations (*SI Appendix*, Fig. S6).

Among studied cells, 318 (50.3%) of B lymphocytes and 346 (56.1%) of monocytes carried at least one mtDNA mutations (Fig. 2*B*). Interestingly, the majority of mtDNA mutation sites were singletons in the studied populations: 390 (82.5%) and 396 (78.4%) mutation sites were observed only once, which, respectively, accounted for 93.8% and 90.2% of all the mutation sites identified in the B lymphocyte and the monocyte populations (*SI Appendix*, Fig. S7). Among the singletons, 360 B lymphocyte singletons and 366 monocyte singletons were still singletons even when two cell populations were considered together. The distribution of VAFs skewed toward high and low ends of the VAF spectrum. Specifically, 24.3% of the observed mtDNA mutation sites had median VAF over 90% in B lymphocytes, which was significantly higher than that (17.5%) in monocytes (two-sided *P*-value = 0.01921. Fig. 2*C*). Given that we only focused on mutations with VAFs higher than 20%, our observations suggested that mtDNA mutations are prevalent at the single-cell level.

Single-Cell mtDNA Mutations are Functionally Important. We observed positive correlations between the number of mtDNA mutation sites located in a specific mtDNA region (genes, D-loop, or intergenic regions) in the cell populations and the size of the region. Specifically, Pearson's correlation is 0.958 (*p*-value < 2.2×10^{-16}) and 0.936 (*p*-value < 2.2×10^{-16}) for the B lymphocytes and the monocytes, respectively. On average, there were 0.026 and 0.028 mutation sites per base pair across the B lymphocytes and the monocytes, respectively. The results suggested that single-cell mtDNA mutations might be generated mostly at random (60).

The majority of the observed mutations are functionally important. In our data, mutations in the protein-coding genes are the most common: 312 B lymphocyte mutations at 271 sites (65.1%) and 328 monocyte mutations at 264 sites (60.1%) were in protein-coding genes (Fig. 3 *B* and *C*). The second most common type of mutations were in rRNA/tRNA-coding genes: 97 mutations at 84 (20.2%) mtDNA sites for B lymphocytes and 113 mutations at 103 mtDNA sites (23.5%) for monocytes were in rRNA-coding genes, while, respectively, 44 mutations at 42 sites (10.1%) and 39 mutations at 33 sites (7.5%) were in tRNA-coding genes. In total, protein-coding and tRNA/rRNA-coding mutation sites accounted for over 90% of all identified sites. Although the D-loop is considered highly variable

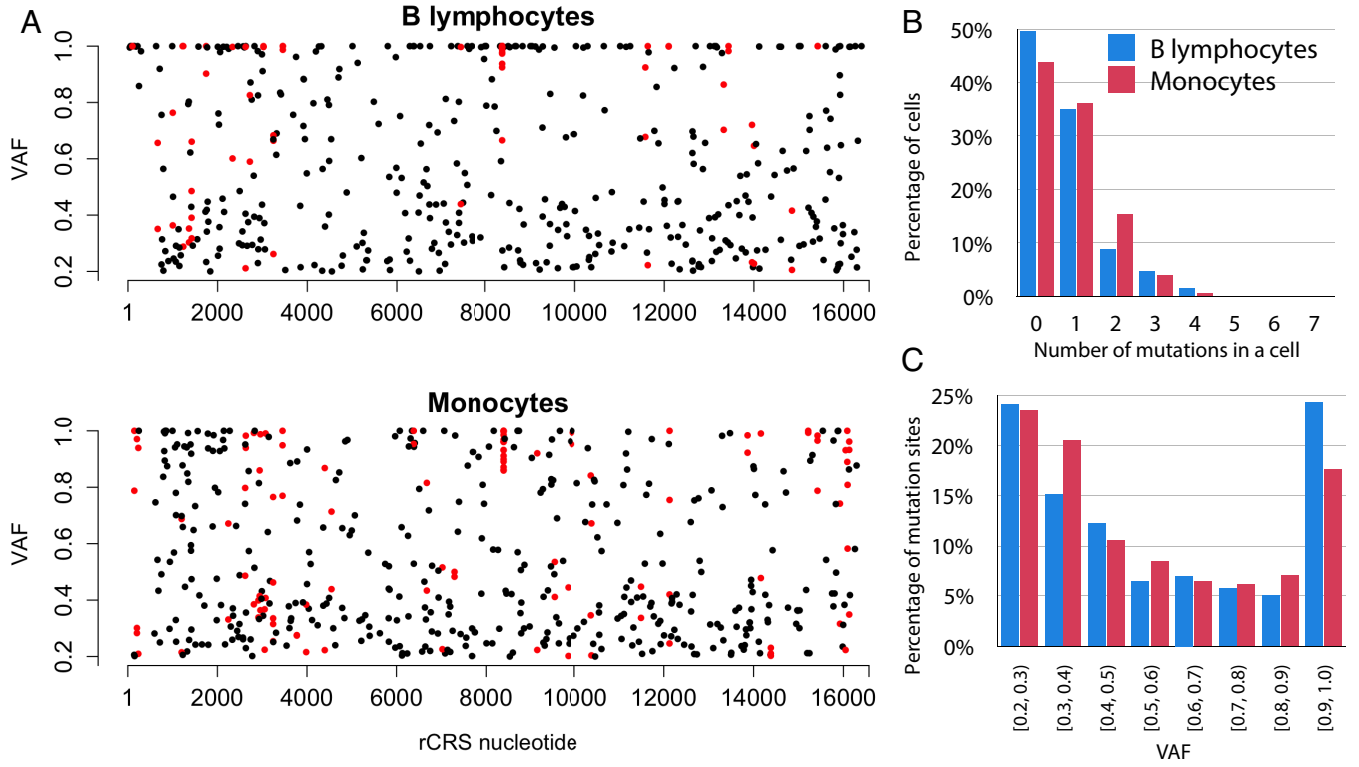

Fig. 2. Single-cell mtDNA mutations are prevalent. (*A*) Distribution of mtDNA mutations (*horizontal* axis) and their VAFs (*vertical* axis). Singletons and shared mutations in the specific cell populations were labeled in black and red, respectively. (*B*) Percentages of cells (*vertical* axis) that carried a specific number of mtDNA mutations (*Horizontal* axis). Chi-squared test of the contingency table of raw cell counts gave *P*-value = 0.006513, indicating statistically significant differences in the mutation burden of the two cell types. (*C*) Percentages of mtDNA mutation sites (*Vertical* axis) that had median VAFs within a specific range of VAFs across all occurrences of mutations on the same sites (shown in *Horizontal* axis). Chi-squared test of the contingency table of raw mutation counts gave *P*-value = 0.1232.

among individuals in human populations (61, 62), we observed only 20 mutations at 19 sites (4.6%) in B lymphocytes and 23 mutations at 14 sites (3.2%) in monocytes in the D-loop. However, as expected, most of the D-loop mutations were found in the two hypervariable regions of the D-loop, HVR1 (rCRS16024-16383), and HRV2 (HVR-II: rCRS57-372, HVR-III: rCRS438-574) (61): 18 and 23 mutations at 17 and 14 sites, respectively, in the B lymphocytes and the monocytes were found in these regions.

In addition, most protein-coding mutation sites were nonsynonymous. Among all identified protein-coding mutation sites, 70.8% (192 sites), 24.0% (65 sites), and 5.2% (14 sites) were nonsynonymous, synonymous, and stop-gain, respectively, in the B lymphocytes, and 74.2% (213 sites), 20.6% (59 sites), and 5.2% (15 sites), respectively, in the monocytes. The Combined Annotation-Dependent Depletion (CADD) scores measure the deleteriousness of single-nucleotide variants (63). Across all nonsynonymous mutations, 52.2% in B lymphocytes and 59.6% in monocytes had CADD scores over 20, which is a threshold for predicted high pathogenicity (24). The percentages were reduced to 44.2% and 40% among nonsynonymous mutations with VAF over 90%, suggesting that purifying selection against highly pathogenic mutations might have occurred in both cell types, but more efficiently in the monocytes than in the B lymphocytes.

Single-Cell mtDNA Mutation Load Varies by Cell Type. Different cell types may have different mtDNA mutation load, for which our data provided an interesting opportunity to investigate. Although both populations had a median of one mutation per cell, the average number of mutations per cell is 0.748 for B lymphocytes and 0.818 for monocytes (Figs. 2*B* and 4*A*). This ~9% difference in mean is statistically significant (two-sided Wilcoxon rank sum test, *P*-value = 0.028; Fig. 4*A*).

We further conducted a resampling-based simulation to study the minimum number of single cells required to represent single-cell mutation load of a single-cell population. Our simulation showed that with as few as 20 cells, the SD across resampled mean mutation load was so large, compared to the real values of mutation load in cell populations, that we were not able to reliably quantify the mutation load (Fig. 4*B*). However, with a sample size of 60 single cells, the differences in mutation load between these two cell populations can be reliably detected in the resampling. Our simulation further showed that SD of the mutation load converged to around 0.05 if we used 260 or more cells (Fig. 4*C*). The SD did not thereafter decrease too much with more cells in the simulation, indicating that we may need as few as about 260 cells to represent the mutation load in a single-cell population (Fig. 4*C*).

Discussions

Massively parallel single-cell mtDNA sequencing provides an avenue to dissect cell-specific mtDNA heteroplasmy. Recent developments in scATAC-seq and scMito-seq techniques signified critical steps toward inferring mtDNA heteroplasmy and clonal relationships in individual cells. However, most existing single-cell mtDNA sequencing methods use expensive commercial kits for single-cell mtDNA amplification and next-generation sequencing (NGS) library preparation. We have developed a cost-effective, targeted-sequencing approach for assessing single-cell mtDNA mutations. It was based on our previously developed approach, STAMP, to assessing single-nucleotide mtDNA mutations from cell population samples (25, 50). In comparison with scMito-seq

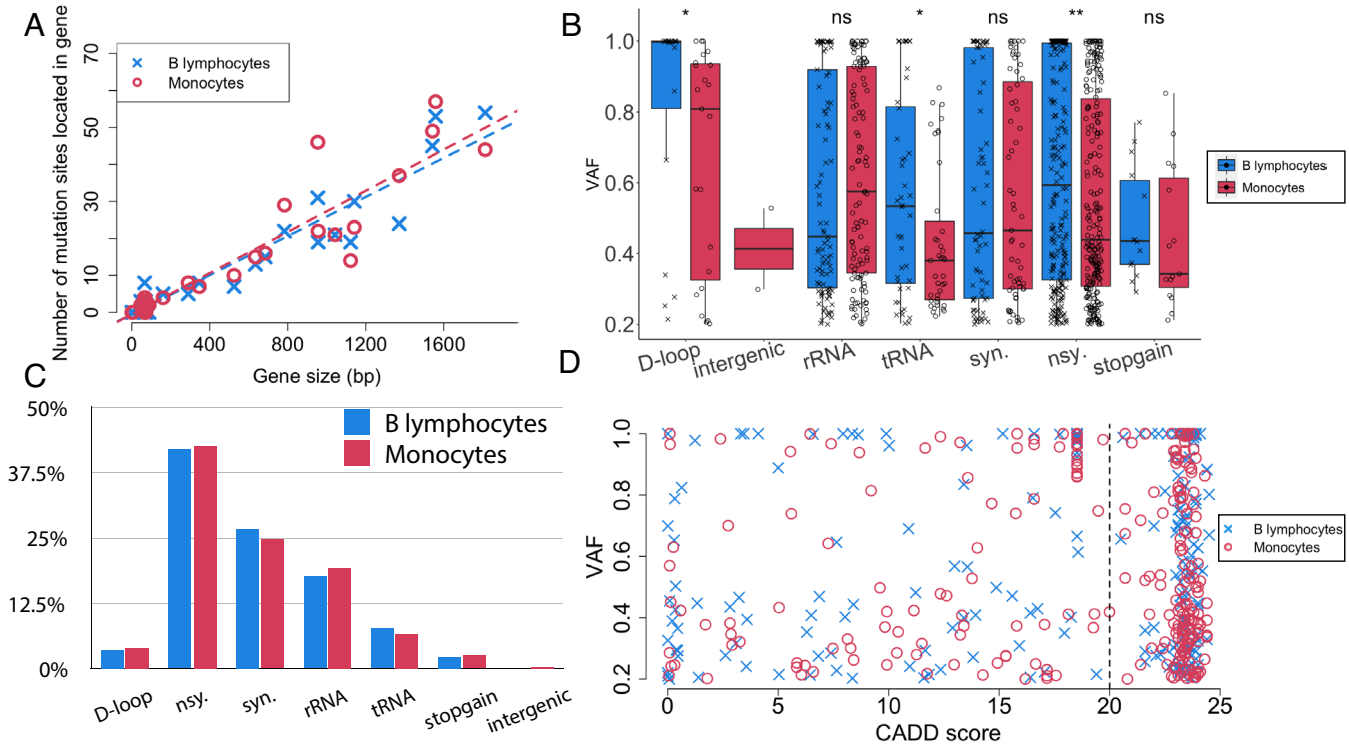

Fig. 3. Single-cell mtDNA mutations are functionally important. (*A*) Correlation between the number of mutations located in an mtDNA region in the populations (*Vertical* axis) and the size of the region (*Horizontal* axis). (*B*) VAF distribution (*Vertical* axis) of different categories (*Horizontal* axis) of mtDNA mutations. Syn.: synonymous mutations. Nsy.: nonsynonymous mutations. Asterisks indicated the significance levels of Wilcoxon two-sided tests. ns: $P > 0.05$, *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, and ****$P \leq 0.0001$. (*C*) The percentage of mutation sites (*Vertical* axis) that belonged to a specific functional category (*Horizontal* axis), normalized by the sequencing depth of coverage. Chi-squared test of the contingency table of raw counts gave *P*-value = 0.3627. Syn.: synonymous mutations. Nsy.: nonsynonymous mutations. (*D*) The VAFs (*Vertical* axis) and CADD score (*Horizontal* axis) of single-cell mtDNA mutations. The dotted vertical line represents a CADD score of 20.

which subjects circular mtDNA to RCA and tagmentation of the concatemeric products by Tn5 transposase, scSTAMP provides more streamlined experimental procedure (i.e., no purification after single-cell lysis) and improved mitochondrial genome alignment rates. We can customize arm sequences of scSTAMP to adapt to mtDNA of different species, while scMito-seq requires specialized commercial kits with lesser-known primer sequences for library preparation. It is important to point out that both methods share a similar flexibility for either high- or low-throughput applications.

Although single-cell mtDNA mutations have been investigated previously in a few studies, these studies did not systematically characterize the single-cell mtDNA mutation spectrum or their functional significance due to either low coverage of the mtDNA genome or small number of cells used in the studies (35, 44, 45). We validated scSTAMP by applying the method independently on two equal aliquots of the same cell that scSTAMP can reliably detect mtDNA mutations of medium to high VAFs (defined as >20%). It is worth noting that the current study is the only one conducting this approach to demonstrate the reliability of single-cell mtDNA sequencing method in our method development stage. With validated scSTAMP, we systematically assessed the patterns and dynamics of single-cell mtDNA mutations in B lymphocytes and monocytes, representing two major lineages of hematopoiesis, in an aged individual. We reported that mtDNA mutations of median to high VAFs were prevalent at the single-cell level and that most of them were singletons in the respective population. We further demonstrated that many of these were nonsynonymous and highly pathogenic. Given that most mutations were singletons in the cell populations, they were unlikely to be detected with cell population-based sequencing. Interestingly, mathematical modeling indicated that it is expected to have many somatic mtDNA singleton mutations given the sampled number of cells (64). Our results using published scATAC-seq data also revealed many functionally important mtDNA singletons (*SI Appendix*, Figs. S9 and S10).

Selection and genetic drift are considered important evolutionary forces shaping the mtDNA mutation spectrum in individual cells (53, 64, 65). Our results are consistent with the possibility that the mtDNA quality control system in cells is less efficient at purging pathogenic mutations with advancing age (66). The prevalence of high-VAF pathogenic mutations that we identified implied weak purifying selection or damaged mtDNA quality control in aged cells (67). Differences in mtDNA copy number in different cell types may also impact the dynamics of mtDNA mutations, because they represent differences in the effective population sizes of mtDNA molecules (29). In this study, we observed higher percentages of fixed or nearly fixed mutations in the B lymphocytes than in the monocytes. Furthermore, there are greater number of mutations and more efficient purifying selection against pathogenic mutations in monocytes than B lymphocytes. All these observations are expected if B lymphocytes have lower mtDNA copy number, and thus lower effective population size, than monocytes (68). Indeed, B lymphocytes are much smaller in size than monocytes (69), supporting possible differences in mtDNA copy numbers between two cell types. It remains to be studied in the future whether there is a relationship among mtDNA copy number, single-cell mtDNA mutation dynamics, and the rate of mitochondrial functional decline in different cell types.

We recently reported an increase in mtDNA mutations with age in the females from the UK10K project (21), which is consistent

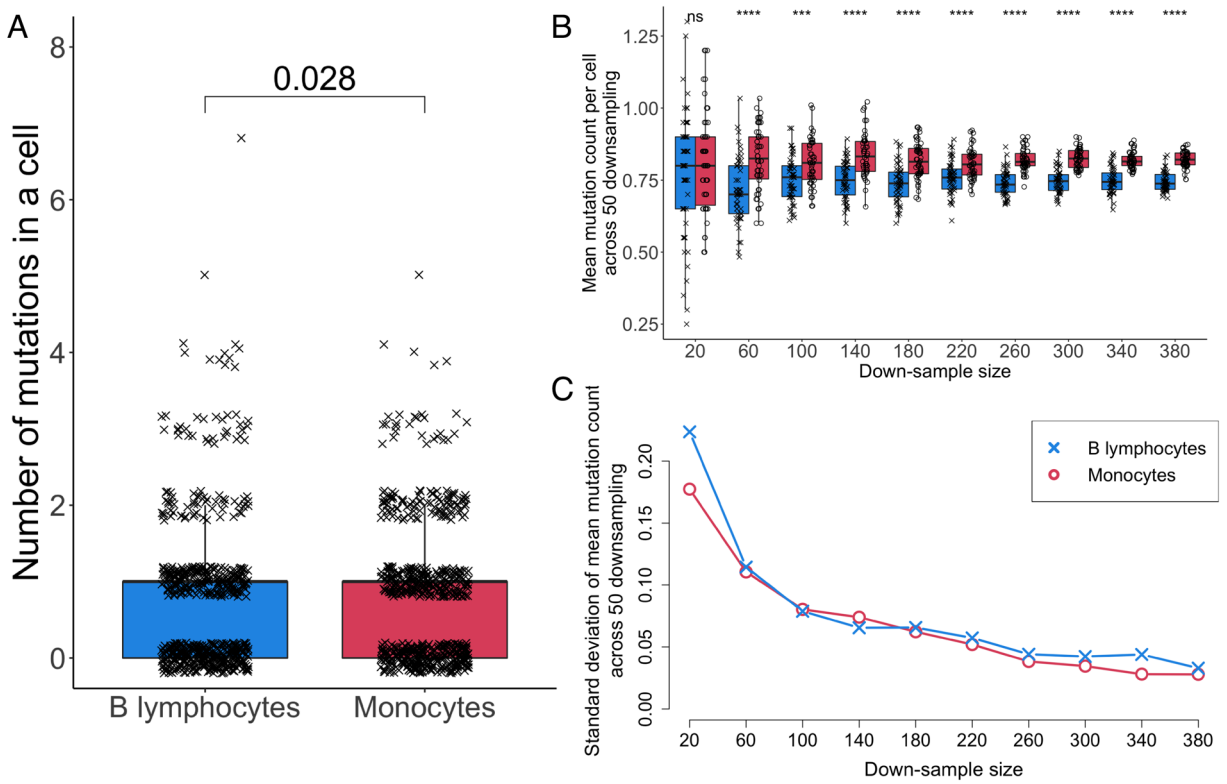

Fig. 4. Single-cell mtDNA mutation load between B lymphocytes and the monocytes. (*A*) Comparison of the number of mutations carried by a cell (*Vertical* axis) between the B lymphocytes and the monocytes. The *P*-value shown was calculated by two-sided, unpaired Wilcoxon rank sum test. (*B*) A down-sampling-based simulation to examine the number of cells required to represent the mutation load. We sampled a specific number of cells (indicated in the *Horizontal* axis) without replacement and computed the mean number of mutations, for 50 times each. Asterisks indicated the significance levels of Wilcoxon two-sided tests. ns: $P > 0.05$, *$P \le 0.05$, **$P \le 0.01$, ***$P \le 0.001$, and ****$P \le 0.0001$. All *P*-values are two-sided and unpaired by the Wilcoxon rank sum test. (*C*) SD of mean mutation load across 50 resamples for each specific number of cells down-sampled.

with other studies (26, 70, 71). However, the resolution of these previous studies was limited by the use of peripheral blood mononuclear cells that were mixtures of different types of cells and by the use of bulk samples that masked mutations at the single-cell level. In this study, using only 24 B lymphocytes and 24 monocytes from the validation sample, we showed that the mutation load of B lymphocytes from the first donor (~60-y-old) was lower than that from the second donor (76-y-old), but the mutation load of monocytes was similar between the two donors (*SI Appendix*, Fig. S8). It is important to note that variation in mutation load calculation increases when the number of cells sampled is small (Fig. 4*C*). Therefore, we need to consider randomness in sampling, cell types, and/or other features related to the donors in interpreting the observation between these two individuals. A systematic characterization of single-cell mtDNA mutations during healthy aging and in different disease settings is warranted.

Aging is a process marked by a progressive degradation of cellular functions (66). Mitochondrial dysfunction is considered a hallmark of aging, where OXPHOS functionality tends to decline along time (72–74). Decline in mitochondrial function was also hypothesized to be causal to many other age-related features (66, 75, 76). However, the mechanisms underlying mitochondrial functional decline with advancing age remain controversial. mtDNA is prone to accumulation of point mutations, insertions and deletions, and large structural arrangements in aging (21, 29, 57–59), which has been suggested to be causal to age-related human multisystem disorders, impaired mitochondrial functions, premature aging, and reduced life span in mice (60, 65, 77, 78). It has been reported that some pathogenic mitochondrial tRNA mutations are associated with the systematic disruption of cellular epigenetic pathways, the severity of which depends on the levels of the mutations (4, 28, 56). Nevertheless, most of these studies were conducted using bulk mtDNA analysis, and the mtDNA mutations observed in cell populations were too low in frequency to be functionally significant. Our results in this study showed the prevalence of high-frequency pathogenic mutations at the single-cell level. Furthermore, most pathogenic mutations were singletons in the cell population, preventing them from being detected in the bulk mtDNA analysis. Given the high percentage of cells harboring these mutations, they could collectively lead to mitochondrial functional decline at the tissue and organism level.

Heteroplasmic pathogenic mtDNA mutations are associated with age-related neurodegeneration, including Alzheimer's disease, but the dynamics of the mitochondrial genome in neurological disorders is largely unexplored. Single-cell analysis of heteroplasmy holds promise for revealing complexities in mtDNA heteroplasmic dynamics. We believe that development of a cost-effective tool, like scSTAMP, as well as large-scale assessment of dynamics of mtDNA mutations at single-cell level will accelerate our understanding of the roles of mtDNA variations in human diseases and facilitate discovery of new therapies for neurodegenerative diseases by targeting mitochondria-related pathways (79, 80).

It is worth noting that there are two limitations in our study. One limitation is that we have only investigated mtDNA single-nucleotide variants, so our results likely underestimated the prevalence of mtDNA mutations given the possible existence of insertions/deletions (indels) and structural variants. In addition, we only studied single-nucleotide variants with VAFs over 20%, which underestimates the functional significance of single-cell

mtDNA mutations because low VAF mutations and larger-scale mutations spanning two or more base pairs were not investigated. Another limitation is that our samples were drawn from only two female donors at their advanced ages. It is possible that the prevalence of single-cell mtDNA mutations was unique to these two females instead of being a norm in human population. We believe this is unlikely because using the published scATAC-seq data, we were also able to identify many highly diverse and functionally important mtDNA mutations at the single-cell level (*SI Appendix*, Figs. S9 and 10). Taken together, our results demonstrate the diversity of single-cell mtDNA mutations that were previously underappreciated and highlight the importance of investigating age-related mtDNA mutation dynamics at the single-cell level and their functional consequence in aging and age-related diseases.

Materials and Methods

Study Samples. A validation sample (buffy coat from human blood, a ~60-y-old female) was used to establish the scSTAMP workflow and assess the performance. A second sample (buffy coat from human blood, a 76-y-old female) was used for analysis of single-cell mtDNA mutation spectrum. Both human samples were collected in the Human Metabolic Research Unit at Cornell University. The study was approved by the Cornell University Institutional Review Board, and all the subjects provided written informed consent in accordance with the Declaration of Helsinki.

Single-Cell Mitochondrial Genome Sequencing. Mononuclear cells, from the buffy coat of peripheral blood, were separated by density gradient centrifugation and washed with phosphate-buffered saline (PBS). The cells were suspended in PBS containing 1% bovine serum albumin (BSA) to 1 × 10^7 cells/mL. PE Anti-Human CD19 (eBioscience), Qdot 705 CD14 antibody (Thermo Fisher Scientific), and LIVE/DEAD fixable dead cell stain kit (Thermo Fisher Scientific) with far red fluorescent reactive dye (Thermo Fisher Scientific Inc) were used for staining of the cells according to the manufacturer's instructions. Single live CD19 and Qdot 705 positive cells were sorted by a BD FACSAria cell sorter directly into individual wells of a 384 well plate predispensed with 1 μL of lysis buffer (200 mM NaOH and 50 mM dithiothreitol). The cells were spun down briefly and stored at −80 °C until ready for processing.

Directly before amplification, single cells were incubated at 65 °C for 10 min to liberate total DNA. After the lysis step, 12 μL of PCR master mix containing 1 μL of 200 mM tricine, 2.6 μL of 5× LongAmp *Taq* reaction buffer, 0.39 μL of 10mM dNTPs, 0.52 μL of each primer (5 μM, *SI Appendix*, *Supplementary Tables*), and 0.52 μL (1.25 units) of LongAmp *Taq* DNA polymerase were added directly to each well. Amplification was performed with the following cycling parameters: an initial denaturation of 94 °C for 30 s; followed by 45 cycles of 94 °C for 15 s, 61 °C for 30 s, 65 °C for 8 min and 20 s, and 72 °C for 30 s. After cycling, there was a final extension step of 72 °C for 10 min.

The amplified product for each cell was diluted 200-fold into 1× TE (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0). 3.2 μL of diluted product was then input into an 8 μL hybridization reaction containing 1× Ampligase buffer and 1× probe-mixture pool. Hybridization was started with denaturation at 95 °C for 10 min, followed by a gradual decrease in temperature of 1 °C per min to 55 °C and hybridized for 20 h in a thermocycler. Then 6 μL gap-filling mix (0.1 mM dNTPs, 0.6M Betaine, 0.1 M $(NH_4)_2SO_4$, 0.5 U of *TSP* DNA polymerase, and 0.5 U of Ampligase in 1× Ampligase buffer) was added to the reaction and incubated at 55 °C for another 20 h for gap filling.

For each capture reaction, captured targets were amplified and barcoded with customized P5i5 indexing primer (*SI Appendix*, *Supplementary Tables*) and P7i7 indexing primer (*SI Appendix*, *Supplementary Tables*). The PCR conditions were as follows: 1 μL of captured product, 1× Phusion HF buffer (NEB), 0.2 mM dNTP, 0.5 μM of each of the indexing primers, and 0.01 U/μL Phusion Hot-Start II DNA polymerase (NEB) in a 50 μL final reaction volume which was incubated at 98 °C for 30 s followed by 25 cycles of 98 °C for 10 s, 65 °C for 15 s, and 72 °C for 15 s and a final extension for 2 min at 72 °C. 10 μL of product from each sample were pooled, and then the pooled libraries were first purified with 0.65× volumes of SPRI beads, eluted in 20 μL of nuclease free water. The corresponding expected size of 550 to 600 bp was gel extracted and resuspended in 15 μL of TE buffer. The pool was quantified by a Qubit dsDNA HS assay kit (Thermo Fisher Scientific).

The pool was sequenced with 2 × 250 paired-end reads, on a HiSeq flow cell with customized sequencing primers. The first sequence read (251 nt) was obtained using the Read 1 primer (*SI Appendix*, *Supplementary Tables*). The first index, located at the 3′ end of the fragment, was sequenced using the i7 Index primer (*SI Appendix*, *Supplementary Tables*). The second index read was then performed to obtain the index sequence at the 5′ end of the fragment using the adapter lawn on the surface of the sequencing flow cells. Finally, the second sequence read (251 nt) was obtained using the Read 2 primer (*SI Appendix*, *Supplementary Tables*). The overall process of cluster generation, sequencing, image processing, demultiplexing, and quality score calculation was performed on the HiSeq 2500.

mtDNA Variant Identification. We used the STAMP toolkit (https://github.com/mtstamp/stamp) developed in (25) to process sequencing reads generated from scSTAMP. A brief workflow was as follows. Ligation arms, extension arms, and molecular barcodes were first trimmed from raw paired-end reads, and the trimmed read pairs were sorted into clusters of capture products based on EL arm sequences. The resulted paired-end reads were then mapped in a first round to the complete human genome (GRCh38 full assembly plus decoy, alternate contigs and HLA sequences, ftp://ftp.1000genomes.ebi.ac.uk) by BWA-MEM (version 0.7.17) (81) and were mapped in a second round to a modified mtDNA sequence with the final 120 bp copied to the start. Reads mapped to the target regions were locally realigned with Freebayes bamleftalign (version 1.1.0) (82), and their base qualities were recalibrated with samtools calmd (version 1.6) (83). For paired-end reads with the same molecular barcode, the base information at corresponding sites was corrected and merged using a Bayesian approach to generate a consensus read representation (25). The same approach was also used to correct and merge base information within the overlapping region of the paired-end reads. Nuclear-mitochondrial segment (NUMT) sequences are determined and removed by the STAMP toolkit (25). Finally, information on consensus reads was output into a pileup file with samtools and was used to call mtDNA variants.

Quality Control and Bioinformatics Annotations. To ensure accurate identification of mtDNA variants, we discarded single-cell libraries with median depth of mtDNA coverage <100×. 82.3% (n = 632) of B lymphocytes and 80.3% (n = 617) of monocytes were retained for downstream analysis. To ensure the quality of mtDNA variant calling results, we stringently filtered the results to retain only mutations with (1) depth of coverage > 100, (2) ≥ 5 supporting reads for mutations, and (3) VAF ≥ 20% relative to the consensus mtDNA sequences of the respective cell populations. The functional impact of mtDNA variants was annotated with the ANNOVAR pipeline (84), and pathogenicity of mtDNA variants was evaluated using CADD (63), both of which were provided by the STAMP pipeline.

Statistical Analysis.
All statistical analyses and plotting were conducted with the R language (version 4.03) (85). Base R plotting and ggplot2 were used for figure generation. Unless otherwise specified, all statistical comparisons between two samples were conducted unpaired, and the P-values were two-sided.

Data, Materials, and Software Availability. The sequencing data have been submitted to the GenBank Sequence Read Archive under the accession number PRJNA804226.

ACKNOWLEDGMENTS. We would like to thank Drs. Douglas Wallace, Andrew Clark, Amy Williams, and Jason Mezey, for their insightful discussions for this project, and Dr. Yiping Wang for his comments on the manuscript. The work was supported by NIH R01GM117190A1 and a start-up fund to Z.G.

Author affiliations: [a]Division of Nutritional Sciences, Cornell University, Ithaca, NY 14853; [b]Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, China; [c]Greater Bay Area Institute of Precision Medicine, Fudan University, Nansha District, Guangzhou 511400, China; and [d]School of Life Sciences, Fudan University, Shanghai 200433, China

1. A. P. West, G. S. Shadel, S. Ghosh, Mitochondria in innate immune responses. *Nat. Rev. Immunol.* **11**, 389–402 (2011).
2. R. Sandhir, A. Halder, A. Sunkaria, Mitochondria as a centrally positioned hub in the innate immune response. *BBA Mol. Basis Dis.* **1863**, 1090–1097 (2017).
3. D. C. Wallace, W. Fan, Energetics, epigenetics, mitochondrial genetics. *Mitochondrion* **10**, 12–31 (2010).
4. M. Wiese, A. J. Bannister, Two genomes, one cell: Mitochondrial-nuclear coordination via epigenetic pathways. *Mol. Metab.* **38**, 100942 (2020).
5. C. Wang, R. J. Youle, The role of mitochondria in apoptosis. *Annu. Rev. Genet.* **43**, 95–118 (2009).
6. J. B. Stewart, P. F. Chinnery, Extreme heterogeneity of human mitochondrial DNA from organelles to populations. *Nat. Rev. Genet.* **22**, 106–118 (2021).
7. C. Haag-Liautard *et al.*, Direct estimation of the mitochondrial DNA mutation rate in Drosophila melanogaster. *PLoS Biol.* **6**, e204 (2008).
8. D. C. Wallace, Mitochondrial DNA mutations in disease and aging. *Environ. Mol. Mutagen.* **51**, 440–450 (2010).
9. R. W. Taylor, D. M. Turnbull, Mitochondrial DNA mutations in human disease. *Nat. Rev. Genet.* **6**, 389–402 (2005).
10. E. A. Schon, S. DiMauro, M. Hirano, Human mitochondrial DNA: Roles of inherited and somatic mutations. *Nat. Rev. Genet.* **13**, 878–890 (2012).
11. A. B. Otten *et al.*, Mutation-specific effects in germline transmission of pathogenic mtDNA variants. *Hum. Reprod.* **33**, 1331–1341 (2018).
12. M. S. Sharpley *et al.*, Heteroplasmy of mouse mtDNA is genetically unstable and results in altered behavior and cognition. *Cell* **151**, 333–343 (2012).
13. M. Keogh, P. F. Chinnery, Hereditary mtDNA heteroplasmy: A baseline for aging? *Cell Metab.* **18**, 463–464 (2013).
14. M. Kong *et al.*, Aging-associated accumulation of mitochondrial DNA mutations in tumor origin. *Life Med.* **338**, 151–159 (2022).
15. D. C. Wallace, D. Chalkia, Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease. *Cold Spring Harb. Perspect. Biol.* **5**, a021220 (2013).
16. M. Li *et al.*, Detecting heteroplasmy from high-throughput sequencing of complete human mitochondrial DNA genomes. *Am. J. Hum. Genet.* **87**, 237–249 (2010).
17. M. X. Sosa *et al.*, Next-generation sequencing of human mitochondrial reference genomes uncovers high heteroplasmy frequency. *PLoS Comput. Biol.* **8**, e1002737 (2012).
18. H. Goto *et al.*, Dynamics of mitochondrial heteroplasmy in three families investigated via a repeatable re-sequencing study. *Genome Biol.* **12**, 1–16 (2011).
19. B. A. Payne *et al.*, Universal heteroplasmy of human mitochondrial DNA. *Hum. Mol. Genet.* **22**, 384–390 (2013).
20. L. C. Greaves *et al.*, Clonal expansion of early to mid-life mitochondrial DNA point mutations drives mitochondrial dysfunction during human ageing. *PLoS Genet.* **10**, e1004620 (2014).
21. R. Zhang, Y. Wang, K. Ye, M. Picard, Z. Gu, Independent impacts of aging on mitochondrial DNA quantity and quality in humans. *BMC Genomics* **18**, 890 (2017).
22. K. Ye, J. Lu, F. Ma, A. Keinan, Z. Gu, Extensive pathogenicity of mitochondrial heteroplasmy in healthy human individuals. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 10654–10659 (2014).
23. R. Zhang, K. Nakahira, X. Guo, A. M. Choi, Z. Gu, Very short mitochondrial DNA fragments and heteroplasmy in human plasma. *Sci. Rep.* **6**, 1–10 (2016).
24. Y. Wang, M. Picard, Z. Gu, Genetic evidence for elevated pathogenicity of mitochondrial DNA heteroplasmy in autism spectrum disorder. *PLoS Genet.* **12**, e1006391 (2016).
25. X. Guo, Y. Wang, R. Zhang, Z. Gu, STAMP: A multiplex sequencing method for simultaneous evaluation of mitochondrial DNA heteroplasmies and content. *NAR Genom. Bioinform.* **2**, lqaa065 (2020).
26. M. Li, R. Schröder, S. Ni, B. Madea, M. Stoneking, Extensive tissue-related and allele-related mtDNA heteroplasmy suggests positive selection for somatic mutations. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 2491–2496 (2015).
27. A. Hübner *et al.*, Sharing of heteroplasmies between human liver lobes varies across the mtDNA genome. *Sci. Rep.* **9**, 1–11 (2019).
28. M. Picard *et al.*, Progressive increase in mtDNA 3243A> G heteroplasmy causes abrupt transcriptional reprogramming. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E4033–E4042 (2014).
29. J. P. Burgstaller *et al.*, Large-scale genetic analysis reveals mammalian mtDNA heteroplasmy dynamics and variance increase through lifetimes and generations. *Nat. Commun.* **9**, 2488 (2018).
30. Y. Wang *et al.*, Association of mitochondrial DNA content, heteroplasmies and inter-generational transmission with autism. *Nat. Commun.* **13**, 3790 (2022).
31. L. C. Greaves *et al.*, 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**, e1003082 (2012).
32. E. Kang *et al.*, Age-related accumulation of somatic mitochondrial DNA mutations in adult-derived human iPSCs. *Cell Stem Cell* **18**, 625–636 (2016).
33. Y.-G. Yao, S. Kajigaya, N. S. Young, Mitochondrial DNA mutations in single human blood cells. *Mutat. Res.* **779**, 68–77 (2015).
34. J. Morris *et al.*, Pervasive within-mitochondrion single-nucleotide variant heteroplasmy as revealed by single-mitochondrion sequencing. *Cell Rep.* **21**, 2706–2713 (2017).
35. L. S. Ludwig *et al.*, Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics. *Cell* **176**, 1325–1339 e1322 (2019).
36. J. Xu *et al.*, Single-cell lineage tracing by endogenous mutations enriched in transposase accessible mitochondrial DNA. *Elife* **8**, e45105 (2019).
37. H. A. Coller *et al.*, High frequency of homoplasmic mitochondrial DNA mutations in human tumors can be explained without selection. *Nat. Genet.* **28**, 147–150 (2001).
38. P. F. Chinnery, D. C. Samuels, Relaxed replication of mtDNA: A model with implications for the expression of disease. *Am. J. Hum. Genet.* **64**, 1158–1165 (1999).
39. J. Elson, D. Samuels, D. Turnbull, P. Chinnery, Random intracellular drift explains the clonal expansion of mitochondrial DNA mutations with age. *Am. J. Hum. Genet.* **68**, 802–806 (2001).
40. W. Wei *et al.*, Germline selection shapes human mitochondrial DNA diversity. *Science* **364**, eaau6520 (2019).
41. Y. He *et al.*, Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. *Nature* **464**, 610–614 (2010).
42. G. Avital *et al.*, Mitochondrial DNA heteroplasmy in diabetes and normal adults: Role of acquired and inherited mutational patterns in twins. *Hum. Mol. Genet.* **21**, 4214–4224 (2012).
43. C. Bi *et al.*, Single-cell individual complete mtDNA sequencing uncovers hidden mitochondrial heterogeneity in human and mouse oocytes. *bioRxiv* [Preprint] (2020). Accessed 10 July 2022 https://doi.org/10.1101/2020.12.28.424537
44. C. A. Lareau *et al.*, Massively parallel single-cell mitochondrial DNA genotyping and chromatin profiling. *Nat. Biotechnol.* **39**, 451–461 (2020), 10.1038/s41587-020-0645-6.
45. E. Jaberi, E. Tresse, K. Grønbæk, J. Weischenfeldt, S. Issazadeh-Navikas, Identification of unique and shared mitochondrial DNA mutations in neurodegeneration and cancer by single-cell mitochondrial DNA structural variation sequencing (MitoSV-seq). *EBioMedicine* **57**, 102868 (2020).
46. E. P. Mimitou *et al.*, Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells. *Nat. Biotechnol.* **39**, 1246–1258 (2021).
47. H. Medini, T. Cohen, D. Mishmar, Mitochondrial gene expression in single cells shape pancreatic beta cells' sub-populations and explain variation in insulin pathway. *Sci. Rep.* **11**, 466 (2021).
48. T. E. Miller *et al.*, Mitochondrial variant enrichment from high-throughput single-cell RNA sequencing resolves clonal populations. *Nat. Biotechnol.* **40**, 1030–1034 (2022).
49. Z. Li *et al.*, Chromatin-accessibility estimation from single-cell ATAC-seq data with scOpen. *Nat. Commun.* **12**, 6386 (2021).
50. Y. Wang, X. Guo, K. Ye, M. Orth, Z. Gu, Accelerated expansion of pathogenic mitochondrial DNA heteroplasmies in Huntington's disease. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2014610118 (2021).
51. M. Kondo, Lymphoid and myeloid lineage commitment in multipotent hematopoietic progenitors. *Immunol. Rev.* **238**, 37–46 (2010).
52. M. Kondo *et al.*, Biology of hematopoietic stem cells and progenitors: Implications for clinical application. *Annu. Rev. Immunol.* **21**, 759–806 (2003).
53. J. B. Stewart, P. F. Chinnery, The dynamics of mitochondrial DNA heteroplasmy: Implications for human health and disease. *Nat. Rev. Genet.* **16**, 530–542 (2015).
54. S. E. Durham, D. C. Samuels, L. M. Cree, P. F. Chinnery, Normal levels of wild-type mitochondrial DNA maintain cytochrome c oxidase activity for two pathogenic mitochondrial DNA mutations but not for m. 3243A→ G. *Am. J. Hum. Genet.* **81**, 189–195 (2007).
55. R. Rossignol *et al.*, Mitochondrial threshold effects. *Biochem. J.* **370**, 751–762 (2003).
56. P. K. Kopinski *et al.*, Regulation of nuclear epigenome by mitochondrial DNA heteroplasmy. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 16028–16035 (2019).
57. A. Linnane, T. Ozawa, S. Marzuki, M. Tanaka, Mitochondrial DNA mutations as an important contributor to ageing and degenerative diseases. *Lancet* **333**, 642–645 (1989).
58. K. Khrapko *et al.*, Cell-by-cell scanning of whole mitochondrial genomes in aged human heart reveals a significant fraction of myocytes with clonally expanded deletions. *Nucleic Acids Res.* **27**, 2434–2441 (1999).
59. A. Bender *et al.*, High levels of mitochondrial DNA deletions in substantia nigra neurons in aging and Parkinson disease. *Nat. Genet.* **38**, 515–517 (2006).
60. M. Vermulst, J. H. Bielas, L. A. Loeb, Quantification of random mutations in the mitochondrial genome. *Methods* **46**, 263–268 (2008).
61. M. Van Oven, M. Kayser, Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum. Mutat.* **30**, E386–E394 (2009).
62. T. J. Nicholls, M. Minczuk, In D-loop: 40 years of mitochondrial 7S DNA. *Exp. Gerontol.* **56**, 175–181 (2014).
63. M. Kircher *et al.*, A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**, 310–315 (2014).
64. H. Li, J. Slone, L. Fei, T. Huang, Mitochondrial DNA variants and common diseases: A mathematical model for the diversity of age-related mtDNA mutations. *Cells* **8**, 608 (2019).
65. G. C. Kujoth *et al.*, Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. *Science* **309**, 481–484 (2005).
66. C. Lopez-Otin, M. A. Blasco, L. Partridge, M. Serrano, G. Kroemer, The hallmarks of aging. *Cell* **153**, 1194–1217 (2013).
67. M. A. Walker *et al.*, Purifying selection against pathogenic mitochondrial DNA in human T cells.*N. Engl. J. Med.* **383**, 1556–1563 (2020).
68. M. Picard, Blood mitochondrial DNA copy number: What are we counting. *Mitochondrion* **60**, 1–11 (2021).
69. S. Wang, K. Mak, L. Chen, M. Chou, C. Ho, Heterogeneity of human blood monocyte: Two subpopulations with different sizes, phenotypes and functions. *Immunology* **77**, 298 (1992).
70. N. Sondheimer *et al.*, Neutral mitochondrial heteroplasmy and the influence of aging. *Hum. Mol. Genet.* **20**, 1653–1659 (2011).
71. M. Li *et al.*, Transmission of human mtDNA heteroplasmy in the genome of the Netherlands families: Support for a variable-size bottleneck. *Genome Res.* **26**, 417–426 (2016).
72. D. R. Green, L. Galluzzi, G. Kroemer, Mitochondria and the autophagy–inflammation–cell death axis in organismal aging. *Science* **333**, 1109–1112 (2011).
73. J. Campisi, Aging, cellular senescence, and cancer. *Annu. Rev. Physiol.* **75**, 685–705 (2013).
74. N. Sun, R. J. Youle, T. Finkel, The mitochondrial basis of aging. *Mol. cell* **61**, 654–666 (2016).
75. J. Y. Jang, A. Blum, J. Liu, T. Finkel, The role of mitochondria in aging. *J. Clin. Invest.* **128**, 3662–3670 (2018).
76. G. Desdín-Micó *et al.*, T cells with dysfunctional mitochondria induce multimorbidity and premature senescence. *Science* **368**, 1371–1376 (2020).
77. A. Trifunovic *et al.*, Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature* **429**, 417–423 (2004).
78. D. C. Wallace, A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: A dawn for evolutionary medicine. *Annu. Rev. Genet.* **39**, 359–407 (2005).
79. M. Scheibye-Knudsen, E. F. Fang, D. L. Croteau, D. M. Wilson III, V. A. Bohr, Protecting the mitochondrial powerhouse. *Trends Cell Biol.* **25**, 158–170 (2015).
80. E. F. Fang *et al.*, Mitophagy inhibits amyloid-β and tau pathology and reverses cognitive deficits in models of Alzheimer's disease. *Nat. Neurosci.* **22**, 401–412 (2019).
81. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv [preprint] (2013). https://doi.org/10.48550/arXiv.1303.3997. Accessed 3 June 2021
82. E. Garrison, G. Marth, Haplotype-based variant detection from short-read sequencing. *arXiv* [Preprint] (2012). https://doi.org/10.48550/arXiv.1207.3907. Accessed 3 June 2021
83. H. Li *et al.*, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
84. K. Wang, M. Li, H. Hakonarson, ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164–e164 (2010).
85. R. C. Team, R: A language and environment for statistical computing (2013). Vienna, Austria. http://www.R-project.org/.