Chimeric mitochondrial RNA transcripts predict mitochondrial genome deletion mutations in mitochondrial genetic diseases and aging

Amy R Vandiver, Allen Herbst, Paul Stothard, et al.

Genome Res. published online November 27, 2024
Access the most recent version at doi:10.1101/gr.279072.124

Chimeric mitochondrial RNA transcripts predict mitochondrial genome deletion mutations in mitochondrial genetic diseases and aging.

Amy R. Vandiver[1,2], Allen Herbst[3*], Paul Stothard[3], Jonathan Wanagat[2,4]

[1]Department of Medicine, Division of Dermatology, UCLA, Los Angeles, California, USA

[2]Veterans Administration Greater Los Angeles Healthcare System, Los Angeles, California, USA

[3]Department of Agricultural, Food, and Nutritional Science, University of Alberta, Edmonton, Alberta, Canada

[4]Department of Medicine, Division of Geriatrics, UCLA, Los Angeles, California, USA

*Current affiliation: US Geological Survey, National Wildlife Health Center, Madison, Wisconsin, USA

Correspondence:

Jonathan Wanagat, MD, PhD

Division of Geriatrics, Department of Medicine

University of California, Los Angeles

650 Charles E. Young Drive South, RM 34-115

Los Angeles, CA 90095

Phone: 310-825-8253

Email: jwanagat@mednet.ucla.edu

Key words: mitochondria, chimeric RNA, aging, mutations

Running title: Chimeric mitochondrial RNA

Abstract

While it is well understood that mitochondrial DNA (mtDNA) deletion mutations cause incurable diseases and contribute to aging, little is known about the transcriptional products that arise from these DNA structural variants. We hypothesized that mitochondrial genomes containing deletion mutations express chimeric mitochondrial RNAs. To test this, we analyzed human and rat RNA sequencing data to identify, quantitate, and characterize chimeric mitochondrial RNAs.

We observed increased chimeric mitochondrial RNA frequency in samples from patients with mitochondrial genetic diseases and in samples from aged humans. The spectrum of chimeric mitochondrial transcripts reflected the known pattern of mtDNA deletion mutations. To test the hypothesis that mtDNA deletions induce chimeric RNA transcripts, we treated 18mo and 34mo rats with guanidinopropionic acid to induce high levels of skeletal muscle mtDNA deletion mutations. With mtDNA deletion induction, we demonstrate that the chimeric mitochondrial transcript frequency also increased and correlated strongly with an orthogonal DNA-based mutation assay performed on identical samples. Further, we show that the frequency of chimeric mitochondrial transcripts predicts expression of both nuclear and mitochondrial genes central to mitochondrial function, demonstrating the utility of these events as metrics of age-induced metabolic change.

Mapping and quantitation of chimeric mitochondrial RNAs provides an accessible, orthogonal approach to DNA-based mutation assays, offers a potential method for identifying mitochondrial pathology in widely accessible datasets, and opens a new area of study in mitochondrial genetics and transcriptomics.

Introduction

Mitochondrial DNA (MtDNA) deletion mutation containing genomes contribute to disease and aging by causing mitochondrial dysfunction and cell death (Pak et al. 2003; Bua et al. 2006; Herbst et al. 2007, 2013; Cheema et al. 2015; Herbst et al. 2016, 2021b) Mitochondrial genetic diseases affect approximately 1 in 5,000 individuals (Ng and Turnbull 2016) with a smaller prevalence of ~1 in 100,000 for individuals with large-scale mtDNA deletion syndromes , mostly commonly Kearns-Sayre syndrome, Pearson syndrome and chronic progressive external ophthalmoplegia (Goldstein and Falk 1993). Disease progression in these individuals correlates with deletion size, deletion frequency within vulnerable tissues (e.g., brain, heart, skeletal muscle), and location of the deletion within the mitochondrial genome (Grady et al. 2014) . Exciting potential therapies are being developed and tested (Pirinen et al. 2020; Ng and Turnbull 2016; Shoop et al. 2023), but none yet are disease modifying or available for routine use. While mitochondrial genetic diseases often present in children and young adults and progress to premature mortality (García-Cazorla et al. 2005; Barends et al. 2016), mtDNA deletion mutations are also implicated in the aging process (Pak et al. 2003; Bua et al. 2006; Herbst et al. 2007, 2013, 2016, 2021b; Cheema et al. 2015). MtDNA deletion mutations are directly linked to multiple hallmarks of aging (Lopez-Otin et al. 2013) including mitochondrial dysfunction, genomic instability, and cellular senescence. The stochastic development of mtDNA deletion mutations has implications for the role of these mutations in somatic mosaicism (Campbell et al. 2015) and aging in post-mitotic tissues. The relative lack of genetic tools to correct or modulate mtDNA deletion mutations and a lack of understanding regarding the mechanisms of mutation accumulation is driving research in mitochondrial genetics to reveal new opportunities for treatment or prevention.

Current methods for mtDNA deletion mutation detection, mapping, and quantitation are exclusively DNA-based and few studies explored the effects on mitochondrial transcription (Lee

et al. 2020; Herbst et al. 2013). Work done prior to the widespread use of RNA sequencing observed chimeric mitochondrial transcripts associated with single mtDNA deletions in disease settings (Shoubridge et al. 1990; Savre-Train et al. 1992). These data suggested to us that unique, chimeric mitochondrial RNA transcripts would be formed by additional mtDNA deletion events (**Figure 1**). We hypothesized that these chimeric transcripts would be readily observable in RNA-seq data from diverse tissues and that the chimeric mitochondrial RNA frequency would be higher in individuals with known mitochondrial genetic diseases and in older individuals.

Chimeric RNAs in the nucleus are transcripts formed by gene fusion or intergenic splicing events and have been studied extensively in cancer and rare nuclear genetic diseases (Sun and Li 2022). Numerous unique chimeric RNAs have been found in these disorders and are often diagnostic biomarkers and therapeutic targets. For example, discovery of the *BCR-ABL* gene fusion (Rowley 1973) and corresponding chimeric transcript(Shtivelman et al. 1985) found in chronic myelogenous leukemia led to the development of BCR-ABL inhibitors that are the mainstay of treatment(Wong and Witte 2004). More recently, with the advent of RNA-seq and transcriptomic analyses, nuclear chimeric RNAs are being found in normal cells (Singh et al. 2020). By contrast, mitochondrial RNA transcripts are typically excluded *in silico* from RNA-seq analyses to allow better detection of lower abundance nuclear transcripts and chimeric mitochondrial RNAs have not previously been reported from such data. For RNA-seq data, numerous prediction algorithms have been devised to identify chimeric RNAs (Sun and Li 2022).

In this study, we aimed to deploy a well-established chimeric RNA prediction algorithm on publicly available and made-for-purpose RNA-seq datasets representing a range of conditions in which mtDNA deletion mutations are pathogenic. We further aimed to relate the chimeric

mtRNAs to age, disease state, gene expression, and, when possible, the underlying mtDNA deletion mutation frequency.

Results

Chimeric mtRNA transcripts are observed in patients with mitochondrial myopathy due to large-scale mtDNA deletion mutations.

To test whether chimeric mitochondrial transcripts corresponding to large deletion events can be found through application of established chimeric RNA detection software, we first analyzed RNA-seq of two human fibroblast lines previously documented to contain low levels of a single clonal genomic deletion (Majora et al. 2009) and a control neonatal dermal fibroblast cell line. Chimeric mtRNA transcripts were identified using STAR-fusion, using the most sensitive parameters to understand the breadth of events present. In control cells, no fusion was identified more than 7 times. By contrast, RNA sequencing of disease fibroblasts demonstrated abundant transcripts (2248 transcripts and 287 transcripts) with fusion sites corresponding to the breakpoints of the mitochondrial “common deletion.” Paired long read sequencing of the same samples demonstrated the presence of a single clonal deletion at this location (**Supplemental Figure 1**).

To test whether chimeric mitochondrial RNA transcripts exist in human transcriptomic data, we next analyzed skeletal muscle RNA-seq data obtained from adult patients with a nuclear mutation that is known to induce high levels of mtDNA deletion mutations (Pirinen et al. 2020). Some of these patients have mutations in the nuclear encoded gene for Twinkle, a mitochondrial DNA helicase. Disruption of Twinkle is known to induce a wide range of mtDNA deletion mutations that, following our hypothesis, would be prone to generating chimeric mitochondrial RNAs (Young and Copeland 2016). When our approach was used to quantify

chimeric mitochondrial RNA in this patient population, we observed chimeric mitochondrial RNAs in all samples. The two subjects with the Twinkle mutations had a 2.4-fold higher chimeric mitochondrial RNA frequency in their skeletal muscle as compared to healthy control subjects, consistent with the known increase in deletion mutations (**Table 1**). When specific fusion locations were considered, a 4.7-fold increase in chimeric mtRNA frequency was observed in major-arc fusions between cytochrome c oxidase subunits and cytochrome b in Twinkle subjects.

The same RNA-seq dataset also contained three individuals (Patients 1, 4, and 5), without Twinkle mutations, who have a single mtDNA deletion mutation species observed on long-extension PCR and agarose gel electrophoresis. The breakpoints of the mtDNA mutation in these subjects had not been sequenced (A. Suomalainen, personal communication). In Patient 1, we identified a single chimeric mitochondrial RNA (MT:9,189 – MT:14,990) at levels 10-20-fold higher than chimeras in the Twinkle subjects. This particular chimeric mitochondrial RNA was present in all biopsy samples from Patient 1. We did not find similarly expanded chimeras in the other two single deletion mutation subjects, possibly because their mtDNA deletion mutations do not create a detectable chimeric mitochondrial RNA, for example, an mt-mRNA fusion between two mt-tRNA genes. Principle component analysis of the chimeric mitochondrial RNAs from control subjects and those with large-scale deletions showed clustering of the control samples separate from the Twinkle subjects (**Figure 2A**). The most influential loadings in this analysis patients included *MT-ND6—MT-TE*, *MT-RNR2—MT-TL2* and *MT-TL1—MT-TL2*.

We next analyzed the characteristics of fusion events represented by at least 7 transcripts, the cut off established by our control data. The fusion sites from control subjects and single deletion subjects clustered in the major arc as seen in the aging rat skeletal muscle (**Figure 2B**,**C**). The fusion sites in the Twinkle subjects were more evenly distributed across the mitochondrial

genome. The distribution of fusion event locations was distinct from the distribution of mtDNA deletion mutations breakpoints reported in long read sequencing of normal aging muscle and in available online mtDNA mutation databases (**Figure 2C**). The fusion event sizes (i.e., the base pairs of DNA between the fusion sites) varied greatly between the subject groups. Control subjects had fusion sizes ranging from less than 1 kb to 13 kb while the fusion sizes for the single deletion subjects predominated in the location of the presumed deletion mutation. The fusion sizes in the Twinkle subjects ranged from 1kb to 11kb with most being 7-9kb (**Figure 2D**).

Chimeric mtRNA transcripts increase with age in human skeletal muscle and brain.

Our interest in mtDNA deletion mutations focuses on their contribution to morbidity and mortality associated with mammalian aging. Aging induces an exponential increase in deletion frequency across a number of tissues, but especially pronounced is the increase in human skeletal muscle (Herbst et al. 2021b, 2021a), where clonal expansion of somatically-derived deletions causes metabolic dysfunction and triggers of cell death (Pak et al. 2003; Bua et al. 2006; Cheema et al. 2015; Herbst et al. 2007, 2016, 2021b). To determine the relevance of chimeric mitochondrial RNA in this setting, we applied our analytical approach to an RNA-seq database from a study of skeletal muscle biopsies from individuals <30yo and >65yo (Kulkarni et al. 2020) . We observed a 1.8-fold increase in average chimeric mitochondrial RNA frequency in skeletal muscle from the older individuals (**Table 2**). mtDNA deletions characterized in single muscle fibers from aged humans often identify deletion breakpoints joining the mtDNA cytochrome c oxidase (COX) subunits 1, 2, or 3 and cytochrome b (Bua et al. 2006). When we examined a subset of Cox1,2,3 to cytb chimeric mitochondrial RNAs, we found a 2.4-fold increase with age (**Table 2**).

When considering events sequenced more than 7 times, The chimeric mitochondrial RNA fusion sites clustered primarily in the mitochondrial major arc in both young and older humans (**Figure 3A,B**). 1.8% of chimeric mitochondrial RNA reads had fusion sites corresponding to the large

direct repeats that flank the human common mtDNA deletion mutation (mtDNAdel4977, **Figure 3C**), the most commonly reported mtDNA deletion event in aging human tissue. 58% of chimeric mitochondrial reads found in muscle data had fusion sites within the mtDNA "contact zone", observed to enrich for mtDNA deletion breakpoints (Shamanskiy et al. 2023). The chimeric mitochondrial RNA sizes averaged ~6kb with a greater mean width in the older subjects (Welch two sample *t*-test, p=0.005)(**Figure 3D**).

We next examined an RNA-seq data set from a study of male human brain samples ages 52 to 93 taken at the time of death with or without a diagnosis of Parkinson's disease(Dumitriu et al. 2016) (Table 3). In the human brain samples, the chimeric mitochondrial RNA fusion sites are distributed throughout the mitochondrial genome, in contrast to human skeletal muscle where they are primarily localized to the major arc (**Figure 4A-B**). The size distribution of chimeric mitochondrial RNA fusion events also differs from skeletal muscle with smaller sized fusions than what we observed in skeletal muscle (median 3627 bp vs 5150 bp, Welch's *t*-test p<0.001) (**Figure 4C**). Many fusion sites overlap with deletion mutations observed with nanopore sequencing of substantia nigra samples (**Figure 4B**) with the exception of breakpoints in the minor arc that are more numerous in the chimeric mitochondrial RNA data. 64% of fusion sites found in brain data are within the mtDNA "contact zone".

Chimeric mitochondrial transcripts are induced with induction of mtDNA deletions

It is established that treatment with guanidinopropionic acid (GPA), a drug known to worsen muscle aging, increases the frequency of age-related mtDNA deletions in skeletal muscle (Herbst et al. 2013). To test the hypothesis that chimeric mitochondrial transcripts are products of mtDNA deletions, we analyzed RNA-seq data obtained from skeletal muscle from rats aged 18mo and 34mo and rats we treated with guanidinopropionic acid (GPA). When STAR-Fusion was used to identify chimeric transcripts in these data (PRJNA793055), we observed numerous

chimeric transcripts with both breakpoints on the mitochondrial genome. This analysis demonstrated an 4.4 -fold age-induced increase in chimeric mtRNA frequency in control mice with a further 2.2-fold increase with GPA treatment in aged mice (**Table 4**), corresponding closely to the previously reported magnitude of mtDNA deletion mutation change measured by DNA-based methods (Herbst et al. 2013).

In the same rat skeletal muscle samples used to generate the RNA-seq data, we measured mtDNA deletion frequency using a DNA-based digital PCR approach that targets the large deletion events that occur in aging (Herbst et al. 2017, 2022). We found a strong correlation between mtDNA deletion mutation frequency and chimeric mitochondrial RNA frequency (linear regression, R^2=0.86, p<0.0001, **Figure 5**).

Principle component analysis of chimeric mitochondrial RNA event frequency (**Figure 6A**) demonstrates increased variability of chimeric mtRNA events in 34-month-old rats with and without GPA treatment. The most influential loadings in the PCA include chimeric mitochondrial RNAs involving *MT-CO1* or *MT-CO2* fused with *MT-CYB*. The fusion sites are primarily within the major arc of the mitochondrial genome (**Figure 6B,C**). For 90% of fusion events, the distance between break points is between 1kb to 10kb with 66% of events being in the range of 6-9kb (**Figure 6D**).

Chimeric mtRNA frequency predicts expression of genes involved in mitochondrial dysfunction in an aging rat model.

Having observed chimeric mitochondrial RNA transcripts in multiple models associated with mitochondrial pathology, we next asked whether the presence of the chimeric transcripts would predict specific gene expression changes involved in disease pathogenesis. For an initial test,

we evaluated gene expression data from aging rat skeletal muscle from which we observed an increase in chimeric mitochondrial RNA associated with aging (PRJNA793055). When we compared each samples' sum total of chimeric mitochondrial RNA transcripts for all rat samples to the global transcriptome, we identified 6104 differentially expressed genes associated with the chimeric mitochondrial RNA frequency. Gene set enrichment analysis of these genes identified significant enrichment for downregulation of multiple pathways involving mitochondrial function, including electron transport chain and tricarboxylic acid cycle as well as mitochondrial translation, while upregulated pathways focused on cell communication and immunoregulatory pathways(**Figure 7A**). When differential gene expression analysis was adjusted for the age and GPA treatment status of the animal, these findings persisted (**Supplementary Figure 2**). While mtDNA encoded genes are seen to be differentially expressed (Figure 7B, 7D), the vast majority (99.8%) of differentially expressed genes identified are encoded in the nuclear genome (Figure 7C, 7E), suggesting that chimeric mtRNA frequency predicts transcriptomic changes outside the mitochondria.

Discussion

We report a class of previously undocumented chimeric mitochondrial RNAs, which increase in multiple conditions associated with mitochondrial pathology and predict mtDNA deletion frequency and altered transcription of genes essential for mitochondrial function. The new class of chimeric mitochondrial RNAs we have discovered provides an exciting window into mitochondrial genomics. The distribution of chimeric mtRNA breakpoints closely matches the breakpoints identified previously by DNA sequencing approaches, while the frequency of chimeric mtRNAs predicts mtDNA deletion mutation frequency measured by PCR-based approaches. These relationships strongly suggest that the chimeric mtRNAs are transcriptional products of the mtDNA deletion events.

Many aspects of the chimeric mitochondrial RNAs strengthen what is known about mtDNA deletion mutations from DNA-based assays such as Southern blot(Corral-Debrinski et al. 1992), whole genome PCR(Tengan and Moraes 1996), high throughput sequencing (Lujan et al. 2020), digital PCR (Herbst et al. 2017) and long read mtDNA sequencing(Vandiver et al. 2022). The chimeric mitochondrial RNAs predominate in the mitochondrial major arc, span large portions of the mitochondrial genome (i.e., 6-8kb), and increase with age in human and rat skeletal muscle and human brain. In the human muscle and brain RNA-seq databases, we observed chimeric mitochondrial RNAs that correspond to the frequently reported mitochondrial “common” or “4977” deletion. The chimeric mitochondrial RNA frequency in the rat skeletal muscle correlates strongly with a DNA-based digital PCR assay, an orthogonal approach, supporting the validity of both approaches and the observed age-induced increases. This newly identified ability to identify pathogenic mtDNA variants in patients from available RNA-seq data is in line with evolving strategies to leverage transcriptomic data to better identify and understand undiagnosed and rare diseases (Lee et al. 2020) and greatly extends the ability to identify these events in difficult to obtain specimens. Using an RNA-based approach extends prior DNA-based findings by indicating a potential post-genomic role for mtDNA deletions that may uncover new mechanisms or points for intervention.

An extensive literature on chimeric RNAs from the nuclear genome in cancer suggests many possible roles for chimeric mitochondrial RNAs(Sun and Li 2022). Chimeric mitochondrial RNAs could interfere with mtDNA replication, transcription, and translation. MtDNA deletion mutations reach very high levels within individual cells(Pak et al. 2003; Bua et al. 2006; Cheema et al. 2015; Herbst et al. 2007, 2013), and in such cells, the chimeric mtRNA would be expected to also accumulate to very high levels where it could deplete nucleotide pools, disrupt mt-mRNA processing, or impair mt-mRNA degradation/turnover pathways. Effects of chimeric mitochondrial RNAs on protein translation could include the sequestration of ribosomes and

tRNAs on these aberrant transcripts. Chimeric mitochondrial RNAs could also bind other mitochondrial RNA transcripts or even leak into the cytoplasm(Kim et al. 2017) and interfere with nuclear transcriptional and translation machinery or trigger inflammatory responses(Hooftman et al. 2023; Zecchini et al. 2023) . For example, when we previously examined the GPA and aging gene expression data, we found limited correlates(Herbst et al. 2023) . By using the embedded chimeric mtRNA frequency, we observed additional correlations that point to disruption of mitochondrial function and metabolism and identified upregulation of inflammatory responses.

Chimeric mitochondrial RNA analyses are complimentary to the extensive literature on mtDNA deletions using DNA-based approaches. The age-induced increases in chimeric mitochondrial RNAs add rigor to DNA-based findings that mtDNA deletions increase with age in mammalian skeletal muscle and brain(Taylor et al. 2014; Herbst et al. 2017; Vandiver et al. 2022, 2023). In addition to verifying prior findings, the use of chimeric mitochondrial RNA to query mitochondrial genome rearrangements offers several advantages to DNA-based methods. The use of RNA transcripts to uncover mutation events leverages the transcriptional amplification of these genomic events that may improve detection of rare events. Finally, it would seem reasonable to exploit all of the biological inferences available from an RNA-seq experiment, especially when the signals are interpretable.

The potential to use RNA sequencing data for identification of mtDNA structural variation is particularly robust due to the abundance of mitochondrial reads in standard RNA sequencing data sets. Mitochondrial transcripts make up 30-40% of reads in standard RNA-seq libraries (Mercer et al. 2011; Yang et al. 2014) and as high as 70-80% in single-cell RNA sequencing experiments(Yang et al. 2014). RNA-seq experiments are often designed to maximize nuclear gene expression data and avoid the sequencing of RNA species such as mitochondrial

transcripts that are considered non-relevant, unimportant or redundant. In single cell sequencing data, a preponderance of mitochondrial transcripts is used to exclude apoptotic cells(Yang et al. 2014). While it may be acceptable to discard or ignore some data that truly has no relevance or possible interpretation, the detection of chimeric mtRNAs in RNA-seq datasets cautions against this bias and demonstrates the importance of sharing and depositing the raw data.

We focused our efforts on interrogating RNA-seq databases that extended our previous studies on aging skeletal muscle. The identification and quantitation of chimeric mtRNAs should be generalizable across organisms, making it potentially useful in comparative biology studies of aging for any organism for which there are annotated mitochondrial genomes or methods developed to search for chimeric mitochondrial RNAs in *de novo* transcriptomic data. Single-cell or single-nucleus RNA-seq characterization of chimeric mitochondrial RNAs could be used to define tissue and cell distributions of chimeric mitochondrial RNAs and the underlying mtDNA deletion mutations and these could be combined with nuclear gene expression data to identify cellular impacts.

While we demonstrate a strong correspondence between chimeric mtRNA and mtDNA mutations, there remain many open questions as to the specific relationship. We observe a strong linear relationship between frequency of chimeric RNAs and mtDNA mutations in our rat model, however, this is not sufficient to suggest that the frequency of chimeric RNAs is a direct report of the frequency of mtDNA deletions. Transcript frequency is a reflection not only of DNA abundance but transcription, transcript maturation and transcript stability. It is possible that the presence of mtDNA deletions and chimeric transcripts could interfere with any of these processes. As such, inferences can not yet be made as to how chimeric mtRNA frequency relates to well-defined characteristics of mtDNA deletions, such as a phenotypic threshold or cellular heteroplasmy. Further, our analysis indicates that a significant level of chimeric mtRNA

are detected even in younger tissues, which are not frequently reported to contain abundant mtDNA deletions. One potential explanation is that low levels of mtDNA deletions are present in younger tissues and are easier to detect in this setting due to transcriptional amplification. Alternatively, there may be background chimeric mtRNA reads present in sequencing data that is not reflective of underlying mtDNA deletions. Studies of other classes of mitochondrial RNAs have called into question the potential for those mitochondrial RNAs to derive from nuclear embedded mitochondrial sequences (NUMTs)(Pozzi and Dowling 2019). The lack of chimeric RNAs observed in a control cell line argues against against NUMTs as a primary source of chimeric mtRNAs. In addition, NUMT origination cannot be accurately excluded using short read sequencing (Albayrak et al. 2016). These questions and more will be important to address as these transcripts are examined in more data sets.

The location of chimeric fusion locations observed is intriguing. While age-associated mtDNA deletions are most frequently observed to occur within the mitochondrial major arc, recent studies have demonstrated deletion events involving other areas of the genome (Bua et al. 2006; Lujan et al. 2020; Vandiver et al. 2023). The tissue specificity of this finding remains an open question. In this initial analysis of chimeric mitochondrial transcripts, fusion sites found in muscle data were primarily localized within the major arc, while brain data showed a broader distribution, suggesting potential differences in the distribution of mtDNA deletions between tissues. The identification of this unstudied class of mitochondrial transcripts will facilitate further studies to clarify these differences and gain insight into tissue specificity of mitochondrial genomic variation.

Limitations of our initial chimeric mtRNA analyses suggest next steps in characterizing chimeric mitochondrial RNAs. Addressing these limitations should include additional definition of the RNA-seq characteristics that affect chimeric mitochondrial RNA detection and quantitation.

Factors such as RNA isolation method, mRNA enrichment, library preparation and read depth need to be validated to enable sensitive detection of chimeric mitochondrial RNA as has been done for general RNA-seq approaches (Conesa et al. 2016). Methods that preserve biological signals arising from the mitochondria (rather than the standard approaches that discard them) will increase the sensitivity of the RNA-seq approach for detecting chimeric mtRNAs and improve chimera quantitation.

Alternative explanations for our findings could include age-related disruption of mitochondrial transcription leading to aberrant RNA splicing events. However, this would not be consistent with the occurrence of the chimeric mitochondrial RNA corresponding to the well-characterized human “common” deletion. We adapted computational methods that were designed for detection of chimeric RNAs in cancer that may need to be further customized for chimeric mitochondrial RNAs. Further, we utilized short read sequencing data, which has limitations for study of fusion events due to mapping challenges, thus long read RNA-seq may be particularly well suited to find these previously undocumented RNA mutants(Workman et al. 2019) .

We have begun identifying, mapping, and quantitating a new class of mitochondrial transcripts that appear to be the direct result of age-induced and disease-associated mtDNA deletion mutations. This approach provides orthogonal validation to prior studies of mtDNA deletion mutations and extends the ability to identify such events across a broader range of samples. Further study of these chimeric mitochondrial RNAs will expand our understanding of the underlying mutational events and open new avenues for disease prevention or therapeutic intervention.

Methods

RNA-seq databases

Publicly available RNA-seq data were downloaded from the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/), accessions GSE129811, PRJNA793055, GSE157585, GSE68719 were used.

Analysis of RNA sequencing data for mitochondrial fusion transcripts

Gene fusion identification:

STAR-Fusion version 1.10.0(Haas et al. 2019) was used to identify candidate fusion transcripts within RNA-Seq datasets. For the rat study, a "CTAT genome lib" reference package was prepared for STAR-Fusion using sequence and feature information for assembly Rnor6.0 downloaded from Ensembl version 104(Harrison et al. 2024). For the human studies, reference files for assembly GRCh38 were used, also from Ensembl version 104. For both rat and human the Ensembl gene transfer format (GTF) files were manually modified prior to running the 'CTAT genome lib' building script, in order to convey that the MT-ATP8 and MT-ATP6 genes are encoded within a single overlapping transcript that, although detected by STAR-Fusion, does not represent a chimeric mitochondrial RNA. Similarly, GTF information for the MT-ND4l and Mt-ND4 genes was updated to reflect that they are normally expressed as a single transcript. Single-end reads were aligned to the rat or human reference genomes using STAR v2.7.11a(Dobin et al. 2013) and STAR-Fusion v1.13.0 was then run on the STAR output to identify fusion candidates. Default STAR-Fusion parameters were used with the exception of the following settings: --min_FFPM 0 was used to prevent filtering of fusion transcripts based on abundance and --no_remove_dups was used to prevent removal of duplicate reads. The --alignIntronMax 100 option was used with STAR, which we have found to improve fusion detection for some deletions.

Comparing gene fusions among samples: Further analysis of fusion locations was done within the R environment(R Core Team 2023)The STAR-Fusion output files describing the candidate

fusions (star-fusion.fusion_candidates.preliminary files) were used in downstream analyses. R scripts were used to parse the files and to enumerate mitochondrial gene fusions within each sample. Specifically, for each observed fusion type (based on genes involved and ignoring the precise boundaries of the fusion) the total number of supporting reads was calculated, using values extracted from the JunctionReadCount column. Subsequently for each dataset a table termed "raw counts" was generated, consisting of samples (rows) and fusion types (columns) with cells containing the summation of the JunctionReadCount values. A second table, termed "FFPM" for "fusion fragments per million total RNA-seq fragments" was generated from the first table by dividing each raw count by the total number of sequenced fragments (in millions) in the corresponding sample. Sample meta data was programmatically added to each table as additional columns, to facilitate further analyses. PCA plots were produced from FFPM tables using the ggfortify R package(Tang et al. 2016). Plots visualizing or comparing fusion site locations were generated from the star-fusion.fusion_candidates.preliminary files, using the reported breakpoint locations. Kernel density of break points was calculated and plotted using base R *density* function. Histograms were plotted using base R *hist* function. For plots examining gene fusion location and fusion sizes, events detected in more than 7 transcripts were considered.

Gene set enrichment analysis: Publicly available fastqs were quasi-mapped and quantified to the GENCODE *Homo Sapiens* GRCh38 all cDNA reference transcriptome using Salmon (v1.9.0)(Patro et al. 2017) in Python 3.9.15. Differential gene expression analysis and gene set enrichment analysis were conducted in the R environment (R v 4.2). FFPM counts were extracted and differential gene expression analysis was performed using DESeq2 (v1.38.1)(Love et al. 2014) . For differential gene identification, a sum of FFPM for all mitochondrial chimeric transcripts was calculated. Regression was then conducted using DESeq2 using both a model considering only this sum, and a model considering this sum and

adjusting for age and GPA treatment status. Significance testing was performed using Wald test, resulting P-values were adjusted for multiple testing using package default. Gene set enrichment analysis was performed on resultant differentially expressed genes using the FGSEA package (v1.24.0)(Korotkevich et al. 2016) and reactome pathywas from reactome.db. Pathways with NES magnitude greater than 2 and at least 10 included transcripts were considered.

All analysis code is available online at https://github.com/paulstothard/chimeric-mitochondrial-RNA-analysis.

Fibroblast RNA preparation

Dermal fibroblasts containing the mitochondrial common deletion were obtained as a gift from Dr. Jean Kruttman (Majora et al. 2009). Control human dermal Fibroblasts were obtained from Fisher Scientific. DNA was extracted using Qiagen DNeasy according to manufacturer's instructions. Control dermal fibroblasts were purchased from Thermo Fisher Scientific (Catalog #004-5C). RNA was extracted with TRIzol (Thermo Fisher Scientific) followed by RNeasy (Qiagen) clean-up, in accordance with manufacturer's instructions. 150 bp, paired-end sequencing was performed on Illumina NextSeq 500 through the UCLA Technology Center for Genomics and Bioinformatics.

Data access

The fibroblast sequencing data generated in this study have been submitted to the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/) under accession number PRJNA1164295.

Analytical code is available in Supplementary materials and at https://github.com/paulstothard/chimeric-mitochondrial-RNA-analysis.

Competing interest statement

The authors declare no competing interests.

Acknowledgements

This material is the result of work supported with resources and the use of facilities at the Greater Los Angeles Veterans Healthcare System. Scientific illustration assistance was provided by Kate Baldwin, PhD (K8Baldwin.com).

Authors' contributions

AH, PS, and JW conceived and designed the experiments. PS developed the algorithms, AV and PS and conducted the bioinformatic analyses. All authors interpreted the data and wrote and approved the manuscript.

Funding

The authors were supported by NIH R01AG055518 (JW), NIH R01AG069924 (JW), NIH K02AG059847 (JW), the Dermatology Foundation (ARV) and the Melanoma Research Alliance (ARV). This research was enabled in part by support provided by the Digital Research Alliance of Canada (alliancecan.ca).

Figure Legends

Figure 1. Diagram outlining the process of mitochondrial transcript processing in wild-type (top) versus mtDNA deletion mutation (bottom) genomes. Different colors denote different electron transport chain complexes or the mitochondrial ribosomal RNAs.

Figure 2. Chimeric mitochondrial RNAs in patients with mitochondrial genetic diseases. A, Principal component analysis of human chimeric mitochondrial RNA data. B, Distribution of chimeric mitochondrial RNAs detected more than 7 times in patients with single large deletions, patients with heterogeneous mtDNA deletions due to nuclear gene Twinkle mutations, and control subjects. C, Location of fusion sites across the human mitochondrial genome and comparison to known deletion breakpoint distributions. Vertical yellow bars denote the location of the human 4977 or "common" deletion. D, Histogram of fusion event sizes of chimeric

mitochondrial RNAs from patients with single large deletions, Twinkle patients, and control subjects.

Figure 3. Chimeric mtRNAs in aging human skeletal muscle. A, PCA plot of human skeletal muscle chimeric mtRNA data. B, Distribution of chimeric mtRNA breakpoints in 14 younger (left) and 14 older subjects (right) for fusions detected more than 7 times. C, location of chimeric ends across the human mitochondrial genome and comparison to known deletion breakpoint distributions. Vertical yellow bars denote the location of the human 4977 or common deletion. D, Histogram of fusion event sizes of chimeric mtRNAs from <30 and >65yo subjects.

Figure 4. Chimeric mitochondrial RNAs in aging human brain. A, chimeric mitochondrial RNA frequency versus age in control subjects or those diagnosed with Parkinson's disease. B, Distribution of chimeric mitochondrial RNA fusion sites for fusions detected more than 7 times in human brain. C, location of fusion sites across the human mitochondrial genome and comparison to known mtDNA deletion mutation breakpoint distributions in human brain. D, Histogram of fusion event sizes in human brain.

Figure 5. MtDNA deletion mutation frequency predicts chimeric mitochondrial RNA frequency in rat skeletal muscle.

Figure 6. Mapping of age- and drug-induced mitochondrial RNA fusion events in rats. A, PCA plot of rat chimeric mtRNA data. B, Distribution of chimeric mitochondrial RNA breakpoints for representative sample from each category mapped onto a diagram of the wild-type mitochondrial genome. C, location of chimeric ends across the rat mitochondrial genome. Vertical yellow bars denote the location of the longest direct repeats (16 base pairs) in the rat mtDNA reference genome (located at positions 8,102 and 12,936). D, Histogram of fusion event sizes for the chimeric mitochondrial RNAs.

Figure 7. Chimeric mitochondrial RNA frequency predicts altered transcription of cellular pathways. A, Gene set enrichment analysis of differentially expressed genes associated with

chimeric mitochondrial RNA frequency in rat skeletal muscle aging. Negative correlations of FFPM to gene expression for nuclear-encoded *SDHB* (B) and *COX5b* (D) and mitochondrially-encoded *MT-COI* (C) and *MT-CYB* (E).

Table 1: Chimeric mtRNA frequency in patients with mitochondrial genetic disease.

Database Experimental group (n)	Average base pairs sequenced	Chimeric mtRNA FFPM	Cox1,2,3-cytb chimeric mtRNA FFPM
Human twinkle muscle (GSE129811)	1.3 ± 0.001e[9]		
Control subjects (8)		44.6 ± 7.47	1.42 ± 0.21
Twinkle subjects (2)		108 ± 33.3	6.66 ± 2.89

n = number of human subjects.

Data shown are mean ± SEM.

Table 2: Chimeric mtRNA frequency in human aging muscle samples

Database Experimental group (n)	Average base pairs sequenced	Chimeric mtRNA FFPM	Cox1,2,3-cytb chimeric mtRNA FFPM
Human aging muscle (GSE157585)	1.4 ± 0.03e[10]		
20-30yo (21)		40.4 ± 4.02*	1.16 ± 0.12**
65-91yo (50)		71.5 ± 5.11*	2.82 ± 0.24**

n = number of human subjects.

Data shown are mean ± SEM.

*P value=6.4E-6 using Welch's Two Sample *t*-test

**P value=7.3E-9 using Welch's Two Sample *t*-test

Table 3: Chimeric mtRNA frequency in human aging brain samples

Database Experimental group (n)	Average base pairs sequenced	Chimeric mtRNA FFPM	Cox1,2,3-cytb chimeric mtRNA FFPM
Human aging brain (GSE68719) 46-97yo (73)	8.5 ± 0.3e9	22.99 ± 1.69	0.93 ± 0.077

n = number of human subjects.

Data shown are mean ± SEM.

528

Table 4: Chimeric mtRNA frequency in rat muscle samples

Database Experimental group (n)	Average base pairs sequenced	Chimeric mtRNA FFPM	Cox1,2,3-cytb chimeric mtRNA FFPM
Rat aging muscle (PRJNA793055)	2.6 ± 0.17e^{9}		
18mo control (5)		0.44 ± 0.12^{a}	0.050 ± 0.035^{a}
18mo GPA treated (5)		0.55 ± 0.10^{a}	0.11 ± 0.026^{a}
34mo control (6)		1.93 ± 0.27^{b}	0.94 ± 0.13^{a}
34mo GPA treated (6)		4.20± 0.80^{c}	2.32 ± 0.45^{b}

n = number of rats.

Data shown are mean ± SEM.

Different superscript letters denote significance at $p < 0.05$ within the column using pairwise t-test.

529

530 **References**

531

532 Albayrak L, Khanipov K, Pimenova M, Golovko G, Rojas M, Pavlidis I, Chumakov S, Aguilar G,
533 Chávez A, Widger WR, et al. 2016. The ability of human nuclear DNA to cause false
534 positive low-abundance heteroplasmy calls varies across the mitochondrial genome.
535 *BMC Genomics* **17**: 1017.

536 Barends M, Verschuren L, Morava E, Nesbitt V, Turnbull D, McFarland R. 2016. Causes of
537 death in adults with mitochondrial disease. *JIMD Rep* **26**: 103–113.

538 Bua E, Johnson J, Herbst A, Delong B, McKenzie D, Salamat S, Aiken JM. 2006. Mitochondrial
539 DNA-deletion mutations accumulate intracellularly to detrimental levels in aged human
540 skeletal muscle fibers. *Am J Hum Genet* **79**: 469–480.

541 Campbell IM, Shaw CA, Stankiewicz P, Lupski JR. 2015. Somatic mosaicism: implications for
542 disease and transmission genetics. *Trends Genet* **31**: 382–392.

543 Cheema N, Herbst A, McKenzie D, Aiken JM. 2015. Apoptosis and necrosis mediate skeletal
544 muscle fiber loss in age-induced mitochondrial enzymatic abnormalities. *Aging Cell* **14**:
545 1085–1093.

546 Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, Szcześniak
547 MW, Gaffney DJ, Elo LL, Zhang X, et al. 2016. A survey of best practices for RNA-seq
548 data analysis. *Genome Biol* **17**: 13.

549 Corral-Debrinski M, Horton T, Lott MT, Shoffner JM, Beal MF, Wallace DC. 1992. Mitochondrial
550 DNA deletions in human brain: regional variability and increase with advanced age. *Nat*
551 *Genet* **2**: 324–329.

552 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras
553 TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.

554 Dumitriu A, Golji J, Labadorf AT, Gao B, Beach TG, Myers RH, Longo KA, Latourelle JC. 2016.
555 Integrative analyses of proteomics and RNA transcriptomics implicate mitochondrial
556 processes, protein folding pathways and GWAS loci in Parkinson disease. *BMC Med*
557 *Genomics* **9**: 5.

558 García-Cazorla A, De Lonlay P, Nassogne MC, Rustin P, Touati G, Saudubray JM. 2005. Long-
559 term follow-up of neonatal mitochondrial cytopathies: a study of 57 patients. *Pediatrics*
560 **116**: 1170–1177.

561 Goldstein A, Falk MJ. 1993. Single large-scale mitochondrial DNA deletion syndromes. In
562 *GeneReviews(®)*, University of Washington, Seattle, Seattle (WA).

563 Grady JP, Campbell G, Ratnaike T, Blakely EL, Falkous G, Nesbitt V, Schaefer AM, McNally
564 RJ, Gorman GS, Taylor RW, et al. 2014. Disease progression in patients with single,
565 large-scale mitochondrial DNA deletions. *Brain* **137**: 323–334.

566 Haas BJ, Dobin A, Li B, Stransky N, Pochet N, Regev A. 2019. Accuracy assessment of fusion
567 transcript detection via read-mapping and de novo fusion transcript assembly-based
568 methods. *Genome Biol* **20**: 213.

569 Harrison PW, Amode MR, Austine-Orimoloye O, Azov AG, Barba M, Barnes I, Becker A,
570 Bennett R, Berry A, Bhai J, et al. 2024. Ensembl 2024. *Nucleic Acids Res* **52**: D891–
571 D899.

572 Herbst A, Aiken JM, Kim C, Gushue D, McKenzie D, Moore TM, Zhou J, Hoang AN, Choi S,
573 Wanagat J. 2023. Age- and time-dependent mitochondrial genotoxic and myopathic
574 effects of beta-guanidinopropionic acid, a creatine analog, on rodent skeletal muscles.
575 *GeroScience* **45**: 555–567.

576 Herbst A, Choi S, Hoang AN, Kim C, Martinez Moreno D, McKenzie D, Aiken JM, Wanagat J.
577 2022. Remdesivir does not affect mitochondrial DNA copy number or deletion mutation
578 frequency in aged male rats: A short report. *PLoS One* **17**: e0271850.

579 Herbst A, Johnson CJ, Hynes K, McKenzie D, Aiken JM. 2013. Mitochondrial biogenesis drives
580 a vicious cycle of metabolic insufficiency and mitochondrial DNA deletion mutation
581 accumulation in aged rat skeletal muscle fibers. *PLoS One* **8**: e59006.

582 Herbst A, Lee CC, Vandiver AR, Aiken JM, McKenzie D, Hoang A, Allison D, Liu N, Wanagat J.
583 2021a. Mitochondrial DNA deletion mutations increase exponentially with age in human
584 skeletal muscle. *Aging Clin Exp Res* **33**: 1811–1820.

585 Herbst A, Pak JW, McKenzie D, Bua E, Bassiouni M, Aiken JM. 2007. Accumulation of
586 mitochondrial DNA deletion mutations in aged muscle fibers: evidence for a causal role
587 in muscle fiber loss. *J Gerontol A Biol Sci Med Sci* **62**: 235–245.

588 Herbst A, Prior SJ, Lee CC, Aiken JM, McKenzie D, Hoang A, Liu N, Chen X, Xun P, Allison DB,
589 et al. 2021b. Skeletal muscle mitochondrial DNA copy number and mitochondrial DNA
590 deletion mutation frequency as predictors of physical performance in older men and
591 women. *Geroscience* **43**: 1253–1264.

592 Herbst A, Wanagat J, Cheema N, Widjaja K, McKenzie D, Aiken JM. 2016. Latent mitochondrial
593 DNA deletion mutations drive muscle fiber loss at old age. *Aging Cell* **15**: 1132–1139.

594 Herbst A, Widjaja K, Nguy B, Lushaj EB, Moore TM, Hevener AL, McKenzie D, Aiken JM,
595 Wanagat J. 2017. Digital PCR Quantitation of Muscle Mitochondrial DNA: Age, Fiber
596 Type, and Mutation-Induced Changes. *J Gerontol A Biol Sci Med Sci* **72**: 1327–1333.

597 Hooftman A, Peace CG, Ryan DG, Day EA, Yang M, McGettrick AF, Yin M, Montano EN, Huo
598 L, Toller-Kawahisa JE, et al. 2023. Macrophage fumarate hydratase restrains mtRNA-
599 mediated interferon production. *Nature* **615**: 490–498.

600 Kim KM, Noh JH, Abdelmohsen K, Gorospe M. 2017. Mitochondrial noncoding RNA transport.
601 *BMB Rep* **50**: 164–174.

602 Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. 2016. Fast gene
603 set enrichment analysis. *bioRxiv*. http://dx.doi.org/10.1101/060012.

604 Kulkarni AS, Peck BD, Walton RG, Kern PA, Mar JC, Windham ST, Bamman MM, Barzilai N,
605 Peterson CA. 2020. Metformin alters skeletal muscle transcriptome adaptations to
606 resistance training in older adults. *Aging (Albany NY)* **12**: 19852–19866.

607 Lee H, Huang AY, Wang L-K, Yoon AJ, Renteria G, Eskin A, Signer RH, Dorrani N, Nieves-
608 Rodriguez S, Wan J, et al. 2020. Diagnostic utility of transcriptome sequencing for rare
609 Mendelian diseases. *Genet Med* **22**: 490–499.

610 Lopez-Otin C, Blasco MA, Partridge L, Serrano M, Kroemer G. 2013. The hallmarks of aging.
611 *Cell* **153**: 1194–1217.

612 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for
613 RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Lujan SA, Longley MJ, Humble MH, Lavender CA, Burkholder A, Blakely EL, Alston CL, Gorman GS, Turnbull DM, McFarland R, et al. 2020. Ultrasensitive deletion detection links mitochondrial DNA replication, disease, and aging. *Genome Biol* **21**: 248.

Majora M, Wittkampf T, Schuermann B, Schneider M, Franke S, Grether-Beck S, Wilichowski E, Bernerd F, Schroeder P, Krutmann J. 2009. Functional consequences of mitochondrial DNA deletions in human skin fibroblasts: increased contractile strength in collagen lattices is due to oxidative stress-induced lysyl oxidase activity. *Am J Pathol* **175**: 1019–1029.

Mercer TR, Neph S, Dinger ME, Crawford J, Smith MA, Shearwood A-MJ, Haugen E, Bracken CP, Rackham O, Stamatoyannopoulos JA, et al. 2011. The human mitochondrial transcriptome. *Cell* **146**: 645–658.

Ng YS, Turnbull DM. 2016. Mitochondrial disease: genetics and management. *J Neurol* **263**: 179–191.

Pak JW, Herbst A, Bua E, Gokey N, McKenzie D, Aiken JM. 2003. Mitochondrial DNA mutations as a fundamental mechanism in physiological declines associated with aging. *Aging Cell* **2**: 1–7.

Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**: 417–419.

Pirinen E, Auranen M, Khan NA, Brilhante V, Urho N, Pessia A, Hakkarainen A, Ulla Heinonen JK, Schmidt MS, Haimilahti K, et al. 2020. Niacin cures systemic NAD+ deficiency and improves muscle performance in adult-onset mitochondrial myopathy. *Cell Metab* **32**: 144.

Pozzi A, Dowling DK. 2019. The genomic origins of small mitochondrial RNAs: Are they transcribed by the mitochondrial DNA or by mitochondrial pseudogenes within the nucleus (NUMTs)? *Genome Biol Evol* **11**: 1883–1896.

R Core Team. 2023. A Language and Environment for Statistical ## Computing_. R Foundation for Statistical Computing. *R Foundation for Statistical Computing*. http://dx.doi.org/https://www.R-project.org/.

Rowley JD. 1973. Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature* **243**: 290–293.

Savre-Train I, Piatyszek MA, Shay JW. 1992. Transcription of deleted mitochondrial DNA in human colon adenocarcinoma cells. *Hum Mol Genet* **1**: 203–204.

Shamanskiy V, Mikhailova AA, Tretiakov EO, Ushakova K, Mikhailova AG, Oreshkov S, Knorre DA, Ree N, Overdevest JB, Lukowski SW, et al. 2023. Secondary structure of the human mitochondrial genome affects formation of deletions. *BMC Biol* **21**: 103.

Shoop WK, Lape J, Trum M, Powell A, Sevigny E, Mischler A, Bacman SR, Fontanesi F, Smith J, Jantz D, et al. 2023. Efficient elimination of MELAS-associated m.3243G mutant mitochondrial DNA by an engineered mitoARCUS nuclease. *Nat Metab* **5**: 2169–2183.

Shoubridge EA, Karpati G, Hastings KE. 1990. Deletion mutants are functionally dominant over wild-type mitochondrial genomes in skeletal muscle fiber segments in mitochondrial disease. *Cell* **62**: 43–49.

Shtivelman E, Lifshitz B, Gale RP, Canaani E. 1985. Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. *Nature* **315**: 550–554.

Singh S, Qin F, Kumar S, Elfman J, Lin E, Pham L-P, Yang A, Li H. 2020. The landscape of chimeric RNAs in non-diseased tissues and cells. *Nucleic Acids Res* **48**: 1764–1778.

Sun Y, Li H. 2022. Chimeric RNAs discovered by RNA sequencing and their roles in cancer and rare genetic diseases. *Genes (Basel)* **13**: 741.

Tang Y, Horikoshi M, Li W. 2016. Ggfortify: Unified interface to visualize statistical results of popular R packages. *R J* **8**: 474.

Taylor SD, Ericson NG, Burton JN, Prolla TA, Silber JR, Shendure J, Bielas JH. 2014. Targeted enrichment and high-resolution digital profiling of mitochondrial DNA deletions in human brain. *Aging Cell* **13**: 29–38.

Tengan CH, Moraes CT. 1996. Detection and analysis of mitochondrial DNA deletions by whole genome PCR. *Biochem Mol Med* **58**: 130–134.

Vandiver AR, Hoang AN, Herbst A, Lee CC, Aiken JM, McKenzie D, Teitell MA, Timp W, Wanagat J. 2023. Nanopore sequencing identifies a higher frequency and expanded spectrum of mitochondrial DNA deletion mutations in human aging. *Aging Cell* **22**: e13842.

Vandiver AR, Pielstick B, Gilpatrick T, Hoang AN, Vernon HJ, Wanagat J, Timp W. 2022. Long read mitochondrial genome sequencing using Cas9-guided adaptor ligation. *Mitochondrion* **65**: 176–183.

Wong S, Witte ON. 2004. The BCR-ABL story: bench to bedside and back. *Annu Rev Immunol* **22**: 247–306.

Workman RE, Tang AD, Tang PS, Jain M, Tyson JR, Razaghi R, Zuzarte PC, Gilpatrick T, Payne A, Quick J, et al. 2019. Nanopore native RNA sequencing of a human poly(A) transcriptome. *Nat Methods* **16**: 1297–1305.

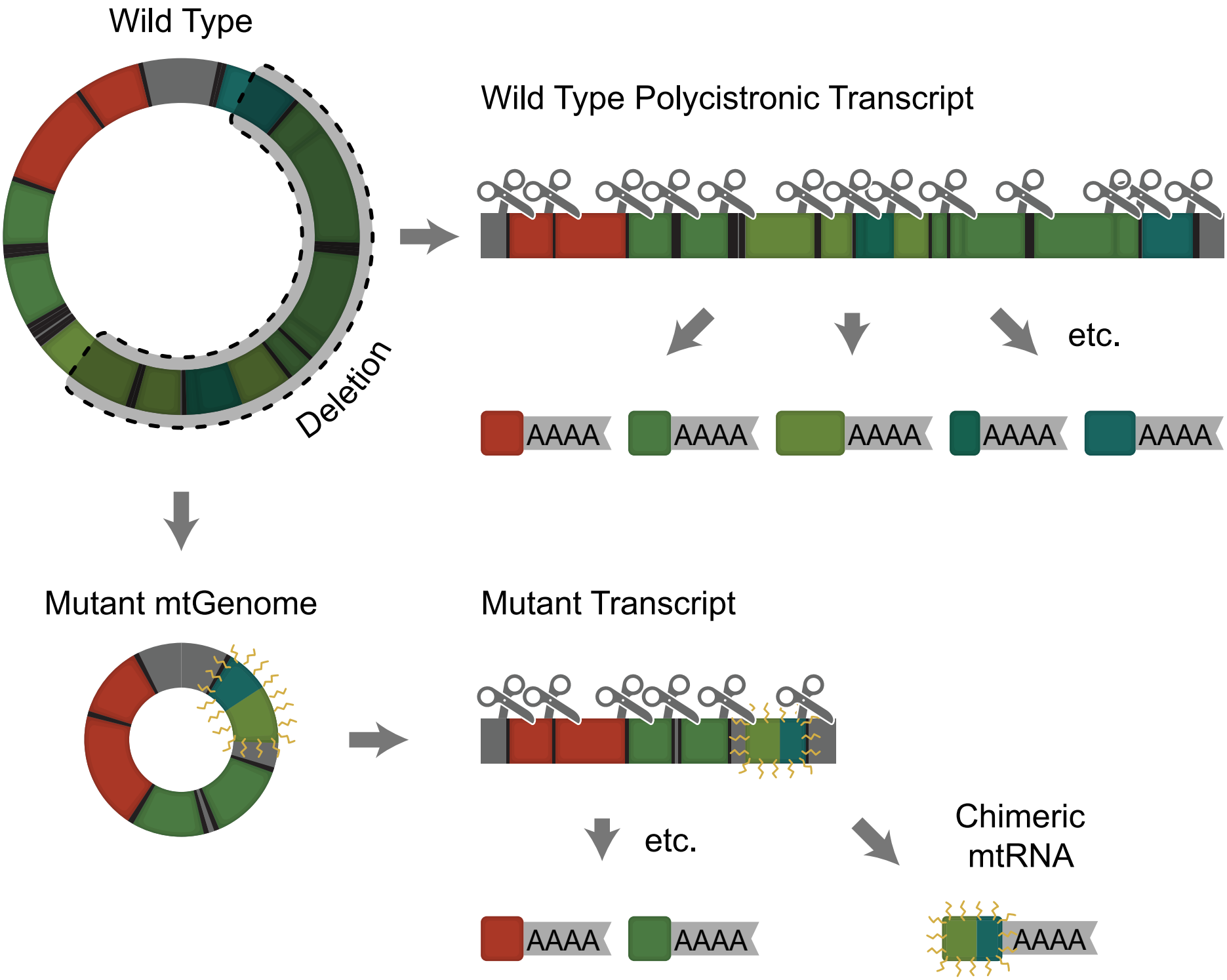
Yang K-C, Yamada KA, Patel AY, Topkara VK, George I, Cheema FH, Ewald GA, Mann DL, Nerbonne JM. 2014. Deep RNA sequencing reveals dynamic regulation of myocardial noncoding RNAs in failing human heart and remodeling with mechanical circulatory support. *Circulation* **129**: 1009–1021.

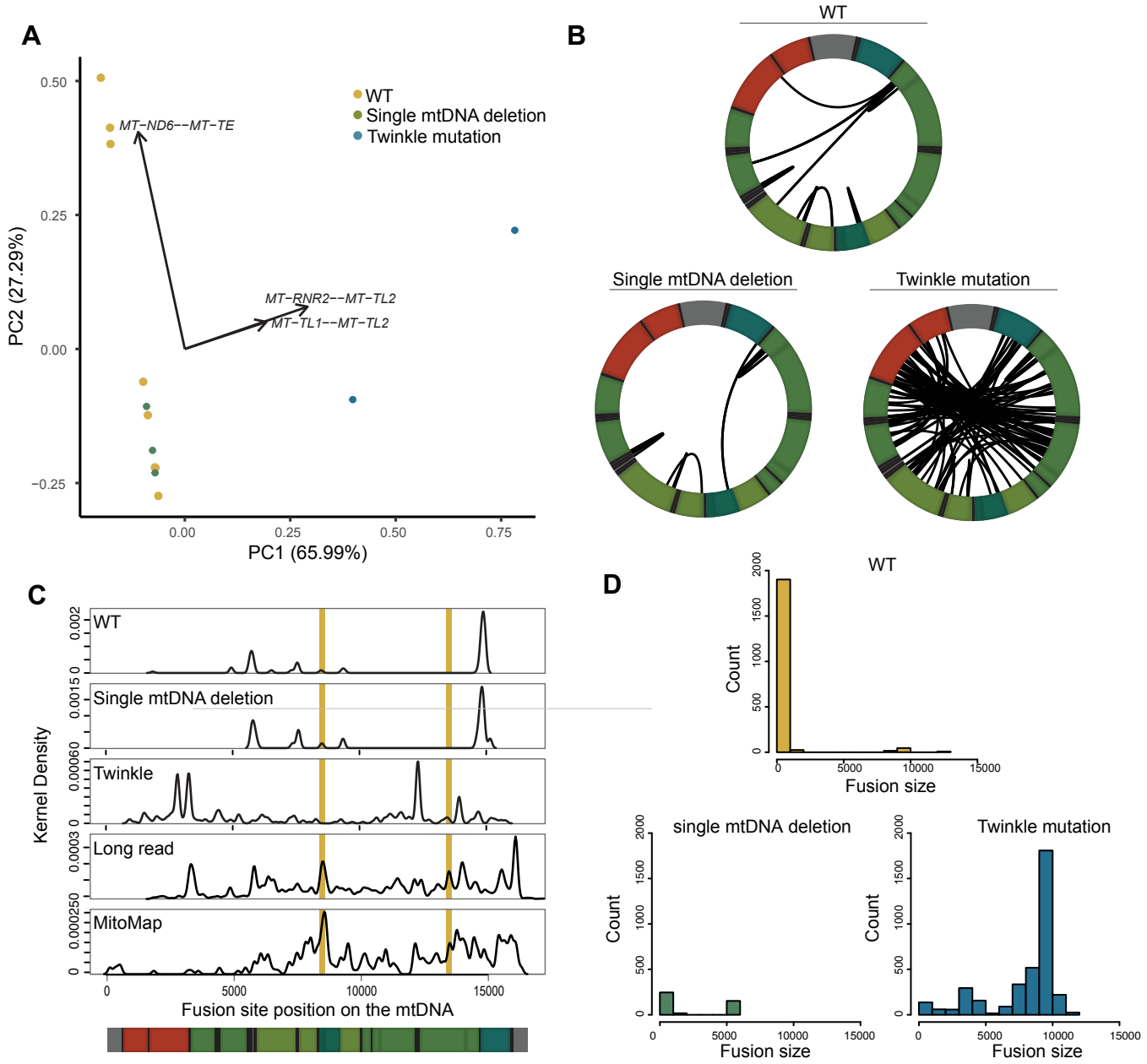
Young MJ, Copeland WC. 2016. Human mitochondrial DNA replication machinery and disease. *Curr Opin Genet Dev* **38**: 52–62.

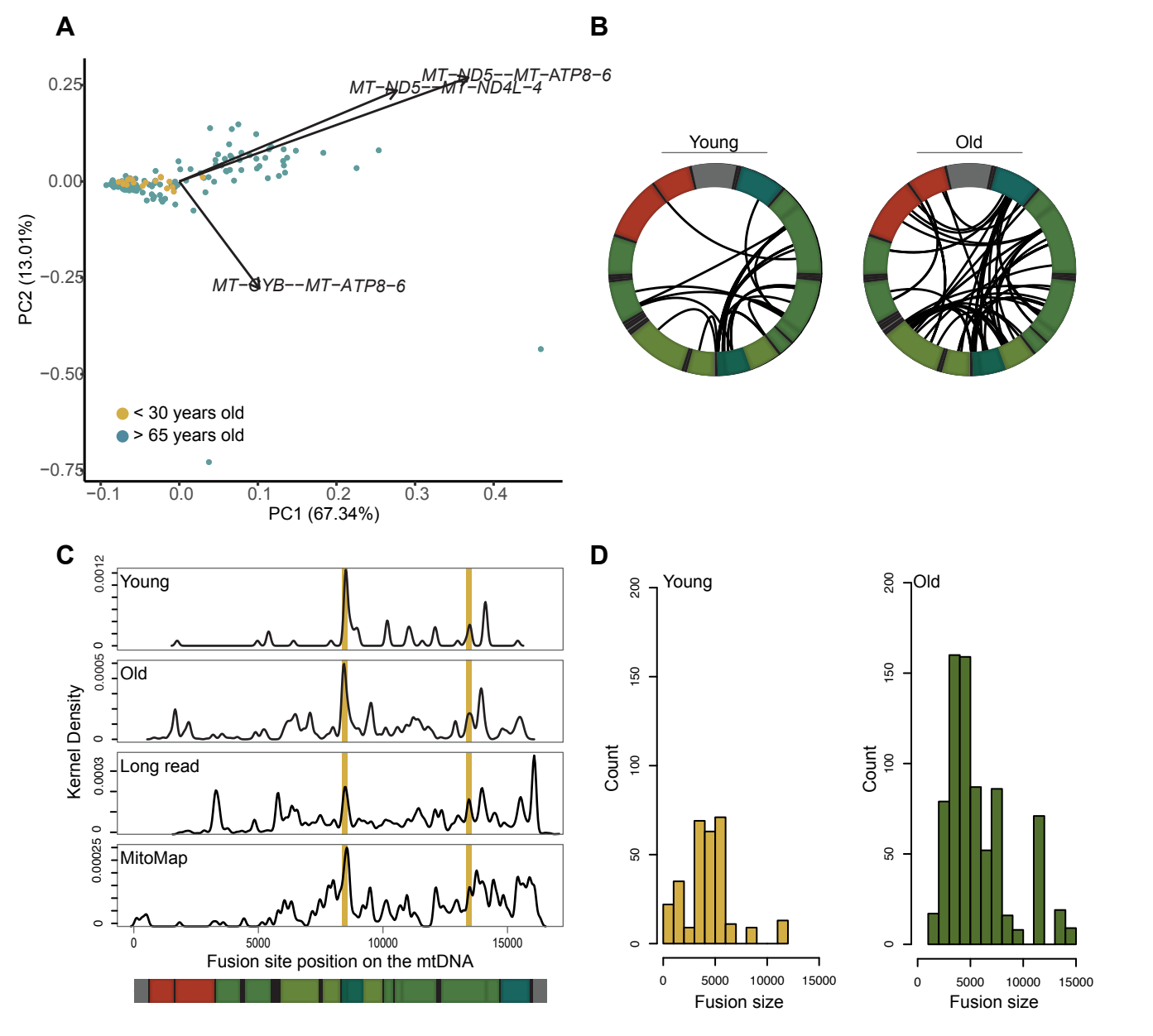
Zecchini V, Paupe V, Herranz-Montoya I, Janssen J, Wortel IMN, Morris JL, Ferguson A, Chowdury SR, Segarra-Mondejar M, Costa ASH, et al. 2023. Fumarate induces vesicular release of mtDNA to drive innate immunity. *Nature* **615**: 499–506.

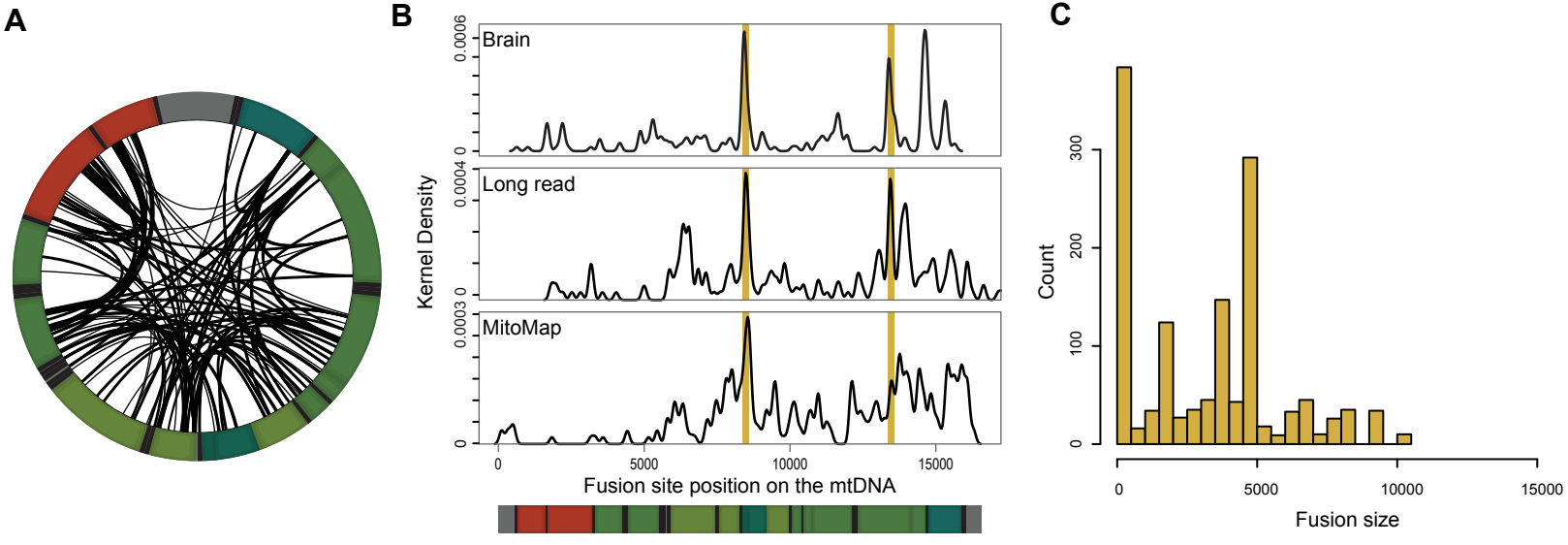
691

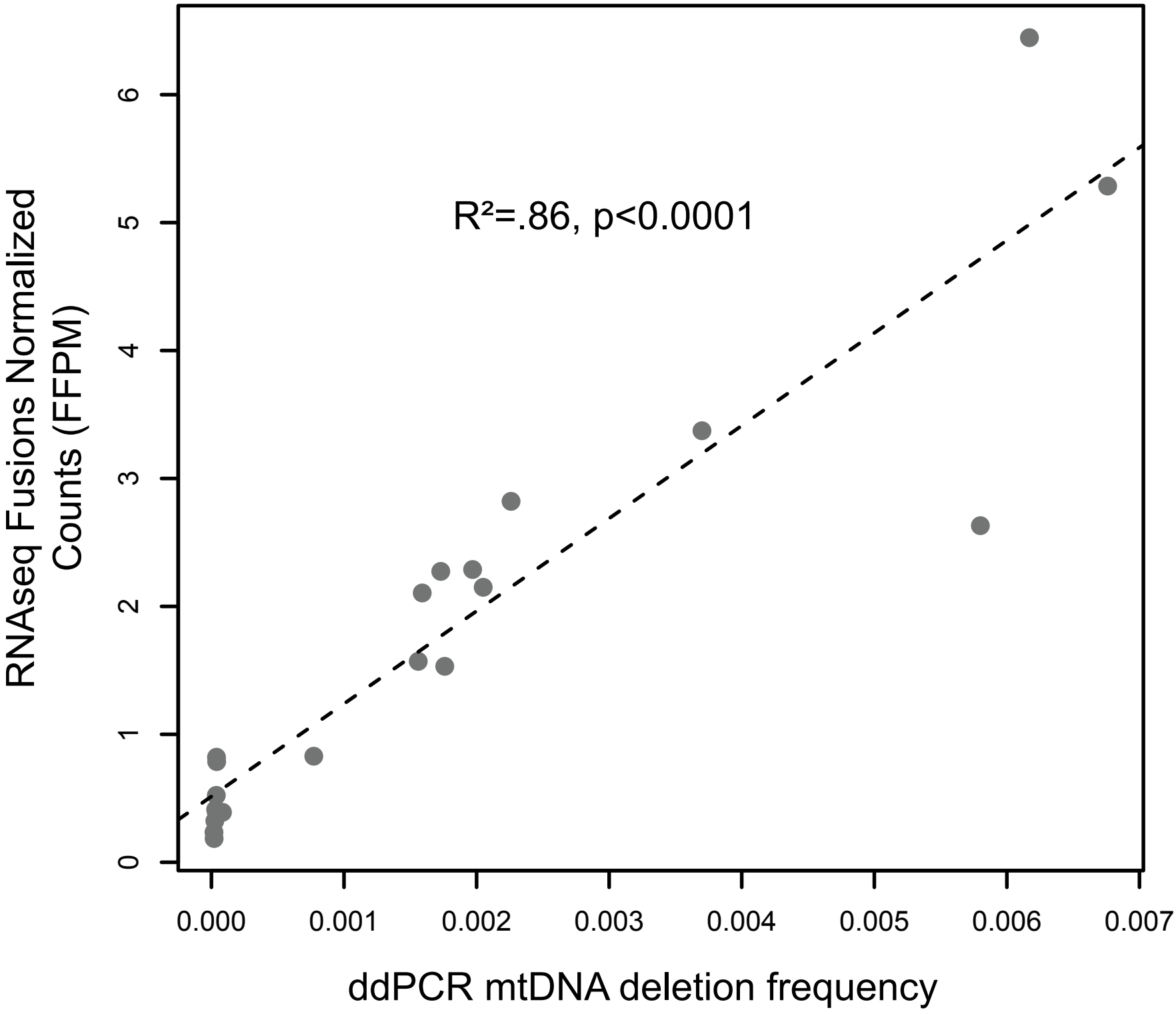
692

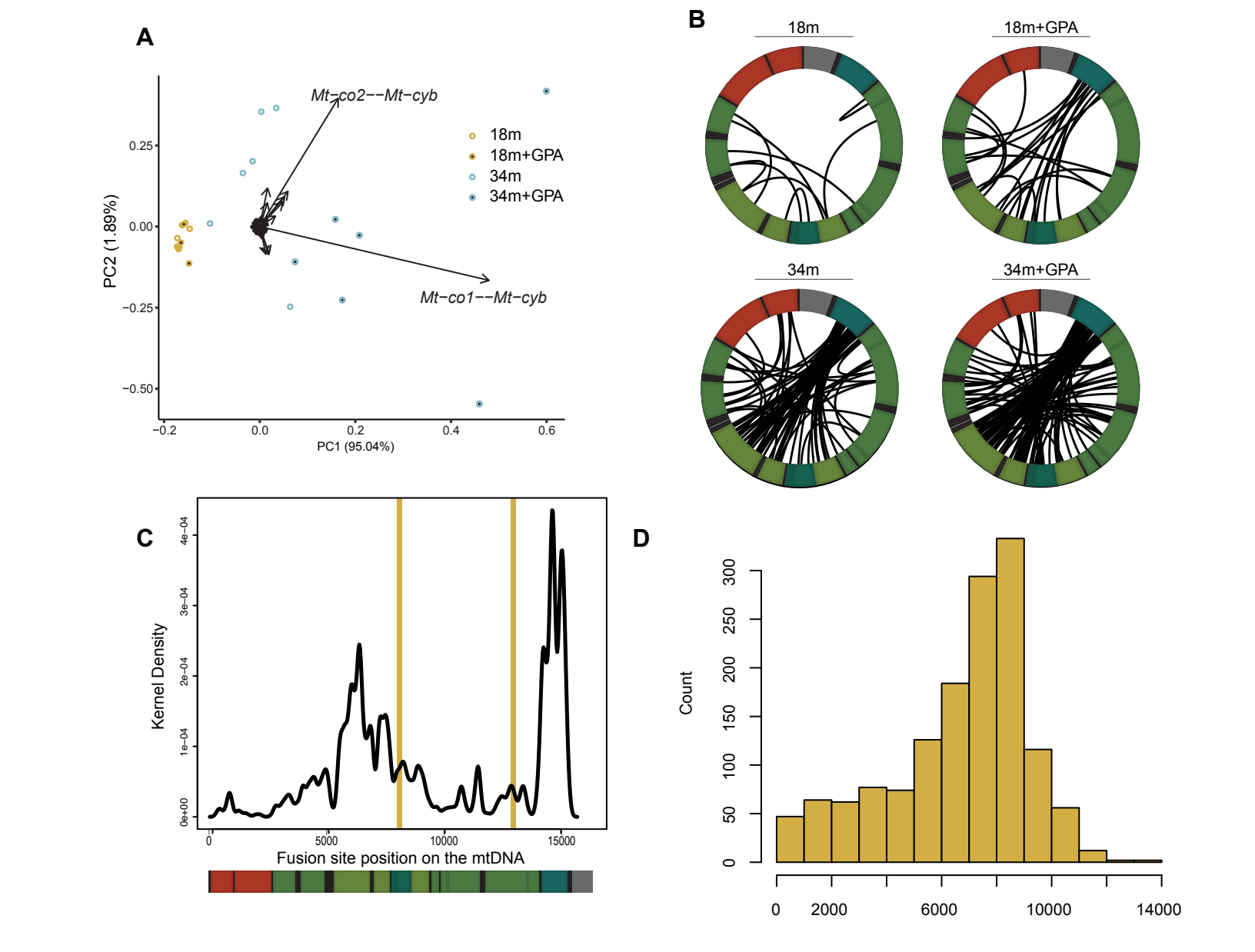

Wild Type
Deletion
Wild Type Polycistronic Transcript
etc.
AAAA
AAAA
AAAA
AAAA
AAAA
Mutant mtGenome
Mutant Transcript
etc.
Chimeric
mtRNA
AAAA
AAAA
AAAA

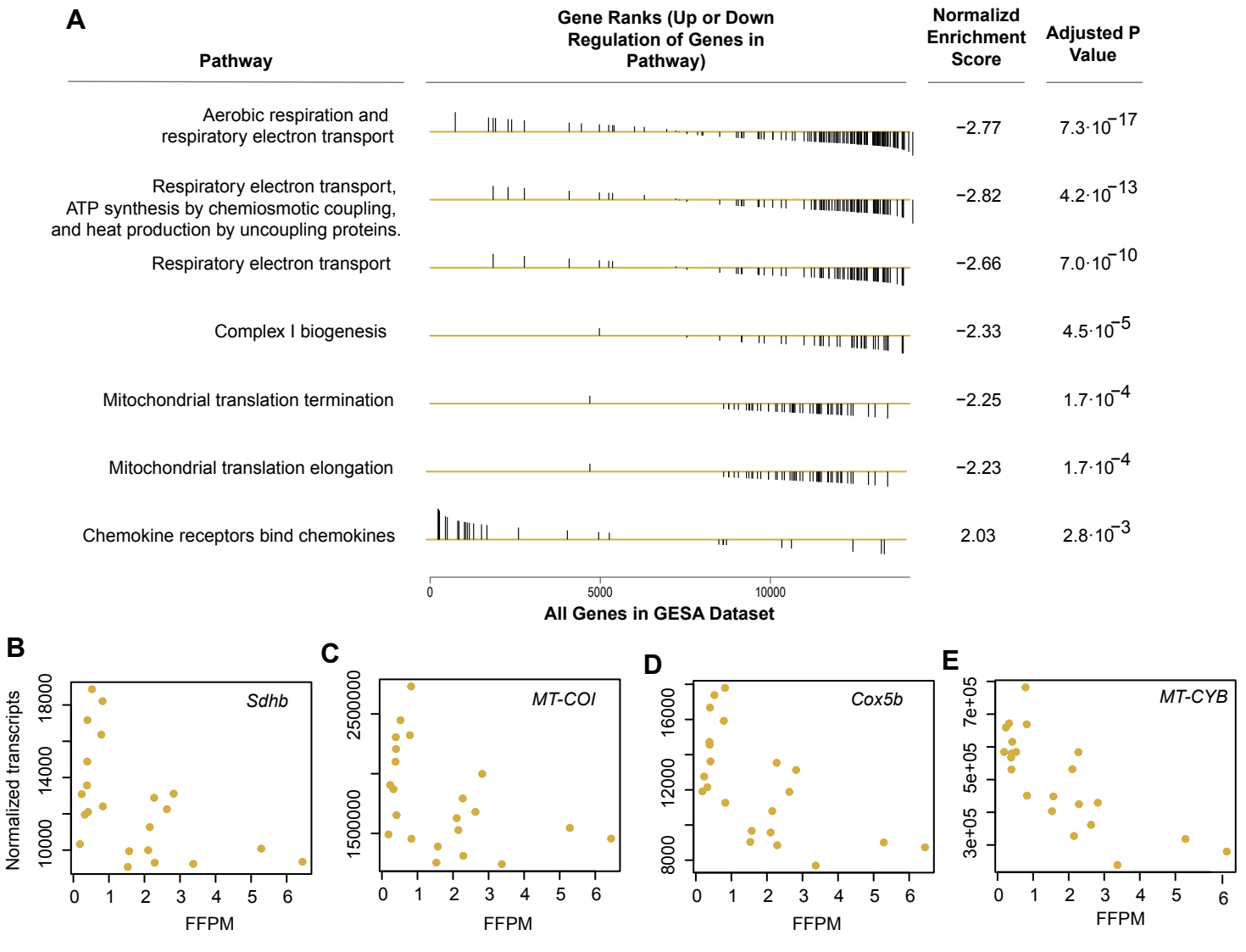

A
WT
Single mtDNA deletion
Twinkle mutation
MT−ND6−−MT−TE
MT−RNR2−−MT−TL2
MT−TL1−−MT−TL2
PC2 (27.29%)
PC1 (65.99%)
B
WT
Single mtDNA deletion
Twinkle mutation
C
WT
Single mtDNA deletion
Twinkle
Long read
MitoMap
Kernel Density
Fusion site position on the mtDNA
D
WT
single mtDNA deletion
Twinkle mutation
Count
Fusion size


A
MT−ND5−−MT−ATP8−6
MT−ND5−−MT−ND4L−4
MT−CYB−−MT−ATP8−6
< 30 years old
> 65 years old
PC2 (13.01%)
PC1 (67.34%)
B
Young
Old
C
Young
Old
Long read
MitoMap
Kernel Density
Fusion site position on the mtDNA
D
Young
Old
Count
Fusion size


A
B
C
Brain
Long read
MitoMap
Kernel Density
0.0006
0.0004
0.0003
0
5000
10000
15000
Fusion site position on the mtDNA
Count
300
200
100
Fusion size


RNAseq Fusions Normalized Counts (FFPM)
0
1
2
3
4
5
6
R²=.86, p<0.0001
0.000
0.001
0.002
0.003
0.004
0.005
0.006
0.007
ddPCR mtDNA deletion frequency


A
Mt−co2−−Mt−cyb
Mt−co1−−Mt−cyb
18m
18m+GPA
34m
34m+GPA
PC2 (1.89%)
PC1 (95.04%)
0.25
0.00
−0.25
−0.50
−0.2
0.0
0.2
0.4
0.6
B
18m
18m+GPA
34m
34m+GPA
C
Kernel Density
0e+00
1e−04
2e−04
3e−04
4e−04
0
5000
10000
15000
Fusion site position on the mtDNA
D
Count
0
50
100
150
200
250
300
0
2000
6000
10000
14000


A
Pathway
Gene Ranks (Up or Down Regulation of Genes in Pathway)
Normalizd Enrichment Score
Adjusted P Value
Aerobic respiration and respiratory electron transport
−2.77
7.3·10−17
Respiratory electron transport, ATP synthesis by chemiosmotic coupling, and heat production by uncoupling proteins.
−2.82
4.2·10−13
Respiratory electron transport
−2.66
7.0·10−10
Complex I biogenesis
−2.33
4.5·10−5
Mitochondrial translation termination
−2.25
1.7·10−4
Mitochondrial translation elongation
−2.23
1.7·10−4
Chemokine receptors bind chemokines
2.03
2.8·10−3
0
5000
10000
All Genes in GESA Dataset
B
Sdhb
Normalized transcripts
10000
14000
18000
0 1 2 3 4 5 6
FFPM
C
MT-COI
1500000
2500000
0 1 2 3 4 5 6
FFPM
D
Cox5b
8000
12000
16000
0 1 2 3 4 5 6
FFPM
E
MT-CYB
3e+05
5e+05
7e+05
0 1 2 3 4 5 6
FFPM