

A concerted increase in readthrough and intron retention drives transposon expression during aging and senescence

Reviewed Preprint

Published from the original preprint after peer review and assessment by eLife.

About eLife's process

Reviewed preprint posted

June 22, 2023 (this version)

Sent for peer review

April 25, 2023

Posted to bioRxiv

January 20, 2023

Kamil Pabis, Diogo Barardo, Kumar Selvarajoo, Jan Gruber, Brian K. Kennedy ✉

Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore • Healthy Longevity Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore • Centre for Healthy Longevity, National University Health System • Bioinformatics Institute, Agency for Science, Technology and Research (A*STAR), Singapore, Singapore • Singapore Institute of Food and Biotechnology Innovation, Agency for Science, Technology and Research (A*STAR), Singapore, Singapore • NUS Synthetic Biology for Clinical and Technological Innovation (SynCTI), National University of Singapore, Singapore, Singapore • School of Biological Sciences, Nanyang Technological University, Singapore, Singapore • Science Divisions, Yale-NUS College, Singapore, 138527, Singapore

 (https://en.wikipedia.org/wiki/Open_access)

 (<https://creativecommons.org/licenses/by/4.0/>)

Abstract

Aging and senescence are characterized by pervasive transcriptional dysfunction, including increased expression of transposons and introns. Our aim was to elucidate mechanisms behind this increased expression. Most transposons are found within genes and introns, with a large minority being close to genes. This raises the possibility that transcriptional readthrough and intron retention are responsible for age-related changes in transposon expression rather than expression of autonomous transposons.

To test this, we compiled public RNA-seq datasets from aged human fibroblasts, replicative and drug-induced senescence in human cells and RNA-seq from aging mice and senescent mouse cells. Indeed, our reanalysis revealed a correlation between transposons expression, intron retention and transcriptional readthrough across samples and within samples. Both intron retention and readthrough increased with aging or cellular senescence and these transcriptional defects were more pronounced in human samples as compared to those of mice.

In support of a causal connection between readthrough and transposon expression, analysis of models showing induced transcriptional readthrough confirmed that they also show elevated transposon expression. Taken together, our data shows that elevated transposon reads during aging seen in various RNA-seq dataset are concomitant with multiple transcriptional defects. Intron retention and transcriptional readthrough are the most likely explanation for the expression of transposable elements that lack a functional promoter.

eLife assessment

This study presents **valuable** findings on the role of transcription readout and intron retention in transposon expression during aging in mammals. While the findings are interesting, the evidence supporting the claims of the authors is **incomplete**, and additional analysis would be required to strengthen the study. The work will be of interest to scientists studying aging, transcription regulation, and epigenetics.

Intro

Repetitive DNA makes up at least 50% of the human genome ([de Koning et al. 2011](#)) and has been linked to genomic instability, cancer and – somewhat less consistently – species longevity ([Pabis et al. 2021](#), [Khristich and Mirkin 2020](#)). Transposons are one particularly abundant class of repetitive sequences found in the nuclear genome of eukaryotes and are strongly associated with cancer-causing structural variants like DNA deletions, insertions or inversions ([Rodriguez-Martin et al. 2020](#)). Although transposons are often considered parasitic, harmful or, at best, neutral recent evidence suggests they may contribute to genomic diversification on evolutionary timescales and to the cellular stress-response by providing transcription factor binding sites ([Villanueva-Cañas et al. 2019](#)).

The major transposon families in the human genome are long interspersed nuclear element (LINEs), short interspersed nuclear element (SINEs), long terminal repeat transposons (LTRs) and DNA transposons comprising 21%, 11%, 8% and 3% of the human genome, respectively ([Kazazian and Moran 2017](#)).

In recent years, transposon expression, particularly of elements belonging to the LINE-1 family, has been hypothesized to be a near universal marker of aging. Several lines of evidence support a link between transposon reactivation and aging ([Gorbunova et al. 2021](#)).

Using RNA-seq, for example, to quantify reads mapping to transposable elements it was shown that expression of these was higher in 10 vs 1 day-old nematodes ([LaRocca, Cavalier, Wahl 2020](#)), 40 vs 10 day-old flies ([Wood et al. 2016](#)) and in muscle and liver of aged mice ([De Cecco et al. 2013](#)). Transposon expression also increases during in vitro senescence ([Colombo et al. 2018](#)) and in fibroblasts isolated from people between the ages of 1 and 96 years ([LaRocca, Cavalier, Wahl 2020](#)), whereas lifespan extending mutations and interventions in mice reduce expression of transposons ([Wahl et al. 2021](#)).

However, not all studies report such consistent age-related increases across all transposon classes ([Ghanam et al. 2019](#)). Moreover, it would be a mistake to conflate changes in RNA-seq-, RNA-, DNA- and protein-based measurements of transposon expression as evidence for one and the same phenomenon.

Quantification of transposon expression using RNA-seq or PCR techniques remains challenging given their repetitive nature and the fact that 98 to 99% of all transposons are co-expressed with neighboring transcriptional units ([Deiningner et al. 2017](#), [Stow et al. 2021](#)). In fact, only a small fraction of LINE-1 and SINE elements, the latter relying on co-transposition by LINE elements, are expressed from a functional promoter. It is believed that such “hot” LINE-1 loci drive most transposition events and genomic instability ([Deiningner et al. 2017](#), [Rodriguez-Martin et al. 2020](#)).

The large number of co-expressed transposons would likely mask any signal from autonomous elements in standard RNA-seq experiments. Therefore, alternative explanations for increased transposon expression in such datasets are needed and these could include age-related changes in transposon adjacent genes e.g. intron retention and transcriptional readthrough.

Most eukaryotic genes contain introns which need to be removed during a complex co-transcriptional process called splicing. Defects in splicing are the cause of many, often neurologic, hereditary conditions ([Scotti and Swanson 2016](#)). One such type of splicing defect is intron retention. Although basal levels of intron retention may be benign or even physiologic, high levels are considered harmful ([Zheng et al. 2020](#)). Aging is accompanied by many changes to splicing patterns ([Wang et al. 2018](#)), including intron retention, which is increased during cellular senescence ([Yao et al. 2020](#)) and in aging drosophila, mouse hippocampus and human prefrontal cortex. In addition, levels in human AD brain tissues are elevated even further than in the aged brain ([Adusumalli et al. 2019](#)).

Transcription is terminated by the cleavage and polyadenylation machinery when the polymerase transcribes through the polyadenylation signal leading to pre-mRNA cleavage, polymerase pausing, conformational change and eventually polymerase detachment ([Rosa-Mercado et al. 2021](#)). However, this termination may fail leading to so called readthrough transcription which increases during stress conditions ([Rosa-Mercado and Steitz 2022](#)) and during cellular senescence, at least for a subset of genes ([Muniz et al. 2017](#)).

Emerging evidence suggests that intron retention, readthrough and transposon expression are linked ([Rosa-Mercado et al. 2021](#), [Hadar et al. 2022](#)), but this has not been studied in the context of aging. In this manuscript we show that intron retention and readthrough taken together can explain most of the apparent increase in transposon expression seen in RNA-seq datasets of aging.

Results

1. Genomic location of transposons implicates them as a marker of multiple transcriptional defects

During aging and senescence expression of all four major transposon families (LTR, DNA, LINE, SINE) increases ([Table S1](#)). However, such a global increase in transposon expression cannot be explained by transposon biology alone since most transposons are ancient, and, having accumulated many mutations, lack a functional promoter ([Özgen et al. 2019](#)). Instead, we reasoned that transposon expression is associated with the expression of nearby genes.

Consistently, out of 4.5 million annotated transposons 60% were located within genes (most of them intronic), 10% within 5 kb of a gene and the remaining 30% were between genes, albeit usually still within 100 kb ([Table 1](#)). For transposons whose expression significantly changed with age or senescence, the number of intronic and downstream-of-gene transposons is even higher.

Table 1.

Fraction of genic or intronic transposons and of transposons downstream (ds) or upstream (us) of genes in two different datasets.

Dataset	Genic	Intronic	ds	us	ds	us	n
all transposons	0.604	0.493	0.045	0.049	0.187	0.174	4520928
Fleischer et al. (aging, significant)	0.857	0.564	0.082	0.025	0.105	0.037	7673
Colombo et al. (senescence, significant)	0.698	0.540	0.107	0.053	0.177	0.118	8730
	within genes		within 5kb		within 100kb		

Based on these findings we studied intronic transposons as a potential marker for age-related intron retention and downstream-of-gene transposons as a marker of age-related readthrough transcription.

2. Elevated transposon expression and intron retention with age

First, we confirmed that age-related transposon expression is indeed elevated as shown by [LaRocca et al. \(2020\)](#). This analysis is based on a dataset originally published by [Fleischer et al. \(2018\)](#), who performed RNA-seq of 143 fibroblasts isolated from donors between the ages of 0 to 96. This dataset will be hereafter referred to as the “aging dataset”. We find that transposon expression increases slowly and non-linearly during aging, with a sudden and sharp increase around 80 years of age ([Fig. 1A](#), [Fig. S1](#)). Next, we showed that intron retention is also elevated with aging as shown by [Yao et al. \(2020\)](#). We find that intron retention increases with age ([Fig. 1B](#)) and that samples with high transposon expression also showed elevated intron retention ([Fig. 1C](#)).

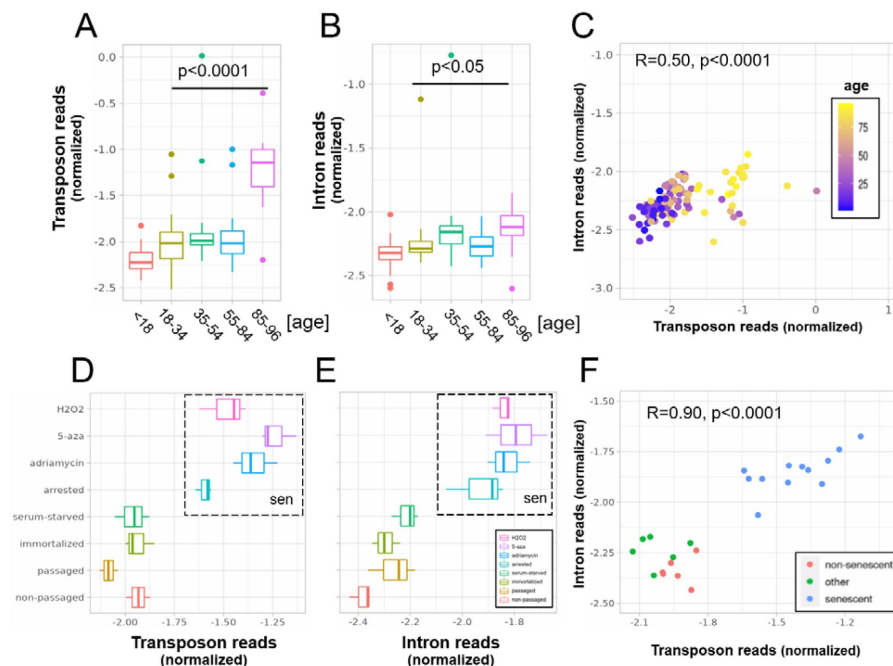


Fig. 1.

A concerted increase in transposon and intron expression with aging and senescence.

Transposon (A) and intron (B) expression is increased in skin fibroblasts isolated from aged individuals. What is more, expression of both transposons and introns shows a significant correlation (C). Similarly, transposon (D) and intron (E) expression is increased in MDAH041 cells induced into senescence (sen) by treatment with drugs or via passaging. Normalized reads from the top 1000 differentially expressed genes,

transposons and introns were used in this analysis.

A) Transposon reads normalized by the expression of adjacent genes plotted across five age groups (adolescent n=32, young n=31, middle-aged n=22, old n=37 and very old n=21).

B) Intron reads normalized by the expression of adjacent genes plotted across five age groups (adolescent n=32, young n=31, middle-aged n=22, old n=37 and very old n=21).

C) Scatterplot between transposon and intron expression (normalized as in A and B) for all 143 samples. Each sample is colored by age.

D) Transposon reads normalized by the expression of adjacent genes plotted across four senescent conditions (H₂O₂, 5-azacytidine, adriamycin, replicative senescence) and four other conditions (serum-starved, immortalized, intermediate passage and early passage). N=3 per group.

E) Intron reads normalized by the expression of adjacent genes plotted across four senescent conditions (H₂O₂, 5-azacytidine, adriamycin, replicative senescence) and four other conditions (serum-starved, immortalized, intermediate passage and early passage). N=3 per group.

F) Scatterplot between transposon and intron expression (normalized as in E and F) for all 24 samples. Each sample is colored by senescence status.

In our analysis we corrected for gene co-expression by normalizing the expression levels of introns and transposons to the expression levels of the closest gene. However, another complementary approach would be to plot and compare the raw, unadjusted log-fold changes for genes, introns, transposons and other elements. In this analysis, we found that, on average, age-related upregulation of transposons and introns is more pronounced than upregulation of genes (**Fig. S2A**), consistent with **Fig. 1A** and **1B**.

Next, we re-analyzed a dataset originally published by **Colombo et al. (2018)** which will be referred to as the “senescence dataset”. Here, we tested whether MDAH041 cells induced into senescence with H₂O₂, 5-azacytidine, adriamycin or serial passaging show elevated transposon expression and intron retention. As expected, all these treatments increased the expression of transposons (**Fig. 1D**, **Fig. S2B**), as documented before in this dataset (**Colombo et al. 2018**). The retention of introns was also increased (**Fig. 1E**) and expression levels of introns and transposons correlated with each other (**Fig. 1F**).

Having established that transposons reads and retained introns are elevated with aging and senescence, we tested whether this is conserved between mouse and human. Transposon expression ($p < 0.0001$) was elevated, and intron retention showed a trend towards elevation ($p = 0.14$), in the liver of 26 month-old mice, suggesting these age-related transcriptional defects might be conserved between species (**Fig. S3A, B**). In contrast, we found that increased expression of transposons and intron retention in models of cellular senescence is not conserved between species. To the contrary, early passage mouse fibroblasts showed elevated transposon expression and intron retention compared to senescent fibroblasts (**Fig. S3C, D**), which contrasts with the mouse liver and human senescence data.

3. Transposon expression as a marker of intron retention

So far these age-related events have been considered independently, but given the localization of transposons being mostly intragenic we reasoned that intron retention could bias the transposon findings. Indeed, we found that samples with high intron retention also show high transposon expression. This is true both in the aging dataset (**Fig. 1C**) and in the

cellular senescence dataset (Fig. 1F), where intronic reads accounted for most of the observed variability in transposon expression (or vice versa). Looking at individual genomic loci, we also found a strong correlation between the log-fold change of a transposon and the log-fold change of the intron it is located in (Fig. S4).

As a more direct test of this hypothesis, we determined the correlation between the counts of retained introns and the counts of intronic transposons for each sample individually. Non-intronic transposons and randomized transposons served as a negative control. Consistent with the log-fold change data, we found a specific correlation between intronic transposon and intron counts in both datasets, with stronger correlations in the senescence dataset (Fig. S5).

Altogether this data suggests that intron retention contributes to the age-related increase in the expression of transposons (Fig. 2).

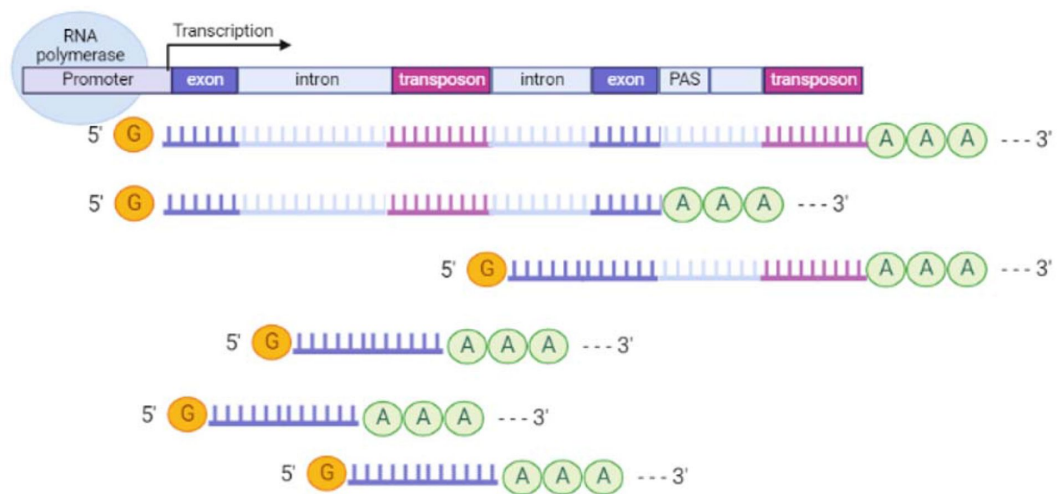


Fig. 2.

A schematic model of transposon expression.

In our model, transcription can give rise to mRNAs and pre-mRNAs that contain retained introns when co-transcriptional splicing is impaired, as seen during aging and senescence, and these often contain transposon sequences. In addition, transcription can give rise to mRNAs and pre-mRNAs that contain transposon sequences towards the 3'-end of the mRNA when co-transcriptional termination at the polyadenylation signal (PAS) is impaired as seen with aging and senescence. Some of these RNAs may be successfully polyadenylated whereas others will be subject to nonsense mediated decay. Image created by Biorender.

4. Expression of extragenic transposons as a marker of readthrough transcription

Most transposons showing increased expression with aging or cellular senescence mapped within genes or introns and their expression changes may be attributed to intron retention. How can we explain expression of the remaining 14 to 30% of transposons? (Table 1)

To answer this question, we analyzed the distribution of transposons mapping to extragenic regions. Out of all the transposons, showing significantly changed expression during aging or cellular senescence, most within 5 kb of genes and they were found predominantly downstream of genes (Fig. 3A). The distribution of transposons in the aging and senescence dataset was very similar except for a wider distribution of senescence-associated transposons within a 100 kb window around genes (Fig. 3B). This contrasts with the almost completely even distribution of randomly permuted transposons.

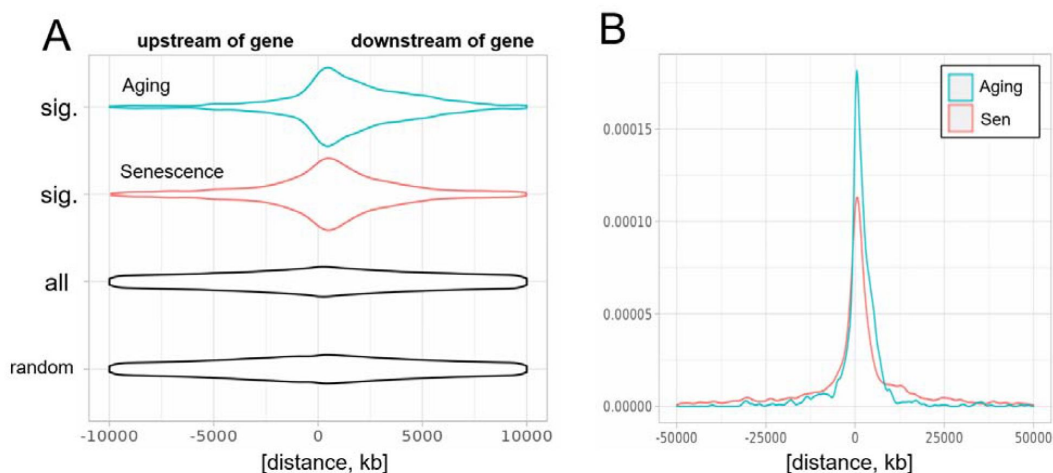


Fig. 3.

Biased distribution of extragenic transposons relative to genes.

(A) Extragenic transposons with significantly (sig.) changed expression during aging or cellular senescence show a biased distribution, when mapped back onto the genome, with a preference towards a 5 kb region at the 3'-end of genes when compared to all annotated transposons (all) or randomly permuted transposons (random). (B) Extragenic transposons whose expression changes during cellular senescence are spread out further from genes compared to aging-associated transposons.

Given the proximity of significantly changed transposons and genes, this suggested that an age-related increase in transcriptional readthrough could explain elevated transposon expression. To test this, we next quantified readthrough in a 10 to 20 kb region downstream of genes. We chose a region not immediately adjacent to genes to avoid biases through basal levels of readthrough, although the data was comparable for other distances (Fig. S7).

Indeed, readthrough transcription was increased with aging (Fig. 4A, B), and in the senescence dataset, levels of readthrough transcription increased with replicative or chemically induced senescence (Fig. 4D, E). Transposon expression correlated with readthrough levels on a per sample basis in both the aging and the senescence dataset (Fig. 4C, F).

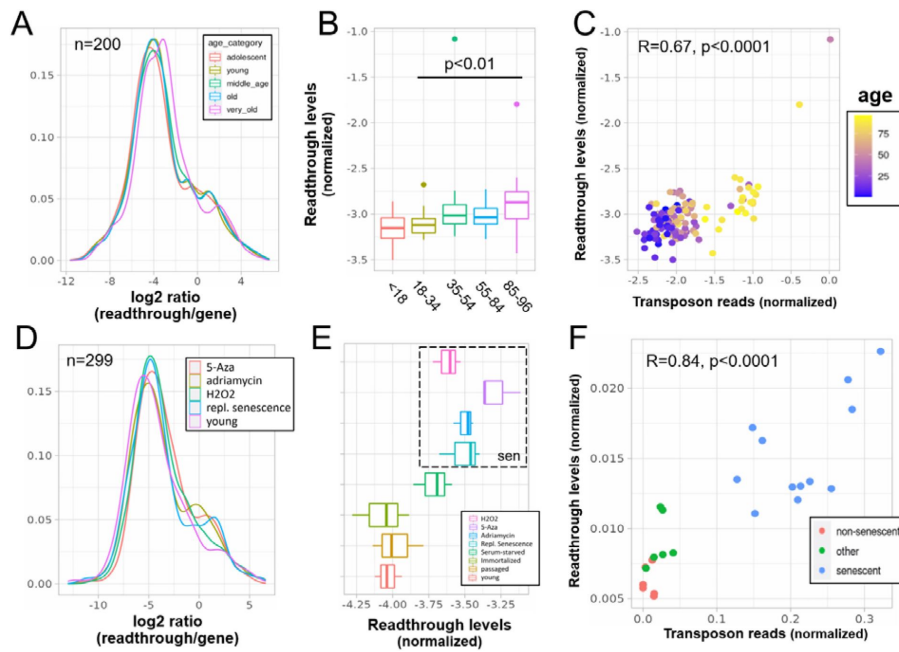


Fig. 4.

Increased readthrough levels with aging and senescence.

Readthrough transcription is increased in fibroblasts isolated from the very old (A, B) and after induction of senescence in vitro (D, E). Readthrough transcription is correlated with transposon expression (C, F).

A) Readthrough was determined in a region 10 to 20 kb downstream of genes for a subset of genes that were at least 20 kb away from the nearest neighboring gene (n=206 regions). The log₂ ratio of readthrough

to gene expression is plotted across five age groups (adolescent n=32, young n=31, middle-aged n=22, old n=37 and very old n=21).

B) As in (A) but data is plotted on a per sample basis.

C) Scatterplot between transposon expression and readthrough levels (normalized as in A and Fig. 1) for all 143 samples. Each sample is colored by age.

D) Readthrough was determined in a region 10 to 20 kb downstream of genes for a subset of genes that were at least 20 kb away from the nearest neighboring gene (n=299 regions). The log₂ ratio of readthrough to gene expression is plotted across four senescent conditions (H₂O₂, 5-Aza, Adriamycin, replicative senescence) and for early passage cells. N=3 per group.

E) As in (D) but data is plotted on a per samples basis and for additional control datasets (serum-starved, immortalized, intermediate passage and early passage). N=3 per group.

F) Scatterplot between transposon expression and readthrough levels (normalized as in A and Fig. 1) for all 24 samples. Each sample is colored by senescence status.

As shown before for intron retention, these findings may not translate to other species. We found that readthrough was unchanged in the liver of 26 month-old mice (p=0.22, Fig. S6A). In addition, readthrough decreased with replicative senescence of mouse fibroblasts, contrary to our expectations (Fig. S6C, D).

Next, we quantified readthrough across a 100 kb region downstream of genes. We found that both aging and senescence lead to readthrough across the whole region, although senescence was associated with more extensive readthrough further downstream of genes (Fig. S7, Fig. 3B). As another line of evidence we quantified transcription upstream of genes, which is called read-in. This kind of transcription is often attributed to widespread readthrough extending from the end of a gene to the start of another (Roth et al. 2020). Consistent with the extensive readthrough levels observed we found elevated read-in transcription during cellular senescence (Fig. S8A) but not aging (Fig. S8B).

Having shown that samples with higher levels of readthrough also have higher transposon expression (per sample correlation), we asked whether transposons are more often than expected located in readthrough regions (per locus correlation). Indeed as expected, in both datasets, transposons significantly changed with age or senescence were more often found in significant readthrough regions compared to all transposons (**Table S2**). Moreover, there was a significant correlation between the log-fold changes of transposons downstream of genes with the log-fold changes of readthrough downstream of the same genes (**Fig. S9**). Transposons located elsewhere showed a weaker correlation with readthrough (**Table S3**).

To further test whether these correlations between log-fold changes are also reflected on the count level, we determined the correlation between readthrough counts and the counts of transposons downstream of genes for each sample individually. Transposons located upstream of genes and randomized transposons served as a negative control. Consistent with the above data, we found a specific correlation between transposon and readthrough counts in both datasets (**Fig. S10**).

Finally, we asked whether readthrough levels are predictive of transposon expression on a per sample basis independently of intron retention. Using the aging dataset with 143 samples we compared a linear model just including intron retention with one also including readthrough and found the second model to perform significantly better ($p < 0.0001$, **Table S4**). In addition, we found that the correlation between readthrough levels and transposon expression remained significant even when correcting for intron retention using partial correlation ($R = 0.53$, $p < 0.0001$). We did not test this in the senescence dataset due to its limited size.

5. Readthrough is a causal inducer of transposon expression

Based on this correlative evidence (**Fig. 4**, **Fig. S9**), we hypothesized that specific inducers of readthrough would also increase transposon expression. To test this, we re-analyzed literature data from studies involving multiple treatments that promote readthrough. These included viral infection, hyperosmotic stress (KCl treatment), heat shock and oxidative stress.

We found that KCl treatment of HEK293 cells for 1 hour ([Rosa-Mercado et al. 2021](#)) indeed strongly induced both readthrough (**Fig. 5A**) and transposon expression (**Fig. 5B**) as determined by RNA-seq and TTLseq. Consistent with this, in our reanalysis of a study using mNET-seq profiling of nascent mRNA ([Bauer et al. 2018](#)), KCl treatment of HEK293 cells also induced readthrough (**Fig. 5D**) and transposon expression (**Fig. 5E**), whereas influenza infection caused an intermediate phenotype. Finally, we analyzed RNA-seq data from NIH-3T3 mouse fibroblasts subjected to three stressors over two hours ([Vilborg et al. 2017](#)). Heatshock caused extensive readthrough (**Fig. S11A**) and transposon expression (**Fig. S11B**), whereas hydrogen peroxide treatment and KCl caused moderate expression of these.

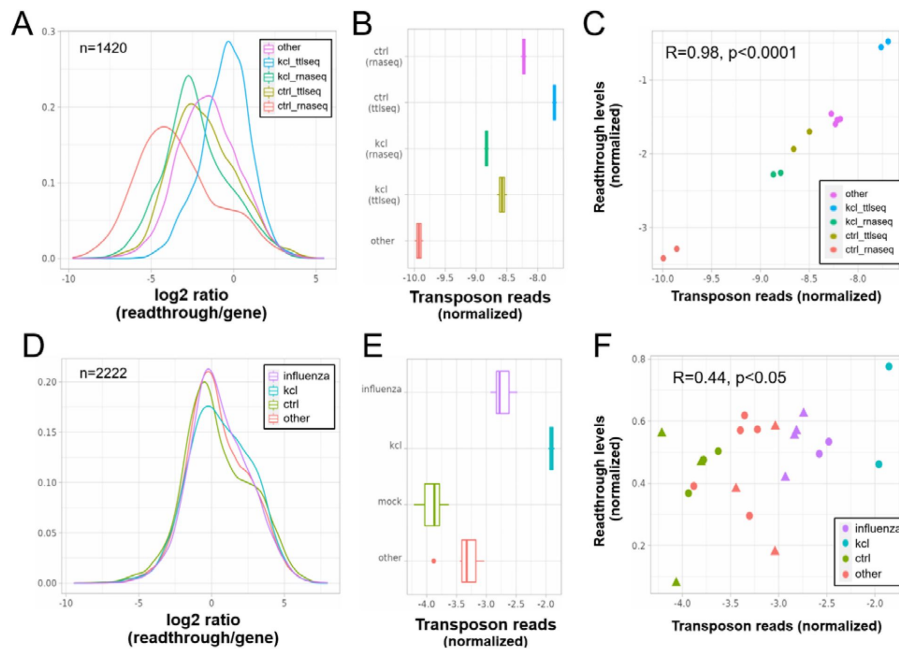


Fig. 5.

Elevated transposon expression after induced readthrough.

Readthrough transcription is increased after KCl treatment of HEK293 cells (A, D) and to a lesser extent during influenza infection (D). These readthrough inducing treatments also promote increased transposon expression (B, E). Giving rise to a correlation between readthrough and transposon expression on a per sample basis (C, F). Data in (A-C) is from [Rosa-Mercado et al. \(2021\)](#). In this study HEK293 cells were subjected to 1 hour of hyperosmotic stress (KCI) after which

RNA-seq and TTLseq were performed. Data in (D-F) is from [Bauer et al. \(2018\)](#). In this study HEK293 cells were subjected to 1 hour of hyperosmotic stress (KCI) after which mNETseq was performed.

A) Readthrough was determined in a region 10 to 20 kb downstream of genes for a subset of genes that was at least 20 kb away from the nearest neighbouring gene (n=1420 regions).

B) To normalize transposon expression counts for each transposon were corrected for the expression of the nearest gene. Normalized transposon counts for each sample are shown as box-whisker plot.

C) Normalized transposon counts (as in B) and readthrough levels are plotted for each sample (n=12).

D) Readthrough was determined in a region 10 to 20 kb downstream of genes for a subset of genes that was at least 20 kb away from the nearest neighbouring gene (n=2222 regions).

E) To normalize transposon expression counts for each transposon were corrected for the expression of the nearest gene. Normalized transposon levels for each sample are shown as box-whisker plot.

F) Normalized transposon counts (as in B) and readthrough counts are plotted for each sample (n=22). The shapes indicate the cell type (circles=HEK293, triangles=A549).

We also reasoned that readthrough should specifically increase the expression of transposons that are close to genes. Consistent with this notion, induced readthrough led to preferential expression of gene proximal transposons (i.e. those within 25 kb of genes), when compared with senescence or aging ([Fig. S12A](#)). Moreover, the fraction of significantly upregulated transposons found outside of genes was higher with induced readthrough than with senescence or aging ([Fig. S12B](#)).

6. Increased expression of autonomous LINE-1 elements with senescence but not aging

Many expressed LINE-1 elements are found within genes (Table S5), which poses a challenge for the quantification of active LINE-1 elements. Given this proximity it is unsurprising that a substantial correlation between the expression of transposons and neighboring genes (Fig. 6) is seen in RNA-seq.

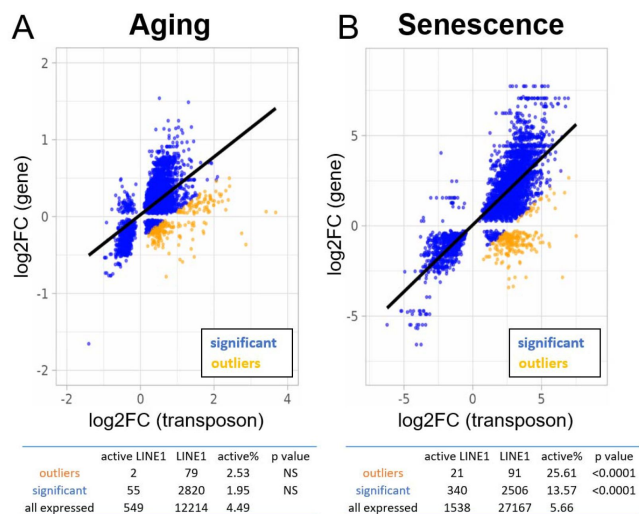


Fig. 6.

Increased expression of active LINE-1 elements with senescence but not aging.

A scatterplot between the log₂-fold change of every transposon ($p < 0.05$) differentially expressed with age or senescence and the log₂-fold change of the nearest gene. The percentage of active LINE-1 elements among transposons with higher-than-expected expression levels is unchanged in aging (A) and increased during senescence (B). For the tables, we compared the percentage of active LINE-1 elements among all LINE-1 elements for 250 transposons with higher-than-expected expression levels (in orange; labeled “outliers”), all significantly changed transposons (blue plus orange; labeled “significant”) and all transposons passing

read filtering (not plotted; labeled “all expressed”). The p-value is based on the comparison with “all expressed”.

To mitigate this problem, we selected 250 transposons that showed higher than expected expression levels compared to the nearest gene. We reasoned that LINE-1 elements with a functional promoter (“active” LINE-1s) would be more likely autonomous and should be enriched among transposons that show a divergent expression profile from their neighboring genes.

We found no enrichment for active LINE-1 elements in the aging dataset (Fig. 6A). In contrast, transposons significantly upregulated with senescence and, among those, especially the ones with higher-than-expected expression, showed significant enrichment for active LINE-1 elements (Fig. 6B). Moreover, we found that almost 100% of active LINE-1 elements were increased with senescence which was higher than for all LINE-1 elements and other transposons (Table S5).

Finally, we redid the same analysis with a different definition of active LINE-1 elements as those potentially encoding functional ORF1p and ORF2p. In this analysis, we included all elements encoding an open reading frame > 900 nt with high sequence similarity to ORF1p and >3400 nt with high sequence similarity to ORF2p ($p < 0.05$ for pairwise alignment). Out of 122626 annotated LINE-1 elements 585 encoded a putative ORF1p and 133 encoded an ORF2p.

ORF1p-encoding LINE-1 elements were significantly enriched among transposons with higher-than-expected expression and among transposons significantly changed with senescence or aging (Table S7). These findings were more robust for senescence than for aging. ORF2p-encoding LINE-1 elements, in contrast, were too rare to be studied reliably as to their enrichment.

Discussion

We show that transposons would serve as a useful biomarker of aging, even if most of the transposon signal in RNA-seq was due to co-expression with nearby transcriptional units, because they capture different kinds of transcriptional dysfunction.

Introns

Intron retention is one prominent example of dysfunctional splicing during aging, but the association with age-related transposon expression has not been evaluated. A recent study by [Gualandi et al. \(2022\)](#) using RNA-seq data from the 1000 Genomes project found a strong per sample correlation between intron and transposon expression in the general population. In line with these results, we showed that intron retention may explain a large fraction of transposon reads in aging and senescence datasets ([Fig. 1](#)).

Readthrough, read-in and transposons

In a study of oncogene-Induced senescence, [Muniz et al. \(2017\)](#) found increased readthrough transcription at 91 so called START loci, which was associated with suppression of neighboring genes through the production of antisense RNAs and several of these antisense RNAs were necessary for the maintenance of the senescent state. Our data is consistent with their model, although we find even more extensive readthrough transcription. Analyzing hundreds of loci we show that senescence promotes transcriptional readthrough leading to the overexpression of intergenic transposons (>25 kb from the nearest gene) and read-in at neighboring genes ([Fig. 3B](#), [Fig. S8B](#)). In contrast, aging-related readthrough was qualitatively different, producing shorter readthrough and no read-in ([Fig. 3B](#), [Fig. S8A](#)).

Our correlative data suggested that readthrough transcription could account for the expression of many transposons that are outside of genes and thus cannot be explained by intron retention ([Fig. 4C](#), [Fig. 4F](#)). To test this hypothesis, we reanalyzed data from models of induced readthrough showing that readthrough is indeed a causal inducer for the expression of gene proximal transposons after hyperosmotic stress and viral infection ([Fig. 5](#)). Given the technical challenges of measuring readthrough and the unreliability of intron retention detection tools ([David et al. 2022](#)) our results are likely underestimating the true magnitude of correlation between readthrough and transposon expression.

In a previous study it was observed that senescence leads to increased expression of transposons and of genes with low abundances ([Zhang et al. 2021](#)). This “leakage” was attributed to loss of H3K9me3 histone marks and increased heterochromatin accessibility, but is also consistent with readthrough transcription, as we observed. Indeed, it has been suggested that readthrough is the most likely cause for apparent intergenic transcription ([Agostini et al. 2021](#)).

Is there genuine expression of autonomous transposons during aging?

The distribution of transposons showing apparent age-related expression within the genome is an important question since intragenic transposons are more likely to be co-expressed with genes. Using Sanger sequencing of long-range amplicons to study replicative senescence, [De Cecco et al. \(2019\)](#) found a large fraction of transposons mapping to

intergenic (69%) regions and few to intragenic/gene-proximal regions (31%). In contrast using RNA-seq data, we find that most transposons are located within genes ($\geq 70\%$; [Table 1](#)), although, there are more intergenic transposons significantly changed with senescence than with aging ([Fig. 3](#)). One explanation for this discrepancy is the mapping approach used. Since their analysis did not use an expectation-maximization algorithm to map back transposons onto the genome, it may underestimate the number of intragenic and gene-adjacent transposons in contrast to our analysis using TELocal.

Given that inactive LINE-1 loci greatly outnumber active loci ([Deininger et al. 2017](#), [Stow et al. 2021](#)), and we saw only limited evidence for active LINE-1 transcription in fibroblasts isolated from donors of varying ages ([Fig. 6A](#)), our data appears to question the role of transposons in organismal aging. An intriguing explanation is provided by [De Cecco et al. \(2019\)](#), who found that LINE-1 ORF1 protein-positive cells colocalize with a small subset of senescent cells. Most likely, such changes would not be obvious from bulk sequencing, perhaps explaining why we did not see elevated LINE-1 expression in the aging dataset. In contrast, the signal would be readily detectable in cultures induced into senescence where most cells do enter senescence, which is consistent with our findings ([Fig. 6B](#)).

As for another explanation, we speculate that LINE-1 elements expressed through intron retention and readthrough are compatible with translation via leaky scanning or re-initiation ([Mouilleron et al. 2015](#)), accounting for increased LINE-1 ORF1 protein levels detected in senescent cells, even if leaky scanning is orders of magnitude less efficient than normal initiation.

Summary

Taken together, our work and existing literature suggest that RNA-seq, and by extension qPCR, of transposable elements overestimates their true expression and their age-related changes. Such data alone is insufficient as support for a “transposon hypothesis of aging” that postulates an increase in autonomous expression of such elements. Instead, a concerted increase in transposon expression, intron retention and transcriptional readthrough points to a global increase in transcriptional dysfunction during aging and senescence, at least in human cells.

Nevertheless, there are several other techniques such as immunoblotting for ORF1 protein, rapid amplification of cDNA ends (RACE), or genomic PCR that may (and do) provide evidence in favor of age-related reactivation of potentially autonomous transposons, specifically LINE-1 elements.

Although we have shown increased transcriptional dysfunction with aging and cellular senescence, so far we lack a model to explain these findings. In this context it is noteworthy that the elongation rate of Pol II is correlated with readthrough transcription ([Fong et al. 2015](#)), intron processing, and may increase with aging ([Debès et al. 2019](#)), potentially providing a unifying explanation.

Methods

Datasets, read preprocessing and alignment

We searched the SRA for suitable datasets that used total RNA extraction because this improves the detection of non-canonical transcripts that may be degraded before successful polyadenylation. Publicly available human RNA-seq datasets of aging and senescence,

respectively, were obtained from [Colombo et al. 2018 \(GSE60340\)](#) and [Fleischer et al. 2018 \(GSE113957\)](#). Human RNA-seq and TTLseq data was obtained from [Rosa-Mercado et al. 2021 \(GSE152059\)](#), human mNET-seq data from [Bauer et al. 2018 \(PRJNA432639\)](#) and mouse RNA-seq data from [Vilborg et al. 2017 \(GSE98906\)](#).

Mouse liver RNA-seq data from ad libitum and dietary restricted mice was obtained from [Hahn et al. 2017 \(GSE92486\)](#) and mouse replicative senescence data from [Wang et al. 2022 \(GSE179880\)](#).

Reads were filtered, adaptor-trimmed and repaired using fastp ([Chen et al. 2018](#)) with a min Phred score of 25. Afterwards reads passing these filters were aligned to the human genome (GRCh38) or mouse genome (GRCm39) using STAR 2.7.8a ([Dobin et al. 2013](#)) in single-pass mode. To optimize the quantification of transposons the STAR parameters `--outFilterMultimapNmax` and `--winAnchorMultimapNmax` were set to 100. The full pipeline with all the configuration files is available on github ([pabisk/aging_transposons](https://github.com/pabisk/aging_transposons)).

Detection of readthrough and intron retention

We employed a modified Automatic Readthrough Transcription Detection (ARTDeco) pipeline, using their provided annotations ([Roth et al. 2020](#)) but with a different quantification approach. Readthrough candidate regions in the ARTDeco pipeline were defined based on the longest transcript to exclude the effects of differential 3'-UTR usage. Genic and readthrough reads were counted using featureCounts from the Rsubread package in 10 kb bins downstream of genes, in a window between 0 to 100 kb. Readthrough regions with low coverage, in close proximity to, or overlapping, nearby protein-coding genes were excluded from the analysis. Intron retention was detected using the intron REtention Analysis and Detector (iREAD), which is a tool that detects retention events based on both splice junction reads and intron expression levels ([Li et al. 2020](#)). All samples were treated as unstranded.

Locus specific detection of transposon reads and transcripts

We used TELocal v1.1.1 ([mhammell-laboratory/TElocal](https://github.com/mhammell-laboratory/TElocal)) to quantify transposon reads, which is functionally similar to TETranscripts ([Jin et al. 2015](#)) except it provides locus-specific expression levels. Both tools use an expectation maximization algorithm to assign ambiguous reads. Annotation files were downloaded from the mhammell lab and from genecode (v39 for human and M27 for mouse). TELocal was also used to quantify transcriptomic counts. Again, all samples were treated as unstranded.

Expression and normalization of genes, introns, transposons and readthrough transcripts

Transcripts and elements with low expression counts were filtered. After filtering, intron, readthrough, read-in or transposon elements were combined with transcript counts for the analysis in DESeq2. For the aging fibroblast dataset age was normalized and centered and the design model considered age+gender+ethnicity. For the induced senescence dataset all four inducers were included together and the remaining treatment groups were considered non-senescent.

In order to study expression of these elements unbiased by gene coexpression, we selected the top 1000 overexpressed elements and divided their read counts by the read counts of the nearest gene. When not specified in the annotation the nearest genes or elements were

identified with the bedtools closest or intersect function, as appropriate. Finally, to provide a single expression value for each sample we used the mean log of the normalized expression values.

Expression of potentially active and protein-coding LINE-1 elements

We defined potentially active LINE-1 elements as those that have a functional promoter (McKerrow and Fenyö 2020). Whereas to define potentially protein-coding LINE-1 elements we predicted open reading frames for each annotated element using the predORF function from the systemPipeR package. Global pairwise alignment to ORF1p (uniprot: Q9UN81) and ORF2p (uniprot: O00370) was performed and only elements with a significant alignment were included for further analyses ($p < 0.05$).

Acknowledgements

We would like to thank Juliane Liepe, Marco Malavolta and Chin-Tong Ong for support and/or their help in improving this manuscript and VitaDAO for financial support.

References

Adusumalli S. , Ngian Z.-K. , Lin W.-Q. , Benoukraf T. , Ong C.-T. (2019) **Increased intron retention is a post-transcriptional signature associated with progressive aging and Alzheimer's disease** *Aging Cell* **18**
<https://doi.org/10.1111/acer.12928>

Agostini F. , Zagalak J. , Attig J. , Ule J. , Luscombe N. M. (2021) **Intergenic RNA mainly derives from nascent transcripts of known genes** *Genome Biology* **22**
<https://doi.org/10.1186/s13059-021-02350-x>

Bauer D. L. V. , Tellier M. , Martínez-Alonso M. , Nojima T. , Proudfoot N. J. , Murphy S. , Fodor E. (2018) **Influenza Virus Mounts a Two-Pronged Attack on Host RNA Polymerase II Transcription** *Cell Reports* **23**:2119–2129
<https://doi.org/10.1016/j.celrep.2018.04.047>

Chen S. , Zhou Y. , Chen Y. , Gu J. (2018) **fastp: An ultra-fast all-in-one FASTQ preprocessor** *Bioinformatics* **34**:i884–i890
<https://doi.org/10.1093/bioinformatics/bty560>

Colombo A. R. , Elias H. K. , Ramsingh G. (2018) **Senescence induction universally activates transposable element expression** *Cell Cycle* **17**:1846–1857
<https://doi.org/10.1080/15384101.2018.1502576>

David J. K. , Maden S. K. , Wood M. A. , Thompson R. F. , Nellore A. (2022) **Retained introns in long RNA-seq reads are not reliably detected in sample-matched short reads (p. 2022.03.11.484016)** *bioRxiv*.
<https://doi.org/10.1101/2022.03.11.484016>

De Cecco M. , Ito T. , Petrashen A. P. , Elias A. E. , Skvir N. J. , Criscione S. W. , Caligiana A. , Broccoli G. , Adney E. M. , Boeke J. D. , Le O. , Beauséjour C. , Ambati J. , Ambati K. , Simon M. , Seluanov A. , Gorbunova V. , Slagboom P. E. , Helfand S. L. , Sedivy J. M. (2019) **L1 drives IFN in senescent cells and promotes age-associated inflammation** *Nature* **566**

<https://doi.org/10.1038/s41586-018-0784-9>

Debès C. , Papadakis A. , Grönke S. , Karalay Ö. , Tain L. , Mizi A. , Nakamura S. , Hahn O. , Weigelt C. , Josipovic N. , Zirkel A. , Brusius I. , Sofiadis K. , Lamprousi M. , Lu Y.-X. , Huang W. , Esmailie R. , Kubacki T. , Späth M. R. , Beyer A. (2022) **Aging-associated changes in transcriptional elongation influence metazoan longevity (p. 719864)** *bioRxiv*.

<https://doi.org/10.1101/719864>

Deininger P. , Morales M. E. , White T. B. , Baddoo M. , Hedges D. J. , Servant G. , Srivastav S. , Smither M. E. , Concha M. , DeHaro D. L. , Flemington E. K. , Belancio V. P. (2017) **A comprehensive approach to expression of L1 loci** *Nucleic Acids Research* **45**

<https://doi.org/10.1093/nar/gkw1067>

Deniz Ö. , Frost J. M. , Branco M. R. (2019) **Regulation of transposable elements by DNA modifications** *Nature Reviews Genetics* **20**

<https://doi.org/10.1038/s41576-019-0106-6>

Dobin A. , Davis C. A. , Schlesinger F. , Drenkow J. , Zaleski C. , Jha S. , Batut P. , Chaisson M. , Gingeras T. R. (2013) **STAR: Ultrafast universal RNA-seq aligner** *Bioinformatics* **29**:15–21

<https://doi.org/10.1093/bioinformatics/bts635>

Fleischer J. G. , Schulte R. , Tsai H. H. , Tyagi S. , Ibarra A. , Shokhirev M. N. , Huang L. , Hetzer M. W. , Navlakha S. (2018) **Predicting age from the transcriptome of human dermal fibroblasts** *Genome Biology* **19**

<https://doi.org/10.1186/s13059-018-1599-6>

Fong N. , Brannan K. , Erickson B. , Kim H. , Cortazar M. A. , Sheridan R. M. , Nguyen T. , Karp S. , Bentley D. L. (2015) **Effects of Transcription Elongation Rate and Xrn2 Exonuclease Activity on RNA Polymerase II Termination Suggest Widespread Kinetic Competition**

Molecular Cell **60**:256–267

<https://doi.org/10.1016/j.molcel.2015.09.026>

Ghanam A. R. , Cao J. , Ouyang X. , Song X. (2019) **New Insights into Chronological Mobility of Retrotransposons In Vivo** *Oxidative Medicine and Cellular Longevity*, 2019

<https://doi.org/10.1155/2019/2818415>

Gorbunova V. , Seluanov A. , Mita P. , McKerrow W. , Fenyö D. , Boeke J. D. , Linker S. B. , Gage F. H. , Kreiling J. A. , Petrashen A. P. , Woodham T. A. , Taylor J. R. , Helfand S. L. , Sedivy J. M. (2021) **The role of retrotransposable elements in ageing and age-associated diseases**

Nature **596**

<https://doi.org/10.1038/s41586-021-03542-y>

Gualandi N. , Iperi C. , Esposito M. , Ansaloni F. , Gustincich S. , Sanges R. (2022) **Meta-Analysis Suggests That Intron Retention Can Affect Quantification of Transposable Elements from RNA-Seq Data** *Biology* **11**

<https://doi.org/10.3390/biology11060826>

Hadar S. , Meller A. , Saida N. , Shalgi R. (2022) **Stress-induced transcriptional readthrough into neighboring genes is linked to intron retention** *IScience* **25**
<https://doi.org/10.1016/j.isci.2022.105543>

Hahn O. , Grönke S. , Stubbs T. M. , Ficzb G. , Hendrich O. , Krueger F. , Andrews S. , Zhang Q. , Wakelam M. J. , Beyer A. , Reik W. , Partridge L. (2017) **Dietary restriction protects from age-associated DNA methylation and induces epigenetic reprogramming of lipid metabolism** *Genome Biology* **18**
<https://doi.org/10.1186/s13059-017-1187-1>

Jin Y. , Tam O. H. , Paniagua E. , Hammell M. (2015) **TEtranscripts: A package for including transposable elements in differential expression analysis of RNA-seq datasets** *Bioinformatics* **31**:3593–3599
<https://doi.org/10.1093/bioinformatics/btv422>

Kazazian H. H. , Moran J. V. (2017) **Mobile DNA in Health and Disease** *New England Journal of Medicine* **377**:361–370
<https://doi.org/10.1056/NEJMra1510092>

Khristich A. N. , Mirkin S. M. (2020) **On the wrong DNA track: Molecular mechanisms of repeat-mediated genome instability** *Journal of Biological Chemistry* **295**:4134–4170
<https://doi.org/10.1074/jbc.REV119.007678>

Koning A. P. J. de , Gu W. , Castoe T. A. , Batzer M. A. , Pollock D. D. (2011) **Repetitive Elements May Comprise Over Two-Thirds of the Human Genome** *PLOS Genetics* **7**
<https://doi.org/10.1371/journal.pgen.1002384>

LaRocca T. J. , Cavalier A. N. , Wahl D. (2020) **Repetitive elements as a transcriptomic marker of aging: Evidence in multiple datasets and models** *Aging Cell* **19**
<https://doi.org/10.1111/accel.13167>

Li H.-D. , Funk C. C. , Price N. D. (2020) **iREAD: A tool for intron retention detection from RNA-seq data** *BMC Genomics* **21**
<https://doi.org/10.1186/s12864-020-6541-0>

McKerrow W. , Fenyö D. (2020) **L1EM: A tool for accurate locus specific LINE-1 RNA quantification** *Bioinformatics* **36**:1167–1173
<https://doi.org/10.1093/bioinformatics/btz724>

Mouilleron H. , Delcourt V. , Roucou X. (2016) **Death of a dogma: Eukaryotic mRNAs can code for more than one protein** *Nucleic Acids Research* **44**:14–23
<https://doi.org/10.1093/nar/gkv1218>

Muniz L. , Deb M. K. , Aguirrebengoa M. , Lazorthes S. , Trouche D. , Nicolas E. (2017) **Control of Gene Expression in Senescence through Transcriptional Read-Through of Convergent Protein-Coding Genes** *Cell Reports* **21**:2433–2446
<https://doi.org/10.1016/j.celrep.2017.11.006>

Pabis K. (2021) **Triplex and other DNA motifs show motif-specific associations with mitochondrial DNA deletions and species lifespan** *Mechanisms of Ageing and Development* **194**

<https://doi.org/10.1016/j.mad.2021.111429>

(1970) **Retained introns in long RNA-seq reads are not reliably detected in sample-matched short reads**

Rodriguez-Martin B. , Alvarez E. G. , Baez-Ortega A. , Zamora J. , Supek F. , Demeulemeester J. , Santamarina M. , Ju Y. S. , Temes J. , Garcia-Souto D. , Detering H. , Li Y. , Rodriguez-Castro J. , Dueso-Barroso A. , Bruzos A. L. , Dentro S. C. , Blanco M. G. , Contino G. , Ardeljan D. , Tubio J. M. C. (2020) **Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition** *Nature Genetics* **52**

<https://doi.org/10.1038/s41588-019-0562-0>

Rosa-Mercado N. A. , Steitz J. A. (2022) **Who let the DoGs out? – Biogenesis of stress-induced readthrough transcripts** *Trends in Biochemical Sciences* **47**:206–217

<https://doi.org/10.1016/j.tibs.2021.08.003>

Rosa-Mercado N. A. , Zimmer J. T. , Apostolidi M. , Rinehart J. , Simon M. D. , Steitz J. A. (2021) **Hyperosmotic stress alters the RNA polymerase II interactome and induces readthrough transcription despite widespread transcriptional repression** *Molecular Cell* **81**:502–513

<https://doi.org/10.1016/j.molcel.2020.12.002>

Roth S. J. , Heinz S. , Benner C. (2020) **ARTDeco: Automatic readthrough transcription detection** *BMC Bioinformatics* **21**

<https://doi.org/10.1186/s12859-020-03551-0>

Scotti M. M. , Swanson M. S. (2016) **RNA mis-splicing in disease** *Nature Reviews Genetics* **17**

<https://doi.org/10.1038/nrg.2015.3>

Stow E. C. , Kaul T. , deHaro D. L. , Dem M. R. , Beletsky A. G. , Morales M. E. , Du Q. , LaRosa A. J. , Yang H. , Smither E. , Baddoo M. , Ungerleider N. , Deininger P. , Belancio V. P. (2021) **Organ-, sex- and age-dependent patterns of endogenous L1 mRNA expression at a single locus resolution** *Nucleic Acids Research* **49**:5813–5831

<https://doi.org/10.1093/nar/gkab369>

Vilborg A. , Sabath N. , Wiesel Y. , Nathans J. , Levy-Adam F. , Yario T. A. , Steitz J. A. , Shalgi R. (2017) **Comparative analysis reveals genomic features of stress-induced transcriptional readthrough** *Proceedings of the National Academy of Sciences* **114**:E8362–E8371

<https://doi.org/10.1073/pnas.1711120114>

Villanueva-Cañas J. L. , Horvath V. , Aguilera L. , González J. (2019) **Diverse families of transposable elements affect the transcriptional regulation of stress-response genes in *Drosophila melanogaster*** *Nucleic Acids Research* **47**:6842–6857

<https://doi.org/10.1093/nar/gkz490>

Wahl D. , Cavalier A. N. , Smith M. , Seals D. R. , LaRocca T. J. (2021) **Healthy Aging Interventions Reduce Repetitive Element Transcripts** *The Journals of Gerontology: Series A* **76**:805–810
<https://doi.org/10.1093/gerona/glaa302>

Wang K. , Wu D. , Zhang H. , Das A. , Basu M. , Malin J. , Cao K. , Hannenhalli S. (2018) **Comprehensive map of age-associated splicing changes across human tissues and their contributions to age-associated diseases** *Scientific Reports* **8**
<https://doi.org/10.1038/s41598-018-29086-2>

Wang Y. , Liu L. , Song Y. , Yu X. , Deng H. (2022) **Unveiling E2F4, TEAD1 and AP-1 as regulatory transcription factors of the replicative senescence program by multi-omics analysis** *Protein & Cell* **13**:742–759
<https://doi.org/10.1007/s13238-021-00894-z>

Wood J. G. , Jones B. C. , Jiang N. , Chang C. , Hosier S. , Wickremesinghe P. , Garcia M. , Hartnett D. A. , Burhenn L. , Neretti N. , Helfand S. L. (2016) **Chromatin-modifying genetic interventions suppress age-associated transposable element activation and extend life span in Drosophila** *Proceedings of the National Academy of Sciences* **113**:11277–11282
<https://doi.org/10.1073/pnas.1604621113>

Yao J. , Ding D. , Li X. , Shen T. , Fu H. , Zhong H. , Wei G. , Ni T. (2020) **Prevalent intron retention fine-tunes gene expression and contributes to cellular senescence** *Aging Cell* **19**
<https://doi.org/10.1111/acer.13276>

Zhang X. , Liu X. , Du Z. , Wei L. , Fang H. , Dong Q. , Niu J. , Li Y. , Gao J. , Zhang M. Q. , Xie W. , Wang X. (2021) **The loss of heterochromatin is associated with multiscale three-dimensional genome reorganization and aberrant transcription during cellular senescence** *Genome Research* **31**:1121–1135
<https://doi.org/10.1101/gr.275235.121>

Zheng J.-T. , Lin C.-X. , Fang Z.-Y. , Li H.-D. (2020) **Intron Retention as a Mode for RNA-Seq Data Analysis** *Frontiers in Genetics*

Author information

Kamil Pabis

Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Healthy Longevity Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Centre for Healthy Longevity, National University Health System

Diogo Barardo

Healthy Longevity Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Centre for Healthy Longevity, National University Health System

Kumar Selvarajoo

Bioinformatics Institute, Agency for Science, Technology and Research (A*STAR), Singapore, Singapore, Singapore Institute of Food and Biotechnology Innovation, Agency for Science, Technology and Research (A*STAR), Singapore, Singapore, NUS Synthetic Biology for Clinical and Technological Innovation (SynCTI), National University of Singapore, Singapore, Singapore, School of Biological Sciences, Nanyang Technological University, Singapore, Singapore
ORCID iD: [0000-0002-0314-9666](https://orcid.org/0000-0002-0314-9666)

Jan Gruber

Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Science Divisions, Yale-NUS College, Singapore, 138527, Singapore

Brian K. Kennedy

Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Healthy Longevity Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Centre for Healthy Longevity, National University Health System

For correspondence: bkennedy@nus.edu.sg

Editors

Reviewing Editor

Pablo Manavella

Universidad Nacional del Litoral-CONICET, Argentina

Senior Editor

Kathryn Cheah

University of Hong Kong, Hong Kong

Reviewer #1 (Public Review):

Transcriptional readthrough, intron retention, and transposon expression have been previously shown to be elevated in mammalian aging and senescence by multiple studies. The current manuscript claims that the increased intron retention and readthrough could completely explain the findings of elevated transposon expression seen in these conditions. To that end, they analyze multiple RNA-seq expression datasets of human aging, human senescence, and mouse aging, and establish a series of correlations between the overall expression of these three entities in all datasets.

While the findings are useful, the strength of the evidence is incomplete, as the individual analyses unfortunately do not support the claims. Specifically, to establish this claim there is a burden of proof on the authors to analyze both intron-by-intron and gene-by-gene, using internal matched regions, and, in addition, thoroughly quantify the extent of transcription of completely intergenic transposons and show that they do not contribute to the increase in aging/senescence. Furthermore, the authors chose to analyze the datasets as unstranded, even though strand information is crucial to their claim, as both introns and readthrough are stranded, and if there is causality, than opposite strand transposons should show no preferential increase in aging/senescence. Finally, there are some unclear figures that do not seem to show what the authors claim. Overall, the study is not convincing.

Major concerns:

1. Why were all datasets treated as unstranded? Strand information seems critical, and should not be discarded. Specifically, stranded information is crucial to increase the confidence in the causality claimed by the authors, since readthrough and intron retention are both strand specific, and therefore should influence only the same strand transposons and not the opposite-strand ones.
2. "Altogether this data suggests that intron retention contributes to the age-related increase in the expression of transposons" - this analysis doesn't demonstrate the claim. In order to prove this they need to show that transposons that are independent of introns are either negligible, or non-changing with age.
3. Additionally, the correct control regions should be intronic regions other than the transposon, which overall contributed to the read counts of the intron.
4. Furthermore, analysis of read spanning intron and partly transposons should more directly show this contribution.
5. "This contrasts with the almost completely even distribution of randomly permuted transposons." How was random permutation of transposons performed? Why is this control not trivial, and why is this a good control?
6. Fig 4: the choice to analyze only the 10kb-20kb region downstream to TSE for readthrough regions has probably reduced the number of regions substantially (there are only 200 left) and to what extent this faithfully represent the overall trend is unclear at this point.
7. Fig. 5B shows the opposite of the authors claims: in the control samples there are more transposon reads than in the KCl samples.
8. "induced readthrough led to preferential expression of gene proximal transposons (i.e. those within 25 kb of genes), when compared with senescence or aging". A convincing analysis would show if there is indeed preferential proximity of induced transposons to TSEs. Since readthrough transcription decays as a function of distance from TSEs, the expression of transposons should show the same trends if indeed simply caused by readthrough. Also, these should be compared to the extent of transposon expression (not induction) in intergenic regions without any readthrough, in these conditions.

Reviewer #2 (Public Review):

In this manuscript, the authors examined the role of transcription readout and intron retention in increasing transcription of transposable elements during aging in mammals. It is assumed that most transposable elements have lost the regulatory elements necessary for transcription activation. Using available RNA-seq datasets, the authors showed that an increase in intron retention and readthrough transcription during aging contributes to an increase in the number of transcripts containing transposable elements.

Previously, it was assumed that the activation of transposable elements during aging is a consequence of a gradual imbalance of transcriptional repression and a decrease in the functionality of heterochromatin (depression of transcription in heterochromatin). Therefore, this is an interesting study with important novel conclusion. However, there are many questions about bioinformatics analysis and the results obtained.

Major comments:

1. In Introduction the authors indicated that only small fraction of LINE-1 and SINE elements are expressed from functional promoters and most of LINE-1 are co-expressed with neighboring transcriptional units. What about other classes of mobile elements (LTR mobile element and transposons)?
2. Results: Why authors considered all classes of mobile elements together? It is likely that most of the LTR containing mobile elements and transposons contain active promoters that are repressed in heterochromatin or by KRAB-C2H2 proteins.
3. Fig. 2. A schematic model of transposon expression is not presented clearly. What is the purpose of showing three identical spliced transcripts?
4. The study analyzed the levels of RNA from cell cultures of human fibroblasts of different ages. The annotation to the dataset indicated that the cells were cultured and maintained. (The cells were cultured in high-glucose (4.5mg/ml) DMEM (Gibco) supplemented with 15% (vol/vol) fetal bovine serum (Gibco), 1X glutamax (Gibco), 1X non-essential amino acids (Gibco) and 1% (vol/vol) penicillin-streptomycin (Gibco). How correct that gene expression levels in cell cultures are the same as in body cells? In cell cultures, transcription is optimized for efficient division and is very different from that of cells in the body. In order to correlate a result on cells with an organism, there must be rigorous evidence that the transcriptomes match.
5. The results obtained for isolated cultures of fibroblasts are transferred to the whole organism, which has not been verified. The conclusions should be more accurate.
6. The full pipeline with all the configuration files IS NOT available on github (pabisk/aging_transposons).
7. Analysis of transcripts passing through repeating regions is a complex matter. There is always a high probability of incorrect mapping of multi-reads to the genome. Things worsen if unpaired short reads are used, as in the study (L=51). Therefore, the authors used the Expectation maximization algorithm to quantify transposon reads. Such an option is possible. But it is necessary to indicate how statistically reliable the calculated levels are. It would be nice to make a similar comparison of TE levels using only unique reads. The density of reads would drop, but in this case it would be possible to avoid the artifacts of the EM algorithm.