


RESEARCH ARTICLE

Transposable element landscapes in aging *Drosophila*

Nachen Yang¹, Satyam P. Srivastav¹[✉], Reazur Rahman²[✉], Qicheng Ma¹[✉], Gargi Dayama¹[✉], Sizheng Li¹, Madoka Chinen³[✉], Elissa P. Lei³[✉], Michael Rosbash², Nelson C. Lau^{1,4*}

1 Boston University School of Medicine, Department of Biochemistry, Boston, Massachusetts, United States of America, **2** Brandeis University, Department of Biology and Howard Hughes Medical Institute, Waltham, Massachusetts, United States of America, **3** Nuclear Organization and Gene Expression Section, NIDDK, NIH, Bethesda, Maryland, United States of America, **4** Boston University Genome Science Institute, Boston, Massachusetts, United States of America

 These authors contributed equally to this work.

 Current address: Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York, United States of America

* nclau@bu.edu


 OPEN ACCESS

Citation: Yang N, Srivastav SP, Rahman R, Ma Q, Dayama G, Li S, et al. (2022) Transposable element landscapes in aging *Drosophila*. PLoS Genet 18(3): e1010024. <https://doi.org/10.1371/journal.pgen.1010024>

Editor: Justin P. Blumenstiel, University of Kansas, UNITED STATES

Received: October 7, 2021

Accepted: January 10, 2022

Published: March 3, 2022

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting Information](#) files. The genomic sequencing data can be accessed from the NCBI SRA as Study Bioproject PRJNA678741.

Funding: This work was supported by NIH grants R01-AG052465 to NCL and a subcontract to MR, and R21-HD088792 and R01-GM135215 to NCL; and Intramural Program of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health (DK015602 to EPL). The funders had no role in study design, data

Abstract

Genetic mechanisms that repress transposable elements (TEs) in young animals decline during aging, as reflected by increased TE expression in aged animals. Does increased TE expression during aging lead to more genomic TE copies in older animals? To address this question, we quantified TE Landscapes (TLs) via whole genome sequencing of young and aged *Drosophila* strains of wild-type and mutant backgrounds. We quantified TLs in whole flies and dissected brains and validated the feasibility of our approach in detecting new TE insertions in aging *Drosophila* genomes when small RNA and RNA interference (RNAi) pathways are compromised. We also describe improved sequencing methods to quantify extra-chromosomal DNA circles (eccDNAs) in *Drosophila* as an additional source of TE copies that accumulate during aging. Lastly, to combat the natural progression of aging-associated TE expression, we show that knocking down *PAF1*, a conserved transcription elongation factor that antagonizes RNAi pathways, may bolster suppression of TEs during aging and extend lifespan. Our study suggests that in addition to a possible influence by different genetic backgrounds, small RNA and RNAi mechanisms may mitigate genomic TL expansion despite the increase in TE transcripts during aging.

Author summary

Transposable elements, also called transposons, are genetic parasites found in all animal genomes. Normally, transposons are compacted away in silent chromatin in young animals. But, as animals age and transposon-silencing defense mechanisms break down, transposon RNAs accumulate to significant levels in old animals like fruit flies. An open question is whether the increased levels of transposon RNAs in older animals also correspond to increased genomic copies of transposons. This study approached this question

collection and analysis, decision to publish, or preparation of the manuscript.

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

by sequencing the whole genomes of young and old wild-type and mutant flies lacking a functional RNA interference (RNAi) pathway, which naturally silences transposon RNAs. Although the wild-type flies with intact RNAi activity had little new accumulation of transposon copies, the sequencing approach was able to detect several transposon accumulation occurrences in some RNAi mutants. In addition, we found that some fly transposon families can also accumulate as extra-chromosomal circular DNA copies. Lastly, we showed that genetically augmenting the expression of RNAi factors can counteract the rising transposon RNA levels in aging and promote longevity. This study improves our understanding of the animal host genome relationship with transposons during natural aging processes.

Introduction

All animal genomes carry the genetic burden of a sizeable reservoir of parasitic elements called transposons or transposable elements (TEs). This TE burden can range from the extreme >70% proportion of the axolotl genome [1,2] to >50% in the human genome [3] to >10% in the *Drosophila melanogaster* genome [4,5]. TEs are selfish invaders of animal genomes with some potential for stimulating more rapid gene regulatory innovations like serving as novel enhancers [6], but more frequently are detrimental to animal fitness when they insert into and disrupt expression of important genes [7]. Therefore, conserved chromatin regulation and RNA-interference (RNAi) pathways must silence TEs to ensure fertility and animal health. However, these genomic defense mechanisms also weaken during animal aging concomitant with observable decreases in genomic integrity in aging cells. This phenomenon has been articulated in the hypothesis of TEs impacting aging as proposed by Gorbunova et al [8], where the decline of genome homeostasis during aging may unleash the detrimental effects of reanimating TE activity.

Initial support for this hypothesis in the model organism *D. melanogaster* came from studies of TE expression increasing in aging flies [9–12]. For example, mutants in chromatin silencing factors and RNAi pathway genes which repress TEs have reduced lifespans [9,10,13–16], whereas dietary restriction and overexpressing the RNAi and chromatin factors can limit TE expression and promote longevity [13,14]. Neurodegeneration modeled in aging flies through overexpressing aggregating proteins like TDP-43 and TAU also leads to elevated TE expression [17–20]. Additionally, there is evidence of a somatic population of Piwi proteins which can serve an additional TE defense mechanism that when mutated leads to shorter lifespan and loss of stem cell maintenance [13,16,21–23].

Beyond flies, mammals also must repress TEs for critical development of germ cells, embryos and neurons. Mammals have a complex, interconnected network of silencing pathways like the axis of SETDB1 [24,25], KAP1 [26–28] and the HUSH complex [29–32]; and its cooperation with histone deacetylases like SIRT6 [33,34] and histone methyltransferases like Suv39h1 and G9A [35–37]. In addition, there are DNA methyltransferases that genetically interact with the piRNA pathway to target TEs for chromatin silencing in mammalian germ cells [38–44]. Although >45% of mammalian genomes are comprised of many TE repeats, the vast majority are inactive with mainly *LINE-1/L1* implicated in somatic genome mosaicism in developing brains and individual neurons [45–53]. *LINE-1/L1* is linked to deleterious novel mutations in tumors and they are activated in cell culture models of cellular senescence [54–59]. Although TE control is clearly important to mammalian health, the large genome sizes and longer lifespans hamper comprehensive assessments of mammalian TEs during aging.

Therefore, in this study we leveraged *Drosophila's* rapid aging, its compact genome and powerful genetic tools as significant advantages for characterizing how TEs may change during normal animal aging. An important goal of our study is to address the debate of whether TEs quantitated from Whole Genome Sequencing (WGS) of *Drosophila* genomes represent true gains in TE genomic load [60]. One bioinformatics program called TEMP [61] has been used extensively in determining TE insertions from *Drosophila* WGS [62,63] but its capacity to distinguish bona fide TE insertions from potential library sequencing artifacts has been re-examined [60]. Noting the high degree of variability in TE insertion calls from various bioinformatics programs applied to *Drosophila* WGS data [64], we therefore developed our own program called the Transposon Insertion & Depletion AnaLyzEr (TIDAL) to identify the tremendous diversity of TEs across various *Drosophila* strains [65]. TIDAL's increased specificity in TE determinations comes from requiring sequencing reads mapping to both sides of genomic locus flanking the TE insertion, with a threshold minimum of 4 reads (i.e. ~2 reads on each TE insertion junction). This specificity was benchmarked against genomic PCR tests [65], and TIDAL has characterized TEs in other *Drosophila* studies of genetic factors regulating TE silencing [66,67].

In this study, we demonstrate how WGS and extrachromosomal circular DNA (eccDNA) sequencing of aged and young flies can report changes in TEs during fly aging particularly in sensitized genetic backgrounds such as in some RNAi mutants. Although TE RNA upregulation is a recurring phenotype of aging wild-type flies, we show that genomic TEs can expand in some strains defective for RNA silencing, either because of their unique genetic backgrounds or because disrupted RNA silencing now allows TEs to expand their genomic DNA (gDNA) copy numbers. We also measure genomic TEs by tissue-specific (i.e. fly brain tissues) gDNA sequencing; and eccDNA accumulation during aging of the *ISO1* strain as an additional feature of the hypothesis of TEs impacting animal aging. Lastly, we show that genetically boosting RNAi activity in aged flies via knockdown of *PAF1* can suppress TE RNAs and extend longevity. Together, these results suggest that, in addition to possible influences of genetic background, the RNAi pathway may buffer genomic alterations by the natural increase of TE RNAs during aging and suggest *PAF1* inhibition in aging animals could be a new target for genetic suppression of TE expression.

Results

Recurring increase of TE RNA expression during fly aging

Although previous studies using certain control wild-type (WT) fly strains showed that TE RNAs were upregulated in aged flies [10,14], we decided to reconfirm this observation for three commonly used WT fly strains that would form the basis of this study. Using our lab's standard rearing conditions, we first determined the aging curves for the *ISO1* strain used for the *D. melanogaster* reference genome sequence [4], an isogenic *w1118* strain that is a common background strain in genetic studies [68], and the *Oregon-R* strain used in a series of functional genomics datasets [69]. We established that these three strains displayed lifespans typical of other WT fly strains (Fig 1A).

We then followed the experimental convention of other studies [10,14] to standardize the comparison of 30-day aged adults versus 5-day young adults, and we performed quantitative RT-PCR on a panel of TEs from total RNAs from females (Fig 1B). We replicated many examples of TE RNAs being upregulated in the aged WT flies' whole bodies but noticed variability in which specific TE families were the most significantly upregulated during aging. For example, *gypsy*, *mdg1*, and *I-element* were upregulated at the RNA level in *ISO1* aged flies, while *copia* and *1731* RNAs were upregulated in *w1118* and *OreR* (Wilcoxon rank sum test,

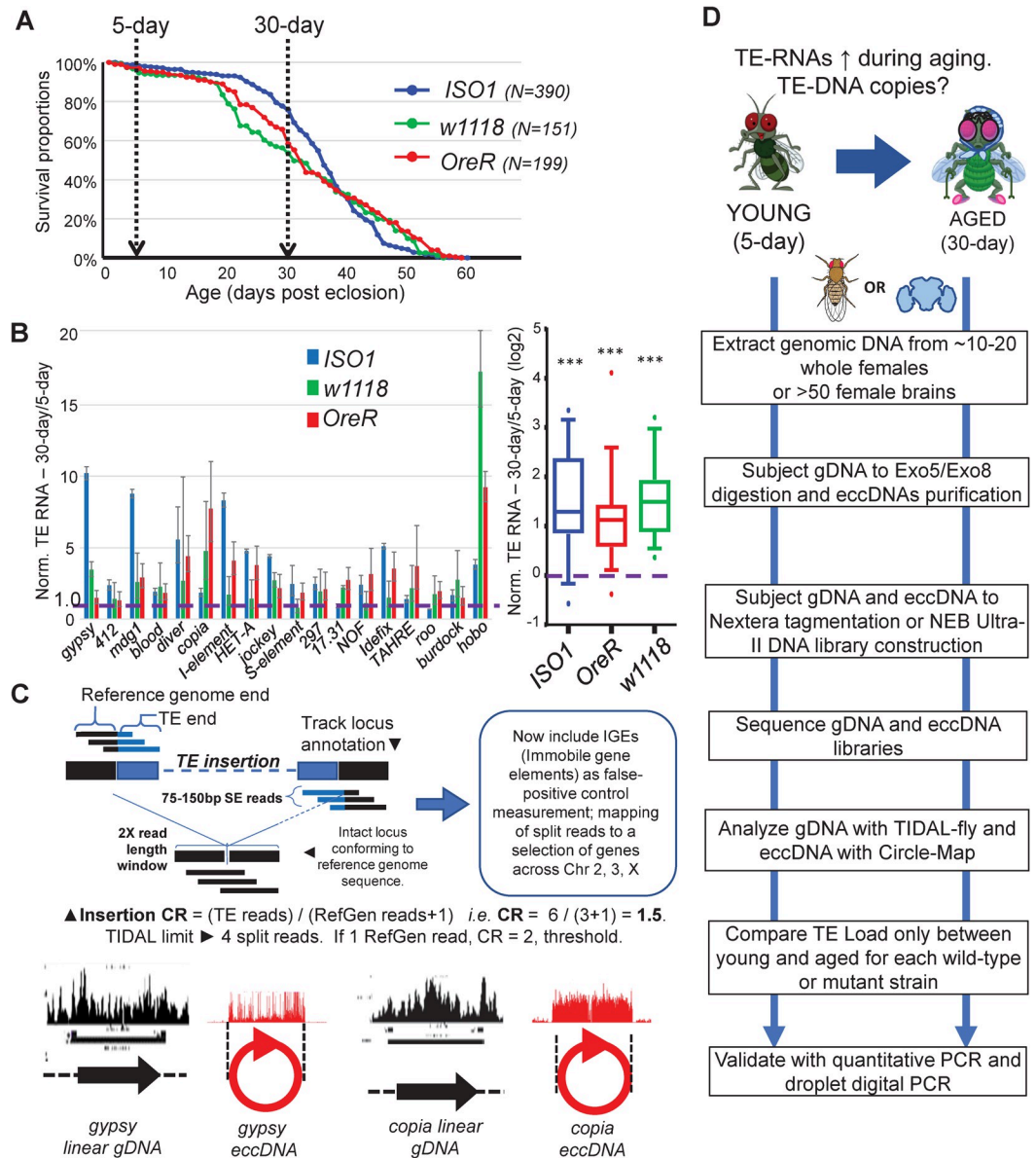


Fig 1. Overview of study to examine whether TE-DNA copy numbers change during fly aging. (A) Survival curves of the three wild-type (WT) fly strains carried out in this study, indicating the selection of 30-day adults as a representative timepoint of aging onset (B) Validation of TE transcript expression increases during fly aging through RT-qPCR of TE RNAs normalized to *Rp49* transcripts from whole female fly bodies. Error bars are propagated standard deviations of delta-CT values from three replicates. Boxplot on the right summarizes the data on the left plot and tests the statistical significance of TE RNA up-regulation by the Wilcoxon test in all three WT strains, p -value < 0.0001. (C) Overview of TE detection strategy from Whole Genome Sequencing (WGS) data using updated TIDAL and extra-chromosomal circular DNA (eccDNAs) detection scripts. The coverage plots for WGS and eccDNA data are illustrated in black diagrams and red diagrams, respectively. (D) Study designed for comparing TE load between 5-day young and 30-day aged flies within each WT and mutant strain.

<https://doi.org/10.1371/journal.pgen.1010024.g001>

$p < 0.001$, Fig 1B). This variability may reflect the inherently distinct TLs between these three strains [65], but the trend holds true that WT adult flies recurrently experience increased TE expression during aging.

Only one recent study we are aware of assessed TLs during fly aging by WGS of enriched $\alpha\beta$ -Kenyon Cell neurons [60] and which argued that various pitfalls obscured the ability to

observe TL increases during fly aging. For example, the study itself discussed that Multiple Displacement Amplification (MDA) required to amplify the minute amount of neuronal gDNA prior to Illumina library construction could contribute to artifactual chimeric molecules that represent false positive TE insertions [60]. Therefore, our more comprehensive effort to examine TLs through direct WGS should add valuable insight to this question.

When determining TL from *Drosophila* WGS datasets, we first considered how two different TE-insertion discovery programs, TEMP [61] and TIDAL [65] can each yield different results from analyses of the same dataset (S1 Fig). Balancing sensitivity against specificity, TIDAL has similar trends as TEMP in revealing the diversity of TLs amongst *Drosophila* samples (S1B Fig) and both are effective at calling germline insertions, but TIDAL avoids false positive predictions that others have contended as somatic TE insertions by requiring read support on both sides of the TE insertion junction (Fig 1C versus S1C Fig, [60]). TIDAL handles this issue differently by computing a Coverage Ratio (CR) score for each TE insertion from pooled sequencing of a small group of flies (Fig 1C), where TE insertion reads are divided by reference genome mapping reads and a pseudocount of 1; such as a CR of 2 that we used as an arbitrary cutoff for indicating deep penetrance of a TE insertion at a given insertion locus [65].

Measuring TE landscapes (TL) by direct WGS and the TIDAL program

To meaningfully compare TL changes during a single generation of aging flies from WGS and to avoid the genomic complications of normalizing against Y-chromosome reads that are exceptionally dense with repeats [70], we only compared samples from within the same strain in small numbers of young versus aged whole female adult flies or female brains (Fig 1D). In our process we extracted a set amount of genomic DNA from 10 flies that allowed for reproducible WGS library construction without requiring MDA or other total DNA amplification methods before library preparation. We then sequenced on the Illumina platform each fly strains' bulk gDNA library to a minimum $> \sim 30$ million 75-bp reads for $> \sim 16X$ fold genomic coverage of the version Dm6 genome assembly (S1 Table). Each library was analyzed identically with the TIDAL program [65] and new TE insertions were counted individually and then normalized against the reads per million measurement to account for sequencing depth differences.

In developing our methodology to examine fly TLs during aging, we recall our previous study showing that each fly strain's unique TLs depends on how inherently distinct its genetic background is from the reference genome strain *ISO1* [65]. Therefore, it was expected that new TE insertions quantified and normalized against each library's sequencing depth would yield the lowest numbers for *ISO1* and the most TE insertion differences in *w1118* and *OreR* (Fig 2A). In order to make equivalent TL comparisons between young and aged fly libraries, the deeper WGS library was downsampled to the sequencing depth where mapped reads were approximately the same as the lower cognate library before TIDAL analysis. We then assessed how reproducible were TL determination with TIDAL by conducting 100 randomized permutations from 50% downsampled *OreR* 5-day libraries to bootstrap TE insertion number calls (Figs 2B, 2C and S2A). The cumulative average number and standard deviation of TE insertion calls after 100 permutations were already reached by 20 permutations, so to economize the demanding computational resources, we applied a 20-permutation bootstrap to all the other WGS libraries reported in the main figures of this study to provide narrow 95% confidence intervals for TL determinations by TIDAL (S2D, S2E and S2F Fig). Finally, we re-analyzed the WGS datasets from [60] with TIDAL, and found that even at the single fly or single neuron level, it is challenging to detect significant TL changes in wild types (S2H and S2I Fig).

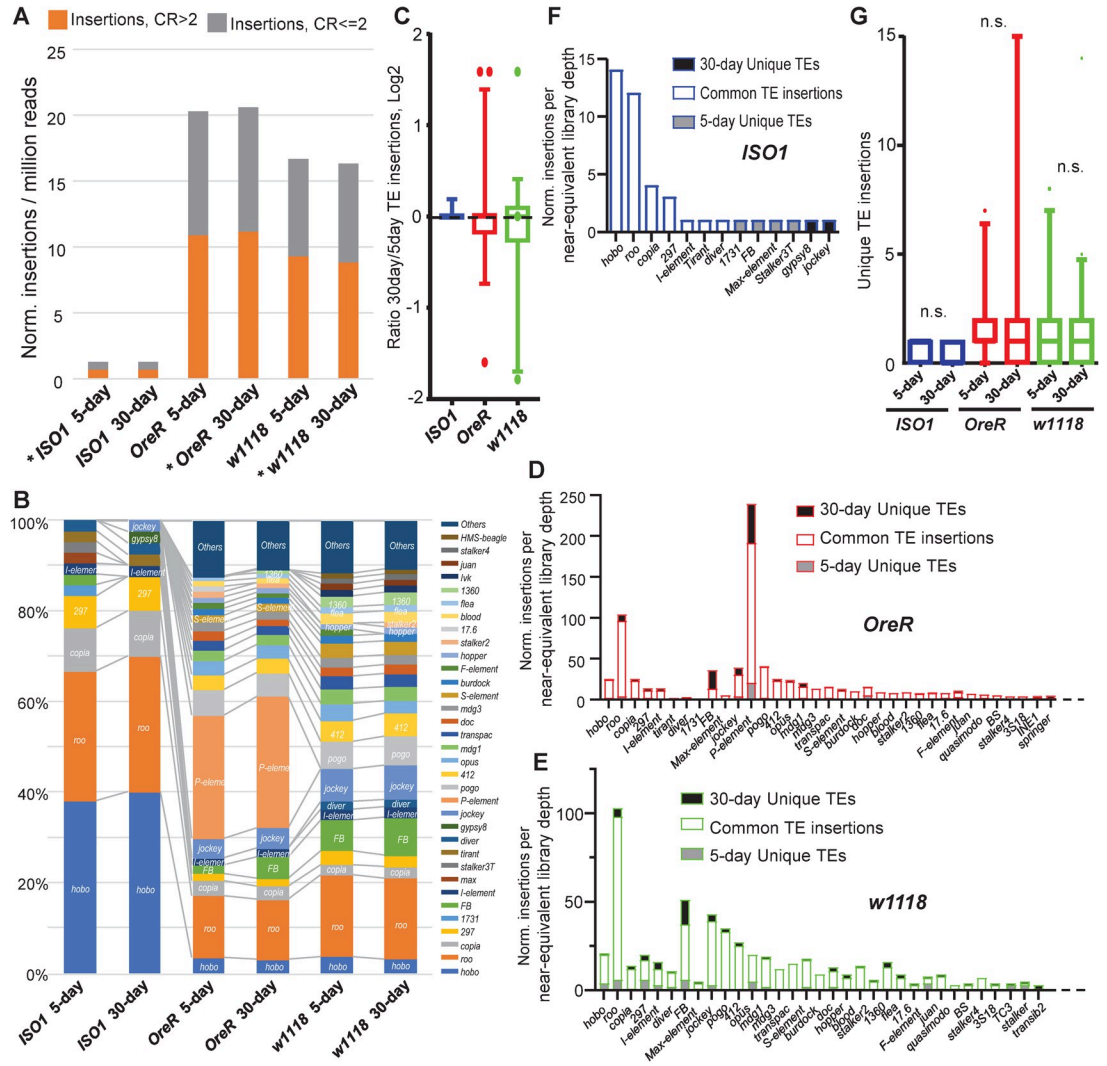


Fig 2. WGS analysis of TE insertion numbers between 5-day young versus 30-day aged wild-type fly strains. (A) Quantification of new TE insertions as compared to the reference genome using the TIDAL-fly program. Categories of total TE insertions broken by the Coverage Ratios (CR) of CR>2 and CR<= 2. (B) Within each strain, TE families' percentages are ordered by the color legend. (C) Ratios of the 30-day versus 5-day of normalized TE insertions from panels D-F. (D-F) Number of unique TE insertions (filled bars) present in 5-day and 30-day relative to common insertions present in both samples (open bar) of *OreR*, *w1118* and *ISO1* fly strains. These panels display only the TE families that were detected by TIDAL be at least 1% of total number of TE families (*i.e.* all the TEs not lumped into the "Others" category of Fig 2B). (G) Boxplot for conducting paired Wilcoxon tests does not show statistical significance for the gain of unique TE insertions in aging WT flies. n.s. = not statistically significant.

<https://doi.org/10.1371/journal.pgen.1010024.g002>

As expected, each of these WT flies TLs displayed completely distinct compositions of new insertions of TE families relative to the Dm6 reference genome sequence (Fig 2B), such as a larger proportion of *hobo* TEs in *ISO1*, major infiltration of *P-elements* in *OreR*, and several more *FB*, *pogo* and *412* TEs in *w1118*. Focusing on the ratio of 30-day to 5-day insertions for the specific TE families making up the bulk of these strains TLs, we could observe most of the TEs remained unchanged during aging whilst a few TEs have some change in copy numbers during aging (outlier dots in Fig 2C representing a few TE families). This was reflected at the total TL level with modest or few TE insertions in 30-day aged *ISO1* flies versus 5-day young flies, with also perplexing total decreases in *w1118* flies (Fig 2A). Since *OreR* and *w1118* have

distinct genetic backgrounds from the *ISO1* reference genome, we further classified the total TE insertions detected by TIDAL to differentiate the “unique” from “common” TE insertions by comparing the genomic location of TE insertions between each 5-day and 30-day sample pairs. When insertions with the same genomic locus are present in both 5-day and 30-day samples, these are noted as “common insertions”; and those TE loci that are present only in one condition but not the other are classified as “unique insertions”. Although the vast majority of the TE insertions were commonly detected by TIDAL in both 5-day and 30-day *w1118* and *OreR* flies (Fig 2D and 2E), there were more TE insertions only detected in these 30-day aged flies than the 5-day young fly genomes. Only a few *hobo* insertions were also only seen in 5-day young *ISO1* flies and were no longer detected in 30-day aged flies (Fig 2F). This observation can be explained by this analysis that only focuses on TE insertion counts as quantile samplings of reads discordant from the reference genome.

New somatic transposition event in a small subset of cells could be overshadowed by a background of unmodified reference sequences and could explain a TE with a low CR score that is sampled in 5-day fly gDNA sample but then missed in the 30-day sample. This is a known limitation of the WGS approach and sacrificing sensitivity to improve specificity in the original TIDAL program [65], resulting in insufficient statistical significance to detect TL increases in WT whole flies (Fig 2G). However, we could dismiss the concern of possible skews in chromosomal structures amongst the fly genomes sequenced in this study by showing predominantly diploid sequence coverage in the FREEC plots that TIDAL automatically generates (S3 Fig) [65].

To tackle the potential limitation of missing reads unmappable to the euchromatic genome sequence, we updated TIDAL to also map to *Drosophila* TE families consensus sequence coverage, and added arbitrarily-selected protein coding genes, analogous to the modification to TEMP to track protein-coding genes as Immobile Gene Elements (IGEs) [60]. We gauged a relatively low average rate of ~12% of IGEs across the libraries sequenced as defined by percentage of false positive TIDAL counts from IGE versus the total number of TIDAL counts [TE+IGE] (S1 Table), whereas these protein coding genes sequencing coverage generally also remained stable between 5-day young and 30-day aged flies (S4A Fig). Tracking TE consensus sequence coverage has the advantage of accounting for all accumulating TE sequences in both the mappable and unassembled and ambiguous-mapping regions of the genome. With this analysis approach, we could detect some change of total TE sequence coverage in *OreR* and *w1118* 30-day aged flies versus 5-day young flies (S4A and S4B Fig, p -value <0.01, Wilcoxon rank-sum tests) whilst our stable protein-coding gene coverage control for these two strains remains the same (S4A and S4B Fig). For *ISO1*, the total TE sequence coverage between 5-day versus 30-day remains the same however there is a slight reduction of protein coding genes coverage in 30-day aged flies (p -value <0.001). However, with these being single-batch sequencing runs, the caveat of more future sequencing experiments may be needed to bolster this result. In addition, both this coverage analysis and the quantile insertions analysis cannot discriminate between a full-length or truncated TE sequence, which we have noted in *P-elements* that can have critically variable transposition activities [71].

Resolving and validating our approaches measuring TE landscapes with RNAi mutants

This unresolved genomics challenge of using short read WGS data for analyzing TE sequences coverage also extends to some limitations in using droplet digital PCR (ddPCR) to precisely quantify genomic TE copies for only the isoforms covered by the short ddPCR amplicons [45] and little change is detected by digital PCR as well (S4C Fig). In questioning the accuracy of

this ddPCR assay in absolute quantification of TE copies, we compared ddPCR results on *P-elements* and *I-elements* versus WGS and TIDAL determinations in two other directly matched gDNA samples (S5A and S5B Fig). The ddPCR copy number measurements were very similar to the WGS and TIDAL determinations, indicating both methodologies are consistent with each other in the quantifications. Furthermore, we replicated a previously reported genetic cross [72] that in one format triggers a large burst of *I-element* transposition in the embryos but in a second format maintains *I-element* silencing (S5C Fig). We reanalyzed the WGS datasets from [72] with TIDAL reporting 505 new *I-element* copies versus the 3732 insertions called by TEMP in that study, with our ddPCR results leaning closer to the TIDAL count (1590 copies, S5D and S5E Fig). These data reaffirm the findings from [72] that the oocyte is the critical battleground between the host and the selfish genetic element.

To explain why aging-associated TL changes seemed muted or were challenging to detect in WT fly strains, we considered two competing hypotheses: (1) non-penetrant TE insertions are masked by multiple unmodified genomic loci within the pools of sequences imposing limitations in WGS and TIDAL analysis versus (2) WT flies retain RNAi defenses like TE-targeting siRNAs [21,22,73–75] and piRNAs [13,76–78] to restrain further TE RNAs accumulating and contributing to genomic transposition. To test these hypotheses, we collected the same 5-day and 30-day aging cohorts from three sets of different mutants in the two main arms of the RNAi pathway in *Drosophila* (Fig 3). We analyzed two independent mutants each in the *piwi*, *aubergine*, and *AGO2* genes and conducted the same whole flies WGS and TIDAL analysis as the WT strains.

Both *aubergine* (*aub*) mutant alleles displayed the greatest number of TE copy increases during aging (Fig 3B, 3E and 3H and S1 Table), with both high penetrance insertions (orange bars, CR>2) as well as less penetrant insertions (grey bars, CR = 2). However, the degree of aged associated total TE copy increases in *piwi* and *AGO2* mutants was more variable between alleles (Fig 3A and 3C). Although various individual TE families appeared to increase in copy number during aging (Fig 3D, 3E and 3F), when the entire TL was considered, then only one of the two alleles displayed a statistically significant increase of TE insertions unique to the aged samples (Fig 3G, 3H and 3I). Although we attribute these TL changes to the severe lack of RNAi suppression, we cannot rule out a potential contribution by the unique genetic background effects in these denoted alleles.

Measuring TL changes in fly brains during aging

Can new TE insertions also be detected in specific tissues where cells that are more permanent and not turned over as frequently, such as the brain? For example, in mammalian neurons, the most active TE *LINE-1/L1* has been implicated in transposing relatively frequently during development to give rise to genomic mosaicism in the brain [45–53]. Given the caveats of having to do prior total DNA amplification from limited gDNA from fly neurons [60], we undertook WGS from at least 50 dissected female brains to provide sufficient nucleic acid for RT-PCR confirmation of neuronal gene expression and WGS of brain DNA (Fig 4).

We successfully generated libraries directly from brains of WT fly strains and *piwi* and *AGO2* mutants without any prior total DNA amplification, and after setting libraries to sequencing depths with equivalent mapped reads by downsampling, we could detect some increases in specific TE families from *OreR* during aging (Fig 4B and 4E). Although there may be piRNA-like small RNAs and *piwi* expression in fly heads [13,21,22], we detected increases in TLs in the *piwi*(*HDRmCherry*) mutant's brains that were similar in magnitude to the WT *OreR* strain (Fig 4C and 4F, $p < 0.01$ and $p < 0.001$ respectively). Additional increases in total TLs were also detectable in the *AGO2* mutants' brains (Fig 4D, 4G and 4J), but with so few

brain samples and rather low increase of TEs in the *piwi(g1)* mutant, this rendered the aggregate test of all the RNAi mutants' brains as not statistically significant enough to ascribe an aging-related TE landscape increase like in whole flies (Fig 4K).

Although we cannot rule out the possible influence of unique genetic backgrounds in these RNAi mutants, we attribute the different TL changes with different penetrance levels to the separate germline versus somatic roles of each RNAi mutant gene. Because both *aubergine* and one of the AGO2 and *piwi* mutants displayed a higher total TE number in aged flies compared to young flies (Fig 3), we suggest that RNAi can normally repress TE transcripts and might

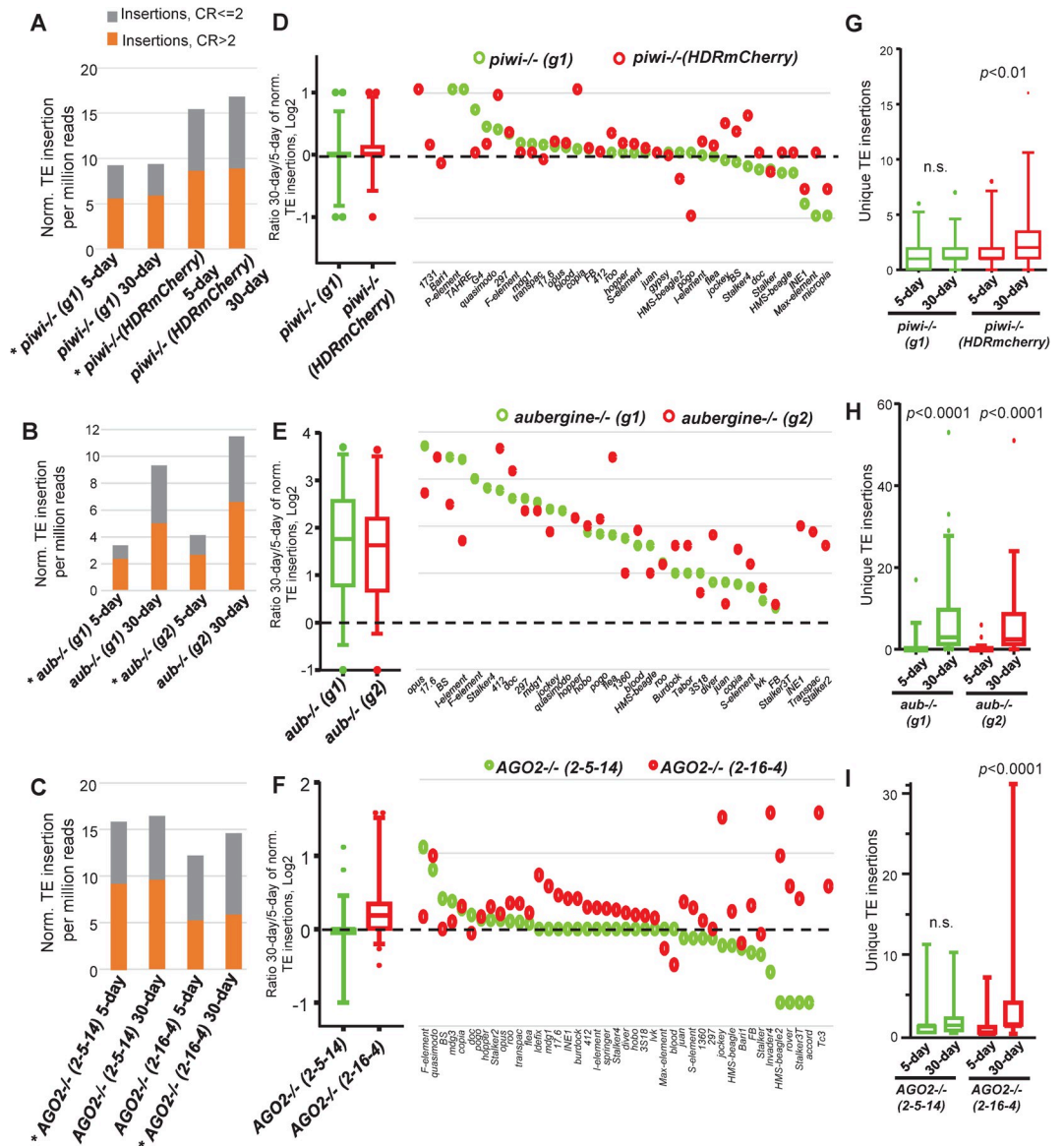


Fig 3. WGS analysis of TE insertion numbers between 5-day young versus 30-day aged RNAi mutant fly strains. (A-C) Quantification of new TE insertions as compared to the reference genome using the TIDAL program. Categories of total TE insertions broken by the coverage ratios (CR) of CR>2 and CR< = 2. Asterisks mark the library that was downsampled to the equivalent depth of the cognate comparison library. (D-F) Left, 5–95 percentile boxplots of ratios of the 30-day versus 5-day of normalized TE insertions from panels A-C. Right, dot plots display only the TE families with significant changes between the two timepoints in at least one of the mutant strains. (G-I) Boxplots for conducting paired Wilcoxon tests for statistical significance in the gain of unique TE insertions in aging RNAi mutant flies. *p*-value from Wilcoxon tests, n.s. = not statistically significant.

<https://doi.org/10.1371/journal.pgen.1010024.g003>

also limit genomic damage during aging. Furthermore, our data show that a direct WGS approach that obviates prior whole DNA amplification can still detect TL changes, albeit only in sensitized genetic backgrounds like the RNAi mutants such as in *aubergine* where the effects may come from ovaries in particular (Fig 3B, 3E and 3H).

Extra-chromosomal circular DNAs (eccDNAs) as an additional genomic cache of increasing TE sequences

In normal and diseased animal cells, there is a cache of eccDNAs that has recently been explored by deep sequencing of DNA that is resistant to extensive exonuclease digestion [79–82]. In certain tumor samples, eccDNAs are implicated in rapid copy-number expansion of oncogenes [83], while ectopic accumulation of DNA in the cytoplasm of senescing cells might trigger aging-associated inflammation responses [84]. Several earlier studies had also found evidence of eccDNAs in *Drosophila*, with the *copia* TE as a prominent example accumulating in certain strains [85–89]. Lastly, eccDNA enriched in TE sequences and other repeats were detected in normal plants and gDNA of human tissues [79,90], which in both of these studies required total DNA amplification prior to library construction to enrich the surviving eccDNAs after exonuclease digestion.

We investigated eccDNAs in *Drosophila* by optimizing our own method to purify enough eccDNAs to directly generate libraries for deep sequencing without requiring prior total DNA amplification (Fig 5A). Furthermore, we used spike-ins of cloning-vector plasmid DNAs into gDNA preparations and magnetic beads for improved recovery and quantitation of eccDNAs for comparing between different samples. To confirm that eccDNA was recovered after two rounds of Exo5 and Exo8 exonuclease digestion steps which only degrade linear but not circular DNA, we conducted PCR with standard primers amplifying linear genes and TEs (F1-R1 primer pairs, S2 Table), and outward-facing primers that either generate an amplicon from a TE eccDNA or tandem genomic copies of the same TE (P10-P11 primer pairs) (Fig 5B). Linear gene amplicons were significantly depleted after exonuclease digestions, while the amplicon for the spike-in plasmid was maintained. Linear TE amplicons were also reduced while eccDNA-targeted amplicons for the *copia* TE was resilient against the exonuclease treatment. Some other TE amplicons with outward-facing primers that were reduced after exonuclease treatment may reflect more tandem copies of these TEs.

Since the regular PCR amplicons for the *copia* eccDNA were readily apparent in WT strains (Fig 5B), we used qPCR to quantify the changes and show that *copia* eccDNA copies were increased >~2-fold in 30-day aged flies compared to 5-day young flies (Fig 5C). This result motivated us to deeply sequence short read libraries generated directly from those eccDNA-enriched samples which did not undergo any total DNA amplification (S3 Table). After confirming from a genome browser view the enrichment of *copia* , *gypsy* , satellite repeats and mitochondrial eccDNA enrichment from these libraries (S6A–S6D Fig), we adapted the TIDAL scripts of mapping reads to the TE families consensus sequences to measure sequencing coverage as well as circular junction spanning reads against *copia* and observed an aging-associated increase in *copia* eccDNA that was consistent with our qPCR results (Fig 5D). We also applied this custom eccDNA quantitation pipeline to all the other *Drosophila* TEs as well as adapting the CIRCLE-Map pipeline previously used to measure mammalian eccDNAs [79] to the *Drosophila* TEs. We then normalized the ratios of the eccDNA-TE counts between 30-day aged and 5-day young flies (Fig 5E). Although the CIRCLE-Map pipeline was more sophisticated at providing a significance “circle score” that we set the cutoff to be >50, our custom eccDNA quantitation pipeline’s results were notably consistent in showing overall that most eccDNAs as TEs were increasing in the libraries of 30-day aged flies (Figs 5F, 5G, S6G and S6H).

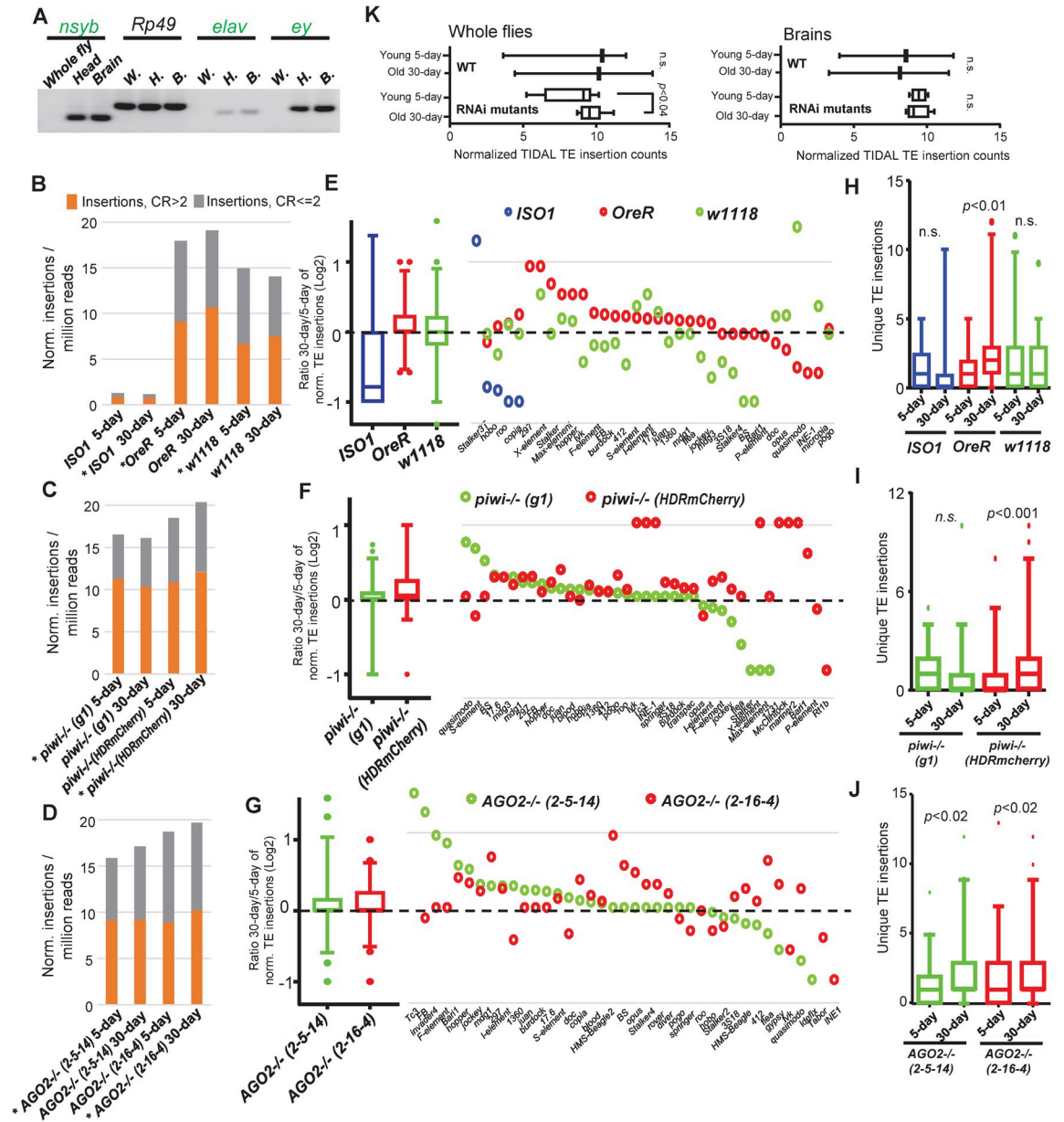


Fig 4. Aging-associated TE landscapes in fly brains of WT and RNAi mutant strains. (A) Validation of fly brain dissections by RT-PCR of brain-specific gene expression. TIDAL analysis of WGS for new TE insertions in the brains of (B) Wild-type (WT) strains, (C) *piwi* mutants, and (D) *AGO2* mutants. The bar graphs on the left represent categories of total TE insertions broken by the coverage ratios (CR) of CR>2 and CR<= 2. Asterisks mark the library that was downsampled to the equivalent depth of the cognate comparison library. (E-G) Left, 5–95 percentile boxplots of ratios of the 30-day versus 5-day of normalized TE insertions from panels B-D. Right dot plots display only the TE families with significant changes between the two timepoints in at least one of the mutant strains. (H-J) Boxplots for conducting paired Wilcoxon tests for statistical significance in the gain of unique TE insertions in aging RNAi mutant fly brains. (K) Boxplots of the grouped differences of paired TE insertion counts between 30-day versus 5-day amongst whole flies and fly brain WGS libraries between WT and RNAi mutants. *p*-value from a one-tailed Wilcoxon rank-sum tests, *n.s.* = not statistically significant. Details of the samples and values in used to build these plots are in [S1 Table](#), which only include the newer CRISPR/genome edited mutants versus the wild-type strains, as discussed in the main figures and discussion.

<https://doi.org/10.1371/journal.pgen.1010024.g004>

However, the additional normalization to the plasmid spike-ins were also informative in moderating eccDNA levels in *w1118* while reaffirming the TE eccDNA increases in *OreR* and *ISO1* (Fig 5F and 5G). The plasmid spike-ins controlled for variations in extraction procedures and

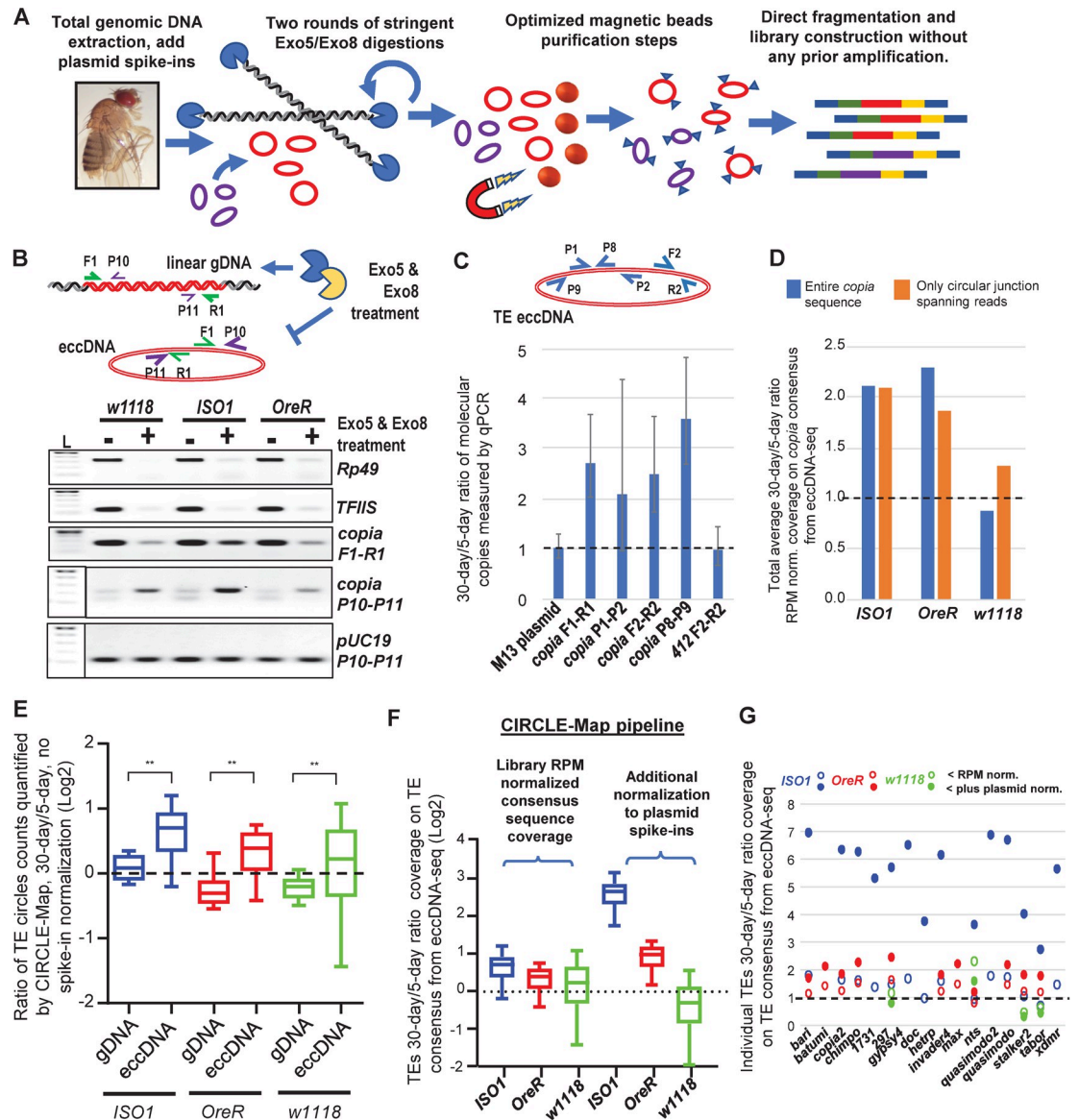


Fig 5. Aging *Drosophila* display increases in TEs existing as extra-chromosomal circular DNA (eccDNA). (A) Diagram of methodology to enrich and purify eccDNAs for direct library construction and sequencing without requiring prior amplification. (B) Genomic PCR from WT flies demonstrating the depletion of linear gDNA and enrichment of eccDNA with TE sequences during exonucleases treatments. The diagram above explains configuration of PCR primers. L = DNA ladder. (C) qPCR validation of spike-in plasmids and *cop*ia eccDNA after exonucleases treatments of *ISO1* gDNA from young versus aged adult flies. The diagram above explains the configuration of PCR primers. (D) Ratio of the read coverage just across the *cop*ia consensus sequence comparing young versus aged flies. (E) Comparison of the ratios of 30-day to 5-day CIRCLE-Map counts of TE circles between total gDNA libraries and eccDNA-enriched libraries without plasmid spike-in normalization. (F) Box plots of 30-day/5-day ratios of read coverage for eccDNA TE sequences rated by the CIRCLE-Map pipeline with a positive “circle score” [79] comparing to additional normalization to plasmid spike in. (G) Dot graph highlighting specific TE eccDNAs whose 30-day/5-day sequencing ratios are normalized to library RPMs (outlined dots) and to the plasmid spike-ins (filled-in dots) from a subset of (F) for TE families that had significant “circle score” >50.

<https://doi.org/10.1371/journal.pgen.1010024.g005>

linker addition steps between samples during the eccDNA library construction. Thus, while *ISO1* TEs did not change much at the chromosomal level during aging, *ISO1* TE copy numbers may instead increase through eccDNA accumulation. Intriguingly, the *ISO1* strain showed the least chromosomal TE changes yet exhibited the greatest increase in TE-eccDNAs in the whole

flies, while the *OreR* and *w1118* strains also showed evidence of TE-eccDNAs accumulating in the brain after normalization to plasmid spike-ins (S6E Fig).

We have confidence in our quantification of TE-eccDNA increases during fly aging because of four reasons. (1) The increase in TE-eccDNAs during fly aging detected by the CIRCLE-Map program was evident in the eccDNA-enriched libraries while being clouded by background in WGS libraries from gDNA (Fig 5E). (2) There was no aging increase and low counts in the eccDNA libraries for IGEs, since these would be in linear gDNA that would be extensively degraded by the exonuclease treatments (S6F Fig). (3) The current results (Round 10) show plasmid spike-in counts as a ratio of the eccDNA library being similar to the ratio of counts for TE-eccDNA (S6I Fig), which contrasts against an earlier eccDNA sequencing run (Round 9) that we did not analyze further because of too much plasmid spike-in. (4) Mitigating concerns of pipetting variability between samples, we observed across multiple samples consistent ratios between the counts for a group of plasmids that were spiked-in prior to exonuclease treatment relative to a fifth plasmid spiked-in even earlier, prior to gDNA extraction (S6J Fig). Together, these metrics will be helpful towards evaluating the success of future eccDNA sequencing efforts, which we can envision to be adding an extra heterologous gDNA spike-in to complement the plasmid spike-in.

Other considerations of TE landscapes in other RNAi mutants

In addition to variations in TLs between WT strains, we also observed differences in TLs between other RNAi mutants that we cannot fully explain. For example, we examined aging-associated TLs from two EMS-induced point mutants of *Dcr-2* (*L811fsx*) and *Dcr-2* (*R416X*) from [91], the nuclease acting upstream of *AGO2* to generate the siRNAs from TE dsRNAs. However, there was inconsistent and contrary TL differences between young and aged *Drosophila* in these *Dcr-2* mutants whole flies and brains (S7A and S7B Fig) as well as in *AGO3* mutants (S7C Fig, [92]). Perhaps these sets of mutants are not as penetrant in the loss of RNAi activity such as the persistence of siRNAs in the two *Dcr2* mutants [91], these *AGO3* mutants actually do not lose fertility nor lose TE repression [92], or there might be potential off target effects contributed by different genetic backgrounds. Furthermore, the analysis of a partially rescuing *AGO2* transgene in the *AGO2* (*2-5-14*) null mutant did lower the initial levels of TE insertion differences noted by TIDAL, but this partial rescue (where the *AGO2* transgene did not reach wild-type expression levels) still did not prevent aging-associated TE increases (S7D Fig). Nevertheless, we propose that RNAi activity must be sustained during aging to mitigate negative effects of increased TE expression in aged flies, a phenotype that has also been frequently observed in mammals [34,55,93,94].

Genetically enhancing RNAi counteracts TE expression during *Drosophila* aging

Although TE expression still increased in WT aging flies, we hypothesized whether endogenous RNAi pathways that still limit genomic TL increases could also be genetically enhanced to mitigate the aging-associated rise of TE RNAs. To test this hypothesis, we first used a ubiquitous *Tubulin-GAL4* driver to overexpress *AGO2* in adults, and as expected, multiple TE RNAs had lowered expression relative to the negative control (Fig 6A). We then used the same driver to overexpress *piwi*, and although there was likely a silencing limit to prevalent *piwi* expression in the ovary, the enhancement of *piwi* expression and TE silencing were much more apparent in the female carcass (Fig 6B). Although the status of the genetic backgrounds and TE expression levels of the parental strains in these GAL4-UAS experiments is uncertain, these data provided a proof of principal that augmenting these RNAi pathways in adults result in improvements in TE silencing.

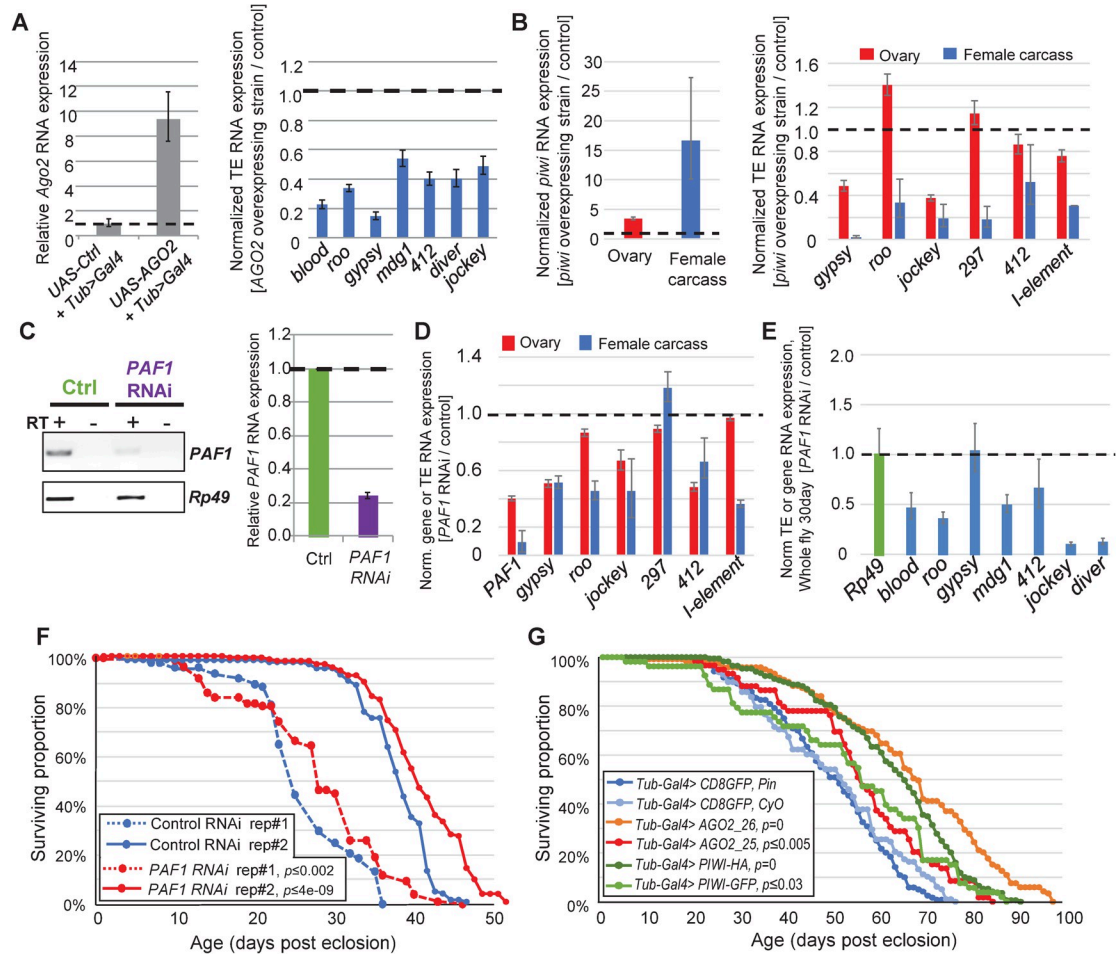


Fig 6. Genetic interventions of TE expression in adult *Drosophila*. (A) Overexpressing AGO2 [*UAS-HA-AGO2/Tub>Gal4*] and (B) overexpressing PIWI [*UAS-3X-HA piwi/+; Tub>Gal4/+*] results in a reduction of TE RNA expression in 5-day young adult *Drosophila*. Left graphs confirm gene overexpression and right graphs detail TE RNA expression measured by RT-qPCR of the target gene compared to the *Rp49* housekeeping gene and with error bars representing propagated standard deviation of triplicate measurements. (C) Adult-specific knockdown of *PAF1* in 5-day young females qualitatively assessed in the gel (left) and RT-qPCR (middle), which reduces TE RNA expression (right) using *Tub>Gal4; PAF1 RNAi*. Examining the effect of TE RNA reduction in the *PAF1* knockdown in the ovary and carcass of 5-day flies (D) and 30-day whole flies (E). (F) Life span comparison between control versus *PAF1 RNAi* knockdown of adult female flies upon raising them at 29°C to release the Gal80^{ts} inhibitor to induce RNAi from the *Tub>GAL4*. *PAF1 RNAi* N = 112,119, *Control RNAi* N = 170,153, rep#1 and rep#2, respectively. (G) Life span comparison between strains overexpressing AGO2 and PIWI and control strains *CD8GFP*. *AGO2_25*, N = 119, *AGO2_26*, N = 59, *PIWI-HA*, N = 169, *PIWI-GFP*, n = 53, *CD8GFP/+;CyO/+*, n = 98, *CD8GFP/+;Pin/+*, N = 120. *p*-values are from the log rank test calculated with the Oasis tool [123].

<https://doi.org/10.1371/journal.pgen.1010024.g006>

However, inhibiting a factor that normally limits RNAi activity would be preferable from a therapeutic standpoint. Examples of endogenous negative regulation of RNAi activity include proteasome-mediated turnover of AGO2 [95], ENRI factors that negatively regulate nuclear RNAi in nematodes [96], and the RNA exosome and *PAF1*'s transcription elongation role modulating RNAi silencing activity on TEs conserved in both fission yeast and flies [97–99]. Even though we were able to use siRNA knockdown of *PAF1* in *Drosophila* OSS cells to demonstrate enhanced TE silencing, we recognized that genetic knockdowns of this essential modulator of RNAi would also have detrimental effects on development like its requirement in ovarian development [97].

So, to circumvent developmental impacts of *PAF1* knockdown in flies, we further combined the temperature-sensitive inhibitor of GAL4 expressed from a second transgene of *Tubulin-Gal80^{ts}* with the *Tubulin-Gal4* driver [100]. This double-transgenic fly could then be crossed to the same *UAS-PAF1-RNAi* line so that flies can develop fully at the permissive temperature of 18°C, and after eclosion be raised at 29°C to trigger the RNAi knockdown of *PAF1* (Fig 6C). Because elevated temperature itself can affect TE silencing activity in flies [101–104], we used an *mCherry-shRNA* strain as a negative control that was also raised at 29°C at the same time as the *PAF1* knockdowns. There was appreciable enhancement of TE silencing in the whole female flies (Fig 6C) with similar levels of TE silencing enhancement between the ovaries and the soma (Fig 6D and 6E). We attribute the increased TE silencing during *PAF1* knockdown to the reduced elongation rate of TE transcripts so that RNAi factors can better engage [97] and not from a global transcription reduction because steady state levels of control gene, *TFIIs*, *AGO2* and *piwi* were not reduced by *PAF1* knockdown (S8A Fig).

We then measured lifespans from two cohorts of flies where *PAF1* knockdown was compared to the *mCherry-shRNA* negative control with the *Tubulin-Gal80^{ts}* and *Tubulin-Gal4* driver cross at 29°C (Fig 6F). Although one cohort measured at the start of the Covid-19 pandemic initially dipped at 2 weeks, the *PAF1 RNAi* knockdown flies ended up living statistically longer than the control, and this lifespan extension was more pronounced and replicated in a second cohort. To support the notion that *PAF1* knockdown was enhancing RNAi activity to extend *Drosophila* lifespan, we also observed a significant lifespan extension from multiple lines overexpressing *AGO2* and *PIWI* (Fig 6G). Although the control strain used for comparisons to *AGO2* and *PIWI* overexpression also express GFP, perceptions of GFP toxicity can be tempered by that fact that these control animals outlive all other WT strains (Fig 1A).

Since we had observed TE landscape activity in the adult fly brain (Fig 4), we also tested a brain-specific driver, *elav-GAL4*, that was effective at triggering *PAF1* knockdown and enhancing TE silencing in the 30-day aged fly brains (S8A Fig). However, this *elav-GAL4* driver that is expressed as early as during embryonic development [105] was likely also reducing *PAF1* to a point that it lowered the overall fitness relative to the control (S8A Fig). To bypass the critical embryonic developmental stages, we were able to finally recombine *Tubulin-Gal80^{ts}* with the *elav-GAL4* and confirmed neuronal-driven GFP expression after 29°C induction (S8B Fig). However, neither *PAF1* knockdown nor TE silencing enhancement could be observed initially from third-instar larvae or eclosed adult 29°C induction, although some reduced *PAF1* and TE expression was eventually observed in 30-day old adults (S8C Fig). The inclusion of *Tubulin-Gal80^{ts}* has somehow tempered the effectiveness of the *elav-GAL4-PAF1 RNAi*, so that neither developmental issues of premature *PAF1* loss nor longevity extension was seen (S8D Fig), but a final experiment using the drug RU486 to trigger *Geneswitch->elav-Gal4-PAF1-RNAi* was able to cause more widespread gene and TE expression knockdown and extend lifespan versus the ethanol control. Together, these data suggest that future spatial and temporally controlled inhibition of *PAF1* to optimal levels may be a relevant avenue of intervening with the aging-associated increase in TE expression in maturing adult animals.

Discussion

In this study we conducted an analysis of WGS approaches towards assessing changing TLs during *Drosophila* aging, and we found that TL increases are detectable in the genomes of aging RNAi mutants, especially *aubergine*. These mutants are viable although others have shown that they have reduced longevity compared to control strains [10,13,16], and our data now confirms that unchecked elevation of TE transcripts can result in quantifiable genomic alterations in a single lifetime of flies. This result was primarily supported by the data in

aubergine flies, suggesting that the germline is the primary location of this activation. However, it was more difficult to detect new TE insertions amongst the gDNA of WT fly strains: we had to focus the TIDAL analyses on specific TE families mobilizing into uniquely-mapping sequences and also count the coverage on TE family consensus sequences (S3 Fig). After showing that an orthogonal quantitation method like ddPCR is consistent with TIDAL's quantitation of TE copy numbers from WGS of *P-elements* and *I-elements* (S4 Fig), our parsimonious conclusion is that despite aging-associated increases in TE expression during fly aging, the RNAi pathway may protect the fly genomes from accumulating new TE insertions, particular from the ovary which can still contribute to the overall longevity of female flies. Furthermore, our study expands the dimension of WGS of TEs by incorporating eccDNA as an *in vivo* cache of accumulating TE DNA sequences (Fig 5).

Despite the compactness and completeness of the *D. melanogaster* genome sequence, technical challenges still remain to fully optimize WGS approaches for quantifying TEs. For example, all current metazoan genome assemblies still suffer from large sequencing gaps in telomeric, centromeric and other repetitive regions that remain challenging to analyze. Meanwhile, long-read sequencing like Nanopore and PacBio that could close these gaps are still less economical and not as accurate as the Illumina sequencing platform [106], yet library construction methods for the Illumina platform require sufficient input material for reproducible generation of sequencing libraries. Single-cell WGS is not yet robust enough nor has total DNA amplification approaches been demonstrated to be unhampered by molecule bias, so our study required pools of genomes and non-amplified input DNA samples to reduce the prior concerns.

In retrospect, our study had to balance several competing factors such as (1) reducing to as few a number of individual animal genomes for gDNA input versus (2) generating WGS libraries that robustly sequence without prior whole DNA amplification versus (3) reducing the presence of artifactual PCR amplicons like IGE insertion events versus (4) achieving nearly equivalent sequencing depths and conditions between samples amidst inherent biochemical and bioinformatics limitations. For instance, preliminary single fly WGS libraries will require further optimization because IGE artifacts were unacceptably higher than our settled approach here to pool 10 whole flies and 50 brains of a particular strain which has the caveat of yielding differences in the normalized TE insertion count number. Future studies will be needed to improve sensitivity and accuracy of WGS analysis of *Drosophila* TE landscapes, because our current data does not capture the single-cell microscopy detection of transposon-based genetic reporters like the *gypsy-TRAP* and *gypsy-CLEVR* that can detect increased transposition activity in aging flies [13–15,18,107]. These reporters have the advantage of low cost and sensitivity of detecting small numbers of cells in a background of nonmodified cells, yet this transgenic construct is also only designed for *gypsy* to insert and activate a fluorescent protein read-out and cannot assess overall genomic TEs.

With the pleiotropic effects of TE expression during aging, it remains a worthy goal to combat this potential impact such as with therapeutic approaches using reverse transcriptase inhibitors to inhibit *LINE1/L1* activity [33], while other studies showed that dietary restriction and prolonged exercise in animals can reduce aging-associated increases in TE expression [14,54,108]. Our study proposes an additional therapeutic target of augmenting the RNAi pathway's response to TEs by inhibiting *PAFI*, which has a conserved impact on limiting RNAi from silencing TE transcripts [97,98]. Perhaps therapeutic siRNAs against *PAFI* transcripts can be hypothesized as a feed-forwarding therapeutic agent to augment RNAi activity in aging animal cells.

A final question to resolve in the future is what cascade of epigenetic and chromatin landscape changes during animal aging leads to increases in TE expression? Given the pleiotropic nature of the animal aging process, we anticipate that there will also be multiple genomic

mechanisms that will vary in impact between different genetic backgrounds. For example, we describe variation amongst three WT *Drosophila* strains in the level of accumulating eccDNAs containing TE sequences (Fig 5), while others have shown increased in polyploidy in adult *Drosophila* brains [109] as well as somatic genome instability in regions of the *Drosophila* genome [110] that might contribute to changes at the level of TE consensus sequence coverages (S3B Fig). Lastly, during fly aging there are also gross-level changes in histone marks typically associated with chromatin silencing [12,14], which may precede the increase of TE expression, so the future extension of this work will be to add epigenetic and chromatin accessibility landscapes to TEs during *Drosophila* aging.

Materials and methods

Drosophila strains, genetic crosses and aging curves

All flies were raised at 25°C on standard cornmeal food. For fly aging analyses, newly eclosed female flies were harvested from bottles and mated with males for two days. These females were then divided into ~20 individuals per vial and flipped to new vials every 2–3 days to mitigate crowding stress according to this protocol [111]. Surviving flies were counted at each flip, and the percentage of cumulative survival rate at each time point was plotted against its corresponding age (date of counting subtracted by date of eclosion).

The isogenized *ISO1* fly strain for the Dm6 reference genome sequence was obtained from Susan Celniker [4]; the *w1118* is an isogenized strain and was a gift from R. Scott Hawley [68]; and the *OreR* from the ModEncode project was a gift from Terry Orr-Weaver [69]. The RNAi pathway null mutant strains *piwi-(g1)*, *aubergine-(g1)*, *aubergine-(g2)*, *AGO3-(g1)* and *AGO3-(g2)* were a gift from Julius Brennecke [92]. An additional mutant strain of *Piwi-[HDR-4xP3-mCherry]* [112] was a gift from Eric Lai. The null *AGO2* mutants deletion strains of *AGO2-[2-5-14]* and *AGO2-[2-16-4]* and *Ago2-WT-rescue* stocks were generated by CRISPR Cas9 approaches as described in [113,114]. The strains with active and inactive *I-elements* and spermless males were a gift from Zhao Zhang [72]. The *UAS-Ago2-HA* strains were a gift from Arno Muller lab [115] with two transgene insertions: *Ago2-25* on Chr3 and *Ago2-26* on Chr2. The *UASp-3xHA-Piwi* was a gift from the Ruth Lehman lab [116] and *UAS-GFP-Piwi* was obtained from Jean-Yves Roignant as a kind gift from J. Brennecke. The *PAF1* knockdown RNAi line was obtained from the Vienna *Drosophila* Resource Center (VDRC#108826) and an *mCherry-shRNA* control line was obtained from the Harvard TRiP resource (BDSC#35785). The driver strains of *Tubulin-Gal4* and *elav-Gal4* and the *UAS-CD8GFP;Pin/CyO* stocks were gifts from Leslie Griffith [117]. The *Geneswitch->elav-Gal4* driver was a gift from the Marr Lab at Brandeis University. In addition to an early *Tubulin-Gal80^{ts}* strain we received from the Griffith lab, we also obtained a second *Tubulin-Gal80^{ts}* stock from the Bloomington *Drosophila* Stock Center (BDSC#7018) and combined with *Tubulin-Gal4* which later we molecularly validated before using it for the downstream aging experiments.

For the *Tubulin-Gal4* induced over-expression, parental cross between *Tubulin-Gal4* with the *UAS-CD8GFP;Pin/CyO* as control versus *UAS-Ago2* or *UAS-PIWI* were set at 25°C and F1 adult females of the control over-expression (*UAS-CD8GFP/+; +/Pin; Tubulin-Gal4/+* and *UAS-CD8GFP/+; +/CyO; Tubulin-Gal4/+*) versus *Ago2* (*UAS-Ago2(26)/+; Tubulin-Gal4/+* and *UAS-Ago2(25)/Tubulin-Gal4*) overexpression or *PIWI* (*UASp-PIWI-3xHA/+; Tubulin-Gal4/+* and *UAS-PIWI-GFP/Tubulin-Gal4*) overexpression were collected and followed at 25°C for life span assay. For the *Tubulin>Gal4* induced knockdown, parental cross between *Tubulin-Gal80^{ts}/+; Tubulin-Gal4/TM6* with *mCherry-shRNA* versus *PAF1-RNAi/CyO* were set at 25°C and F1 adult females with correct genotype were collected respectively and transferred to 29°C incubator and assayed for the entire life span assay. For the neuronal *elav-Gal4* induced

knockdown, parental cross between the *elavGal4/CyO* driver with *mCherry-shRNA* versus *Paf1-RNAi/CyO* flies were set at 25°C and F1 flies with correct genotype were collected and assayed for the life span assay.

For the combinatory *elav-Gal4* with *tubulin-Gal80^{ts}* experiment, *elav-Gal4* was first recombined with *tubulin-Gal80^{ts}* (since both transgenes are located on Chr2) by screening from at least 20 individual F1 recombinants backcrossed to double balancers. Genotyping on F2 recombinants against both Gal4 and Gal80^{ts} primers confirmed the success of obtaining 1 line containing both transgenes. This recombinant line was further functionally validated by crossing with a *UAS-CD8GFP* stock, of which the F1 progenies with genotype of *UAS-CD8GFP/+; elav-Gal4, tub-Gal80^{ts}/+* glow only after shifting to 29°C. These *elav-Gal4, tubulin-Gal80^{ts}/CyO; +/TM6* males were crossed with *mCherry-shRNA* versus *PAF1-RNAi/CyO* virgins and F1 adult females with correct genotypes were harvested and transferred to 29°C for the life span assay. Replicate 1 and 2 have identical genotype except one set (of both control versus *PAF1 RNAi*) are heterozygous of *TM6* instead of wildtype on Chr3.

For the Geneswitch assay, parental cross between *Geneswitch->elav-Gal4* and *Paf1 RNAi/CyO* were set at 25°C and 200 ul 100% EtOH versus 500 uM RU486 (dissolved in 100% EtOH) were added to fly vials with third instar crawling larval (generally at day 7 after initial cross set day). F1 adult females were collected and transferred periodically to the EtOH versus RU486 (same volume and concentration as added during the third instar larvae stage) fly food throughout the life span assay and counted until age up to 95 days.

To quantify *I-element* copies by ddPCR, fly cross schemes from [72] were replicated (S5C Fig). Parental crosses between the *w¹¹¹⁸* strain with active *I-elements* and *w^k* strain with inactive *I-elements* were performed reciprocally to generate many virgin F1 females where one strain enables *I-element* transposition (“invaded” from *w^k* as the maternal parent) versus a control that maintains *I-element* silencing (*w¹¹¹⁸* as the maternal parent). These F1 females were then crossed to sperm-less males that were obtained as F1 male progenies from the parental cross of *w¹¹¹⁸* virgin females with XY attached male. F2 oocytes were collected overnight and DNA was extracted for ddPCR against the *I-element* and *Rp49*.

Fly brain isolation, genomic DNA extraction, WGS library construction and deep sequencing

Fly brains were dissected from at least 50 females per age group, following a procedure laid out in [118]. Eye disks and other tissues were removed from heads with forceps, and brain lobes were dissected into tubes with ice-cold PBS before freezing once at -20°C. Whole female flies and fly brains were homogenized in a standard DNA digestion buffer (1% SDS, 50 mM Tris-HCl pH 8.0, 100 mM EDTA, 100 mM NaCl, 0.5mg/ml Proteinase K) overnight at 50°C, and then extracted using standard phenol chloroform extraction, ethanol precipitation, and resuspending gDNA pellets in pure water.

WGS of whole flies began with the circa 2014 Nextera Tn5 tagmentation kit (Illumina) using an input of 50 ng gDNA and outputs were purified with AMPure XP beads (Beckman Coulter). WGS libraries were quality controlled with the high-sensitivity DNA kit on the Bioanalyzer (Agilent), selecting for size distributions of 300bp to 1kb and concentrations over 1 nM. Multiplexed libraries were sequenced on Illumina Nextseq500 high-output flow cells using 75 bp paired-end and single end reads. All WGS libraries were sequenced to a minimum depth of 30 million reads (S1 and S2 Tables). After determining that some whole fly libraries made using NEBNext Ultra-II DNA library prep kit for Illumina (NEB) were as complete and has better yields than the then discontinued Nextera kit, we completed the fly brain gDNA libraries with the NEBNext kit and sequenced them to similar depths as above.

RNA extraction, quantitative RT-PCR, digital droplet PCR (dd-PCR) and TE copy number estimation

Total RNA was extracted from 5–10 female flies harvested at corresponding age with TRI-reagent (MRC, Inc.). Reverse transcription (RT) was performed using random primers, ProtoScript II (NEB), and 1 μ g input of total RNA. Quantitative PCR (qPCR) with the Luna Sybr-Green master-mix (NEB) using primer sequences in [S2 Table](#) and 2 μ L of a 1:10 dilution of the cDNA.

Relative changes in gene expression were calculated using the $2^{-\Delta\Delta C_t}$ method with *Rp49* as a housekeeping gene for normalization. Briefly, the ΔC_t value difference between TE target and housekeeping gene (*Rp49*) was calculated for both experiment groups (e.g. knockdown conditions) and control groups; and the difference between these two ΔC_t values ($\Delta\Delta C_t$ experiment- $\Delta\Delta C_t$ control) was further calculated to obtain the $\Delta\Delta C_t$ value. Relative fold change values (from experiment to control) was calculated from the exponent of 2 to the power of negative $\Delta\Delta C_t$ value.

Droplet digital PCR (ddPCR) was conducted on a QX200 instrument with the Evagreen assay reaction (Biorad). Copy number measurements from specific TE primers ([S2 Table](#)) were normalized to *Rp49* as a diploid gene, starting first at 2 ng of gDNA as input per 20 μ L ddPCR for droplet generation for most TEs. For TEs with very high copy numbers that saturate the droplets, input gDNA was diluted further into the ddPCR mix prior to droplet generation. At least 10,000 droplets were required to achieve good statistical estimation of the concentration calculated by Poisson distribution using Quantasoft Analysis Pro (Biorad). TE copy numbers per genome was determined by dividing against half of the measured *Rp49* copies.

Extracellular circular DNA isolation and sequencing

To quantify eccDNAs during fly aging, 30 female flies were harvested from 5-days and 30-days post eclosion, and a fixed amount of pre-extraction plasmids was added prior to cell lysis: ~80 pg of ~7kb-pGL3-DmPiwipro1 and ~50 pg of ~11kb pJDS246-CarhCas9 (pJC9F3). About 30 μ g of total gDNA was recovered from using MasterPure Complete DNA and RNA Purification kit (Lucigen), and 0.5 μ g-1 μ g gDNA was checked on a 1% agarose gel for integrity and quality. Good gDNA primarily migrated at >10kb and lack of RNA contamination that would result in bands at the lower molecular weight ranges. 20 μ g gDNA were added to a 40 μ L of a second post-extraction plasmid cocktail: (1ng/ μ L of the 2.7kb pUC19, 0.1ng/ μ L of the 3.5kb pMaxGFP, 0.01ng/ μ L of the 5.2kb pGSH0 and 0.001ng/ μ L of the 6.3kb pCENPm3) and split equally to two reactions: Exo5/8 non-treated control versus Exo5/8 treated samples. We conducted a first round of Exo5/Exo8 (NEB) treatment at 37°C overnight, then an additional 2-hour treatment with freshly replenished buffer, ATP and enzymes. The reaction was stopped and purified using AMPure XP beads (Beckman Coulter) and eluted in 50 μ L of water.

To check the efficiency of Exo5/8 treatment, 10 μ L of the eluate from untreated versus treated samples were loaded on 1% agarose gel to visualize the completeness of digestion of gDNA. We quality controlled Exo5/8 treatments by performing qPCR against *Rp49*, ND5 and various plasmid primers including pUC19, pGL3piwipro and pCas9 and Ct values were compared between untreated versus treated samples. Mitochondria was not a reliable circular molecule because of the high variability of ND5 Ct values across multiple sample preps. Comparing between treated and untreated sample, the plasmids Ct values were generally stable (<2 Ct difference), and much higher for *Rp49* (>5 Ct difference) indicating the Exo5/8 treatments were effective at removing linear chromosomal DNA and not affecting the circular plasmids. Half of the Exo5/8 treated sample (25 μ L out of 50 μ L purified elute) was used as template for library construction using NEBNext UltraII library prep kit as stated above.

For eccDNA sequencing from brains, 200 female brains were dissected and added with half the volume of pre-extraction plasmids as whole flies, and gDNA concentration was measured by the Qubit 4 Fluorometer (ThermoFisher). To 100 ng of brain gDNA, we mixed 20 μ l of the post-extraction plasmids spike-in cocktail and a tenth of the Exo5/8 enzyme as whole flies gDNA. At least 10 million eccDNA reads were required for analysis. Libraries were pair-end sequenced at 36 bp by 36 bp on a Nextseq550 flow cell (Illumina).

TIDAL updates with total TE consensus and gene mapping strategies

TE insertion analysis was carried out with an updated version of our previously developed TIDAL program (original code available on the Github repository at: <https://github.com/laulabbrandeis/TIDAL>) [65]. In this study, the updated version of TIDALv1.2 is also posted to Github at (<https://github.com/laulabbumc/TIDAL1.2>). These scripts carry out the analysis run the same way as the original TIDAL, but we incorporated two additional features. First, for the euchromatic TE insertions we selected 100 arbitrarily selected protein coding gene (Immobile gene elements IGEs) that are computed along with consensus TE sequence to benchmark noise in detection of genetic elements. The algorithm used to identify transposon insertion sites based on consensus transposon sequence is then applied on these 100 IGE sequence to determine their insertion sites. Second, for the total reads mapped to consensus TE sequences, here we added 100 IGEs are computed by mapping reads with bowtie2 using parameters “—sensitive—end-to-end” and custom shell, Perl, C-code, and R-code scripts all accessible from (<https://github.com/laulabbumc/TIDAL1.2>).

TEMP v1.05 code was acquired from the GitHub repository at: (<https://github.com/JialiUMassWengLab/TEMP>), and was run with default parameters except “-x 30, -m 3 -f 500”. These parameters were chosen to ensure that TEMP results are consistent with analysis shown in [60,62].

Downsampling and bootstrapping of TIDALv1.2 results

For each pair of 5d versus 30d samples, number of mapped reads were calculated by removing unmapped portion from the total reads (S1 Table) and the ratio of mapped reads between the lower depth versus deeper depth sample was calculated and round up to 4 significant digits. This ratio was used as the fraction number parameter by the Seqtk tool [119] for randomly selecting a fraction from the deeper depth library to the similar amount of the lower depth library. Mapped reads were computed and compared between each set of the libraries, leading to an average of an insignificant difference of less than 3000 reads.

For the initial bootstrapping of a testing *OreR* 30-day whole fly library for 100 permutations, Seqtk tool [119] was iterated 100 times with seed numbers set equal to the iteration number and fraction number set equal to 0.5. These 100 random downsampled fastq files were subsequently batch processed by TIDALv1.2 and TE insertion numbers for each permutation were counted. We plot the number of TE insertions per million of mapped reads for each permutation (S2A Fig) and calculated the standard deviation between each number of permutations run (increased number of permutations from 1 to 100, S2C Fig). We also plotted the cumulative average of TE insertions per million of mapped reads from increasing number of permutations (S2B Fig). Finally, we chose to set at 20 permutations for analyzing normalized TE insertions per million of mapped reads for all the libraries with error bar plotting as the 95% Confidence Interval.

After downsampling to equalize the library sizes between 5-day and 30-day samples, the differences in total number of TE insertion calls between the 30-day and 5-day samples were reassessed and recorded in S1 Table.

Bioinformatics counting of eccDNA from TEs and spike-in plasmids using a custom pipeline and CIRCLE-Map program

In our first look at the eccDNA reads, we inputted them into an existing bioinformatics pipeline already developed for mapping *Drosophila* small RNA counts to TEs [120]. Reads were first checked by the Cutadapt program to see if adaptor sequences at the 3' end needed to be removed, and then we indexed the reads to the *Drosophila* genome assembly file by running BWA version 1 [121] and formatdb from NCBI. Using Bowtie1 with 2 mismatches [122], reads were mapped to genome to get the genic and intergenic counts using the genome GTF file. The total number of reads mapped to the *Drosophila* genome was derived by subtracting the total number of reads not mapped to the *Drosophila* genome from the total number of reads. The total number of mapped reads was used as the basis for normalization of TE counts and spike-in plasmid counts.

Plasmid sequences were treated as linear entries in the FASTA file database similar to the TE family consensus sequences. The raw read counts from TE mapping were further normalized by the total number of reads mapping to the *Drosophila* Dm6 genome assembly. For spike-in plasmid counting, because several plasmids share the same backbone with different inserts, read frequencies were normalized by the total plasmid mapping sites as well as by the total number of *Drosophila* genome-mapping reads. Normalization using spike-ins was calculated as $[(x/y) * 10^6] / z$ where x is the read counts, y is the total number of aligned reads from a given library and z is spike-in read counts.

To execute the CIRCLE-Map program for repeats [79], we indexed the *Drosophila* genome FASTA file by BWA. We then used the MEM algorithm under BWA to align reads against the *Drosophila* genome FASTA file. Next, we sorted the reads by alignment position within the resulting BAM file and indexed the resulting BAM file. Finally, we detected the circles by calling CIRCLE-Map program. The CIRCLE-Map program for repeats yields an output for reads with two high scoring alignments as these ones are indicative of circles formed from regions with homology.

For calculating the ratio of individual circular TE coverage between 5-day versus 30-day, the cumulative average of the mean base coverage from all types of circular DNA coordinates for each type of TE was calculated and the ratio was further normalized by the respective total reads after trimming (bottom panel, S3 Table). For normalization by spike-ins, the ratio was further normalized by the mean base coverage ratio of the pre-extraction plasmid pGL3-DmPiwipro1, which was consistently detected by the CIRCLE-Map program throughout all samples with top ranked circle scores.

Supporting information

S1 Fig. Variability in TE Landscapes (TLs) between TEMP and TIDAL programs applied to the same set of whole genome sequencing (WGS) libraries. (A) The WGS libraries from [63] study where TEMP was used to determine TE Landscapes has different results from the subsequent [60] study as well as being distinct from our re-analysis of the datasets in this study with TIDAL. (B) Fewer but more confident determinations of TE landscapes using TIDAL to reanalyze the WGS data originally analyzed by TEMP in [62]. (C) TIDAL re-analysis of the WGS data from [60], which only detects the TE insertion marked by the maroon dot because this is supported by pair end reads pairs and encompassing split reads shown in the example genome browser snapshots to the right. All the orange dots are likely false positives that TEMP will report but TIDAL is designed to filter. (D) Boxplot of TIDAL TE counts normalized to library depth of RPMs, and then the ratio of the normalized TE count between 21 day and 2–4

day flies from the *w1xHar* cross.
(PDF)

S2 Fig. Bootstrapping analysis of TE insertion determinations by TIDAL from WGS data.

(A) Using the *OreR* whole females WGS library data and downsampling to 50% depth, 100 permutations of randomly-downsampled WGS data were run through TIDAL and TE insertions at different Coverage Ratios (CR) as well as Immobile Genetic Elements (IGEs). (B) Cumulative average of TIDAL TE insertion determinations with increasing number of permutations. (C) Standard deviation of TIDAL TE insertions reach an optimally low minimum at 20 permutations in this *OreR* bootstrapping test. (D–G) Bar charts showing the average TIDAL TE insertion determinations with 95% confidence interval marked by the error bars for each determination after subjecting each of these libraries to a 50% downsampling and 20 permutations bootstrapping test to evaluate for potential sequencing sampling noise during TIDAL analysis. Each cohort corresponds to other analyses from Figs 2, 3 and 4. (H) Re-analysis of [60] WGS from fly brains with TIDAL and sequencing coverage analysis. Libraries grouped by triplicates with average TE insertions normalized by library depth and standard deviation plotted. One library each in young fly sample from the 250nt long read library is a major outlier causing the wide standard deviation. (I) Boxplots of the ratio of 30-day versus 5-day coverage of individual TE families (left) and 100 arbitrarily selected protein coding genes also used for IGE analysis (right) from concatenated libraries from all brain samples.
(PDF)

S3 Fig. Control-FREEC plots of WGS coverage from fly strains used to measure TE profiles during fly aging.

(A) Wild type fly lines libraries. (B) *Aubergine* and *AGO3* mutants libraries. (C) *Piwi* mutants libraries. (D) *AGO2* mutants libraries. (E) *Dcr2* mutants libraries.
(PDF)

S4 Fig. Location-independent, sequencing coverage-based comparison of genomic TE loads during *Drosophila* aging.

(A) Total TE RPM (top) and average TE (middle) read coverage for TE consensus families versus (bottom) the average read coverage for 100 arbitrarily selected protein-coding genes also used for IGE analysis were examined between 5-day and 30-day old female *Drosophila* wild-type strains. Wilcoxon rank-sum tests applied, with p -value < 0.05 (*), < 0.01 (**), < 0.001 (***). (B) Boxplot distributions of the ratios of sequencing coverage between 5-day and 30-day old flies for individual TE families (top), protein-coding genes (middle) and normalized TE coverage by gene coverage (bottom). (C) Individual TEs per genome quantified by droplet digital PCR in *ISO1* (top), *OreR* (middle) and *w1118* (bottom) strains from a standard input gDNA from 5-day versus 30-day old whole females. Error bars are the propagated 95% confidence intervals from the absolute quantitation based upon the Poisson distribution. (D) Boxplot of the data in (C) confirms the lack of statistical significance in TE copy number difference between young and old WT flies.
(PDF)

S5 Fig. Supporting WGS/TIDAL TE insertion counts with droplet digital PCR (ddPCR).

(A) *P-elements* copies amongst high and low copy strains determined by WGS/TIDAL in Srivastav et al, 2019 are validated by droplet digital PCR. (B) *I-element* copies from *w1118* and *ISO1* from WGS/TIDAL are also validated by droplet digital PCR. (C) Genetic scheme for repeating the natural bursts of *I-element* transposition during oogenesis, derived from Wang et al, 2018. (D) The qPCR showing the fold higher copy number of *I-element* in the active parental strain versus the control strain using two concentrations of input DNA. (E) ddPCR absolute quantification of *I-element* copies per haploid genome from P0 females, F1 females, and F2 oocytes of Active *I-element* versus control strains. All ddPCR error bars represent the

propagated 95% confidence interval of the Poisson distribution used by the ddPCR quantitation algorithm.

(PDF)

S6 Fig. Visualizing enrichment of TEs, satellite repeats, and mitochondrial DNA from eccDNA-seq. Genome browser plots for (A) *copia*, (B) *gypsy*, (C) a Chr3R satellite repeat, and (D) *Drosophila* mitochondrial DNA loci with plots of WGS from gDNA on top in black, in the middle are red tracks of eccDNA sequences as RPM counts not normalized to plasmid spike in, and on the bottom are the Repeatmasker tracks. (E) EccDNAs with significant “circle scores” from the brains of *OreR* and *w1118* show increases during aging. (F) Ratio of 30-day/5-day reads from eccDNA libraries mapped to the list of 100 IGEs in the TIDAL v1.2 pipeline. (G) Box plots of 30-day/5-day ratios of read coverage for eccDNA TE sequences rated by our own custom quantitation pipeline that uses a TE-mapping scripts previously used for small RNA analysis. (H) Dot graph highlighting specific TE eccDNAs whose 30-day/5-day sequencing ratios are normalized to the RPM library size (open) or further normalized to the plasmid spike-ins (closed) from (G). (I) Boxplot showing an earlier Round-9 eccDNA library contained too much of the plasmid spike-ins to be useful for TE eccDNA analysis, compared to the subsequent Round-10 eccDNA library where plasmid spike-in reads are comparable to levels of TE eccDNA reads. (J) Relatively even levels of the individual plasmid spike-ins between samples evaluated against a fifth plasmid that was added to each set of flies before gDNA extraction, indicating that pipetting differences are not largely distorting the eccDNA quantifications.

(PDF)

S7 Fig. Additional assessment of TE landscapes in other RNAi mutants. Quantification of new TE insertions as compared to the reference genome at different Coverage Ratios (CR) using the TIDAL-fly program from whole flies (A) and brains (B) of two *dicer-2* (*Dcr-2*) mutants. Additional TE insertion quantifications of (C) *AGO3* null mutants and (D) transgenic rescue of the *AGO2* gene into the *AGO2* (*2-5-14*) null mutant. Asterisks mark the library that was down-sampled to the equivalent depth of the cognate comparison library.

(PDF)

S8 Fig. Trials with brain-specific RNAi knockdown of *PAF1* with *elav-GAL4*. (A) Expressed as early as embryos, the *Elav-Gal4* driving *PAF1* knockdown in brains of adult *Drosophila* causes TE reduction in left graph without affecting control genes in the right graph. However, the early start of *elav-Gal4-PAF1-RNAi* also caused lifespan reduction in lifespan curve below ($p < 0.001$) (B) Genomic PCR (top) and functional confirmation of the recombinant *elav-Gal4, Gal80^{ts}* stock (bottom). Crosses to include *Gal80^{ts}* with *elav-Gal4* enabled later adult as well as third-instar larvae induction as visualized by neuronal GFP expression. (C) Despite now using heat shock of *elav-Gal4, Gal80^{ts}* to drive neuron-specific *PAF1* RNAi, the knockdown was ineffective at 5-days post-eclosion (top graphs) with just some knockdown at 30-days post eclosion (bottom graph). (D) As a result, there is no significant lifespan extension with the *elav-Gal4, Gal80^{ts}* system. Control RNAi [*mCherry-shRNA/+; elavGal4,tubGal80^{ts}/+*] and *PAF1* RNAi [*elavGal4,tubGal80^{ts}/PAF1RNAi*] females were induced at 29°C since day 1 adult. Both Ctrl and *PAF1* RNAi replicates are identical in genotype except replicate 2 are both heterozygous for the TM6 balancer on Chr3. (E) Additional knockdown of *PAF1* using the *Geneswitch* -> *elav-Gal4* system from third instar larvae with the drug RU486 versus ethanol carrier showed broader overall gene knockdown, and this resulted in a slightly protracted life-span extension. Chisq = 4.1 on 1 degrees of freedom, $p < 0.05$.

(PDF)

S1 Table. Statistics of the rounds of whole genome sequencing (WGS) using TIDALv1.2.
(PDF)

S2 Table. Oligonucleotide primers used in this study.
(PDF)

S3 Table. Statistics for the eccDNA sequencing.
(PDF)

Acknowledgments

We thank Michael Marr, Dianne Schwarz and Michael Blower for comments on the manuscript. We thank Julie Vienne and Thomas Lontis for advice on fly aging assays and lifespan analyses. We also thank multiple laboratories in the *Drosophila* research community for sharing fly stocks as detailed in the Materials and Methods section.

Author Contributions

Conceptualization: Nelson C. Lau.

Data curation: Reazur Rahman, Gargi Dayama, Sizheng Li, Nelson C. Lau.

Formal analysis: Nachen Yang, Reazur Rahman, Qicheng Ma, Gargi Dayama, Sizheng Li, Nelson C. Lau.

Funding acquisition: Michael Rosbash, Nelson C. Lau.

Investigation: Nachen Yang, Satyam P. Srivastav, Reazur Rahman, Nelson C. Lau.

Methodology: Nachen Yang, Satyam P. Srivastav, Nelson C. Lau.

Project administration: Nelson C. Lau.

Resources: Madoka Chinen, Elissa P. Lei, Michael Rosbash, Nelson C. Lau.

Software: Reazur Rahman, Qicheng Ma, Gargi Dayama, Sizheng Li, Nelson C. Lau.

Supervision: Elissa P. Lei, Michael Rosbash, Nelson C. Lau.

Validation: Nelson C. Lau.

Visualization: Nelson C. Lau.

Writing – original draft: Nachen Yang, Nelson C. Lau.

Writing – review & editing: Nachen Yang, Elissa P. Lei, Michael Rosbash, Nelson C. Lau.

References

1. Smith JJ, Timoshevskaya N, Timoshevskiy VA, Keinath MC, Hardy D, Voss SR. A chromosome-scale assembly of the axolotl genome. *Genome research*. 2019; 29(2):317–24. Epub 2019/01/27. <https://doi.org/10.1101/gr.241901.118> PMID: 30679309; PubMed Central PMCID: PMC6360810.
2. Nowoshilow S, Schloissnig S, Fei JF, Dahl A, Pang AWC, Pippel M, et al. The axolotl genome and the evolution of key tissue formation regulators. *Nature*. 2018; 554(7690):50–5. Epub 2018/01/25. <https://doi.org/10.1038/nature25458> PMID: 29364872.
3. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001; 409(6822):860–921. <https://doi.org/10.1038/35057062> PMID: 11237011.
4. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, et al. The genome sequence of *Drosophila melanogaster*. *Science*. 2000; 287(5461):2185–95. <https://doi.org/10.1126/science.287.5461.2185> PMID: 10731132.

5. Kaminker JS, Bergman CM, Kronmiller B, Carlson J, Svirskas R, Patel S, et al. The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. *Genome biology*. 2002; 3(12):RESEARCH0084. <https://doi.org/10.1186/gb-2002-3-12-research0084> PMID: 12537573; PubMed Central PMCID: PMC151186.
6. Chuong EB, Elde NC, Feschotte C. Regulatory activities of transposable elements: from conflicts to benefits. *Nature reviews Genetics*. 2017; 18(2):71–86. <https://doi.org/10.1038/nrg.2016.139> PMID: 27867194; PubMed Central PMCID: PMC5498291.
7. Wells JN, Feschotte C. A Field Guide to Eukaryotic Transposable Elements. *Annual review of genetics*. 2020. Epub 2020/09/22. <https://doi.org/10.1146/annurev-genet-040620-022145> PMID: 32955944.
8. Gorbunova V, Boeke JD, Helfand SL, Sedivy JM. Human Genomics. Sleeping dogs of the genome. *Science*. 2014; 346(6214):1187–8. <https://doi.org/10.1126/science.aaa3177> PMID: 25477445; PubMed Central PMCID: PMC4312280.
9. Chen H, Zheng X, Xiao D, Zheng Y. Age-associated de-repression of retrotransposons in the *Drosophila* fat body, its potential cause and consequence. *Aging Cell*. 2016; 15(3):542–52. <https://doi.org/10.1111/accel.12465> PMID: 27072046; PubMed Central PMCID: PMC4854910.
10. Li W, Prazak L, Chatterjee N, Gruninger S, Krug L, Theodorou D, et al. Activation of transposable elements during aging and neuronal decline in *Drosophila*. *Nature neuroscience*. 2013; 16(5):529–31. <https://doi.org/10.1038/nn.3368> PMID: 23563579; PubMed Central PMCID: PMC3821974.
11. Wood JG, Helfand SL. Chromatin structure and transposable elements in organismal aging. *Frontiers in genetics*. 2013; 4:274. <https://doi.org/10.3389/fgene.2013.00274> PMID: 24363663; PubMed Central PMCID: PMC3849598.
12. Wood JG, Hillenmeyer S, Lawrence C, Chang C, Hosier S, Lightfoot W, et al. Chromatin remodeling in the aging genome of *Drosophila*. *Aging Cell*. 2010; 9(6):971–8. <https://doi.org/10.1111/j.1474-9726.2010.00624.x> PMID: 20961390; PubMed Central PMCID: PMC2980570.
13. Jones BC, Wood JG, Chang C, Tam AD, Franklin MJ, Siegel ER, et al. A somatic piRNA pathway in the *Drosophila* fat body ensures metabolic homeostasis and normal lifespan. *Nature communications*. 2016; 7:13856. <https://doi.org/10.1038/ncomms13856> PMID: 28000665; PubMed Central PMCID: PMC5187580.
14. Wood JG, Jones BC, Jiang N, Chang C, Hosier S, Wickremesinghe P, et al. Chromatin-modifying genetic interventions suppress age-associated transposable element activation and extend life span in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*. 2016; 113(40):11277–82. <https://doi.org/10.1073/pnas.1604621113> PMID: 27621458; PubMed Central PMCID: PMC5056045.
15. Chang YH, Keegan RM, Prazak L, Dubnau J. Cellular labeling of endogenous retrovirus replication (CLEVR) reveals de novo insertions of the gypsy retrotransposable element in cell culture and in both neurons and glial cells of aging fruit flies. *PLoS biology*. 2019; 17(5):e3000278. Epub 2019/05/17. <https://doi.org/10.1371/journal.pbio.3000278> PMID: 31095565; PubMed Central PMCID: PMC6541305.
16. Sousa-Victor P, Ayyaz A, Hayashi R, Qi Y, Madden DT, Lunyak VV, et al. Piwi Is Required to Limit Exhaustion of Aging Somatic Stem Cells. *Cell reports*. 2017; 20(11):2527–37. Epub 2017/09/14. <https://doi.org/10.1016/j.celrep.2017.08.059> PMID: 28903034; PubMed Central PMCID: PMC5901960.
17. Chang YH, Dubnau J. The Gypsy Endogenous Retrovirus Drives Non-Cell-Autonomous Propagation in a *Drosophila* TDP-43 Model of Neurodegeneration. *Curr Biol*. 2019; 29(19):3135–52 e4. Epub 2019/09/10. <https://doi.org/10.1016/j.cub.2019.07.071> PMID: 31495585; PubMed Central PMCID: PMC6783360.
18. Krug L, Chatterjee N, Borges-Monroy R, Hearn S, Liao WW, Morrill K, et al. Retrotransposon activation contributes to neurodegeneration in a *Drosophila* TDP-43 model of ALS. *PLoS genetics*. 2017; 13(3):e1006635. <https://doi.org/10.1371/journal.pgen.1006635> PMID: 28301478; PubMed Central PMCID: PMC5354250.
19. Sun W, Samimi H, Gamez M, Zare H, Frost B. Pathogenic tau-induced piRNA depletion promotes neuronal death through transposable element dysregulation in neurodegenerative tauopathies. *Nature neuroscience*. 2018; 21(8):1038–48. Epub 2018/07/25. <https://doi.org/10.1038/s41593-018-0194-1> PMID: 30038280; PubMed Central PMCID: PMC6095477.
20. Guo C, Jeong HH, Hsieh YC, Klein HU, Bennett DA, De Jager PL, et al. Tau Activates Transposable Elements in Alzheimer's Disease. *Cell reports*. 2018; 23(10):2874–80. Epub 2018/06/07. <https://doi.org/10.1016/j.celrep.2018.05.004> PMID: 29874575; PubMed Central PMCID: PMC6181645.
21. Mirkovic-Hosle M, Forstemann K. Transposon defense by endo-siRNAs, piRNAs and somatic piRNAs in *Drosophila*: contributions of Loqs-PD and R2D2. *PloS one*. 2014; 9(1):e84994. Epub 2014/01/24.

- <https://doi.org/10.1371/journal.pone.0084994> PMID: 24454776; PubMed Central PMCID: PMC3890300.
22. Ghildiyal M, Seitz H, Horwich MD, Li C, Du T, Lee S, et al. Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. *Science*. 2008; 320(5879):1077–81. Epub 2008/04/12. 1157396 [pii] <https://doi.org/10.1126/science.1157396> PMID: 18403677.
 23. Lin KY, Wang WD, Lin CH, Rastegari E, Su YH, Chang YT, et al. Piwi reduction in the aged niche eliminates germline stem cells via Toll-GSK3 signaling. *Nature communications*. 2020; 11(1):3147. Epub 2020/06/21. <https://doi.org/10.1038/s41467-020-16858-6> PMID: 32561720; PubMed Central PMCID: PMC7305233.
 24. Kato M, Takemoto K, Shinkai Y. A somatic role for the histone methyltransferase Setdb1 in endogenous retrovirus silencing. *Nature communications*. 2018; 9(1):1683. Epub 2018/04/29. <https://doi.org/10.1038/s41467-018-04132-9> PMID: 29703894; PubMed Central PMCID: PMC5923290.
 25. Matsui T, Leung D, Miyashita H, Maksakova IA, Miyachi H, Kimura H, et al. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature*. 2010; 464(7290):927–31. Epub 2010/02/19. <https://doi.org/10.1038/nature08858> PMID: 20164836.
 26. Ecco G, Cassano M, Kaulzaric A, Duc J, Coluccio A, Offner S, et al. Transposable Elements and Their KRAB-ZFP Controllers Regulate Gene Expression in Adult Tissues. *Developmental cell*. 2016; 36(6):611–23. Epub 2016/03/24. <https://doi.org/10.1016/j.devcel.2016.02.024> PMID: 27003935; PubMed Central PMCID: PMC4896391.
 27. Rowe HM, Jakobsson J, Mesnard D, Rougemont J, Reynard S, Aktas T, et al. KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature*. 2010; 463(7278):237–40. Epub 2010/01/16. <https://doi.org/10.1038/nature08674> PMID: 20075919.
 28. Castro-Diaz N, Ecco G, Coluccio A, Kapopoulou A, Yazdanpanah B, Friedli M, et al. Evolutionally dynamic L1 regulation in embryonic stem cells. *Genes & development*. 2014; 28(13):1397–409. Epub 2014/06/19. <https://doi.org/10.1101/gad.241661.114> PMID: 24939876; PubMed Central PMCID: PMC4083085.
 29. Tunbak H, Enriquez-Gasca R, Tie CHC, Gould PA, Mlcochova P, Gupta RK, et al. The HUSH complex is a gatekeeper of type I interferon through epigenetic regulation of LINE-1s. *Nature communications*. 2020; 11(1):5387. Epub 2020/11/05. <https://doi.org/10.1038/s41467-020-19170-5> PMID: 33144593; PubMed Central PMCID: PMC7609715.
 30. Robbez-Masson L, Tie CHC, Conde L, Tunbak H, Husovsky C, Tchasovnikarova IA, et al. The HUSH complex cooperates with TRIM28 to repress young retrotransposons and new genes. *Genome research*. 2018; 28(6):836–45. Epub 2018/05/08. <https://doi.org/10.1101/gr.228171.117> PMID: 29728366; PubMed Central PMCID: PMC5991525.
 31. Liu N, Lee CH, Swigut T, Grow E, Gu B, Bassik MC, et al. Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators. *Nature*. 2018; 553(7687):228–32. Epub 2017/12/07. <https://doi.org/10.1038/nature25179> PMID: 29211708; PubMed Central PMCID: PMC5774979.
 32. Tchasovnikarova IA, Timms RT, Matheson NJ, Wals K, Antrobus R, Gottgens B, et al. GENE SILENCING. Epigenetic silencing by the HUSH complex mediates position-effect variegation in human cells. *Science*. 2015; 348(6242):1481–5. Epub 2015/05/30. <https://doi.org/10.1126/science.aaa7227> PMID: 26022416; PubMed Central PMCID: PMC4487827.
 33. Simon M, Van Meter M, Ablaeva J, Ke Z, Gonzalez RS, Taguchi T, et al. LINE1 Derepression in Aged Wild-Type and SIRT6-Deficient Mice Drives Inflammation. *Cell metabolism*. 2019; 29(4):871–85 e5. Epub 2019/03/12. <https://doi.org/10.1016/j.cmet.2019.02.014> PMID: 30853213; PubMed Central PMCID: PMC6449196.
 34. Van Meter M, Kashyap M, Rezazadeh S, Geneva AJ, Morello TD, Seluanov A, et al. SIRT6 represses LINE1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age. *Nature communications*. 2014; 5:5011. <https://doi.org/10.1038/ncomms6011> PMID: 25247314; PubMed Central PMCID: PMC4185372.
 35. Bulut-Karslioglu A, De La Rosa-Velazquez IA, Ramirez F, Barenboim M, Onishi-Seebacher M, Arand J, et al. Suv39h-dependent H3K9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells. *Molecular cell*. 2014; 55(2):277–90. Epub 2014/07/02. <https://doi.org/10.1016/j.molcel.2014.05.029> PMID: 24981170.
 36. Leung DC, Dong KB, Maksakova IA, Goyal P, Appanah R, Lee S, et al. Lysine methyltransferase G9a is required for de novo DNA methylation and the establishment, but not the maintenance, of proviral silencing. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108(14):5718–23. Epub 2011/03/24. <https://doi.org/10.1073/pnas.1014660108> PMID: 21427230; PubMed Central PMCID: PMC3078371.
 37. Vogel MJ, Guelen L, de Wit E, Peric-Hupkes D, Loden M, Talhout W, et al. Human heterochromatin proteins form large domains containing KRAB-ZNF genes. *Genome research*. 2006; 16(12):1493–

504. Epub 2006/10/14. <https://doi.org/10.1101/gr.5391806> PMID: 17038565; PubMed Central PMCID: PMC1665633.
38. Schopp T, Zoch A, Berrens RV, Auchynnikava T, Kabayama Y, Vasiliauskaite L, et al. TEX15 is an essential executor of MIWI2-directed transposon DNA methylation and silencing. *Nature communications*. 2020; 11(1):3739. Epub 2020/07/29. <https://doi.org/10.1038/s41467-020-17372-5> PMID: 32719317; PubMed Central PMCID: PMC7385494.
39. Zoch A, Auchynnikava T, Berrens RV, Kabayama Y, Schopp T, Heep M, et al. SPOCD1 is an essential executor of piRNA-directed de novo DNA methylation. *Nature*. 2020; 584(7822):635–9. Epub 2020/07/17. <https://doi.org/10.1038/s41586-020-2557-5> PMID: 32674113.
40. Carmell MA, Girard A, van de Kant HJ, Bourc'his D, Bestor TH, de Rooij DG, et al. MIWI2 Is Essential for Spermatogenesis and Repression of Transposons in the Mouse Male Germline. *Developmental cell*. 2007. <https://doi.org/10.1016/j.devcel.2007.03.001> PMID: 17395546.
41. Kuramochi-Miyagawa S, Watanabe T, Gotoh K, Totoki Y, Toyoda A, Ikawa M, et al. DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes & development*. 2008; 22(7):908–17. Epub 2008/04/03. 22/7/908 [pii] <https://doi.org/10.1101/gad.1640708> PMID: 18381894; PubMed Central PMCID: PMC2279202.
42. Shoji M, Tanaka T, Hosokawa M, Reuter M, Stark A, Kato Y, et al. The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline. *Developmental cell*. 2009; 17(6):775–87. Epub 2010/01/12. S1534-5807(09)00434-1 [pii] <https://doi.org/10.1016/j.devcel.2009.10.012> PMID: 20059948.
43. Zamudio N, Barau J, Teissandier A, Walter M, Borsos M, Servant N, et al. DNA methylation restrains transposons from adopting a chromatin signature permissive for meiotic recombination. *Genes & development*. 2015; 29(12):1256–70. Epub 2015/06/26. <https://doi.org/10.1101/gad.257840.114> PMID: 26109049; PubMed Central PMCID: PMC4495397.
44. Pastor WA, Stroud H, Nee K, Liu W, Pezic D, Manakov S, et al. MORC1 represses transposable elements in the mouse male germline. *Nature communications*. 2014; 5:5795. Epub 2014/12/17. <https://doi.org/10.1038/ncomms6795> PMID: 25503965; PubMed Central PMCID: PMC4268658.
45. Bedrosian TA, Quayle C, Novaresi N, Gage FH. Early life experience drives structural variation of neural genomes in mice. *Science*. 2018; 359(6382):1395–9. <https://doi.org/10.1126/science.aah3378> PMID: 29567711.
46. Erwin JA, Paquola AC, Singer T, Gallina I, Novotny M, Quayle C, et al. L1-associated genomic regions are deleted in somatic cells of the healthy human brain. *Nature neuroscience*. 2016; 19(12):1583–91. <https://doi.org/10.1038/nn.4388> PMID: 27618310; PubMed Central PMCID: PMC5127747.
47. Marchetto MC, Narvaiza I, Denli AM, Benner C, Lazzarini TA, Nathanson JL, et al. Differential L1 regulation in pluripotent stem cells of humans and apes. *Nature*. 2013; 503(7477):525–9. <https://doi.org/10.1038/nature12686> PMID: 24153179.
48. Muotri AR, Chu VT, Marchetto MC, Deng W, Moran JV, Gage FH. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature*. 2005; 435(7044):903–10. <https://doi.org/10.1038/nature03663> PMID: 15959507.
49. Richardson SR, Gerdes P, Gerhardt DJ, Sanchez-Luque FJ, Bodea GO, Munoz-Lopez M, et al. Heritable L1 retrotransposition in the mouse primordial germline and early embryo. *Genome research*. 2017; 27(8):1395–405. <https://doi.org/10.1101/gr.219022.116> PMID: 28483779; PubMed Central PMCID: PMC5538555.
50. Baillie JK, Barnett MW, Upton KR, Gerhardt DJ, Richmond TA, De Sapio F, et al. Somatic retrotransposition alters the genetic landscape of the human brain. *Nature*. 2011; 479(7374):534–7. <https://doi.org/10.1038/nature10531> PMID: 22037309; PubMed Central PMCID: PMC3224101.
51. Evrony GD, Lee E, Park PJ, Walsh CA. Resolving rates of mutation in the brain using single-neuron genomics. *eLife*. 2016; 5. <https://doi.org/10.7554/eLife.12966> PMID: 26901440; PubMed Central PMCID: PMC4805530.
52. Evrony GD, Lee E, Mehta BK, Benjamini Y, Johnson RM, Cai X, et al. Cell lineage analysis in human brain using endogenous retroelements. *Neuron*. 2015; 85(1):49–59. <https://doi.org/10.1016/j.neuron.2014.12.028> PMID: 25569347; PubMed Central PMCID: PMC4299461.
53. Poduri A, Evrony GD, Cai X, Walsh CA. Somatic mutation, genomic variation, and neurological disease. *Science*. 2013; 341(6141):1237758. <https://doi.org/10.1126/science.1237758> PMID: 23828942; PubMed Central PMCID: PMC3909954.
54. De Cecco M, Criscione SW, Peterson AL, Neretti N, Sedivy JM, Kreiling JA. Transposable elements become active and mobile in the genomes of aging mammalian somatic tissues. *Aging (Albany NY)*. 2013; 5(12):867–83. <https://doi.org/10.18632/aging.100621> PMID: 24323947; PubMed Central PMCID: PMC3883704.

55. De Cecco M, Criscione SW, Peckham EJ, Hillenmeyer S, Hamm EA, Manivannan J, et al. Genomes of replicatively senescent cells undergo global epigenetic changes leading to gene silencing and activation of transposable elements. *Aging Cell*. 2013; 12(2):247–56. <https://doi.org/10.1111/accel.12047> PMID: 23360310; PubMed Central PMCID: PMC3618682.
56. De Cecco M, Ito T, Petrashen AP, Elias AE, Skvir NJ, Criscione SW, et al. L1 drives IFN in senescent cells and promotes age-associated inflammation. *Nature*. 2019; 566(7742):73–8. Epub 2019/02/08. <https://doi.org/10.1038/s41586-018-0784-9> PMID: 30728521; PubMed Central PMCID: PMC6519963.
57. deHaro D, Kines KJ, Sokolowski M, Dauchy RT, Strevva VA, Hill SM, et al. Regulation of L1 expression and retrotransposition by melatonin and its receptor: implications for cancer risk associated with light exposure at night. *Nucleic acids research*. 2014; 42(12):7694–707. Epub 2014/06/11. <https://doi.org/10.1093/nar/gku503> PMID: 24914052; PubMed Central PMCID: PMC4081101.
58. Rodic N, Sharma R, Sharma R, Zampella J, Dai L, Taylor MS, et al. Long interspersed element-1 protein expression is a hallmark of many human cancers. *Am J Pathol*. 2014; 184(5):1280–6. Epub 2014/03/13. <https://doi.org/10.1016/j.ajpath.2014.01.007> PMID: 24607009; PubMed Central PMCID: PMC4005969.
59. Huang CR, Schneider AM, Lu Y, Niranjana T, Shen P, Robinson MA, et al. Mobile interspersed repeats are major structural variants in the human genome. *Cell*. 2010; 141(7):1171–82. Epub 2010/07/07. <https://doi.org/10.1016/j.cell.2010.05.026> PMID: 20602999; PubMed Central PMCID: PMC2943426.
60. Treiber CD, Waddell S. Resolving the prevalence of somatic transposition in *Drosophila*. *eLife*. 2017; 6. <https://doi.org/10.7554/eLife.28297> PMID: 28742021; PubMed Central PMCID: PMC5553932.
61. Zhuang J, Wang J, Theurkauf W, Weng Z. TEMP: a computational method for analyzing transposable element polymorphism in populations. *Nucleic acids research*. 2014; 42(11):6826–38. <https://doi.org/10.1093/nar/gku323> PMID: 24753423; PubMed Central PMCID: PMC4066757.
62. Perrat PN, DasGupta S, Wang J, Theurkauf W, Weng Z, Rosbash M, et al. Transposition-driven genomic heterogeneity in the *Drosophila* brain. *Science*. 2013; 340(6128):91–5. <https://doi.org/10.1126/science.1231965> PMID: 23559253; PubMed Central PMCID: PMC3887341.
63. Khurana JS, Wang J, Xu J, Koppetsch BS, Thomson TC, Nowosielska A, et al. Adaptation to P element transposon invasion in *Drosophila melanogaster*. *Cell*. 2011; 147(7):1551–63. Epub 2011/12/27. S0092-8674(11)01436-X [pii] <https://doi.org/10.1016/j.cell.2011.11.042> PMID: 22196730; PubMed Central PMCID: PMC3246748.
64. Sytnikova YA, Rahman R, Chirn GW, Clark JP, Lau NC. Transposable element dynamics and PIWI regulation impacts lncRNA and gene expression diversity in *Drosophila* ovarian cell cultures. *Genome research*. 2014; 24(12):1977–90. <https://doi.org/10.1101/gr.178129.114> PMID: 25267525; PubMed Central PMCID: PMC4248314.
65. Rahman R, Chirn GW, Kanodia A, Sytnikova YA, Brembs B, Bergman CM, et al. Unique transposon landscapes are pervasive across *Drosophila melanogaster* genomes. *Nucleic acids research*. 2015; 43(22):10655–72. <https://doi.org/10.1093/nar/gkv1193> PMID: 26578579; PubMed Central PMCID: PMC4678822.
66. Penke TJ, McKay DJ, Strahl BD, Matera AG, Duronio RJ. Direct interrogation of the role of H3K9 in metazoan heterochromatin function. *Genes & development*. 2016; 30(16):1866–80. Epub 2016/08/28. <https://doi.org/10.1101/gad.286278.116> PMID: 27566777; PubMed Central PMCID: PMC5024684.
67. Ninova M, Chen YA, Godneeva B, Rogers AK, Luo Y, Fejes Toth K, et al. Su(var)2-10 and the SUMO Pathway Link piRNA-Guided Target Recognition to Chromatin Silencing. *Molecular cell*. 2020; 77(3):556–70 e6. Epub 2020/01/07. <https://doi.org/10.1016/j.molcel.2019.11.012> PMID: 31901446; PubMed Central PMCID: PMC7007863.
68. Miller DE, Smith CB, Kazemi NY, Cockrell AJ, Arvanitakas AV, Blumenstiel JP, et al. Whole-Genome Analysis of Individual Meiotic Events in *Drosophila melanogaster* Reveals That Noncrossover Gene Conversions Are Insensitive to Interference and the Centromere Effect. *Genetics*. 2016; 203(1):159–71. <https://doi.org/10.1534/genetics.115.186486> PMID: 26944917; PubMed Central PMCID: PMC4858771.
69. mod EC, Roy S, Ernst J, Kharchenko PV, Kheradpour P, Negre N, et al. Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science*. 2010; 330(6012):1787–97. <https://doi.org/10.1126/science.1198374> PMID: 21177974; PubMed Central PMCID: PMC3192495.
70. Brown EJ, Nguyen AH, Bachtrog D. The Y chromosome may contribute to sex-specific ageing in *Drosophila*. *Nat Ecol Evol*. 2020; 4(6):853–62. Epub 2020/04/22. <https://doi.org/10.1038/s41559-020-1179-5> PMID: 32313175; PubMed Central PMCID: PMC7274899.
71. Srivastav SP, Rahman R, Ma Q, Pierre J, Bandyopadhyay S, Lau NC. Har-P, a short P-element variant, weaponizes P-transposase to severely impair *Drosophila* development. *eLife*. 2019; 8. <https://doi.org/10.7554/eLife.49948> PMID: 31845649; PubMed Central PMCID: PMC6917496.

72. Wang L, Dou K, Moon S, Tan FJ, Zhang ZZ. Hijacking Oogenesis Enables Massive Propagation of LINE and Retroviral Transposons. *Cell*. 2018; 174(5):1082–94 e12. <https://doi.org/10.1016/j.cell.2018.06.040> PMID: 30057117.
73. Okamura K, Balla S, Martin R, Liu N, Lai EC. Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in *Drosophila melanogaster*. *Nat Struct Mol Biol*. 2008; 15(9):998. Epub 2008/09/05. nsmb0908-998c [pii] <https://doi.org/10.1038/nsmb0908-998c> PMID: 18769470.
74. Kawamura Y, Saito K, Kin T, Ono Y, Asai K, Sunohara T, et al. *Drosophila* endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature*. 2008; 453(7196):793–7. Epub 2008/05/09. nature06938 [pii] <https://doi.org/10.1038/nature06938> PMID: 18463636.
75. Czech B, Malone CD, Zhou R, Stark A, Schlingeheyde C, Dus M, et al. An endogenous small interfering RNA pathway in *Drosophila*. *Nature*. 2008; 453(7196):798–802. <https://doi.org/10.1038/nature07007> PMID: 18463631.
76. Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, et al. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell*. 2007; 128(6):1089–103. Epub 2007/03/10. S0092-8674(07)00257-7 [pii] <https://doi.org/10.1016/j.cell.2007.01.043> PMID: 17346786.
77. Saito K, Nishida KM, Mori T, Kawamura Y, Miyoshi K, Nagami T, et al. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes & development*. 2006; 20(16):2214–22. Epub 2006/08/03. gad.1454806 [pii] <https://doi.org/10.1101/gad.1454806> PMID: 16882972; PubMed Central PMCID: PMC1553205.
78. Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science*. 2006; 313(5785):320–4. <https://doi.org/10.1126/science.1129333> PMID: 16809489.
79. Moller HD, Mohiyuddin M, Prada-Luengo I, Sailani MR, Halling JF, Plomgaard P, et al. Circular DNA elements of chromosomal origin are common in healthy human somatic tissue. *Nature communications*. 2018; 9(1):1069. <https://doi.org/10.1038/s41467-018-03369-8> PMID: 29540679; PubMed Central PMCID: PMC5852086.
80. Shmookler Reis RJ, Lumpkin CK Jr., McGill JR, Riabowol KT, Goldstein S. Genome alteration during in vitro and in vivo aging: amplification of extrachromosomal circular DNA molecules containing a chromosomal sequence of variable repeat frequency. *Cold Spring Harbor symposia on quantitative biology*. 1983; 47 Pt 2:1135–9. <https://doi.org/10.1101/sqb.1983.047.01.128> PMID: 6574863.
81. Shibata Y, Kumar P, Layer R, Willcox S, Gagan JR, Griffith JD, et al. Extrachromosomal microDNAs and chromosomal microdeletions in normal tissues. *Science*. 2012; 336(6077):82–6. <https://doi.org/10.1126/science.1213307> PMID: 22403181; PubMed Central PMCID: PMC3703515.
82. Misra R, Shih A, Rush M, Wong E, Schmid CW. Cloned extrachromosomal circular DNA copies of the human transposable element THE-1 are related predominantly to a single type of family member. *J Mol Biol*. 1987; 196(2):233–43. [https://doi.org/10.1016/0022-2836\(87\)90687-5](https://doi.org/10.1016/0022-2836(87)90687-5) PMID: 2821286.
83. Turner KM, Deshpande V, Beyter D, Koga T, Rusert J, Lee C, et al. Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. *Nature*. 2017; 543(7643):122–5. <https://doi.org/10.1038/nature21356> PMID: 28178237; PubMed Central PMCID: PMC5334176.
84. Dillon LW, Kumar P, Shibata Y, Wang YH, Willcox S, Griffith JD, et al. Production of Extrachromosomal MicroDNAs Is Linked to Mismatch Repair Pathways and Transcriptional Activity. *Cell reports*. 2015; 11(11):1749–59. <https://doi.org/10.1016/j.celrep.2015.05.020> PMID: 26051933; PubMed Central PMCID: PMC4481157.
85. Vu W, Nuzhdin S. Genetic variation of copia suppression in *Drosophila melanogaster*. *Heredity (Edinb)*. 2011; 106(2):207–17. <https://doi.org/10.1038/hdy.2010.41> PMID: 20606692; PubMed Central PMCID: PMC3183883.
86. Flavell AJ, Ish-Horowicz D. Extrachromosomal circular copies of the eukaryotic transposable element copia in cultured *Drosophila* cells. *Nature*. 1981; 292(5824):591–5. <https://doi.org/10.1038/292591a0> PMID: 6265802.
87. Cohen S, Yacobi K, Segal D. Extrachromosomal circular DNA of tandemly repeated genomic sequences in *Drosophila*. *Genome research*. 2003; 13(6A):1133–45. <https://doi.org/10.1101/gr.907603> PMID: 12799349; PubMed Central PMCID: PMC403641.
88. Mossie KG, Young MW, Varmus HE. Extrachromosomal DNA forms of copia-like transposable elements, F elements and middle repetitive DNA sequences in *Drosophila melanogaster*. Variation in cultured cells and embryos. *J Mol Biol*. 1985; 182(1):31–43. [https://doi.org/10.1016/0022-2836\(85\)90025-7](https://doi.org/10.1016/0022-2836(85)90025-7) PMID: 2582138.
89. Ilyin YV, Schuppe NG, Lyubomirskaya NV, Gorelova TV, Arkhipova IR. Circular copies of mobile dispersed genetic elements in cultured *Drosophila melanogaster* cells. *Nucleic acids research*. 1984; 12

- (19):7517–31. <https://doi.org/10.1093/nar/12.19.7517> PMID: 6093043; PubMed Central PMCID: PMC320178.
90. Lanciano S, Carpentier MC, Llauro C, Jobet E, Robakowska-Hyzorek D, Lasserre E, et al. Sequencing the extrachromosomal circular mobilome reveals retrotransposon activity in plants. *PLoS genetics*. 2017; 13(2):e1006630. <https://doi.org/10.1371/journal.pgen.1006630> PMID: 28212378; PubMed Central PMCID: PMC5338827.
 91. Lee YS, Nakahara K, Pham JW, Kim K, He Z, Sontheimer EJ, et al. Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell*. 2004; 117(1):69–81. Epub 2004/04/07. S0092867404002612 [pii]. [https://doi.org/10.1016/s0092-8674\(04\)00261-2](https://doi.org/10.1016/s0092-8674(04)00261-2) PMID: 15066283.
 92. Senti KA, Jurczak D, Sachidanandam R, Brennecke J. piRNA-guided slicing of transposon transcripts enforces their transcriptional silencing via specifying the nuclear piRNA repertoire. *Genes & development*. 2015; 29(16):1747–62. <https://doi.org/10.1101/gad.267252.115> PMID: 26302790; PubMed Central PMCID: PMC4561483.
 93. Belancio VP, Roy-Engel AM, Pochampally RR, Deininger P. Somatic expression of LINE-1 elements in human tissues. *Nucleic acids research*. 2010; 38(12):3909–22. Epub 2010/03/11. <https://doi.org/10.1093/nar/gkq132> PMID: 20215437; PubMed Central PMCID: PMC2896524.
 94. Mumford PW, Romero MA, Osburn SC, Roberson PA, Vann CG, Mobley CB, et al. Skeletal muscle LINE-1 retrotransposon activity is upregulated in older versus younger rats. *Am J Physiol Regul Integr Comp Physiol*. 2019; 317(3):R397–R406. Epub 2019/06/13. <https://doi.org/10.1152/ajpregu.00110.2019> PMID: 31188650.
 95. Chinen M, Lei EP. *Drosophila* Argonaute2 turnover is regulated by the ubiquitin proteasome pathway. *Biochem Biophys Res Commun*. 2017; 483(3):951–7. Epub 2017/01/15. <https://doi.org/10.1016/j.bbrc.2017.01.039> PMID: 28087276; PubMed Central PMCID: PMC5279065.
 96. Lewis A, Berkuyrek AC, Greiner A, Sawh AN, Vashisht A, Merrett S, et al. A Family of Argonaute-Interacting Proteins Gates Nuclear RNAi. *Molecular cell*. 2020; 78(5):862–75 e8. Epub 2020/04/30. <https://doi.org/10.1016/j.molcel.2020.04.007> PMID: 32348780.
 97. Clark JP, Rahman R, Yang N, Yang LH, Lau NC. *Drosophila* PAF1 Modulates PIWI/piRNA Silencing Capacity. *Curr Biol*. 2017; 27(17):2718–26 e4. <https://doi.org/10.1016/j.cub.2017.07.052> PMID: 28844648; PubMed Central PMCID: PMC5611886.
 98. Kowalik KM, Shimada Y, Flury V, Stadler MB, Batki J, Buhler M. The Paf1 complex represses small-RNA-mediated epigenetic gene silencing. *Nature*. 2015; 520(7546):248–52. <https://doi.org/10.1038/nature14337> PMID: 25807481; PubMed Central PMCID: PMC4398878.
 99. Yamanaka S, Mehta S, Reyes-Turcu FE, Zhuang F, Fuchs RT, Rong Y, et al. RNAi triggered by specialized machinery silences developmental genes and retrotransposons. *Nature*. 2013; 493(7433):557–60. <https://doi.org/10.1038/nature11716> PMID: 23151475; PubMed Central PMCID: PMC3554839.
 100. Pfeiffer BD, Ngo TT, Hibbard KL, Murphy C, Jenett A, Truman JW, et al. Refinement of tools for targeted gene expression in *Drosophila*. *Genetics*. 2010; 186(2):735–55. <https://doi.org/10.1534/genetics.110.119917> PMID: 20697123; PubMed Central PMCID: PMC2942869.
 101. Fast I, Rosenkranz D. Temperature-dependent small RNA expression in *Drosophila melanogaster*. *RNA biology*. 2018; 15(3):308–13. Epub 2018/01/19. <https://doi.org/10.1080/15476286.2018.1429881> PMID: 29345184; PubMed Central PMCID: PMC5927726.
 102. Kelleher ES, Jaweria J, Akoma U, Ortega L, Tang W. QTL mapping of natural variation reveals that the developmental regulator *bruno* reduces tolerance to P-element transposition in the *Drosophila* female germline. *PLoS biology*. 2018; 16(10):e2006040. <https://doi.org/10.1371/journal.pbio.2006040> PMID: 30376574; PubMed Central PMCID: PMC6207299.
 103. Moon S, Cassani M, Lin YA, Wang L, Dou K, Zhang ZZ. A Robust Transposon-Endogenizing Response from Germline Stem Cells. *Developmental cell*. 2018; 47(5):660–71 e3. <https://doi.org/10.1016/j.devcel.2018.10.011> PMID: 30393075.
 104. Genencher B, Durdevic Z, Hanna K, Zinkl D, Mobin MB, Senturk N, et al. Mutations in Cytosine-5 tRNA Methyltransferases Impact Mobile Element Expression and Genome Stability at Specific DNA Repeats. *Cell reports*. 2018; 22(7):1861–74. Epub 2018/02/15. <https://doi.org/10.1016/j.celrep.2018.01.061> PMID: 29444437.
 105. Robinow S, White K. The locus *elav* of *Drosophila melanogaster* is expressed in neurons at all developmental stages. *Developmental biology*. 1988; 126(2):294–303. Epub 1988/04/01. [https://doi.org/10.1016/0012-1606\(88\)90139-x](https://doi.org/10.1016/0012-1606(88)90139-x) PMID: 3127258.
 106. Amarasinghe SL, Su S, Dong X, Zappia L, Ritchie ME, Gouli Q. Opportunities and challenges in long-read sequencing data analysis. *Genome biology*. 2020; 21(1):30. Epub 2020/02/09. <https://doi.org/10.1186/s13059-020-1935-5> PMID: 32033565; PubMed Central PMCID: PMC7006217.

107. Li W, Jin Y, Prazak L, Hammell M, Dubnau J. Transposable elements in TDP-43-mediated neurodegenerative disorders. *PloS one*. 2012; 7(9):e44099. <https://doi.org/10.1371/journal.pone.0044099> PMID: 22957047; PubMed Central PMCID: PMC3434193.
108. Romero MA, Mumford PW, Roberson PA, Osburn SC, Parry HA, Kavazis AN, et al. Five months of voluntary wheel running downregulates skeletal muscle LINE-1 gene expression in rats. *Am J Physiol Cell Physiol*. 2019; 317(6):C1313–C23. Epub 2019/10/17. <https://doi.org/10.1152/ajpcell.00301.2019> PMID: 31618076.
109. Nandakumar S, Grushko O, Buttitta LA. Polyploidy in the adult *Drosophila* brain. *eLife*. 2020; 9. Epub 2020/08/26. <https://doi.org/10.7554/eLife.54385> PMID: 32840209; PubMed Central PMCID: PMC7447450.
110. Yarosh W, Spradling AC. Incomplete replication generates somatic DNA alterations within *Drosophila* polytene salivary gland cells. *Genes & development*. 2014; 28(16):1840–55. <https://doi.org/10.1101/gad.245811.114> PMID: 25128500; PubMed Central PMCID: PMC4197960.
111. Vienne J, Spann R, Guo F, Rosbash M. Age-Related Reduction of Recovery Sleep and Arousal Threshold in *Drosophila*. *Sleep*. 2016; 39(8):1613–24. Epub 2016/06/17. <https://doi.org/10.5665/sleep.6032> PMID: 27306274; PubMed Central PMCID: PMC4945321.
112. Ren X, Yang Z, Mao D, Chang Z, Qiao HH, Wang X, et al. Performance of the Cas9 nickase system in *Drosophila melanogaster*. *G3 (Bethesda)*. 2014; 4(10):1955–62. Epub 2014/08/17. <https://doi.org/10.1534/g3.114.013821> PMID: 25128437; PubMed Central PMCID: PMC4199701.
113. Lee SK, Xue Y, Shen W, Zhang Y, Joo Y, Ahmad M, et al. Topoisomerase 3beta interacts with RNAi machinery to promote heterochromatin formation and transcriptional silencing in *Drosophila*. *Nature communications*. 2018; 9(1):4946. Epub 2018/11/25. <https://doi.org/10.1038/s41467-018-07101-4> PMID: 30470739; PubMed Central PMCID: PMC6251927.
114. Nazer E, Dale RK, Chinen M, Radmanesh B, Lei EP. Argonaute2 and LaminB modulate gene expression by controlling chromatin topology. *PLoS genetics*. 2018; 14(3):e1007276. Epub 2018/03/13. <https://doi.org/10.1371/journal.pgen.1007276> PMID: 29529026; PubMed Central PMCID: PMC5864089.
115. Hain D, Bettencourt BR, Okamura K, Csorba T, Meyer W, Jin Z, et al. Natural variation of the amino-terminal glutamine-rich domain in *Drosophila argonaute2* is not associated with developmental defects. *PloS one*. 2010; 5(12):e15264. <https://doi.org/10.1371/journal.pone.0015264> PMID: 21253006; PubMed Central PMCID: PMC3002974.
116. Zamparini AL, Davis MY, Malone CD, Vieira E, Zavadii J, Sachidanandam R, et al. Vreteno, a gonad-specific protein, is essential for germline development and primary piRNA biogenesis in *Drosophila*. *Development*. 2011; 138(18):4039–50. Epub 2011/08/13. dev.069187 [pii] <https://doi.org/10.1242/dev.069187> PMID: 21831924; PubMed Central PMCID: PMC3160098.
117. Shang Y, Griffith LC, Rosbash M. Light-arousal and circadian photoreception circuits intersect at the large PDF cells of the *Drosophila* brain. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105(50):19587–94. <https://doi.org/10.1073/pnas.0809577105> PMID: 19060186; PubMed Central PMCID: PMC2596742.
118. Nagoshi E, Sugino K, Kula E, Okazaki E, Tachibana T, Nelson S, et al. Dissecting differential gene expression within the circadian neuronal circuit of *Drosophila*. *Nature neuroscience*. 2010; 13(1):60–8. Epub 2009/12/08. <https://doi.org/10.1038/nn.2451> PMID: 19966839; PubMed Central PMCID: PMC3878269.
119. Li H. Seqtk is a fast and lightweight tool for processing sequences in the FASTA or FASTQ format. <https://github.com/lh3/seqtk/releases/tag/v1.3>. 2018.
120. Chirn GW, Rahman R, Sytnikova YA, Matts JA, Zeng M, Gerlach D, et al. Conserved piRNA Expression from a Distinct Set of piRNA Cluster Loci in Eutherian Mammals. *PLoS genetics*. 2015; 11(11):e1005652. <https://doi.org/10.1371/journal.pgen.1005652> PMID: 26588211; PubMed Central PMCID: PMC4654475.
121. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010; 26(5):589–95. <https://doi.org/10.1093/bioinformatics/btp698> PMID: 20080505; PubMed Central PMCID: PMC2828108.
122. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology*. 2009; 10(3):R25. <https://doi.org/10.1186/gb-2009-10-3-r25> PMID: 19261174; PubMed Central PMCID: PMC2690996.
123. Han SK, Lee D, Lee H, Kim D, Son HG, Yang JS, et al. OASIS 2: online application for survival analysis 2 with features for the analysis of maximal lifespan and healthspan in aging research. *Oncotarget*. 2016; 7(35):56147–52. Epub 2016/08/17. <https://doi.org/10.18632/oncotarget.11269> PMID: 27528229; PubMed Central PMCID: PMC5302902.