Title: Proper control of R-loop homeostasis is required for maintenance of gene expression and neuronal function during aging

Running title: R-loop dysregulation in the aging eye

Authors: Juan Jauregui-Lozano[1], Alyssa Easton[2], Spencer Escobedo[1], Nadia A. Lanman[3,4], Vikki M. Weake[1,4] and Hana Hall[1,5]

[1] Department of Biochemistry, Purdue University, West Lafayette, Indiana

[2] Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, Indiana

[3] Department of Comparative Pathobiology, College of Veterinary Medicine, Purdue University, West Lafayette, IN.

[4] Purdue University Center for Cancer Research, Purdue University, West Lafayette, Indiana

[5] Purdue Institute for Integrative Neuroscience, Purdue University, West Lafayette, IN.

Correspondence

Hana Hall, Department of Biochemistry, Purdue University, 175 S. University Street, West Lafayette, IN 47907, USA, Email: hallh@purdue.edu

Funding information: Research reported in this publication was supported by the National Eye Institute of the NIH under Award Number 1R21EY031024-01 (to V.M.W. and H.H.) and 2R01EY024905 to V.M.W.

ABSTRACT

Age-related loss of cellular function and increased cell death are characteristic hallmarks of aging. While defects in gene expression and RNA metabolism have been linked with age-associated human neuropathies, it is not clear how the changes that occur during aging contribute to loss of gene expression homeostasis. R-loops are DNA-RNA hybrids that typically form co-transcriptionally via annealing of the nascent RNA to the template DNA strand, displacing the non-template DNA strand. Dysregulation of R-loop homeostasis has been associated with both transcriptional impairment and genome instability. Importantly, a growing body of evidence links R-loop accumulation with cellular dysfunction, increased cell death and chronic disease onset. Here, we characterized the R-loop landscape in aging *Drosophila melanogaster* photoreceptor neurons. Our data shows that transcribed genes in *Drosophila* photoreceptor neurons accumulate R-loops during aging. Further, our data reveals an association between age-related R-loop accumulation and decreased expression of long and highly expressed genes. Lastly, we show that photoreceptor-specific depletion of Top3β, a DNA/RNA topoisomerase associated with R-loop resolution, leads to both downregulation of of long genes with neuronal function and decreased visual response in flies. Together, our studies present novel data showing increased levels of R-loop in aging photoreceptor neurons, highlighting the link between dysregulation of R-loop homeostasis, gene expression and visual function.

KEYWORDS: *Drosophila*, eye, aging, R-loop, transcription, visual, photoreceptors, neurons.

1 INTRODUCTION

Aging is a process characterized by a time-dependent decline in physiological homeostasis that eventually leads to a loss of organismal function and increased incidence of death (Partridge & Mangel, 1999). Characteristic functional changes include loss of gene and protein expression, mitochondrial dysfunction, cellular senescence, and stem cell exhaustion (López-Otín et al., 2013a). Aging is also a major contributor to development of many chronic diseases including ocular disease (Bonnel et al., 2003; Curcio, 2001; Gao & Hollyfield, 1992; Weale, 1998). Age-related vision loss and maculopathy have been associated with decreased density of retinal cells, including photoreceptors (Panda-Jonas et al., 1995). Specifically, the age-related decline in photoreceptors affects predominantly rods rather than cones (Jackson et al., 2002). In addition, emerging evidence links age-related neurological diseases, including retinal neuropathies, to defects in gene expression and RNA metabolism (Parapuram et al., 2010). Nonetheless, the molecular mechanisms that contribute to the age-associated susceptibility of the eye to disease development are poorly understood.

R-loops are three-stranded nucleic acid structures consisting of an RNA-DNA hybrid and a misplaced single-strand of DNA(Aguilera & García-Muse, 2012). They typically form during transcription in organisms ranging from yeast to humans (Aguilera & Gomez-Gonzalez, 2017) and play a significant role in normal cellular physiology, being required for the initiation of mitochondrial replication and class switch recombination (Mackay et al., 2020). Moreover, recent studies suggest that R-loops can dynamically regulate gene expression (Niehrs & Luke, 2020). Due to their enrichment at gene termini, R-loops also modulate gene expression by preventing DNA methylation or limiting transcription factor access to promoters and facilitating efficient transcription termination at 3’-ends (Boque-Sastre et al., 2015; Skourti-Stathaki et al., 2011, 2014). Moreover, some genomic loci are more susceptible to R-loop formation: for example, R-loops correlate positively with increased transcription, gene length, and GC content

(Chan et al., 2014; Ginno et al., 2012; Gómez-González et al., 2011). Besides DNA sequence, R-loop formation is also determined by DNA topology. Topoisomerases are highly conserved enzymes that resolve topological stress during transcription and replication (McKinnon, 2016). Importantly, recent studies have shown that topoisomerases are critical for proper neuronal function (McKinnon, 2016). Topoisomerase 3β (Top3β) is a member of the type IA superfamily of topoisomerases, which unwind negatively supercoiled DNA formed during transcription and replication, an activity that prevents R-loop formation (Drolet, 2006; Drolet et al., 1994, 1995). Loss of Top3β is associated with neurological disorders (Mackay et al., 2020) and has been shown to reduce lifespan in mice (Kwan & Wang, 2001).

Although R-loops are normal biological structures, their persistent formation is a major source of spontaneous DNA damage that can lead to transcriptional dysregulation and genome instability (Aguilera & García-Muse, 2012), two early hallmarks of aging (López-Otín et al., 2013a). However, our understanding of the link between formation of R-loops and their impact on gene expression and cellular function during aging are quite limited. Due to high levels of transcription and alternative splicing, retinal cells and particularly photoreceptor neurons may be highly sensitive to RNA metabolism dysregulation (Fernandopulle et al., 2021). *Drosophila* compound eyes contain approximately 800 units called ommatidia with each consisting of 20 cells, including eight photoreceptor (PR) neurons (Carthew, 2007). The six outer PRs (R1-R6) expressing Rhodopsin 1 (Rh1) are mainly responsible for black and white vision and motion, and are similar to human rods. The inner PRs (R7 and R8) express Rh3/4 and Rh5/6, respectively, are responsible for color vision (Hardie, 1985) and resemble human cones. To characterize how aging impacts the genomic R-loop landscape in *Drosophila* photoreceptors, we isolated genetically labeled outer PRs using our recently improved nuclei immuno-enrichment (NIE) method (Jauregui-Lozano et al., 2021a) and performed MapR coupled with next generation sequencing (Yan & Sarma, 2020a) to map the genomic location of R-loops.

Here, we show that R-loop levels in photoreceptor neurons increase with age and that age-associated changes in the R-loop landscape are highly dynamic. Further, we show that genes enriched for genomic characteristic associated with R-loops, such as transcript levels and GC content, decrease expression with age. Finally, we show that depletion of DNA/RNA topoisomerase Top3β results in increased R-loop levels, reduced expression of long genes with neuronal function and reduced visual response. Together, our data show that aging is associated with increased levels of R-loops that correlate with decreased gene expression and visual function in photoreceptor neurons.

2 RESULTS

2.1 Aging photoreceptor neurons show increased global levels of R-loops that correlate with loss of function and precede age-associated retinal degeneration

To examine the global levels of R-loops in photoreceptor neurons, we tagged the outer nuclear membrane of R1-R6 with GFP fused to the KASH domain of Msp300 protein using the Rh1-Gal4 driver, then isolated outer PR nuclei from the head homogenate with our tissue-specific (NIE method (Jauregui-Lozano et al., 2021b). Next, we aged flies over the course of 50 days post-eclosion (emergence from the pupae), extracted genomic DNA from isolated PR nuclei at three time points (Figure 1A). We then assessed the global R-loop signal with a DNA slot blot assay using the S9.6 antibody, which recognizes RNA:DNA hybrids (Boguslawski et al., 1986). Specificity of S9.6 antibody towards RNA-DNA hybrids was shown by pre-treatment of DNA with ribonuclease H1 (RNase H1) that resulted in significant decrease of S9.6 signal (Figure 1B). Signal quantification showed over 20% increase in R-loop levels in PRs isolated from middle-aged, 30-day old flies (Rauser et al., 2005)(Rauser et al., 2005) as compared to that in young, 10-day old flies. This trend continued with a significant increase of nearly 80% in global R-loop levels at day 50 (Figure 1B-C). These observations suggest that R-loops start accumulating early during photoreceptor aging, at a time point where flies show decreased visual function

(Hall et al., 2017). Importantly, using optical neutralization, which measures photoreceptor structural integrity with light microscopy, we observed no retinal degeneration by middle age, with a stochastic loss of rhabdomeres occurring after day 40 (Figure 1, Suppl. 1). Thus, our data shows that process of R-loop accumulation during aging precedes age-related retinal degeneration.

2.2 Mapping genome-wide distribution of R-loops in aging PR neurons

To determine the genomic landscape of R-loops in aging PRs, we coupled our NIE approach with MapR, a recently published R-loop mapping strategy based on the specificity of RNase H1 enzyme to RNA:DNA hybrids, combined with the micrococcal nuclease (MNase)-based CUT&RUN technology (Yan et al., 2019a). MapR uses a recombinant mutant form of MNase-fused RNase H1, which binds but does not degrade the RNA moiety within an RNA-DNA hybrid (ΔRH). Upon binding of the RNA-DNA hybrid by ΔRH-MNase, MNase activation by addition Ca^{2+} addition results in cleavage of surrounding DNA and a subsequent release of R-loop associated DNA, which is used for library preparation coupled with high-throughput sequencing (Figure 2A). Surprisingly, we found that coupling NIE-purified photoreceptor nuclei with the standard MapR protocol (Yan & Sarma, 2020a) yielded signal over genic regions resembling MNase-seq rather than R-loop specific enrichment. Our data showed MapR signal depletion around the Transcription Start Site (TSS) of genic regions (Suppl. Figure 2A), suggesting that our samples were being over-digested by MNase. We therefore modified the standard MapR protocol based on the recently published CUT&RUN protocol, which incorporates high calcium-low salt washing steps and decreased digestion time; both conditions have been shown to improve data quality for other applications (Meers, 2019). To validate the quality of our modified protocol, we compared our MapR data in *Drosophila* PRs to the original MapR data obtained in human HEK293T cells (Yan et al., 2019a) (Suppl. Figure 2B) and found that metagene profiles over the gene bodies showed similar R-loop distribution, with signal enrichment around the TSS

and the Transcription Termination Site (TTS) (Suppl. Figure 2A). Furthermore, the MapR signal was not detectable in control samples pre-treated with RNase H1 enzyme as shown by metagene plot and heatmap ranking genes based on MapR signal (Suppl. Figure 2C). These results were corroborated by genome browser analysis showing a significant MapR signal reduction across selected genomic regions (Suppl. Figure 2D), thus validating the specificity and high quality of our modified MapR method. As expected, the R-loop signal was greatly reduced in non-expressed genes compared to expressed genes (Suppl. Figure 2E).

We performed MapR in aging PRs at day 10, 30 and 50 post-eclosion using three independent biological replicates which generated at least $3.5x10^7$ uniquely mapped fragments per sample (Table 1). Spearman's correlation analysis based on read distribution over a 1000-bp binned genome, revealed a strong positive association amongst the biological replicates (Spearman's $p \geq 0.96$). In contrast, when we compared the samples between the age time points, we observed lower positive association between day 10 and day 50 ($p \geq 0.93$), suggesting that the R-loop landscape changed during aging (Figure 2B). This result was also observed by Principal Component Analysis (PCA) of the normalized R-loop distribution, which revealed that 53.8% variance amongst the samples for all biological replicates was attributed to age (Figure 2C). Notably, while samples clustered by age, the similarity between biological replicates decreased with age, suggesting that aging is associated with increased heterogeneity in R-loop distribution.

Next, we performed peak calling using the Model-based Analysis for ChIP-Seq (MACS2) algorithm (Y. Zhang et al., 2008). To account for the differences in the number of called peaks dependent on number of mapped fragments (Suppl. Figure 2F), we down-sampled each bam file to the same number of mapped fragments ($3.5x10^6$) and called the peaks. We found that peak distribution was stably maintained during aging, with approximately 60% of peaks being annotated to promoters (TSS ± 2kb) and approximately 25% of peaks annotated to introns (Figure 2D); this is consistent with previous reports of genome-wide distribution of R-loops

(Aguilera & Gomez-Gonzalez, 2017). Further, evaluation of the Fraction of Reads in Peaks (FRiP) score, which measures the quality of signal enrichment as defined by modENCODE (Landt et al., 2012), revealed consisted FRiP scores (<0.37) for all samples (Figure 2E). Taken together, we successfully applied the MapR method to photoreceptors isolated from the whole organism to producehigh quality R-loop mapping data, showing that the genome-wide R-loop landscape of the PR undergoes age-associated changes.

2.3 Age-associated changes in R-loop levels in PR neurons are highly dynamic

Since R-loops typically form co-transcriptionally, we next examined global and locus-specific distribution of age-associated changes in R-loops across actively transcribed genes, defined as having more than seven transcripts per million (TPM). First, we analyzed global R-loop signal over gene bodies for actively expressed genes and compared the average signal across the gene, as counts per million (CPM). As expected, R-loop signal was enriched mainly over TSS and towards the 3' ends of genes across all age time points (Figure 3A). However, the landscape of R-loop distribution changed with age; while R-loop levels decreased over TSS, they increased over gene bodies and 3' ends of genes in old PR when compared to young PRs.

We next characterized the direction and temporal pattern of the altered R-loop landscape in more detail, thus, obtaining a ratio of the CPM-normalized counts between day 30 and day 10 samples (Early aging) and day 50 and day 30 samples (Late aging) across all expressed genes. Fold change data was then log transformed to center "no-change" signal around 0 (Figure 3B-bottom). Metaplot analysis of R-loop fold change over gene bodies showed that there was a substantial gain in R-loops at TSS across the expressed genes early in aging (Figure 3B-left). However, the dynamics of R-loop distribution changed dramatically late in aging when we observed decreased R-loop signal over TSS with concomitant R-loop gain over promoters and across gene bodies (Figure 3B-right). Thus, our data suggest that R-loop homeostasis is

198 regulated in spatiotemporal manner during aging, rather than in steady and progressive age-
199 dependent changes across the genome.

200 To further evaluate how R-loop genomic coverage changed with age, we asked whether the
201 age-associated increase in R-loops could be a consequence of broadening of the peaks.
202 Supporting this hypothesis, metaplots of CPM-normalized MapR signal aligned to peak center
203 between time points showed that overall, peaks increased width with age (Figure 3C). Next, we
204 quantified global R-loop coverage, as defined by the sum of peak width for each time point.
205 Notably, R-loop peaks covered approximately 18.7 megabases (Mb) of the genome at day 50,
206 compared to 18.1 Mb at day 10, showing a modest but significant increase in coverage during
207 aging (t-test, $p<0.022$) (Figure 3D). Supporting this data, violin plots depicting the peak width for
208 all peaks revealed a slight but consistent increase in peak width at day 30 and day 50 as
209 compared to that at day 10 (Suppl. Figure 3B). Given the 180 Mb total size of the *Drosophila*
210 genome (Adams, 2000), our data showed that in *Drosophila* photoreceptor neurons, R-loops
211 covered approximately 10% of the genome, which is similar to the genomic R-loop coverage
212 obtained from other organisms, including mammals (Manzo et al., 2018; Sanz et al., 2016a;
213 Villarreal et al., 2020).

214 To identify age-regulated R-loop peaks, we quantified MACS2-called peaks and obtained
215 differential R-loop signal (DRS) using a csaw and EdgeR pipeline (Lun & Smyth, 2016;
216 Robinson et al., 2010). DRS analysis revealed that either in early or late aging, the number of
217 peaks that increased or decreased signal, as measured by fold change in CPM, was stably
218 maintained ($p\text{-adj}<0.05$, $|FC|>1.2$), with 13% and 18% of peaks with DRS in early and late
219 aging, respectively (Supp. Figure 3B), suggesting that the landscape of age-regulated R-loops
220 was heterogeneous. To further evaluate the genomic distribution of age-regulated R-loops, we
221 analyzed MapR peak occupancy changes during aging based on their distribution over genic
222 loci, specifically at the Promoter-TSS, 5' untranslated regions (UTR), exons, introns, 3'UTR and

TSS regions (Figure 3E-top). Early in aging, there was a gain in the number of R-loop peaks at 5' UTR, exons and TTS of genic regions. While introns showed a loss in the number of R-loop peaks, there was no difference in the peak number over promoter-TSS and 3' UTR regions early in aging. In contrast, promoter-TSS and 5' UTR regions lost R-loop peaks late in aging with no significant difference in the number of peaks at TTS. Both, exons and introns showed similar trends of R-loop peak gains and losses, respectively. Collectively, R-loop distributions exhibited highly dynamic changes during aging across different regions of actively transcribed genes, suggesting differential regulation of R-loop homeostasis associated with different stages of transcription.

2.4 Age-related R-loop accumulation is associated with specific genomic features and decreased expression of long and highly expressed genes

R-loop formation is typically associated with specific genic characteristics such as gene expression level, torsional stress, and GC content (Allison & Wang, 2019; Chedin & Benham, 2020). To determine whether any of these features were associated with increased changes in R-loops in aging PR neurons, we compared the fold change in R-loop signal around TSS (±3 kb) and TTS (±3 kb) at day 50 relative to day 10, as previously described (see section 2.3). Heatmap plots of photoreceptor-expressed genes ranked based on their fold change in the R-loop signal showed that majority of the TSS-associated R-loops increased with age, while only approximately 50% genes had an increase in R-loop signal around TTS. (Figure 4A). Notably, genes with R-loop gains around the TSS had higher GC content than genes with R-loop losses (Wilcoxon-test, $p<2.2x10^{-5}$), with no statistically significant enrichment for long or highly expressed genes (Figure 4B). However, the genes with age-related R-loop gains around the TTS were enriched for GC rich, long and highly expressed genes (Kruskal-Wallis, $p<4.5x10^{-2}$, $p<4.5x10^{-4}$, and $p<2.1x10^{-9}$, respectively) (Figure 4C). Taken together, these data show that age-related R-loop formation is associated with specific genomic features and enriched at long

and highly expressed genes. Since the photoreceptor-specific genes that decrease expression with age are enriched for long and highly expressed genes (Hall, 2017), we next investigated if age-related transcriptome changes had significant enrichment for genomic features associated with R-loop hotspots. Transcriptome profiling of PRs isolated from flies at day 10 and 50 revealed that 1512 genes were age-regulated (16%), with 351 genes (23%) decreasing expression and 1161 genes (76%) increasing expression (p-adj<0.05, |FC|>1.5) (Figure 4D).

Since persistent formation of R-loops can inhibit transcription (Belotserkovskii et al., 2010; Sanz et al., 2016b), we next asked whether genes that changed expression during aging were enriched for genomic characteristics associated with increased R-loop formation, such as GC skew, gene length, and transcription levels. To do this, we generated Receiver Operating Characteristic (ROC) curves to assess the ability of each individual genomic feature to identify whether a gene would change expression with age. Notably, we found that genes that decreased expression with age were enriched for long and highly expressed genes as shown by area under the curve analysis (AUC=0.62 and 0.80, respectively) (Figure 4E). In contrast, genes that increased expression with age did not show any specific genomic characteristics. Taken together, these observations suggest that R-loop accumulation correlated with decreased transcription of long and highly expressed genes in aging PRs, while there was no correlation between the genes that increased expression and R-loop formation.

2.5 Loss of *Drosophila* Top3β leads to increase in R-loop levels

The neuronal transcriptome is enriched for long and highly expressed genes, that undergo high level of torsional stress during transcription (King et al., 2013; Liu & Wang, 1987). To solve DNA and RNA topological problems, cells use conserved topoisomerase enzymes that play a critical role in a wide range of fundamental metabolic processes in the genome (Pommier et al., 2016; Wang, 2002). Top3β is a highly conserved, dual-activity topoisomerase in animals that can change the topology of both DNA and RNA (Xu et al., 2013). Top3β can unwind negatively

supercoiled DNA that forms during transcription, an activity that prevents formation of R-loops (Chedin & Benham, 2020). Loss of Top3β function is associated with increased R-loop levels in mammalian cells (X. Yang et al., 2000; T. Zhang et al., 2020) and has been shown to reduce lifespan in mice (Kwan & Wang, 2001). In addition, mutations in Top3β are linked to neurological disorders, thus highlighting the critical role of Top3β in neuronal function (Joo et al., 2020). To determine the role of *Drosophila* Top3β in regulation of R-loop homeostasis, we first depleted *Top3β* with ubiquitous RNAi in larvae (*tubP-Gal4>UAS-RNAi*). Using DNA slot blot and RNA-DNA-specific antibody S9.6, we detected a 10% increase in R-loop levels in Top3β-depleted samples as compared to a control expressing non-specific RNAi (Figure 5A). Pre-treatment of DNA samples with RNase H1 led to a complete loss of the signal (Figure 5A), showing the specificity of the signal for RNA-DNA hybrids. qRT-PCR analysis of Top3β transcript levels showed approximately 80% reduction in Top3β-depleted samples as compared to a control, thus validating the efficiency of the knockdown (Figure. 5C). Taken together, these data show the conserved role of Top3β in R-loop homeostasis in *Drosophila*.

Since loss of Top3β in *Drosophila* and mice leads to several neuronal phenotypes, such as disruption of synapse formation and behavioral impairments, we were next interested to see if depletion of Top3β specifically in PR neurons had any impact on visual function. Like most flying insects, *Drosophila* move towards light, thus exhibiting positive phototaxis (Choe & Clandinin, 2005). Importantly, we and others showed that positive phototaxis declines with age in flies (Carbone et al., 2016; Grotewiel et al., 2005; Hall et al., 2017; Simon et al., 2006). To assess visual behavior, we depleted Top3β specifically in photoreceptors using *Rh1-Gal4>UAS-RNAi* and performed phototaxis assays at days 10 and 30 post-eclosion. As expected, there was approximately 15% decrease in positive phototaxis in the control flies between day 10 and day 30 (Wilcoxon test, p-value<0.037) (Figure 5D). Notably, while flies with PR-specific depletion of Top3β showed no difference in the phototactic response at day 10, they showed approximately

60% decrease in visual behavior at day 30 as compared to a control (Wilcoxon test, p-value<0.47, and <0.015, respectively) (Figure 5D). Importantly, this decrease in visual behavior was not due to the loss of PR neurons, as optical neutralization showed no retinal degeneration in Rh1>siTop3β flies at day 30 post-eclosion (Figure 5E). Taken together, these data suggest that Top3β is required for maintenance of proper visual function in aging *Drosophila* photoreceptor neurons.

2.6 Top3β regulates expression of a subset of long genes associated with neuronal function in photoreceptors

Given the role of Top3β in the resolution of torsional stress during transcription, we next hypothesized that Top3β might be required to regulate the expression of genes with neuronal function. To test this hypothesis, we analyzed the transcriptome of PR neurons depleted for Top3β in *Rh1-Gal4>UAS-RNAi; UAS-GFP^{KASH}* flies at day 30 using our NIE protocol. Differential expression analysis using DESeq2 (Love et al., 2014), between Top3β-RNAi and control, revealed that 1% of genes were regulated by Top3β (66 out 6500, p-adj<0.05, |FC|>1.4) (Figure 6A). Additionally, quantitative analysis of gene length based on whether a gene was age-regulated revealed that genes with decreased expression in Top3β-depleted PRs were highly enriched for long genes relative to genes that increased or did not change expression (Wilcoxon test, $p<1.2e^{-3}$ and $<2.4x10^{-9}$, respectively) (Figure 6B). Gene Ontology (GO) enrichment analysis of these Top3β-dependent genes revealed that genes with decreased expression were highly enriched for genes with neuronal functions, as shown by the gene concept network analysis (Cnetplot) (Figure 6C). These genes included *Tenascin major* (*Ten-m; FBgn0004449*) and *Tenascin accessory* (*Ten-a; FBgn0267001*), which form a transmembrane heterodimer involved in synapsis regulation (Mosca et al., 2012; Sun & Xie, 2012), *Tripartite motif containing 9* (*Trim9; FBgn0051721)*, a E3 ubiquitin ligase involved in neurogenesis, axon guidance, and eye development (Akin & Zipursky, 2016; Morikawa et al., 2011), and *knockout* (*ko; FBgn0020294),*

a storkhead-box protein involved in axon guidance (Hartmann et al., 1997). Importantly, GO term analysis of up-regulated genes did not lead to any significant biological category enrichment, highlighting the role of Top3β in maintaining transcript levels of long genes.

Lastly, we sought to evaluate if Top3β-dependent genes were also mis-regulated during natural aging in PR neurons. Gene length analysis of age-regulated genes revealed that genes with decreased expression were longer relative to genes that increased or did not change expression, between day 50 and 10 (Wilcoxon test, $p<2.2x10^{-16}$) (Figure 6D), consistent with our ROC analysis. Strikingly, genes that increased expression during aging were also slightly longer in average than the genes that did not change expression, suggesting that photoreceptor neurons might possess additional mechanisms that positively regulate the expression of a subset of long genes independent of Top3β (Figure 6D). To further assess if Top3β-dependent genes that decreased expression with age were associated with neuronal function, we compared the aging transcriptome (day 50 vs day 10) to Top3β-dependent genes. We identified a significant number of long Top3β-dependent genes (length < 80 kB) with decreased expression in aging PRs that are linked with neuronal function (Figure 6E). These include genes such as *eyeless,* a transcription factor involved in retinal determination gene network *(ey;* FBgn0005558*)* (Graw, 2017)*, fruitless,* a transcription factor involved in neuronal circuits *(fru;* FBgn0004652*)* (Yamamoto et al., 1998)*, Anion exchanger 2,* an anion:anion antiporter *(Ae2;* FBgn0036043*)* (Bosman et al., 1996, 1998)*, and Tenascin major(Ten-M;* FBgn0004449*)* (DePew et al., 2019).

Further, analysis of MapR fold change signal over gene bodies during aging for Top3β-dependent genes revealed that a subset of genes that are transcriptionally regulated by Top3β accumulated R-loops with age (Fig. 6F). In addition, detailed analysis of our recent proteomic studies in the aging *Drosophila* eye (Hall et al., 2021) revealed a statistically significant 20% decrease in Top3β protein levels during aging (Fig. 6G).

Collectively, our data show that Top3β is required to maintain expression of a specific subset of genes with neuronal function that tend to be very long and thus are most likely sensitive to loss of topoisomerase activity due to high levels of torsional stress. This suggests that during aging, proper levels of Top3β are required to maintain R-loop homeostasis and expression of genes important for visual function.

3 DISCUSSION

R-loops have been shown to form in wide range of organisms, including humans (Gomez-Gonzalez & Aguilera, 2019). While R-loops were previously considered to be mere byproducts of transcription, it has been demonstrated that R-loops play a significant physiological role in cellular biology. Notably, there is a growing body of evidence that links R-loop accumulation to transcriptional imbalance and genomic instability, two main hallmarks of aging (López-Otín et al., 2013b). Furthermore, dysregulation of R-loop homeostasis has been linked to human pathologies, including neurodegeneration (Groh & Gromak, 2014). Since age is the main risk factor for many neurodegenerative diseases, our current study focused on characterizing the changes in R-loop landscape induced during aging and evaluating the impact of R-loops on the gene expression, specifically in *Drosophila* photoreceptor neurons.

Characterization of global R-loop levels in aging PRs showed that there was a significant increase in R-loops by middle age, with an additional increase late in aging. To further evaluate the R-loop distribution genome-wide, we optimized a recently published R-loop mapping strategy, called MapR, coupled with high-throughput sequencing. Since this method involves incubation of isolated nuclei/cells with a recombinant mutant form of RNaseH1 tethered with MNase and thus does not require modification of the organismal genome, it is suitable for studies in whole animals. Our data demonstrated that R-loops covered approximately 10% of the *Drosophila* PR genome, which is similar to the reported genomic distribution of R-loops in other organisms (Sanz et al., 2016b; Wahba et al., 2016). Moreover, as expected, R-loops were

enriched over known genomic hot-spots such gene promoters and terminators (Niehrs & Luke, 2020). Our data further revealed that age-associated changes in R-loop occupancy were highly complex and regulated in a spatiotemporal manner, exemplified by differential R-loop distribution across the genome during early and late aging. Specifically, we found that some regions within gene bodies were more prone to changes in R-loop occupancy than others. Even by middle age, we observed significant gains in R-loop levels mainly over TSS of expressed genes. In contrast, the R-loop landscape changed substantially late in aging, when we observed most R-loop gains over the gene bodies and towards the TTS. R-loops are known to play a key physiological role in transcription regulation due to their presence at promoters and terminators, where they regulate transcription initiation and termination, respectively (Niehrs & Luke, 2020). However, if R-loops are not removed efficiently, they can negatively impact transcription (Belotserkovskii et al., 2010; Sanz et al., 2016b). Indeed, our transcriptome profiling reveals that by middle age, flies show decreased expression of genes with neuronal function, inversely correlating with increased levels of R-loops during aging. Furthermore, we found that while exons showed an overall increase in R-loops with age, introns showed the opposite trend. Because aging is associated with altered regulation of splicing (Stegeman et al., 2018) and genes with a role in RNA processing including splicing decrease expression during aging (Hall et al., 2017), imbalance in splicing regulation may result in disproportional gains of R-loops over the exons of transcribed genes.

Detailed analysis of genomic features for genes that accumulated R-loops around either the TSS or the TTS revealed a bias for GC% rich genes. Recent systematic analysis of the aging transcriptome of multiple organisms and cell types showed a positive correlation between transcriptional downregulation and specific genomic features, such as gene length and GC content (Stoeger et al., 2019). Thus, R-loop accumulation over specific genomic regions and

their contribution to the regulation of the aging transcriptome might be conserved in other organisms.

One of the major driving forces of R-loop accumulation is topological stress that arises from under- and over-winding of the double helix during transcription and replication (McKinnon, 2016). Notably, long genes accumulate high topological stress during transcription and loss of topoisomerase activity has been shown to preferentially inhibit expression of long genes (Joshi et al., 2012; Keszthelyi et al., 2016). Since age-associated R-loop gains were particularly localized at long and highly expressed genes, we sought to further explore the impact of decreased topoisomerase activity on the maintenance of photoreceptor neuron homeostasis. We focused on the DNA/RNA topoisomerase Top3β because loss of Top3β function is associated with increased R-loop levels in mammalian cells (Y. Yang et al., 2014; T. Zhang et al., 2020). In addition, mutations in Top3β are linked to neurological disorders, thus highlighting the critical role of Top3β in neuronal function (Joo et al., 2020). Notably, we demonstrated that normal Top3β levels were required for maintenance of neuronal function, as shown by an age-associated decrease in visual behavior upon photoreceptor-specific downregulation of Top3β. In addition, depletion of Top3β in PRs leads to decreased expression of a subset of long genes with neuronal function, some of which also gained R-loop levels with age. Because proteomic analysis of *Drosophila* eyes shows decreased protein levels of Top3β with age, our data suggests that Top3β may function in maintenance of R-loop homeostasis and gene expression required for proper neuronal function.

Aging is characterized by a progressive and time dependent deterioration of an organism that eventually leads to loss of cellular function and increased cell death (Partridge & Mangel, 1999). The hallmarks of aging include alteration in gene expression, changes to the epigenome, loss of protein homeostasis, and increased cellular senescence (López-Otín et al., 2013b). Our current studies demonstrate a novel finding: *Drosophila* photoreceptor neurons accumulate R-loops

during aging, and these age-associated changes in R-loop occupancy are highly dynamic and regulated in a spatiotemporal manner. Specifically, gene bodies show significant gains in R-loops late in aging, and long, highly-expressed genes are especially prone to R-loop accumulation. Furthermore, the results of this study highlight the role of Top3β in maintenance of R-loop homeostasis and proper expression of a specific subset of long genes with neuronal function during aging.

4 EXPERIMENTAL PROCEDURES

Fly maintenance

Flies were raised in 12:12 h light:dark cycle at 25°C on standard fly food (Lewis, 1960). For aging experiments, flies were collected over three days post-eclosion and maintained in population cages with a density of ~1000 flies/cage with fresh fly food change every two days. Male flies were aged to the specified age, collected and flash-frozen in liquid nitrogen at Zeitgeber time 6 (-/+ 1 hour).

Nuclei Immuno-Enrichment (NIE) protocol

NIE was performed as described previously (Jauregui-Lozano, 2021). Detailed protocol can be found at dx.doi.org/10.17504/protocols.io.buiqnudw

Slot blot

DNA was extracted using Quick DNA MiniPrep Plus Kit (Zymo Research, Catalog #D4068) following the instructions for tissue homogenization. DNA concentration was measured using Qubit dsDNA settings. 35-50 ng of DNA was loaded on a Hybod-N^+ membrane (GE Healthcare, Catalog #RPN119B) using slot blot apparatus. After drying, the membrane was blocked using 5% milk in PBST, then incubated with anti-S9.6 antibody in PBST (Millipore Sigma, Catalog # MABE1095), followed by goat anti-mouse HRP antibody (BioRad, Catalog #*170-6516*). For loading control, membranes were denatured in 0.5 M NaOH, 1.5 M NaCl for 10 min at RT, followed by 10 min-wash in PBST and neutralized in 1.5 M NaCl, 0.5 M TrisCl for 10 min at RT. The

membrane was then washed with PBST for 10 min, blocked in 5% milk PBST for 30 min, and incubated with mouse anti single-stranded DNA antibody (MilliporeSigma, Catalog #MAB3034), followed by goat-anti-mouse HRP antibody (BioRad, Catalog#*170-6516*). To test the S9.6 antibody specificity, half of the DNA samples were treated with RNase H enzyme (10U) at 37°C for at least 3 HR or over/night.

GST-ΔRh-MNase and GST-MNase protein purification

Protein overexpression and purification were performed as described previously (Yan & Sarma, 2020a). *Protein overexpression:* BL21 (DE3) competent *E. coli* cells (NEB, Ipswich, MA, Catalog #C2527H) were transformed with 10 ng of either pGEX-6p-1-GST-MNase or pGEX-6p-1-GST-ΔRNH-MNase plasmid (Addgene, Watertown, MA, Catalog #136291 and 136292, respectively). Transformed bacteria were grown in 500 mL of standard LB media (*LB (Luria-Bertani) Liquid Medium*, n.d.) supplemented with 100 µM Carbenicillin at 37°C. Once optical density reached 0.5 at 600 nm, protein expression was induced with 1 mM IPTG and cultures were grown at 37°C for additional 3 hours under constant rotation. Cells were pelleted using at 4ºC, 8000 RPM for 10 minutes. *Protein purification:* Bacterial pellets were resuspended in ice-cold 1X PBS buffer (ThermoFisher, Waltham, MA, Catalog #70011-044) and sonicated using a Branson Digital Sonifier in five 15sec-45sec ON/OFF cycles. Pierce™ Glutathione Magnetic Agarose Beads (Thermo Fisher, Catalog #78601) were used to purify GST-tagged MNase and RNAseH-MNase recombinant proteins according to manufacturer's instructions.

Improved MapR

To profile genome-wide distribution of R-loops, we followed the MapR protocol (Yan et al., 2019) with some modifications according to the improved CUT&RUN protocol to decrease background MNase cleavage and digestion (Meers et al., 2019). Briefly, isolated nuclei were washed with 1 mL of Digitonin-containing wash buffer (20 mM HEPES-NaOH, 150 mM NaCl, 0.5 mM Spermidine, 0.02% Digitonin) freshly supplemented with EDTA-free cOmplete protease

inhibitors (Sigma-Aldrich, St. Louis, MO, Catalog #11873580001). Nuclei were resuspended in 150 µl of Digitonin-containing wash buffer and 1 µM GST-ΔRh-MNAse or GST-MNase was added to a final concentration followed by one-hour incubation at 4°C with constant rotation. Nuclei were washed three times with 500 uL Digitonin-containing wash buffer, then washed one time with 1 mL of Low-Salt Rinse Buffer (20 mM HEPES, pH7.5, 0.5 mM spermidine, 0.05% Digitonin) freshly supplemented with EDTA-free complete protease inhibitor. Nuclei were resuspended in 200 µL of ice-cold calcium containing Incubation buffer (3.5 mM HEPES pH 7.5, 10 mM CaCl2, 0.05% Digitonin) and placed on wet ice for 60 seconds. Upon removal of supernatant, nuclei were resuspended in 150 µL of EGTA-STOP buffer (170 mM NaCl, 20 mM EGTA, 0.05% Digitonin, 20 µg/ml glycogen, 25 µg/ml RNase A), followed by 30-minute incubation at 37°C. DNA was extracted using Quick-DNA Microprep Kit (Zymo Research, Irvine, CA, Catalog #D4074). DNA was quantified with Qubit 1X HS DNA (ThermoScientific, Catalog#Q33203) and 2 ng of purified DNA was used to make sequencing libraries with Tecan Ovation Ultralow V2 DNA-Seq Library Preparation Kit-Unique Dual Indexes (Tecan, Switzerland, Catalog #9149-A01). Up to 16 libraries were pooled in one lane for paired-end 150 bp Illumina HiSeq sequencing.

RNA-seq

Isolated nuclei were resuspended in 100 µL TRI reagent (Zymo Research, Irvine CA, Catalog #R2050-1-200) and incubated at RT for 1 hour, followed by RNA extraction using the Direct-zol™ RNA Microprep (Zymo Research, Catalog #R2061). Purified RNA was quantified with the Qubit™ RNA HS Assay Kit (ThermoFisher, Catalog #Q32852) as per the manufacturers' instructions**.** cDNA libraries were prepared with 10 ng of nuclear RNA using Ovation SoLo RNA-seq System including *Drosophila*-specific anyDeplete technology for rRNA depletion (Tecan, Redwood City, CA, Catalog #0502-32). Up to 16 libraries were pooled in one lane for paired-end 150 bp Illumina HiSeq sequencing.

Quantitative PCR (qRT-PCR)

cDNA was synthesized using 300 ng of RNA using EpiScript RNase H- Reverse Transcriptase (Lucigen, Middleton, WI, Catalog #ERT12910K). qRT-PCR was performed using Bullseye EvaGreen qPCR 2X master mix-ROX (Midsci, Valley Park, MO, Catalog #BEQPCR-R) and using the following primers for *eIF-1a* forward 5'-GCTGGGCAACGGTCGTCTGGAGGC-3' and reverse 5'-CGTCTTCAGGTTCCTGGCCTCGTCCGG-3'; and for *Top3β* forward 5'-GAATGGGCGCGCGGTCGGGTC-3' and reverse 5'-CGCATCAGTTCGACGGTGTTCAGTGCC-3'.

Optic Neutralization

Flies were anesthetized on a CO_2 pad and glued to a glass slide. Live rhabdomeres were imaged using brightfield light microscopy in an Olympus BX51 microscope, as described previously (Stegeman et al., 2018).

Bioinformatic analysis

Raw reads were trimmed using Trimmomatic version 0.39 (Bolger et al., 2014) to filter out low quality reads (Q>30) and clean adapter reads. Cleaned reads were aligned to the *Drosophila melanogaster* genome (BDGP Release 6 + ISO1 MT/dm6 from UCSC) using splicing-aware aligner STAR version 1.3 (Dobin et al., 2013) for RNA-seq and Bowtie2 version 2.3.5.1 (Langmead & Salzberg, 2012) for MapR using –sensitive settings. Samtools version 1.8 (Li et al., 2009) was used to obtain, sort and index BAM files. For genome browser inspection as well as further analyses, bigwig files were generated by normalizing datasets to count-per-million CPM coverage tracks with *deepTools* version 3.1.1 (Ramírez et al., 2014) using --*normalizeUsing CPM* settings. Spearman's correlation scores were calculated using subpackages *multiBigwigSummary* and *plotCorrelation* as part of deepTools. Metaplots and genomic distribution heatmaps were generated with *computeMatrix, plotHeatmap* and *plotProfile* deepTool subpackages. MapR narrow and broad peaks were obtained using MACS2 version 2.1.2 (Y. Zhang et al., 2008) with standard settings. Peak overlap and genomic distribution of peaks was determined using R package ChIPseeker (v1.26.2) (Yu et al., 2015)

and Homer (v4.11) Quantitative peak analysis was performed using GenomicRanges (v1.42.0) (Lawrence et al., 2013), csaw (v1.24.3) (Lun & Smyth, 2016) and edgeR (v3.13) (Robinson et al., 2010). R analysis was run in RStudio (v1.4.1106). Differential gene expression (DGE) analysis was performed using DESeq2 (Love et al., 2014). To increase the stringency of analysis, we generated shrunk fold change estimates within DESeq2 using the lfcShrink function (Zhu et al., 2019).

Phototaxis

Phototaxis assay was performed as previously described (Stegeman et al., 2018). Briefly, 30 male flies were placed in a vial, tapped down and placed into the opening of the Maze apparatus. After letting them adjust to the darkness for 10 min, flies were given 30 seconds to choose between light and dark vial. Light preference index was then calculated as proportion of flies that chose light out of total number of flies.

Graph plots

Bar-plots were generated using the R package ggplot2 (v3.3.3), and scripts used for RNA-seq analysis and plot generation are available upon request.

DATA AVAILABILITY

RNA-seq expression data and MapR mapping data are accessible through Gene Expression Omnibus repository under series accession numbers GSE174488, GSE174491 and GSE174515, respectively.

ACKNOWLEDGMENTS

Fly stocks from the Bloomington Drosophila Stock Center and information from FlyBase were used in this study.

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

AUTHOR CONTRIBUTIONS

J.J-L. performed all the NGS experiments and the bioinformatics analysis. N.A.L. and V.M.W. assisted with data analysis. A.N.E. and H.H. performed the slot blots and Western blots. S. E. performed the phototaxis assay. J.J-L. and H.H. wrote the manuscript with input from other authors. H.H. supervised the project and conceived the study.

ORCID

Juan Jauregui-Lozano https://orcid.org/0000-0001-6633-9152
Alyssa N Easton https://orcid.org/0000-0002-0972-3630
Spencer E Escobedo https://orcid.org/0000-0002-9026-9104
Nadia A Lanman https://orcid.org/0000-0002-1819-7070
Vikki M. Weake https://orcid.org/0000-0002-5933-9952
Hana Hall https://orcid.org/0000-0002-7544-4437

HOW TO CITE THIS ARTICLE

Jauregui-Lozano J, Easton AN, Escobedo SE, Lanman NA, Weake VM, Hall H. Proper control of R-loop homeostasis is required for maintenance of gene expression and neuronal function during aging

REFERENCES

Aguilera, A., & García-Muse, T. (2012). R Loops: From Transcription Byproducts to Threats to Genome Stability. *Molecular Cell*, *46*(2), 115–124. https://doi.org/10.1016/j.molcel.2012.04.009

Aguilera, A., & Gomez-Gonzalez, B. (2017). DNA-RNA hybrids: The risks of DNA breakage during transcription. *Nat Struct Mol Biol*, *24*(5), 439–443. https://doi.org/10.1038/nsmb.3395

573 Akin, O., & Zipursky, S. L. (2016). Frazzled promotes growth cone attachment at the source of a
574 Netrin gradient in the Drosophila visual system. *ELife*, *5*, e20762.
575 https://doi.org/10.7554/eLife.20762
576 Allison, D. F., & Wang, G. G. (2019). R-loops: Formation, function, and relevance to cell stress.
577 *Cell Stress*, *3*(2), 38–46. https://doi.org/10.15698/cst2019.02.175
578 Belotserkovskii, B. P., Liu, R., Tornaletti, S., Krasilnikova, M. M., Mirkin, S. M., & Hanawalt, P.
579 C. (2010). Mechanisms and implications of transcription blockage by guanine-rich DNA
580 sequences. *Proceedings of the National Academy of Sciences*, *107*(29), 12816–12821.
581 https://doi.org/10.1073/pnas.1007580107
582 Boguslawski, S. J., Smith, D. E., Michalak, M. A., Mickelson, K. E., Yehle, C. O., Patterson, W.
583 L., & Carrico, R. J. (1986). Characterization of monoclonal antibody to DNA · RNA and
584 its application to immunodetection of hybrids. *Journal of Immunological Methods*, *89*(1),
585 123–130. https://doi.org/10.1016/0022-1759(86)90040-2
586 Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina
587 sequence data. *Bioinformatics*, *30*(15), 2114–2120.
588 https://doi.org/10.1093/bioinformatics/btu170
589 Bonnel, S., Mohand-Said, S., & Sahel, J.-A. (2003). The aging of the retina. *Experimental*
590 *Gerontology*, *38*(8), 825–831. https://doi.org/10.1016/S0531-5565(03)00093-7
591 Boque-Sastre, R., Soler, M., Oliveira-Mateos, C., Portela, A., Moutinho, C., Sayols, S.,
592 Villanueva, A., Esteller, M., & Guil, S. (2015). Head-to-head antisense transcription and
593 R-loop formation promotes transcriptional activation. *Proceedings of the National*
594 *Academy of Sciences*, *112*(18), 5785–5790.
595 Bosman, G. J., Engbersen, A., Vollaard, C. H., Bartholomeus, I. G., Pistorius, A. M., Renkawek,
596 K., & De Grip, W. J. (1996). Implications of aging- and degeneration-related changes in
597 anion exchange proteins for the maintenance of neuronal homeostasis. *Cellular and*
598 *Molecular Biology (Noisy-Le-Grand, France)*, *42*(7), 905–918.

599 Bosman, G. J., Renkawek, K., Van Workum, F. P., Bartholomeus, I. G., Marini, S., & De Grip,
600 W. J. (1998). Neuronal anion exchange proteins in Alzheimer's disease pathology.
601 *Journal of Neural Transmission. Supplementum*, *54*, 248–257.
602 Carbone, M. A., Yamamoto, A., Huang, W., Lyman, R. A., Meadors, T. B., Yamamoto, R.,
603 Anholt, R. R. H., & Mackay, T. F. C. (2016). Genetic architecture of natural variation in
604 visual senescence in Drosophila. *Proceedings of the National Academy of Sciences*,
605 *113*(43), E6620–E6629. https://doi.org/10.1073/pnas.1613833113
606 Carthew, R. W. (2007). Pattern formation in the Drosophila eye. *Current Opinion in Genetics &*
607 *Development*, *17*(4), 309–313. https://doi.org/10.1016/j.gde.2007.05.001
608 Chan, Y. A., Aristizabal, M. J., Lu, P. Y. T., Luo, Z., Hamza, A., Kobor, M. S., Stirling, P. C., &
609 Hieter, P. (2014). Genome-Wide Profiling of Yeast DNA:RNA Hybrid Prone Sites with
610 DRIP-Chip. *PLOS Genetics*, *10*(4), e1004288.
611 https://doi.org/10.1371/journal.pgen.1004288
612 Chedin, F., & Benham, C. J. (2020). Emerging roles for R-loop structures in the management of
613 topological stress. *The Journal of Biological Chemistry*, *295*(14), 4684–4695.
614 https://doi.org/10.1074/jbc.REV119.006364
615 Choe, K., & Clandinin, T. R. (2005). Thinking about Visual Behavior; Learning about
616 Photoreceptor Function. In *Current Topics in Developmental Biology* (Vol. 69, pp. 187–
617 213). Academic Press. https://doi.org/10.1016/S0070-2153(05)69007-2
618 Curcio, C. A. (2001). Photoreceptor topography in ageing and age-related maculopathy. *Eye*,
619 *15*(3), 376–383. https://doi.org/10.1038/eye.2001.140
620 DePew, A. T., Aimino, M. A., & Mosca, T. J. (2019). The Tenets of Teneurin: Conserved
621 Mechanisms Regulate Diverse Developmental Processes in the Drosophila Nervous
622 System. *Frontiers in Neuroscience*, *13*. https://doi.org/10.3389/fnins.2019.00027

623 Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson,
624 M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*,
625 *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635
626 Drolet, M. (2006). Growth inhibition mediated by excess negative supercoiling: The interplay
627 between transcription elongation, R-loop formation and DNA topology. *Molecular*
628 *Microbiology*, *59*(3), 723–730. https://doi.org/10.1111/j.1365-2958.2005.05006.x
629 Drolet, M., Bi, X., & Liu, L. F. (1994). Hypernegative supercoiling of the DNA template during
630 transcription elongation in vitro. *J Biol Chem*, *269*(3), 2068–2074.
631 Drolet, M., Phoenix, P., Menzel, R., Massé, E., Liu, L. F., & Crouch, R. J. (1995).
632 Overexpression of RNase H partially complements the growth defect of an Escherichia
633 coli delta topA mutant: R-loop formation is a major problem in the absence of DNA
634 topoisomerase I. *Proceedings of the National Academy of Sciences*, *92*(8), 3526–3530.
635 https://doi.org/10.1073/pnas.92.8.3526
636 Fernandopulle, M. S., Lippincott-Schwartz, J., & Ward, M. E. (2021). RNA transport and local
637 translation in neurodevelopmental and neurodegenerative disease. *Nature*
638 *Neuroscience*, 1–11. https://doi.org/10.1038/s41593-020-00785-2
639 Gao, H., & Hollyfield, J. G. (1992). Aging of the human retina. Differential loss of neurons and
640 retinal pigment epithelial cells. *Investigative Ophthalmology & Visual Science*, *33*(1), 1–
641 17.
642 Ginno, P. A., Lott, P. L., Christensen, H. C., Korf, I., & Chédin, F. (2012). R-Loop Formation Is a
643 Distinctive Characteristic of Unmethylated Human CpG Island Promoters. *Molecular*
644 *Cell*, *45*(6), 814–825. https://doi.org/10.1016/j.molcel.2012.01.017
645 Gomez-Gonzalez, B., & Aguilera, A. (2019). Transcription-mediated replication hindrance: A
646 major driver of genome instability. *Genes Dev*. https://doi.org/10.1101/gad.324517.119
647 Gómez-González, B., García-Rubio, M., Bermejo, R., Gaillard, H., Shirahige, K., Marín, A.,
648 Foiani, M., & Aguilera, A. (2011). Genome-wide function of THO/TREX in active genes

649 prevents R-loop-dependent replication obstacles. *The EMBO Journal*, *30*(15), 3106–
650 3119. https://doi.org/10.1038/emboj.2011.206
651 Graw, J. (2017). From eyeless to neurological diseases. *Experimental Eye Research*, *156*, 5–9.
652 https://doi.org/10.1016/j.exer.2015.11.006
653 Groh, M., & Gromak, N. (2014). Out of Balance: R-loops in Human Disease. *PLOS Genetics*,
654 *10*(9), e1004630. https://doi.org/10.1371/journal.pgen.1004630
655 Grotewiel, M. S., Martin, I., Bhandari, P., & Cook-Wiens, E. (2005). Functional senescence in
656 Drosophila melanogaster. *Ageing Research Reviews*, *4*(3), 372–397.
657 https://doi.org/10.1016/j.arr.2005.04.001
658 Hall, H., Cooper, B. R., Qi, G., Wijeratne, A. B., Mosley, A. L., & Weake, V. M. (2021).
659 Quantitative proteomic and metabolomic profiling reveals altered mitochondrial
660 metabolism and folate biosynthesis pathways in the aging Drosophila eye. *BioRxiv*,
661 2021.06.02.446766. https://doi.org/10.1101/2021.06.02.446766
662 Hall, H., Medina, P., Cooper, D. A., Escobedo, S. E., Rounds, J., Brennan, K. J., Vincent, C.,
663 Miura, P., Doerge, R., & Weake, V. M. (2017). Transcriptome profiling of aging
664 Drosophila photoreceptors reveals gene expression trends that correlate with visual
665 senescence. *BMC Genomics*, *18*(1), 894. https://doi.org/10.1186/s12864-017-4304-3
666 Hardie, R. C. (1985). Functional Organization of the Fly Retina. In H. Autrum, D. Ottoson, E. R.
667 Perl, R. F. Schmidt, H. Shimazu, & W. D. Willis (Eds.), *Progress in Sensory Physiology*
668 (pp. 1–79). Springer. https://doi.org/10.1007/978-3-642-70408-6_1
669 Hartmann, C., Landgraf, M., Bate, M., & Jäckle, H. (1997). Krüppel target gene knockout
670 participates in the proper innervation of a specific set of Drosophila larval muscles. *The*
671 *EMBO Journal*, *16*(17), 5299–5309. https://doi.org/10.1093/emboj/16.17.5299
672 Jackson, G. R., Owsley, C., & Curcio, C. A. (2002). Photoreceptor degeneration and dysfunction
673 in aging and age-related maculopathy. *Ageing Research Reviews*, *1*(3), 381–396.
674 https://doi.org/10.1016/S1568-1637(02)00007-7

675 Jauregui-Lozano, J., Bakhle, K., & Weake, V. M. (2021a). In vivo tissue-specific chromatin
676 profiling in Drosophila melanogaster using GFP-tagged nuclei. *Genetics*, *iyab079*.
677 https://doi.org/10.1093/genetics/iyab079
678 Jauregui-Lozano, J., Bakhle, K., & Weake, V. M. (2021b). In vivo tissue-specific chromatin
679 profiling in Drosophila melanogaster using GFP-tagged nuclei. *Genetics*, *iyab079*.
680 https://doi.org/10.1093/genetics/iyab079
681 Joo, Y., Xue, Y., Wang, Y., McDevitt, R. A., Sah, N., Bossi, S., Su, S., Lee, S. K., Peng, W., Xie,
682 A., Zhang, Y., Ding, Y., Ku, W. L., Ghosh, S., Fishbein, K., Shen, W., Spencer, R.,
683 Becker, K., Zhao, K., … Wang, W. (2020). Topoisomerase 3β knockout mice show
684 transcriptional and behavioural impairments associated with neurogenesis and synaptic
685 plasticity. *Nature Communications*, *11*(1), 3143. https://doi.org/10.1038/s41467-020-
686 16884-4
687 Joshi, R. S., Piña, B., & Roca, J. (2012). Topoisomerase II is required for the production of long
688 Pol II gene transcripts in yeast. *Nucleic Acids Research*, *40*(16), 7907–7915.
689 https://doi.org/10.1093/nar/gks626
690 Keszthelyi, A., Minchell, N. E., & Baxter, J. (2016). The Causes and Consequences of
691 Topological Stress during DNA Replication. *Genes*, *7*(12).
692 https://doi.org/10.3390/genes7120134
693 King, I. F., Yandava, C. N., Mabb, A. M., Hsiao, J. S., Huang, H.-S., Pearson, B. L., Calabrese,
694 J. M., Starmer, J., Parker, J. S., Magnuson, T., Chamberlain, S. J., Philpot, B. D., &
695 Zylka, M. J. (2013). Topoisomerases facilitate transcription of long genes linked to
696 autism. *Nature*, *501*, 58. https://doi.org/10.1038/nature12504
697 Kwan, K. Y., & Wang, J. C. (2001). Mice lacking DNA topoisomerase IIIβ develop to maturity but
698 show a reduced mean lifespan. *Proceedings of the National Academy of Sciences*,
699 *98*(10), 5717–5721. https://doi.org/10.1073/pnas.101132498

Landt, S. G., Marinov, G. K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S., Bernstein, B. E., Bickel, P., Brown, J. B., Cayting, P., Chen, Y., DeSalvo, G., Epstein, C., Fisher-Aylor, K. I., Euskirchen, G., Gerstein, M., Gertz, J., Hartemink, A. J., Hoffman, M. M., … Snyder, M. (2012). ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Research*, *22*(9), 1813–1831. https://doi.org/10.1101/gr.136184.111

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M. T., & Carey, V. J. (2013). Software for Computing and Annotating Genomic Ranges. *PLOS Computational Biology*, *9*(8), e1003118. https://doi.org/10.1371/journal.pcbi.1003118

LB (Luria-Bertani) liquid medium. (n.d.). Retrieved March 16, 2021, from http://cshprotocols.cshlp.org/content/2006/1/pdb.rec8141.full?text_only=

Lewis, E. (1960). A new standard food medium. *Drosophila Information Service*, *34*(117), 1–55.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liu, L. F., & Wang, J. C. (1987). Supercoiling of the DNA template during transcription. *Proceedings of the National Academy of Sciences*, *84*(20), 7024–7027.

López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M., & Kroemer, G. (2013a). The Hallmarks of Aging. *Cell*, *153*(6), 1194–1217. https://doi.org/10.1016/j.cell.2013.05.039

López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M., & Kroemer, G. (2013b). The Hallmarks of Aging. *Cell*, *153*(6), 1194–1217. https://doi.org/10.1016/j.cell.2013.05.039

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

726 Lun, A. T. L., & Smyth, G. K. (2016). csaw: A Bioconductor package for differential binding
727 analysis of ChIP-seq data using sliding windows. *Nucleic Acids Research*, *44*(5), e45.
728 https://doi.org/10.1093/nar/gkv1191
729 Ma, J., & Weake, V. M. (2014). Affinity-based Isolation of Tagged Nuclei from Drosophila
730 Tissues for Gene Expression Analysis. *JoVE (Journal of Visualized Experiments)*, *85*,
731 e51418. https://doi.org/10.3791/51418
732 Mackay, R. P., Xu, Q., & Weinberger, P. M. (2020). R-Loop Physiology and Pathology: A Brief
733 Review. *DNA and Cell Biology*, *39*(11), 1914–1925.
734 https://doi.org/10.1089/dna.2020.5906
735 Manzo, S. G., Hartono, S. R., Sanz, L. A., Marinello, J., De Biasi, S., Cossarizza, A., Capranico,
736 G., & Chedin, F. (2018). DNA Topoisomerase I differentially modulates R-loops across
737 the human genome. *Genome Biology*, *19*(1), 100. https://doi.org/10.1186/s13059-018-
738 1478-1
739 McKinnon, P. J. (2016). Topoisomerases and the regulation of neural function. *Nature Reviews.*
740 *Neuroscience*, *17*(11), 673–679. https://doi.org/10.1038/nrn.2016.101
741 Meers, M. P., Bryson, T. D., Henikoff, J. G., & Henikoff, S. (2019). Improved CUT&RUN
742 chromatin profiling tools. *ELife*, *8*, e46314. https://doi.org/10.7554/eLife.46314
743 Morikawa, R. K., Kanamori, T., Yasunaga, K., & Emoto, K. (2011). Different levels of the
744 Tripartite motif protein, Anomalies in sensory axon patterning (Asap), regulate distinct
745 axonal projections of Drosophila sensory neurons. *Proceedings of the National Academy*
746 *of Sciences*, *108*(48), 19389–19394.
747 Mosca, T. J., Hong, W., Dani, V. S., Favaloro, V., & Luo, L. (2012). Trans-synaptic Teneurin
748 signalling in neuromuscular synapse organization and target choice. *Nature*, *484*(7393),
749 237–241. https://doi.org/10.1038/nature10923

Niehrs, C., & Luke, B. (2020). Regulatory R-loops as facilitators of gene expression and genome stability. *Nature Reviews Molecular Cell Biology*, 1–12. https://doi.org/10.1038/s41580-019-0206-3

Panda-Jonas, S., Jonas, J. B., & Jakobczyk-Zmija, M. (1995). Retinal Photoreceptor Density Decreases with Age. *Ophthalmology*, *102*(12), 1853–1859. https://doi.org/10.1016/S0161-6420(95)30784-1

Parapuram, S. K., Cojocaru, R. I., Chang, J. R., Khanna, R., Brooks, M., Othman, M., Zareparsi, S., Khan, N. W., Gotoh, N., Cogliati, T., & Swaroop, A. (2010). Distinct signature of altered homeostasis in aging rod photoreceptors: Implications for retinal diseases. *PloS One*, *5*(11), e13885. https://doi.org/10.1371/journal.pone.0013885

Partridge, L., & Mangel, M. (1999). Messages from mortality: The evolution of death rates in the old. *Trends in Ecology & Evolution*, *14*(11), 438–442. https://doi.org/10.1016/S0169-5347(99)01646-8

Pommier, Y., Sun, Y., Huang, S. N., & Nitiss, J. L. (2016). Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nature Reviews Molecular Cell Biology*, *17*, 703. https://doi.org/10.1038/nrm.2016.111

Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A., & Manke, T. (2014). deepTools: A flexible platform for exploring deep-sequencing data. *Nucleic Acids Research*, *42*(Web Server issue), W187–W191. https://doi.org/10.1093/nar/gku365

Rauser, C. L., Abdel-Aal, Y., Shieh, J. A., Suen, C. W., Mueller, L. D., & Rose, M. R. (2005). Lifelong heterogeneity in fecundity is insufficient to explain late-life fecundity plateaus in Drosophila melanogaster. *Experimental Gerontology*, *40*(8), 660–670. https://doi.org/10.1016/j.exger.2005.06.006

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

776 Sanz, L. A., Hartono, S. R., Lim, Y. W., Steyaert, S., Rajpurkar, A., Ginno, P. A., Xu, X., &
777 Chédin, F. (2016a). Prevalent, Dynamic, and Conserved R-Loop Structures Associate
778 with Specific Epigenomic Signatures in Mammals. *Molecular Cell*, *63*(1), 167–178.
779 https://doi.org/10.1016/j.molcel.2016.05.032
780 Sanz, L. A., Hartono, S. R., Lim, Y. W., Steyaert, S., Rajpurkar, A., Ginno, P. A., Xu, X., &
781 Chédin, F. (2016b). Prevalent, Dynamic, and Conserved R-Loop Structures Associate
782 with Specific Epigenomic Signatures in Mammals. *Molecular Cell*, *63*(1), 167–178.
783 https://doi.org/10.1016/j.molcel.2016.05.032
784 Simon, A. F., Liang, D. T., & Krantz, D. E. (2006). Differential decline in behavioral performance
785 of Drosophila melanogaster with age. *Mechanisms of Ageing and Development*, *127*(7),
786 647–651. https://doi.org/10.1016/j.mad.2006.02.006
787 Skourti-Stathaki, K., Kamieniarz-Gdula, K., & Proudfoot, N. J. (2014). R-loops induce repressive
788 chromatin marks over mammalian gene terminators. *Nature*, *516*(7531), 436–439.
789 https://doi.org/10.1038/nature13787
790 Skourti-Stathaki, K., Proudfoot, N. J., & Gromak, N. (2011). Human Senataxin Resolves
791 RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent
792 Termination. *Molecular Cell*, *42*(6), 794–805.
793 https://doi.org/10.1016/j.molcel.2011.04.026
794 Stegeman, R., Hall, H., Escobedo, S. E., Chang, H. C., & Weake, V. M. (2018). Proper splicing
795 contributes to visual function in the aging Drosophila eye. *Aging Cell*, *17*(5), e12817.
796 https://doi.org/10.1111/acel.12817
797 Stoeger, T., Grant, R. A., McQuattie-Pimentel, A. C., Anekalla, K., Liu, S. S., Tejedor-Navarro,
798 H., Singer, B. D., Abdala-Valencia, H., Schwake, M., Tetreault, M.-P., Perlman, H.,
799 Balch, W. E., Chandel, N., Ridge, K., Sznajder, J. I., Morimoto, R. I., Misharin, A. V.,
800 Budinger, G. R. S., & Amaral, L. A. N. (2019). Aging is associated with a systemic
801 length-driven transcriptome imbalance. *BioRxiv*, 691154. https://doi.org/10.1101/691154

802 Sun, M., & Xie, W. (2012). Cell adhesion molecules in Drosophila synapse development and
803 function. *Science China Life Sciences*, *55*(1), 20–26. https://doi.org/10.1007/s11427-
804 012-4273-3
805 Villarreal, O. D., Mersaoui, S. Y., Yu, Z., Masson, J.-Y., & Richard, S. (2020). Genome-wide R-
806 loop analysis defines unique roles for DDX5, XRN2, and PRMT5 in DNA/RNA hybrid
807 resolution. *Life Science Alliance*, *3*(10). https://doi.org/10.26508/lsa.202000762
808 Wahba, L., Costantino, L., Tan, F. J., Zimmer, A., & Koshland, D. (2016). S1-DRIP-seq
809 identifies high expression and polyA tracts as major contributors to R-loop formation.
810 *Genes & Development*, *30*(11), 1327–1338. https://doi.org/10.1101/gad.280834.116
811 Wang, J. C. (2002). Cellular roles of DNA topoisomerases: A molecular perspective. *Nature*
812 *Reviews Molecular Cell Biology*, *3*(6), 430–440. https://doi.org/10.1038/nrm831
813 Weale, R. (1998). The Eye within the Framework of Human Senescence: Biological Decline and
814 Morbidity. *Ophthalmic Research*, *30*(2), 59–73. https://doi.org/10.1159/000055456
815 Xu, D., Shen, W., Guo, R., Xue, Y., Peng, W., Sima, J., Yang, J., Sharov, A., Srikantan, S.,
816 Yang, J., Fox, D., Qian, Y., Martindale, J. L., Piao, Y., Machamer, J., Joshi, S. R.,
817 Mohanty, S., Shaw, A. C., Lloyd, T. E., … Wang, W. (2013). Top3β is an RNA
818 topoisomerase that works with Fragile X syndrome protein to promote synapse
819 formation. *Nature Neuroscience*, *16*(9), 1238–1247. https://doi.org/10.1038/nn.3479
820 Yamamoto, D., Fujitani, K., Usui, K., Ito, H., & Nakano, Y. (1998). From behavior to
821 development: Genes for sexual behavior define the neuronal sexual switch in
822 Drosophila. *Mechanisms of Development*, *73*(2), 135–146.
823 https://doi.org/10.1016/S0925-4773(98)00042-2
824 Yan, Q., & Sarma, K. (2020a). MapR: A Method for Identifying Native R-Loops Genome Wide.
825 *Current Protocols in Molecular Biology*, *130*(1), e113. https://doi.org/10.1002/cpmb.113
826 Yan, Q., & Sarma, K. (2020b). MapR: A Method for Identifying Native R-Loops Genome Wide.
827 *Current Protocols in Molecular Biology*, *130*(1), e113. https://doi.org/10.1002/cpmb.113

828 Yan, Q., Shields, E. J., Bonasio, R., & Sarma, K. (2019a). Mapping Native R-Loops Genome-
829 wide Using a Targeted Nuclease Approach. *Cell Reports*, *29*(5), 1369-1380.e5.
830 https://doi.org/10.1016/j.celrep.2019.09.052
831 Yan, Q., Shields, E. J., Bonasio, R., & Sarma, K. (2019b). Mapping Native R-Loops Genome-
832 wide Using a Targeted Nuclease Approach. *Cell Reports*, *29*(5), 1369-1380.e5.
833 https://doi.org/10.1016/j.celrep.2019.09.052
834 Yang, X., Li, W., Prescott, E. D., Burden, S. J., & Wang, J. C. (2000). DNA Topoisomerase IIβ
835 and Neural Development. *Science*, *287*(5450), 131–134.
836 https://doi.org/10.1126/science.287.5450.131
837 Yang, Y., McBride, K. M., Hensley, S., Lu, Y., Chedin, F., & Bedford, M. T. (2014). Arginine
838 Methylation Facilitates the Recruitment of TOP3β to Chromatin to Prevent R Loop
839 Accumulation. *Molecular Cell*, *53*(3), 484–497.
840 https://doi.org/10.1016/j.molcel.2014.01.011
841 Yu, G., Wang, L.-G., & He, Q.-Y. (2015). ChIPseeker: An R/Bioconductor package for ChIP
842 peak annotation, comparison and visualization. *Bioinformatics*, *31*(14), 2382–2383.
843 https://doi.org/10.1093/bioinformatics/btv145
844 Zhang, T., Wallis, M., Petrovic, V., Challis, J., Kalitsis, P., & Hudson, D. F. (2020). Loss of
845 TOP3β leads to increased R-loop formation and genome instability. *Open Biology*, *9*(12),
846 190222. https://doi.org/10.1098/rsob.190222
847 Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C.,
848 Myers, R. M., Brown, M., Li, W., & Liu, X. S. (2008). Model-based analysis of ChIP-Seq
849 (MACS). *Genome Biology*, *9*(9), R137. https://doi.org/10.1186/gb-2008-9-9-r137
850 Zhu, A., Ibrahim, J. G., & Love, M. I. (2019). Heavy-tailed prior distributions for sequence count
851 data: Removing the noise and preserving large differences. *Bioinformatics*, *35*(12),
852 2084–2092. https://doi.org/10.1093/bioinformatics/bty895
853

5 FIGURE LEGENDS

Table 1. Summary of uniquely mapped fragments for Aging MapR samples to *Drosophila melanogaster* dm6 genome.

Figure 1. Aging photoreceptor neurons show increased global levels of R-loops that correlate with loss of function and precede age-associated retinal degeneration

(a) Experimental scheme to detect global levels of R-loops in aging photoreceptor neurons

(b) Top: Schematic diagram of the cellular localization of the GFP^{KASH} protein. Dark blue lines represent each lipid layer from the nuclear membrane. Bottom: Schematic diagram of the ommatidia, or structural subunit in the Drosophila compound eye. Each ommatidia is composed of 8 photoreceptor neurons, labeled R1 to R8. Outer photoreceptors express the *ninaE* (Rh1) gene.

(c) Slot blot analysis of R-loop levels from photoreceptor nuclei at day 10, day 30 and day 50 post-eclosion treated with (right) or without (left) RNAse H1. Slot blots were performed using S9.6 antibody (top) and ssDNA for loading control (bottom).

(d) Quantification of S9.6 slot blot signal in aging PRs. S9.6 signal at day 30 and day 50 is normalized to day 10 signal. Values above 1 represent increase signal relative to Day 10. Mean + Standard Deviation (SD), (n=2).

Figure 2. Mapping genome-wide distribution of R-loops in aging PR neurons

(a) Schematic diagram of R-loop mapping technique used in this study (MapR). Immuno-enriched nuclei are incubated with ΔRNAseH-MNase (ΔRH-MNase), where ΔRH binds to R-loops. Enzymatic activation of MNase results in cleavage of surrounding DNA and subsequent R-loop enriched DNA release, which is purified and used for sequencing library preparation.

(b) Spearman correlation heatmap of MapR read distribution over 1000-bp binned genome. Scores between 0 and 1 shown in each box correspond to Spearman's rank score.

(c) Principal component analysis (PCA) of Aging MapR samples based on read distribution over 1000-bp binned genome.

(d) Genomic distribution of MapR peaks during aging. Promoter is defined as the region -2/+2 kb around TSS defined by RefSeq.

(e) Fraction of Reads in Peaks (FRiP) scores for Aging MapR samples. Scores above 0.3 are commonly associated with high quality ChIP-seq datasets as defined my modENCODE standards.

Figure 3. Age-associated changes in R-loop levels in PR neurons are highly dynamic

(a) Metaplot of CPM-normalized MapR signal over gene bodies for all genes across timepoints. Signal is an average obtained from three independent biological replicates per timepoint. TSS indicates Transcription Start Site and TES indicates Transcription Termination Site.

(b) MapR fold change metaplot (top) and heatmap (bottom) for early aging (left) and late aging (right).

(c) Genomic browser inspection of six genes that have either gain (left) or loss (right) of R-loops during aging.

(d) Metaplot of MapR signal around called peaks during aging.

(e) Boxplot of genomic coverage as defined by the total sum of peak width obtained at each time point. Peaks that mapped to scaffold or non-defined chromosomes were excluded from analysis. We used Wilcoxon Rank-Sum test to compare pair-wise differences in the distribution of genomic coverage amongst ages, (n=3).

(f) R-loop signal fold change of quantified peaks in early and late aging as defined by differential occupancy quantified by CPM using csaw and DESeq2. Peaks were annotated to genomic features using Homer and divided based on annotation.

Figure 4. Age-related R-loop accumulation is associated with specific genomic features and decreased expression of long and highly expressed genes

(a) Heatmap showing log_2 ratios of MapR signal around the TSS (left) or TTS (right), comparing day 50 to day 10. Genes are ranked based on their fold change value and divided in four groups based on their position on the heatmap.

(b) Boxplot analysis of GC content, gene length and expression levels for each group of genes divided in four groups based on the MapR fold changes around TSS. P-value is obtained using Wilcoxon test.

(c) as in (b) but for MapR fold changes around the TTS. P-value is obtained using Wilcoxon test.

(d) Volcano plot representing Differentially expressed genes between Day 50 and D10. Differentially expressed genes are obtained using DESeq2 (adjusted p-value < 0.05, |FC|>1.5).

(e) Receiver operating characteristic (ROC) curves for genomic features (GC content, Gene Length and Expression levels) that identify age-regulated genes. AUC stands for Area Under the Curve. Red and blue lines correspond to genes that are up- and down-regulated with age, respectively

Figure 5. Loss of *Drosophila* Top3β leads to increase in R-loop levels

(a) Slot blot analysis of R-loop levels from 3rd instar larvae ubiquitously expressing a siRNA against mCherry (siControl) or against Top3β (siTop3β). Samples are treated with (right) or without (left) RNAse H1. Slot blots were performed using S9.6 antibody (top) and ssDNA for loading control (bottom).

(b) Quantification of S9.6 slot blot of siTop3β and siControl 3rd instar larvae. S9.6 signal is normalized to ssDNA slot blot signal, (n=2).

(c) Bar plots of quantitative PCR validating the downregulation of Top3β upon ubiquitous expression of an siRNA (BDSC#31480). P-value is obtained using T-test, (n=3).

(d) Box plots showing the light preference indices (positive phototaxis) for Rh1>GFP^{KASH},mCherry-RNAI (referred as siControl) or Rh1>GFP^{KASH}, Top3β-RNAi (siTop3β) at day 10 and 30 (6 biological replicates for each time point or RNAi, total 24 experiments; 27 - 33 male flies/experiment). P value obtained using Wilcoxon test.

(e) Optic neutralization of siControl and siTop3β at day 10 and 30 post eclosion. Retinal degeneration (RD) scores were obtained by blindly quantifying 5 biological replicates. Score of 0% means there was no observable loss of rhabdomere or ommatidia.

Figure 6. Top3β regulates expression of a subset of long genes associated with neuronal function in photoreceptor neurons

(a) Volcano plot representing differentially expressed genes (DEGs) between siTop3β and siControl-expressing photoreceptors at day 30 post eclosion. DEGs obtained using DESeq2 (adjusted p-value < 0.05, |FC|>1.5). Size of each point reflect the gene length of the whole gene as defined as most upstream TSS and most downstream TTS.

(b) Box plots showing the gene length (as $\log_2$-transformed bp) for genes identified as down-, up- or not regulated using DESeq2 in siTop3β photoreceptors relative to siControl (adjusted p-value < 0.05, |FC|>1.5). P value obtained using Wilcoxon test.

(c) Gene concept network analysis (Cnetplot) of genes downregulated in siTop3β photoreceptors relative to siControl. Gene length in kilobases is shown next to each gene.

(d) Box plots showing the gene length (as log_2-transformed bp) for genes identified as down-, up- or not regulated using DESeq2 in day 50 photoreceptors relative to day 10. P value is obtained using Wilcoxon test.

(e) Scatter plot comparison of genes that change expression with age and genes that change upon downregulation of Top3β. Red and blue represent genes that are up- or down-regulated with age, respectively. Light blue box represents genes that are downregulated in siTop3β photoreceptors relative to siControl.

(f) Heatmap showing log_2 ratios of aging MapR fold change signal over the gene body for genes that were transcriptionally dysregulated upon downregulation of Top3β, or "Top3β-dependent genes". Genes are ranked based on their fold change value.

(g) Comparison of Top3β protein levels in aging eyes from Rh1>GFP^{KASH} flies, as measured by normalized mass spectrometry abundance. Proteomic samples were prepared from 10- and 40-day old flies, (n=4). Raw data taken from (Hall et al., 2021).

SUPPLEMENTAL FIGURES

Supplemental Figure 1

(a) Aging optic neutralization time course of Rh1>GFP^{KASH} and Rh1>mCherry-RNAi, Rh1>GFP^{KASH}, male flies at day 10, 20, 30, 40 and 50 post eclosion. Retinal degeneration (RD) scores were obtained by blindly quantifying 5 biological replicates.

Supplemental Figure 2

(a) Metaplot of CPM-normalized MapR signal over gene bodies for all genes comparing the standard and modified MapR protocol when protocol is performed using *Drosophila* photoreceptor nuclei.

(b) Metaplot of CPM-normalized MapR signal over gene bodies in HEK293T cells (Data from original MapR paper (Yan & Sarma, 2020b)).

(c) Metaplot (top) and heatmap (bottom) of CPM-normalized MapR signal over gene bodies for samples that were treated with (right) or without (left) RNAseH1. For heatmap, genes are ranked based on their MapR signal.

(d) Genome browser inspection (IGV) of a selected genomic region of samples used for metaplots and heatmaps in Figure S2C. Signal is normalized to CPM and both lanes are shown in the same scale.

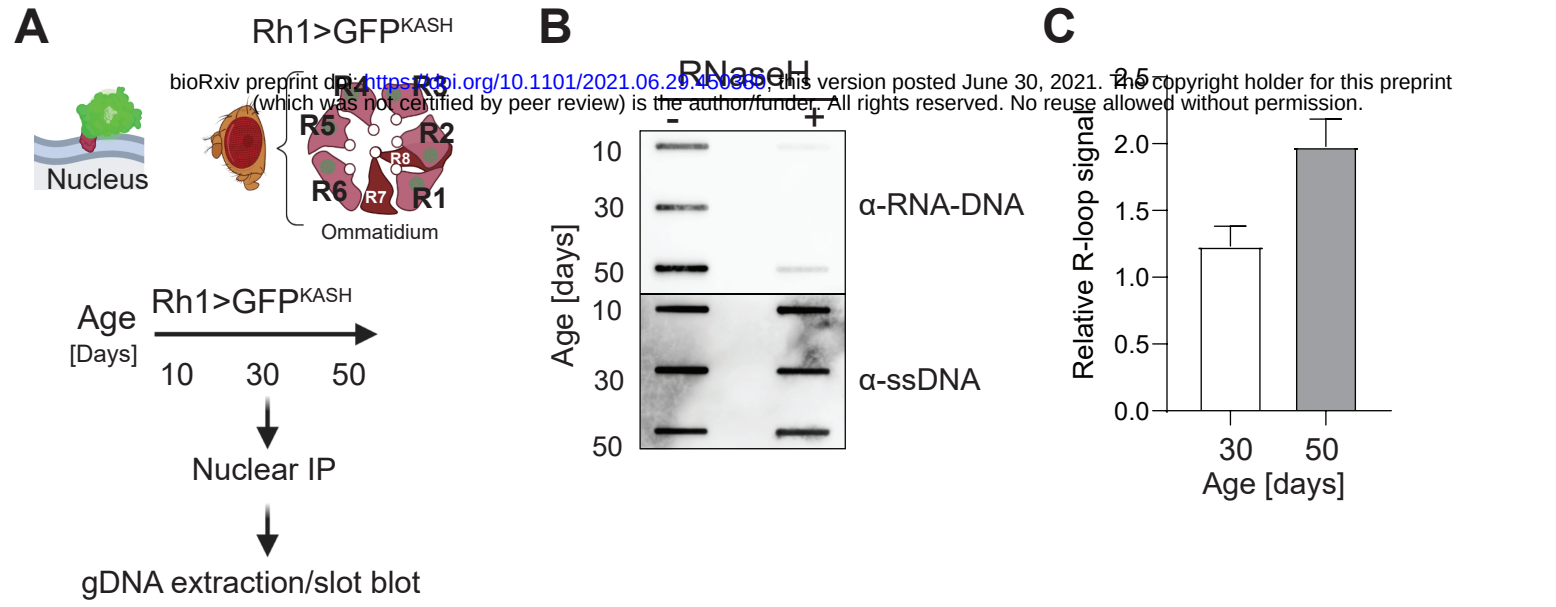
(e) Metaplot of CPM-normalized MapR signal over gene bodies of genes with higher (dark blue) or lower (green) transcript-per-million (TPM) values of 7, obtained with RNA-seq.

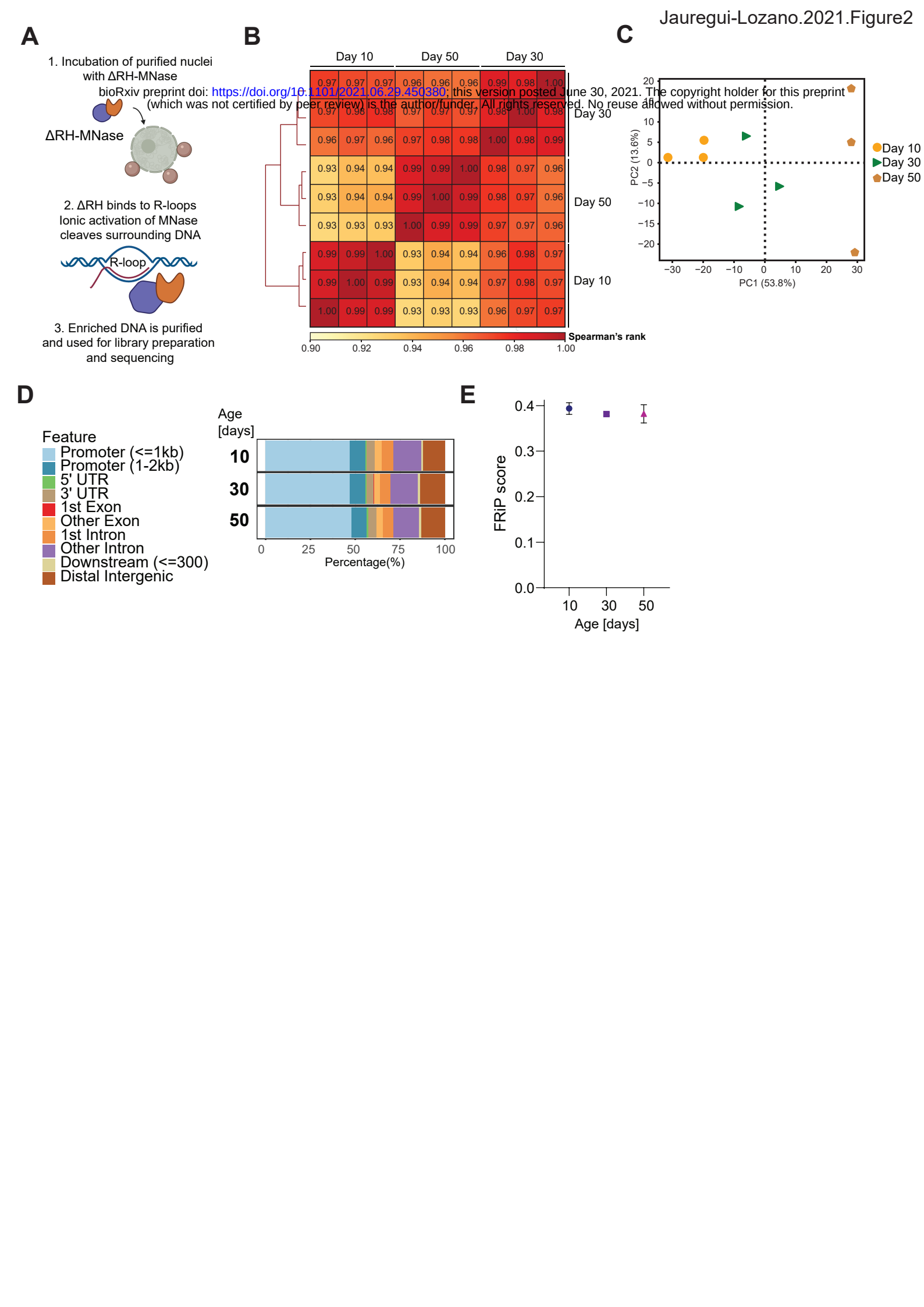
(f) Number of MapR peaks called using MACS2 based on sequencing depth. Bam file is down-sampled to 1, 5, 10, 20, 40 and 50 million mapped fragments and peaks are called.

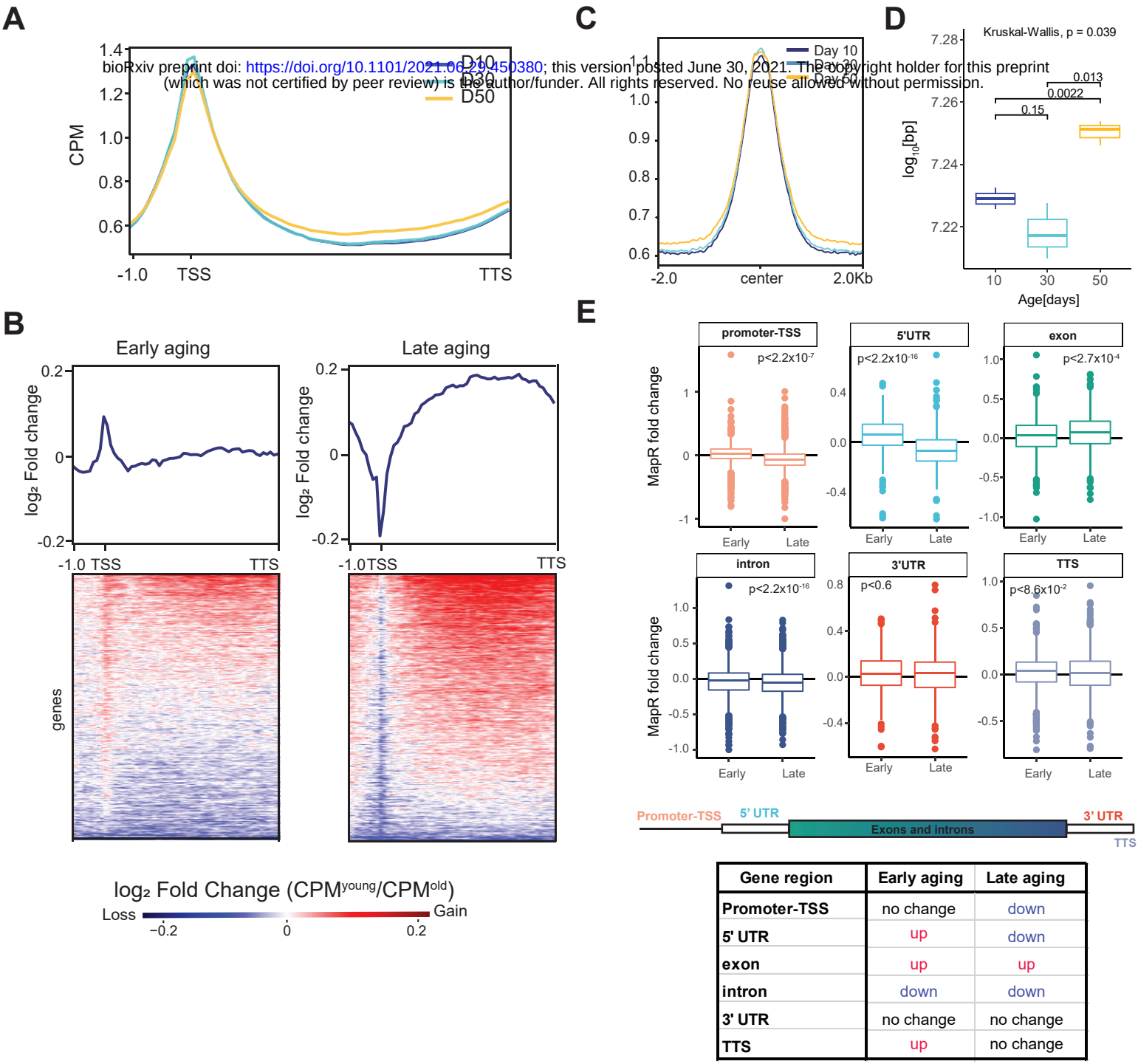
Supplemental Figure 3.

(a) Violin plots showing the peak width distribution for each timepoint. Black line is placed at the calculated mean for peak width at day 10.

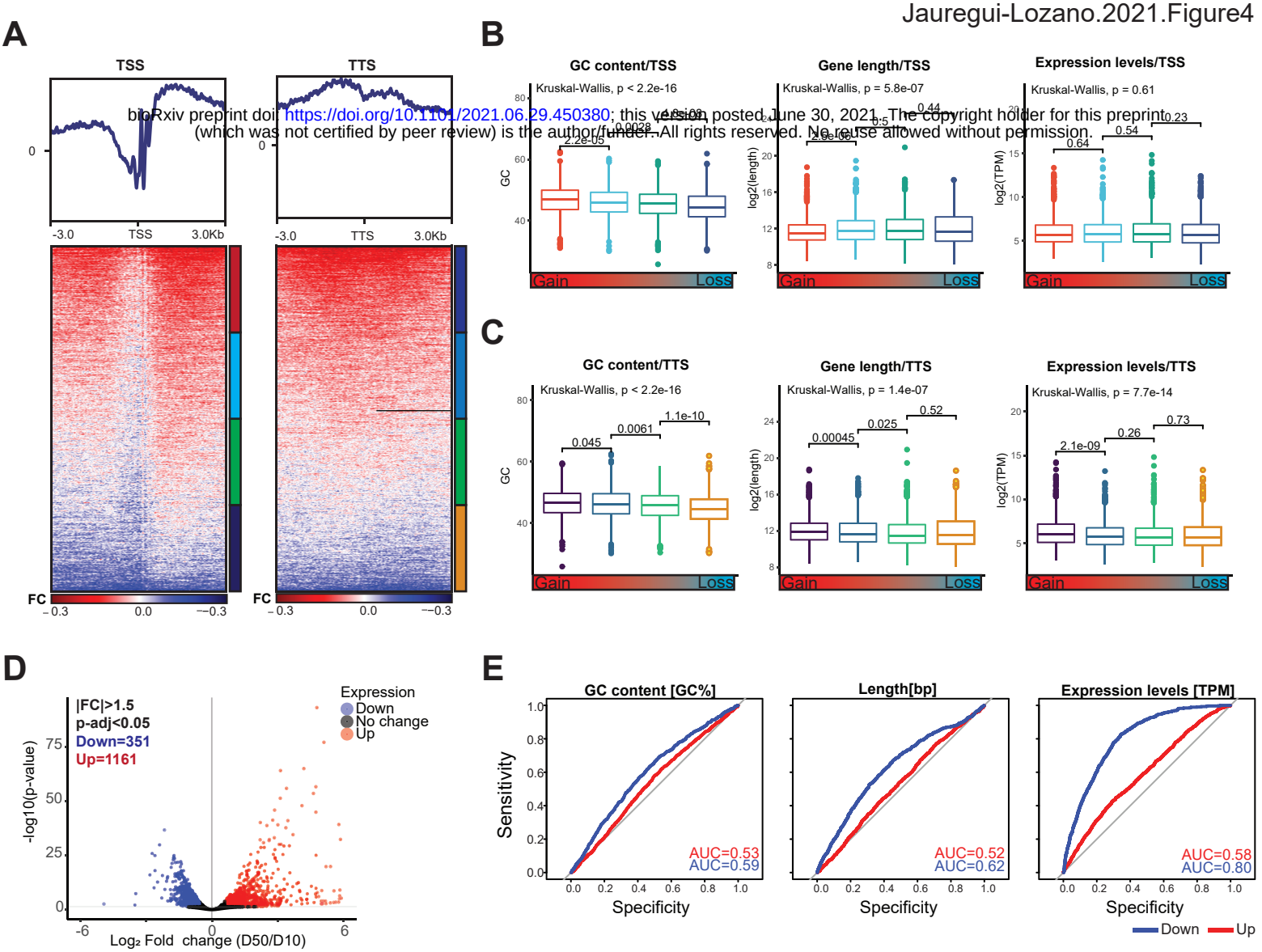
(c) Brand-Altman (MA) plot showing the correlation between $\log_2$-transformed R-loop signal fold change and peak occupancy, as defined by counts per million (CPM) for early (top) and late (bottom) aging. Differential peaks are defined as having a False Discovery Rate (FDR) lower than 5%, and fold change greater than 1.2.

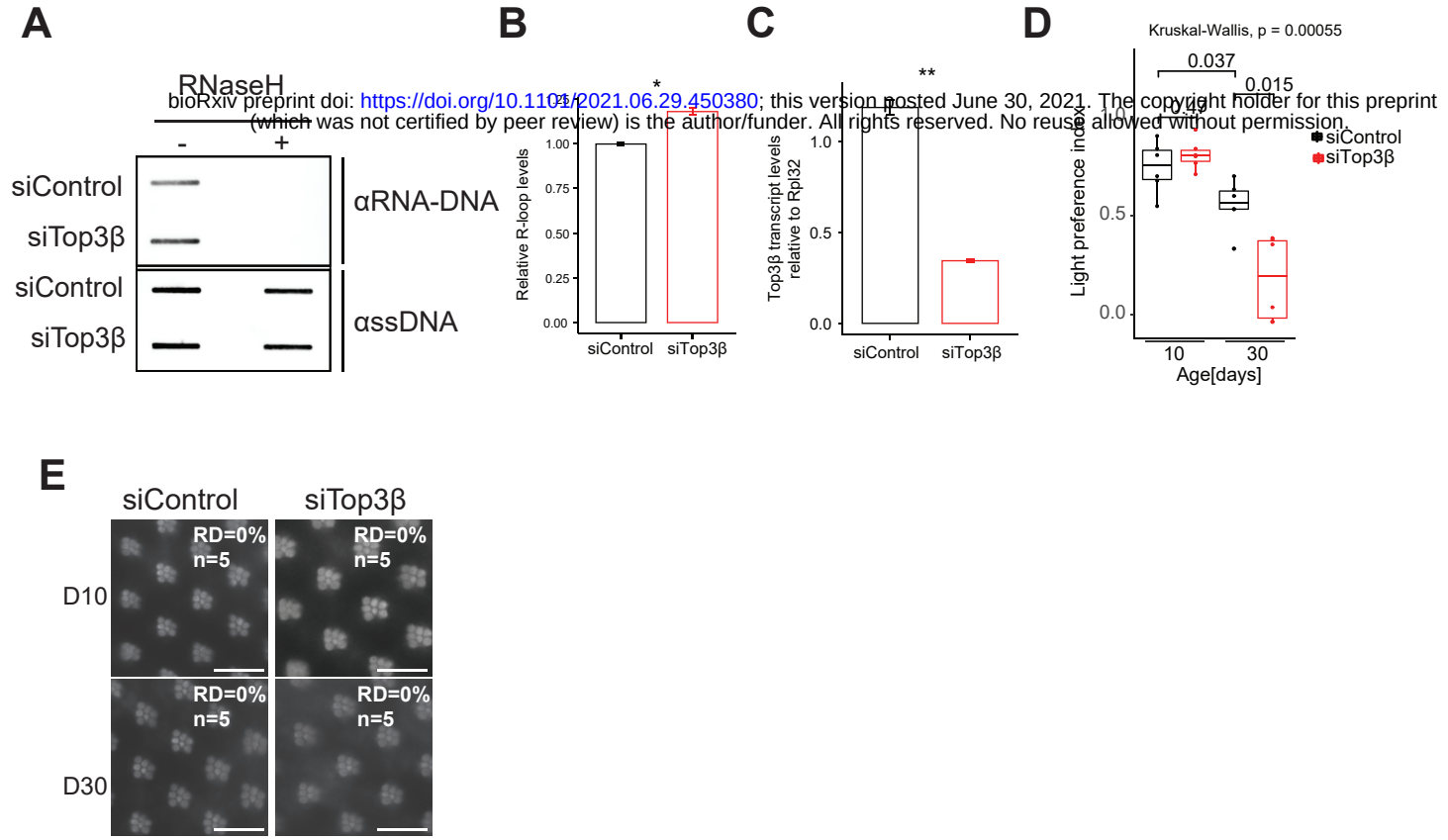


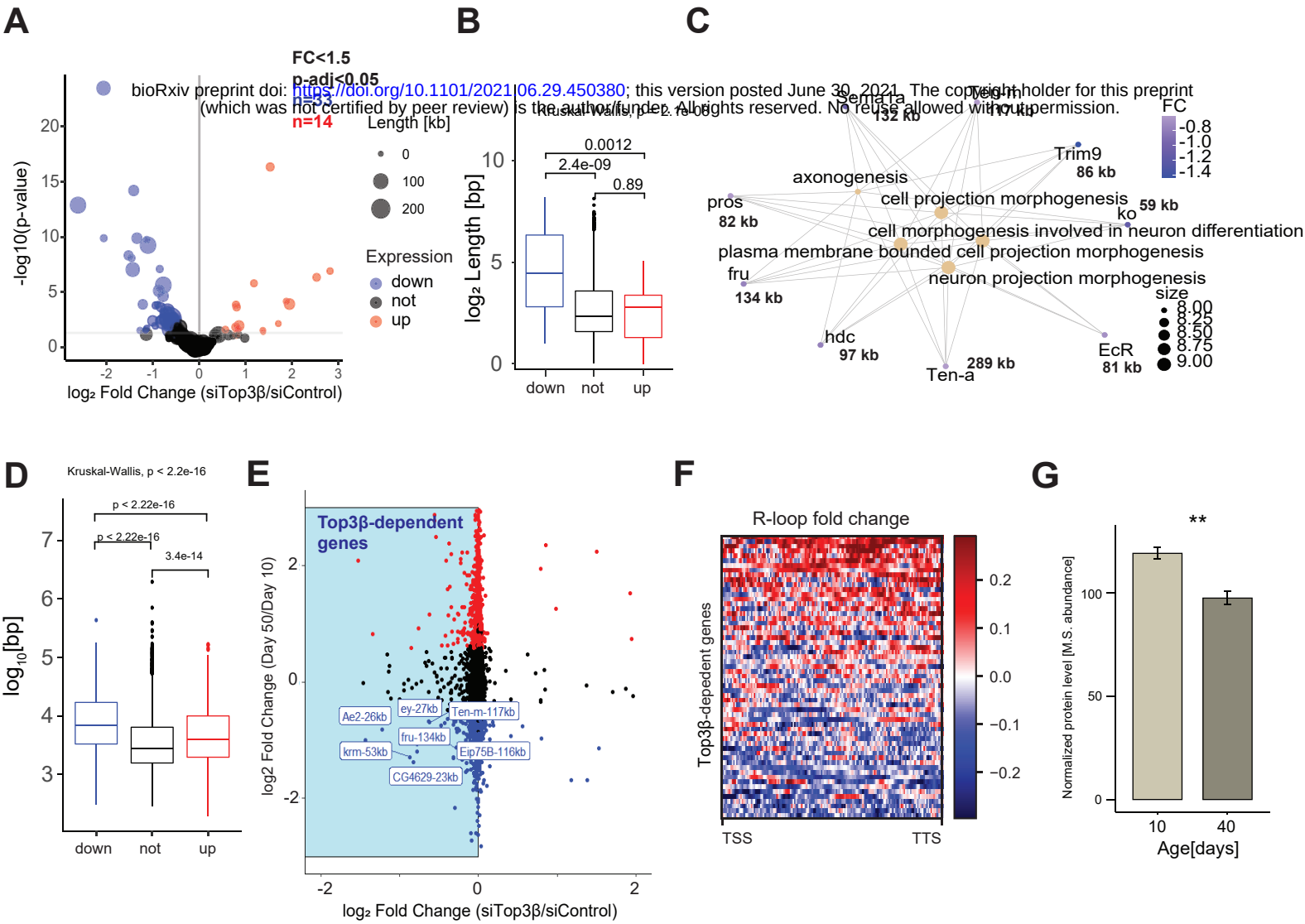




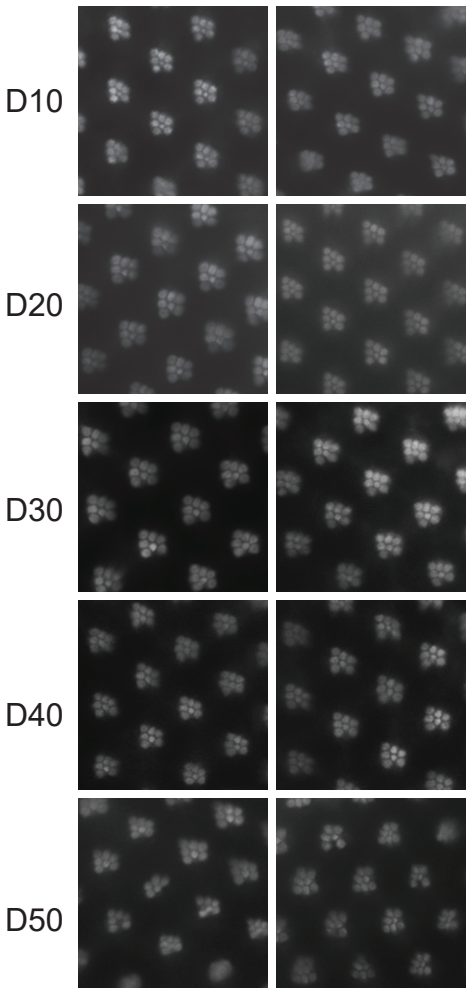
Gene region	Early aging	Late aging
Promoter-TSS	no change	down
5' UTR	up	down
exon	up	up
intron	down	down
3' UTR	no change	no change
TTS	up	no change

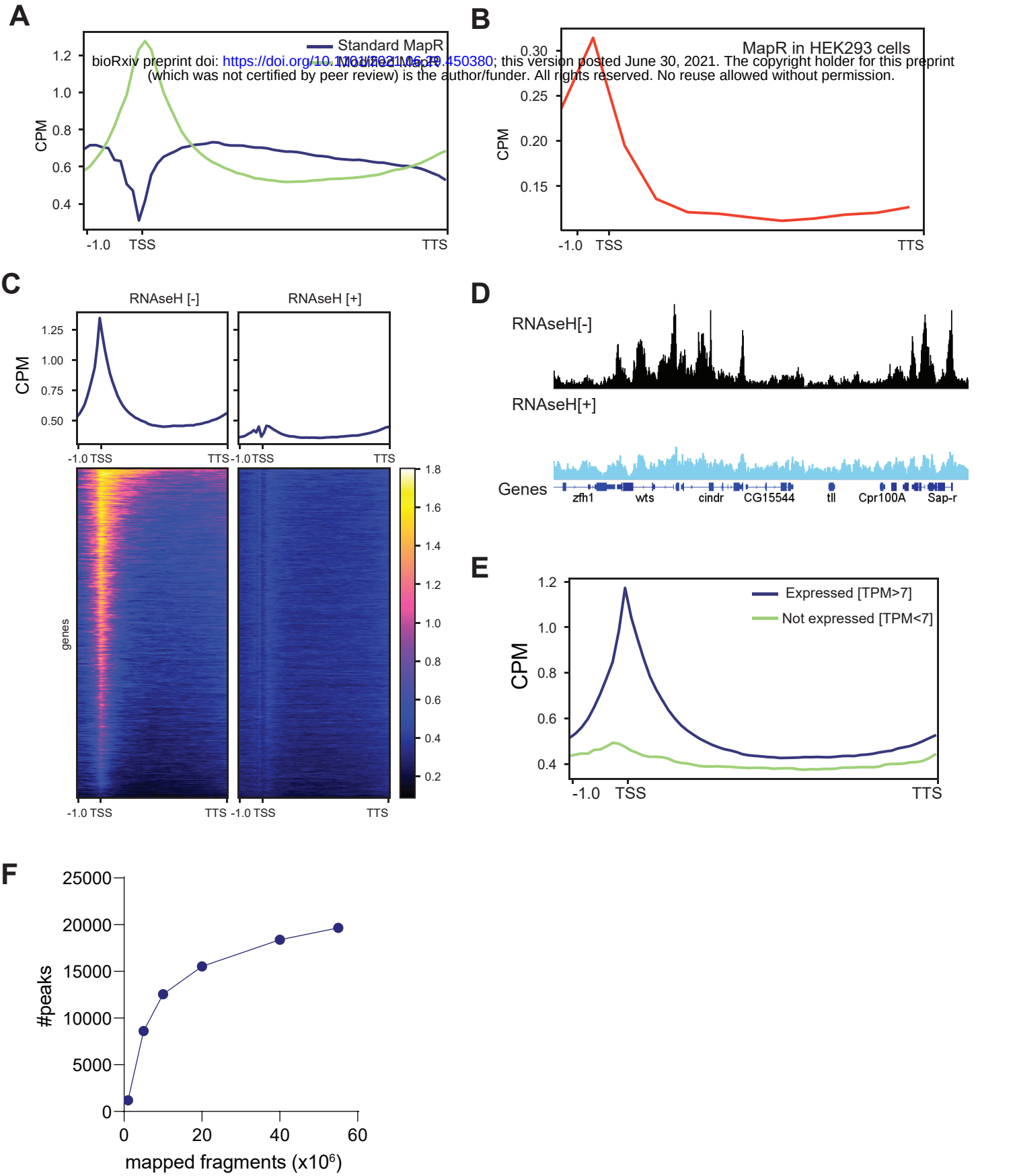


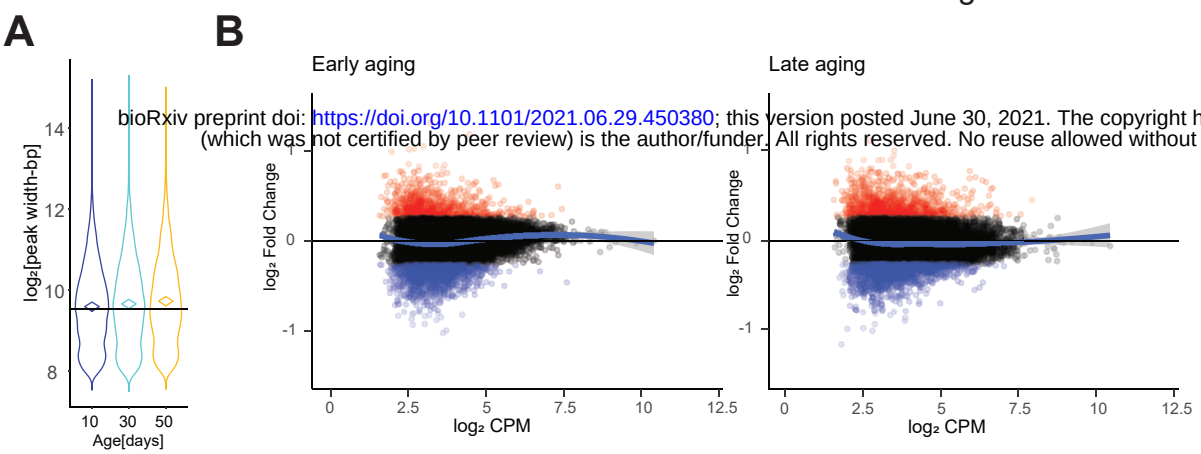




A







Sample	# mapped fragments[x10^6]
D10-1	54784167
D10-2	47815925
D10-2	53697476
D30-1	44366144
D30-2	45811335
D30-3	35823854
D50-1	46970803
D50-2	42207867
D50-3	49960154