Gene body DNA hydroxymethylation restricts the magnitude of transcriptional changes during aging

James R. Occean[1], Na Yang[1], Yan Sun[2], Marshall S. Dawkins[1], Rachel Munk[1], Cedric Belair[1], Showkat Dar[1], Carlos Anerillas[1], Lin Wang[1], Changyou Shi[1], Christopher Dunn[3], Michel Bernier[4], Nathan L. Price[4], Julie S. Kim[2], Chang-Yi Cui[1], Jinshui Fan[5], Moitrayee Bhattacharyya[6], Supriyo De[5], Manolis Maragkakis[1], Rafael deCabo[4], Simone Sidoli[2], Payel Sen[1].

[1]Laboratory of Genetics and Genomics, National Institute on Aging, NIH, Baltimore, MD

[2]Department of Biochemistry, Albert Einstein School of Medicine, Bronx, NY

[3]Flow Cytometry Unit, National Institute on Aging, NIH, Baltimore, MD

[4]Translational Gerontology Branch, National Institute on Aging, NIH, Baltimore, MD

[5]Computational Biology and Genomics Core, Laboratory of Genetics and Genomics, National Institute on Aging, NIH, Baltimore, MD

[6]Department of Pharmacology, Yale University, New Haven, CT

Correspondence: payel.sen@nih.gov

Abstract

DNA hydroxymethylation (5hmC) is the most abundant oxidative derivative of DNA methylation (5mC) and is typically enriched at enhancers and gene bodies of transcriptionally active and tissue-specific genes. Although aberrant genomic 5hmC has been implicated in many age-related diseases, the functional role of the modification in aging remains largely unknown. Here, we report that 5hmC is stably enriched in multiple aged organs. Using the liver and cerebellum as model organs, we show that 5hmC accumulates in gene bodies associated with tissue-specific function and thereby restricts the magnitude of gene expression changes during aging. Mechanistically, we found that 5hmC decreases binding affinity of splicing factors compared to unmodified cytosine and 5mC, and is correlated with age-related alternative splicing events, suggesting RNA splicing as a potential mediator of 5hmC's transcriptionally restrictive function. Furthermore, we show that various age-related contexts, such as prolonged quiescence and senescence, are partially responsible for driving the accumulation of 5hmC with age. We provide evidence that this age-related function is conserved in mouse and human tissues, and further show that the modification is altered by regimens known to modulate lifespan. Our findings reveal that 5hmC is a regulator of tissue-specific function and may play a role in regulating longevity.

Introduction

Aging is characterized by progressive deterioration of physiological function and is the main risk factor for morbidity and mortality. Aging tissues display altered transcriptional landscapes [1-4], consistent with an important role for gene regulation in aging. Indeed, epigenetic alterations, mechanisms that shape chromatin architecture and modulate gene expression patterns, are a cardinal feature of aging [5-7] and age-related diseases, including cancer [8], cardiovascular disease [9], neurodegeneration [10], obesity [11], and type 2 diabetes [11]. Considering the reversible nature of epigenetic marks, they are potential therapeutic targets to ameliorate age-related functional decline.

In this study, we investigated the role of an understudied epigenetic mark, DNA hydroxymethylation (5-hydroxymethylcytosine, 5hmC), in aging. 5hmC is the most abundant oxidative derivative of 5-methylcytosine (5mC), a better-known biomarker of age [12-14]. The oxidation of 5mC to 5hmC is mediated by the ten-eleven translocation (TET) Fe^{2+} 2-oxoglutarate dioxygenase family of enzymes [15, 16]. Several environmental factors may contribute to the formation of 5hmC. For example, (i) reactive oxygen species (ROS) can form 5hmC via abstraction of an H-atom from the methyl group of 5mC [17]; (ii) vitamin C (but not other antioxidants) can induce DNA demethylation and increase 5hmC by acting as a cofactor for TET enzymes, facilitating the reduction of Fe^{3+} to Fe^{2+} [18, 19]; (iii) prolonged quiescence can also increase the level of genomic 5hmC over time. The latter is hinted by findings showing that DNA replication (for example in cancer) dilutes 5hmC [20-22], and is supported by our study. Furthermore, TET proteins have been shown to be more prone to oxidizing 5mC to 5hmC but not 5hmC to its oxidative product, 5-formylcytosine (5fC), fostering accumulation of 5hmC over time [23]. Moreover, activity of the TET enzymes is dependent on α-ketoglutarate, a metabolite generated in the tricarboxylic acid (TCA) cycle and previously reported to promote longevity in various organisms [24-26]. Recent work has also shown that TET enzymes, namely TET1 and TET2, are required for the age-related protective and regenerative effects of OSK (*Oct4, Sox2*, *and Klf4*)-induced partial reprogramming in mouse retinal ganglion cells [27]. These studies strongly suggest a function for 5hmC in aging.

5hmC was formerly regarded as an intermediary step in the DNA demethylation pathway. Importantly, previously established bisulfite-based techniques (e.g., methylation arrays, reduced representation bisulfite

sequencing, whole-genome bisulfite sequencing, etc.) cannot distinguish 5mC from 5hmC [28], thereby confounding interpretations and potentially minimizing the contribution of this modification. Studies show that 5hmC recruits interacting proteins (or "readers") unique from 5mC [29]. Some of these 5hmC interactors, e.g., Uhrf2 (Ubiquitin Like With PHD And Ring Finger Domains 2), were also shown to be tissue-specific and dynamically regulated during differentiation [29]. Concomitantly, multi-tissue studies have documented the presence of 5hmC at enhancer regions and gene bodies of tissue-specific and transcriptionally active genes [30-36], suggesting an active regulatory role for the modification.

Here, we surveyed global levels of 5hmC in multiple young and old organs. We report that 5hmC accumulates in aged tissues due to prolonged quiescence. Using murine liver and cerebellum as model organs, we show that 5hmC plays a critical role in restricting the magnitude of transcriptional changes during aging, especially over tissue-specific genes. We validate this finding in published 5hmC datasets from humans. Lastly, we show that this transcriptional restriction function might be detrimental in the response to stress, and that reduction of 5hmC under stress conditions is important for longevity-promoting interventions.

Results

Global accumulation of 5hmC in aging heart, lung, and liver

We initiated this study by performing a global assessment of 5mC and 5hmC in genomic DNA (gDNA) of young (~2-5 months, n=4) and old (~18-20 months, n=4) C57BL/6JN mice tissues of both sexes, including brain (cortex, cerebellum, hippocampus), heart, lung, liver, spleen, kidney, and muscle (**Fig. 1A, Supp. Table 1 & 2**). As reported previously, brain showed the highest absolute levels of 5hmC (~0.5%) [37] by liquid chromatography tandem mass spectrometry (LC-MS/MS) followed by muscle (0.3%), liver (0.14%), kidney (0.13%), heart (0.12%), lung (0.11%), and spleen (0.09%). Of note, the brain and heart are largely post-mitotic, and the liver, kidney and lung are mildly proliferative upon injury, while the spleen is a site of active proliferation; thus, levels of 5hmC, to some degree, correlate with tissue proliferative capacity. There were no significant differences in 5mC levels between old and young for any of the tissues measured; however, a significant increase in 5hmC levels was detected in the old liver (mean difference = 0.20%), old heart (mean

difference = 0.11%), and old lung (mean difference = 0.08%). The global increase of 5hmC with age without detectable differences in 5mC is concordant with previously reported LC-MS/MS results from mouse liver tissue [38] and does not preclude a reduction of 5mC at local sites (as we show below). Because the liver showed the greatest mean difference in 5hmC with age, we opted to further investigate 5hmC in this tissue. We first corroborated this finding using two orthogonal approaches. We detected significantly higher 5hmC signal in the old by dot blot analysis, using an antibody against 5hmC and gDNA from mouse liver tissue (**Fig. 1B**). Similarly, immunofluorescence microscopy showed a higher 5hmC signal in the nuclei of old liver tissue sections compared to young (**Fig. 1C**). Importantly, this increase of 5hmC with age occurs in both sexes. Taken together, these data indicate a global increase of 5hmC in multiple aged tissues.

5hmC accumulates at genic regions associated with hepatic metabolism during aging

We next investigated the genomic localization of 5hmC using hydroxymethylated DNA immunoprecipitation followed by sequencing (hMeDIP-seq). The hMeDIP-seq was performed on gDNA isolated from the liver of young and old C57BL/6JN mice by immunoprecipitating (IP) with an antibody targeting 5hmC (n=4 biological replicates per group, **Supp. Fig. 1A, Supp. Table 1**). 10% of gDNA was kept as input and did not go through IP. IP success was verified by qPCR analysis of exogenous hydroxymethylated and unmodified DNA spike-in controls that were added before samples underwent IP (**Supp. Fig. 1B**). The sequencing reads were aligned to the GRCm38/mm10 genome assembly and were not significantly different between young and old IP or input samples in terms of sequencing depth, mapped fragments, or fragment length (**Supp. Fig. 1C-F**).

To identify the primary source of variance in genome-wide 5hmC signal, we performed principal component analysis (PCA) on the RPKM (reads per kilobase per million) normalized and input subtracted 5hmC samples. Principal component 1 (PC1) accounted for 79.4% of the variation in 5hmC signal and clearly distinguished the samples by age, designating age as the main contributor to variability in genome-wide 5hmC signal (**Fig. 2A**). We next called 5hmC peaks for young and old IP samples (with corresponding input as control) using MACS2 (FDR cutoff or q = 0.001) [39]. A significantly higher number of peaks were identified in the old (mean=126,934) compared to young (mean = 115,583) (**Fig. 2B**). These peaks tended to be broader, with

higher tag density, but lower mean signal in the old samples (**Fig. 2C**). Correlation heatmap using the peaks called showed high replicate reproducibility (**Fig. 2D**).

To identify age-related differentially hydroxymethylated regions (DHMRs), we performed DiffBind analysis [40] using peaks obtained from MACS2. We identified 7,585 peaks with significantly higher 5hmC enrichment in the old (hyper DHMRs, fold change [FC] ≥ 2, FDR < 0.05) and 863 peaks with significantly higher enrichment in young samples (hypo DHMRs, FC ≤ -2, FDR < 0.05) (**Fig. 2E, Supp. Table 3**). 5hmC signal at the center of the hyper DHMRs showed reproducibly higher signal for old compared to young, and vice versa for hypo DHMRs (**Fig. 2F**). **Fig. 2G** shows genome browser view examples of 5hmC signal detected in hyper and hypo DHMRs.

To gain insight into the functional pathways associated with the DHMRs, we performed Gene Ontology (GO) analysis using Genomic Regions Enrichment of Annotations Tool (GREAT) [41]. Pathways enriched for hypo and hyper DHMRs were largely constituted of metabolic and mitochondrial-related terms, including small molecule, lipid, and carboxylic acid metabolic process as well as mitochondrion organization and transport (**Fig. 2H**). The similarity in the GO terms between age-related hypo and hyper DHMRs indicates 5hmC undergoes dynamic changes at genomic regions associated with metabolic and mitochondrial function. It is noteworthy that the liver is a key regulator of whole-body metabolism, and that 5hmC has been previously reported to mark tissue-specific genes [30, 31, 35], suggesting that age-related differences in 5hmC may, in general, occur at tissue-specific genes. Consistent with this notion, we assessed transcription factors (TFs) associated with the top 500 genes marked by 5hmC in young and old using landscape in-silico subtraction analysis (LISA) [42], which revealed motifs for several TFs that have been previously implicated in liver-specific function, including peroxisome proliferator-activated receptors (PPARγ) [43], T-box transcription factor 3 (TBX3) [44, 45], and GLIS Family Zinc Finger 2 (GLIS2) [46] (**Supp. Fig. 1G**). We then annotated the DHMRs to CpG and genic features, which revealed that most changes were occurring at regions farthest (>4kb) from CpG islands (interCGIs) and in gene bodies, primarily intronic regions (**Fig. 2I**). Furthermore, when we traced 5hmC signal from the transcription start site (TSS) to the transcription end site (TES) of all known mm10 genes, we observed a pronounced and significant accumulation of 5hmC at gene bodies in the aged liver (**Fig. 2J**).

Collectively, these analyses show that 5hmC accumulates in the aged liver genome-wide and undergoes dynamic changes at gene bodies and regions associated with liver-specific function, primarily hepatic metabolism.

Gene body 5hmC restricts the magnitude of transcriptional changes during aging

The age-related accumulation of 5hmC at gene bodies and metabolism-related loci in the liver prompted us to investigate the functional role of gene body 5hmC. We analyzed RNA-seq data that we previously generated from the livers of young (~2 months and n=3) and old (~18 months and n=3) mice of both sexes (to be reported elsewhere, Yang et al. *Molecular Cell*, under revision**, Supp. Table 1**). We first assessed the association between gene body 5hmC and gene expression "*within* each age group and *across* different genes" by tracing 5hmC signal over the gene body of protein-coding genes with low, intermediate, and high expression, ranked according to the averaged DESeq2 normalized mRNA count (**Supp. Table 4**). Consistent with previous studies [30-33, 35], we observed a positive association between gene body 5hmC and gene expression for both young and old age groups (**Fig. 3A**).

We next interrogated the role of gene body 5hmC in transcriptional changes "*across* age groups". Spearman's rank-order correlation revealed a significant negative correlation between mean gene body 5hmC signal and mRNA fold change (old vs young) (**Supp. Fig. 1H**, panels 1 and 2, note the negative rho). Similarly, we also identified a significant negative correlation between 5hmC fold change (old vs young) and mRNA fold change (**Supp. Fig. 1H**, panel 3, note the negative rho). These correlational analyses suggest that increased 5hmC levels may downregulate transcriptional changes during aging which we investigated further (see below).

Interestingly, we noted that, in the scatter plots (**Supp. Fig. 1H,** panels 1 and 2), genes with lower mean 5hmC signal are broadly dispersed along the y-axis (mRNA fold change) and as 5hmC increases, the dispersion in mRNA fold change decreases and the values center around y=0 (no difference in mRNA between old and young). We confirmed that this distribution was dependent on 5hmC signal by generating random 5hmC values for each gene and correlating them with corresponding mRNA fold change. As shown in **Supp. Fig. 1H** (panel 4), this correlation was not statistically significant and did not mimic the trend in dispersion

observed between young and old 5hmC levels and mRNA fold change. To validate this finding, we ranked all detectable protein-coding genes into 100 groups (~169-192 genes per group) based on increasing mean gene body 5hmC signal. We computed the variance (measure of dispersion) in mRNA fold change for all 100 groups. As speculated, there was a strong negative correlation between increasing 5hmC signal and variance in mRNA fold change for both young and old age groups (**Fig. 3B**). These observations led us to speculate that 5hmC may be associated with lower dispersion and magnitude of transcriptional changes with age. Alternatively, since 5hmC is positively correlated with gene expression within an age group (**Fig. 3A**), the low dispersion with high 5hmC could simply reflect the lower variance due to high expression and less noise in these genes. We later address this point by assigning a functional role for 5hmC in this process using liver regeneration as a method to reduce 5hmC and observing an outcome on transcriptional changes (see **Fig. 5**). Together, our findings suggest a potentially strong regulatory function for gene body 5hmC on mRNA fold change, wherein genes with increased levels of 5hmC show lower variability in transcriptional changes with age.

To visualize the negative correlation between 5hmC change and transcriptional change with age, we sought to compare gene expression "*across* age groups and *within* genes". We ranked all genes based on either fold change (old vs young, **Fig. 3C, Supp. Table 4**) or absolute fold change neglecting direction (|old vs young|, **Fig. 3E, Supp. Table 4**). We then traced the gene body 5hmC signal over the bottom 33% (lower fold change, n=6,340) or top 33% (higher fold change, n=6,339) of genes (**Fig. 3D & 3F**). We found that both genes with relatively low mRNA fold change with age (**Fig. 3D**, left panel) and lower magnitude of mRNA fold change with age (**Fig. 3F**, left panel), had pronounced and significantly higher 5hmC in the old, whereas genes that gained expression with age (**Fig. 3D**, right panel) and underwent dramatic transcriptional changes with age (**Fig. 3F**, right panel) had an overall low 5hmC profile but still experienced a modest, and significant, increase in 5hmC with age. In sum, these data indicate a dual role for 5hmC in (1) restricting the variability and magnitude of gene expression changes with age, and (2) a propensity to repress transcription. The former is predicted to be potentially protective, preserving homeostatic expression of tissue-specific genes. The latter is likely to be detrimental, as it could decrease tissue-specific functions with age. Notably, in lower model

organisms lacking 5hmC (*Caenorhabditis elegans* and *Drosophila melanogaster*), the age-related transcriptionally restrictive function has been reported for another gene body epigenetic modification, H3K36me3, in that case, global reduction of H3K36me3 leads to higher variability in age-dependent transcriptional changes and shorter lifespan [47].

Age-related differences in 5hmC occur without detectable differences in 5mC

Since 5hmC is a product of 5mC oxidation, we inquired whether an age-related transcriptional restriction function for 5hmC would also be apparent for 5mC. Accordingly, we profiled 5mC genome-wide in gDNA isolated from liver tissue of the same mice as the hMeDIP-seq using methylated DNA immunoprecipitation followed by sequencing (MeDIP-seq) with an antibody targeting 5mC (QC metrics in **Supp. Fig. 2A**-**F**). In PCA, PC1 accounted for 31.8% of the variability in genome-wide 5mC signal and showed a modest clustering of the samples by age (**Supp. Fig. 3A**). Using the same MACS2 parameters as the hMeDIP-seq (q = 0.001), we called a greater mean number of peaks in old (mean = 32,435) than young (mean = 20,471.5); however, this difference was not statistically significant (**Supp. Fig. 3B**). The peaks in the old samples tended to be broader, with higher tag density, and higher mean signal (**Supp. Fig. 3C**). A correlation heatmap using peak calls did not show clear clustering of samples by age (**Supp. Fig. 3D, contrast to Fig. 2D**).

Differential analysis of old vs young using the same DiffBind parameters and threshold as the liver hMeDIP-seq resulted in only 1 peak surviving FDR < 0.05 with a modest fold change ($\log_2$ FC = 0.0035). To enable downstream comparison with the liver hMeDIP-seq data, we relaxed the threshold for statistical significance and classified differentially methylated regions (DMRs) using $p < 0.05$. Using these parameters, we obtained 1,809 total DMRs, 1,515 with greater enrichment in old ($\log_2$ FC > 0, $p < 0.05$) and 294 with greater enrichment in young ($\log_2$ FC < 0, $p < 0.05$) (**Supp. Fig. 3E, Supp. Table 3**). Verification of the 5mC signal at the center of the hyper DMRs showed reproducibly higher signal for old but not young, and vice versa for the hypo DMRs (**Supp. Fig. 3F**). Moreover, 5mC signal at the liver *DHMRs* (from **Fig. 2E**) showed that hyper DHMRs had overall higher 5mC signal than hypo DHMRs, indicating that 5mC to 5hmC conversion contributed to the higher 5hmC enrichment with age (**Supp. Fig. 3G**). Example genome browser views of hyper and hypo DMRs are shown for individual replicates in **Supp. Fig. 3H.** In contrast to the liver DHMRs

which were prominently enriched for metabolic terms (**Fig. 2H**), GO analysis of the liver DMRs using GREAT revealed terms related to immune function and transport such as calcium ion import (**Supp. Fig. 3I**). In fact, previous work comparing 5mC and 5hmC has shown that 5hmC is a better marker of tissue-specific genes than 5mC [31], so it is reasonable that the liver DHMRs would be mostly associated with liver-specific function (i.e., metabolic function) as opposed to the DMRs. CpG and genic annotations of the DMRs were comparable with the DHMRs, wherein most DMRs were associated with interCGIs and gene bodies, primarily intronic regions (**Supp. Fig. 3J**). 5mC signal at gene bodies of all mm10 genes did not show significant differences in 5mC with age (**Supp. Fig. 3K**) as shown for 5hmC (**Fig. 2J**). Overall, these results confirm previous reports that 5hmC is a better marker of tissue-specific genes than 5mC and show that age-related differences in 5hmC occur without pronounced detectable differences in 5mC at these locations.

We next assessed whether 5mC also predicted the magnitude of transcriptional changes with age by tracing 5mC signal at gene bodies of the top and bottom 33% of genes with relatively lower and higher magnitude of transcriptional changes with age (**Fig. 3G**). In contrast to what we observed for 5hmC patterns (**Fig. 3E & 3F**), we found no significant differences in 5mC signal for either group of genes. However, genes that underwent lower magnitude of transcriptional changes with age trended towards lower 5mC signal in the old, in agreement with the increase of 5hmC in the old.

In general, the genes that underwent dramatic transcriptional changes with age and displayed lower 5hmC signal (**Fig. 3E**, right panel), tended to have significantly shorter 5’ UTRs (untranslated regions), 3’ UTRs, transcript length, CDS (coding sequence) length, and had fewer of exons compared to the genes that were transcriptionally restricted with age (**Fig. 3H**). This suggests that relatively long genes may be more prone to 5hmC accumulation. Finally, heatmaps of 5hmC gene body enrichment and mRNA fold change for the liver-DHMR GO terms “mitochondrion organization” and “fatty acid metabolic process” recapitulated 5hmC’s transcriptionally restrictive function with age (**Fig. 3I**).

Overall, our results show that genes that undergo low mRNA fold change and low magnitude of transcriptional changes with age are characterized by a pronounced and significant accumulation of gene body 5hmC during aging but with a modest and non-significant decrease of 5mC. This suggests that although 5hmC

is catalyzed from 5mC, gene body 5hmC levels are stably enriched in aged tissues and may possibly exert a greater influence on age-related transcriptional changes.

Alternative splicing mediates 5hmC's transcriptionally restrictive function through decreased binding of splicing factors

DNA modifications have been shown to influence transcriptional activity via recruitment of proteins that alter chromatin architecture or transcription factor binding [48]. Accordingly, we sought to identify 5hmC-protein interactors as a potential mechanism through which the modification might restrict the magnitude of transcriptional changes with age. We performed oligonucleotide mass spectrometry using nuclear extracts, prepared from young (n=4) and old (n=4) mouse liver tissue from both sexes, and three 20 bp DNA oligos that were either unmodified (C), methylated (5mC), or hydroxymethylated (5hmC) at the cytosine nucleotides (**Fig. 4A, Supp. Table 1 & 5**). The oligos were designed using an endogenous DNA sequence around the summit of two hyper DHMRs, one associated with the GO term "mitochondrion organization" (oligo 1) and one associated with the GO term "fatty acid metabolic process" (oligo 2) from the liver DHMR GREAT analysis (**Fig. 2H)**. Genome browser views of the regions used to design the oligos are shown in **Supp. Fig. 4A**. Following streptavidin capture, bound peptides from the DNA-pull downs as well as 10% input were identified by mass spectrometry.

We identified 221 proteins in the nuclear extract from the old mouse liver samples that were significantly enriched in the oligo 1 5hmC pull-down and 759 that were depleted (5hmC vs input) (**Fig. 4B, Supp. Table 6**). For oligo 2, 245 proteins were enriched in the 5hmC pull-down and 1027 were depleted (**Fig. 4C**). 49% (154) of the proteins were found to be commonly enriched in both oligo 1 and oligo 2 5hmC pull-downs (**Supp. Fig. 4B**, top), while 53% (619) were found to be commonly depleted (**Supp. Fig. 4B**, bottom). The common 5hmC-enriched proteins were associated with biological processes relating to translation, transcription, and chromatin regulation, while the commonly depleted proteins were associated with metabolic processes and RNA splicing (**Supp. Fig. 4C**).

We next assessed the specificity of protein interactions for 5hmC, by measuring differences in abundance of these proteins in 5hmC, 5mC, or C pull-downs in the old, as a potential mechanism underlying

the relationship between 5hmC and transcriptional restriction or propensity for repression. For oligo 1, we identified 3 proteins that were significantly depleted for the 5hmC modification compared to 5mC and C, which included the NuA4 histone acetyltransferase complex associated protein, MORF4L2 (**Supp. Fig. 4D**, in red). 12 proteins were significantly enriched for 5hmC, two of which are known to be involved in the endoplasmic reticulum stress response (HYOU1 [49] and TMEM259 [50], **Supp. Fig. 4D**, in red), a condition previously linked to changes in chromatin architecture [51]. By contrast in the old samples, we found 90 proteins that were significantly enriched in the 5hmC oligo 1 pull-down and 49 that were depleted compared to 5mC and C oligos in the young (**Supp. Fig. 4E**). GO terms for the proteins depleted in the 5hmC pull-downs in the old, included RNA splicing, mRNA processing, and chromatin silencing, while the proteins enriched in the 5hmC pull-down were associated with translation, metabolic and mitochondrial processes (**Fig. 4D**).

For oligo 2, we identified 3 proteins (PISD, G3BP1, PA2G4) that were significantly enriched in the 5hmC modification compared to 5mC and C in the old (**Supp. Fig. 4F**). Notably, G3BP1 is a DNA and RNA-binding protein implicated in mRNA decay [52] and previously reported to regulate steady state levels of mRNAs with highly structured 3'UTRs [53]. Thus, 5hmC's binding to G3BP1 would agree with the negative correlation observed between 5hmC and age-related mRNA fold change in (**Supp. Fig. 3H**, panels 1-3) as well as its modest repressive function (**Fig. 3C & 3D**). Moreover, GO analysis of the proteins depleted for 5hmC for oligo 2 were similar to oligo 1 in showing enrichment of GO terms associated with RNA splicing, mRNA processing, and chromatin regulation (**Fig. 4E,** top panel). Additionally, comparisons between 5hmC oligo in the old and 5mC and C in the young revealed that proteins depleted for 5hmC were involved in mRNA processing and multiple RNA splicing-related GO terms (**Supp. Fig. 4G & Fig. 4E,** bottom panel)**.** Altogether, these results show that 5hmC consistently has decreased affinity for proteins involved in splicing compared to 5mC and C within aged tissues and during aging.

The prominent depletion of RNA splicing-related proteins in the aged liver for the majority of the 5hmC oligo pull-downs, coupled with the pronounced age-related changes in 5hmC at intronic regions (**Fig. 2H**) motivated us to investigate alternative splicing during aging. Using rMATS (Multivariate Analysis of Transcript Splicing) [54], we detected several alternative splicing events between old and young liver samples in bulk RNA-

seq data (**Fig. 4F**, top panel, **Supp. Table 7**). We found that genes with relatively higher gene body 5hmC in the old experienced the most alternative splicing events with age (**Fig. 4F**, middle panel). In agreement, genes that underwent the lowest magnitude of transcriptional changes with age also experienced the most age-related alternative splicing events (**Fig. 4F**, bottom panel).

Although isolated alternative splicing events can be detected with short-read sequencing (for example with rMATS), comprehensive analysis of splice isoforms is limited. With nanopore sequencing, where read length is equal to fragment length, entire transcripts can be sequenced in single reads giving a more detailed view of isoform diversity. We thus performed direct RNA-seq (dRNA-seq) with nanopore sequencing (Oxford Nanopore Technologies) on the MinION platform with livers from young (~2 months, n=4) and old (~18 months, n=4) C57BL/6JN mice of both sexes (**Supp. Table 1**). dRNA-seq is free from PCR bias and can provide information on expression, splicing isoforms, transcript length and poly A length. PCA plots of normalized transcript counts clearly segregated the samples by age (**Supp. Fig. 4H**) like our short-read data (reported elsewhere, Yang et al, *Molecular Cell*, under revision). A comparison of the differential isoform usage from long-read data (**Supp. Table 8**) showed similar trends to rMATS results (**Supp. Fig. 4F)**, thus providing additional validation of increased alternative splicing events in genes marked by high 5hmC in old. i.e., genes showing lower absolute fold change with age (**Fig. 4G**). A few examples of these alternative splicing events are shown in **Supp. Fig. 4I**. A survey of the mean transcript lengths and poly A lengths in these genes showed a small but significant decrease in overall transcript length (mean difference = -45.5 bp) but slight increase in poly A length (mean difference = 13.1) in the old (**Fig. 4H**, top left and bottom left plots, **Supp. Table 9**). The length difference could indicate a prevalence of short isoforms or ongoing decay in old tissue. By contrast, genes undergoing dramatic changes in gene expression with age showed no evidence of transcript length change (**Fig. 4H**, top right plot, **Supp. Table 9**). The poly A length of this set of genes however, showed a very prominent increase (mean difference = 48.5) in the old, suggesting that the corresponding transcripts are relatively stabilized. Several studies have shown that splicing factors can influence transcriptional elongation rates by modulating pol II elongation, and vice versa [55-57]. Splicing factors have also been shown to regulate steady state mRNA levels by producing alternatively spliced transcripts targeted for degradation by the

nonsense-mediated decay pathway [58-61]. Together, these events could help fine-tune homeostatic expression of tissue-specific genes. Thus, our data suggest a potential role for RNA splicing in mediating 5hmC's age-related transcriptionally restrictive function.

Quiescence and senescence drive the increase of 5hmC with age and impact cellular function

We next sought to address the mechanisms responsible for the increase of 5hmC in the aged liver. RNA-seq data showed no significant differences in mRNA levels of 5hmC regulators, i.e., *Tet1*, *Tet2*, and *Tet3*, between old and young (**Fig. 5A**). We did, however, detect a significant age-related decrease in a gene with indirect links to 5hmC, *Idh3b* (isocitrate dehydrogenase 3, beta subunit). However, because isocitrate dehydrogenases are responsible for converting isocitrate to α-ketoglutarate, a cofactor for TET enzymes, their decrease in expression contradicts the increase of 5hmC in the aged liver. We further investigated whether TET activity was altered in the aged liver and found no significant differences (**Fig. 5B**). These results indicate that the increase of 5hmC with age is not a simple consequence of cognate enzyme abundances or activity.

Studies have shown that TET proteins have higher enzymatic activity towards 5mC than 5hmC, and therefore are more prone to write than erase 5hmC from the genome [23]. By contrast, DNA replication has been reported to dilute 5hmC [20, 21]. This led us to hypothesize that in the aged liver, where cells are not actively dividing, are in prolonged quiescence in a nutrient-rich environment (a state that we have previously called hyper-quiescence, to be reported elsewhere, Yang et al. *Molecular Cell*, under revision), this balance is not well-maintained, thus leading to genomic increases of 5hmC. To test this hypothesis, we induced quiescence in HepG2 cell lines in two ways: by contact inhibition, a model that mimics the state of non-dividing organs in two dimensions, and by serum starvation. We detected a significant increase of 5hmC with longer times in quiescence induced by contact inhibition (**Fig. 5C)** but not serum starvation (**Supp. Fig. 5A**), suggesting that the hyper-quiescent state contributes to the accumulation of 5hmC.

Given that 5hmC marks mitochondrial and metabolic genes in the aged liver, coupled with the observed negative correlation between 5hmC and mRNA fold change (**Supp. Fig. 1H**), we further assessed whether the increase in 5hmC would directly influence metabolic function. We increased 5hmC by contact inhibition-induced quiescence and observed significantly lower total ATP production compared to proliferating cells (**Fig.**

5D). Notably, there were no significant differences in the mitochondrial membrane potential of the quiescent cells compared to proliferating cells (**Fig. 5E**), suggesting that the reduced ATP levels with high 5hmC is not due to a loss of mitochondrial membrane integrity but possibly downregulation of mRNA levels.

Having observed that cell cycle arrest can drive the increase of 5hmC, we next inquired whether 5hmC levels might also be higher in senescence, a state of stable cell cycle arrest. Senescent cells are known to accumulate in aged tissues, contributing to age-related decline, and thus may partially contribute to the increase of 5hmC with age (**Fig. 1**). We induced senescence in human WI-38 cells by treatment with etoposide (ETIS), ionizing (γ) radiation (IRIS), or oxidative stress (OSIS) for 10 days. Cells were confirmed to be in the senescence state by the level of senescence-associated beta-galactosidase (SA-β-gal) staining (**Supp. Fig. 5B**), increased expression of known senescence-associated markers p16, p21, and IL6; downregulation of Lamin B1 (**Supp. Fig. 5C**); and reduced BrdU labeling (**Supp. Fig. 5D**). We found that senescent cells, either ETIS, IRIS, or OSIS, had significantly higher levels of genomic 5hmC compared to proliferating controls (**Fig. 5F**). Our results suggest that the accumulation of senescent cells with age may contribute to the increase of 5hmC in the aged liver.

Given the contribution of prolonged cell-cycle arrest (either quiescence or senescence) to the age-related increase of 5hmC, we next asked whether disrupting quiescence by prompting liver regeneration in the aged liver would reduce global 5hmC levels *in vivo*. We performed 70% partial hepatectomy in young and old mice of both sexes (n=3 biological replicates per timepoint, **Fig. 5G**) and collected liver samples pre-surgery and 48, 72, 96, 120, and 240 h post-surgery, as reported elsewhere (Yang et al. *Molecular Cell,* under revision). Results from LC-MS/MS showed a significant combined effect of age and regeneration on relative 5hmC levels. Importantly, we observed a progressive dilution of the age-accumulated 5hmC with liver regeneration, supporting previous data on DNA replication-induced reduction of 5hmC levels [20-22]. More interestingly, the dilution of 5hmC after regeneration significantly increased the magnitude of transcriptional changes in genes that otherwise undergo low mRNA absolute fold change with age (**Fig. 5I, Supp. Table 1**, RNA-seq of before and after regeneration liver samples to be reported elsewhere, Yang et al, *Molecular Cell*, under revision), consistent with 5hmC's transcription restriction function. Opposingly, genes that undergo high

mRNA absolute fold change with age, and “all” genes, showed a significant decrease in transcriptional amplitude and mean gene expression after regeneration (**Fig. 5J & 5K**). These results are congruent with the notion that the regenerated liver is “younger”, and therefore shows only modest age-related transcriptional changes compared to its pre-surgery old counterpart. More importantly, this *in vivo* perturbation of 5hmC levels assigns a pivotal role for 5hmC in regulating magnitude of transcriptional changes with age.

Previous studies have shown that ROS can promote conversion 5mC to 5hmC [17], and we thus also wondered whether age-related oxidative stress might also contribute to the accumulation of 5hmC in aged liver. Indeed, dihydroethidium (DHE) staining of mouse liver sections revealed significantly higher ROS signal in old nuclei compared to young (**Supp. Fig. 5E & 5F**). To directly assess whether ROS influences genomic 5hmC, we non-specifically induced cellular ROS in HepG2 cells using an acute 2 h treatment with 600 µM hydrogen peroxide (H_2O_2), which was shown to be sufficient to increase ROS levels (**Supp. Fig. 5G**). We also sought to decrease H_2O_2-induced ROS by either 24 h sequential or co-treatment with a radical scavenger, N-acetylcysteine amide (NAC), shown to reduce H_2O_2-induced ROS at 2 mM concentration (**Supp. Fig. 5H**). We found that acute H_2O_2 treatment did not significantly increase 5hmC signal, and reducing ROS, via sequential or co-treatment with NAC, did not influence genomic 5hmC levels (**Supp. Fig. 5I**). Chronic 24 h treatment with H_2O_2 also had no effect on 5hmC production (**Supp. Fig. 5J**). Thus, in our hands, ROS itself was not sufficient to increase global 5hmC levels.

Since quiescence and senescence can influence metabolic and mitochondrial function independent of changes in 5hmC, we opted to assess the functional effects of 5hmC using vitamin C (ascorbic acid). Vitamin C can drive the increase of 5hmC by acting as a co-factor for the TET proteins and enhancing oxidation of 5mC to 5hmC [18, 19]. Indeed, we observed a positive association between vitamin C treatment and 5hmC production (**Fig. 5L**) in HepG2 cells as reported previously in other models [18, 62]. As with contact inhibition (**Fig. 5C-E**), when we increased 5hmC by vitamin C, we observed a significant decrease in ATP production (**Fig. 5M**) but no significant differences in mitochondrial membrane potential (**Fig. 5N**). Overall, our data suggest that age-related contexts such as quiescence and senescence, or vitamin C treatment, but not ROS, can increase 5hmC and downregulate tissue-specific function.

5hmC's transcriptionally restrictive function extends to mouse cerebellum

5hmC has been reported to be highly abundant in brain tissue [63], corroborated in our data (**Fig. 1A**). We wondered whether the transcriptionally restrictive function for 5hmC observed in the liver (**Fig. 3D & 3E**) extended to the brain, even though brain regions did not show a significant global increase of 5hmC with age (**Fig. 1A**). We profiled 5hmC genome-wide using hMeDIP-seq and generated corresponding RNA-seq data for the cerebellum from young and old C57BL/6JN mice of both sexes (n=4 biological replicates per group, **Supp. Table 1**, QC metrics in **Supp. Fig. 6A-F**). In PCA, PC1 accounted for 64.8% of the variation in genome-wide 5hmC signal and clustered the samples by age (**Supp. Fig. 7A**). We called a greater mean number of peaks in old samples (mean = 145,112.3) than young (mean = 142,966.8) using MACS2 (q = 0.001), however this difference was not statistically significant (**Supp. Fig. 7B**). These peaks tended to be broader, with higher tag density, and higher mean signal in the old samples (**Supp. Fig. 7C**). A correlation heatmap using peak calls also clustered the samples by age (**Supp. Fig. 7D**), although the differences were not as sharp as with the liver (**Fig. 2D**).

Differential analysis of old vs young resulted in only 60 peaks surviving FDR < 0.05 (58 with higher enrichment in old and 2 with lower enrichment). To enable downstream comparison with the liver hMeDIP-seq and MeDIP-seq data, we relaxed the threshold for statistical significance and classified cerebellum DHMRs using $p < 0.05$, as previously done for the liver MeDIP-seq data (**Supp. Fig. 3E**). We obtained 11,703 total DHMRs in the cerebellum, 9,359 with greater enrichment in old ($\log_2$ FC > 0, $p < 0.05$) and 2,344 with greater enrichment in young ($\log_2$ FC < 0, $p < 0.05$) (**Supp. Fig. 7E, Supp. Table 10**). Verification of the 5hmC signal at the center of the hyper DHMRs showed reproducibly higher signal for old but not young, and vice versa for the hypo DHMRs (**Supp. Fig. 7F**). Example genome browser views of hyper and hypo DHMRs in the cerebellum are also shown for individual replicates in **Supp. Fig. 7G.** In accord with the idea that age-related differences in 5hmC primarily occur at regions associated with tissue-specific function (**Fig. 2H**), we observed that the cerebellum DHMRs were enriched for GO terms relating to cerebellum (i.e., "neuromuscular process controlling balance") and brain function (i.e., "hindbrain morphogenesis" and "synaptic processes") (**Supp. Fig. 7H**). Annotations of the cerebellum DHMRs were also comparable with the liver DHMRs and DMRs, wherein

most were associated with interCGIs and gene bodies, primarily intronic regions (**Supp. Fig. 7I**). In contrast to the liver, 5hmC signal at the gene bodies of all mm10 genes trended towards an increase with age, but the differences were not statistically significant (**Supp. Fig. 7J**).

Consistent with previous literature [30-33, 35], we observed a positive association between 5hmC and gene expression for both young and old age groups (**Supp. Fig. 8A**). We then traced the gene body 5hmC signal over the bottom or top 33% of genes with low and high magnitude of transcriptional changes with age (**Supp. Table 10**). We found that genes that remain relatively unchanged in expression with age had significantly higher 5hmC signal in the old, whereas genes that underwent dramatic transcriptional changes with age had no significant differences in 5hmC (**Supp. Fig. 8B**). Similar to the liver, genes that underwent dramatic transcriptional changes with age tended to harbor significantly shorter 5'UTRs, 3'UTRs, transcript length, CDS length, and had fewer exons compared to the genes that underwent low transcriptional changes with age (**Supp. Fig. 8C**). As examples of the age-related transcriptionally restrictive function for 5hmC, **Supp. Fig. 8D** shows cerebellum gene body 5hmC enrichment and corresponding magnitude of mRNA fold change with age for the cerebellum DHMR-associated GO terms "establishment of vesicle localization" and "neuromuscular process controlling balance".

Collectively, these results reiterate that 5hmC undergoes genome-wide changes at gene bodies and regions associated with tissue-specific function and serves a common transcriptionally restrictive function in mouse tissues during aging.

Human tissues also show 5hmC-mediated transcriptional restriction

We next investigated whether 5hmC's transcription restriction function extended to human tissues—thus indicating a conserved epigenetic regulation in aging tissues. Accordingly, we mined publicly available human RNA-seq data from post-mortem human tissues, brain (cortex), heart (left ventricle), and liver, from the Genotype-Tissue Expression (GTEx) project [64] and two published human 5hmC datasets (hmC-CATCH-seq [31] and 5hmC-Seal [30]) for the corresponding tissues (**Fig. 6A**). Tissue-specific genes were obtained from the Human Protein Atlas, defined as genes with at least four-fold higher expression in the tissue compared to any other tissues ("tissue enriched") or the average of all other tissues ("tissue enhanced") [65]. We hypothesized

that genes that undergo a relatively higher magnitude of transcriptional change with age (regardless of direction) would be marked by relatively lower levels of 5hmC, whereas tissue-specific genes (i.e., a combination of tissue enriched and tissue enhanced), which are typically enriched for 5hmC, would undergo relatively lower or no detectable transcriptional changes with age.

To mitigate possibilities of disease-related epigenetic and transcriptional changes, we filtered out donors from the GTEx datasets that experienced a slow death, defined by death after a long illness with more than 1 day of a terminal phase. Sample size distributions by age groups and sex for the GTEx tissues used in this study are shown in **Supp. Fig. 9A**. Overall, we included RNA-seq data from 193 brain donors, 393 heart donors, and 207 liver donors. Most donors or each tissue type were male.

To identify genes that undergo significant transcriptional changes during aging, we employed ImpulseDE2 [66], an R Bioconductor package designed to model longitudinal datasets, i.e., RNA-seq, and capture permanent and temporal gene expression trajectories. Sex was included as a covariate in all models. For the brain, ImpulseDE2 identified 5,625 genes that underwent significant monotonous expression trajectories with age (5,574 downregulated and 51 upregulated) and one gene that was transiently upregulated with age (**Supp. Fig. 9B**, left panel, **Supp. Table 11**). Given our interest in the magnitude of transcriptional changes with age, rather than direction, we classified all the age-related permanently and transiently changing genes as “brain-differential” genes. In contrast to the brain-differential genes, “brain-specific” genes did not undergo significant transcriptional changes with age (**Supp. Fig. 9B**, middle panel). As expected, GO analysis of the brain-enriched genes revealed biological processes associated with nervous system and synaptic transmission, while the differential genes were enriched for catabolic processes and transport (**Supp. Fig. 9B**, right panel).

In heart tissue, ImpulseDE2 identified 8,163 genes that underwent significant monotonic expression changes with age (8,078 downregulated, 85 upregulated) and 35 genes with transient expression changes (33 upregulated, 2 downregulated) (**Supp. Fig. 9C**, left panel). We detected no significant transcriptional changes in heart-specific genes with age (**Supp. Fig. 9C**, middle panel). GO analysis showed that the heart-specific genes were enriched for biological processes associated with cardiovascular function, while the heart-

differential genes were associated with protein transport, transcriptional and chromatin regulation, and mRNA splicing (**Supp. Fig. 9C**, right panel).

Lastly, in liver tissue, ImpulseDE2 identified 1,852 genes that underwent significant monotonous expression changes with age (1,789 downregulated, 63 upregulated) and 3 genes that underwent transient expression changes (1 upregulated, 2 downregulated). The liver-specific genes also showed no significant changes in expression with age (**Supp. Fig. 9D**, middle panel) and were associated with metabolic processes, while the liver-differential genes were associated with cell division, migration, and shape (**Supp. Fig. 9D**, right panel).

We next sought to assess gene body 5hmC levels of the tissue-specific and tissue-differential genes. Using human 5hmC data obtained from He et al [31], we observed that genes with significant transcriptional changes in the brain during aging had relatively lower gene body 5hmC signal compared to the tissue-specific genes, with the exception of one sample (**Fig. 6B**). In heart and liver tissue, we detected significantly higher 5hmC levels in the tissue-specific genes compared to the differential genes (**Fig. 6C & 6D**). Human 5hmC data from Cui et al [30] only showed significant differences between brain-specific and brain-differential genes for one sample (**Fig. 6E**). However, all heart and liver samples showed significantly higher 5hmC levels in the tissue-specific genes compared to the differential genes, in agreement with data from He et al [31] (**Fig. 6F & 6G**). Collectively, these results suggest that 5hmC may be a common regulator of transcriptional restriction in aged mammalian tissues, with significant differences in the heart and liver, and to a lesser extent, in the brain.

5hmC is downregulated in response to high-fat diet and disulfiram

Finally, given the potential implications for 5hmC enrichment at regions associated with metabolic processes in the aged liver, we investigated whether levels of this modification were detrimental under stress conditions known to promote metabolic dysfunctions, such as high-fat diet (HFD) consumption. Using C57BL/6JN mice of both sexes, we assessed the effects of HFD (60% kcal fat, n=6) compared to standard diet (SD, 10% kcal fat, n=6) on global 5hmC levels in the liver (**Supp. Fig. 10A)**. We observed an overall significant decrease of 5hmC signal in HFD livers (**Supp. Fig. 10B & 10C**); however, the decrease was strikingly sex-specific. Females on HFD had significantly lower 5hmC levels compared to SD, whereas males on HFD had no

significant differences in 5hmC levels (**Supp. Fig. 10D**). Several studies have reported on the sex-specific differences in metabolic response to HFD [67-69]. In general, females are typically more protected, compared to males, from the immediate adverse effects of HFD. Thus, considering 5hmC's role in restricting transcriptional changes with age, it is tempting to speculate that the downregulation of global 5hmC levels in the females might permit transcriptional flexibility to accommodate the HFD consumption.

We thus inquired whether treatments known to reduce HFD-induced metabolic dysfunction would also alter global 5hmC levels. Disulfiram (DSF), an FDA-approved drug used to treat alcoholism, was previously shown to harbor metabolic protective effects against HFD by reducing weight gain, liver steatosis, and promoting insulin responsiveness [70]. We assessed global 5hmC levels in liver gDNA of four different groups of C57BL/6JN mice of both sexes obtained from Bernier et al [70]. The mice were fed either a HFD (3 months) then switched to SD (3 months), HFD for the entire duration of the study (6 months), HFD with low dosage (100 mg/kg body weight/day) of DSF (HFD-L, 6 months), or HFD with high dosage (200 mg/kg body weight/day) of DSF (HFD-H, 6 months) (schematic, **Supp. Fig. 10E**). Overall, we did not observe significant differences in 5hmC between the SD switch group and the HFD group (**Supp. Fig. 10G**), although the latter tended to decrease 5hmC levels. Interestingly, however, HFD-L and HFD-H treatment groups had significantly lower 5hmC signal compared to SD switch and HFD groups. We further assessed whether these effects were sex-specific and observed a stronger dose-response relationship for the DSF as well as significantly lower 5hmC in the DSF groups compared to SD and HFD groups in the males (**Supp. Fig. 10G**, right panel). Females, however, had comparable levels of 5hmC among the HFD, HFD-L, and HFD-H treatment groups. Altogether, these results suggest that altered 5hmC levels may be a potential mechanism underlying sex-specific response to HFD and that downregulation of the modification may have long-term beneficial effects against HFD, possibly by promoting transcriptional flexibility.

Discussion

We have surveyed global 5mC/5hmC levels and found that 5hmC globally accumulates in multiple aged tissues without detectable differences in 5mC (**Fig. 1**). We performed genome-wide profiling of 5hmC in mouse

liver and cerebellum and identified specific increases in 5hmC at gene bodies associated with tissue-specific function (**Fig. 2** and **Supp. Fig. 6**), in agreement with previous studies showing that 5hmC marks tissue-specific genes [30, 31, 35]. Surprisingly, we observed that the age-related accumulation of 5hmC is not driven directly by differences in expression or enzymatic activity of TET proteins, but rather age-related contexts such as prolonged quiescence and senescence (**Fig. 5 & Supp. Fig. 5**). We further observed that regions gaining 5hmC with age were marked by 5mC in the young (**Supp. Fig. 3G**) consistent with a progressive conversion of 5mC to 5hmC over time.

It is unclear why age-related differences in 5hmC occur primarily at tissue-specific genes. A proposed mechanism for 5hmC's tissue-specific localization posits that tissue-specific TFs recruit TET proteins to tissue-specific genes, which then traverse the gene body alongside transcription elongation complexes to oxidize 5mC to 5hmC [71, 72]. We identified TFs from LISA analysis (**Supp. Fig. 1G**) that bind genes with high 5hmC signal and are known TET interactors. For example, PPARγ has been previously shown to interact with TET1 and direct the local increase of 5hmC levels [73]. Direct evidence for TET interactions with the other TFs is currently lacking, however, knockdown of ZIC2 has been shown to decrease the binding affinity of TET1 in primed epiblast [74]. Knockdown of Glis2 has also been shown to decrease 5hmC levels in human embryonic stem cells [75], while Tbx3 over-expression results in higher 5hmC levels [76]. These data suggest that 5hmC's enrichment in tissue-specific genes in the liver may be due to recruitment of TET proteins by TFs. Our data thus provides preliminary evidence for a proposed model wherein TET proteins interact with tissue-specific TFs, resulting in 5hmC localization to tissue-specific genes [71, 72] (**Supp. Fig. 1G**).

Given that a functional role for 5hmC in aging is largely unknown, our primary focus in this study was to dissect the role of 5hmC in transcriptional changes during aging. We found that although 5hmC is typically enriched in gene bodies of transcriptionally active and tissue-specific genes, the modification has an evolutionarily conserved function to restrict the magnitude of transcriptional changes during aging (**Fig. 3, 6, Supp. Fig. 1 & 8**).

5hmC is catalyzed from 5mC and we indeed observed opposing 5mC/5hmC levels at genes with stable gene expression changes during aging (**Fig. 3F & 3G**), however the changes in 5mC were modest and non-

significant, suggesting that changes in 5hmC may be a better predictor of age-related transcriptional changes. In support of our finding, *Tet1/3* double-deficient mice have been shown to exhibit lower transcriptional fidelity in early embryogenesis [77]. Suggestively, transcriptional dysregulation is a hallmark of cancer, an age-related disease with extensively documented TET mutations as well as loss of 5hmC [78-80]. Perhaps the function of 5hmC may be primarily protective, maintaining stable gene expression changes with age, but we also note that 5hmC would have to be downregulated to enable proper transcriptional responses under certain conditions such as HFD and drug treatment (**Supp. Fig. 10**). We therefore speculate that while the mark promotes transcriptional restriction during aging, it may also hinder the transcriptional flexibility necessary during stress.

We recall above that another epigenetic modification, H3K36me3, has been previously shown to restrict the magnitude of age-related transcriptional changes in two model organisms that generally lack 5hmC, *C. elegans* and *D. melanogaster* [47]. Both H3K36me3 and 5hmC are typically localized at gene bodies, though H3K36me3 levels have been reported to decline with age [81]. Possibly 5hmC may have evolved as an additional regulator of transcriptional restriction in more complex organisms.

We investigated 5hmC-interacting proteins and identified several factors involved in translation, transcription, and chromatin accessibility (**Fig. 4B & 4C, Supp. Fig. 4B & 4C**). Interestingly, we observed that 5hmC has lower binding affinity for splicing-associated factors (**Fig. 4D & 4E**) and is positively associated with age-related splicing events (**Fig. 4F & 4G**). Similarly, genes with a lower magnitude of transcriptional changes with age exhibited higher age-related splicing events, as shown by both short-read (**Fig. 4F**) and long-read sequencing (**Fig. 4G**). We speculate that 5hmC might exert its transcriptionally restrictive function during aging by repelling splicing-associated factors; future work aims to test this notion.

Lastly, we showed that increases in 5hmC are also associated with lower gene expression in old samples (**Fig. 3C & 3D, Supp. Fig. 1H**). We demonstrate this by leveraging 5hmC's tissue-specific localization and link 5hmC production to downregulation of ATP production in HepG2 cells (**Fig. 5C-E & 5L-N**).

Overall, our work elucidates an important and previously unrecognized function of a relatively new epigenetic modification, 5hmC, in aging. Aging is generally associated with global decline of cellular and organ function and accompanied transcriptional changes. 5hmC supports homeostasis in an aged environment by

imposing transcription restriction of tissue-specific genes. However, prolonged periods of restriction may eventually downregulate tissue-specific functions (**Fig. 7**). Furthermore, increased 5hmC can be detrimental in the face of stress and needs to be reduced to respond to environmental cues and in turn allow proper transcriptional responses.

Figure Titles and Legends

Figure 1: Global accumulation of 5hmC in aging heart, lung, and liver

(**A**) Schematic of experimental procedure and tissues used to profile global levels of 5mC and 5hmC by DNA mass spectrometry. Barplots show mean ± standard error of mean. * = $p < 0.05$ by unpaired t-test. (**B**) Dot blot for 5hmC using increasing amounts of young and old liver gDNA with signal quantifications on the right. + control is 200 ng of young mouse hippocampus gDNA and – control is water. Data are summarized as mean ± standard deviation. ** = q < 0.01 and *** = q < 0.001 by multiple unpaired t-test with FDR correction (two-stage step-up (Benjamini, Krieger, and Yekutieli [82]). (**C**) Immunofluorescence microscopy for 5hmC using young and old sex-matched liver sections with mean 5hmC signal intensity per nuclei quantifications on the right. *** = $p <$ 0.001 by Welch's t-test.

Figure 2: **5hmC accumulates at genic regions associated with hepatic metabolism during aging**

(**A**) Principal component analysis (PCA) plot obtained using input subtracted 5hmC bigWig files. (**B**) Number of 5hmC peaks called for young and old liver samples by MACS2 at q = 0.001. Data are summarized as mean ± standard deviation. * = $p < 0.05$ by Welch's t-test. (**C**) Violin plots (with boxplots) of peak characteristics size in bp (left), sum (middle), and mean (right) of signal. For $\log_2$(size) mean, young = 8.76, old = 8.98; $\log_2$(sum), young = 13.14, old = 13.22; $\log_2$(mean), young = 4.37, old = 4.24. *** = $p < 0.001$ by Welch's t-test. (**D**) Correlation heatmap of 5hmC peaks called by MACS2 produced using DiffBind. (**E**) Volcano plot of differentially hydroxymethylated regions (DHMRs) in the liver identified by DiffBind using FDR < 0.05. Hypo DHMRs (fold change ≤ -2) are peaks with less enrichment in the old and hyper DHMRs (fold change ≥ 2) are peaks with higher enrichment in the old. (**F**) Metaplots of 5hmC signal at the DHMRs. (**G**) Example genome browser tracks for liver hyper DHMRs (*Rbm47* and *Noct*) and hypo DHMRs (*Tsku* and *Klf15*). (**H**) Gene

ontology (GO) terms associated with the DHMRs from (E) using GREAT. The top 5 biological process terms with FDR < 0.05 are shown. (**I**) Pie charts showing CpG and genic/intergenic annotations of the DHMRs from (E). (**J**) Metaplots of 5hmC signal over the gene bodies of all mm10 genes with quantifications for each replicate on the side. *** = $p < 0.001$ by Welch's t-test.

Figure 3: Gene body 5hmC restricts the magnitude of transcriptional changes during aging

(**A**) Metaplots of 5hmC signal over the gene bodies of genes with low (n = 6,340), intermediate (n = 6,339), and high (n= 6,339) average mRNA counts for young (left) and old (right) samples. (**B**) Correlation between young (left) and old (right) gene body 5hmC signal ranking and variance in mRNA fold change (old vs young). (**C**) Boxplots showing mRNA fold change (old vs young) for genes in the low and high "mRNA fold change with age" category. (**D**) Metaplots of liver 5hmC signal over bodies of genes in (C) with low (left) and high (right) mRNA fold change with age. Quantifications are depicted below. ** = p < 0.01 and *** = p < 0.001 by Welch's t-test. (**E**) Same as (C) except for genes in the low and high "|mRNA fold change| with age" category. (**F**) Same as (D) except metaplots of liver 5hmC signal over bodies of genes in (E). (**G**) Metaplots of liver 5mC signal over gene bodies of genes in (E) with low (left) and high (right) |mRNA fold change| with age. Quantifications are depicted below. N/S = $p \geq 0.05$ by Welch's t-test. (**H**) Boxplots showing distribution of 5'UTR, 3'UTR, transcript, and CDS length as well number of exons for the for the genes in (D) with low (left) and high (right) |mRNA fold change| with age. *** = $p < 0.001$ by Welch's t-test. (**I**) Heatmaps of two GO terms from Fig. 2H showing gene body 5hmC signal and corresponding |mRNA fold change (old vs young)|. Heatmap is sorted by decrease in magnitude of mRNA fold change between old and young.

Figure 4: Alternative splicing mediates 5hmC's transcriptionally restrictive function through decreased binding of splicing factors

(**A**) Schematic of oligo mass-spec experimental procedure. (**B**) Volcano plots showing differential proteins identified between 5hmC oligo 1 pull-down in the old and input. Significant interactors are colored by some of the top prominent enriched terms, chromatin, metabolism, splicing, and translation. (**C**) Volcano plots showing differential proteins identified between 5hmC oligo 2 pull-down in the old and input. (**D**) GO terms associated with proteins depleted or enriched for oligo 1 in the old 5hmC vs young C and mC comparison. (**E**) GO terms

associated with proteins depleted or enriched for oligo 2 in the old 5hmC vs old C and mC comparison (top panel) and the old 5hmC vs young C and mC comparison (bottom panel). (**F**) Number of differential splicing events detected in RNA-seq data between old and young samples (top panel). Number of events and the unique number of genes are indicated. Number of differential splicing events grouped by increasing gene body 5hmC signal in old (middle panel). Number of differential splicing events grouped by increasing magnitude of mRNA fold change with age (bottom panel). (**G**) Bar plots showing differential isoform usage for genes with low or high gene body 5hmC levels or low and high mRNA absolute fold change at different *p*-value thresholds. (**H**) Transcript length (top) and poly A length (bottom) distribution for genes that undergo low or high mRNA absolute fold change with age. *** = $p < 0.001$ by Mann-Whitney test.

Figure 5: Quiescence and senescence drive the increase of 5hmC with age and impact cellular function

(**A**) Normalized mRNA counts in young and old liver samples for enzymes that can potentially modulate 5hmC levels. (**B**) TET activity assay in young and old liver lysate. (**C**) Dot blot for 5hmC signal in contact inhibition-induced quiescent HepG2 cells. + control is 200 ng of young mouse hippocampus gDNA and – control is water. Quantifications are depicted below. ** = $p < 0.01$, *** = $p < 0.001$ by two-way ANOVA with Tukey's multiple comparisons test. (**D**) ATP production assay using 14-day contact inhibited quiescent HepG2 cells. Data are summarized as mean ± standard deviation from 3 biological replicates with 3 technical replicates each. *** = $p < 0.001$ by Welch's t-test. (**E**) TMRM mitoprobe assay with proliferating cells and cells that have been contact-inhibited for 14 days. Data are summarized as mean ± standard deviation. (**F**) Dot blot of 5hmC using proliferating and irradiation-, etoposide-, and oxidative stress-induced senescent cells. + control is 50 ng of young mouse hippocampus gDNA and – control is water. Quantifications are shown below. ** = $p < 0.01$, *** = $p < 0.001$ by two-way ANOVA with Tukey's multiple comparisons test. (**G**) Schematic for 70% partial hepatectomy procedure (**H**) Relative 5hmC signal for young and old liver samples at the indicated timepoints post-surgery. Data are summarized as mean ± standard deviation. ** represents significance ($p < 0.01$) of the interaction effect of age x timepoints on 5hmC levels relative to old pre-surgery group. (**I**) Boxplot showing the transcriptional amplitudes before and after regeneration for the group of genes that showed low mRNA absolute fold change with age. (**J**) Same as (I) except for the group of genes that showed high mRNA absolute

fold change with age. (**K**) Same as (I) and (J) except for all genes. For (I-K), *** = p < 0.001 by Mann-Whitney test. (**L**) Dot blot for 5hmC using gDNA isolated from HepG2 cells treated with 0-300 µM vitamin C for 24 h. + control represent 200 ng of young mouse hippocampus gDNA and – control is water. Quantifications are depicted below. ** = *p* <0.01 and *** = *p* < 0.001 by two-way ANOVA with Tukey's multiple comparisons test. (**M**) ATP production assay using HepG2 cells treated with 150 uM vitamin C for 24 h. Data are summarized as mean ± standard deviation from 3 biological replicates with 3 technical replicates each. ** = *p* <0.01 by Welch's t-test. (**N**) TMRM mitoprobe assay with control cells or cells treated with 150 µM vitamin C for 24 h. For E and N, an unpaired t-test yielded no significant results.

Figure 6: Human tissues also show 5hmC-mediated transcriptional restriction

(**A**) Schematic outline of procedure, age groups, and tissues chosen from the GTEx project. (**B**) Metaplots of 5hmC signal from He et al. [31] over gene bodies of brain-differential and brain-specific genes, (**C**) heart-differential and heart-specific genes, and (**D**) liver-differential and liver-specific genes. ** = *p* < 0.01 and *** = *p* < 0.001 by ANOVA. (**E**) Metaplots of 5hmC signal from Cui et al. [30] over gene bodies of brain-differential and brain-specific genes, (**F**) heart-differential and heart-specific genes, and (**G**) liver-differential and liver-specific genes. ** = *p* < 0.01 and *** = *p* < 0.001 by ANOVA.

Figure 7: Overall model

Model showing the transcription restriction function and propensity to downregulate tissue-specific function of 5hmC during aging.

Methods

Cell lines and culture conditions

HepG2 (ATCC, human male) cells were cultured in a 37°C 5% CO_2 and 20% O_2 humidified incubator with Dulbecco's Modified Eagles Medium (DMEM, Gibco) supplemented with 10% Fetal Bovine Serum (FBS, Thermo Fisher) and 1% penicillin/streptomycin (Penn/Strep, Thermo Fisher). WI-38 (Coriell Institute, human female) cells were cultured in a 37°C 5% CO_2 and 20% O_2 humidified incubator with DMEM (Gibco) supplemented with 10% heat-inactivated FBS (Gibco), 0.5% Penn/Strep (Gibco), sodium pyruvate (Gibco), and

non-essential amino acids (Gibco).

Animals

This study was approved by the Animal Care and Use Committee of the NIA in Baltimore, MD under Animal Study Protocol number 481-LGG-2022 (all except disulfiram experiments) and 444-TGB-2016 (disulfiram experiments). Young and old inbred C57BL6/JN mice of both sexes were acquired from the NIA aged rodent colony (https://ros.nia.nih.gov/) and housed in rooms that were maintained at 22.2 ± 1 °C and 30-70% humidity. For disulfiram experiments, C57BL/6J mice (Jackson Laboratory stock #000664) were single housed under temperature-controlled conditions with 12 h light/12 h dark cycle with ad libitum access to house chow (2018 Teklad Global 18% Protein Rodent Diet 2018S, Harlan Teklad) and water. Routine tests were performed to ensure that mice are pathogen-free and sentinel cages maintained and tested according to American Association for Accreditation of Laboratory Animal Care (AAALAC) criteria. The age and sex information are available in Supp. Table 1.

High fat diet and disulfiram treatment

For mice on high fat diet, two regimens were used. In the first regimen, the diet was applied early in life and for a short duration. Three C57BL/6J male mice (Jackson Laboratory stock #000664) and three C57BL/6J female mice (Jackson Laboratory stock #000664) at 6 weeks of age were fed with Research Diets, Inc. D12492i (60 kcal% fat) diet up to 11 weeks of age. An equal number of control mice of the same genotype and sex were fed with Research Diets, Inc. D12450Bi (10 kcal% fat) up to 11 weeks of age. Mice were then sacrificed, and their livers harvested. In the second regimen, diet was applied later and for longer duration. Beginning at 9 months of age, mice were maintained on a high fat diet consisting of AIN-93G modified to provide 60% of calories from fat (HFD; carbohydrate:protein:fat ratio of 16:23:61) for the next 3 months, after which animals were randomly divided into four groups (n = 9, 6 males and 3 females). Group 1, was continued on HFD diet; Group 2, was switched to standard AIN-93G diet (SD, carbohydrate:protein:fat ratio of 64:19:17), Group 3 was fed HFD supplemented with a low dose of disulfiram (100 mg/kg body weight/day DSF; HFDL); and Group 4 was fed HFD supplemented with a high dose of disulfiram (200 mg/kg body weight/day DSF; HFDH). All animals were sacrificed after an additional 3 months of treatment and livers were collected for further analysis. DSF (Sigma-

Aldrich, St-Louis, MO) was included in HFD-modified AIN-93G diet (Dyets, Inc., Bethlehem, PA) at a concentration of 2.33 g/kg (low dose, HFDL) and 2.67 g/kg (high dose, HFDH), respectively.

Partial hepatectomy surgery

70% partial hepatectomy was performed in accordance with the guidelines from Mitchell et al. [83] and previously reported elsewhere (Yang et al. *Molecular Cell*, under revision). Briefly, liver lobes were removed and labeled as “pre-surgery”. After indicated post-surgery time, animals were sacrificed by carbon dioxide asphyxiation and cervical location. The liver was dissected and frozen in isopentane chilled with liquid nitrogen and stored in -80°C.

Induction of quiescence

To induce quiescence of HepG2 cells by serum starvation, ~70% confluent plates of HepG2 cells were changed to DMEM without FBS for 1-3 days. To induce quiescence by contact inhibition, cells were allowed to grow until they reached 100% confluency. The cultures were maintained for 2, 7 or 14 days with media change (with serum) every two days. Proliferating controls were included for both serum starvation and contact inhibition experiments.

Induction of senescence

WI-38 cells were maintained at low population doubling (PD) levels for proliferating conditions. Cellular senescence was triggered by different means. Etoposide-induced senescence (ETIS) was achieved by culturing for 10 days in the presence of etoposide (Selleckchem) at 50 µM, with medium refreshed every 3 days. Ionizing (γ) radiation-induced senescence (IRIS) was achieved by exposing cells to 15 Gray (Gy) followed by culturing for 10 days. Oxidative stress-induced senescence (OSIS) was achieved by adding 0.75 mM H_2O_2 directly to cells in complete medium and replacing with fresh medium 2 h later.

RT-qPCR for senescent cells

Cells were lysed in either Tri-Reagent (Invitrogen) or RLT buffer (Qiagen), and the lysate was processed with the QIAcube (Qiagen) to purify total RNA, which was then reverse-transcribed (RT) to create cDNA using Maxima reverse transcriptase (Thermo Fisher Scientific) and random hexamers. Real-time, quantitative (q)PCR analysis was then performed using SYBR Green mix (Kapa Biosystems), and the relative expression

was determined by the 2-ΔΔCt method on a QuantStudio 6 qPCR machine (Thermo Fisher). The levels of mRNAs were normalized to human *ACTB*.

BrdU assay

~125,000 cells were seeded in 6-well plates and incubated with BrdU diluted in DMEM with 10% FBS for 24 hours. BrdU incorporation was measured following the manufacturer's protocol (Cell Signaling Technology). Briefly, cells were fixed and denatured before the addition of anti-BrdU mouse monoclonal antibody. BrdU incorporation was detected by measuring absorbance at 450 nm using a GloMax plate reader (Promega).

Induction and neutralization of ROS in cell culture

To induce acute ROS, ~1 x 10^6 HepG2 cells were treated with 600 µM H_2O_2 (Sigma) in serum-free DMEM for 2 h then changed into fresh media (without serum) for 24 h. To neutralize ROS, HepG2 cells were either co-treated or sequentially treated with H_2O_2 and N-acetylcysteine amide (NAC). For co-treatment, cells were first treated with 600 µM H_2O_2 and 2 mM NAC in serum-free DMEM for 2 h then washed twice and incubated with 2 mM NAC in serum-free DMEM for 24 h. Sequential treatment was performed by first treating cells with 600 µM H_2O_2 in serum-free DMEM for 2 h then washed twice and incubated with 2 mM NAC in serum-free DMEM for 24 h. DHE staining was also performed in parallel for the ROS neutralization experiments, using separate wells, as described in the "ROS detection in cells and tissue" section. To induce chronic ROS, ~1 x 10^6 HepG2 cells were treated with 20 µM H_2O_2 in serum-free DMEM for 24 h. As control, cells were kept in serum-free DMEM with DMSO for the duration of treatment.

Treatment with Vitamin C

To induce production of 5hmC by vitamin C, ~1 x 10^6 HepG2 cells were treated with vitamin C (Sodium L-ascorbate, Sigma) at concentrations of 30, 75, 150, and 300 µM in DMEM with serum for 24 h.

Genomic DNA isolation for dot blots

gDNA was isolated from ~25 mg of frozen tissue or ~1 x 10^6 HepG2 cells suspended in 200 µL of PBS using the Quick-DNA Miniprep Plus kit (Zymo Research) following the manufacturer's protocol with an overnight (for frozen tissue) or 10 min (cells) proteinase K digestion. The amount and quality of extracted DNA was assessed using Qubit HS assay kit (Thermo Fisher) and NanoDrop One (Thermo Fisher).

Quantification of cytosine modification by mass spec

gDNA was isolated from ~25 mg tissue following instructions on the Quick-DNA Miniprep Plus Kit using an overnight proteinase K digestion. The amount of DNA was quantified using a Nanodrop and samples were verified to have a 260/280 of >1.8 and a 230/260 of ≥2. ~2.5ug of DNA in a volume of 130 µL was sheared to ~800bp using a S220 focused ultrasonicator (Covaris) and the following parameters: peak incident power 105, duty factor 5%, cycles per burst 200, treatment time: 50s. The shearing was verified by phenol-chloroform purification followed by ethanol precipitation and running on a 1% agarose gel.

Cytosine methylation and hydroxymethylation were quantified using a protocol modified from Sun et al. [84, 85]. Briefly, DNA was digested into single nucleosides by using the Nucleoside Digestion Mix (New England Biolabs) enzyme cocktail at 37°C for 2 h. 5mC was identified and quantified by nano-liquid chromatography coupled online with tandem mass spectrometry (nLC-MS/MS). Using a Dionex RSLC Ultimate 3000 (Thermo Scientific, San Jose, CA, USA), nLC was configured with a 300 µm ID x 0.5 cm C18 trap column (Dionex, Thermo Scientific) and a 75 µm ID x 25 cm Reprosil-Pur C18-AQ (3 µm; Dr. Maisch GmbH, Germany) analytical nano-column were used to identify and quantify the absolute value of 5mC and 5hmC as shown in **Fig. 1A**. nLC was configured with a two-column system consisting of a 75 µm ID x 1 cm poly-graphitic carbon resin (PGC, HyperCarb, Thermo Scientific) trap column and a 75 µm ID x 25 cm PGC analytical nano-column for relative 5hmC comparison between young and old mice in liver samples shown in **Fig. 5G**. Except for the C18 trap column (cartridge from Thermo Scientific), all other columns were packed in-house. nLC was coupled online to an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific). The spray voltage was set to 2.3 kV and the temperature of the heated capillary was set to 275 °C. PGC columns setup was used for relative 5mC and 5hmC analysis. The full scan range of 110−1200 m/z was acquired in the Orbitrap at a resolution 120,000. Targeted scans were performed for MS/MS fragmentation using an HCD energy of 30 V and acquired in the Orbitrap at a resolution of 7,500. To accurately quantitate the absolute value, C18 columns are used, and the instrument was optimized to fragment the protonated nucleosides deoxycytidine (dC), deoxy-methylcytidine (dmC) and deoxy- hydroxymethylcytosine (dhmC) into protonated nucleobases C, 5mC, and 5hmC with m/z at 112.0505, 126.0662, and 142.0611, respectively. The full scan range was 110−600 m/z

acquired in the Orbitrap at a resolution 120,000. The source fragmentation energy was set at 30 V and RF lens % was set at 50, which gives more than 90% generation of nucleobases. To accurately quantify 5mC and 5hmC, a calibration curve of 5mC% and 5hmC% were constructed by analyzing samples with varying amount of 5mC and 5hmC standard in the presence of constant C standard and were used to correct the observed 5mC% and 5hmC% from the real samples. Quantification was obtained by extracting the ion chromatograms of C, 5mC and 5hmC using Skyline software, respectively. 5mC and 5hmC levels were calculated by dividing the area under the curve of the given species by the total area of all (un)modified C quantified.

Immunofluorescence

Fresh liver tissues were fixed with 4% paraformaldehyde and immersed in 20% sucrose solution at 4°C overnight. The tissues were embedded in OCT compound, frozen at -80°C, and then cut into 12 µm sections onto positively charged slides in a cryostat chamber. The sections were permeabilized with 0.2% Triton X-100 in Tris Buffered Saline (TBS) for 5 min at room temperature. Antigen retrieval steps were performed based on previous publication [86]. Briefly, sections were treated with 2 N hydrochloric (HCl) acid in PBS for 30 min in a 37°C incubator. After denaturation, sections were neutralized in two successive rounds with 0.1 M Tris-HCl (pH: 7.5) in PBS for 5 min. The sections were blocked for 1 h at room temperature with 2% normal goat serum (Vector Biolabs) then incubated with 2 µg/mL dilution 5hmC antibody (Active Motif, 39092) overnight at 4°C in a humidified chamber. After three rounds of washes with TBS (supplemented with 0.1% Tween-20, TBST; Pierce), the sections were incubated with a secondary antibody conjugated to a fluorescent dye (Thermo Fisher, A-11008) for 1 h at room temperature. Lastly, sections were stained with 5 µg/mL DAPI in TBS for 1 h at room temperature in a humidified chamber. Following washes with TBS, the sections were mounted with Epredia Lab Vision PermaFluor Aqueous Mounting Medium (Fisher Scientific) then photographed using a Zeiss LSM 710 confocal microscope. Intensities were quantified using ImageJ v1.51g [87].

Dot blot

The dot blot assay was adapted from a previous publication [15]. 2 µL of gDNA (containing indicated amounts of DNA) was denatured with 0.5 N NaOH for 15 minutes then spotted on a nitrocellulose blotting membrane (GE Healthcare) and cross-linked at 120,000 microjoules for 20 min using a Stratalinker® UV Crosslinker 1800

(Stratagene). As a negative control, we included 2 µL of water. Unless otherwise specified, 2 µL containing indicated amounts of gDNA isolated from a young mouse hippocampus was used as positive control. The membrane was then blocked in 5% skimmed milk in TBS containing 0.1% Tween 20 (TBST) for 1 h at room temperature followed by incubation with 1:10,000 dilution of 5hmC antibody (Active Motif, 39069) overnight at 4°C. After three rounds of washes, the membrane was incubated with 1:10,000 dilution of HRP-conjugated anti-rabbit IgG (BioRad, 1706515) for 1 h at room temperature. Digital ECL substrate solution (Kindle Biosciences) was added to the membrane before detection with a ChemiDoc MP Imaging System (BioRad). The dot blot intensity was quantified using ImageJ v1.51g [87].

ROS detection in tissue sections

Fresh liver tissues were embedded in OCT compound without fixation, frozen at -80°C, then cut into 12 µm sections onto positively charged microscope slides (Fisher Scientific) in a cryostat chamber. Slides were rinsed with H_2O for 30s then incubated with 50 µM dihydroethidium staining solution (DHE, Thermo Fisher) in 1x PBS at 37°C for 30 min in a dark humidified chamber. Slides were washed twice then stained with 5 µg/mL DAPI in 1x PBS for 30 min at room temperature. The sections were mounted with Epredia Lab Vision PermaFluor Aqueous Mounting Medium (Fisher Scientific) then photographed using a Zeiss LSM 710 confocal microscope. Intensities were quantified using ImageJ v1.51g [87].

For the H_2O_2 treatments, ~70-80% confluent HepG2 cells in µ-Slides (Ibidi, 80826) were treated with H_2O_2 (Sigma) in serum-free DMEM for 2 h at concentrations of 0, 10, 50, 100, 200, and 600 µM. Cells were washed twice with PBS and then incubated with 8 µM of DHE (Thermo Fisher) in serum-free DMEM for 15 min at 37°C. After washes, cells were incubated with DAPI (1:1000) at 37°C for 30 min and then photographed using a Zeiss LSM 710 confocal microscope.

ATP assay

Proliferating, quiescent (contact-inhibited for 14 days) or H_2O_2 treated HepG2 cells were trypsinized, counted, and 1 million cells were pelleted. The cell pellets were resuspended in 350 µL DMEM medium, and 100 µL was distributed to 3 wells of a 96-well plate for each sample. 100 µL of CellTiter-Glo reagent (Promega) was then added to each well containing cell suspension. The contents were mixed for 2 minutes on an orbital shaker to

induce cell lysis. The plate was then incubated at room temperature for 10 min to stabilize the luminescent signal and recordings were taken on a GloMax Discover System (Promega).

TMRM assay

After contact inhibition or treatment of HepG2 cells with vitamin C, cells were resuspended in 1 mL PBS at approximately 1 x 10^6 cells/ml. TMRM was added to a final concentration of 20 nM and incubated for 30 mins at 37 °C, 5% CO_2. For CCCP control samples, CCCP was added to a final concentration of 50 nM to the cells, incubated for 5 mins at 37 °C, 5% CO_2 and then treated with 20 nM TMRM reagent for 30 mins. Cells were analyzed on a BD Symphony flow cytometer with 561 nm excitation**.**

Nuclei preparation

Nuclei preparations were performed by douncing frozen liver tissue in nuclei preparation buffer containing 10 mM Tris-HCl (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$, 0.1% Tween 20, 0.1% NP-40, 0.01% digitonin, 1 mM BSA, and supplemented with 1x Halt protease and phosphatase inhibitor cocktail (Thermo Fisher) and 1 mM sodium butyrate. The resulting homogenate was filtered through a 30 µm cell strainer then washed using wash buffer containing 10 mM Tris-HCl (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$, 0.1% Tween 20, 1% BSA, and supplemented with 1x Halt protease and phosphatase inhibitor cocktail (Thermo Fisher) and 1 mM sodium butyrate to stop lysis. After centrifugation, the nuclei pellet was washed thrice in nuclei suspension buffer containing PBS, 2% BSA, 3 mM $MgCl_2$, and supplemented with 1x Halt protease and phosphatase inhibitor cocktail (Thermo Fisher) and 1 mM sodium butyrate.

TET activity assays

Nuclei preparations were performed as mentioned above. The nuclei pellets were lysed in nuclei lysis buffer containing 10 mM Tris-HCl (pH 7.4), 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% sodium-deoxycholate, 0.5% N-lauroylsarcosine, and supplemented with 1x Halt protease and phosphatase inhibitor cocktail (Thermo Fisher) and 1 mM sodium butyrate then sheared to <500 bp using a Covaris S220 Ultrasonicator (peak incident power 140, 200 cycles per burst, duty factor 5%, 10 min). The protein was quantified using the Pierce™ BCA Protein Assay Kit (Thermo Fisher) and ~10 µg total protein was used to measure TET activity using the TET Hydroxylase Activity Quantification Kit (Abcam) following the manufacturer's protocol.

DNA oligo pulldown mass spectrometry

Nuclei preparations were performed as mentioned above. As described in [29], the nuclei pellet was then lysed on ice for 90 minutes in 2 volumes of nuclei lysis buffer containing 420 mM NaCl, 20 mM HEPES, 20% v/v glycerol, 2 mM $MgCl_2$, 0.2 mM EDTA, 0.1% NP40, 0.5 mM DTT, and supplemented with 1x Halt protease and phosphatase inhibitor cocktail (Thermo Fisher) and 1 mM sodium butyrate. Prior to protein quantification, samples were pre-cleared using washed Dynabeads MyOne C-1 beads (Invitrogen) suspended in nuclei lysis buffer. DNA pull-downs were performed as described in [29] with 10 ug of DNA oligo for each pull-down (unmodified, methylated, and hydroxymethylated; GenScript) and 400 ug of nuclear extract. 10% lysate was saved as “input”.

Proteins were eluted from beads at room temperature with mixing for 20 min and at 65°C for 10 min on a thermomixer (800 RPM) in two successive rounds using 100 µL of biotin elution buffer (12.5 mM D-biotin, 7.5 mM HEPES pH 7.5, 75 mM NaCl, 1.5 mM EDTA, 0.15% SDS, 0.075% sarkosyl, and 0.02% sodium deoxycholate). The two eluents were pooled and precipitated overnight at 4°C with chilled trichloroacetic acid (25% v/v). Proteins were pelleted at 16,000g in 4°C for 30 min to remove supernatant and washed once with ice-cold acetone then pelleted again at 16,000g in 4°C for 5 min and air dried for 1 min. An S-Trap column cleanup was performed by loading the protein pellet resuspended in 165 µL of added loading buffer (90% methanol and 10 mM sodium bicarbonate pH 8.0) onto an S-Trap Micro Spin Column (Protifi). The column was washed twice with 150 µL of loading buffer and gently centrifuged before overnight digestion at 37°C with 0.1 µg/µL of trypsin (Promega) in 50 mM ammonium bicarbonate. After a gentle centrifuge, peptides were first eluted with 40 µL of 0.2% aqueous formic acid then eluted again with 35 µL of 50% acetonitrile containing 0.2% formic acid. Eluted peptides were dried in DNA120 SpeedVac (Thermo Fisher) with no heat.

Prior to mass spectrometry analysis, samples were desalted using a 96-well plate filter (Orochem) packed with 1 mg of Oasis HLB C-18 resin (Waters). Briefly, the samples were resuspended in 100 µl of 0.1% TFA and loaded onto the HLB resin, which was previously equilibrated using 100 µl of the same buffer. After washing with 100 µl of 0.1% trifluorosacetic acid (TFA), the samples were eluted with a buffer containing 70 µl of 60% acetonitrile and 0.1% TFA and then dried in a vacuum centrifuge.

Samples were resuspended in 10 µl of 0.1% TFA and loaded onto a Dionex RSLC Ultimate 300 (Thermo Scientific), coupled online with an Orbitrap Fusion Lumos (Thermo Scientific). Chromatographic separation was performed with a two-column system, consisting of a C-18 trap cartridge (300 µm ID, 5 mm length) and a picofrit analytical column (75 µm ID, 25 cm length) packed in-house with reversed-phase Repro-Sil Pur C18-AQ 3 µm resin. To analyze the proteome, peptides were separated using a 60 min gradient from 4-30% buffer B (buffer A: 0.1% formic acid, buffer B: 80% acetonitrile + 0.1% formic acid) at a flow rate of 300 nl/min. The mass spectrometer was set to acquire spectra in a data-dependent acquisition (DDA) mode. Briefly, the full MS scan was set to 300-1200 m/z in the orbitrap with a resolution of 120,000 (at 200 m/z) and an AGC target of 5x10e5. MS/MS was performed in the ion trap using the top speed mode (2 secs), an AGC target of 1x10e4 and an HCD collision energy of 35.

Proteome raw files were searched using Proteome Discoverer software (v2.4, Thermo Scientific) using SEQUEST search engine and the SwissProt mouse database. The search for total proteome included variable modification of N-terminal acetylation, and fixed modification of carbamidomethyl cysteine. Trypsin was specified as the digestive enzyme with up to 2 missed cleavages allowed. Mass tolerance was set to 10 pm for precursor ions and 0.2 Da for product ions. Peptide and protein false discovery rate was set to 1%. Following the search, data was processed as described by previously [88]. Briefly, proteins were $\log_2$ transformed, normalized by the average value of each sample and missing values were imputed using a normal distribution 2 standard deviations lower than the mean. Statistical regulation was assessed using heteroscedastic t-test (if p-value < 0.05). Data distribution was assumed to be normal but this was not formally tested.

RNA isolation, RT-qPCR, and RNA-sequencing

Liver RNA-seq data was previously generated and reported elsewhere in Yang et al. *Molecular Cell* (under revision). For cerebellum, RNA was isolated from frozen tissue by homogenization in Trizol followed by isopropanol precipitation. The RNA was further purified using RNeasy columns (Qiagen). An on-column DNase I digestion was performed during the purification step to remove gDNA. The RNA amount and integrity were confirmed using the Qubit RNA HS Assay Kit and RNA IQ Assay (Thermo Fisher) respectively. Total RNA (~700 ng) was used to make RNA-seq libraries following the Zymo-Seq Ribo-free Total RNA Library Kit (Zymo

Research) instructions with dual indexing. The pooled RNA-seq libraries were subjected to two rounds of 50 bp paired end sequencing on a NextSeq 550 platform using a P2 100-cycle kit (Illumina).

Direct RNA-seq with nanopore sequencing

RNA integrity was assessed using Qubit RNA IQ assay kit (Thermo Fisher) before library preparation. 25-50 µg of total RNA was used to make libraries using the direct RNA sequencing kit (Oxford Nanopore Technologies) as previously described with modifications [89]. Briefly, after selection of poly(A) RNAs using Oligo d(T)25 Magnetic Beads (New England Bioloabs), 15 pmoles of REL5 adapter (/5Bio/rArArUrGrArUrArCrGrGrCrGrArCrCrArCrCrGrArGrArUrCrUrArCrArCrUrCrUrUrUrCrCrCrUrArCrArCrGrArCrGrCrUrCrUrUrCrCrGrArUrCrU) was ligated to the 5′ ends of poly(A)-purified RNAs using T4 RNA ligase 1 (New England Bioloabs) for 3 hours at 37°C. 750 ng of REL-ligated poly(A) RNAs was used for library preparation according to manufacturer's protocol (Oxford Nanopore Technologies). Final libraries were quantified using Qubit 1X dsDNA High Sensitivity (HS) assay kit (Thermo Fisher) and sequenced on a MinION device using R9.4.1 flow cells (Oxford Nanopore Technologies).

Methyl and hydroxymethyl DNA immunoprecipitation sequencing (MeDIP-seq and hMeDIP-seq)

The hMeDIP assay was performed using the MagMeDIP-seq Package (Diagenode) following the manufacturer's protocol with a mouse antibody against 5hmC (Diagenode, C15200200-50) or 5mC (Diagenode, 33D3 clone, C15200081). Briefly, 1.2 µg of gDNA was sonicated into ~200 bp fragments using the S220 focused ultrasonicator (Covaris). To maximize IP yield for hMeDIP-seq, samples were processed in duplicates then pooled before IPure purification. Prior to immunoprecipitation, samples were spiked with hydroxymethylated, methylated and unmethylated internal DNA controls. IP efficiency and success was verified by qPCR targeting internal DNA controls. The DNA amount was quantified by Qubit HS DNA kit and the fragment size was assessed on a 2100 BioAnalyzer using a DNA HS kit (Agilent). Individual libraries for immunoprecipitated DNA and 10% input were dual indexed (NEBNext Multiplex Oligos, unique dual indices, New England Biolabs), PCR amplified and then pooled into equimolar amounts and further quantified using the NEBNext Library Quant Kit (New England Biolabs). The pooled library was subjected to 50 bp paired-end sequencing on the Illumina NextSeq 2000 platform.

Antibodies and oligos

All antibodies and oligos used in this study are listed in **Supp. Table 12**.

Bioinformatic analysis

hmeDIP-seq and meDIP-seq analysis

hMeDIP-seq and MeDIP-seq sequencing reads were de-multiplexed using bcl2fastq/2.20.0 and adapter trimmed using trimgalore/0.6.6. FastQ quality was assessed using FastQC/ 0.11.9 [90]. Reads were aligned to the mouse reference genome (assembly GRCm38/mm10) using bowtie/2-2.4.4 then filtered for a minimum mapping quality of 10 using samtools/1.9 [91]. BAM files were then sorted and additionally filtered for uniquely mapped and non-duplicate reads using sambamba/0.7.1 [92]. Encyclopedia of DNA Elements (ENCODE) blacklisted regions (mm10) [93] were filtered from BAM files using bedtools/2.30.0. RPKM (reads per kilobase per million mapped reads) normalized bigWig files were generated by first indexing BAM files using samtools/1.9 followed by conversion with the bamCoverage function of deeptools/3.5.0 [94]. 5hmC samples were input subtracted using the bigWigCompare function of deepTools/3.5.0. Broad and narrow peaks were called on 5hmC BAM files (with corresponding input as control) at a q-value of 0.001 using macs/2.2.7.1 [39]. Differentially enriched peaks were identified using DiffBind/3.0.15 [40] with the summits parameter set at 110. Random data was generated (**Fig. 3B**) by using the R seq function with ranges comparable to young and old gene body 5hmC signal (-24.74 to 58.90, increments of .01) and then randomly sampled with replacement using the R sample function.

RNA-seq analysis

Sequencing reads were processed as reported elsewhere (Yang et al. *Molecular Cell,* under revision). Briefly, reads were de-multiplexed using bcl2fastq/2.20.0, adapter trimmed using trimmomatic/0.39 [95], and then quality assessed using FastQC/0.11.9 [96]. Reads were aligned to the mouse reference genome, GRCm38/mm10, using STAR/2.7.5b [97] and the resulting BAM files were sorted and indexed using samtools/1.10 [98]. BAM files were filtered for duplicates and alignments with a minimum mapping quality of 10 using picard/2.20.8 and samtools/1.10 [98], respectively. Gene counts were estimated using the featureCounts function of the Rsubread R package/2.6.4 [99]. DESeq2/1.30.1 [100] was used to perform count normalization and differential gene

expression analysis between old and young samples. Genes were ranked and categorized according to DESeq2 normalized mRNA levels (averaged within groups) or absolute fold-change ratio |old vs young| using the ntile function of dplyr/1.0.7 (**Fig. 3B**).

To identify age-related alternative splicing events, reads were aligned to the mouse reference genome (GRCm38/mm10) using STAR/2.7.5b [97] with additional parameters (--outSAMstrandField intronMotif --outSAMattrIHstart 0 –alignSoftClipAtReferenceEnds No). rmats/4.1.1 [101] was employed on bam files with parameters --readLength 37 --variable-read-length and --cstat 0.0001.

Nanopore sequencing data analysis

Nanopore dRNA-seq data were basecalled using Guppy/6.1.2. Reads were subsequently mapped to mouse genome mm10 using minimap2/2.24 [102] with parameters -a -x splice -k 12 -u b --secondary=no. Basecalled reads were also separately aligned against the mouse transcriptome (Ensembl version 92) using -a -x map-ont -k 12 -u f --secondary=no. FLAIR/v1.7.0 [103] was used to identify and quantify novel transcripts. To create a unified database of existing and newly identified transcripts all sequenced samples were pooled at the flair collapse step, as suggested by the authors. Internally, DESeq2 [104] and DRIMSeq [105] were used for differential expression and differential isoform usage calculations respectively. Poly(A) tail lengths were extracted from sequenced reads using the nanopolish polya package [106]. Only poly(A) tail lengths that passed the software quality control scores and that were tagged as “PASS” were used in our analysis.

To identify transcripts with systematic whole-molecule or poly(A) length changes across experimental conditions, we employed in-house scripts that use linear mixed models to compare replicates using the library as a random effect (https://github.com/maragkakislab/nanoplen).

Human data analysis

GTEx RNA-seq analysis. Gene read counts for the brain (cortex), heart (left ventricle), and liver were obtained from the GTEx portal (V8 release). GTEx donors that experienced a slow death, defined by death after a long illness with more than 1 day of a terminal phase (DTHHRDY = 4), were filtered out. RNA-seq data was filtered for protein-coding genes with an average mRNA count ≥ 10. ImpulseDE2/0.99/10 [66] was employed in R/4.0.5 to identify transient and monotonous gene expression changes with age. The “vecConfounders”

parameter was used to account for donor sex. An FDR-corrected *p*-value cutoff of 0.05 was used for statistical significance.

Human 5hmC data. Human 5hmC data for corresponding tissues were mined from published datasets, He et al. [31] and Cui et al. [30]. Tissue-specific genes were obtained from the Human Protein Atlas and defined as genes with at least four-fold higher expression in the tissue compared to any other tissues ("tissue enriched") or the average of all other tissues ("tissue enhanced") [65]. Bed files for the ImpulseDE2 differential genes and tissue-specific genes were generated using the University of California, Santa Cruz (UCSC) Genome Browser (Table Browser) with the appropriate genome assembly corresponding to the 5hmC data (hg38 for He et al. [31] and hg19 for Cui et al. [30]).

PCA plots

hMeDIP-seq or MeDIP-seq PCA plots were generated using the plotPCA function of deepTools/3.5.0 [94] using multiBigwig summary files.

Area under the curve (AUC) calculation

To obtain genome coverage information (AUC) across regions of interest, we used the summary function of bwtool/1.0 [107] with additional parameter "-with-sum".

Annotation

Genomic annotations were performed using annotatr/1.16.0 [108] with "mm10_cpgs", "mm10_basicgenes", "mm10_genes_intergenic", and "mm10_genes_intronexonboundaries" as annotations.

Gene ontology analysis

GO analysis for the DHMRs and DMRs were performed using GREAT/4.0.4 [41] with the mm10 genome as background and default association rule settings. The top 5 significant biological process category terms are reported (ranked according to FDR). For the GTEx data, GO analysis was performed using DAVID/6.8 with *Mus musculus* genes as background. The top biological process category terms with $p < 0.05$ are reported (ranked according to *p*-value).

Genome browser tracks

bigWig files for individual and pooled (across replicates) samples were used to generate genome browser tracks via the UCSC Genome Browser using either custom tracks or track hubs.

Heatmaps

All heatmaps, except for the peak correlation heatmaps, were generated using the ComplexHeatmap/3.16 package in R with row standardized values. DiffBind was used to generate peak correlation heatmaps based on peaksets obtained from macs/2.2.7.1.

Motifs

Motif analysis to identify transcription factors was performed using LISA (http://lisa.cistrome.org/).

Hypergeometric test

Hypergeometric tests for the venn diagrams were performed using EVenn [109].

Data availability

All genome-wide datasets have been submitted to the Gene Expression Omnibus portal (GEO: GSE221124) and will be publicly available as of the date of publication. Raw mass spec data are deposited at chorusproject.org/1795.

Code availability

All code used in this study is at https://github.com/PSenlab/Occean_2022.

Acknowledgements

We wish to acknowledge the National Institute on Aging Intramural Research Program (NIA IRP), National Institutes of Health (NIH), for financial support. This work was funded by grant NIH ZIA AG000679 to P.S. Y.S. and S.S. gratefully acknowledge the Leukemia Research Foundation, AFAR, the Japan Agency for Medical Research and Development, Merck, Deerfield, Einstein-Mount Sinai Pilot Diabetes grant, and NIH grants 1S10OD030286-01 and P30CA01333047. We thank the Comparative Medicine Section at NIA for providing support with animal experiments and Elin Lehrmann for GEO upload. We appreciate Myriam Gorospe,

Weidong Wang and David Schlessinger, for critical reading of the manuscript. We thank Radhika Patnala and Arne Fabritius for illustrations. This work utilized the computational resources of the NIH HPC Biowulf cluster (http://hpc.nih.gov).

Author Contributions

P.S. conceptualized the project. J.O. and P.S. wrote the manuscript. J.O. performed wet lab experiments with help from N.Y., M. S. D., R.M., L.W. and C.Y.S. Y.S., J.S.K. and S.S. ran and analyzed DNA and oligo pulldown mass-spec data. N.Y. and C.B. performed direct RNA-seq experiments. S.D. and E.M. analyzed direct RNA-seq data. C.A. provided senescent cells and performed senescence measures. J.O., N.Y. and P.S. performed bioinformatics analyses except direct RNA-seq. C.Y.C. assisted in DHE staining and mouse dissections. M.B. (Bernier), N.L.P. and R.dC. provided livers from disulfiram-treated animals. C.D. helped with TMRM flow cytometry assay. M.B. (Bhattacharyya) aided with oligo mass-spec experiments. J.F. and S.D. provided sequencing support.

Competing Interests

The authors declare no competing interests.

References

1. Kayo, T., Allison, D.B., Weindruch, R. & Prolla, T.A. Influences of aging and caloric restriction on the transcriptional profile of skeletal muscle from rhesus monkeys. *Proc Natl Acad Sci U S A* **98**, 5093-5098 (2001).
2. Linford, N.J. *et al.* Transcriptional response to aging and caloric restriction in heart and adipose tissue. *Aging Cell* **6**, 673-688 (2007).
3. Roy, A.K. *et al.* Impacts of transcriptional regulation on aging and senescence. *Ageing research reviews* **1**, 367-380 (2002).
4. Soreq, L. *et al.* Major Shifts in Glial Regional Identity Are a Transcriptional Hallmark of Human Brain Aging. *Cell Reports* **18**, 557-570 (2017).
5. Booth, Lauren N. & Brunet, A. The Aging Epigenome. *Molecular Cell* **62**, 728-744 (2016).
6. Sen, P., Shah, P.P., Nativio, R. & Berger, S.L. Epigenetic Mechanisms of Longevity and Aging. *Cell* **166**, 822-839 (2016).
7. Yang, N. & Sen, P. The senescent cell epigenome. *Aging (Albany NY)* **10**, 3590 (2018).
8. Esteller, M. Epigenetics in cancer. *New England Journal of Medicine* **358**, 1148-1159 (2008).
9. Herman, A.B., Occean, J.R. & Sen, P. Epigenetic dysregulation in cardiovascular aging and disease. *J Cardiovasc Aging* **1**, 10 (2021).

10. Hwang, J.Y., Aromolaran, K.A. & Zukin, R.S. The emerging field of epigenetics in neurodegeneration and neuroprotection. *Nat Rev Neurosci* **18**, 347-361 (2017).
11. Ling, C. & Rönn, T. Epigenetics in Human Obesity and Type 2 Diabetes. *Cell Metabolism* **29**, 1028-1044 (2019).
12. Field, A.E. *et al.* DNA Methylation Clocks in Aging: Categories, Causes, and Consequences. *Molecular Cell* **71**, 882-895 (2018).
13. Horvath, S. & Raj, K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nature Reviews Genetics* **19**, 371-384 (2018).
14. Minteer, C. *et al.* Tick tock, tick tock: Mouse culture and tissue aging captured by an epigenetic clock. *Aging Cell* **21**, e13553 (2022).
15. Ito, S. *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129-1133 (2010).
16. Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930-935 (2009).
17. Madugundu, G.S., Cadet, J. & Wagner, J.R. Hydroxyl-radical-induced oxidation of 5-methylcytosine in isolated and cellular DNA. *Nucleic Acids Research* **42**, 7450-7460 (2014).
18. Blaschke, K. *et al.* Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* **500**, 222-226 (2013).
19. Yin, R. *et al.* Ascorbic Acid Enhances Tet-Mediated 5-Methylcytosine Oxidation and Promotes DNA Demethylation in Mammals. *Journal of the American Chemical Society* **135**, 10396-10403 (2013).
20. Inoue, A. & Zhang, Y. Replication-Dependent Loss of 5-Hydroxymethylcytosine in Mouse Preimplantation Embryos. *Science* **334**, 194-194 (2011).
21. Mooijman, D., Dey, S.S., Boisset, J.-C., Crosetto, N. & van Oudenaarden, A. Single-cell 5hmC sequencing reveals chromosome-wide cell-to-cell variability and enables lineage reconstruction. *Nature Biotechnology* **34**, 852-856 (2016).
22. Li, W. & Liu, M. Distribution of 5-hydroxymethylcytosine in different human tissues. *J Nucleic Acids* **2011**, 870726 (2011).
23. Hu, L. *et al.* Structural insight into substrate preference for TET-mediated oxidation. *Nature* **527**, 118-122 (2015).
24. Asadi Shahmirzadi, A. *et al.* Alpha-Ketoglutarate, an Endogenous Metabolite, Extends Lifespan and Compresses Morbidity in Aging Mice. *Cell Metab* **32**, 447-456.e446 (2020).
25. Chin, R.M. *et al.* The metabolite α-ketoglutarate extends lifespan by inhibiting ATP synthase and TOR. *Nature* **510**, 397-401 (2014).
26. Su, Y. *et al.* Alpha-ketoglutarate extends Drosophila lifespan by inhibiting mTOR and activating AMPK. *Aging (Albany NY)* **11**, 4183 (2019).
27. Lu, Y. *et al.* Reprogramming to recover youthful epigenetic information and restore vision. *Nature* **588**, 124-129 (2020).
28. Nestor, C., Ruzov, A., Meehan, R.R. & Dunican, D.S. Enzymatic approaches and bisulfite sequencing cannot distinguish between 5-methylcytosine and 5-hydroxymethylcytosine in DNA. *BioTechniques* **48**, 317-319 (2010).
29. Spruijt, C.G. *et al.* Dynamic readers for 5-(hydroxy) methylcytosine and its oxidized derivatives. *Cell* **152**, 1146-1159 (2013).
30. Cui, X.-L. *et al.* A human tissue map of 5-hydroxymethylcytosines exhibits tissue specificity through gene and enhancer modulation. *Nature Communications* **11**, 6161 (2020).
31. He, B. *et al.* Tissue-specific 5-hydroxymethylcytosine landscape of the human genome. *Nature Communications* **12**, 4249 (2021).
32. Lin, I.-H., Chen, Y.-F. & Hsu, M.-T. Correlated 5-hydroxymethylcytosine (5hmC) and gene expression profiles underpin gene and organ-specific epigenetic regulation in adult mouse brain and liver. *PloS one* **12**, e0170779 (2017).
33. Mellén, M., Ayata, P., Dewell, S., Kriaucionis, S. & Heintz, N. MeCP2 Binds to 5hmC Enriched within Active Genes and Accessible Chromatin in the Nervous System. *Cell* **151**, 1417-1430 (2012).

34. Nativio, R. *et al.* An integrated multi-omics approach identifies epigenetic alterations associated with Alzheimer's disease. *Nature Genetics* **52**, 1024-1035 (2020).
35. Nestor, C.E. *et al.* Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Res* **22**, 467-477 (2012).
36. Stroud, H., Feng, S., Morey Kinney, S., Pradhan, S. & Jacobsen, S.E. 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biology* **12**, R54 (2011).
37. Kriaucionis, S. & Heintz, N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**, 929-930 (2009).
38. Tammen, S.A. *et al.* Aging alters hepatic DNA hydroxymethylation, as measured by liquid chromatography/mass spectrometry. *J Cancer Prev* **19**, 301-308 (2014).
39. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**, 1-9 (2008).
40. Stark, R. & Brown, G. DiffBind: differential binding analysis of ChIP-Seq peak data. *R package version* **100** (2011).
41. McLean, C.Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnology* **28**, 495-501 (2010).
42. Qin, Q. *et al.* Lisa: inferring transcriptional regulators through integrative modeling of public chromatin accessibility and ChIP-seq data. *Genome Biology* **21**, 32 (2020).
43. Brown, J.D. & Plutzky, J. Peroxisome proliferator-activated receptors as transcriptional nodal points and therapeutic targets. *Circulation* **115**, 518-533 (2007).
44. Liang, B. *et al.* TBX3 functions as a tumor suppressor downstream of activated CTNNB1 mutants during hepatocarcinogenesis. *J Hepatol* **75**, 120-131 (2021).
45. Mukherjee, S., French, D.L. & Gadue, P. Loss of TBX3 enhances pancreatic progenitor generation from human pluripotent stem cells. *Stem Cell Reports* **16**, 2617-2627 (2021).
46. Loft, A. *et al.* Liver-fibrosis-activated transcriptional networks govern hepatocyte reprogramming and intra-hepatic communication. *Cell Metabolism* **33**, 1685-1700.e1689 (2021).
47. Pu, M. *et al.* Trimethylation of Lys36 on H3 restricts gene expression change during aging and impacts life span. *Genes & development* **29**, 718-731 (2015).
48. Nan, X., Campoy, F.J. & Bird, A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* **88**, 471-481 (1997).
49. Salminen, A., Kaarniranta, K. & Kauppinen, A. ER stress activates immunosuppressive network: implications for aging and Alzheimer's disease. *J Mol Med (Berl)* **98**, 633-650 (2020).
50. Zhu, B. *et al.* ER-associated degradation regulates Alzheimer's amyloid pathology and memory function by modulating γ-secretase activity. *Nature Communications* **8**, 1-15 (2017).
51. Dicks, N., Gutierrez, K., Michalak, M., Bordignon, V. & Agellon, L.B. Endoplasmic reticulum stress, genome damage, and cancer. *Front Oncol* **5**, 11 (2015).
52. Tourrière, H. *et al.* RasGAP-associated endoribonuclease G3Bp: selective RNA degradation and phosphorylation-dependent localization. *Mol Cell Biol* **21**, 7747-7760 (2001).
53. Fischer, J.W., Busa, V.F., Shao, Y. & Leung, A.K.L. Structure-Mediated RNA Decay by UPF1 and G3BP1. *Mol Cell* **78**, 70-84.e76 (2020).
54. Shen, S. *et al.* rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proceedings of the National Academy of Sciences* **111**, E5593-E5601 (2014).
55. Fong, Y.W. & Zhou, Q. Stimulatory effect of splicing factors on transcriptional elongation. *Nature* **414**, 929-933 (2001).
56. Kornblihtt, A.R., de la Mata, M., Fededa, J.P., Munoz, M.J. & Nogues, G. Multiple links between transcription and splicing. *Rna* **10**, 1489-1498 (2004).
57. Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S. & Fu, X.-D. The splicing factor SC35 has an active role in transcriptional elongation. *Nature structural & molecular biology* **15**, 819-826 (2008).
58. da Costa, P.J., Menezes, J. & Romão, L. The role of alternative splicing coupled to nonsense-mediated mRNA decay in human disease. *The International Journal of Biochemistry & Cell Biology* **91**, 168-175 (2017).

59. Lareau, L.F., Brooks, A.N., Soergel, D.A., Meng, Q. & Brenner, S.E. The coupling of alternative splicing and nonsense-mediated mRNA decay. *Adv Exp Med Biol* **623**, 190-211 (2007).
60. Lewis, B.P., Green, R.E. & Brenner, S.E. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc Natl Acad Sci U S A* **100**, 189-192 (2003).
61. Ni, J.Z. *et al.* Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes Dev* **21**, 708-718 (2007).
62. Yin, R. *et al.* Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. *J Am Chem Soc* **135**, 10396-10403 (2013).
63. Globisch, D. *et al.* Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS One* **5**, e15367 (2010).
64. Lonsdale, J. *et al.* The genotype-tissue expression (GTEx) project. *Nature genetics* **45**, 580-585 (2013).
65. Uhlén, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science (New York, NY)* **347**, 1260419-1260419 (2015).
66. Fischer, D.S., Theis, F.J. & Yosef, N. Impulse model-based differential expression analysis of time course sequencing data. *Nucleic acids research* **46**, e119-e119 (2018).
67. Daly, C.M. *et al.* Sex differences in response to a high fat, high sucrose diet in both the gut microbiome and hypothalamic astrocytes and microglia. *Nutr Neurosci* **25**, 321-335 (2022).
68. Iena, F.M. *et al.* Sex-specific effect of high-fat diet on glycerol metabolism in murine adipose tissue and liver. *Frontiers in endocrinology* **11**, 577650 (2020).
69. Oraha, J., Enriquez, R.F., Herzog, H. & Lee, N.J. Sex-specific changes in metabolism during the transition from chow to high-fat diet feeding are abolished in response to dieting in C57BL/6J mice. *International Journal of Obesity* **46**, 1749-1758 (2022).
70. Bernier, M. *et al.* Disulfiram Treatment Normalizes Body Weight in Obese Mice. *Cell Metab* **32**, 203-214.e204 (2020).
71. Laird, A., Thomson, J.P., Harrison, D.J. & Meehan, R.R. 5-hydroxymethylcytosine profiling as an indicator of cellular state. *Epigenomics* **5**, 655-669 (2013).
72. Pfeifer, G.P. & Szabó, P.E. Gene body profiles of 5-hydroxymethylcytosine: potential origin, function and use as a cancer biomarker. *Epigenomics* **10**, 1029-1032 (2018).
73. Fujiki, K. *et al.* PPARγ-induced PARylation promotes local DNA demethylation by production of 5-hydroxymethylcytosine. *Nature Communications* **4**, 2262 (2013).
74. Luo, X. *et al.* Coordination of germ layer lineage choice by TET1 during primed pluripotency. *Genes Dev* **34**, 598-618 (2020).
75. Pells, S. *et al.* Novel Human Embryonic Stem Cell Regulators Identified by Conserved and Distinct CpG Island Methylation State. *PLoS One* **10**, e0131102 (2015).
76. Dan, J. *et al.* Roles for Tbx3 in regulation of two-cell state and telomere elongation in mouse ES cells. *Scientific Reports* **3**, 3492 (2013).
77. Kang, J. *et al.* Simultaneous deletion of the methylcytosine oxidases Tet1 and Tet3 increases transcriptome variability in early embryogenesis. *Proceedings of the National Academy of Sciences* **112**, E4236-E4245 (2015).
78. Bradner, J.E., Hnisz, D. & Young, R.A. Transcriptional Addiction in Cancer. *Cell* **168**, 629-643 (2017).
79. Lio, C.-W.J., Yuita, H. & Rao, A. Dysregulation of the TET family of epigenetic regulators in lymphoid and myeloid malignancies. *Blood* **134**, 1487-1497 (2019).
80. Rasmussen, K.D. & Helin, K. Role of TET enzymes in DNA methylation, development, and cancer. *Genes & development* **30**, 733-750 (2016).
81. Sen, P. *et al.* H3K36 methylation promotes longevity by enhancing transcriptional fidelity. *Genes Dev* **29**, 1362-1376 (2015).
82. Benjamini, Y., Krieger, A.M. & Yekutieli, D. Adaptive linear step-up procedures that control the false discovery rate. *Biometrika* **93**, 491-507 (2006).
83. Mitchell, C. & Willenbring, H. A reproducible and well-tolerated method for 2/3 partial hepatectomy in mice. *Nat Protoc* **3**, 1167-1170 (2008).

84. Sun, Y., Stransky, S., Aguilan, J., Brenowitz, M. & Sidoli, S. DNA methylation and hydroxymethylation analysis using a high throughput and low bias direct injection mass spectrometry platform. *MethodsX* **8**, 101585 (2021).
85. Sun, Y. *et al.* High throughput and low bias DNA methylation and hydroxymethylation analysis by direct injection mass spectrometry. *Anal Chim Acta* **1180**, 338880 (2021).
86. Singh, R.K., Diaz, P.E., Binette, F. & Nasonkin, I.O. Immunohistochemical Detection of 5-Methylcytosine and 5-Hydroxymethylcytosine in Developing and Postmitotic Mouse Retina. *J Vis Exp* (2018).
87. Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nature methods* **9**, 671-675 (2012).
88. Aguilan, J.T., Kulej, K. & Sidoli, S. Guide for protein fold change and p-value calculation for non-experts in proteomics. *Mol Omics* **16**, 573-582 (2020).
89. Ibrahim, F., Oppelt, J., Maragkakis, M. & Mourelatos, Z. TERA-Seq: true end-to-end sequencing of native RNA molecules for transcriptome characterization. *Nucleic Acids Res* **49**, e115 (2021).
90. Andrews, S. FastQC a Quality Control Tool for High Throughput Sequence Data [Online]. (2010).
91. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
92. Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032-2034 (2015).
93. Amemiya, H.M., Kundaje, A. & Boyle, A.P. The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Scientific Reports* **9**, 9354 (2019).
94. Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic acids research* **42**, W187-W191 (2014).
95. Bolger, A.M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).
96. Andrews, S. (Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom, 2010).
97. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
98. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
99. Liao, Y., Smyth, G.K. & Shi, W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res* **47**, e47 (2019).
100. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**, 1-21 (2014).
101. Shen, S. *et al.* rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc Natl Acad Sci U S A* **111**, E5593-5601 (2014).
102. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094-3100 (2018).
103. Tang, A.D. *et al.* Full-length transcript characterization of SF3B1 mutation in chronic lymphocytic leukemia reveals downregulation of retained introns. *Nat Commun* **11**, 1438 (2020).
104. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
105. Nowicka, M. & Robinson, M.D. DRIMSeq: a Dirichlet-multinomial framework for multivariate count outcomes in genomics. *F1000Res* **5**, 1356 (2016).
106. Workman, R.E. *et al.* Nanopore native RNA sequencing of a human poly(A) transcriptome. *Nat Methods* **16**, 1297-1305 (2019).
107. Pohl, A. & Beato, M. bwtool: a tool for bigWig files. *Bioinformatics* **30**, 1618-1619 (2014).
108. Cavalcante, R.G. & Sartor, M.A. Annotatr: genomic regions in context. *Bioinformatics* **33**, 2381-2383 (2017).
109. Chen, T., Zhang, H., Liu, Y., Liu, Y.-X. & Huang, L. EVenn: Easy to create repeatable and editable Venn diagrams and Venn networks online. *Journal of Genetics and Genomics* **48**, 863-866 (2021).

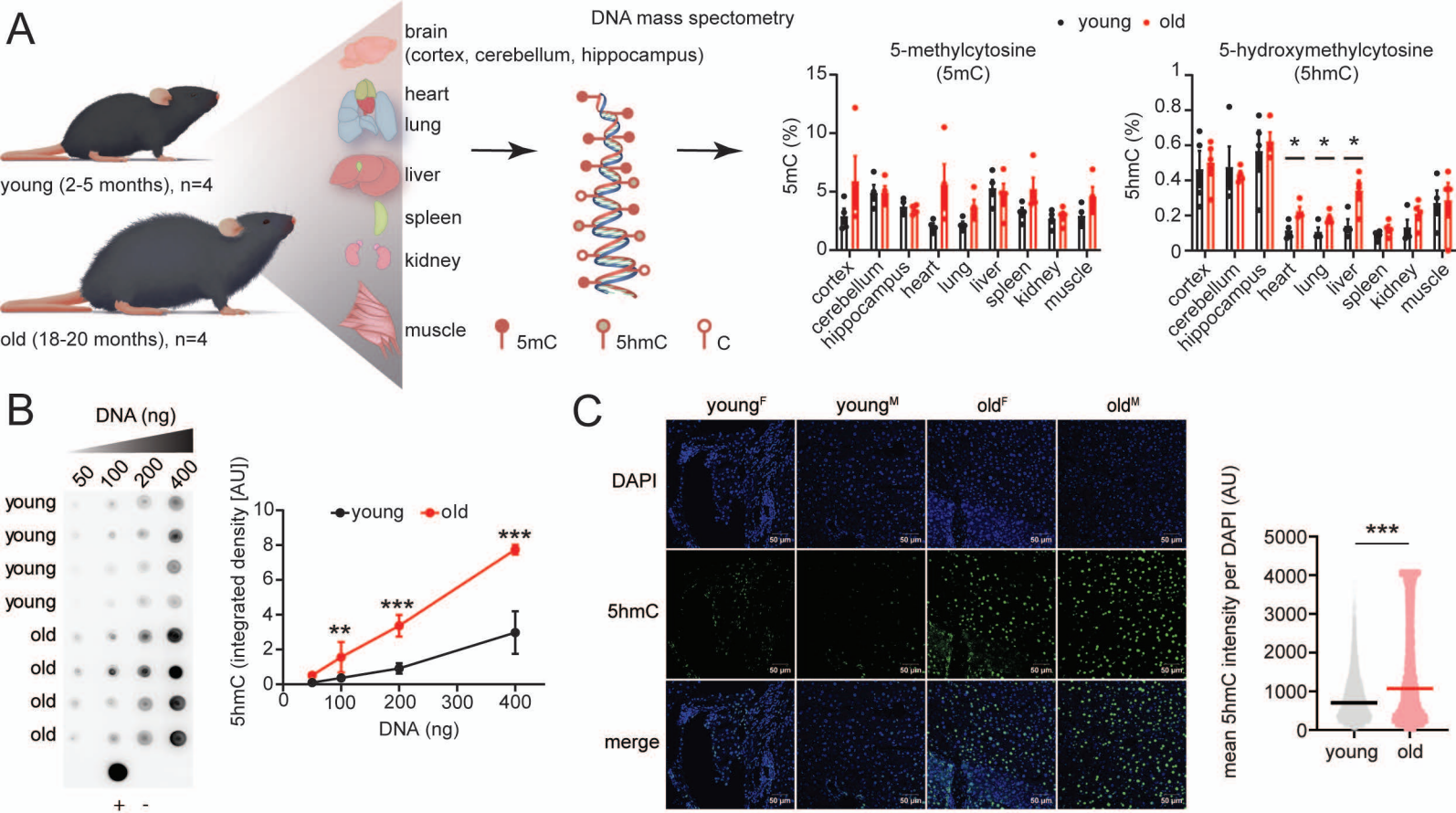
A
young (2-5 months), n=4
old (18-20 months), n=4
brain
(cortex, cerebellum, hippocampus)
heart
lung
liver
spleen
kidney
muscle
DNA mass spectometry
5mC
5hmC
C
young
old
5-methylcytosine
(5mC)
5mC (%)
5-hydroxymethylcytosine
(5hmC)
5hmC (%)
cortex
cerebellum
hippocampus
heart
lung
liver
spleen
kidney
muscle
B
DNA (ng)
50
100
200
400
young
young
young
young
old
old
old
old
+
-
5hmC (integrated density [AU])
DNA (ng)
young
old
**
***
***
C
young^F
young^M
old^F
old^M
DAPI
5hmC
merge
mean 5hmC intensity per DAPI (AU)
***
young
old

Figure 1

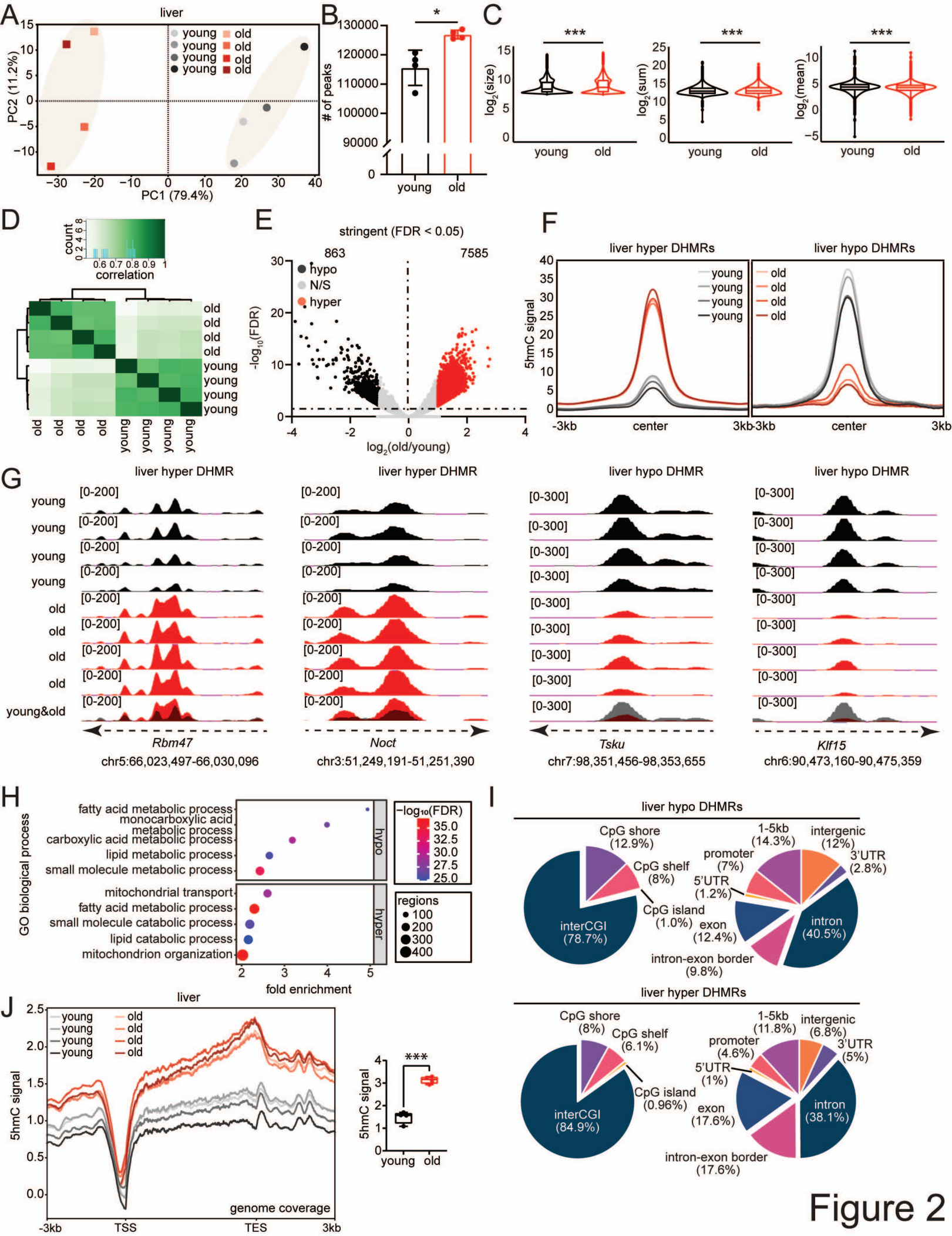
A
liver
young
young
young
young
old
old
old
old
PC2 (11.2%)
PC1 (79.4%)
B
# of peaks
young
old
C
log2(size)
log2(sum)
log2(mean)
young
old
D
count
correlation
old
young
E
stringent (FDR < 0.05)
863
7585
hypo
N/S
hyper
-log10(FDR)
log2(old/young)
F
liver hyper DHMRs
liver hypo DHMRs
5hmC signal
-3kb
center
3kb
G
liver hyper DHMR
liver hypo DHMR
young
young
young
young
old
old
old
old
young&old
[0-200]
[0-300]
Rbm47
chr5:66,023,497-66,030,096
Noct
chr3:51,249,191-51,251,390
Tsku
chr7:98,351,456-98,353,655
Klf15
chr6:90,473,160-90,475,359
H
GO biological process
fatty acid metabolic process
monocarboxylic acid metabolic process
carboxylic acid metabolic process
lipid metabolic process
small molecule metabolic process
mitochondrial transport
fatty acid metabolic process
small molecule catabolic process
lipid catabolic process
mitochondrion organization
hypo
hyper
-log10(FDR)
35.0
32.5
30.0
27.5
25.0
regions
100
200
300
400
fold enrichment
I
liver hypo DHMRs
CpG shore (12.9%)
CpG shelf (8%)
CpG island (1.0%)
interCGI (78.7%)
1-5kb (14.3%)
intergenic (12%)
promoter (7%)
3'UTR (2.8%)
5'UTR (1.2%)
exon (12.4%)
intron (40.5%)
intron-exon border (9.8%)
liver hyper DHMRs
CpG shore (8%)
CpG shelf (6.1%)
CpG island (0.96%)
interCGI (84.9%)
1-5kb (11.8%)
intergenic (6.8%)
promoter (4.6%)
3'UTR (5%)
5'UTR (1%)
exon (17.6%)
intron (38.1%)
intron-exon border (17.6%)
J
liver
young
young
young
young
old
old
old
old
5hmC signal
genome coverage
-3kb
TSS
TES
3kb
young
old

Figure 2

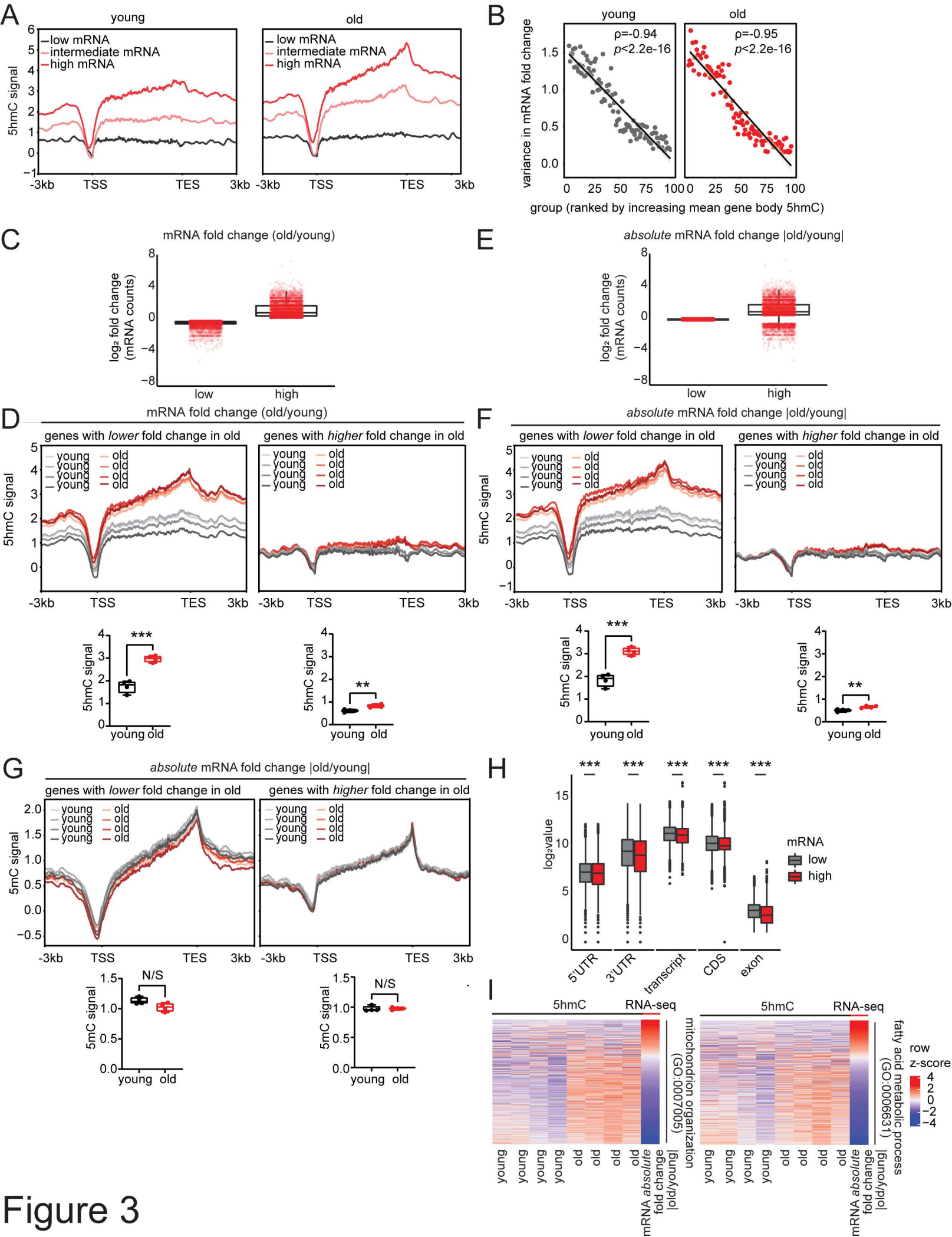
A
young
old
low mRNA
intermediate mRNA
high mRNA
5hmC signal
-3kb
TSS
TES
3kb
B
young
old
ρ=-0.94
p<2.2e-16
ρ=-0.95
p<2.2e-16
variance in mRNA fold change
group (ranked by increasing mean gene body 5hmC)
C
mRNA fold change (old/young)
log₂ fold change (mRNA counts)
low
high
E
absolute mRNA fold change |old/young|
D
genes with lower fold change in old
genes with higher fold change in old
young
old
***
**
F
G
absolute mRNA fold change |old/young|
5mC signal
N/S
H
log₂value
mRNA
low
high
5'UTR
3'UTR
transcript
CDS
exon
I
5hmC
RNA-seq
mitochondrion organization (GO:0007005)
fatty acid metabolic process (GO:0006631)
mRNA absolute fold change |old/young|
row z-score

Figure 3

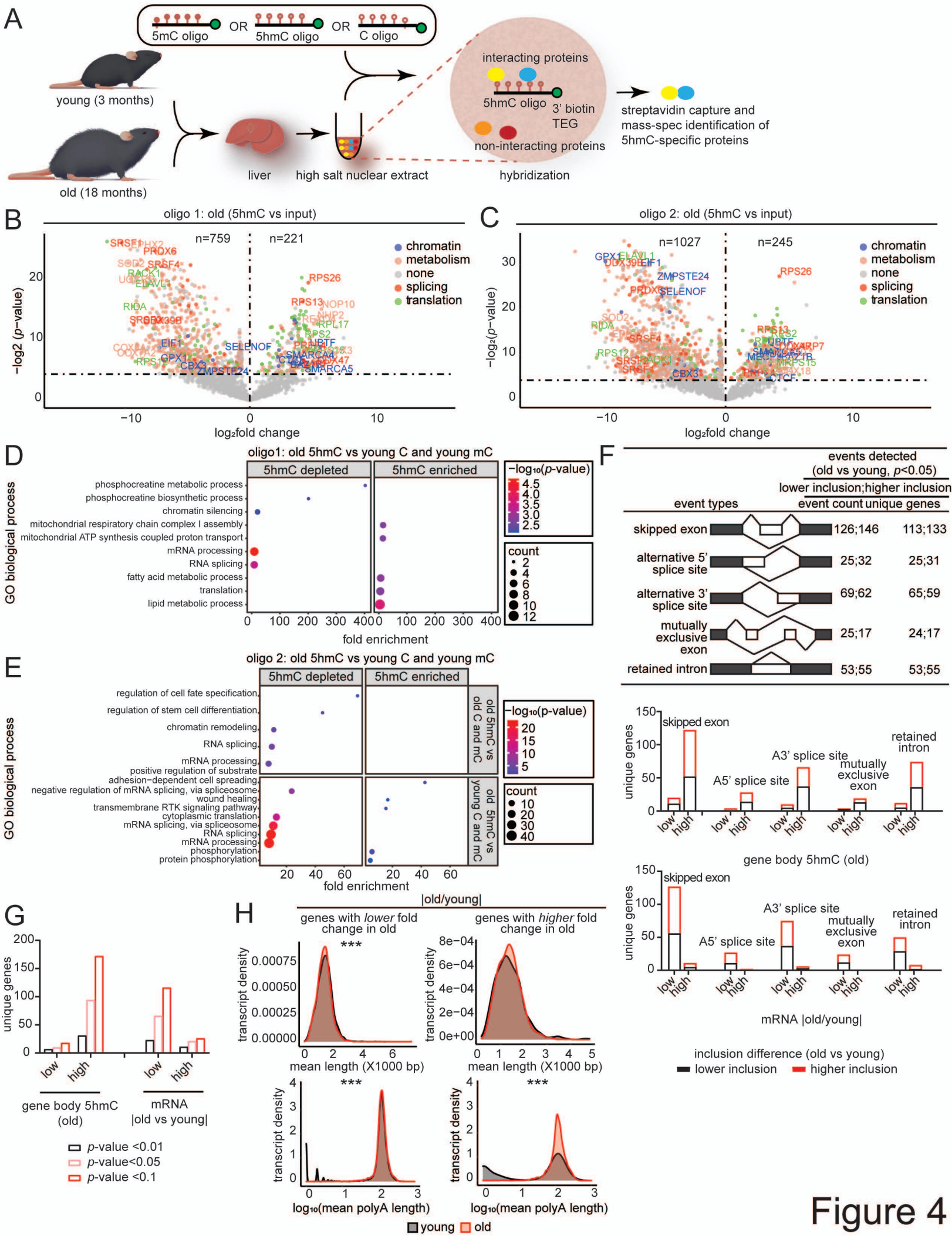
A
5mC oligo
OR
5hmC oligo
OR
C oligo
young (3 months)
old (18 months)
liver
high salt nuclear extract
interacting proteins
5hmC oligo
3' biotin TEG
non-interacting proteins
hybridization
streptavidin capture and mass-spec identification of 5hmC-specific proteins
B
oligo 1: old (5hmC vs input)
n=759
n=221
chromatin
metabolism
none
splicing
translation
−log2 (p−value)
log2fold change
C
oligo 2: old (5hmC vs input)
n=1027
n=245
−log2(p−value)
log2fold change
D
oligo1: old 5hmC vs young C and young mC
5hmC depleted
5hmC enriched
GO biological process
phosphocreatine metabolic process
phosphocreatine biosynthetic process
chromatin silencing
mitochondrial respiratory chain complex I assembly
mitochondrial ATP synthesis coupled proton transport
mRNA processing
RNA splicing
fatty acid metabolic process
translation
lipid metabolic process
−log10(p-value)
count
fold enrichment
E
oligo 2: old 5hmC vs young C and young mC
5hmC depleted
5hmC enriched
old 5hmC vs old C and mC
old 5hmC vs young C and mC
regulation of cell fate specification
regulation of stem cell differentiation
chromatin remodeling
RNA splicing
mRNA processing
positive regulation of substrate adhesion−dependent cell spreading
negative regulation of mRNA splicing, via spliceosome
wound healing
transmembrane RTK signaling pathway
cytoplasmic translation
mRNA splicing, via spliceosome
RNA splicing
mRNA processing
phosphorylation
protein phosphorylation
−log10(p-value)
count
fold enrichment
F
events detected (old vs young, p<0.05)
lower inclusion;higher inclusion
event types
event count
unique genes
skipped exon 126;146 113;133
alternative 5' splice site 25;32 25;31
alternative 3' splice site 69;62 65;59
mutually exclusive exon 25;17 24;17
retained intron 53;55 53;55
unique genes
skipped exon
A5' splice site
A3' splice site
mutually exclusive exon
retained intron
low
high
gene body 5hmC (old)
mRNA |old/young|
inclusion difference (old vs young)
lower inclusion
higher inclusion
G
unique genes
low
high
gene body 5hmC (old)
mRNA |old vs young|
p-value <0.01
p-value<0.05
p-value <0.1
H
|old/young|
genes with lower fold change in old
genes with higher fold change in old
transcript density
mean length (X1000 bp)
log10(mean polyA length)
young
old

Figure 4

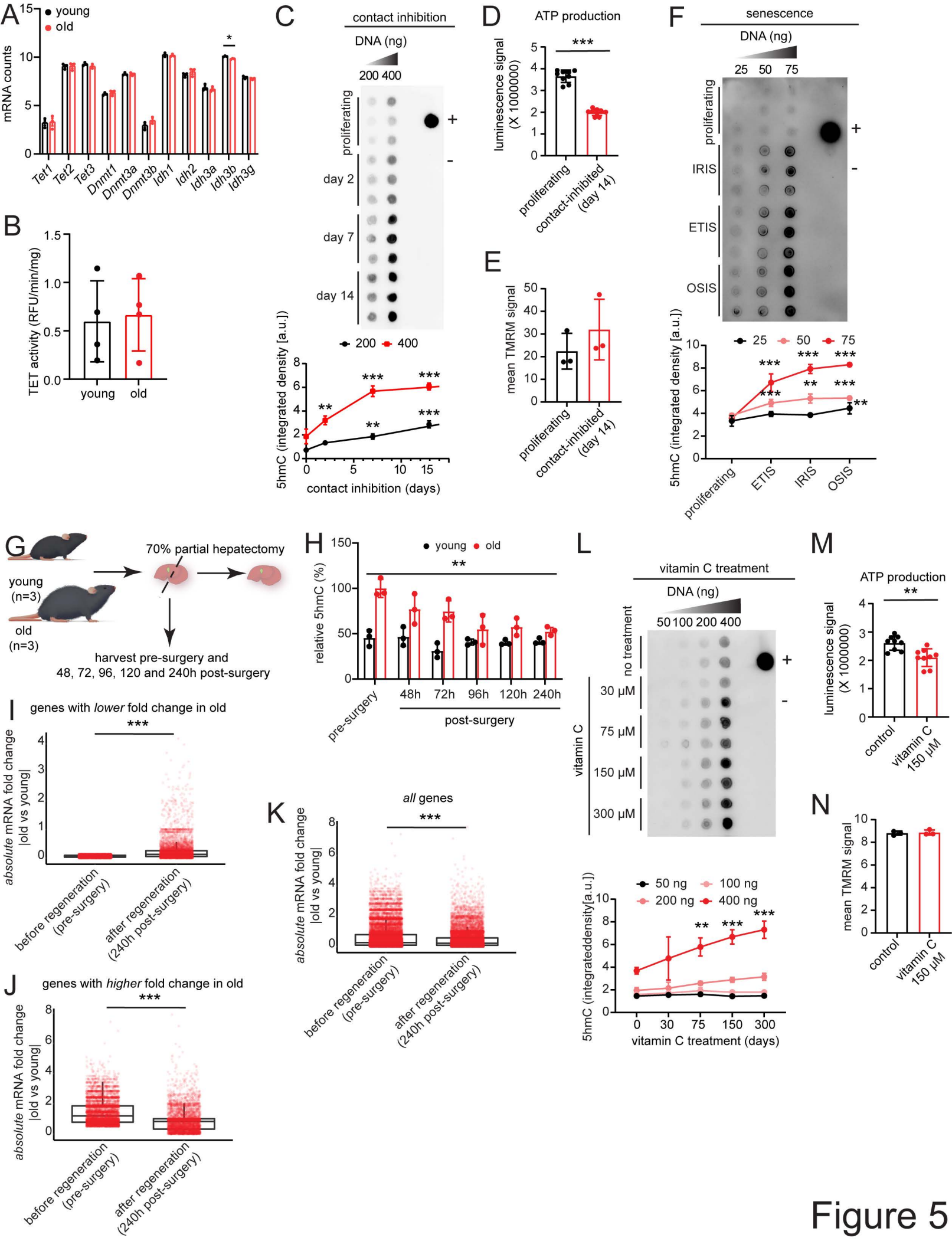
A
young
old
mRNA counts
Tet1
Tet2
Tet3
Dnmt1
Dnmt3a
Dnmt3b
Idh1
Idh2
Idh3a
Idh3b
Idh3g
B
TET activity (RFU/min/mg)
young
old
C
contact inhibition
DNA (ng)
200 400
proliferating
day 2
day 7
day 14
5hmC (integrated density [a.u.])
contact inhibition (days)
D
ATP production
luminescence signal (X 1000000)
proliferating
contact-inhibited (day 14)
E
mean TMRM signal
F
senescence
25 50 75
IRIS
ETIS
OSIS
G
young (n=3)
old (n=3)
70% partial hepatectomy
harvest pre-surgery and 48, 72, 96, 120 and 240h post-surgery
H
relative 5hmC (%)
pre-surgery
48h
72h
96h
120h
240h
post-surgery
I
genes with lower fold change in old
absolute mRNA fold change |old vs young|
before regeneration (pre-surgery)
after regeneration (240h post-surgery)
J
genes with higher fold change in old
K
all genes
L
vitamin C treatment
no treatment
30 μM
75 μM
150 μM
300 μM
vitamin C
50 ng
100 ng
200 ng
400 ng
5hmC (integrated density [a.u.])
vitamin C treatment (days)
M
control
vitamin C 150 μM
N

Figure 5

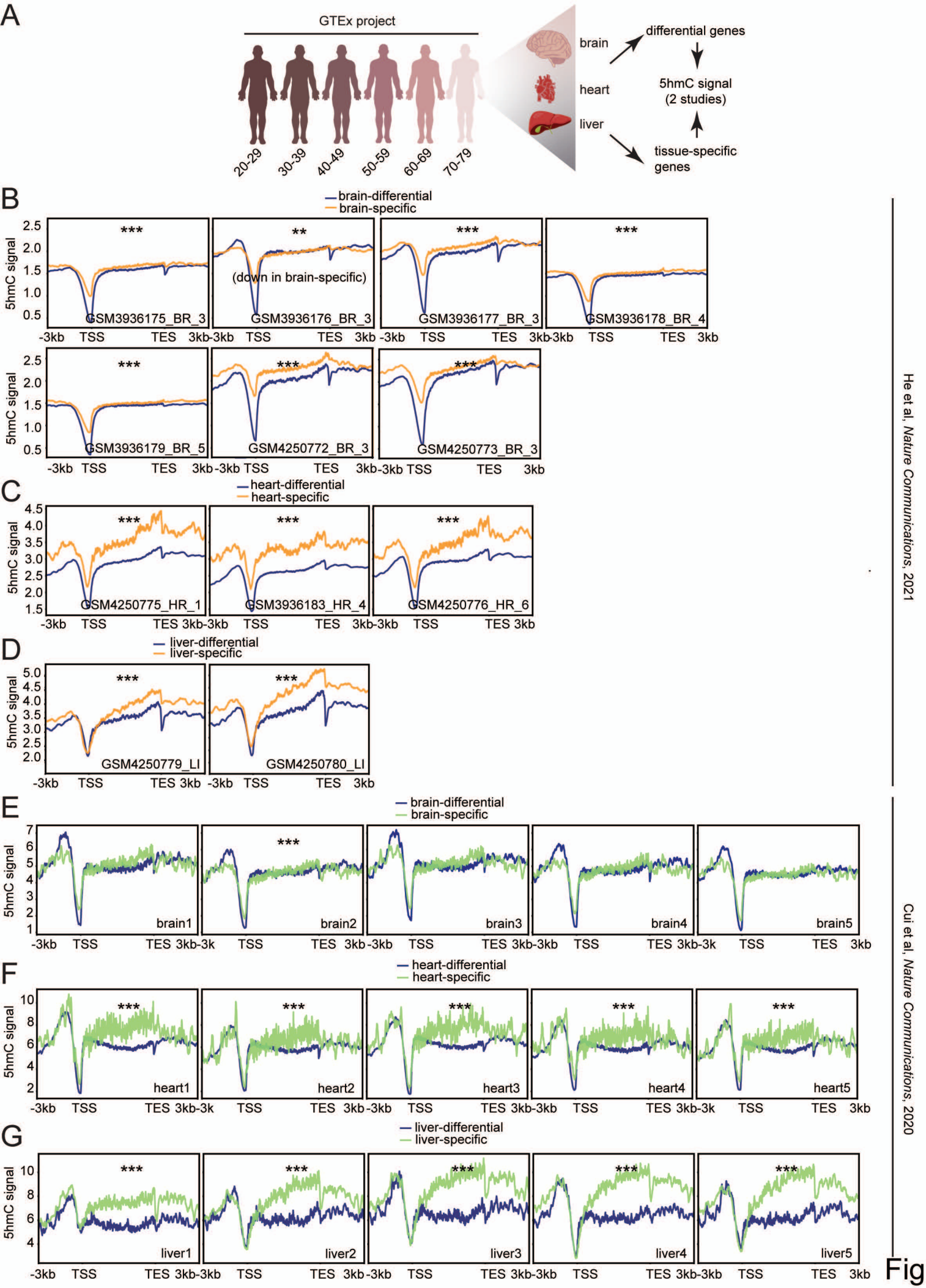
A
GTEx project
20-29
30-39
40-49
50-59
60-69
70-79
brain
heart
liver
differential genes
5hmC signal
(2 studies)
tissue-specific
genes
B
brain-differential
brain-specific
5hmC signal
***
**
(down in brain-specific)
GSM3936175_BR_3
GSM3936176_BR_3
GSM3936177_BR_3
GSM3936178_BR_4
GSM3936179_BR_5
GSM4250772_BR_3
GSM4250773_BR_3
-3kb TSS TES 3kb
C
heart-differential
heart-specific
GSM4250775_HR_1
GSM3936183_HR_4
GSM4250776_HR_6
D
liver-differential
liver-specific
GSM4250779_LI
GSM4250780_LI
E
brain-differential
brain-specific
brain1
brain2
brain3
brain4
brain5
F
heart-differential
heart-specific
heart1
heart2
heart3
heart4
heart5
G
liver-differential
liver-specific
liver1
liver2
liver3
liver4
liver5
He et al, Nature Communications, 2021
Cui et al, Nature Communications, 2020

Figure 6

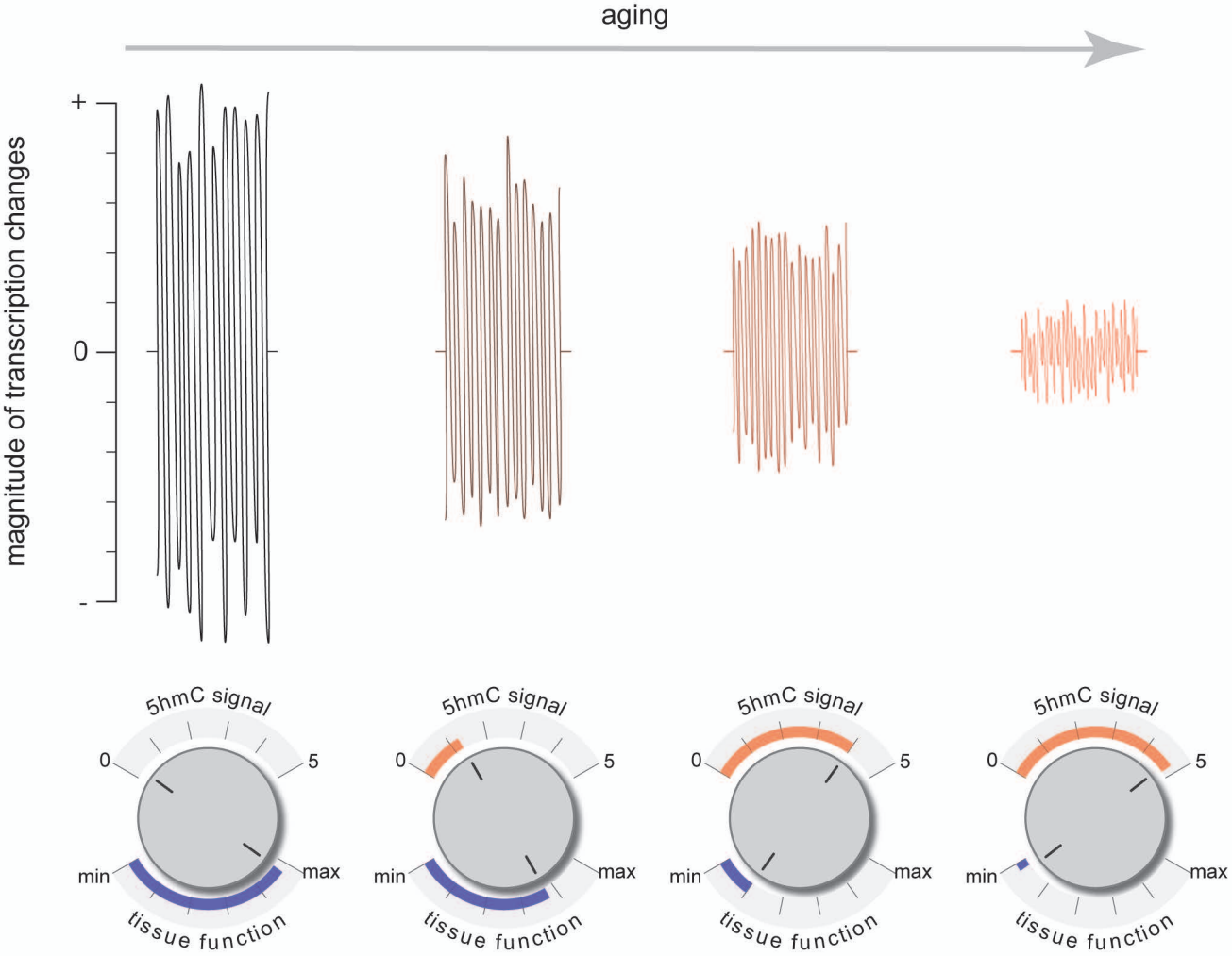
aging
magnitude of transcription changes
+
0
-
5hmC signal
0
5
min
max
tissue function
5hmC signal
0
5
min
max
tissue function
5hmC signal
0
5
min
max
tissue function
5hmC signal
0
5
min
max
tissue function

Figure 7