Healthy aging interventions reduce non-coding repetitive element transcripts

Devin Wahl[1,2], Alyssa N. Cavalier[1,2], Meghan Smith[1,2], Douglas R. Seals[3] and Thomas J. LaRocca[1,2,4]

1. Center for Healthy Aging, Colorado State University; Fort Collins, CO 80523
2. Department of Health and Exercise Science, Colorado State University; Fort Collins, CO 80523
3. Department of Integrative Physiology, University of Colorado Boulder; Boulder, CO 80309
4. Author for correspondence: Tom.LaRocca@colostate.edu

ABSTRACT

Transcripts from non-coding repetitive elements (RE) in the genome may be involved in aging. However, they are often ignored in transcriptome studies on healthspan and lifespan, and their role in healthy aging interventions has not been characterized. Here, we analyze RE in RNA-seq datasets from mice subjected to robust healthspan- and lifespan-increasing interventions including calorie restriction, rapamycin, acarbose, 17-α-estradiol, and Protandim. We also examine RE transcripts in long-lived transgenic mice, and in mice subjected to high-fat diet, and we use RNA-seq to investigate the influence of aerobic exercise on RE transcripts with aging in humans. We find that: 1) healthy aging interventions/behaviors globally reduce RE transcripts, whereas aging and age-accelerating treatments increase RE expression; and 2) reduced RE expression with healthy aging interventions is associated with biological/physiological processes mechanistically linked with aging. Thus, RE transcript dysregulation and suppression are likely novel mechanisms underlying aging and healthy aging interventions, respectively.

INTRODUCTION

Older age is the greatest risk factor for the development of most chronic diseases (1). Accordingly, recent large-scale 'omics' studies have aimed to characterize novel genes and biological pathways that influence aging, and to identify related interventions (e.g., pharmaceutical compounds, exercise, nutrition) that increase longevity and healthspan (2, 3). Indeed, advances in transcriptomics (e.g., RNA-seq) have led to important insight on many genes and pathways linked with 'the hallmarks of aging' and broader health outcomes (4). However, most of these studies have focused on coding sequences—a small fraction of the genome. Non-coding, repetitive elements (RE, >60% of the genome) have been particularly neglected as 'junk DNA' (5), despite growing evidence that they have many important biological functions (6).

RE include DNA transposons, retrotransposons, tandem repeats, satellites and terminal repeats (7). A major fraction of RE, mainly DNA transposons and retrotransposons (e.g., LINEs, SINEs, LTRs), are transposable elements (TE) with the ability to propagate, multiply and change genomic position (6). Most RE/TE are in genomic regions that are chromatinized and suppressed (inactive), but recent reports show that certain TE become active during aging, perhaps due to reduced chromatin architecture/stability (e.g., histone dysregulation) (8). Activation of these specific TE may contribute to aging by causing genomic and/or cellular damage/stress (e.g., inflammation) (9). However, we recently reported that aging is associated with a progressive, *global* increase in transcripts from most RE (not only TE) in model organisms and humans (10). This global dysregulation of RE may have an important, more general role in aging, as RE transcripts have been linked with other key hallmarks of aging including oxidative stress and cellular senescence (11). In fact, it has been suggested that RE dysregulation itself may be an important hallmark of aging (12). If so, a logical prediction would be that interventions that increase health/lifespan and reduce hallmarks of aging (e.g., calorie restriction [CR], select pharmacological agents and exercise) should also suppress RE/TE. Limited evidence suggests this may be true for CR and certain TE in *Drosophila* (13), but global RE/TE expression has not been well studied in this context.

RESULTS AND DISCUSSION

To determine if global RE transcript suppression might be a mechanism underlying healthy aging interventions, we first analyzed RE in an RNA-seq dataset on livers from young and old mice and old mice subjected to life-long (24 months) CR (14). We found a small, but significant age-related increase in most major RE transcript types in this dataset, consistent with our previous work (15) and others' (16). However, this effect was significantly attenuated with CR (**Fig. 1A**). Based on this novel evidence of RE suppression by CR (arguably the strongest health/lifespan-enhancing intervention), we looked to confirm our results in an additional, large dataset including RNA-seq on livers from mice subjected to different durations of CR and pharmacological interventions known to increase health/lifespan (rapamycin, acarbose, 17-α-estradiol [17aE], and Protandim), as well as data on transgenic, long-lived mice (2) (**Fig. 1 and Supplementary Data**). We found that long-term (8 months) CR caused a significant, global reduction in RE transcripts (**Fig. 1B**). Furthermore, we found that both long-term (8 months) rapamycin and acarbose treatments were associated with a comparable, broad reduction in RE transcripts (**Fig. 1B**), consistent with the notion that these compounds are 'calorie restriction mimetics' and may act via similar pathways (17). This effect was particularly clear when we examined RE/TE reductions by major sub-type (**Fig. 1C**). Short-term (2 months) interventions with other healthy aging compounds influenced RE transcript levels to various degrees, although reductions were more pronounced with CR and Protandim, which is thought to activate endogenous antioxidant defenses (**Fig. 1D**). Interestingly, the authors of the original study (2) observed similar variability in gene expression patterns, suggesting time/treatment-specific transcriptome effects. We also found a significant influence of growth hormone receptor knockout (GHRKO, a transgenic longevity model) on the main RE transcript types (**Fig. 1E**). Moreover, in a separate dataset (18), we found that high-fat diet (HFD, a common 'pro-aging' intervention) significantly *increased* all major RE/TE (**Fig. 1F**). Collectively, these results support the idea that global RE transcript levels are linked with healthspan/lifespan, as they are reduced by most "gold standard" anti-aging interventions and increased by adverse, pro-aging treatments.

Next, we examined similarities in the effects of healthy aging interventions on RE by sub-type/family (**Fig. 2 and Supplementary Data**). Again, we observed variable patterns of RE family expression with short-term treatments (**Fig. 2A**). With long-term treatment, CR and rapamycin influenced RE/TE families most similarly, and most transcripts were decreased with all treatments (**Fig. 2B**), and in GHRKO mice (**Fig. 2C**). We next determined which specific RE transcripts were commonly decreased/increased among all treatments. Despite the variable RE/family expression patterns noted above, short-term treatments modulated many of the same transcripts (**Fig. 2D and Supplementary**

Data). Long-term treatments also decreased/increased a large number of common transcripts (518 and 92, respectively) (**Fig. 2E and Supplementary Data**). Consistent with the idea that global RE modulation is linked with healthy aging interventions, we did not notice any particular enrichment for specific RE/TE types in these common transcripts. However, we did note that endogenous retrovirus (ERV) RE transcripts were the most decreased with all long-term treatments and in GHRKO mice. Interestingly, ERVs have been implicated in aging and several diseases of aging including neurodegenerative disorders (19, 20), suggesting that these RE could be an important therapeutic target.

Identifying potentially targetable biological mechanisms linking reduced RE expression with healthy aging interventions will require future experiments. However, to provide initial insight, we examined correlations among gene and RE expression patterns in mice subjected to long-term CR, rapamycin and acarbose (treatments that reduced RE transcripts the most). To do this, we conducted a weighted gene correlation network analysis (WGCNA) on both gene and RE transcript counts (**Fig. 3 and Supplementary Data**). Although gene/RE signatures across interventions were not strikingly similar, we identified one WGCNA module (green) that decreased significantly with all interventions (**Fig. 3A**). This module contained numerous DNA transposons and several LINE, ERV and LTR transcripts. A gene ontology (GO) analysis of the module also showed significant enrichment for many biological processes, including several linked with aging and disease (**Fig. 3B**). In fact, the most specific GO terms included protein deacetylation, DNA repair and immune activation/response pathways. The other gene/RE modules that decreased with CR and/or rapamycin were also enriched for specific GO terms including DNA repair, DNA/RNA processing, histone modifications and stress responses (**Fig. 3B**). These exploratory analyses do not definitively link reduced RE expression with such processes, but they are consistent with current thinking that age-related RE transcript accumulation could cause DNA damage (21) and immune activation/inflammation (22), and that RE dysregulation may be due to age-associated changes in chromatin/histones (23).

There is little or no RNA-seq data on true long-term CR or healthy aging compounds in older humans, as these are challenging clinical interventions to conduct (24). However, one well-studied intervention/behavior associated with increased healthspan and biological effects similar to CR is aerobic exercise (25, 26). Others have studied the transcriptomic effects of exercise interventions on select tissues (27), but proportionally, no studies have been nearly as long as an 8-month mouse intervention (~25% of the animal's life). Therefore, as initial proof of concept, we conducted a cross-sectional study to determine if long-term exposure to this healthy aging behavior has the potential to

reduce RE transcripts (**Fig. 4 and Supplementary Data**). We performed RNA-seq and gene/RE expression analyses on peripheral blood mononuclear cells (PBMC) from: 1) young and older sedentary adults; and 2) older habitually (≥5 years) exercising adults (**Supplementary Table**). Consistent with other reports (28), we found that older age was associated with altered PBMC gene expression, but these changes were largely attenuated in exercising older adults (**Fig. 4A**). Moreover, in support of the idea that healthy aging interventions/behaviors may reduce RE expression in humans, we observed a clear increase in global RE transcript levels in older sedentary adult PBMC, but this effect was strongly attenuated with exercise (**Fig. 4B**). We also found that maximal aerobic exercise capacity (VO_2 max) was inversely related to a composite count of RE that are significantly increased with aging (**Fig. 4C**), suggesting that greater aerobic fitness (and perhaps exposure to aerobic exercise) is directly linked with reduced RE expression. Interestingly, VO_2 max is considered a key physiological predictor of longevity in humans (29), further demonstrating that RE may have an important role in human healthspan/lifespan.

Collectively, our results support the growing idea that global RE dysregulation may be an important mechanism of aging (and not simply an adverse effect of the process). Reversing age-related RE transcript accumulation may be necessary for healthy aging, as our present findings show that health/lifespan-enhancing interventions consistently reduce RE expression. Indeed, we and others have reported age-associated increases in most types of RE transcripts (15, 16, 30), and this suggests a fundamental cellular mismanagement of RE with aging, which could have numerous deleterious effects. For example, our current study suggests that histone modifications, DNA damage and immune/inflammatory responses may be linked with RE dysregulation. Future investigations are needed to determine how and if these processes are specifically connected, as this could lead to novel strategies for additional, potentially complementary healthy aging interventions.

MATERIALS AND METHODS

RNA-seq datasets and availability

The data that support these findings can be found on the Gene Expression Omnibus under accession numbers (GEO): GSE92486 (long term CR), GSE131901 (8-month and 2-month CR and pharmacological treatments), GSE87565 (HFD), and GSE153100 (human exercise).

Bioinformatics analyses

RE transcripts were quantified using TEtranscripts (31) and the RepEnrich2 algorithm (32, 33) as previously described (15), in order to confirm similar findings with RE different analysis platforms. Briefly, reads were trimmed, quality filtered with *fastp* (34), and then aligned to the genome (mm10 *Mus musculus* or Hg38 *Homo sapiens*) using Bowtie (RepEnrich) or the STAR aligner (TEtranscripts) (35). RE transcripts were then quantified using either TEtranscripts or RepEnrich, which are pipelines to quantify RE transcripts by individual total counts, class, and family. Gene expression counts were extracted from bam alignment files produced during TEtranscripts analyses, and differential expression analyses of both RE and genes were performed using Deseq2 software (36). WGCNA was performed according to standard procedures outlined by the analysis pipeline's authors (37) using normalized gene and RE counts for all samples, and a minimum module size of 300 to capture broader groups of RE that correlated with sample traits (specific interventions). GO analyses of genes in the WGCNA modules were performed using the GOrilla algorithm (38), and specific GO modules were identified as terminal nodes in the directed acyclic graph produced by this program.

Human subjects and RNA-seq samples

RNA-seq was performed on PBMC from twelve healthy young (18-22 years) and older (62-74 years) adults. Subjects were non-obese, non-smokers and healthy as assessed by medical history, physical examination, blood chemistries and exercise ECG, and small groups were selected to match characteristics as closely as possible. Young (n=5, 2 male) and older (n=5, 2 m) sedentary subjects performed no regular exercise (< 2 days/week, < 30 min/day), whereas older exercising subjects (n=4, 1 m) performed regular vigorous aerobic exercise (≥ 5 days/week, > 45 min/day) for the previous ≥ 5 years. The study conformed to the Declaration of Helsinki; all procedures were approved by the Institutional Review Board of the University of Colorado Boulder, and written informed consent was obtained from all subjects. Maximal oxygen consumption (VO_2max) was assessed during treadmill exercise as previously described (39) and basic clinical measurements (e.g., blood pressure) were performed using standard techniques. PBMC were isolated from whole

blood by traditional Ficoll gradient centrifugation, and RNA-seq and gene expression analyses were performed using standard methods as previously described (15, 40). Briefly, snap-frozen PBMC pellets were lysed in Trizol (Thermo), and RNA was recovered using a spin column kit (Direct-Zol, Zymo Research) that included a DNase I treatment to remove genomic DNA. Total RNA libraries were generated using Illumina Ribo-Zero kits to deplete ribosomal RNA, and libraries were sequenced on an Illumina NovaSeq 6000 platform to produce >40 M 150-bp single-end fastq reads per sample. Gene and RE expression analyses were performed as described above.

Statistical Analyses

Differential expression of RE transcripts was quantified using the Deseq2 software as previously described using size factors to account for library size differences among samples (40). Chi-square analyses and heatmaps of increased/decreased RE and WGCNA modules were constructed using GraphPad Prism software, and Venn diagrams were generated using jvenn (41).

Acknowledgments

This work is funded by the National Institute on Aging, National Institutes of Health (AG060302).

Author contributions

D.W. designed the study, wrote the paper, generated and analyzed data, and provided conceptual insight; A.N.C. analyzed data, provided conceptual insight, and edited the paper; M.S., edited paper and provided conceptual insight; D.R.S. provided human PBMC samples, edited the paper, and provided conceptual insight; T.J.L. designed the study, wrote the paper, analyzed data and provided conceptual insight.

FIGURES AND LEGENDS

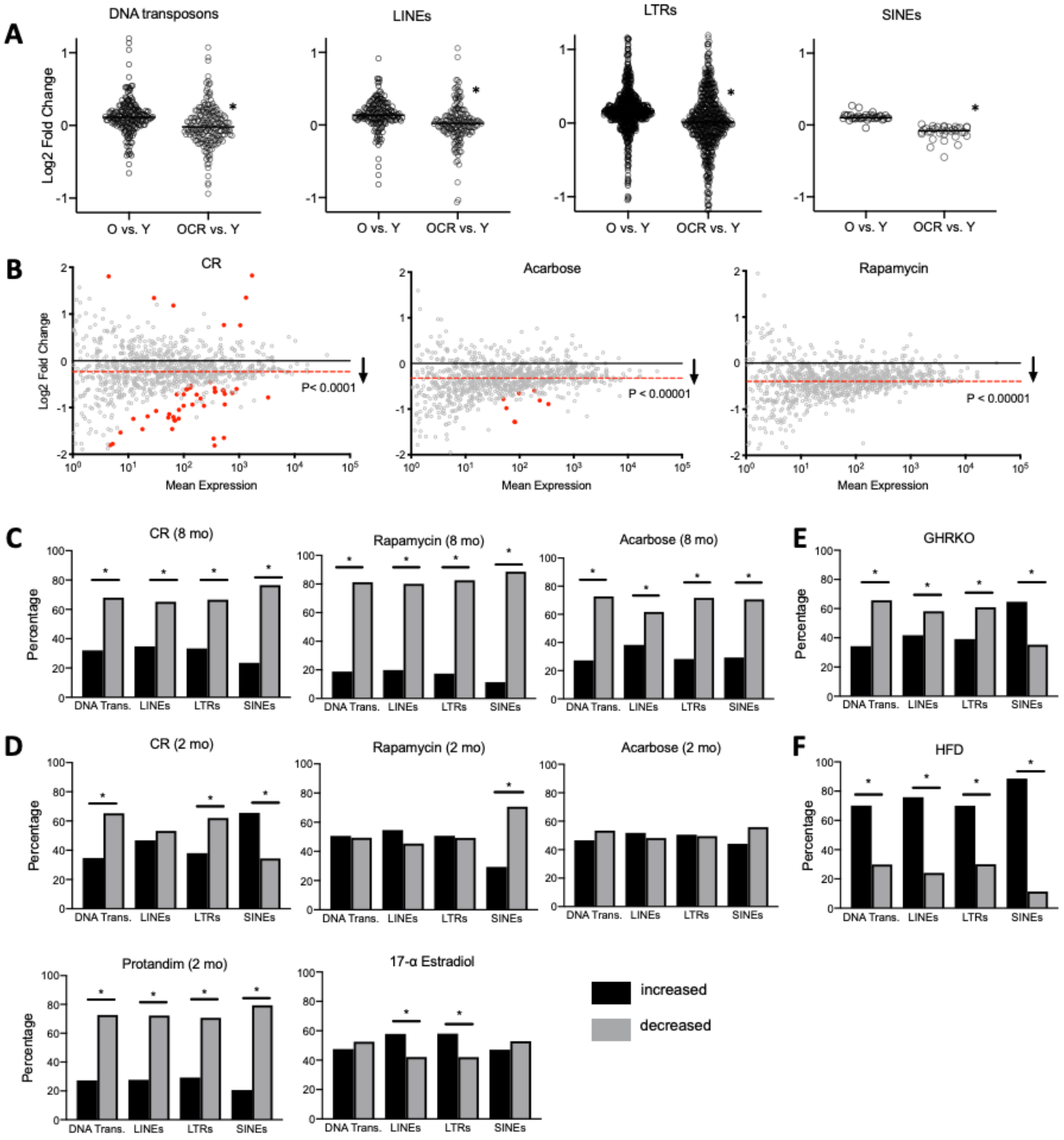
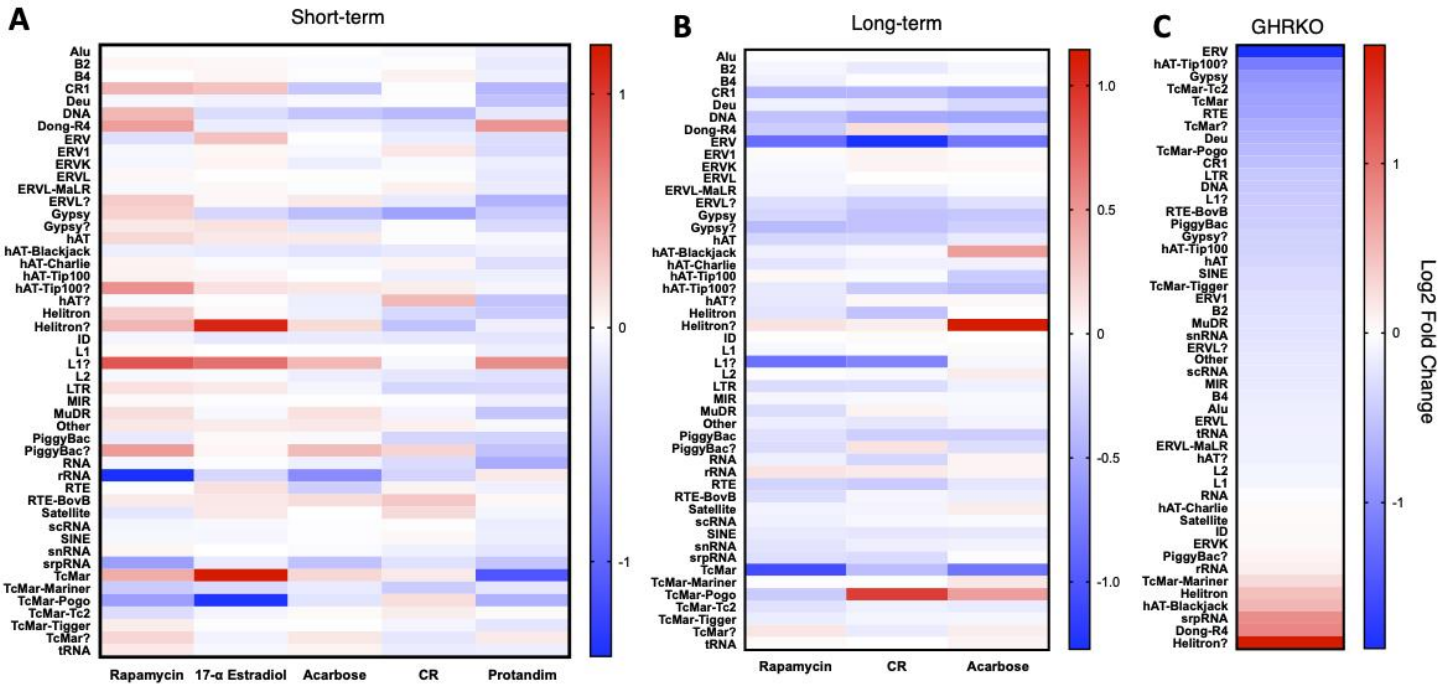


Figure 1. Healthy aging interventions reduce RE transcripts. (A) Age-related increases in the major types of RE transcripts in old (O) or old calorie-restricted (OCR) vs. young (Y) mice. N=3/group, *p<0.0001, two-tailed t-test. All individual RE shown. **(B)** MA plots showing significant decreases in global RE transcripts with long-term (8-month) healthy aging interventions. Significantly reduced or increased RE (FDR<0.1, identified using Deseq2) shown in red, and average transcript reduction indicated by red line. Likelihood of increased/decreased distribution calculated by chi-square analysis. **(C-F)** Percentage of RE transcripts by type increased or decreased with long- or short-term (2-month) interventions (*p<0.05, chi-square analysis). All relevant data/samples from datasets GSE92486, GSE131901 and GSE87565 were used for analyses (N=3-6 mice/group), and raw data are provided in the supplementary data file.



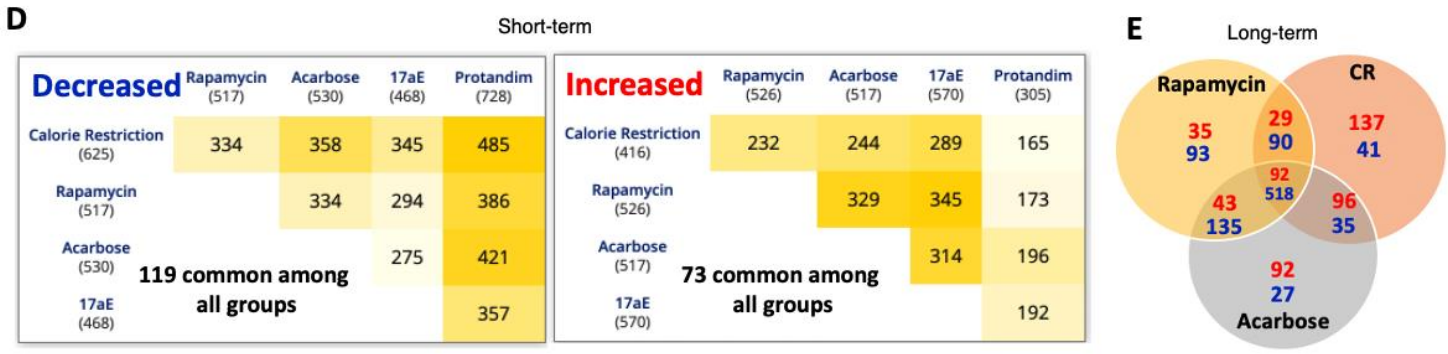


Figure 2. Healthy aging interventions reduce most RE families and many similar, individual RE transcripts. (A-C) Heatmaps showing RE transcript families increased (red) and decreased (blue) by short-term interventions, long-term interventions, and in GHRKO mice. **(D)** Pairwise intersections showing the number of common decreased or increased RE transcripts with short-term interventions. **(E)** Venn diagrams showing the number of common decreased (blue) or increased (red) RE transcripts with long-term interventions. All relevant data/samples from dataset GSE131901 were used for analyses (N=3-6 mice/group), and raw data are provided in the supplementary data file.

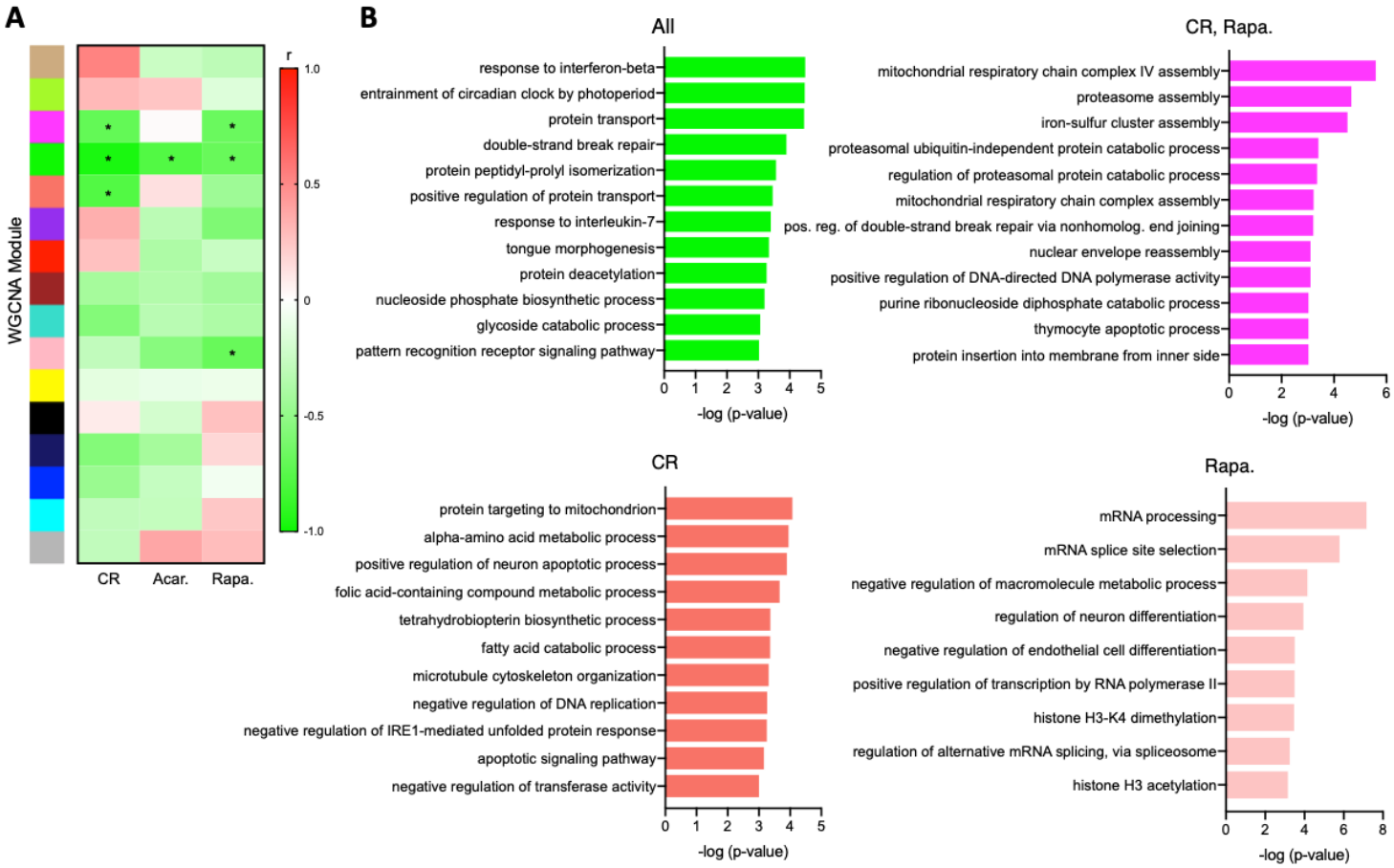


Figure 3. Gene expression patterns associated with RE transcript reductions. (A) WGCNA analysis heatmap of gene/RE modules influenced by long-term interventions (*modules that changed significantly with interventions, P<0.001 in WGCNA). **(B)** Most specific biological processes (gene ontology terms) in each significant WGCNA module. Exact p-values are noted in the supplementary data file, and all available samples/data from dataset GSE131901 were used for analyses (N=6 mice/group).

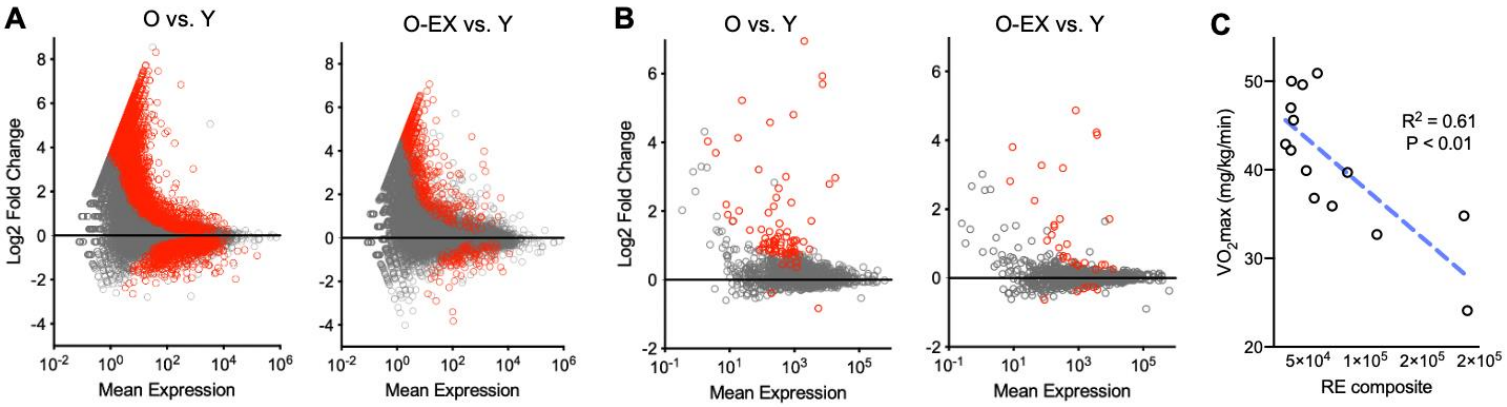


Figure 4. Long-term exercise is associated with reduced RE transcript expression in humans. **(A)** MA plots showing gene expression differences in peripheral blood mononuclear cells of older (O) vs. young (Y) sedentary, and older exercising (O-EX) vs. Y adults. N=4-5 matched samples per group. Note the ~2500 transcripts significantly increased/decreased with aging but largely reversed with exercise (red data points, FDR<0.1, identified using Deseq2). **(B)** MA plots showing RE transcript levels in the same samples/subjects. Note general upward shift and numerous significantly increased RE transcripts with aging (red, FDR<0.1, identified using Deseq2) that are largely reversed with exercise. **(C)** Correlation between maximal aerobic exercise capacity (VO_2 max) and composite count of RE transcripts significantly increased with aging (O vs. Y) in all subjects. Human subjects characteristics in supplementary data file.

References

1. Niccoli T, Partridge L. Ageing as a risk factor for disease. Curr Biol. 2012;22(17):R741-52.
2. Tyshkovskiy A, Bozaykut P, Borodinova AA, Gerashchenko MV, Ables GP, Garratt M, et al. Identification and Application of Gene Expression Signatures Associated with Lifespan Extension. Cell Metab. 2019;30(3):573-93.e8.
3. Wahl D, Anderson RM, Le Couteur DG. Anti-aging therapies, cognitive impairment and dementia. J Gerontol A Biol Sci Med Sci. 2019.
4. Lopez-Otin C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. Cell. 2013;153(6):1194-217.
5. Bodega B, Orlando V. Repetitive elements dynamics in cell identity programming, maintenance and disease. Curr Opin Cell Biol. 2014;31:67-73.
6. Bourque G, Burns KH, Gehring M, Gorbunova V, Seluanov A, Hammell M, et al. Ten things you should know about transposable elements. Genome Biol. 2018;19(1):199.
7. Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. Nat Rev Genet. 2009;10(10):691-703.
8. Goodier JL. Restricting retrotransposons: a review. Mob DNA. 2016;7:16.
9. Saleh A, Macia A, Muotri AR. Transposable Elements, Inflammation, and Neurological Disease. Front Neurol. 2019;10:894.
10. LaRocca TJ, Cavalier AN, Wahl D. Repetitive elements as a transcriptomic marker of aging: Evidence in multiple datasets and models. Aging Cell. 2020.
11. Maxwell PH. What might retrotransposons teach us about aging? Curr Genet. 2016;62(2):277-82.
12. Cardelli M. The epigenetic alterations of endogenous retroelements in aging. Mech Ageing Dev. 2018;174:30-46.
13. Wood JG, Jones BC, Jiang N, Chang C, Hosier S, Wickremesinghe P, et al. Chromatin-modifying genetic interventions suppress age-associated transposable element activation and extend life span in Drosophila. Proc Natl Acad Sci U S A. 2016;113(40):11277-82.
14. Hahn O, Gronke S, Stubbs TM, Ficz G, Hendrich O, Krueger F, et al. Dietary restriction protects from age-associated DNA methylation and induces epigenetic reprogramming of lipid metabolism. Genome Biol. 2017;18(1):56.
15. LaRocca TJ, Cavalier AN, Wahl D. Repetitive elements as a transcriptomic marker of aging: evidence in multiple datasets and models. Aging Cell. In Press.
16. De Cecco M, Criscione SW, Peterson AL, Neretti N, Sedivy JM, Kreiling JA. Transposable elements become active and mobile in the genomes of aging mammalian somatic tissues. Aging (Albany NY). 2013;5(12):867-83.
17. Kim DH, Bang E, Jung HJ, Noh SG, Yu BP, Choi YJ, et al. Anti-aging Effects of Calorie Restriction (CR) and CR Mimetics based on the Senoinflammation Concept. Nutrients. 2020;12(2).
18. Siersbaek M, Varticovski L, Yang S, Baek S, Nielsen R, Mandrup S, et al. High fat diet-induced changes of mouse hepatic transcription and enhancer activity can be reversed by subsequent weight loss. Sci Rep. 2017;7:40220.
19. Balestrieri E, Matteucci C, Cipriani C, Grelli S, Ricceri L, Calamandrei G, et al. Endogenous Retroviruses Activity as a Molecular Signature of Neurodevelopmental Disorders. Int J Mol Sci. 2019;20(23).
20. Meyer TJ, Rosenkrantz JL, Carbone L, Chavez SL. Endogenous Retroviruses: With Us and against Us. Front Chem. 2017;5:23.
21. Hedges DJ, Deininger PL. Inviting instability: Transposable elements, double-strand breaks, and the maintenance of genome integrity. Mutat Res. 2007;616(1-2):46-59.

22. Saldi TK, Gonzales PK, LaRocca TJ, Link CD. Neurodegeneration, Heterochromatin, and Double-Stranded RNA. J Exp Neurosci. 2019;13:1179069519830697.
23. Booth LN, Brunet A. The Aging Epigenome. Mol Cell. 2016;62(5):728-44.
24. Di Francesco A, Di Germanio C, Bernier M, de Cabo R. A time to fast. Science. 2018;362(6416):770-5.
25. Wahl D, Solon-Biet SM, Cogger VC, Fontana L, Simpson SJ, Le Couteur DG, et al. Aging, lifestyle and dementia. Neurobiol Dis. 2019;130:104481.
26. Huffman DM, Schafer MJ, LeBrasseur NK. Energetic interventions for healthspan and resiliency with aging. Exp Gerontol. 2016;86:73-83.
27. Widmann M, Niess AM, Munz B. Physical Exercise and Epigenetic Modifications in Skeletal Muscle. Sports Med. 2019;49(4):509-23.
28. Gano LB, Donato AJ, Pierce GL, Pasha HM, Magerko KA, Roeca C, et al. Increased proinflammatory and oxidant gene expression in circulating mononuclear cells in older adults: amelioration by habitual exercise. Physiol Genomics. 2011;43(14):895-902.
29. Strasser B, Burtscher M. Survival of the fittest: VO2max, a key predictor of longevity? Front Biosci (Landmark Ed). 2018;23:1505-16.
30. Sedivy JM, Kreiling JA, Neretti N, De Cecco M, Criscione SW, Hofmann JW, et al. Death by transposition - the enemy within? Bioessays. 2013;35(12):1035-43.
31. Jin Y, Tam OH, Paniagua E, Hammell M. TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. Bioinformatics. 2015;31(22):3593-9.
32. Criscione SW, Zhang Y, Thompson W, Sedivy JM, Neretti N. Transcriptional landscape of repetitive elements in normal and cancer human cells. BMC Genomics. 2014;15(1):583.
33. Zhang Y-J, Guo L, Gonzales PK, Gendron TF, Wu Y, Jansen-West K, et al. Heterochromatin anomalies and double-stranded RNA accumulation underlie C9orf72 poly(PR) toxicity. Science (New York, NY). 2019;363(6428).
34. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884-i90.
35. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.
36. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
37. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.
38. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009;10:48.
39. LaRocca TJ, Seals DR, Pierce GL. Leukocyte telomere length is preserved with aging in endurance exercise-trained adults and related to maximal aerobic capacity. Mech Ageing Dev. 2010;131(2):165-7.
40. LaRocca TJ, Mariani A, Watkins LR, Link CD. TDP-43 knockdown causes innate immune activation via protein kinase R in astrocytes. Neurobiol Dis. 2019;132:104514.
41. Bardou P, Mariette J, Escudié F, Djemiel C, Klopp C. jvenn: an interactive Venn diagram viewer. BMC Bioinformatics. 2014;15(1):293.