The opposite aging effect to single cell transcriptome profile among cell subsets.

Daigo Okada[1*]

1 Center for Genomic Medicine, Graduate School of Medicine, Kyoto University, Kyoto 606-8507, Japan

*** Correspondence: dokada@genome.med.kyoto-u.ac.jp**

Abstract

Comparing transcriptome profiling between younger and older samples reveals genes related to aging and provides insight into the biological functions affected by aging. Recent research has identified sex, tissue, and cell type-specific age-related changes in gene expression. This study reports the overall picture of the opposite aging effect, in which aging increases gene expression in one cell subset and decreases it in another cell subset. Using the Tabula Muris Senes dataset, a large public single-cell RNA sequencing dataset from mice, we compared the effects of aging in different cell subsets. As a result, the opposite aging effect was observed widely in genome-wide genes, particularly enriched in genes related to ribosomal function and translation. The opposite aging effect was observed in the known aging-related genes. Furthermore, the opposite aging effect was observed in the transcriptome diversity quantified by the number of expressed genes and the Shannon entropy. This study highlights the importance of considering the cell subset when intervening with aging-related genes.

Introduction

Comparing transcriptome profiles between younger and older samples is a powerful approach to identifying genes associated with aging. Comprehensive identification of age-related changes in gene expression can provide a detailed description of biological functions that are altered by aging. While changes in gene expression that are common to all cellular aging have been studied [9, 22, 30], sex [1, 12], tissue [25], and cell type [7, 30] specific changes in aging-related gene expression have been identified.

Recent studies on expression quantitative trait loci (eQTL) have reported cases in which the effects of genetic factors on gene expression act in opposite directions in different cell subsets. For example, the eQTL with opposite positive and negative effects on gene expression levels between different tissues has been reported and called the opposite eQTL effect [16]. The integrated analyses of single-cell RNA-seq (scRNA-seq) and genomic data have also identified eQTLs with opposite effects on different cell types [19, 29]. These findings suggest that the effect of genetic perturbation on the transcriptome is cell subset-dependent.

In this study, we analyzed a large public dataset of mouse scRNA-seq and reported the landscape of opposite effects of aging among cell subsets on the transcriptome. The report of this study highlights the importance of considering the cell subsets when intervening with aging-related genes.

Method

Data and preprocessing

We analyzed data from the Tabula Muris Senes dataset, which is a large public single-cell RNA-seq dataset from mice of different ages [5]. We used the TabulaMurisSenisData package in R to obtain pre-processed data after log transformation [23]. The Tabula Muris Senes project provided two different mouse single-cell RNA-seq datasets from different experimental designs (FACS or droplet). FACS dataset contains the gene expression data of 22966 genes $\times$ 110824 cells and Droplet dataset contains that of 20138 genes $\times$ 245389 cells. We analyzed each of the two data sets to validate the stability of our results. Cells derived from 3-month-old mice were defined as the young group, and cells derived from 18-month-old or older mice were defined as the old group. In this analysis, cells derived from different cell types, sexes, and tissues were treated as different cell subsets. Cell subsets containing more than 25 cells in both young and old groups were used for subsequent analysis. Finally, 189 cell subsets in the FACS dataset and 99 cell subsets in the droplet dataset were used in this analysis. A list of cell types used can be found in Supplementary Table 1.

Quantifying the opposite aging effect on gene expression

In each cell subset, we applied Welch's t-test to the gene expression levels in the pre-processed scRNA-seq dataset to compare the differential expression of each gene between the old and young groups. We applied Bonferroni correction of P values to assess statistical significance and counted the number of cell subsets for which a significant increase or decrease was observed for each gene. In this step, Bonferroni correction was performed for each of the datasets, and the total number of tests is the number of all cell subsets and gene combinations.

We defined the Opposite Effect Score to detect the situation where aging increases gene expression in one cell subset and decreases it in another cell subset. Let N_{up} and N_{down} be the number of cell subsets where the expression level of a gene increased or decreased with aging, respectively. Here, the following Opposite Effect Score can be an indicator of the strength of the opposite aging effect for the gene.

$$\text{Opposite Effect Score} = \frac{min(N_{up}, N_{down})}{\text{total number of subsets}}$$

Genes for which N_{up} and N_{down} were both > 0 were considered as genes with the opposite aging effect. Genes for which $N_{up} > 0$ and $N_{down} = 0$ were considered as "UP-only genes". Genes with $N_{down} > 0$ and $N_{up} = 0$ were considered "DOWN-only genes". We performed Gene Ontology (Biological Process) enrichment analysis on genes with Opposite Effect Score > 0.05 using the WebGestaltR package in R [11]. The database was set to geneontology_Biological_Process_noRedundant. The false discovery rate (FDR) threshold was set to 0.05.

Evaluation of the opposite aging effect on transcriptome diversity

We investigated the opposite aging effect on transcriptome diversity. The simplest measure of transcriptome diversity is the number of genes expressed in a cell. How the number of expressed genes changes with age is controversial. For example, while the number of expressed genes decreases with age in mouse oocytes [3], it increases in mouse blood [2].

Another method to quantify the transcriptome diversity is Shannon entropy [14]. This method can take into account not only zero or non-zero expression of genes but also continuous expression level.

Transcriptome diversity based on Shannon entropy is calculated as follows. The normalization of the sum of expression levels to 1 can be treated as the relative frequency of transcription of genes. Let p_i be the relative frequency of transcription of the i-th gene in a single cell ($i = 1, 2, ., N$), the diversity (H) of the transcriptome in each tissue can be quantified by applying the Shannon entropy equation as

$$H = -\sum_{i=1}^{N} p_i \log_2 p_i$$

Transcriptome diversity based on Shannon entropy is associated with biological variables such as cancer [15]. It can also be used to search for optimal drug concentrations in cell culture [20]. Recently, it is also used in single-cell RNA-seq data analysis [6, 10]. Although it can be used as a good biological feature of the overall transcriptome profile, the relationship between Shannon entropy-based transcriptome diversity and aging is unclear.

We calculated the log2 of the number of expressed genes and Shannon entropy for single cells and examined their age-related change for each cell subset. We evaluated the difference in transcriptome diversity scores between the young and old groups for each cell subset using Welch's t-test. Under the Bonferroni criterion, the threshold for the P-value was set at 0.05 / the number of subsets. We also calculated Cohen's D, a representative effect size measure of the difference in means between groups, using the effsize package in R [27].

Result

We analyzed the Tabula Muris Senes dataset to investigate the landscape of the opposite aging effect on gene expression among cell subsets. Figure 1A shows a scatterplot of the number of cell subsets with increased (Up) and decreased (Down) expression. While there are genes that show only up or down, we also find genes that show reversed aging changes in multiple subsets. We observed the opposite aging effect on gene expression in 15.1 % and 9.7 % of the genes in two datasets, respectively (Figure 1B). The results suggest opposite aging effect was widely observed in genome-wide gene expression profiles.

The histogram of opposite effect scores of genome-wide genes suggests that a strong opposite aging effect is observed in some genes (Figure 2A). The genes with the top ten opposite effect scores are shown in Figure 2B. In Gene Ontology analysis, the 14 and 18 Gene ontology terms were significantly enriched (FDR < 0.05) (Figure 3). Among these terms, GO:0022613 (ribonucleoprotein complex biogenesis) and GO:0002181 (cytoplasmic translation) were the gene ontology terms with the top two enrichment ratio in both datasets. These results suggest that the direction of age-related change of ribosome-associated gene expression depends on cell subsets. The N_{up}, N_{down}, and opposite effect scores for all genes in each dataset are available in Supplementary Tables S2 and S3, respectively.

We also examined the presence or absence of the opposite aging effect in the known 183 aging-associated genes obtained from the GenAge database [26]. In 37.5 and 21.2 % of the GenAge genes, the opposite aging effect was observed in two datasets, respectively (Figure 4A). In the 59 genes (32.2 %), the opposite aging effect was observed in at least one data set. The genes with the top ten opposite aging scores are

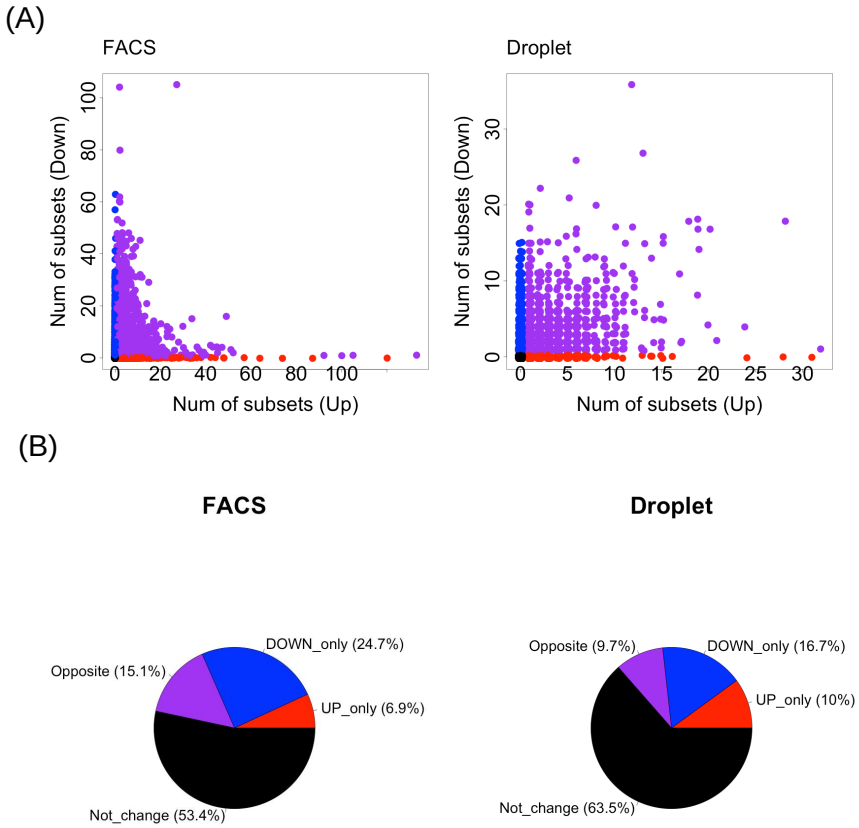

Figure 1. (A)Scatterplots of the number of subsets whose gene expression increased (N_{up}) or decreased (N_{down}) with aging for genome-wide genes. The genes of "UP only " ($N_{up} > 0, N_{down} = 0$), "DOWN only" ($N_{up} = 0, N_{down} > 0$), "Opposite" ($N_{up} > 0, N_{down} > 0$), "Not change" ($N_{up} = 0, N_{down} = 0$) are colored by red, blue and purple, respectively. Data points were jitterd to prevent points from being plotted on top of each other. (B)Percentage of the genes with "UP only", "DOWN only", "Opposite", and "Not change" among genome-wide genes.

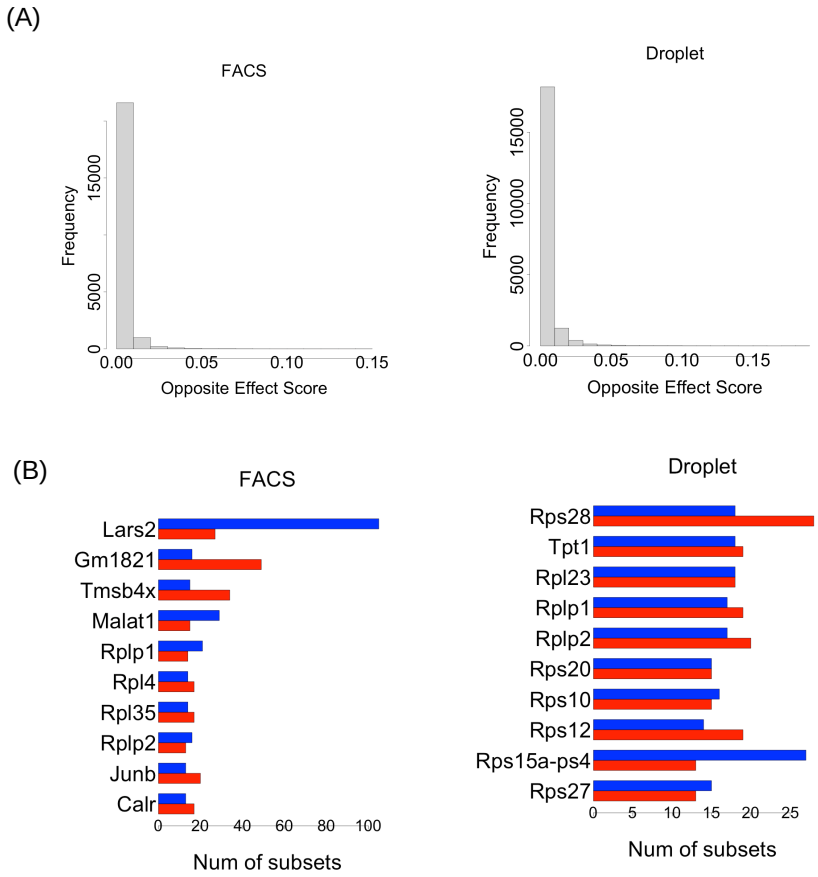

Figure 2. (A) Histogram of opposite effect score for genome-wide genes. (B)The genes with the top ten opposite effect scores. The red bar represents the number of subsets whose gene expression increased (N_{up}). The blue bar represents the number of subsets whose gene expression decreased (N_{down}).

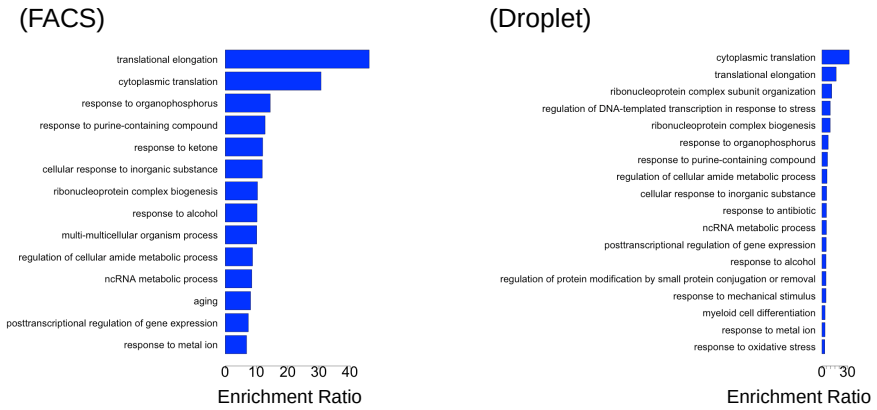

Figure 3. The Results of enrichment analysis to Gene Ontology (Biological Process) for the gene sets with the opposite effect score > 0.05. The X-axis represents the enrichment ratio of each gene ontology term.

shown in Figure 4B. *Lmna*, *Mt1*, and *Cebpb* were in the top ten in both data sets and their expression has been shown in Figure 5 as an example. *Lmna* encodes a nuclear envelope protein involved in maintaining nuclear structure and regulating transcription and is known to be related to aging and longevity [8, 13]. *Mt1* is a gene involved in metal ion metabolism, and its cardiac-specific over-expression extends lifespan [28]. *Cebpb* is a gene involved in fat metabolism, and replacement of the *Cebpa* gene with *Cebpb* is reported to extend life span [4].

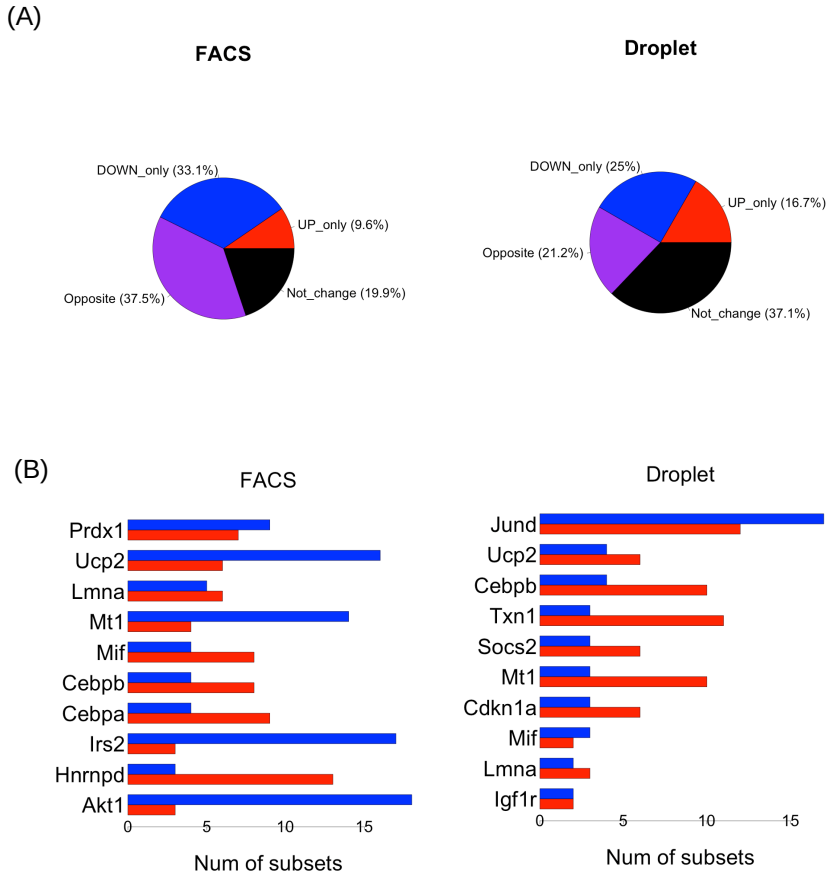

Figure 4. opposite aging effect on the known mouse age-related genes registered in GenAge database. (A)Percentage of the genes with "Up only" ($N_{up} > 0, N_{down} = 0$), "Down only" ($N_{up} = 0, N_{down} > 0$), "Opposite" ($N_{up} > 0, N_{down} > 0$), and "Not change" ($N_{up} = 0, N_{down} = 0$) among GenAge genes. (B) The GenAge genes with the top ten opposite effect scores.The red bar represents the number of subsets whose gene expression increased (N_{up}). The blue bar represents the number of subsets whose gene expression decreased (N_{down}).

Figure 6 shows a volcano plot of age-related change in transcriptome diversity. The opposite aging effect was observed in transcriptome diversity. The number of cell subsets with decreasing transcriptome diversity with aging was greater than that of the subset with increasing transcriptome diversity. This observation is true for both FACS and droplet datasets and on both the number of expressed genes and Shannon entropy. The results suggest that not only the direction of expression of individual genes but also the direction of aging changes in transcriptome diversity depends on the cell subset. The P-values and Cohens'D value for the analysis in each subset in the two data sets

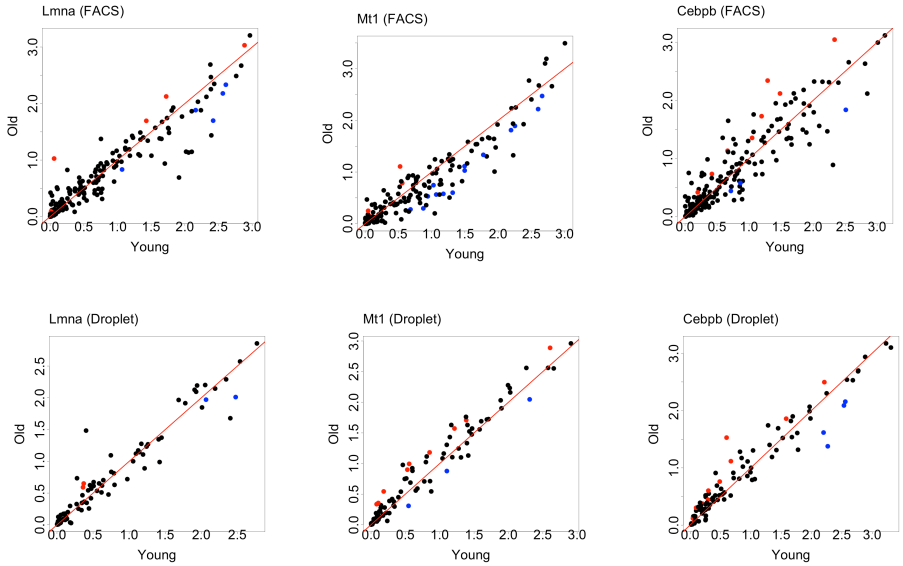

Figure 5. Scatterplots of mean expression in the young group (x-axis) and old group (y-axis) for *Lmna*, *Mt1*, and *Cebpb* in the all cell subsets. Cell subsets with statistically significant differences identified are plotted in red (Up) or blue (Down). The red line represents y=x.

are available in Supplementary Tables S4 - S7.

Discussion

In this study, we analyzed the Tabala Murirs Senis dataset and reported the landscape of the opposite aging effect on transcriptome profiles among cell subsets. The opposite aging effect was widely observed in the aging changes both in gene expression level and in the transcriptome diversity quantified by the number of expressed genes and the Shannon entropy.

The opposite aging effect on transcriptome can reflect a deviation from the optimal gene expression state in each cell subset. The optimal gene expression state for each cell subset is maintained in youth, but shifts to an abnormal state with age may be responsible for the dysfunctions associated with aging. The changes in gene expression associated with aging can be described as a drift in an arbitrary direction rather than an active drift in a specific direction. Gene expression is optimally regulated in each cell subset by gene regulatory networks. Each cell subset has a different steady state of gene expression due to different epigenomic states and external environments. It has been suggested that the disturbance of gene expression not only due to genetic effects by the opposite eQTL effect but also due to aging is dependent on biological conditions.

The opposite aging effect was widely observed in the genome-wide genes. In particular, the opposite aging effect enriched the genes associated with basic cellular functions such as ribosomal function and translation. It has been known that ribosomal function decreases with aging [17, 18, 21, 24]. Destabilization of the regulation of ribosome-related gene expression, represented by the oppoisite aging effect, may lead to a decrease in cellular ribosomal function.

In addition, about 30% of the known aging-related genes were observed in either of the two large data sets. If different subsets differ in the direction of aging changes in gene expression, then over-expression of the gene leads to repress aging-related changes

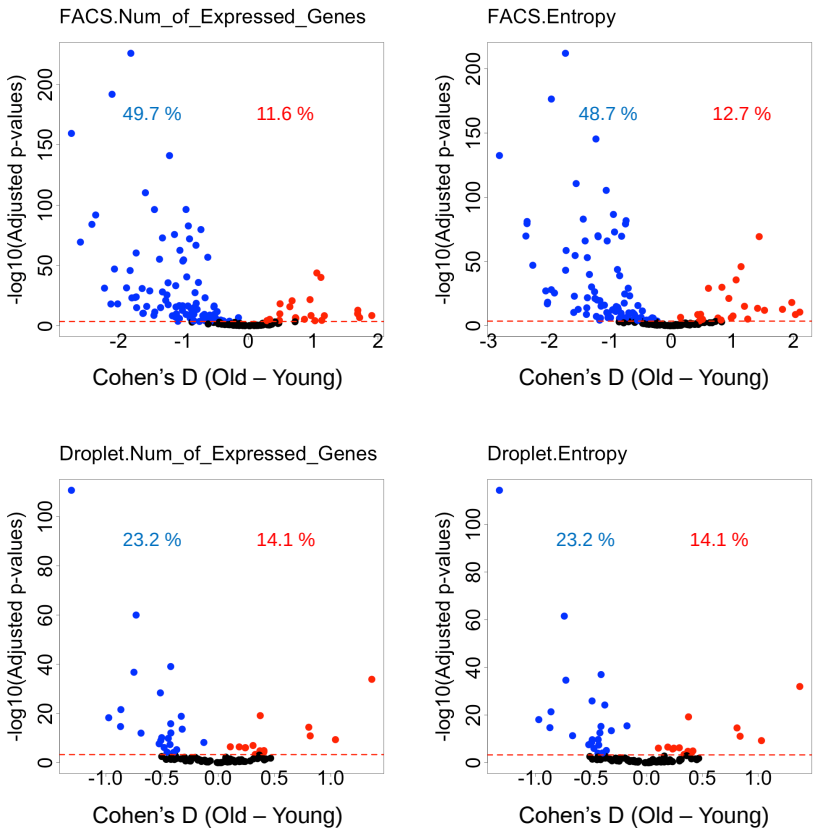

Figure 6. Volcano plot of age-related changes in transcriptome diversity quantified with log2(the number of expressed genes) or Shannon entropy. Each point represents one cell subset. The x-axis is the effect size of the difference between the old group and the young group as quantified by Cohen's D. The y-axis represents -log10 (P value). Cell subsets for which statistically significant differences were detected are plotted in red (Old > Young) or blue (Old < Young) with percentages.

in one cell subset but rather accelerete changes in another subset. This study highlights the importance of considering the cell subset when intervening with aging-related genes.

Conclusion

This study reports the overall picture of the opposite aging effect, in which aging increases gene expression in one cell subset and decreases it in another cell subset. The opposite aging effect among cell subsets is widely observed in the age-related change of each gene expression level and transcriptome diversity.

Declarations

Funding

Not applicable.

Conflict of interest

The author has no conflict of interest to be declared.

Ethics approval

Not applicable.

Data and Code availability

This study used public data only. The code used in this study is available at (https://github.com/DaigoOkada/oppagingeffect).

References

1. N. C. Berchtold, D. H. Cribbs, P. D. Coleman, J. Rogers, E. Head, R. Kim, T. Beach, C. Miller, J. Troncoso, J. Q. Trojanowski, et al. Gene expression changes in the course of normal brain aging are sexually dimorphic. *Proceedings of the National Academy of Sciences*, 105(40):15605–15610, 2008.

2. M. J. Bonder, S. J. Clark, F. Krueger, S. Luo, J. A. de Sousa, A. M. Hashtroud, T. M. Stubbs, A.-K. Stark, S. Rulands, O. Stegle, et al. Single cell dna methylation ageing in mouse blood. *bioRxiv*, pages 2023–01, 2023.

3. J. Castillo-Fernandez, E. Herrera-Puerta, H. Demond, S. J. Clark, C. W. Hanna, M. Hemberger, and G. Kelsey. Increased transcriptome variation and localised dna methylation changes in oocytes from aged mice revealed by parallel single-cell analysis. *Aging Cell*, 19(12):e13278, 2020.

4. C.-H. Chiu, W.-D. Lin, S.-Y. Huang, and Y.-H. Lee. Effect of a c/ebp gene replacement on mitochondrial biogenesis in fat cells. *Genes & development*, 18(16):1970–1975, 2004.

5. T. T. M. Consortium. A single-cell transcriptomic atlas characterizes ageing tissues in the mouse. *Nature*, 583(7817):590–595, 2020.

6. O. Gandrillon, M. Gaillard, T. Espinasse, N. B. Garnier, C. Dussiau, O. Kosmider, and P. Sujobert. Entropy as a measure of variability and stemness in single-cell transcriptomics. *Current Opinion in Systems Biology*, 27:100348, 2021.

7. S. Ham and S.-J. V. Lee. Advances in transcriptome analysis of human brain aging. *Experimental and Molecular Medicine*, 52:1787 – 1797, 2020.

8. H.-T. Huang, D. Wade, D. Bilbao, G. S. Gaidosh, R. Shiekhattar, and M. E. Figueroa. Age-acquired downregulation of lmna leads to epigenetic deregulation and altered hspc function. *Blood*, 138:3280, 2021.

9. A. Jaffe, J. Shin, L. Collado-Torres, J. Leek, R. Tao, C. Li, Y. Gao, Y. Jia, B. Maher, T. Hyde, J. Kleinman, and D. Weinberger. Developmental regulation of human cortex transcription and its clinical relevance at base resolution. *Nature neuroscience*, 18:154 – 161, 2014.

10. S. Kannan, M. Farid, B. L. Lin, M. Miyamoto, and C. Kwon. Transcriptomic entropy benchmarks stem cell-derived cardiomyocyte maturation against endogenous tissue at single cell level. *PLoS computational biology*, 17(9):e1009305, 2021.

11. Y. Liao, J. Wang, E. J. Jaehnig, Z. Shi, and B. Zhang. Webgestalt 2019: gene set analysis toolkit with revamped uis and apis. *Nucleic acids research*, 47(W1):W199–W205, 2019.

12. D. Liu, M. A. Sartor, G. A. Nader, E. E. Pistilli, L. Tanton, C. Lilly, L. Gutmann, H. B. IglayReger, P. S. Visich, E. P. Hoffman, et al. Microarray analysis reveals novel features of the muscle aging process in men and women. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 68(9):1035–1044, 2013.

13. I. C. Lopez-Mejia, M. De Toledo, C. Chavey, L. Lapasset, P. Cavelier, C. Lopez-Herrera, K. Chebli, P. Fort, G. Beranger, L. Fajas, et al. Antagonistic functions of lmna isoforms in energy expenditure and lifespan. *EMBO reports*, 15(5):529–539, 2014.

14. O. Martínez and M. H. Reyes-Valdés. Defining diversity, specialization, and gene specificity in transcriptomes through information theory. *Proceedings of the National Academy of Sciences*, 105(28):9709–9714, 2008.

15. O. Martínez, M. H. Reyes-Valdés, and L. Herrera-Estrella. Cancer reduces transcriptome specialization. *PloS one*, 5(5):e10398, 2010.

16. A. Mizuno and Y. Okada. Biological characterization of expression quantitative trait loci (eqtls) showing tissue-specific opposite directional effects. *European Journal of Human Genetics*, 27(11):1745–1756, 2019.

17. N. Mori, D. Mizuno, and S. Goto. Conservation of ribosomal fidelity during ageing. *Mechanisms of Ageing and Development*, 10:379–398, 1979.

18. K. Nakano and H. Sidransky. Age-related changes in ribosomal profiles and in vitro protein synthesis in skeletal muscle during fasting and subsequent refeeding of rats. *The Journal of nutrition*, 108(3):399–409, 1978.

19. A. Nathan, S. Asgari, K. Ishigaki, C. Valencia, T. Amariuta, Y. Luo, J. I. Beynor, Y. Baglaenko, S. Suliman, A. L. Price, et al. Single-cell eqtl models reveal dynamic t cell state dependence of disease loci. *Nature*, 606(7912):120–128, 2022.

20. N. Ogata, T. Kozaki, T. Yokoyama, T. Hata, and K. Iwabuchi. Comparison between the amount of environmental change and the amount of transcriptome change. *PLoS One*, 10(12):e0144822, 2015.

21. S. I. Rattan. Synthesis, modifications, and turnover of proteins during aging. *Experimental gerontology*, 31(1-2):33–47, 1996.

22. M. B. Sleiman, P. Jha, R. Houtkooper, R. W. Williams, X. Wang, and J. Auwerx. The gene-regulatory footprint of aging highlights conserved central regulators. *Cell Reports*, 32, 2020.

23. C. Soneson, D. Machlab, F. Marini, and S. Astrologo. *TabulaMurisSenisData: Bulk and single-cell RNA-seq data from the Tabula Muris Senis project*, 2023. R package version 1.8.0.

24. M. J. Stec, D. Mayhew, and M. Bamman. The effects of age and resistance loading on skeletal muscle ribosome biogenesis. *Journal of applied physiology*, 119 8:851–7, 2015.

25. R. Stegeman and V. Weake. Transcriptional signatures of aging. *Journal of molecular biology*, 429 16:2427–2437, 2017.

26. R. Tacutu, D. Thornton, E. Johnson, A. Budovsky, D. Barardo, T. Craig, E. Diana, G. Lehmann, D. Toren, J. Wang, et al. Human ageing genomic resources: new and updated databases. *Nucleic acids research*, 46(D1):D1083–D1090, 2018.

27. M. Torchiano. *effsize: Efficient Effect Size Computation*, 2020. R package version 0.8.1.

28. X. Yang, T. A. Doser, C. X. Fang, J. M. Nunn, R. Janardhanan, M. Zhu, N. Sreejayan, M. T. Quirin, J. Ren, X. Yang, et al. Metallothionein prolongs survival and antagonizes senescence-associated cardiomyocyte diastolic dysfunction: role of oxidative stress. *The FASEB Journal*, 20(7):1024–1026, 2006.

29. S. Yazar, J. Alquicira-Hernandez, K. Wing, A. Senabouth, M. G. Gordon, S. Andersen, Q. Lu, A. Rowson, T. R. Taylor, L. Clarke, et al. Single-cell eqtl mapping identifies cell type–specific genetic control of autoimmune disease. *Science*, 376(6589):eabf3041, 2022.

30. M. J. Zhang, A. O. Pisco, S. Darmanis, and J. Zou. Mouse aging cell atlas analysis reveals global and cell type-specific aging signatures. *Elife*, 10:e62293, 2021.