A tissue ubiquitous gene set for cellular senescence

Yilong Qu[1], Zhicheng Ji[1,*]

[1]Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, NC, USA
[*]Corresponding author. E-mail: zhicheng.ji@duke.edu

Abstract

An accurate gene set for cellular senescence is crucial for identifying and studying senescent cells in single-cell RNA-seq datasets. We integrated nine existing senescence gene sets and identified a core senescence gene set comprising four genes: CDKN1A, CDKN2A, IL6, and CDKN2B. We found that these genes are ubiquitously associated with cellular senescence across human and mouse tissues. Using this gene set, we identified cell types enriched with senescent cells and cell-cell communication targets and pathways associated with cellular senescence in human and mouse single-cell datasets.

Main

Cellular senescence is a fundamental biological process in which cells enter a state of permanent cell cycle arrest, losing their ability to proliferate[1,2]. This state serves primarily as a defense mechanism against the uncontrolled cell division observed in cancer. However, the accumulation of senescent cells over time may contribute to the decline in physical and functional integrity of tissues and organs, playing a significant role in aging[3,4]. Additionally, senescent cells are implicated in the development of various age-related diseases, such as osteoarthritis, pulmonary fibrosis, and Alzheimer's disease[5-8]. These cells release pro-inflammatory proteins that can alter the cellular microenvironment, affecting neighboring cells and contributing to chronic inflammation[9]. Given its profound impact on human health and potential as a therapeutic target, it is important to understand cellular senescence and identify its associated biomarkers.

Single-cell RNA-sequencing (scRNA-seq) technologies hold enormous potential for studying the molecular characteristics of senescent cells and their interactions with other cells. However, classical biomarkers for cellular senescence, such as senescence-associated β-galactosidase (SA-β-gal), require cytochemical staining[10] and cannot be directly applied to data from high-throughput sequencing. Instead, cellular senescence in scRNA-seq is often identified by examining whether genes in a senescence gene set (SnG) are highly expressed in a cell[11,12]. Therefore, the reliability of the SnG significantly impacts the

accuracy of identifying senescent cells. Several SnGs have been compiled in previous studies, including CSgene[13], CellAge[14], SeneQuest[2], GenAge[15], SenMayo[12], the GO term “Cellular Senescence”[16], the KEGG pathway “Cellular Senescence"[17], and the Reactome pathways of “cellular senescence” and “Senescence Associated Secretory Phenotypes (SASP)”[18]. Strikingly, we found significant discrepancies across these SnGs. Larger SnGs such as CellAge and CSGene identify many more genes than smaller SnGs such as the GO term and SenMayo (Figure 1a). The Jaccard index, which measures the degree of overlap between two SnGs, is extremely low except for the one between two SnGs by Reactome pathways (Figure 1b). Out of the 1,855 senescence-related genes identified by any of the SnG, 1,115 (60.1%) were reported only in one SnG, and only 44 (2.4%) were reported in at least five SnGs (Figure 1c). A potential reason is that these SnGs were compiled mainly through literature curation by different research groups, and the discrepancy could result from varying scopes and inclusion criteria in the literature search process. The high level of discrepancy indicates limited reliability for a large proportion of genes in each SnG. Moreover, many senescence genes are specific to one tissue type[19], and existing studies do not clarify if these SnGs can be applied generally across various tissue types. For instance, SenMayo[12] was evaluated only in brain and bone marrow tissues. Additionally, senescence-associated genes can either induce or inhibit cellular senescence. This direction is crucial for accurately evaluating the level of cellular senescence but is not recorded by most SnGs. A reliable and tissue-ubiquitous SnG is still needed to facilitate the identification of cellular senescence in scRNA-seq datasets across various tissues.

To address this issue, we performed a meta-analysis of the nine existing SnGs to identify a reliable and tissue-ubiquitous gene set for cellular senescence (Supplementary Table S1). Among genes reported by more than half of the nine gene sets, four genes were reported by eight out of the nine gene sets, including CDKN1A, CDKN2A, IL6, and CDKN2B (Figure 1d). These genes were reported to induce senescence by both CellAge and SeneQuest, two gene sets that recorded the direction of senescent genes. Among these, CDKN1A and CDKN2A, which encode the p21 and p16 proteins respectively, are well-known markers of senescence[20]. Although CDKN1A and CDKN2A were not included in the SenMayo gene set, they were treated as gold-standard senescence genes and excluded for evaluation purposes. For genes that appear in at least five gene sets, almost all show consistent directions across SnGs (Figure 1d).

We then systematically evaluated the reliability and tissue ubiquity of senescent genes reported by different numbers of SnGs in human and mouse. While it is not possible to directly assess a gene’s association with cellular senescence across various tissues due to the lack of experimental techniques and data, we used the gene’s association with age as the primary evidence, given that cellular senescence is a

hallmark of aging[4]. We collected GTEx[21] bulk RNA-seq data from 30 human tissues and performed a correlation analysis between gene expression levels and ages of samples within each tissue (Figure 1e, Methods, Supplementary Table S2). For each gene, we calculated the proportion of tissues with positive associations with age and the proportion with negative associations, and selected the greater of these two proportions. (Figure 2a). We also calculated the proportion of tissues with significant associations with age (Figure 2b). For genes with consistent senescence direction (induction or inhibition) across all SnGs, we calculated the proportion of tissues where the signs of age associations agree with the senescence direction reported by the SnGs (Figure 2c). Additionally, senescent cells are enriched with cells in the G1 phase[12] since senescence inhibits cells from entering the S or M phases of the cell cycle[22]. Therefore, as the secondary evidence, we collected scRNA-seq data from 24 human tissues from the Tabula Sapiens[23] and 23 mouse tissues from the Tabula Muris[24], and calculated the enrichment of cells in the G1 phase, comparing cells expressing or not expressing a gene within each tissue (Methods, Figure 2d). Finally, since proteins with similar functions are more likely to interact with each other[25], genes associated with cellular senescence are likely to have higher protein-protein interaction (PPI) scores in a PPI network. As the tertiary evidence, we calculated PPI scores for genes appearing in different numbers of SnGs for both human and mouse using the STRING PPI database[26] (Methods, Figure 2e).

Figures 2a-e demonstrate that genes reported by a greater number of SnGs have higher values in all five metrics, suggesting that genes with more substantial literature support exhibit greater reliability and tissue ubiquity for cellular senescence. Specifically, genes reported by at least eight SnGs, which are all senescence-inducing (Figure 1d), show positive associations with age in an average of 84.9% human tissues (Figure 2a,2c), are significantly associated with age in an average of 55.4% human tissues (Figure 2b), have averaged log odds ratios for G1 cell enrichment of 0.659 in human and 0.601 in mouse (Figure 2d), and have PPI scores of 5 in both human and mouse (Figure 2e). These values are significantly higher than those for genes reported by no or one gene set (Figure 2a-d). In comparison, genes reported by fewer than eight SnGs do not always show significantly higher values in these metrics (Figures 2a-d). Such evidence suggests that the four genes reported by at least eight SnGs, namely CDKN1A, CDKN2A, IL6, and CDKN2B, constitute a core senescence gene set that can be broadly applied to different tissue types in both human and mouse.

Using this core senescence gene set, we systematically identified and studied senescent cells (SnCs) in the Tabula Sapiens and Tabula Muris scRNA-seq data. We defined SnCs as cells expressing at least three of the four core senescence genes, since CDKN2A is expressed at low levels in single-cell data, even in senescent cells[12]. The proportion of SnCs within each cell type is highly conserved between human and

mouse (correlation=0.804, Figures 2g-h) and are relatively low (Supplementary Table S3). In human, 79.2% of tissues and 80% of cell types have SnC proportions below 5%. In mouse, 95.7% of tissues and 100% of cell types have SnC proportions below 5%. SnCs are most enriched in fibroblasts in both human and mouse[27], with senescent fibroblasts present in 77.8% of human tissues and 75% of mouse tissues. Other cell types, such as epithelial cells and stromal cells, were also reported in the literature as being associated with senescence[12,27]. We then analyzed cell-cell communications between SnCs and other cell types with CellChat[28]. Figure 2i-j show example interaction plots in the human trachea tissue. T cells and epithelial cells are among the top interaction targets of SnCs[12,29] (Figure 2k), consistent across human and mouse (correlation=0.417, Figure 2l, Supplementary Table S4). Several key pathways known as hallmarks of cellular senescence are consistently enriched across tissues (Figure 2m), including COLLAGEN, which is responsible for extracellular matrix remodeling and enlarged phenotypes of senescent cells[30-32]; LAMININ, corresponding to the loss of laminin B1 that compromises nuclear integrity in senescent cells[1,33]; MHC-I, which is elevated in senescent cells[34,35]; and MIF, which is involved in senescence-associated inflammation and oxidative stress response pathways[36,37]. These pathways are again consistent across human and mouse (correlation=0.816, Figure 2n, Supplementary Table S5).

In summary, we compiled a tissue-ubiquitous gene set associated with cellular senescence and extensively evaluated it across 41 tissues from both human and mouse. We identified cell types enriched with senescent cells and cell-cell communication targets and pathways associated with cellular senescence, with conservation observed across human and mouse tissues. This gene set offers a convenient tool for identifying senescent cells in scRNA-seq data, particularly in tissues and cell types with limited prior knowledge of cellular senescence.

Methods

Curation of senescence gene sets

Gene sets of CSgene and SenMayo were downloaded from the original publications. Gene sets of CellAge and GenAge were downloaded from the Human Ageing Genomic Resources (HAGR)[38]. GO term of Cellular Senescence was downloaded from MSigDB[39,40]. Reactome terms of Cellular Senescence and Senescence Associated Secretory Phenotype were downloaded from R package ReactomeContentService4R version 1.10.0[41]. KEGG pathway of Cellular Senescence was downloaded from R package KEGGREST version 1.42.0[42]. Gene set of SeneQuest was downloaded from the

SeneQuest database[4], and a gene was retained if it had at least 15 publication records and at least 80% of the publication records had an agreeing direction of senescence induction or inhibition. The complete assembly of SnGs can be found in Supplementary table S1. To apply these gene sets to mouse bulk and single-cell RNA-seq analysis, genes were converted to their mouse homologs by the Mouse Genome Informatics database (MGI)[43].

GTEx bulk RNA-seq analysis

Gene expression matrices of Genotype-Tissue Expression (GTEx) bulk RNA-seq were downloaded from GTEx portal[21]. Tissues with less than 100 samples were filtered out. Age information of each sample, which was provided as an interval by GTEx, was converted into a numerical value by averaging the two end points of the age interval. Within each tissue, genes with >1 log2(TPM + 1) values in at least 10% of the samples were retained. Linear regression was performed within each tissue using R package limma version 3.54.2[44]. The response variables were the gene expression raw counts, and the independent variables were age and sex of samples. P-values were adjusted for multiple testing by the BH procedure[45], and adjusted p-values less than 0.05 were considered to be statistically significant.

Single-cell RNA-seq analysis

Single-cell RNA-seq datasets of Tabula Sapiens[23] and Tabula Muris[24] dataset were downloaded from the original publications. Single-cell data were analyzed using Seurat version 4.3.0. Specifically, raw counts were log-normalized by the "NormalizeData" function. The original cell type annotations of both datasets were reannotated to have a matching level of annotations (Supplementary Table S6). Cell cycle was determined using the "CellCycleScoring" function. Genes expressed in less than 5% of the cells were filtered out in each human or mouse tissue.

We defined the G1 score for a gene as follows to account for the possibility that a gene can either induce or inhibit senescence.

$$A = number\ of\ cells\ in\ G1\ cell\ cycle\ and\ with\ positive\ expression\ of\ the\ gene$$
$$B = number\ of\ cells\ in\ G1\ cell\ cycle\ and\ with\ zero\ expression\ of\ the\ gene$$
$$C = number\ of\ cells\ not\ in\ G1\ cell\ cycle\ and\ with\ positive\ expression\ of\ the\ gene$$
$$D = number\ of\ cells\ not\ in\ G1\ cell\ cycle\ and\ with\ zero\ expression\ of\ the\ gene$$
$$G1\ score = log(max\{\frac{A/B}{C/D}, \frac{C/D}{A/B}\})$$

We used CellChat version 1.6.1[28] to perform cell-cell communication analysis between SnCs and other cell types. The interaction strengths between cell types and the interaction strengths of signaling pathways for SnCs were calculated using the standard CellChat pipeline in each tissue. To compare across different tissues, the interaction strength between a cell type and SnC was divided by the sum of interaction strengths of all cell types outgoing from SnCs. Similarly, the interaction strength of a signaling pathway outgoing from SnCs was divided by the sum of interaction strengths of all significant signaling pathways outgoing from SnCs.

Acknowledgments

The study was supported by the National Institutes of Health under Award Number 1U54AG075936-01.

Author Contributions

Z.J. conceived the study. Y.Q. collected the data and conducted the analysis. Z.J. and Y.Q. wrote the manuscript.

Competing Interests

All authors declare no competing interests.

Supplementary Information

Supplementary Table S1: Integrative results of nine existing SnGs. The first column "Gene_symbol" shows the human gene symbols. The following nine columns indicate whether a gene was reported by an SnG. The column "number of occurrences" counts the number of SnGs reporting a gene. The following three columns show the senescence directions reported by the three SnGs that have direction information ("up" stands for inducing senescence and "down" stands for inhibiting senescence). If an entry is empty in the last three columns, it indicates that the gene's direction was not reported.

Supplementary Table S2: Differential gene expression results in GTEx bulk-RNA-seq with respect to age. Each sheet corresponds to the analysis results for a tissue specified by the sheet name. The first column in each sheet represents gene symbols, followed by logFC (log 2 fold change), AveExpr (average expression levels), t (t statistics), P.value (p-values), adj.P.Val (p-values adjusted by the BH procedure), and B (B-statistics).

Supplementary Table S3: Proportion of SnCs in each pair of tissue (columns) and cell type (rows) in the Tabula Sapiens (human) and Tabula Muris (mouse) single-cell dataset. "NA" indicates the cell type is not present in the corresponding tissue.

Supplementary Table S4: Interaction strengths from SnCs towards other cell types (columns) in each tissue (rows). “NA” indicates the cell type is not present in the corresponding tissue.

Supplementary Table S5: Interaction strength of each signaling pathway (columns) for interactions where SnCs were senders in each tissue (rows).

Supplementary Table S6: Cell type annotation conversion for the Tabula Sapiens (human) and Tabula Muris (mouse) single-cell datasets. Column “ct_ts” and “cleaned” in the first sheet indicate the original cell_ontology_class annotations in the Tabula Sapiens and the cleaned cell types; Column “ct_tm” and “cleaned” in the second sheet indicate the same for the Tabula Muris dataset.

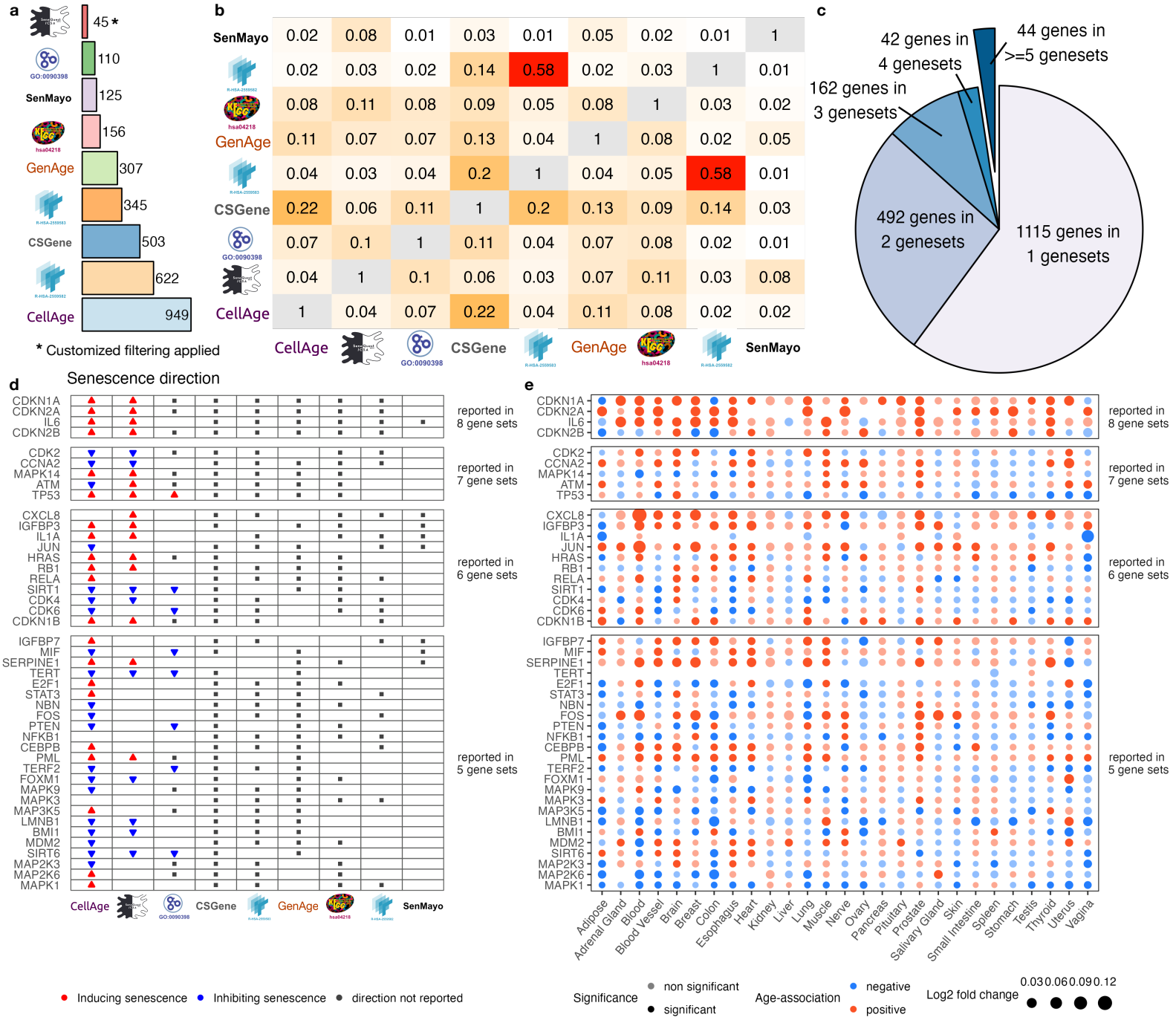

Fig.1: a, Numbers of genes in each SnG. **b**, Jaccard index between each pair of SnGs. **c**, Number of genes reported by different numbers of SnGs. **d**, SnG memberships and senescence directions of 44 genes that are reported by at least five SnGs. **e**, Association between gene expression and age in GTEx bulk RNA-seq data for the 44 genes. Colors indicate the directions of age associations. Transparency levels indicate statistical significancy. Sizes of dots indicate log2 fold changes.

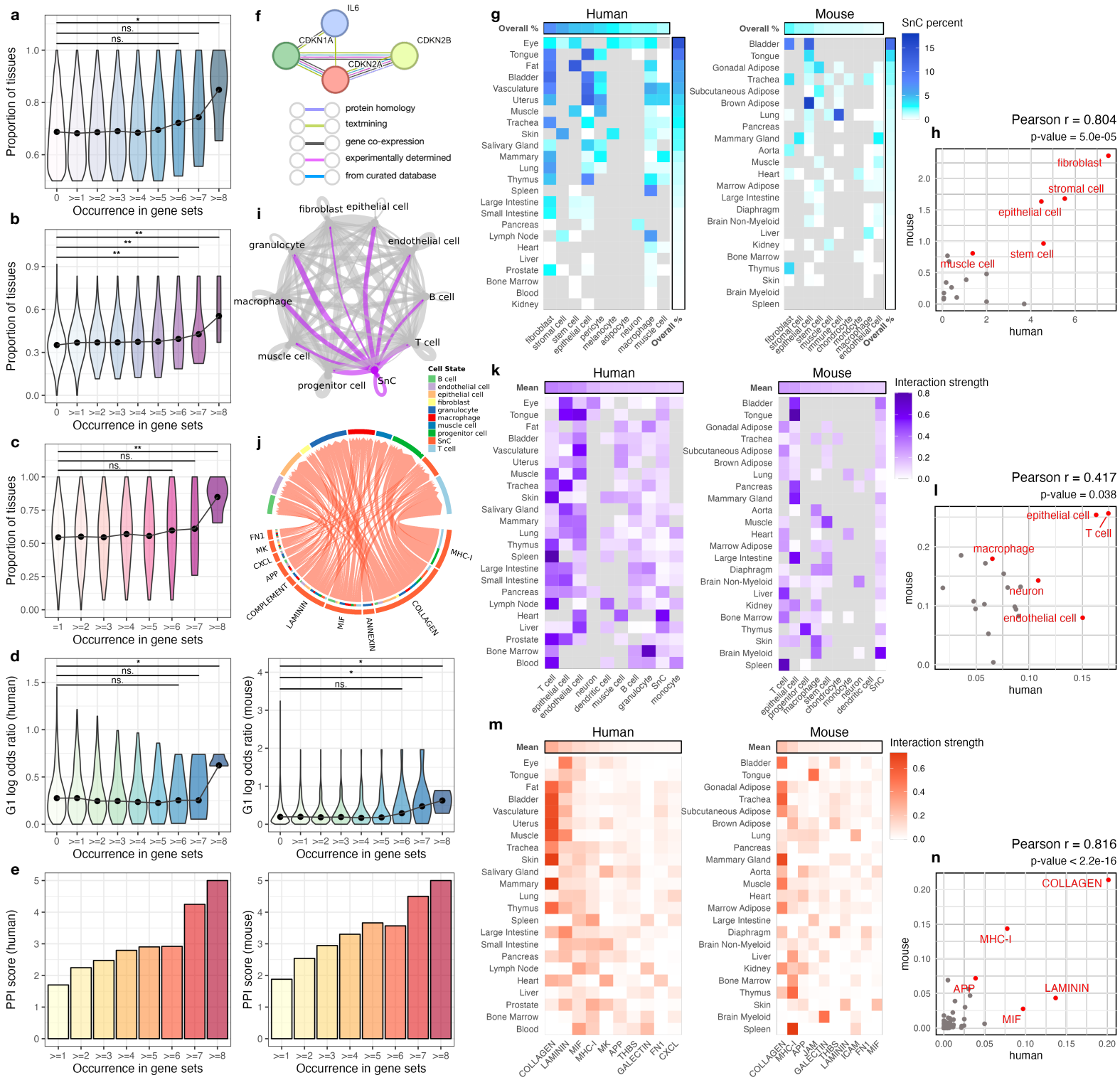

Fig2. a-e, Proportion of tissues with the same signs of associations with age (**a**), proportion of tissues with statistically significant associations with age (**b**), Proportion of tissues with agreeing signs of age associations and SnG senescence directions (**c**), log odds ratios for G1 cell cycle enrichment in human (left) and mouse (right) (**d**), and PPI scores in human (left) and mouse (right) (**e**). "*" means p-value < 0.05 and p-value > 0.01. "**" means p-value < 0.01 and p-value > 0.001. "ns." means not statistically significant. Dots stand for averaged numbers. **f**, Human PPI network of the core senescence gene set. **g**, Proportion of SnCs in each tissue and cell type in human (left) and mouse (right). Overall proportion is calculated as SnC proportion across all cell types within a tissue (rows) or across all tissues within a cell type (column). **h**, Overall SnC proportions of each cell type in human and mouse. Top five cell types are highlighted in red. **i**, CellChat interaction targets of SnCs in human trachea tissue. **j**, CellChat interaction pathways in human trachea tissue. **k**, SnC interaction target strengths in each tissue and cell type in human (left) and mouse (right). **l**, Averaged SnC interaction target strengths of each cell type in human and mouse. Top five cell types are highlighted in red. **m**, SnC interaction pathway strengths in each tissue and cell type in human (left) and mouse (right). **n**, Averaged SnC interaction pathway strengths of each pathway in human and mouse. Top five pathways are highlighted in red.

Reference

1 Hernandez-Segura, A., Nehme, J. & Demaria, M. Hallmarks of Cellular Senescence. *Trends Cell Biol* **28**, 436-453 (2018). https://doi.org/10.1016/j.tcb.2018.02.001

2 Gorgoulis, V. *et al.* Cellular Senescence: Defining a Path Forward. *Cell* **179**, 813-827 (2019). https://doi.org/10.1016/j.cell.2019.10.005

3 Lopez-Otin, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. Hallmarks of aging: An expanding universe. *Cell* **186**, 243-278 (2023). https://doi.org/10.1016/j.cell.2022.11.001

4 Lopez-Otin, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The hallmarks of aging. *Cell* **153**, 1194-1217 (2013). https://doi.org/10.1016/j.cell.2013.05.039

5 Coryell, P. R., Diekman, B. O. & Loeser, R. F. Mechanisms and therapeutic implications of cellular senescence in osteoarthritis. *Nat Rev Rheumatol* **17**, 47-57 (2021). https://doi.org/10.1038/s41584-020-00533-7

6 Martinez-Cue, C. & Rueda, N. Cellular Senescence in Neurodegenerative Diseases. *Front Cell Neurosci* **14**, 16 (2020). https://doi.org/10.3389/fncel.2020.00016

7 Saez-Atienzar, S. & Masliah, E. Cellular senescence and Alzheimer disease: the egg and the chicken scenario. *Nat Rev Neurosci* **21**, 433-444 (2020). https://doi.org/10.1038/s41583-020-0325-z

8 Schafer, M. J. *et al.* Cellular senescence mediates fibrotic pulmonary disease. *Nat Commun* **8**, 14532 (2017). https://doi.org/10.1038/ncomms14532

9 Freund, A., Orjalo, A. V., Desprez, P. Y. & Campisi, J. Inflammatory networks during cellular senescence: causes and consequences. *Trends Mol Med* **16**, 238-246 (2010). https://doi.org/10.1016/j.molmed.2010.03.003

10 Debacq-Chainiaux, F., Erusalimsky, J. D., Campisi, J. & Toussaint, O. Protocols to detect senescence-associated beta-galactosidase (SA-betagal) activity, a biomarker of senescent cells in culture and in vivo. *Nat Protoc* **4**, 1798-1806 (2009). https://doi.org/10.1038/nprot.2009.191

11 Cohn, R. L., Gasek, N. S., Kuchel, G. A. & Xu, M. The heterogeneity of cellular senescence: insights at the single-cell level. *Trends Cell Biol* **33**, 9-17 (2023). https://doi.org/10.1016/j.tcb.2022.04.011

12 Saul, D. *et al.* A new gene set identifies senescent cells and predicts senescence-associated pathways across tissues. *Nat Commun* **13**, 4827 (2022). https://doi.org/10.1038/s41467-022-32552-1

13 Zhao, M., Chen, L. & Qu, H. CSGene: a literature-based database for cell senescence genes and its application to identify critical cell aging pathways and associated diseases. *Cell Death Dis* **7** (2016). https://doi.org/10.1038/cddis.2015.414

14 Avelar, R. A. *et al.* A multidimensional systems biology analysis of cellular senescence in aging and disease. *Genome Biol* **21**, 91 (2020). https://doi.org/10.1186/s13059-020-01990-9

15 de Magalhaes, J. P., Curado, J. & Church, G. M. Meta-analysis of age-related gene expression profiles identifies common signatures of aging. *Bioinformatics* **25**, 875-881 (2009). https://doi.org/10.1093/bioinformatics/btp073

16 Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25-29 (2000). https://doi.org/10.1038/75556
17 Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27-30 (2000). https://doi.org/10.1093/nar/28.1.27
18 Gillespie, M. *et al.* The reactome pathway knowledgebase 2022. *Nucleic Acids Res* **50**, D687-D692 (2022). https://doi.org/10.1093/nar/gkab1028
19 Yamamoto, R. *et al.* Tissue-specific impacts of aging and genetics on gene expression patterns in humans. *Nat Commun* **13**, 5803 (2022). https://doi.org/10.1038/s41467-022-33509-0
20 SenNet, C. NIH SenNet Consortium to map senescent cells throughout the human lifespan to understand physiological health. *Nat Aging* **2**, 1090-1100 (2022). https://doi.org/10.1038/s43587-022-00326-5
21 Consortium, G. T. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580-585 (2013). https://doi.org/10.1038/ng.2653
22 Kumari, R. & Jat, P. Mechanisms of Cellular Senescence: Cell Cycle Arrest and Senescence Associated Secretory Phenotype. *Front Cell Dev Biol* **09** (2021). https://doi.org/10.3389/fcell.2021.645593
23 Jones, R. C. *et al.* The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. *Science* **376**, 711-+ (2022). https://doi.org/10.1126/science.abl4896
24 Schaum, N. *et al.* Single-cell transcriptomics of 20 mouse organs creates a. *Nature* **562**, 367-+ (2018). https://doi.org/10.1038/s41586-018-0590-4
25 Rao, V. S., Srinivas, K., Sujini, G. N. & Kumar, G. N. Protein-protein interaction detection: methods and analysis. *Int J Proteomics* **2014**, 147648 (2014). https://doi.org/10.1155/2014/147648
26 Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* **47**, D607-D613 (2019). https://doi.org/10.1093/nar/gky1131
27 Xu, P. *et al.* The landscape of human tissue and cell type specific expression and co-regulation of senescence genes. *Mol Neurodegener* **17**, 5 (2022). https://doi.org/10.1186/s13024-021-00507-7
28 Jin, S. *et al.* Inference and analysis of cell-cell communication using CellChat. *Nat Commun* **12**, 1088 (2021). https://doi.org/10.1038/s41467-021-21246-9
29 Dhirachaikulpanich, D., Lagger, C., Chatsirisupachai, K., de Magalhaes, J. P. & Paraoan, L. Intercellular communication analysis of the human retinal pigment epithelial and choroidal cells predicts pathways associated with aging, cellular senescence and age-related macular degeneration. *Front Aging Neurosci* **14**, 1016293 (2022). https://doi.org/10.3389/fnagi.2022.1016293
30 Blokland, K. E. C., Pouwels, S. D., Schuliga, M., Knight, D. A. & Burgess, J. K. Regulation of cellular senescence by extracellular matrix during chronic fibrotic diseases. *Clin Sci (Lond)* **134**, 2681-2706 (2020). https://doi.org/10.1042/CS20190893
31 Levi, N., Papismadov, N., Solomonov, I., Sagi, I. & Krizhanovsky, V. The ECM path of senescence in aging: components and modifiers. *FEBS J* **287**, 2636-2646 (2020). https://doi.org/10.1111/febs.15282

32 Mavrogonatou, E., Papadopoulou, A., Pratsinis, H. & Kletsas, D. Senescence-associated alterations in the extracellular matrix: deciphering their role in the regulation of cellular function. *Am J Physiol Cell Physiol* **325**, C633-C647 (2023). https://doi.org/10.1152/ajpcell.00178.2023

33 Adam Freund 1, R.-M. L., Marco Demaria, Judith Campisi. Lamin B1 loss is a senescence-associated biomarker. *Molecular biology of the cell* (2012). https://doi.org/10.1091/mbc.E11-10-0884

34 Pereira, B. I. *et al.* Senescent cells evade immune clearance via HLA-E-mediated NK and CD8(+) T cell inhibition. *Nat Commun* **10**, 2387 (2019). https://doi.org/10.1038/s41467-019-10335-5

35 Hernandez-Mercado, E. *et al.* Increased CD47 and MHC Class I Inhibitory Signals Expression in Senescent CD1 Primary Mouse Lung Fibroblasts. *Int J Mol Sci* **22** (2021). https://doi.org/10.3390/ijms221910215

36 Zhang, Y. *et al.* Macrophage migration inhibitory factor rejuvenates aged human mesenchymal stem cells and improves myocardial repair. *Aging (Albany NY)* **11**, 12641-12660 (2019). https://doi.org/10.18632/aging.102592

37 Welford, S. M. *et al.* HIF1alpha delays premature senescence through the activation of MIF. *Genes Dev* **20**, 3366-3371 (2006). https://doi.org/10.1101/gad.1471106

38 Tacutu, R. *et al.* Human Ageing Genomic Resources: new and updated databases. *Nucleic Acids Res* **46**, D1083-D1090 (2018). https://doi.org/10.1093/nar/gkx1042

39 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550 (2005). https://doi.org/10.1073/pnas.0506580102

40 Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739-1740 (2011). https://doi.org/10.1093/bioinformatics/btr260

41 Chi-Lam Poon, J. C., Solomon Shorser, Joel Weiser, Robin Haw, Lincoln Stein. R interface to the reactome graph database [version 1; not peer reviewed]. *F1000Research 2021* **10(ISCB Comm J):721** (2021). https://doi.org/https://doi.org/10.7490/f1000research.1118690.1

42 KEGGREST: Client-side REST access to the Kyoto Encyclopedia of Genes and Genomes (KEGG) (2023).

43 Blake, J. A. *et al.* Mouse Genome Database (MGD): Knowledgebase for mouse-human comparative biology. *Nucleic Acids Res* **49**, D981-D987 (2021). https://doi.org/10.1093/nar/gkaa1083

44 Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015). https://doi.org/10.1093/nar/gkv007

45 Yoav Benjamini, Y. H. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society* (1995). https://doi.org/https://doi.org/10.1111/j.2517-6161.1995.tb02031.x