A transcriptome-based single-cell biological age model and resource for tissue-specific aging measures

Shulin Mao, Jiayu Su, Longteng Wang, et al.

Genome Res. published online July 31, 2023
Access the most recent version at doi:10.1101/gr.277491.122

A transcriptome-based single-cell biological age model and resource for tissue-specific aging measures

Shulin Mao[1, 2, #], Jiayu Su[2, 3, #], Longteng Wang[2, 4], Xiaochen Bo[5], Cheng Li[2, 6, *], and Hebing Chen[5, *]

1. Yuanpei College, Peking University, Beijing 100871, China.

2. Center for Bioinformatics, School of Life Sciences, Peking University, Beijing 100871, China.

3. Department of Systems Biology, Columbia University, New York, NY, 10032, USA.

4. School of Life Sciences, Joint Graduate Program of Peking-Tsinghua-NIBS, Peking University, Beijing 100871, China.

5. Institute of Health Service and Transfusion Medicine, Beijing 100850, China.

6. Center for Statistical Science, Peking University, Beijing 100871, China.

[#] Equal contribution

* Correspondence to: chenhb@bmi.ac.cn and cheng_li@pku.edu.cn

Running Title: A tissue-specific single-cell biological age model

Keywords: Aging; Biological age; Single-cell transcriptomics

Abstract

Accurately measuring biological age is crucial for improving healthcare for the elderly population. However, the complexity of aging biology poses challenges in how to robustly estimate aging and in how to interpret the biological significance of the traits used for estimation. Here we present SCALE, a statistical pipeline that quantifies biological aging in different tissues using explainable features learned from literature and single-cell transcriptomic data. Applying SCALE to the “Mouse Aging Cell Atlas” (*Tabula Muris Senis*) data, we identified tissue-level transcriptomic aging programs for over 20 murine tissues and created a multi-tissue resource of mouse quantitative aging-associated genes. We observe that SCALE correlates well with other age indicators, such as the accumulation of somatic mutations, and can distinguish subtle differences in aging even in cells of the same chronological age. We further compared SCALE with other transcriptomic and methylation ‘clocks’ in data of aging muscle stem cells, Alzheimer’s disease, and heterochronic parabiosis. Our results confirm that SCALE is more generalizable and reliable in assessing biological aging in aging-related diseases and rejuvenating interventions. Overall, SCALE represents a valuable advancement in our ability to measure aging accurately, robustly, and interpretably in single cells.

Introduction

Aging is the gradual decline of biological and physiological functions that occurs in most living organisms over time. It is a complex process involving cumulative changes and damage at all levels, rendering it the primary risk factor of various human diseases, including cardiovascular disease, neurodegeneration, and cancer (Lopez-Otin et al. 2013). By 2050, elderly individuals over 60 years old will make up more than 20% of the global population (Harper 2014). As the burden of aging related diseases significantly increases in an aging world, there is a growing need to better understand the underlying biological process, quantify the impacts of aging, and provide rejuvenating interventions (Partridge et al. 2018; Scott et al. 2021).

People of the same chronological age have different aging states, which can be monitored using various biomarkers (Belsky et al. 2015). These markers are usually measurable indicators of a particular outcome or source of aging, such as phenotypical measures like frailty and molecular measures like DNA methylation dynamics (Schumacher et al. 2021; López-Otín et al. 2023). While informative, they are not always quantitatively predictive of an individual's true biological age, nor are they easy to obtain. The advancement of high-throughput screening platforms and extensive longitudinal studies has greatly facilitated the search for new non-invasive and quantitative biomarkers of aging. For instance, high-throughput sequencing allows unbiased multi-omics profiling of DNA, RNA, and epigenetic changes during aging, providing a comprehensive view of senescence at tissue and single-cell levels (Solovev et al. 2020; Aging Atlas 2021). These omics datasets contain vast and noisy measurements of potential candidate markers and, consequently, require carefully designed computational models to identify and extract predictive signals from the data.

However, construction of such models is often highly degenerate, yielding little overlap of identified biomarkers between studies and thus making results difficult to interpret (Thompson et al. 2018; Galkin et al. 2020).

Among the many computational algorithms, linear regression and its variants have been widely used to select aging-related biomarkers and build “aging clocks”, i.e., predictors of chronological age and biological age, in various omics datasets and aging systems (Hannum et al. 2013; Horvath 2013; Peters et al. 2015; Levine et al. 2018; Lu et al. 2019; Belsky et al. 2020; Moaddel et al. 2021; Buckley et al. 2023). To reduce overfitting and increase interpretability, regularization techniques are often utilized to shrink coefficients towards zero and highlight a sparse set of relevant features. In particular, the success of the first cross-tissue DNA methylation clock has popularized Elastic Net, a model that minimizes L1 and L2 penalties combined (Friedman et al. 2010; Horvath 2013).

Nonetheless, aging is a complicated phenomenon that impacts every aspect of an organism, and its roots and consequences spread across thousands of targets in a context-specific manner. There is likely no one-size-fits-all predictor of aging. Even if such a solution exists, the presence of multicollinearity makes it challenging to distinguish the true biologically relevant signatures driving the relationship with age. Since omics data often contains hundreds of thousands of redundant features but only one target variable to predict (chronological age), it is, therefore, conceptually and technically challenging to construct a consensus set of aging biomarkers *de novo* across datasets, let alone across modalities (Rutledge et al. 2022). Moreover, varying data quality and batch effects between sequencing platforms and cohorts may bias the feature selection and thus limit the

generalizability of trained aging clocks.

Most of the existing omics-based aging clocks have been constructed using data from bulk tissues, which neglects the variations in cell compositions and cell-to-cell aging heterogeneity. To gain a more detailed and nuanced view of cell-type-specific molecular changes during aging, several studies have applied machine-learning models to single-cell transcriptomics and DNA methylation data (Trapp et al. 2021; Buckley et al. 2023). Despite their success in predicting chronological age within specific training contexts, these clocks are constrained by their applicability to a limited number of cell types and tissues. Their generalizability to other cell types and disease data, particularly in cases with ambiguous cell-type identities, remains uncertain. Additionally, problems like data sparsity and batch effects are more pronounced in single-cell omics data, further complicating the identification of consensus aging markers and interpretation of model results. Furthermore, as chronological age is often the only available measure of biological age, it becomes critical to determine whether the features learned from single-cell omics data can capture other dimensions of biological aging.

Here, we developed SCALE (Single-Cell Aging Level Estimator), a tissue-specific measure that quantifies aging in individual cells using single-cell RNA-sequencing (scRNA-seq) data (Fig. 1). Unlike other clocks, SCALE adopts a knowledge-based forward feature selection strategy to prioritize genes that are well-recognized to be associated with aging, making the results more explainable and less affected by technical noise. We applied SCALE to the *Tabula Muris Senis* dataset (Tabula Muris 2020) to identify tissue-level transcriptomic aging programs for over 20

murine tissues and organs (Supplemental Fig. S1). We further evaluated the effectiveness of SCALE in accurately reflecting biological age under various conditions, including aging-related diseases and anti-aging interventions.

Results

Knowledge-based mining of transcriptomic markers of aging in humans and mice

Extracting an informative subset of aging-related features out of high-dimensional transcriptomic data is challenging, due to the intricate nature of aging itself as well as the interdependencies between genes. Existing regularization-based regression models (i.e., Elastic Net) for building aging clocks struggle with collinear factors, which results in unstable feature selection and fluctuating coefficients. These complications are amplified when the model is trained on or applied to different datasets (Supplemental Fig. S2A-C), undermining model interpretability and performance in biological applications (Supplemental Fig. S2D). To address this, we developed an alternative approach that aims to incorporate more well-documented aging genes and other genes with complementary information.

We began by creating “Aging Map” (Fig. 2; Supplemental Fig. S3; http://sysomics.com/AgingMap/), a manually curated online database of human and mouse aging-related genes. Aging Map first integrated six existing databases focused on different aspects of aging, including "Gene Ontology (GO)" (The Gene Ontology Consortium et al. 2000), "CSGene" (Zhao et al. 2016), "Aging Atlas" (Aging Atlas 2021), "HAGR/GenAge" (Tacutu et al. 2018), "AgeFactDB" (Huhne et al. 2014), and

"Longevity Map" (Budovsky et al. 2013) (Fig. 2A, B). We also extracted and manually curated additional aging-related genes by mining PubMed abstracts to integrate more recent research efforts (Methods). This identified 636 literature-confirmed aging genes in humans and 460 genes in mice, including 162 and 49 novel hits that mostly encode histone proteins. Using Gene Ontology (GO) analysis, we further verified that human and mouse literature-based genes were significantly enriched in well-established aging processes, such as cellular senescence, cell-cycle regulation, and telomere organization (Fig. 2C, D; Supplemental Fig. S3D, E). Together, Aging Map contains 1,119 and 1,459 curated human and mouse aging genes, respectively, covering almost all scales of aging ranging from molecular damage to genetic predisposition. Cross-species comparison revealed modest overlap between known human and mouse aging genes, suggesting both conservation of core senescence pathways and fundamental differences in aging between mice and humans (Fig. 2E).

Aging-associated genes can alternatively be identified in a data-driven manner (Finotello and Di Camillo 2015). However, most data-derived gene discoveries have not been tested experimentally or on independent cohorts, making it difficult to discern whether a gene is truly aging-related or merely a reflection of technical noise (Rutledge et al. 2022). By integrating knowledge from the literature, Aging Map enables the prioritization of functional important aging-associated genes from a vast pool of candidates. Here, we collected genes differentially expressed with age (age-DEGs) from three large longitudinal cohorts, including a pan-tissue study featuring 26 tissues from GTEx (Peters et al. 2015; Harris et al. 2017; Chatsirisupachai et al. 2019). These studies showed little agreement, in part due to differences in study scope, sample source, and transcriptomic profiling

technology, underscoring the need for more robust knowledge-based approaches to select aging markers (Fig. 2F). We then cross-referenced Aging Map human genes with the three age-DEG sets. While the majority of data-driven age-DEGs did not have direct experimental support from existing aging literature, Aging Map helped identify five consensus genes (*CTSC*, *FOXO1*, *LMNA*, *PLAUR*, *RPS6KA1*) that have long been recognized for their roles in longevity (Fig. 2G).

Moving beyond individual aging genes, we explored the potential of using holistic transcriptomic features as quantitative aging biomarkers. Previous studies have reported increased overall RNA abundance in aged murine hematopoietic stem cells (HSCs) (Flohr Svendsen et al. 2021). To test whether the finding generalizes to other systems, we utilized the *Tabula Muris Senis* scRNA-seq data of over 20 tissues and organs (Tabula Muris 2020). Here, we defined global expression of a cell as the log-transformed mean counts per gene and compared it across age groups sequenced in the same batch within each sample using a linear model (Methods). We observed a consistent significant positive correlation between global expression and chronological age in 27 out of the 37 analyzed samples. Two samples (droplet-based murine fat and limb muscle) showed the opposite trend (Fig. 2H; Supplemental Fig. S4), which was likely caused by measurement variability.

Another established phenomenon in aging is the increase in stochasticity and expression noise as cells gradually lose control over gene regulation (Martinez-Jimenez et al. 2017). To quantify the degree of expression uncertainty in each cell, we used entropy, a thermodynamic measure widely applied to describe cell differentiation potentials (MacArthur and Lemischka 2013; Chen and Teschendorff 2019) (Methods). Cells with low entropy are generally more deterministic and express

a narrower spectrum of genes. We found that entropy was significantly different across age groups in all samples (Fig. 2H; Supplemental Fig. S5). Nonetheless, we did not observe a monotonic increase of entropy over chronological age. As the regulation of gene expression is highly dynamic and comprehensive, further investigation is necessary to fully delineate the age-associated quantitative changes in transcriptional fidelity from noise generated by technical limitations and other biological processes. It is also possible to extend the analysis to other high-level systems-based transcriptomic signatures such as activities of gene regulatory networks and pathways using similar approaches.

Collectively, Aging Map provides a pan-tissue set of high-confidence genes associated with various aspects of aging, which can help prioritize targets in data-driven aging research. However, the substantial proportion of unannotated age-DEGs implies that there likely remain many more aging genes yet to be confirmed by experiments (Fig. 2F, G). Many of them, indeed, could represent tissue-specific aging programs driven by stress factors and cell-cell interactions in the local environment (Zhang et al. 2021). Furthermore, not all age associations are quantitative predictors of aging. To address these limitations, we proposed a knowledge-based machine-learning approach to identify novel quantitative aging markers from transcriptomic datasets.

Identification of tissue-specific genes as quantitative aging predictors using guided forward selection

To increase the statistical power of our study, we utilized scRNA-seq, which provides a larger number of data points and higher granularity compared to bulk RNA-seq. We analyzed over 20

183 murine tissues and organs from the *Tabula Muris Senis* dataset to extract tissue specific aging
184 programs (Fig. 3A; Supplemental Fig. S1). Some tissues were sequenced by two different platforms
185 (droplet-based 3' biased and FACS-based full-length sequencing) at different time points, and were
186 thus analyzed separately, resulting in a total of 37 samples from two to six age groups for
187 investigation. We assumed a generalized linear relationship between aging gene expression and
188 chronological age and used Elastic Net to select a sparse subset of predictive genes. In particular,
189 we employed a guided forward selection strategy that iteratively incorporates new genes to improve
190 regression performance while considering the expression of known aging-associated genes to
191 enhance the robustness and the interpretability of feature selection (Methods; Supplemental Fig. S6,
192 Fig. S7A). Based on this method, we generated tissue-specific aging gene sets for downstream
193 analysis (Supplemental Table 1).
194
195 Compared with genes directly selected by regression-based methods (Buckley et al. 2023), our
196 aging gene sets exhibited a significantly higher overlap with well-documented aging genes (genes in
197 Aging Map, Fig. 3B), suggesting an increased likelihood of identifying genes involved in the
198 mechanistic basis of aging. In addition, only seven out of 100 of our brain aging genes were
199 previously reported as cell-type-specific aging genes in brain cells (Buckley et al. 2023). This
200 demonstrates that our approach uncovers novel aging programs shared across cell types, rather
201 than simply combining cell-type-level aging programs (Fig. 3C). Furthermore, most aging genes
202 selected in different tissues were tissue-specific (Supplemental Fig. S7B), aligning with our
203 expectation of aging heterogeneity across tissues (Rutledge et al. 2022). Although there was little
204 overlap in the aging genes across tissues, we found that these genes often contributed to the same

biological processes (Supplemental Fig. S8A, B). In particular, these tissue-specific genes were frequently involved in “aging”, “ATP metabolic process”, and the regulations of protein translation and various other catabolic and metabolic processes, suggesting the importance and relative conservation of some aging programs across murine tissues and organs.

Among the pan-tissue aging genes (intersections of tissue-specific gene sets) identified in this study, some, such as *Igf1* and *Bcl2*, have been confirmed in previous experiments (Junnila et al. 2013; Fernández Á et al. 2018; Rutledge et al. 2022). One of the novel discoveries, *Lars2,* was found to be broadly down-regulated across mouse tissues during aging (Supplemental Fig. S9). Clinical reports show that loss-of-function mutations in the *Lars2* gene are present in 1/3 of patients with Perrault syndrome, which can cause hearing loss, motor and sensory neurons damage, decreased learning ability, and ovarian dysfunction (Kosaki et al. 2018). We therefore speculate that decreased expression of *Lars2* may induce neurological and reproductive system damage, resulting in the aging phenotype. Another example is *Rpl13a*. We found it to be up-regulated in most tissues during aging (Supplemental Fig. S10). Studies in yeast have shown that individuals carrying *Rpl13a* loss-of-function mutations have longer lifespans than wild-type yeasts (Steffen et al. 2008). Also, high expression level of *Rpl13a* could negatively affect translation regulation and protein biogenesis, leading to discordance between transcriptome and proteome and thereby inducing cellular aging (Janssens et al. 2015).

Aging quantification at single-cell resolution using SCALE

With tissue-specific aging genes selected in the previous section, we constructed a statistical

measure named SCALE (Single-Cell Aging Level Estimator) to quantify the relative aging status of each cell residing in the same tissue (Methods). SCALE weights each aging gene according to its expression level, the direction in which its expression changes with chronological age and the proportion of cells expressing the gene.

Applying the model to the *Tabula Muris Senis* dataset, we observed a positive correlation between SCALE scores and chronological age across all tissues (Fig. 3D-F; Supplemental Fig. S11). Pairwise comparisons between chronological age groups in each tissue confirmed that 130 out of the total 134 pairs showed significant differences in their SCALE scores (*t*-test, p-value < 0.05). The remaining four pairs may have insufficient statistical power due to limitations in the dataset or individual cell variability. To further validate our findings, we constructed empirical null distributions of the correlation by randomly selecting genes and replacing the aging genes in our model (Methods). SCALE outperformed the null in 35 of the 37 samples (adjusted p-value < 0.05; Supplemental Fig. S12A), suggesting that the aging genes selected by SCALE are highly informative and still predictive even in the absence of trained regression models.

To demonstrate SCALE's robustness against technical artifacts, we conducted additional sensitivity tests. We downsampled the scRNA-seq data to simulate dropout events and reduce available aging signals. Even after zeroing-out expression of 50% of the genes, which effectively reduced the sequencing depth by half, the SCALE score still presented impressive resilience, with a Pearson's correlation of over 0.8 (Supplemental Fig. S12B-D). We also experimented with the number of aging genes used to build SCALE and found the score to be highly correlated (median Pearson's

correlation > 0.5, Supplemental Fig. S13).

The accumulation of somatic mutations in cells due to various internal and external stress during aging is a renowned hallmark of biological aging (Lodato et al. 2018; Zhang et al. 2019; Brazhnik et al. 2020). To evaluate SCALE's predictive performance regarding biological age, we conducted mutation analysis on samples in which full-length RNA coverage was available for exonic somatic mutation detection (Tabula Muris 2020) (Fig. 3G, Supplemental Fig. S14). In all 24 tissues, SCALE scores demonstrated a significant positive linear relationship with the mean number of somatic mutations in genes (p-value < 0.05, Fig. 3H; Supplemental Fig. S15). Since both the SCALE score and the mutation number also positively correlate with chronological age, we regressed out the effect of chronological age and examined whether SCALE can reflect complementary aspects of biological aging not captured by chronological age. Indeed, we found that the residual SCALE scores (ΔSCALE) were still significantly positively associated with the mean number of somatic mutations in 54 out of the total 72 age groups (75%) in 24 tissues (p-value < 0.05, Fig. 3I, J; Supplemental Fig. S16), indicating that SCALE can distinguish nuanced differences in biological aging within the same chronological age group. Samples in which ΔSCALE had no significant positive association with mutations were mostly from 3-month-old mice (11 out of the total 18 groups), implying that the number of mutations in young mice may not accurately reflect aging status.

SCALE outperforms other single-cell aging clocks in assessing impacts of disease and rejuvenating interventions on aging

Aging is often accompanied by skeletal muscle atrophy and subsequent decline in mobility, which can be attributed to the decreased regenerative capacity of muscle stem cells (Blau et al. 2015; Shcherbina et al. 2020). To test whether our tissue-level model records the aging of stem cells, we applied the SCALE model trained on the *Tabula Muris Senis* limb muscle tissue to an external multi-omics dataset of mouse muscle satellite cells (Hernando-Herraez et al. 2019) (Fig. 4A). Although the two datasets differ substantially in sequencing technology and data quality, SCALE generalized well in the external stem cell data and recovered the aging difference between cells from 2-month-old and 24-month-old mice (p-value = 0.016; Fig. 4B). Leveraging the simultaneously measured single-cell DNA methylation data, we compared SCALE with scAge, an epigenetic aging clock that exploits methylation-age associations learned from bulk data to estimate age at the single-cell level (Trapp et al. 2021). We found that scAge predicted more outliers with extreme age values and had difficulty differentiating cells from different age groups (p-value = 0.053; Fig. 4C, Methods), probably because of sparsity and CpG coverage variation in single-cell DNA methylation data. This suggests that SCALE may be more stable and reliable in capturing biological age when applied to noisy single-cell omics data.

Alzheimer's disease (AD) is an aging-related degenerative neurological disorder that affects vision, hearing, speech, and memory (Ballard et al. 2011). The risk of Alzheimer's disease increases with chronological age, with most cases occurring in people over 65 (Trevisan et al. 2019). To test if SCALE could reflect disease-associated aging changes, we applied the model trained on the *Tabula*

Muris Senis brain tissue to a single-nucleus RNA-seq dataset of brains from AD mouse model (Zhou et al. 2020) (Fig. 4D; Supplemental Fig. S17A). SCALE scores were significantly higher in both brain microglia cells and non-microglia cells of AD mice than those in wild-type mice (Fig. 4E; Supplemental Fig. S17B), which is consistent with the previous finding that AD-associated genes are up-regulated during aging and early-onset AD (Zhou et al. 2020). To further benchmark the robustness of SCALE over other measures, we evaluated the performance of a collection of cell-type specific transcriptomic clocks built on scRNA-seq data of the mouse brain neurogenic region (Buckley et al. 2023). Out of the three clocks where the corresponding cell type was presented in the AD data (microglia, endothelial and oligodendrocytes), two failed to distinguish cells from AD mice versus from the control group (Fig. 4F, Supplemental Fig. S17C, D). Additionally, the clocks of Buckley et al. were highly cell-type-specific and could not generalize across cell types. When applied to three other abundant brain cell types (neurons, astrocytes, and OPCs) in the AD dataset, five out of the six clocks failed to faithfully distinguish AD and WT cells (Supplemental Fig. S17F). On the contrary, the tissue-level SCALE model correctly reflected the significant aging differences between the AD and WT groups (Supplemental Fig. S17E). Together, our results indicate better generalizability of SCALE and highlight the importance of model robustness in protecting aging quantifiers from batch effects.

Heterochronic parabiosis (joining the circulatory systems of two individuals of different ages) has been recognized as an effective way to counteract aging in older animals and improve various physiological functions, including tissue regeneration, bone repair and cognition improvement (Castellano et al. 2015). To investigate whether SCALE could capture the effect of rejuvenating

311 interventions, we applied it to a multi-tissue scRNA-seq dataset of mice receiving the heterochronic
312 parabiosis surgery (Palovics et al. 2022). We observed a significant decrease of SCALE scores in
313 old mice and a significant increase in young mice following five weeks of the parabiosis procedure in
314 19 of the total 20 tissues (p-value < 0.01; Fig. 4G, H; Supplemental Fig. S18), confirming that
315 SCALE can recapitulate the aging order of cells in old and young, control and heterochronic mice. In
316 brain samples, we again compared our model to Buckley's clocks. While the two approaches each
317 identified distinct aging programs at the tissue and cell-type levels, respectively (Fig. 3C), both
318 achieved good performance in quantifying rejuvenation on matching data (Fig. 4H, Supplemental
319 Fig. S18, S19). However, similar to the AD results, we also observed that these cell-type-specific
320 clocks had low predictive accuracy when applied to previously unseen cell types (Fig. 4I,
321 Supplemental Fig. S20A). Moreover, the models of Buckley et al. require explicit cell-type
322 annotations and are only trained for a few brain cell types, which constrains their applicability in
323 aging-related disease study and other scenarios with ambiguous cell type identity. In contrast, while
324 SCALE, too, was not trained on the intervention dataset, it clearly distinguished intervention effects
325 in all cell types (Fig. 4H, Supplemental Fig. S18), including cell types where Buckley's models
326 underperformed (Supplemental Fig. S20B).

327 **Generalizing mouse SCALE models to human and rat data**

328 Although our SCALE models were trained on mouse tissues, we speculated that the conservation of
329 the core aging programs (Fig. 2E), as well as the generalizability of SCALE, might enable us to
330 accurately measure aging across species. To test this hypothesis, we replaced selected mouse
331 aging genes with corresponding homologs in other species and similarly calculated the SCALE

scores in scRNA-seq data from humans and rats.

Our first experiment involved a human brain dataset with seven chronological age groups (Hodge et al., 2019). We found that SCALE was able to distinguish different human age groups (linear regression, p-value < 0.001, Supplemental Fig. S21). We then applied SCALE to a multi-tissue scRNA-seq atlas of rats undergoing aging and calorie restriction (CR), which is known to be one of the most effective anti-aging interventions (Ma et al. 2020). We observed that in the brown adipose tissue and muscle, cells from young rats had significantly lower SCALE scores than those from older individuals. Furthermore, CR significantly reduced the SCALE scores of cells in older individuals compared to those in the control (Supplemental Fig. S22). However, we did not observe similar trends in other rat tissues, suggesting that cellular responses to aging and CR are tissue- and species-specific. In general, we recognize that training SCALE on one species and applying it to another has many limitations, as there may be some discrepancies in aging genes and pathways between species (Fig. 2E). Therefore, if possible, we recommend training individual SCALE models directly in the species of interest.

SCALE revealed cell-type-specific aging patterns across tissues

SCALE's ability to learn and quantify single-cell aging programs within a tissue enables us to compare aging patterns of different cell types residing within the same tissue. Stem cells have the potential to proliferate and differentiate, which plays a significant role in rejuvenation and regeneration. We observed that SCALE scores of stem cells were usually lower than those of other cells. For instance, the SCALE scores of adipose mesenchymal stem cells in all age groups were

significantly lower than that of other cells in the adipose tissue (Fig. 5A). Additionally, we observed that leukocytes in the pancreas were significantly older than others in the same tissue (Fig. 5B). We suspected that this finding was related to the inflammatory response and therefore decided to systematically examine immune cell populations across tissues.

The immune system is vital for organisms to resist infections and maintain homeostasis and is associated with aging. In particular, lymphocytes play a major role in chronic inflammation, and defects in their functional responses have been linked with known aging phenotypes (Nikolich-Zugich 2018). We found that, in most tissues of young mice (1-month-old and 3-month-old), the SCALE scores of B cells and T cells were significantly higher than the scores of the whole tissues where they were located. While in old mice (21-month-old, 24- month-old, and 30-month-old), there became more tissues where the SCALE scores of B cells and T cells were lower than the tissue background (Fig. 5C, D), possibly indicating a slower aging rate in these cells. To determine if it was a B cell- and T cell-specific phenomenon, we investigated the SCALE scores of the endothelial cells, which also exist in multiple tissues. However, we did not observe a similar pattern in endothelial cells (Fig. 5E). One possible explanation for the different aging patterns between B and T cells and other cells in the same tissue is that immune cells are highly mobile and are constantly renewed by HSCs rather than being produced locally.

To further explore how different cell types change during aging in various tissue environments, we calculated the SCALE scores for all 145 cell types in the mouse samples. Generally, we found that the correlation between SCALE scores and chronological age varies across different cell types,

even within the same tissue. This prompted us to use the correlation coefficient as an indicator of aging programs. For example, a low correlation between the SCALE score and chronological age within a cell type could imply stronger unique cell-type aging characteristics. However, different tissues provide distinct environments with various stressors, and thus have different aging programs (Supplemental Fig. S7B; Supplemental Table 1).

To compare one cell type across tissues, we first need to assess the aging baseline in each tissue. We focused on 13 cell types that were detected in at least five samples and investigated if their aging characteristics (i.e., the strength of the correlation between SCALE scores and chronological age) were related to their tissue location. Our results indicated that cells located in tissues with higher correlation coefficients (i.e., stronger aging programs) generally exhibit higher coefficients at the cell-type level, suggesting a tissue influence (Fig. 5F).

Still, deviations in cell-type correlations from the tissue background suggest distinct aging characteristics in several cell types. To infer the significance of deviations (i.e., the influence of cell-type-specific aging characteristics) from tissue baseline, we developed a permutation-based approach and classified cell types in a given tissue into three categories (Methods): 1) cell types with a significantly higher correlation than all other cells in the same tissue; 2) cell types with a comparable correlation coefficient to the tissue as a whole; and 3) cell types with a significantly lower correlation coefficient (Fig. 5G). These categories have straight-forward interpretations. For example, aging in cells of category 3 may have unique aging characteristics that cause the deviation from the tissue baseline (category 2).

Using this metric, we investigated whether the aging pattern is consistent across different tissues for each cell type. We observed that, for a given cell type, it tended to be classified in the same category of the aforementioned three types across diverse tissue environments. For example, immune cells (B cells, T cells, CD4+T cells, CD8+T cells, macrophages, and NK cells) had significantly lower correlation coefficients than the tissue as a whole (category 3) in most non-immune organs where they were detected in our data (Fig. 5H). Furthermore, we found that other cell types that were present in multiple tissues, such as endothelial cells and epithelial cells, also tended to belong to the same category across tissues (Supplemental Fig. S23). These findings indicate that inherent characteristics of cell types play an important role in shaping cell aging patterns, even when situated in different tissue environments.

Discussion

Here we show that tissue-specific aging programs can be learned from scRNA-seq data and applied to describe aging heterogeneity within single-cell populations. From the Aging Map, a curated database of aging-associated genes, we extended our search for transcriptomic features that could quantitatively predict chronological age, generating a multi-tissue resource of mouse genes as aging biomarkers. We then developed a statistical measure of aging called SCALE, which we applied to scRNA-seq datasets from various murine tissues and organs to evaluate differences in biological aging. As an internal validation, we compared SCALE to the number of somatic mutations and confirmed that SCALE can distinguish aging statuses beyond just predicting chronological age. When quantifying the aging and anti-aging effects of diseases and rejuvenating interventions in

external data, SCALE exhibits high robustness to usual challenges in single-cell omics data such as batch effects and sparsity. We further considered the SCALE scores of each cell type and observed conserved cell-type-specific aging patterns across tissues. In conclusion, our study provides a useful tool for revealing tissue-level aging processes and evaluating the benefits of potential interventions.

Our results suggest that SCALE has great potential in various applications (Fig. 6). 1) Predicting relative biological age and helping to avoid bias when using chronological age as the target in aging studies. 2) Prioritizing tissue-specific aging genes and advancing our understanding of aging biology. 3) Assessing the efficacy of rejuvenation interventions. 4) Predicting AD and other aging-related diseases. 5) Quantifying aging at the single-cell level, identifying aging cell populations and aiding in the development of stem cell therapies. 6) Being integrated as a functional health assessment tool for the elderly.

SCALE is an aging quantifier that explicitly makes interpretability a priority and addresses the multilinearity problem using knowledge from literature. By encouraging the inclusion of well-recognized and experimentally confirmed aging genes (Fig. 3B), SCALE is more explainable and robust to spurious correlation and technical variability, as demonstrated in the cross-dataset validation in aging-related diseases and rejuvenating interventions data (Fig. 4). Indeed, we showed that SCALE became less interpretable (reduced overlap with well-documented aging-associated genes) (Supplemental Fig. S24A) and less accurate (subpar performance on rejuvenation data) (Supplemental Fig. S24B, C) when the Aging Map initialization was replaced by randomly selected

genes in the iterative feature selection step. This confirms the importance of including validated aging genes during training. Still, caution should be exercised when interpreting the SCALE scores. While SCALE and other cellular measures of aging utilize known aging-associated pathways, such as the accumulation of somatic mutations or DNA methylation, they are most likely only capturing a small aspect of biological aging. It remains an open question whether the predictive power of SCALE and other quantifiers implies causation, and whether a change in the SCALE score reflects the shifts most relevant for therapeutic purposes. In the future, it is therefore necessary to test out aging clocks on more data in larger cohorts and expand our understanding of anti-aging interventions in a quantitative manner.

From a method development perspective, SCALE can be further improved in many ways. 1) Firstly, all single-cell-level aging clocks, including SCALE, face a common problem: there are often more cells in the training dataset than individuals, meaning the model treats the chronological age as discrete rather than a variable continuously sampled from the population. While the problem may be mitigated by using carefully designed cohorts with larger sample sizes, it is also possible to increase the granularity of age by computationally infer unobserved sample points. As a proof-of-concept experiment, we showed that optimal transport, a mathematical framework that measures the distance between distributions (also known as the Wasserstein distance), can be applied to smoothly interpolate between discrete chronological age groups and simulate the expected distribution of the feature of interest (e.g., the SCALE score) at a new time point (Methods; Supplemental Fig. S25). 2) Secondly, SCALE improves robustness by down-weighting genes that are more affected by sampling noise in scRNA-seq data. It is however possible to directly model

dropouts as missing data and incorporate uncertainty when building the predictor (Ibrahim et al. 2005). Alternatively, technical noise and outliers can be removed through data imputation before training and applying the model (Hou et al. 2020). 3) Thirdly, an increase or decrease in expression during aging does not mean that the gene is responsible for aging. Although our guided forward selection approach encourages selection of well-documented aging genes, it does not directly model causality. Improving the quality of casual aging gene databases and applying casual inference methodology may help to build better aging models. 4) Finally, a better measure of biological aging is likely to be an ensemble of clocks operating at different scales (Belsky et al. 2017; Li et al. 2020). This requires integrating aging quantifiers trained on various data type and aspects of aging, including multi-omics clocks, frailty index, iAGE (Sayed et al. 2021), 3-D facial-image (Xia et al. 2020), etc. Designing such an integration approach in an interpretable way is critical for understanding fundamental aging biology and for guiding rejuvenation strategies.

Methods

Collection of single-cell aging omics datasets

In this study, we collected and analyzed publicly available single-cell aging RNA-seq datasets across multiple tissues and organs in mice, rats, and humans, covering aging, rejuvenations, and age-related diseases. These datasets include the *Tabula Muris Senis* (TMS) dataset of mice (Tabula Muris 2020), the heterochronic parabiosis datasets of mice (Palovics et al. 2022), the single-nucleus dataset of mouse Alzheimer's disease (AD) model (Zhou et al. 2020), and the caloric restriction (CR) dataset of rats (Ma et al. 2020). To compare transcriptomic aging clocks with epigenetic clocks, we additionally collected and analyzed the scM&T-seq data of mouse muscle stem cells that contains

simultaneously measured single-cell DNA methylation and transcriptome profiles (Hernando-Herraez et al. 2019).

Specifically, for the TMS and heterochronic parabiosis datasets, we downloaded the preprocessed and annotated H5AD files from the corresponding data portals (https://figshare.com/articles/Processed_files_to_use_with_scanpy_/8273102 and https://figshare.com/projects/Molecular_hallmarks_of_heterochronic_parabiosis_at_single_cell_resolution/127628, respectively). For the other two single-cell mouse brain aging datasets (Dulken et al. 2019; Buckley et al. 2023), we downloaded the preprocessed and annotated files from https://doi.org/10.5281/zenodo.7145399.

For the AD dataset and the CR dataset, we downloaded the count and barcode matrices from GEO using the accession numbers GSE137869 and GSE140511, respectively. For the scM&T-seq dataset, we downloaded the raw expression and methylation profiles from GEO using the accession numbers GSE121364 and GSE121436, respectively.

Single-cell RNA-seq data preprocessing

We used the R package Seurat version 3.1.5 (Stuart et al. 2019) to analyze all scRNA-seq data. Since different aging datasets were generated by different platforms for different biological systems, we generally followed the preprocessing steps reported in the original publications for each dataset and compared (e.g., disease vs healthy control) within datasets only. For datasets where preprocessed data were available, such as TMS, we proceeded with the provided preprocessed and annotated data in downstream analyses. For datasets with raw data and annotations only (AD, CR,

scM&T-seq), according to the methods and codes from their original articles, we used Seurat to conduct quality control, normalization (using method “LogNormalize”, set “scale.factor” = 10000), and dimension reduction (PCA, t-SNE, and UMAP) for visualization. The heterochronic parabiosis datasets provide filtered raw counts data, which were then processed following the same normalization and visualization steps.

Searching aging-related genes in mice and humans from existing databases and the literature

We first downloaded and combined five existing aging gene databases, namely "CSGene" (Zhao et al. 2016), "Aging Atlas" (Aging Atlas 2021), "HAGR" (Tacutu et al. 2018), "AgeFactDB" (Huhne et al. 2014), and "Longevity Map" (Budovsky et al. 2013), for mouse and human genes individually. Next, we searched for genes associated with aging-related Gene Ontology (GO) terms, including “Aging”, “Cell Age”, “Cellular Senescence”, “Regulation of Cell Aging”, and “Replicative Senescence” (The Gene Ontology Consortium et al. 2000; The Gene Ontology Consortium 2021). To incorporate more recent research efforts, we screened PubMed abstracts for additional aging-associated genes that have been verified experimentally. Specifically, we first searched for potential aging-related genes through the NCBI gene database (https://www.ncbi.nlm.nih.gov/gene/) using the keywords “aging” and “senescence”. For each of the potential genes, we searched through PubMed using the gene name and “aging” or “cell senescence” as keywords. We then downloaded the abstracts of the top 20 related articles and curated them manually to generate the final list of literature-based, experimentally verified aging-related genes. Together, we created Aging Map (http://sysomics.com/AgingMap/), a manually curated online database of human and mouse aging-

related genes.

Global gene expression

To measure the global transcriptional activity of a cell, we defined the global gene expression as the log-transformed mean counts per gene in a cell:

$$Global\ gene\ expression(cell) = log\left(\frac{raw\ counts\ number}{detected\ gene\ number} + 1\right).$$

Note that the global gene expression measure depends both on cellular translational activities and the sequencing depth. In the absence of spike-in controls, we only compared the global expression value across samples generated by the same sequencing platform (e.g., droplet-based scRNA-seq) within the same dataset (e.g., TMS tissues) to mitigate potential variability in sequencing depth. We evaluated the relationship between global gene expression and chronological age (month) in each condition using linear regression and assessed the statistical significance of the coefficient using two-sided Student's *t*-test.

Single-cell entropy

We quantified single-cell transcriptional variability using Shannon entropy. Specifically, in each sample, we first removed genes that were expressed in less than 5% of cells. Then, in each cell, we divided the total *n* expressed genes into *m* groups with equal group intervals based on their expression value. The number of gene groups *m* was determined by the Rice rule and was the same for all cells. Further, we considered each gene as a categorical random variable with *m* potential outcomes and, in each cell, calculated the event probability as the proportion of genes in each of the *m* expression group. The single-cell Shannon entropy can then be defined as

$$Entropy(cell) = \sum_{i}^{m} Pr(i;\ cell) \times logPr(i; cell).$$

Here $Pr(i;\ cell)$ is the frequency of the i-th gene expression group in the given cell. We again calculated the single-cell entropy for each cell in each sample and evaluated the association between entropy and chronological age using the Kruskal–Wallis test.

Selecting tissue-specific aging genes

The first step of the SCALE pipeline is to select aging-associated genes for each tissue using guided forward feature selection. At each iteration, we fitted an Elastic Net model using the R package *glmnet* version 3.0.2 (Friedman et al. 2010) to evaluate the predictive performance of the current aging gene set. To avoid overfitting and underfitting, we divided all samples into training and test sets, tuned model hyperparameters using 10-fold cross-validation on the training set, and internally evaluated the prediction accuracy on the test set. Our procedure can be summarized into the following steps: 1) Remove genes that were not expressed in a sample (i.e., detected in less than 5% of cells); 2) Perform an initial Elastic Net regression to rank gene by their predictive performance on chronological age; 3) Starting with Aging Map genes as the initial gene set (seed), at each iteration, add more genes to the set according to their rankings and fit a new Elastic Net model against chronological age using raw expression counts of genes in the new set; 4) Repeat step 3 until the mean square error (MSE) decreases to a turning point (i.e., defined using the Elbow method or by loss convergence); 5) Select the final aging genes in the last regression model. In our study, we selected the top 100 genes (by coefficient) of each tissue for downstream analysis (Supplemental Table 1).

563 **Calculating the SCALE score**

564 After aging gene selection, for each tissue we now have a set of aging genes with known
565 expression direction of age association (up-regulated or down-regulated during aging). For each
566 single cell in a given tissue, we calculated the SCALE score using the weighted average of
567 normalized expression of genes in the tissue-specific aging gene set:

568 $$SCALE\ score(cell) = \sum_{g\ \in\ aging\ gene\ set} Sign_g \times Prop(g;\ t(cell)) \times Expr(g; cell).$$

569 Here the $Sign_g$ represents the age-association direction (+1 for up-regulated genes and -1 for down-
570 regulated genes) learned from the training data in the gene selection step, $Prop(g;\ t(cell))$ is the
571 proportion of cells in the target dataset that expressed gene *g*, and $Expr(g; cell)$ is the *Z*-score
572 (zero-centered, variance-standardized) of the normalized gene expression level after preprocessing.

573

574 Different from the final Elastic Net prediction and other chronological-age-based aging quantifiers,
575 SCALE generates a relative aging score for each cell based on tissue-specific aging programs.
576 Alternatively, we can view SCALE as a prediction model that replaces regression coefficients
577 learned during feature selection with gene-specific weights to mitigate dropout and other batch
578 effects. Our benchmark demonstrated that the final SCALE score was less sensitive to the training
579 set and other technical variability, therefore more generalizable across datasets (Supplemental Fig.
580 S2). We also hypothesized that the slight decoupling of SCALE and the chronological age help it
581 better capture the biological aging effects in real-world applications.

582 **Evaluating the statistical significance of the SCALE score**

583 We constructed a permutation test to estimate the amount of aging-related information captured by

584 the selected set of aging genes and the statistical significance of SCALE scores in each sample. We
585 first replaced the aging gene set with the same number (100 by default) of expressed genes, which
586 were randomly selected through bootstrap sampling without replacement. Next, we applied the
587 same procedure used to calculate the SCALE score to compute a score based on these genes. By
588 repeating this process 1,000 times, we obtained a permutation-based null distribution of the SCALE
589 score and used it to calculate the empirical p-value of any observed score.

590 **Testing the robustness of the SCALE score**

591 Initially, we set raw count values of given proportions (ranging from 10% to 90%) of genes to zero.
592 Next, we re-normalized the expression matrix to adjust for varying sequencing depths (using
593 method “LogNormalize” in Seurat, set “scale.factor” = 10000), and then re-computed SCALE scores
594 of cells. Finally, we determined the correlation between the original scores and the newly computed
595 ones.

596 **Evaluating the statistical significance of the correlation between the SCALE score and**
597 **chronological age**

598 We again utilized permutation as an empirical statistical approach to compute the null distribution of
599 the correlation coefficient between the SCALE score and chronological age in each set of cells. For
600 each set of cells to inspect (e.g., T cells in the limb muscle), we randomly sampled the same
601 number of cells from the same sample (e.g., limb muscle). Then, we calculated the correlation
602 coefficient between SCALE scores and chronological ages of randomly selected cells. By repeating
603 this process 1,000 times, we obtained a permutation-based null distribution of the correlation

coefficient and used it to calculate the empirical p-value of the observed correlation in the set of interest. We further categorized the set of interest into three groups based on whether its correlation is significantly higher or lower than the tissue background (p-value < 0.05; Fig. 5F).

Single-cell methylation data analysis and epigenetic age calculation

We followed the pipeline and scripts provided by the original publication (https://github.com/alex-trapp/scAge) (Trapp et al. 2021) to reconstruct the scAge model and analyzed the single-cell methylation data of muscle stem cells (GSE121436) (Hernando-Herraez et al. 2019) to predict the single-cell epigenetic ages. To compare the robustness against extreme outliers, we did not filter out cells with lower CpG coverage, as opposed to the quality control steps described in the original paper.

Interpolation of SCALE scores by optimal transport

Optimal transport is a mathematical framework that can be applied to describe the distance between probabilistic distributions. The goal of interpolation is to infer the distribution of a variable of interest at an unobserved data point given the distributions observed at the start and end points. Starting with the observed distributions of SCALE scores at various chronological ages, here we aimed to interpolate the SCALE score distribution at the middle point of a given time interval where no real observation was available. We assumed that the distribution of the SCALE score at the middle point of an interval was the optimal transportation barycenter of distributions at the start and end points. In the one-dimensional case, we solved the barycenter problem using iterative Bregman algorithm with entropic regularization (Benamou et al. 2015).

Software availability

The SCALE pipeline and associated codes and scripts for analyses described in this study are available on GitHub (https://github.com/ChengLiLab/SCALE) and as Supplemental Codes.

Competing interest statement

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work was supported by National Key Research and Development Program of China (2021YFA1100300 to C.L.), National Natural Science Foundation of China (32288102 to C.L., 32025006 to C.L., 31871266 to C.L., 62173338 to H.C.), and Beijing Nova Program of Science and Technology (20220484198 to H.C.). Part of the data analysis was performed on the High Performance Computing Platform of the Center for Life Sciences, Peking University. J.S. was supported by the Edward P. Evans Center for Myelodysplastic Syndromes at Columbia University. We thank Yao Yao for critical comments on the manuscript. We would also like to thank the aging research community for generating and making datasets available online.

Author contributions: C.L. and H.C. conceptualized and supervised the study. S.M. collected resources and created the Aging Map database. S.M., J.S., and L.W. performed data analysis and interpreted the results. S.M. and J.S. prepared and revised the manuscript. All authors read and approved the manuscript.

References

Aging Atlas C. 2021. Aging Atlas: a multi-omics database for aging biology. *Nucleic Acids Res* **49**: D825-D830.

Ballard C, Gauthier S, Corbett A, Brayne C, Aarsland D, Jones E. 2011. Alzheimer's disease. *Lancet* **377**: 1019-1031.

Belsky DW, Caspi A, Arseneault L, Baccarelli A, Corcoran DL, Gao X, Hannon E, Harrington HL, Rasmussen LJ, Houts R et al. 2020. Quantification of the pace of biological aging in humans through a blood test, the DunedinPoAm DNA methylation algorithm. *Elife* **9**.

Belsky DW, Caspi A, Houts R, Cohen HJ, Corcoran DL, Danese A, Harrington H, Israel S, Levine ME, Schaefer JD et al. 2015. Quantification of biological aging in young adults. *Proc Natl Acad Sci U S A* **112**: E4104-4110.

Belsky DW, Moffitt TE, Cohen AA, Corcoran DL, Levine ME, Prinz JA, Schaefer J, Sugden K, Williams B, Poulton R et al. 2017. Eleven Telomere, Epigenetic Clock, and Biomarker-Composite Quantifications of Biological Aging: Do They Measure the Same Thing? *American Journal of Epidemiology* **187**: 1220-1230.

Benamou J-D, Carlier G, Cuturi M, Nenna L, Peyré G. 2015. Iterative Bregman Projections for Regularized Transportation Problems. *SIAM Journal on Scientific Computing* **37**: A1111-A1138.

Blau HM, Cosgrove BD, Ho AT. 2015. The central role of muscle stem cells in regenerative failure with aging. *Nature medicine* **21**: 854-862.

Brazhnik K, Sun S, Alani O, Kinkhabwala M, Wolkoff AW, Maslov AY, Dong X, Vijg J. 2020. Single-cell analysis reveals different age-related somatic mutation profiles between stem and differentiated cells in human liver. *Sci Adv* **6**: eaax2659.

Buckley MT, Sun ED, George BM, Liu L, Schaum N, Xu L, Reyes JM, Goodell MA, Weissman IL, Wyss-Coray T et al. 2023. Cell-type-specific aging clocks to quantify aging and rejuvenation in neurogenic regions of the brain. *Nature Aging* **3**: 121-137.

Budovsky A, Craig T, Wang J, Tacutu R, Csordas A, Lourenco J, Fraifeld VE, de Magalhaes JP. 2013. LongevityMap: a database of human genetic variants associated with longevity. *Trends Genet* **29**: 559-560.

Castellano JM, Kirby ED, Wyss-Coray T. 2015. Blood-Borne Revitalization of the Aged Brain. *JAMA Neurol* **72**: 1191-1194.

Chatsirisupachai K, Palmer D, Ferreira S, de Magalhães JP. 2019. A human tissue-specific transcriptomic analysis reveals a complex relationship between aging, cancer, and cellular senescence. *Aging Cell* **18**: e13041.

Chen W, Teschendorff AE. 2019. Estimating Differentiation Potency of Single Cells Using Single-Cell Entropy (SCENT). *Methods Mol Biol* **1935**: 125-139.

Dulken BW, Buckley MT, Navarro Negredo P, Saligrama N, Cayrol R, Leeman DS, George BM, Boutet SC, Hebestreit K, Pluvinage JV et al. 2019. Single-cell analysis reveals T cell infiltration in old neurogenic niches. *Nature* **571**: 205-210.

Fernández Á F, Sebti S, Wei Y, Zou Z, Shi M, McMillan KL, He C, Ting T, Liu Y, Chiang WC et al. 2018. Disruption of the beclin 1-BCL2 autophagy regulatory complex promotes longevity in

684 mice. *Nature* **558**: 136-140.
685 Finotello F, Di Camillo B. 2015. Measuring differential gene expression with RNA-seq: challenges
686 and strategies for data analysis. *Brief Funct Genomics* **14**: 130-142.
687 Flohr Svendsen A, Yang D, Kim K, Lazare S, Skinder N, Zwart E, Mura-Meszaros A, Ausema A, von
688 Eyss B, de Haan G et al. 2021. A comprehensive transcriptome signature of murine
689 hematopoietic stem cell aging. *Blood* **138**: 439-451.
690 Friedman J, Hastie T, Tibshirani R. 2010. Regularization Paths for Generalized Linear Models via
691 Coordinate Descent. *J Stat Softw* **33**: 1-22.
692 Galkin F, Mamoshina P, Aliper A, de Magalhaes JP, Gladyshev VN, Zhavoronkov A. 2020.
693 Biohorology and biomarkers of aging: Current state-of-the-art, challenges and opportunities.
694 *Ageing Research Reviews* **60**.
695 Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, Klotzle B, Bibikova M, Fan JB, Gao Y
696 et al. 2013. Genome-wide methylation profiles reveal quantitative views of human aging
697 rates. *Mol Cell* **49**: 359-367.
698 Harper S. 2014. Economic and social implications of aging societies. *Science* **346**: 587-591.
699 Harris SE, Riggio V, Evenden L, Gilchrist T, McCafferty S, Murphy L, Wrobel N, Taylor AM, Corley J,
700 Pattie A et al. 2017. Age-related gene expression changes, and transcriptome wide
701 association study of physical and cognitive aging traits, in the Lothian Birth Cohort 1936.
702 *Aging (Albany NY)* **9**: 2489-2503.
703 Hernando-Herraez I, Evano B, Stubbs T, Commere P-H, Jan Bonder M, Clark S, Andrews S,
704 Tajbakhsh S, Reik W. 2019. Ageing affects DNA methylation drift and transcriptional cell-to-
705 cell variability in mouse muscle stem cells. *Nat Commun* **10**: 4361.
706 Horvath S. 2013. DNA methylation age of human tissues and cell types. *Genome Biol* **14**: R115.
707 Hou W, Ji Z, Ji H, Hicks SC. 2020. A systematic evaluation of single-cell RNA-sequencing
708 imputation methods. *Genome Biol* **21**: 218.
709 Huhne R, Thalheim T, Suhnel J. 2014. AgeFactDB--the JenAge Ageing Factor Database--towards
710 data integration in ageing research. *Nucleic Acids Res* **42**: D892-896.
711 Ibrahim JG, Chen MH, Lipsitz SR, Herring AH. 2005. Missing-data methods for generalized linear
712 models: A comparative review. *J Am Stat Assoc* **100**: 332-346.
713 Janssens GE, Meinema AC, Gonzalez J, Wolters JC, Schmidt A, Guryev V, Bischoff R, Wit EC,
714 Veenhoff LM, Heinemann M. 2015. Protein biogenesis machinery is a driver of replicative
715 aging in yeast. *Elife* **4**: e08527.
716 Junnila RK, List EO, Berryman DE, Murrey JW, Kopchick JJ. 2013. The GH/IGF-1 axis in ageing
717 and longevity. *Nat Rev Endocrinol* **9**: 366-376.
718 Kosaki R, Horikawa R, Fujii E, Kosaki K. 2018. Biallelic mutations in LARS2 can cause Perrault
719 syndrome type 2 with neurologic symptoms. *Am J Med Genet A* **176**: 404-408.
720 Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, Hou L, Baccarelli AA, Stewart JD,
721 Li Y et al. 2018. An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany
722 NY)* **10**: 573-591.
723 Li X, Ploner A, Wang Y, Magnusson PK, Reynolds C, Finkel D, Pedersen NL, Jylhava J, Hagg S.
724 2020. Longitudinal trajectories, correlations and mortality associations of nine biological ages
725 across 20-years follow-up. *Elife* **9**.
726 Lodato MA, Rodin RE, Bohrson CL, Coulter ME, Barton AR, Kwon M, Sherman MA, Vitzthum CM,
727 Luquette LJ, Yandava CN et al. 2018. Aging and neurodegeneration are associated with

increased mutations in single human neurons. *Science* **359**: 555-559.

Lopez-Otin C, Blasco MA, Partridge L, Serrano M, Kroemer G. 2013. The hallmarks of aging. *Cell* **153**: 1194-1217.

López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. 2023. Hallmarks of aging: An expanding universe. *Cell*.

Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, Hou L, Baccarelli AA, Li Y, Stewart JD et al. 2019. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging (Albany NY)* **11**: 303-327.

Ma S, Sun S, Geng L, Song M, Wang W, Ye Y, Ji Q, Zou Z, Wang S, He X et al. 2020. Caloric Restriction Reprograms the Single-Cell Transcriptional Landscape of Rattus Norvegicus Aging. *Cell* **180**: 984-1001 e1022.

MacArthur BD, Lemischka IR. 2013. Statistical mechanics of pluripotency. *Cell* **154**: 484-489.

Martinez-Jimenez CP, Eling N, Chen H-C, Vallejos CA, Kolodziejczyk AA, Connor F, Stojic L, Rayner TF, Stubbington MJT, Teichmann SA et al. 2017. Aging increases cell-to-cell transcriptional variability upon immune stimulation. *Science* **355**: 1433-1436.

Moaddel R, Ubaida-Mohien C, Tanaka T, Lyashkov A, Basisty N, Schilling B, Semba RD, Franceschi C, Gorospe M, Ferrucci L. 2021. Proteomics in aging research: A roadmap to clinical, translational research. *Aging Cell* **20**: e13325.

Nikolich-Zugich J. 2018. The twilight of immunity: emerging concepts in aging of the immune system. *Nat Immunol* **19**: 10-19.

Palovics R, Keller A, Schaum N, Tan W, Fehlmann T, Borja M, Kern F, Bonanno L, Calcuttawala K, Webber J et al. 2022. Molecular hallmarks of heterochronic parabiosis at single-cell resolution. *Nature* **603**: 309-314.

Partridge L, Deelen J, Slagboom PE. 2018. Facing up to the global challenges of ageing. *Nature* **561**: 45-56.

Peters MJ, Joehanes R, Pilling LC, Schurmann C, Conneely KN, Powell J, Reinmaa E, Sutphin GL, Zhernakova A, Schramm K et al. 2015. The transcriptional landscape of age in human peripheral blood. *Nat Commun* **6**: 8570.

Rutledge J, Oh H, Wyss-Coray T. 2022. Measuring biological age using omics data. *Nat Rev Genet* **23**: 715-727.

Sayed N, Huang Y, Nguyen K, Krejciova-Rajaniemi Z, Grawe AP, Gao T, Tibshirani R, Hastie T, Alpert A, Cui L et al. 2021. An inflammatory aging clock (iAge) based on deep learning tracks multimorbidity, immunosenescence, frailty and cardiovascular aging. *Nature Aging* **1**: 598-615.

Schumacher B, Pothof J, Vijg J, Hoeijmakers JHJ. 2021. The central role of DNA damage in the ageing process. *Nature* **592**: 695-703.

Scott AJ, Ellison M, Sinclair DA. 2021. The economic value of targeting aging. *Nature Aging* **1**: 616-623.

Shcherbina A, Larouche J, Fraczek P, Yang BA, Brown LA, Markworth JF, Chung CH, Khaliq M, de Silva K, Choi JJ. 2020. Dissecting murine muscle stem cell aging through regeneration using integrative genomic analysis. *Cell Rep* **32**: 107964.

Solovev I, Shaposhnikov M, Moskalev A. 2020. Multi-omics approaches to human biological age estimation. *Mech Ageing Dev* **185**: 111192.

Steffen KK, MacKay VL, Kerr EO, Tsuchiya M, Hu D, Fox LA, Dang N, Johnston ED, Oakes JA,

Tchao BN et al. 2008. Yeast life span extension by depletion of 60s ribosomal subunits is mediated by Gcn4. *Cell* **133**: 292-302.

Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 3rd, Hao Y, Stoeckius M, Smibert P, Satija R. 2019. Comprehensive Integration of Single-Cell Data. *Cell* **177**: 1888-1902 e1821.

Tabula Muris C. 2020. A single-cell transcriptomic atlas characterizes ageing tissues in the mouse. *Nature* **583**: 590-595.

Tacutu R, Thornton D, Johnson E, Budovsky A, Barardo D, Craig T, Diana E, Lehmann G, Toren D, Wang J et al. 2018. Human Ageing Genomic Resources: new and updated databases. *Nucleic Acids Res* **46**: D1083-D1090.

The Gene Ontology Consortium. 2021. The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res* **49**: D325-D334.

The Gene Ontology Consortium, Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS et al. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**: 25-29.

Thompson MJ, Chwialkowska K, Rubbi L, Lusis AJ, Davis RC, Srivastava A, Korstanje R, Churchill GA, Horvath S, Pellegrini M. 2018. A multi-tissue full lifespan epigenetic clock for mice. *Aging-Us* **10**: 2832-2854.

Trapp A, Kerepesi C, Gladyshev VN. 2021. Profiling epigenetic age in single cells. *Nature Aging* **1**: 1189-1201.

Xia X, Chen X, Wu G, Li F, Wang Y, Chen Y, Chen M, Wang X, Chen W, Xian B et al. 2020. Three-dimensional facial-image analysis to predict heterogeneity of the human ageing rate and the impact of lifestyle. *Nat Metab* **2**: 946-957.

Zhang L, Dong X, Lee M, Maslov AY, Wang T, Vijg J. 2019. Single-cell whole-genome sequencing reveals the functional landscape of somatic mutations in B lymphocytes across the human lifespan. *Proc Natl Acad Sci U S A* **116**: 9014-9019.

Zhang MJ, Pisco AO, Darmanis S, Zou J. 2021. Mouse aging cell atlas analysis reveals global and cell type-specific aging signatures. *Elife* **10**.

Zhao M, Chen L, Qu H. 2016. 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**: e2053.

Zhou Y, Song WM, Andhey PS, Swain A, Levy T, Miller KR, Poliani PL, Cominelli M, Grover S, Gilfillan S et al. 2020. Human and mouse single-nucleus transcriptomics reveal TREM2-dependent and TREM2-independent cellular responses in Alzheimer's disease. *Nat Med* **26**: 131-142.

Figure Legends

Figure 1. Schematic illustration of SCALE and the study design.

We designed SCALE to select interpretable aging-related genes and quantify tissue-specific transcriptomic aging in a robust and easily generalizable manner. In a nutshell, SCALE starts with a

core set of manually curated aging-associated genes (Aging Map, Top Left), iteratively expands the list to include additional features that improve age prediction (Top Right), and eventually uses the gene set to quantify aging for each cell through weighted average of gene expression (Bottom Left). By explicitly introducing Aging Map genes as seeds to guide feature selection, SCALE encourages the inclusion of both well-documented genes and genes that contain complementary information of aging. SCALE also benefits from its indirect modeling of chronological age, where the latter is only used to select genes rather than as the final prediction target. This makes SCALE less susceptible to technical noise and batch effects prevalent in single-cell RNA-seq data, allowing it to better capture biological age and the nuanced aging and anti-aging effects of disease and interventions (Bottom Right).

Figure 2. Aging Map: a manually curated database of aging-related transcriptomic features.

(A) Summary of data sources used to build Aging Map, grouped by the scales of aging information covered by each collection.

(B) Number of aging-related genes collected from each data source.

(C) Characteristics of the 636 literature-confirmed aging-related genes in human. Left: overlaps with other databases listed in (A). Novel: not covered in any of the existing database. Singleton: covered by one database. >=2: covered by at least two databases. Right: GO Enrichment analysis showing the top10 GO terms associated with the literature-based aging genes, colored by the proportion of GO gene sets present in the list.

(D) Similar to (C) but showing the results for 460 literature-confirmed aging-related genes in mouse.

(E) Comparisons of human and mouse Aging Map genes. Left: Venn diagram of homologous human and mouse genes. Right: GO Enrichment analysis showing the top10 GO terms associated with aging genes shared between human and mouse, colored by the proportion of GO gene sets present in the list.

(F) Venn diagram comparing human age-associated genes identified in three large cohort studies. See Methods for a full description of each dataset.

(G) Overlaps between human Aging Map genes and the three data-driven gene lists in (F). The five genes shared by all sets: *CTSC, FOXO1, LMNA, PLAUR, RPS6KA1.*

(H) Heatmap of correlation between chronological age and two high-level transcriptomic features (Global Gene Expression and Single Cell Entropy) in each dataset. *** shows adjusted p-value < 0.001; ** shows adjusted p-value < 0.01; and * shows adjusted p-value < 0.05 using linear regression for Global Gene Expression and Kruskal–Wallis test for Single Cell Entropy. All p-values were corrected by the Benjamini- Hochberg procedure.

Figure 3. Tissue-specific quantification of aging in single-cell transcriptomic data using knowledge- guided feature selection.

(A) Illustration of the SCALE pipeline applied to the *Tabula Muris Senis* data.

(B) Boxplot showing overlaps between genes in Aging Map and aging genes selected by two methods, namely aging clocks of Buckley et al and SCALE. Each data point represents the gene set selected for a particular cell type (Buckley et al) or tissue (SCALE). ***p-value < 0.001 using a two-sided unpaired Student's *t*-test.

(C) UpSet plot illustrating the intersection of gene sets used by the six cell-type-specific chronological aging clocks of Buckley et al and the SCALE brain model. Genes selected by SCALE are highlighted in orange. (D) Uniform manifold approximation and projection (UMAP) plot of all cells in Kidney. Different colors show different cell types.

(E) SCALE scores of kidney cells in each age group. Data are presented as the mean ± s.d. ***p-value < 0.001 using a two-sided unpaired Student's *t*-test.

(F) SCALE scores of limb muscle cells in each age group. Data are presented as the mean ± s.d. ***p-value < 0.001 using a two-sided unpaired Student's *t*-test.

(G) Boxplot showing the mutation burden (inferred from full-length scRNA-seq data) of limb muscle cells in each age group. Data are presented as the mean ± s.d. ***p-value < 0.001 using a two-sided unpaired Student's *t*-test.

(H) Relationship between the mean number of somatic mutations in genes and the SCALE score of limb muscle cells. Cells were divided into 10 groups based on their SCALE scores (from low to high) for better visualization. The yellow line shows the linear regression model results fitted to the data with the corresponding R and p-value. Data are presented as the mean ± s.d.

(I) Relationship between the mean number of somatic mutations in genes and the SCALE score residual (after regressing out chronological age) of limb muscle cells. Cells were divided into 10 groups based on their SCALE score residuals (from low to high) for better visualization. Data are presented as the mean ± s.e.. R and p-value were calculated by fitting linear regression model for each age group individually.

(J) Pie chart showing multi-tissue, multi-age-group correlation analysis results between the mean number of somatic mutations in genes and the SCALE score residual. "Decreased" represents

875 mutation number significantly decreases with the SCALE score residuals. “Increased” represents
876 mutation number significantly increases with the SCALE score residuals. “Not significant” represents
877 mutation number is not significantly linearly correlated with the SCALE score residuals.
878
879 **Figure 4. SCALE outperforms other single-cell aging clocks in assessing impacts of disease**
880 **and rejuvenating interventions on aging**
881 (A) Uniform manifold approximation and projection (UMAP) plot of muscle stem cells. Different
882 colors show different chronological ages.
883 (B) SCALE scores of muscle stem cells (using single-cell transcriptomic profiles from single-cell
884 M&T-seq). Data are presented as the mean ± s.d.
885 (C) ScAge predicted age of muscle stem cells (using single-cell DNA methylation data from single-
886 cell M&T- seq). Data are presented as the mean ± s.d.
887 (D) Uniform manifold approximation and projection (UMAP) plot of brain microglia cells from wild
888 type mice (normal, blue) and AD mouse models (AD, red).
889 (E) SCALE scores of brain microglia cells from wild type mice (normal, blue) and AD mouse models
890 (AD, red).
891 (F) Predicted chronological age of brain microglia cells by the method reported by Buckley MT, et al.
892 (G) Uniform manifold approximation and projection (UMAP) plot of brain cells from old mice (green),
893 young mice (magenta), old heterochronic mice (orange), and young heterochronic mice (purple).
894 (H) SCALE scores of brain cells from parabiosis mice. p-values using a two-sided unpaired
895 Student’s *t*-test are shown.

(I) Predicted chronological age of brain cell types (ependymal cell, macrophage, neuron, oligo pre cell, pericyte, and T cell) lacking clocks trained by Buckley MT, et al. This box plot shows the results calculated by the aNSC_NPC clock, which incorrectly predicted cells from young heterochronic mice with lower ages then these of young mice and was not able to distinguish cells from old heterochronic mice and old mice, developed by Buckley MT, et al. p-values using a two-sided unpaired Student's *t*-test are shown.

Figure 5. SCALE revealed cell-type specific aging patterns across tissues

(A) SCALE scores of mesenchymal stem cells from adipose cells in subcutaneous adipose tissue (SCAT). Data are presented as the mean ± s.d. ***p-value < 0.001 using a two-sided unpaired Student's *t*-test.

(B) SCALE scores of Leukocyte in pancreas. Data are presented as the mean ± s.d. ***p-value < 0.001 using a two-sided unpaired Student's *t*-test.

(C) Differences between the SCALE scores of T cells and the tissues in which they are located. "Older" means that the SCALE scores of T cells are significantly larger (p-value < 0.05) than all cells in the tissue. "Younger" means that the SCALE scores of T cells are significantly smaller (p-value < 0.05) than all cells in the tissue. p-values were calculated using two-sided unpaired Student's *t*-tests.

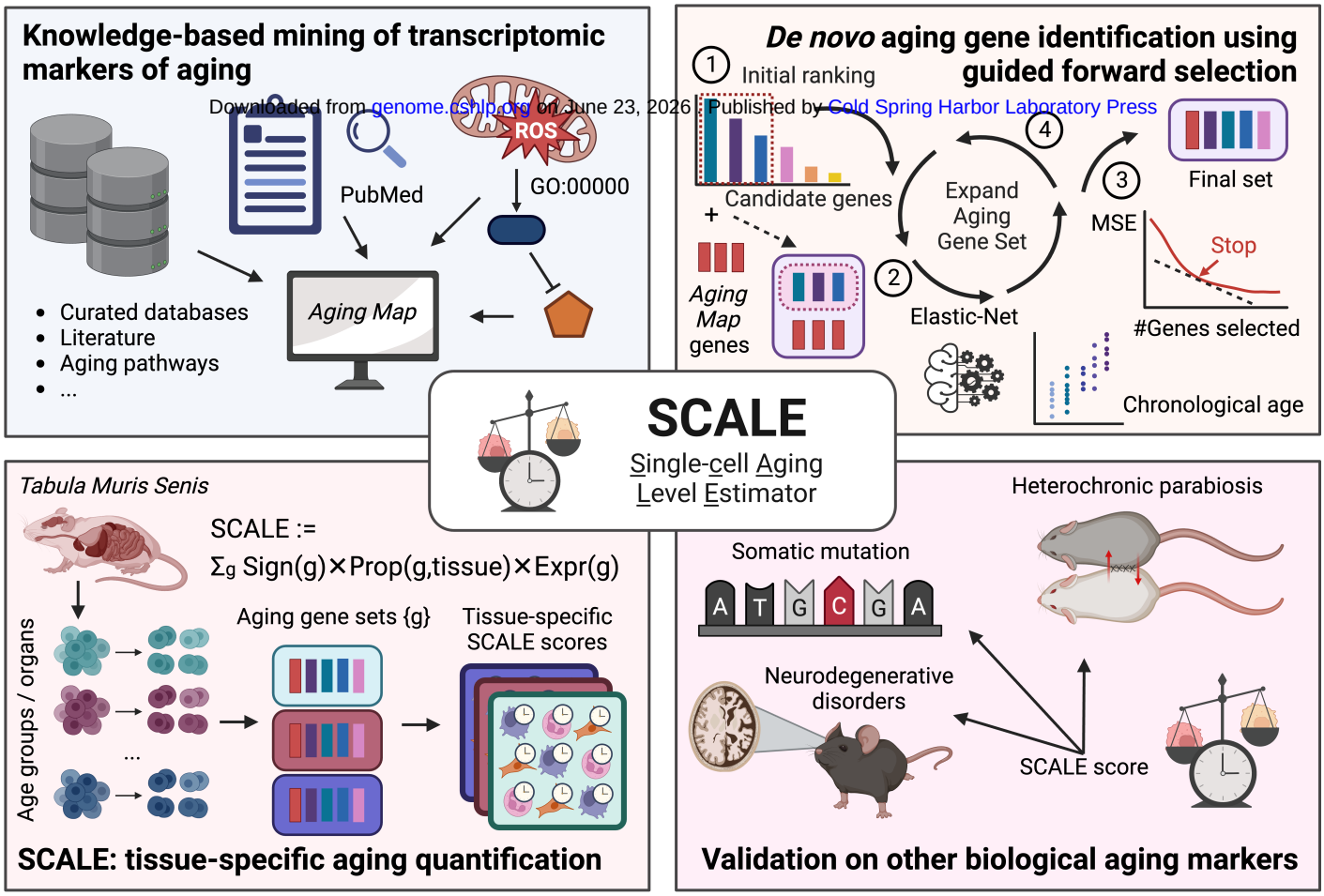
(D) Differences between the SCALE scores of B cells and the tissues in which they are located. "Older" means that the SCALE scores of B cells are significantly larger (p-value < 0.05) than all cells in the tissue. "Younger" means that the SCALE scores of B cells are significantly smaller (p-value < 0.05) than all cells in the tissue. p-values were calculated using two-sided unpaired Student's *t*-tests.

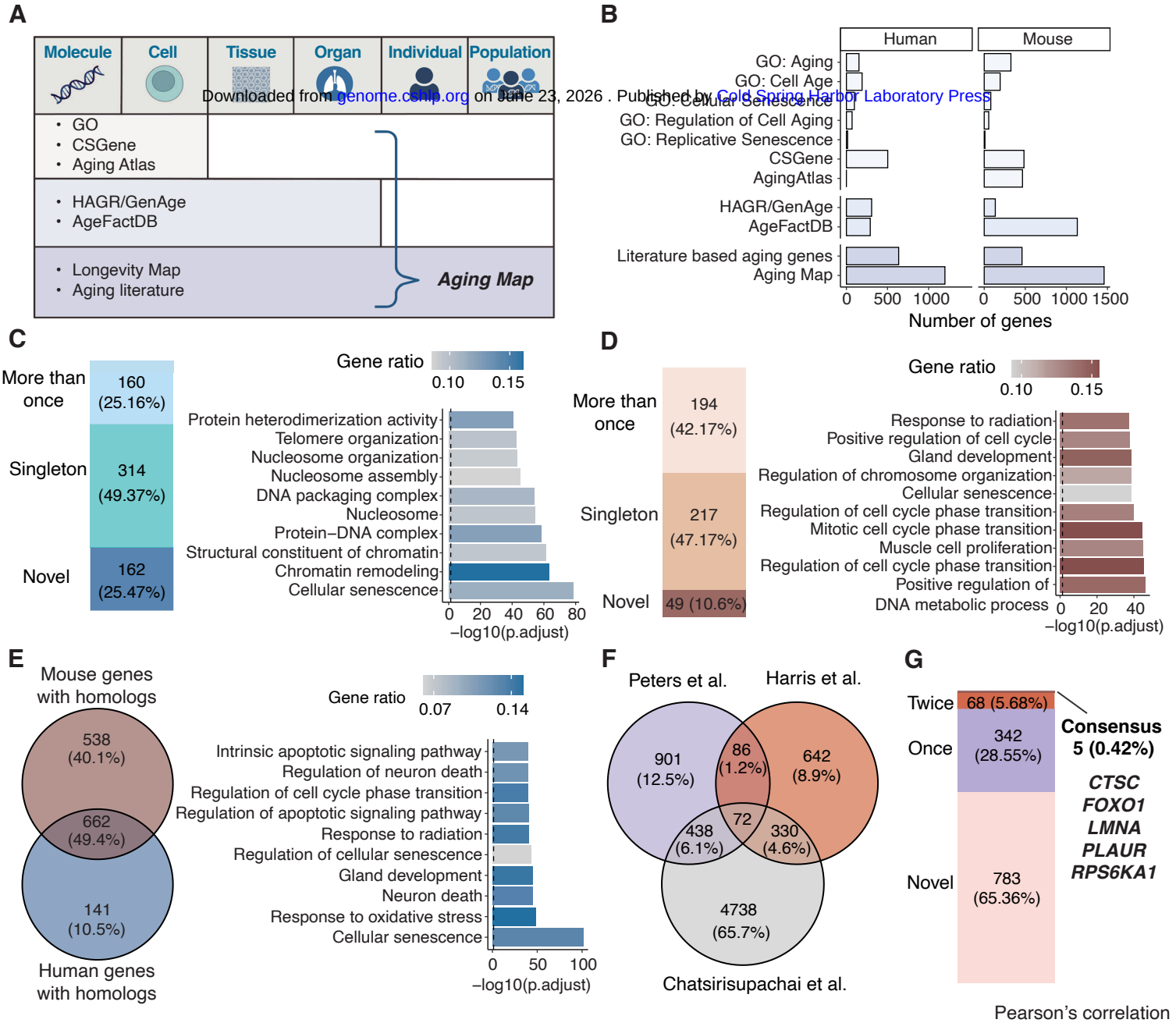
(E) Differences between the SCALE scores of endothelial cells and the tissues in which they are located. “Older” means that the SCALE scores of endothelial cells are significantly larger (p-value < 0.05) than all cells in the tissue. “Younger” means that the SCALE scores of endothelial cells are significantly smaller (p- value < 0.05) than all cells in the tissue. p-values were calculated using two-sided unpaired Student’s *t*-tests. (F) Schematic overview of the relationship between SCALE scores and chronological age. For each cell type or all cells in one tissue, we calculated the correlation coefficient to define the relationship between the SCALE score and chronological age. We observed cell-type-specific aging characteristics and tissue environment may influence cell-type aging patterns.

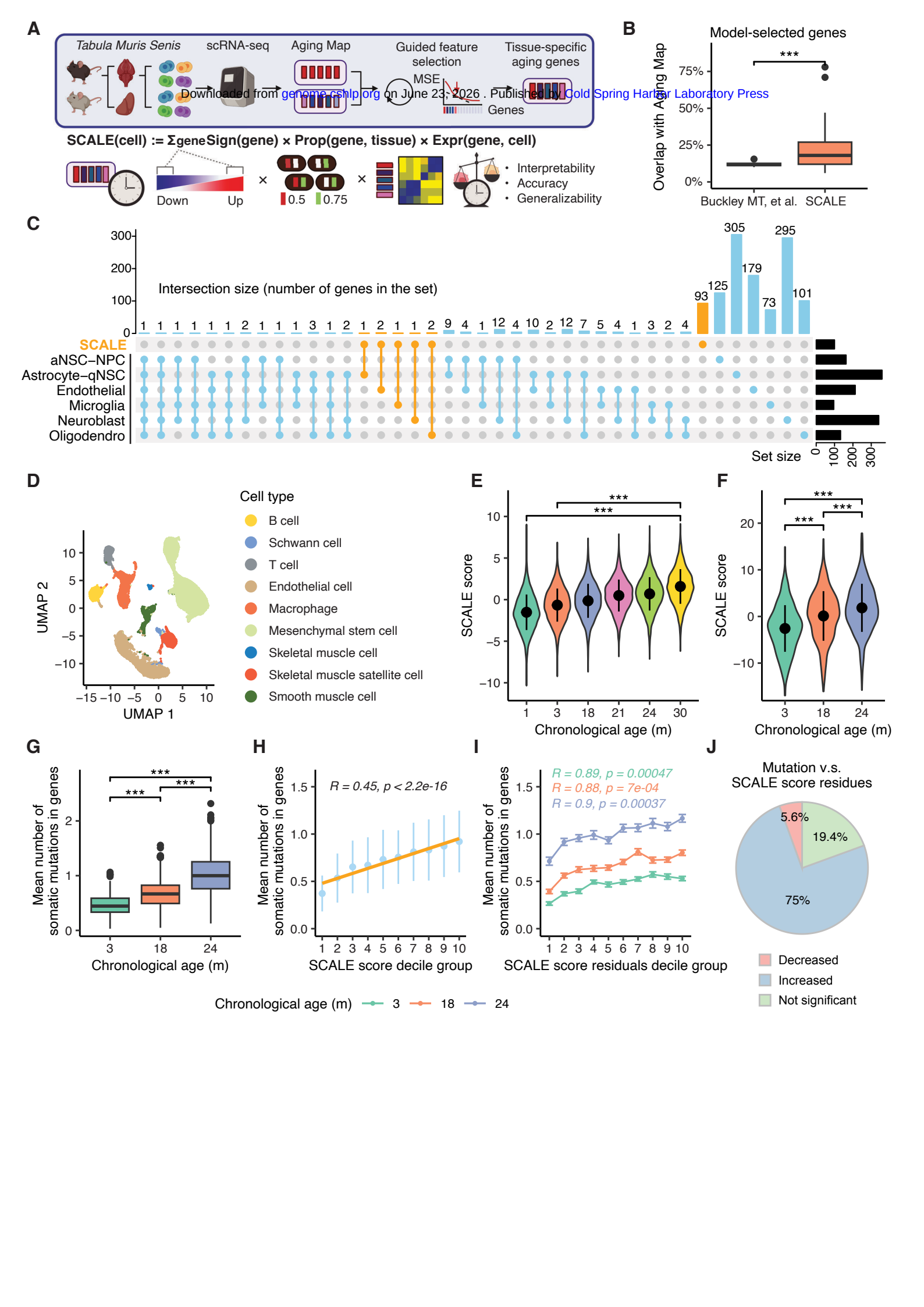
(G) Correlation coefficient between SCALE scores and chronological age of each cell type. The size of the circle shows the Pearson’s correlation coefficient. p-values were calculated using a permutation-based test. Orange indicates that the coefficient is significantly larger, and blue indicates that the coefficient is significantly smaller (p-value < 0.05).

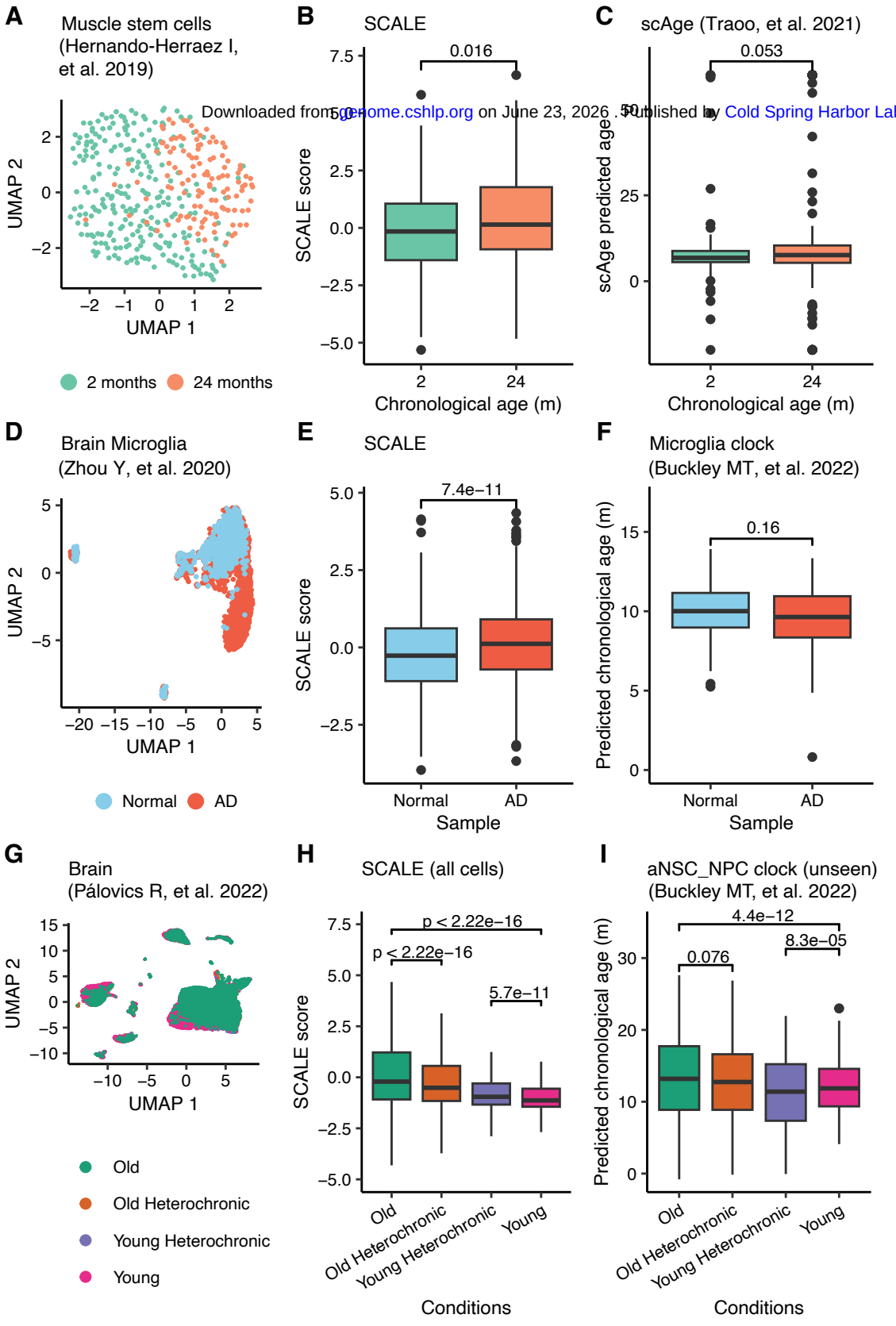
(H) Relationship between the correlation coefficient of each cell type and the correlation coefficient of their counterpart tissues. The R value and p-value are shown.

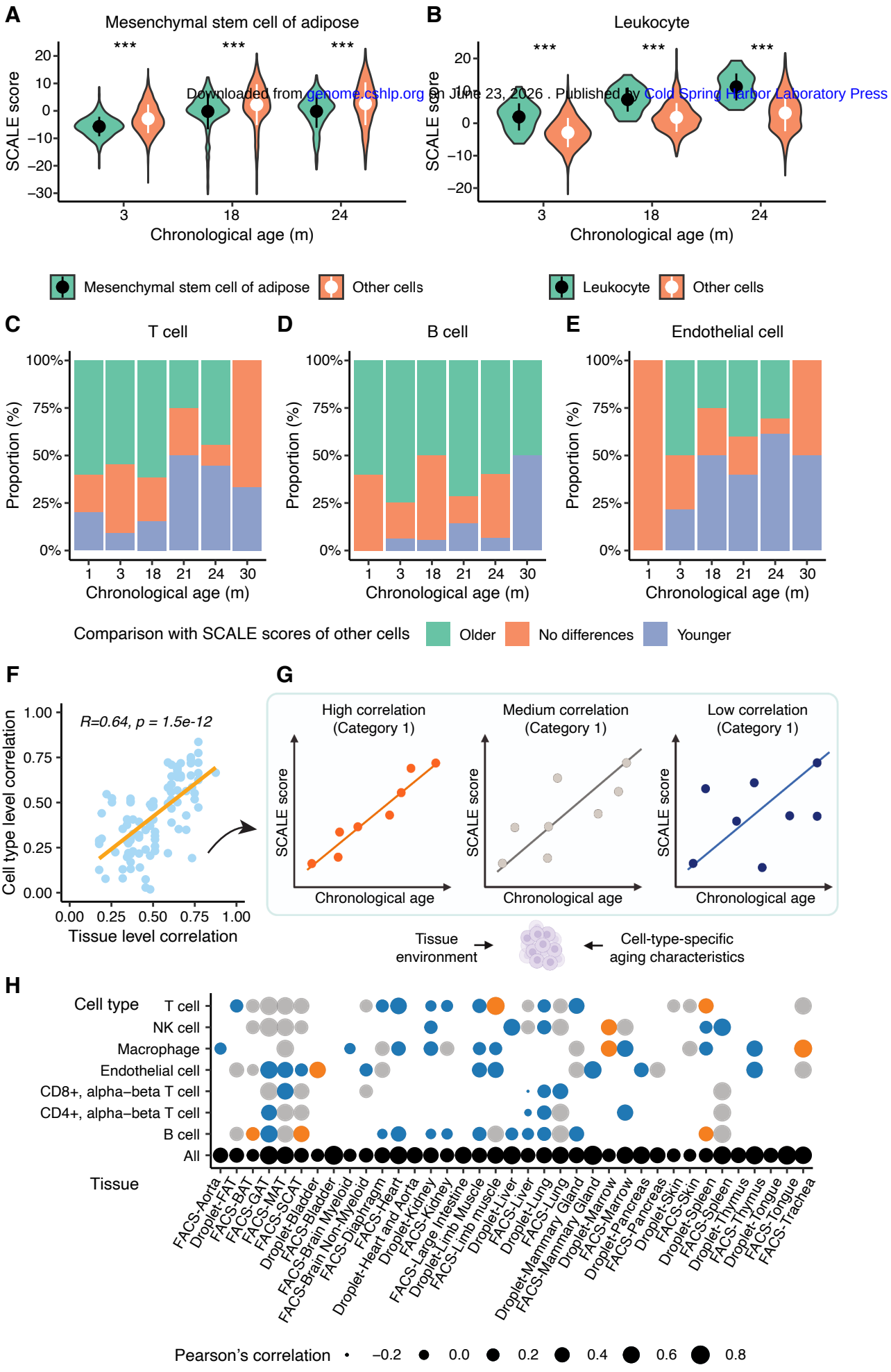
Figure 6. Potential applications of SCALE





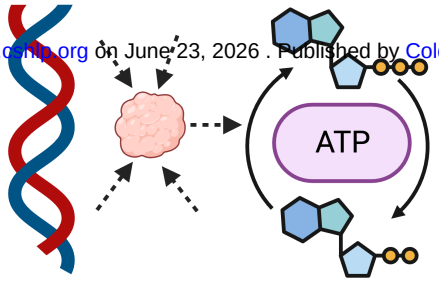




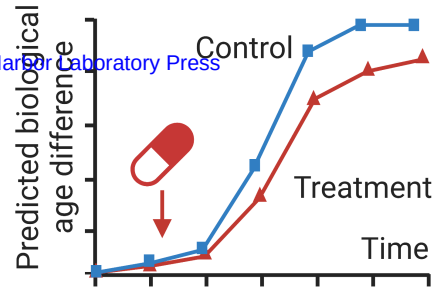


- **Predict relative biological age**

- **Tissue-specific aging programs**



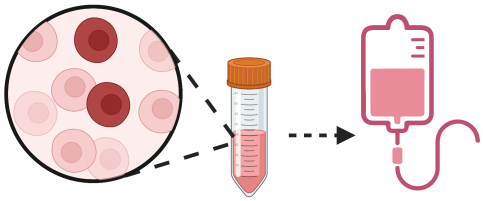
- **Anti-aging interventions**



- **Predict age-related diseases**

Risks: AD, muscle loss ...

- **Stem cell aging and therapies**



Aged cells

- **Functional health assessment**