Single-cell and spatial detection of senescent cells using DeepScence

Yilong Qu[1], Runze Dong[2,3], Liangcai Gu[2,3], Cliburn Chan[1], Jichun Xie[1,4], Carolyn Glass[5,6], Xiao-Fan Wang[7], Andrew B Nixon[8], and Zhicheng Ji[1,†]

[1]Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, NC, USA.
[2]Department of Biochemistry, University of Washington, Seattle, WA, USA.
[3]Institute for Protein Design, University of Washington, Seattle, WA, USA.
[4]Department of Mathematics, Duke University, Durham, NC, USA.
[5]Department of Pathology, Duke University School of Medicine, Durham, NC, USA.
[6]Department of Surgery, Duke University School of Medicine, Durham, NC, USA.
[7]Department of Pharmacology and Cancer Biology, Duke University School of Medicine, Durham, NC, USA.
[8]Department of Medicine, Duke University School of Medicine, Durham, NC, USA.
[†]Corresponding author. E-mail: zhicheng.ji@duke.edu

ABSTRACT

Accurately identifying senescent cells is essential for studying their spatial and molecular features. We developed DeepScence, a method based on deep neural networks, to identify senescent cells in single-cell and spatial transcriptomics data. DeepScence is based on CoreScence, a senescence-associated gene set we curated that incorporates information from multiple published gene sets. We demonstrate that DeepScence can accurately identify senescent cells in single-cell gene expression data collected both *in vitro* and *in vivo*, as well as in spatial transcriptomics data generated by different platforms, substantially outperforming existing methods.

Introductions

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 primarily serves as a defense mechanism against the uncontrolled cell division observed in cancer. However, the accumulation of senescent cells (SnCs) over time may contribute to the decline in the physical and functional integrity of tissues and organs, playing a significant role in aging [3, 4]. Additionally, SnCs are implicated in the development of various age-related diseases, such as osteoarthritis, pulmonary fibrosis, and Alzheimer's disease [5–8]. SnCs secrete pro-inflammatory proteins that alter the cellular microenvironment, affecting neighboring cells and contributing to chronic inflammation [9]. Given its profound impact on human health and its potential as a therapeutic target, it is important to understand cellular senescence and identify its associated biomarkers.

Given the relatively rare presence of SnCs in tissues [10], technologies capable of achieving single-cell or near single-cell resolution, such as single-cell RNA sequencing (scRNA-seq) and spatial transcriptomics (ST), have shown great promise in uncovering the molecular and spatial features of SnCs [10–12]. Several methods have been developed to identify SnCs in scRNA-seq data. However, their reliability remains questionable due to various challenges. The first type of method uses a single gene marker, such as CDKN1A (the gene that encodes the p21 protein) or CDKN2A (the gene that encodes the p16 protein), to identify SnCs[13, 14]. If the gene shows positive expression in a cell, the cell is identified as an SnC. However, an SnC may not exhibit positive expression of the marker gene due to dropout events [15], and the results of such methods can be highly unreliable due to the high levels of noise and sparsity in scRNA-seq data.

The second type of method scores the senescence level of a cell by ranking the genes in a senescence gene set (SnG) using methods such as AUCell and ssGSEA [16, 17]. However, these methods lack the ability to capture the complex nonlinear and combinatorial relationships between a cell's senescence level and the expression of genes in the SnG. Furthermore, identifying a reliable SnG is challenging, and we will demonstrate in this study that there is very little agreement across published SnGs. Additionally, genes in an SnG can either induce or inhibit senescence, but this directionality is ignored by these methods.

A recently published method, SenCID [18], is the third type of method that trains supervised machine learning models to predict SnCs. SenCID trains a support vector machine (SVM) using published bulk RNA-seq data from both normal and senescent cells. The trained model can predict SnCs in both single-cell and bulk RNA-seq datasets. However, a major limitation

of SenCID is that its training data were collected from *in vitro* studies. It is known that the molecular characteristics of cellular senescence can differ between *in vitro* and *in vivo* conditions[19]. Therefore, it is questionable whether the trained SenCID models can reliably predict SnCs in *in vivo* datasets. Additionally, trained SenCID models may not be directly applicable to datasets generated by ST technologies, such as 10x Xenium, which have different data distributions and number of profiled genes compared to scRNA-seq.

To address these issues, we developed DeepScence, an unsupervised machine learning model based on an autoencoder architecture for identifying SnCs. Autoencoders have been successfully applied to scRNA-seq data for dimensionality reduction, denoising, and imputation [20–22]. We designed a customized autoencoder that leverages CoreScence, a core senescence gene set we compiled, to efficiently capture senescence-related information. We systematically evaluated the performance of DeepScence and found that it accurately identifies SnCs in both *in vivo* and *in vitro* settings, as well as in both scRNA-seq and ST datasets. DeepScence substantially outperforms the three types of existing methods, including SenCID. DeepScence paves the way for further studies on the spatial and molecular features of SnCs in various contexts.

Results

Large discrepancies across existing SnGs

A reliable senescence gene set (SnG) is crucial for identifying senescent cells (SnCs). We surveyed nine published SnGs, including SenMayo [23], SenSig [24], CSGene [25], SeneQuest [2], CellAge [26], GenAge [27], SASP-related genes from De Cecco et al. [28], inflammatory network genes in senescence from Freund et al. [9], and the transcriptome signature of senescence from Casella et al. [29]. Strikingly, we found a large discrepancy across these SnGs. First, the number of genes included in each SnG varies greatly (Figure 1a). Larger SnGs, such as SenSig, contain more than a thousand genes, whereas smaller ones, like the Casella et al. SnG, include fewer than a hundred genes. Second, there is an extremely low level of agreement across SnGs (Figure 1b). The Jaccard index, which measures the degree of overlap between two sets, is below 0.2 for nearly all SnG pairs. Third, a gene reported by one SnG may not be consistently found in another. Out of the 2,966 senescence-related genes identified by at least one SnG, 2,052 (69.2%) were reported in only one SnG, and only 39 (1.3%) were reported in at least five SnGs (Figure 1c).

Since these SnGs were primarily compiled through literature curation by different research groups, the discrepancies could stem from variations in the scope and inclusion criteria used during the literature search process. As a result, SnC identification methods that rely on a single SnG may produce biased and irreproducible outcomes.

CoreScence: a core senescence gene set

To address this discrepancy, we compiled a new senescence gene set, CoreScence, which consists of 39 genes reported by at least five published gene sets (Figure 2a, Supplementary Table S1). The rationale is that genes consistently reported by multiple gene sets are less likely to be influenced by inclusion bias and are more likely to be associated with senescence. CoreScence includes several canonical marker genes for cellular senescence, such as CDKN1A and CDKN2A [1, 10]. It is important to note that, in addition to senescence, genes in CoreScence may also be involved in other functions or pathways, such as immunity or development. This will be accounted for in the DeepScence model described below.

To validate CoreScence, we collected information on genes with differential expression between senescent and non-senescent cells across various cell lines from published bulk RNA-seq studies [29, 30] (Figure 2b). We found that genes consistently reported by a greater number of published SnGs exhibit stronger differential signals between senescent and non-senescent cells (Figure 2c-d), suggesting that the genes in CoreScence are more likely to be genuinely associated with senescence.

DeepScence: a deep neural network for identifying SnCs

Building upon CoreScence, we developed DeepScence, a machine learning model based on an autoencoder architecture for identifying SnCs (Figure 3a, Methods). The input to DeepScence is a gene expression count matrix of the genes in the CoreScence gene set, and we assume that these counts follow zero-inflated negative binomial (ZINB) distributions [31]. The bottleneck layer of the autoencoder is designed to consist of two neurons that are almost uncorrelated. One neuron is responsible for capturing senescence-related information, while the other captures information unrelated to senescence. The output of DeepScence consists of three components corresponding to the three parameters of the ZINB distribution. After model fitting, the continuous value of the neuron capturing senescence information, referred to as the senescence score, is the final output of DeepScence. The continuous senescence scores can be optionally binarized through a permutation-based procedure to classify cells as either senescent or non-senescent (Methods).

DeepScence accurately identifies SnCs in *in vitro* scRNA-seq datasets

We applied DeepScence and three existing types of SnC identification methods (Methods) to five *in vitro* studies [30, 32–35]. In each study, scRNA-seq was performed on cell lines both before and after senescence induction. We first evaluated the performance of the continuous senescence scores generated by each method, using senescence induction as the gold standard. Figure 3b shows the receiver operating characteristic (ROC) curves for the top-performing methods in each dataset, and Figure 3c shows the area under the ROC curve (AUROC) as an evaluation metric. Both DeepScence and SenCID demonstrated the best overall performance, with AUROCs exceeding 0.9 across all datasets. DeepScence exhibited slightly better performance than SenCID in every dataset. For some SnGs, gene ranking-based methods performed well overall but failed in certain datasets with AUROCs below 0.8. Other SnGs showed poor performance using these methods. Single marker gene-based SnC identification methods also performed poorly.

For each method, we further classified cells into two discrete categories, senescent and non-senescent, by binarizing the continuous senescence scores (Methods). We evaluated the performance of these binarized groups using accuracy and F1 scores (Figure 3d). Once again, DeepScence and SenCID were the top-performing methods, with DeepScence slightly outperforming SenCID. The other methods showed substantially worse performance in both accuracy and F1 scores.

DeepScence outperforms existing methods in *in vivo* scRNA-seq datasets

We next evaluated all methods for identifying SnCs in *in vivo* scRNA-seq datasets. We collected four datasets in which scRNA-seq was performed on human and mouse tissues *in vivo* [36–40]. Confirmed by β-gal staining, these studies found that senescent cells were more enriched in certain contexts, such as fibrosis, for specific cell types. We again used AUROC as a metric to assess the performance of each method.

Figure 4a shows the overall performance of each method. Figure 4b shows the ROC plots for methods that performed best in the *in vitro* datasets. DeepScence is the only method that has a reliable overall performance, with a 0.91 averaged AUROC. The averaged AUROC of all other methods drop considerably, with the second best method reaching only 0.78. Moreover, methods other than DeepScence fail in at least one cell type with AUROC smaller than 0.6. SenCID, that has a performance comparable to DeepScence in *in vitro* datasets, perform poorly in *in vivo* datasets, with AUROC less than or equal to 0.2 in five cell types. The reason is that SenCID assigns substantially lower senescence scores to the cell population with enriched senescent cells, contradictory with experimental findings. In comparison, the distributions of senescence scores assigned by DeepScence agree with the enrichment of senescent cells in all cases (Figure 4c).

DeepScence outperforms existing methods in spatial transcriptomics datasets

We further applied DeepScence and SenCID, the two top-performing methods in *in vitro* datasets, to ST samples of mouse muscle generated by 10x Visium [37, 41]. The tissue slices were subjected to injury induced by cardiotoxin (CTX) or notexin, and ST data were collected at 2, 5, and 10 days post-injury. The injured regions were distinguishable from healthy regions by histology (Figure 5a), showing higher β-gal intensities, indicating that SnCs were more enriched in the injured regions. DeepScence assigned higher senescence scores to Visium spots in the injured regions compared to the healthy regions (Figure 5b,c), achieving high AUROCs (Figure 5f), which aligns with the experimental findings. In contrast, SenCID assigned higher senescence scores to Visium spots in the healthy regions (Figure 5d,e), resulting in much lower AUROCs (Figure 5f). This example further demonstrates DeepScence's superior performance in *in vivo* settings.

Finally, we applied both methods to synthetic 10x Xenium datasets. Unlike 10x Visium, which profiles nearly the entire transcriptome, 10x Xenium, similar to other ST technologies based on fluorescence in situ hybridization (FISH), can only profile hundreds of prespecified genes [42]. For evaluation purposes, we synthesized 10x Xenium datasets from *in vitro* scRNA-seq data by subsetting a number of genes comparable to actual 10x Xenium data, while adding a small number of senescence-associated genes to the default 10x gene panel (Methods).

DeepScence performs consistently well across different datasets and with varying numbers of genes in the gene panel, whereas SenCID's performance drops considerably (Figure 5g). These results demonstrate the generalizability of DeepScence to data types beyond scRNA-seq. Moreover, DeepScence requires the addition of only 10 senescence-associated genes to achieve strong performance, allowing for the exploration of other biological pathways and functions during gene panel design.

Conclusions

In this study, we developed DeepScence, a deep-learning framework for identifying senescent cells in single-cell and spatial transcriptomics data. DeepScence is built upon CoreScence, a new senescence gene set we compiled to resolve the large discrepancies across existing gene sets. We have demonstrated that DeepScence accurately identifies senescent cells in both *in vitro* and *in vivo* scRNA-seq datasets, as well as in ST data generated by 10x Visium and 10x Xenium. In contrast, existing methods perform poorly in some scenarios. DeepScence is highly flexible and, in principle, can be applied to other types of data, such as proteomics.

Methods

Construction and evaluation of CoreScence

Collection of published SnGs

We collected nine senescence gene sets from published studies. SenMayo was collected from Saul et al. [23], CSGene from Zhao et al. [25], SenSig from Cherry et al. [24], and SASP genes from De Cecco et al. [28]. The gene set for the inflammatory network in cellular senescence was collected from Freund et al. [9], and the transcriptome signature of cellular senescence was collected from Casella et al. [29]. CellAge and GenAge were obtained from the Human Ageing Genomic Resources (HAGR) [26, 27]. The SeneQuest gene set was downloaded from the SeneQuest database [2], and a gene was retained either if it was reported by at least 15 publications, or if it was reported by at least four publications with at least 70% agreement on whether the gene induces or inhibits senescence. This criterion ensures that both frequently reported genes and genes with clear directionality are included.

Collection of differentially expressed genes from bulk RNA-seq datasets

We collected information on genes with differential expression between senescent and non-senescent cells from two published bulk RNA-seq datasets. The differential gene data, provided by the original studies, were directly downloaded from the Gene Expression Omnibus (GEO) under accessions GSE130727 and GSE175533. The downloaded data include log fold changes, p-values, and p-values adjusted for multiple testing. The first study [29] (GSE130727) includes HEAC, HUVEC, IMR-90, and WI-38 cell lines, with senescence induced by doxorubicin treatment, ionized radiation, oncogene induction, and replicative senescence via the Hayflick limit. The second study [30] (GSE175533) includes the WI-38 cell line, where senescence was induced by Ras, replicative senescence via the Hayflick limit, and increased cell density.

DeepScence model

DeepScence utilizes a Zero-Inflated Negative Binomial (ZINB) autoencoder to learn meaningful low-dimensional representations of scRNA-seq datasets and to score each cell for senescence based on its gene expression profile. The details of the DeepScence model are described below.

Input data

DeepScence takes as input a properly filtered expression count matrix. This matrix is first denoised using DCA (version 0.3.1) [20] with default parameters. The denoised expression count matrix is then library size normalized and log-transformed using the `scanpy` (version 1.9.8) functions `sc.pp.normalize_total` and `sc.pp.log1p` with default parameters. The log-normalized expression matrix is then subsetted to retain only the genes in the CoreScence gene set. Finally, the subsetted expression matrix is scaled using the `scanpy` `sc.pp.scale` function, so that the gene expression values have zero mean and unit standard deviation for each gene across all cells.

Architecture

Let $\boldsymbol{X}$ denote the denoised and subsetted gene expression count matrix, and let $\boldsymbol{X'}$ denote the denoised, log-normalized, subsetted, and scaled expression matrix. DeepScence employs an autoencoder architecture similar to DCA [20]. Specifically, $\boldsymbol{X'}$ is encoded into a bottleneck layer consisting of two nodes, and then decoded into three output layers. Each output layer matches the size of the input $\boldsymbol{X'}$ and represents the estimated dropout, mean, and dispersion parameters of the ZINB distribution. Since the input $\boldsymbol{X'}$ is normalized by library size, the encoded 2-dimensional representation, as well as the estimated dropout, mean, and dispersion matrices, are not affected by library sizes. The model can be specified in the formulation below.

$$
\begin{aligned}
\boldsymbol{E} &= \mathrm{ReLU}(\boldsymbol{X'}W_{\boldsymbol{E}}) \\
\boldsymbol{B} &= \tanh(\boldsymbol{E}W_{\boldsymbol{B}}) \\
\boldsymbol{D} &= \mathrm{ReLU}(\boldsymbol{B}W_{\boldsymbol{D}}) \\
\boldsymbol{\Pi} &= \mathrm{sigmoid}(\boldsymbol{D}W_{\boldsymbol{\Pi}}) \\
\boldsymbol{M'} &= \exp(\boldsymbol{D}W_{\boldsymbol{M'}}) \\
\boldsymbol{\Theta} &= \exp(\boldsymbol{D}W_{\boldsymbol{\Theta}})
\end{aligned}
$$

Here, $\boldsymbol{E}$ represents the encoder layer with 32 neurons, $\boldsymbol{B}$ is the bottleneck layer with 2 neurons, and $\boldsymbol{D}$ is the decoder layer with 32 neurons. $\boldsymbol{\Pi}$, $\boldsymbol{M'}$, and $\boldsymbol{\Theta}$ are the output dropout, mean, and dispersion corresponding to the ZINB distribution. $W_{\boldsymbol{E}}$, $W_{\boldsymbol{B}}$, $W_{\boldsymbol{D}}$, $W_{\boldsymbol{\Pi}}$, $W_{\boldsymbol{M'}}$, and $W_{\boldsymbol{\Theta}}$ are the parameters to be estimated during the model fitting process.

Objective function

The objective function of the model above, which will be minimized during the model fitting process, is defined as:

$$\boldsymbol{L} = -\frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\log\left(\mathrm{Z}(x_{ij};\pi_{ij},M_{ij},\theta_{ij})\right) + \lambda\rho(\mathbf{h_1},\mathbf{h_2})^2$$

Here, Z represents the probability density function of the ZINB distribution. I and J represent the total number of CoreScence genes and cells, respectively. x_{ij}, π_{ij}, M_{ij}, and θ_{ij} represent the gene expression count, dropout, mean, and dispersion for ith gene and jth cell, respectively. $\mathbf{h}_1$ and $\mathbf{h}_2$ are two vectors representing the values of the two neurons in the bottleneck layer, before tanh activation. $\rho(\mathbf{h_1},\mathbf{h_2})$ represents the Pearson correlation coefficient between $\mathbf{h}_1$ and $\mathbf{h}_2$.

The objective function of DeepScence consists of two components. The first component is the negative log-likelihood of the ZINB distribution. Minimizing this term ensures that the autoencoder learns essential information from the input gene expression count matrix. Before calculating this term, note that the estimated mean matrix, $\boldsymbol{M'}$, is converted to $\boldsymbol{M}$ by multiplying it by the size factors. The second component is the Pearson correlation coefficient between two neurons in $\boldsymbol{H}$. Minimizing this term ensures that the two neurons in $\boldsymbol{H}$ are almost uncorrelated, allowing them to capture information related to senescence and unrelated information, respectively.

The objective function is a weighted average of the two components, controlled by a positive hyperparameter λ. A λ value that is too low may cause the two neurons in the bottleneck layer to become correlated, diluting the senescence-related information captured by each neuron. Conversely, a λ value that is too high may lead the model to overly focus on minimizing correlation, neglecting the ZINB component that learns from the data. We empirically observed the model's behavior under different λ values and found that $\lambda = 1$ consistently results in a low absolute Pearson correlation ($|\rho| < 0.1$), while still generating desirable outcomes. Therefore, we set the default value of λ to 1, although users can adjust this value if needed.

Model training

By default, DeepScence is trained for 300 epochs without mini-batching at a learning rate of 0.005 using the Adam optimizer[43]. During each epoch, 10% of the cells are used as a validation set, and the validation loss is recorded based on the ZINB loss for these cells. An early stopping mechanism with learning rate reduction is implemented: if the validation loss does not improve for 10 consecutive epochs, the learning rate is halved. If there is no improvement for 30 epochs, the entire training process is stopped.

Output

After model fitting, DeepScence outputs the values of the neuron responsible for capturing senescence information in the bottleneck layer as continuous senescence scores. Since the roles of the two neurons are undetermined during model fitting, DeepScence uses a *post hoc* approach to identify the neuron that captures senescence information. Specifically, for each of the two neurons, DeepScence calculates the absolute value of the Pearson correlation coefficient between the neuron's values and the expression of each gene in CoreScence, based on $\boldsymbol{X'}$. The neuron with the higher average absolute correlation across all CoreScence genes is identified as the one capturing senescence information.

Finally, to ensure that the output senescence score is positively correlated with the cells' senescence levels, DeepScence calculates the Pearson correlation between the output score and the expression of the CDKN1A gene, a canonical marker of senescence. If the correlation is negative, the sign of the output score is flipped.

Binarization of the continuous score

Cellular senescence, a continuous process by nature [44], is best characterized by a continuous score, such as that provided by DeepScence. However, DeepScence also offers the option to convert the continuous senescence score into a binary score, which can be useful for certain downstream analyses, such as comparing the proportion of SnCs across samples or cell types. The binary score is obtained through a two-step process.

In the first step, an uncertainty score is calculated for each cell by comparing its continuous senescence score to a null distribution. Specifically, for each cell in the expression matrix $\boldsymbol{X'}$, the gene expression values are randomly permuted across all CoreScence genes. The permuted gene expression values are then input into the DeepScence autoencoder, already fitted with the original data, and the value of the neuron that captures senescence information is recalculated. This process is repeated 50 times to generate 50 permuted scores. The cell's uncertainty score is defined as the proportion of times the permuted score exceeds the cell's original DeepScence score. A senescence score cutoff is determined as the smallest value such that less than 1% of cells with senescence scores above the cutoff have uncertainty scores greater than 0.5.

In the second step, a mixture of normal distributions is fitted to the continuous senescence scores of all cells using the `GaussianMixture` function from the Python `sklearn` package (version 1.3.2). The number of normal components is selected from 2 to 10 based on the lowest Akaike Information Criterion (AIC). Cells are assigned to the normal components according to posterior probabilities. Among the normal distributions whose 5% quantiles are greater than the senescence score

cutoff defined in the previous step, the distribution with the smallest mean is selected. Senescent cells are defined as those assigned to this distribution or any distribution with a mean larger than that of the selected distribution.

Competing methods

Single marker gene approach

For the scRNA-seq dataset, the gene expression count matrix was normalized by library size and then log-transformed using the functions `sc.pp.normalize_total` and `sc.pp.log1p` with default settings from the Python package `scanpy`. The log-normalized expression values of individual genes, such as CDKN1A and CDKN2A, were used as continuous senescence scores, with cells exhibiting higher expression levels being more likely to be senescent. For binarization, cells with positive gene expression values were classified as senescent, while cells with zero expression values were classified as non-senescent.

Gene ranking approach based on SnGs

For the scRNA-seq dataset, the log-normalized gene expression matrix was obtained as described above. Using each of the published SnGs as input, we then calculated continuous senescence scores with the `gsva` function and the ssGSEA method in the R package `GSVA` (version 1.52.0), as well as the `AUCell_run` function in the R package `AUCell` (version 1.26.0). Note that for SnGs with directional information, only genes that induce senescence were used as input.

`GSVA` and `AUCell` do not provide functions to binarize the continuous senescence scores. However, the original SenMayo study [23], which uses the ssGSEA method, treats the top 10% of cells with the highest senescence scores as senescent cells. Therefore, we only evaluated the binarized scores for the SenMayo gene set using the ssGSEA method. The top 10% of cells with the highest senescence scores were predicted as senescent cells, while the remaining cells were predicted as non-senescent.

SenCID

For each scRNA-seq dataset, the Python package SenCID (version 1.0.0) was used to obtain both continuous and binarized senescence scores with default settings, using the gene expression count matrix as input.

Evaluating SnC identification methods

Collection of in vitro scRNA-seq datasets

We collected *in vitro* scRNA-seq datasets from five studies that profiled both senescent and non-senescent cells.

In the first study [30], the gene expression count matrix for the WI-38 cell line was downloaded from GEO with accession GSE175533. The downloaded matrix had already been filtered, and no additional cell filtering was performed. Cells with the highest population doubling (PDL = 50) were considered senescent cells (SnCs), while negative controls (hTERT) were considered non-senescent.

In the second study [32], the gene expression count matrix for the HCA2 cell line was downloaded from GEO with accession GSE119807. Following the same cell filtering criteria as in [32], cells with more than 500 detected genes and a mitochondrial percentage of less than 10% were included. Cells labeled as "senescence" in the metadata provided by the original study were considered SnCs, while those labeled as 'LowPDCtrl' were considered non-senescent.

In the third study [34], the gene expression count matrix for the IMR-90 cell line was downloaded from GEO with accession GSE115301. Following the same cell filtering criteria as in [34], cells with less than 15% of reads mapped to mitochondrial genes and with positive expression in at least 2,500 genes were retained. Cells labeled as "RIS" (radiation-induced senescence) in the downloaded metadata were considered SnCs, while those labeled as "Growing" were considered non-senescent.

In the fourth study [35], the gene expression count matrix for the second IMR-90 cell line was downloaded from GEO with accession GSE94980. The downloaded matrix had already been filtered, and no additional cell filtering was performed. Cells infected with the reprogramming factors OSKM (indicated by "OSKM" in cell names) were considered SnCs, while the controls (indicated by "Vector" in cell names) were treated as non-senescent.

In the fifth study [33], the gene expression count matrix for the HUVEC cell line was downloaded from GEO with accession GSE102090. The downloaded matrix had already been filtered, and no additional cell filtering was performed. As detailed in the original publication, cells with names ending in the suffix "-2", indicating the second batch, were considered senescent SnCs, while cells from the first batch were considered non-senescent.

For datasets with more senescent cells than non-senescent cells, which is unlikely in *in vivo* settings, we randomly removed SnCs to equalize the numbers of SnCs and non-senescent cells. Finally, we retained genes with non-zero expression in at least 1% of the cells in each dataset.

Collection of in vivo scRNA-seq datasets

We collected *in vivo* scRNA-seq datasets from four studies that profiled both cells from diseased and normal conditions. In each study, β-gal staining was performed on certain cell types to confirm that SnCs were more enriched in cells from diseased conditions. We included only those cell types with β-gal staining information.

In the first study [36], the gene expression count matrix and cell type annotations for mouse testes cells were downloaded from GEO with accession GSE183625. We included only Leydig cells, in which β-gal staining was performed in the original study. Following the same cell filtering criteria as in [36], cells with less than 20% of reads mapped to mitochondrial genes and with positive expression in 200 to 7,000 genes were retained. Gene names were converted to their human homologs using the conversion table available on the Mouse Genome Informatics (MGI) website [45].

In the second study [37], the gene expression count matrix and cell type annotations for mouse muscle cells were downloaded from GEO with accession GSE214892. The matrix had already been filtered, and no additional cell filtering was performed. All cell types were included in this study. Gene names were converted to their human homologs using the conversion table available on the Mouse Genome Informatics (MGI) website [45].

In the third study [38], the gene expression count matrix and cell type annotations for human lung cells were downloaded from GEO with accession GSE190889. We included only AT1 cells, in which β-gal staining was performed in the original study. The matrix had already been filtered, and no additional cell filtering was performed.

In the fourth study [39], the gene expression count matrix for human oral mucosa cells was downloaded from GEO with accession GSE164241. Following the same cell filtering criteria as in [39], cells with less than 10% of reads mapped to mitochondrial genes, positive expression in 200 to 5,000 genes, and a total number of reads between 1,000 and 25,000 were retained. The original study did not provide cell type annotations. To perform cell type annotation, we first applied log normalization to the filtered expression count matrix using the Seurat function `NormalizeData`. We then ran PCA using the Seurat functions `FindVariableFeatures`, `ScaleData`, and `RunPCA` with default parameters. Cell clustering was performed using the Seurat function `FindNeighbors` with the top 15 PCs, followed by `FindClusters` with a resolution of 0.8. Cell type annotation was then performed using the R package GPTCelltype [46], with the tissue name set to 'human oral mucosa' and the model set to 'gpt-4'. We included only fibroblasts, in which β-gal staining was performed in the original study.

Finally, for all datasets, we retained genes with positive expression in at least 1% of the retained cells.

Collection of 10x Visium datasets

We downloaded 10x Visium ST datasets of mouse muscle tissues collected 2 and 5 days post-notexin-injury and 10 days post CTX-injury[41]. The gene expression count matrix, along with the annotation indicating whether a spatial spot was in the injured or healthy regions, was downloaded from Dryad (DOI: 10.5061/dryad.t4b8gtj34) and from GEO with accession number GSE125305, respectively. The matrix had already been filtered, so no additional cell filtering was performed. Gene names were converted to their human homologs using the conversion table available on the Mouse Genome Informatics (MGI) website [45].

Generation of simulated 10x Xenium data

We downloaded the 'Human Multi-Tissue and Cancer Panel' from the 10x pre-designed panels, which includes 377 human genes. Since the original panel contained very few senescence-associated genes, we expanded it by adding randomly selected senescence-associated genes. For DeepScence, the genes were randomly selected from CoreScence, and for SenCID, the genes were randomly selected from the list of genes used in SenCID's SVM model. For each cell line in the *in vitro* studies, we then subsetted the expression count matrix to include only the genes from the expanded panel.

Evaluation details

To evaluate the continuous senescence scores, ROC curves and AUROC were generated using the R package `pROC` (version 1.18.5).

To evaluate the binarized senescence scores, we used the functions `accuracy_score` and `f1_score` from the Python package `scikit-learn` (version 1.1.3). Let TP, TN, FP, and FN represent the true positives, true negatives, false positives, and false negatives, respectively. The accuracy is defined as:

$$\frac{TP+TN}{TP+TN+FP+FN}$$

The F1 score is defined as:

$$2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

where

$$\text{Precision} = \frac{TP}{TP+FP}, \quad \text{Recall} = \frac{TP}{TP+FN}$$

Acknowledgments

Y.Q., C.C., J.X., C.G., X.W., A.N., and Z.J. were funded by the National Institutes of Health under award number U54AG075936. Y.Q. and Z.J. were funded by the National Institutes of Health under award number R35GM154865. R.D. and L.G. were funded by the National Institutes of Health under award number UH3CA268096.

Author contributions

All authors contributed to the conceptualization of the study. Y.Q. developed the DeepScence software and performed the analysis. Y.Q. and Z.J. jointly wrote the manuscript.

Competing interests

A.N. receives funding from Genentech, Genmab, MedImmune/AstraZeneca, and Seattle Genetics. A.N. is a consult of Sanofi, Promega, and Leap Therapeutics.

Code availability

The DeepScence Python package is freely available on GitHub: https://github.com/anthony-qu/DeepScence. The analysis code for reproducing the results can also be accessed on GitHub: https://github.com/anthony-qu/DeepScence_analysis.

Data availability

All data analyzed in this study were obtained from previously published studies.

Information on differentially expressed genes from the two bulk RNA-seq studies was downloaded from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6698740/bin/gkz555_supplemental_files.zip and https://cdn.elifesciences.org/articles/70283/elife-70283-fig1-data2-v3.xlsx.

The scRNA-seq data with *in vitro* senescence inductions were downloaded from GEO. Specifically, the WI-38 dataset was downloaded from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE175533, the HCA2 dataset from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE119807, the two IMR-90 datasets from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE115301 and https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE94980, and the HUVEC dataset from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE102090.

The *in vivo* scRNA-seq data were downloaded from GEO. Specifically, the mouse testes EAO dataset was downloaded from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE183625, the mouse muscle injury data from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE214892, the human lung IPF data from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE190889, and the human oral mucosa data from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE164241.

The 10x Visium ST dataset of mouse muscle injury by notexin injection was downloaded from https://datadryad.org/stash/dataset/doi:10.5061/dryad.t4b8gtj34. The ST dataset of injury by CTX injection was downloaded from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE215305.

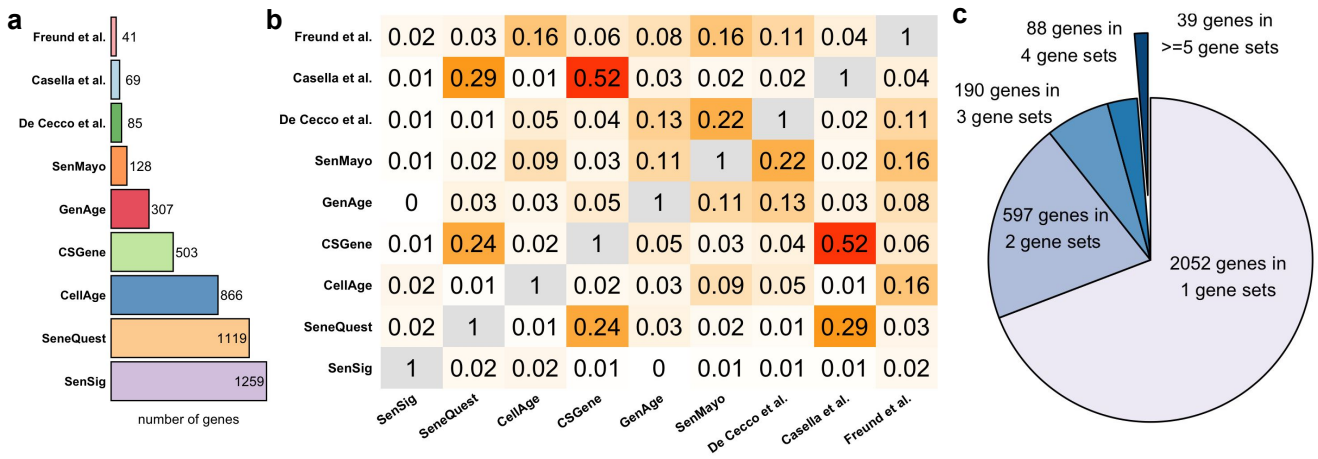

Figure 1. a, Number of genes in each SnG. **b,** Jaccard index between each pair of SnGs. **c,** Number of genes reported by different numbers of SnGs.

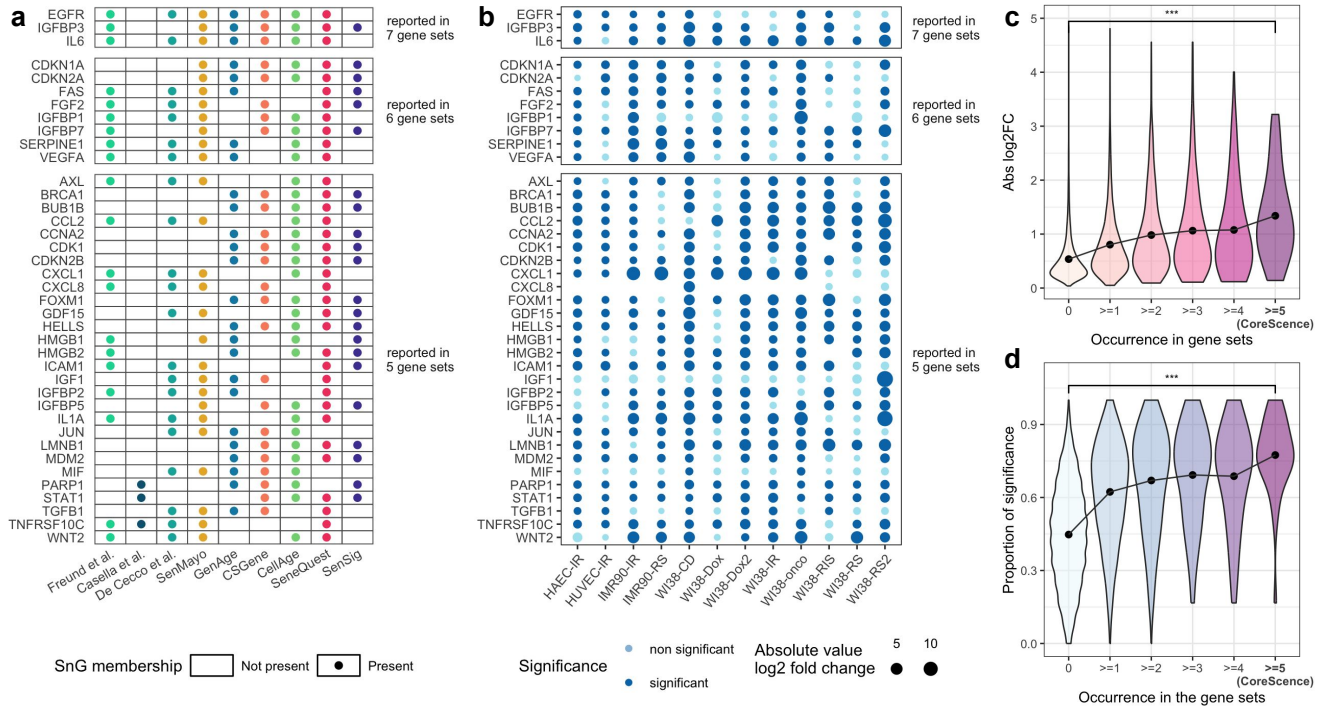

Figure 2. a, The CoreScence genes and their SnG memberships. **b,** Differential expression in bulk RNA-seq data for the CoreScence genes. Colors indicate statistical significance, and the sizes of the dots represent the absolute values of log2 fold change. Column names indicate the cell line and senescence induction methods. IR refers to ionizing radiation; RS refers to replicative senescence; CD refers to cell density; Dox refers to Doxorubicin; RIS refers to Ras-induced senescence; onco refers to oncogene-induced senescence. **c,** Averaged absolute values of log2 fold change across DEG comparisons (y-axis) and the number of SnGs reporting a gene (x-axis). Wilcoxon test was conducted to compare the two distributions. "***" indicates p-value < 0.001. Black dots represent the mean values. **d,** Proportion of DEG comparisons with statistically significant differential expression (y-axis) and the number of SnGs reporting a gene (x-axis). Wilcoxon test was conducted to compare the two distributions. "***" indicates p-value < 0.001. Black dots represent the mean values.

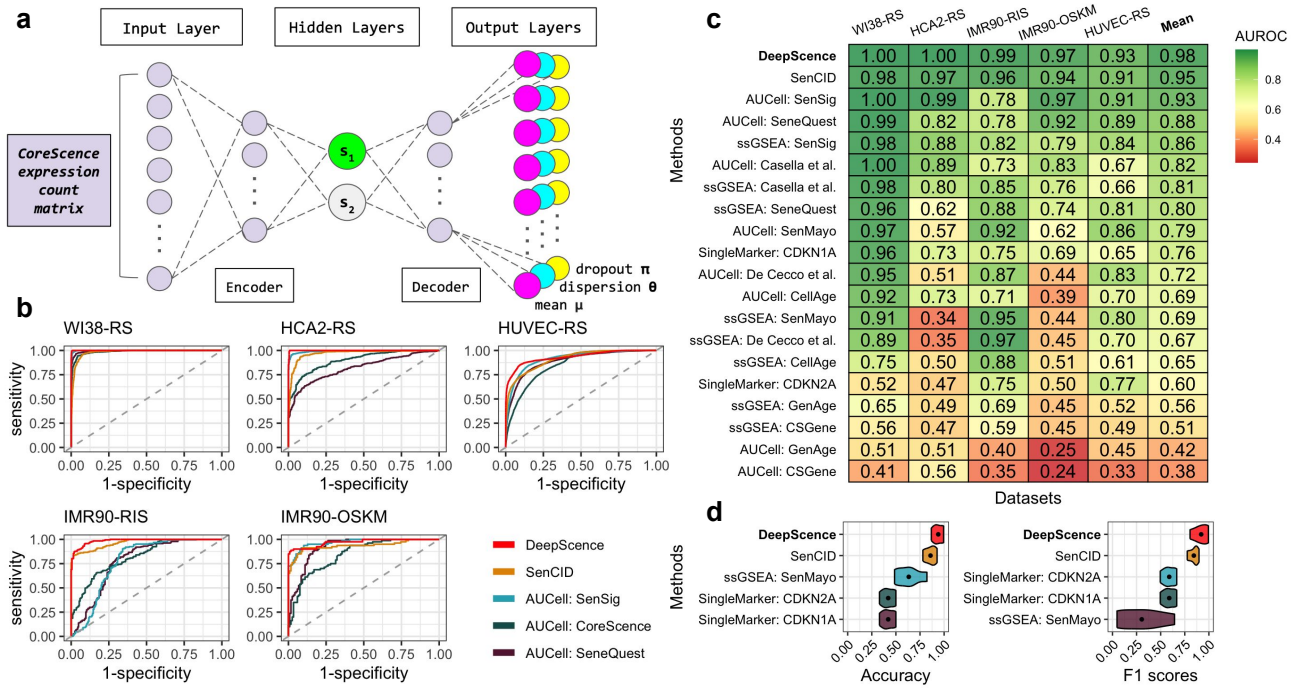

Figure 3. a, DeepScence model architecture. The green neuron, s1, in the bottleneck layer represents the neuron that captures senescence score. **b,** ROC curves of top-performing methods across *in vitro* datasets. **c,** AUROCs for all methods across *in vitro* datasets. Methods are ordered in decreasing order by average AUROCs. **d,** Accuracy and F1 scores across *in vitro* datasets, comparing DeepScence with existing methods capable of binarization. Dots in the violin plots represent mean values.

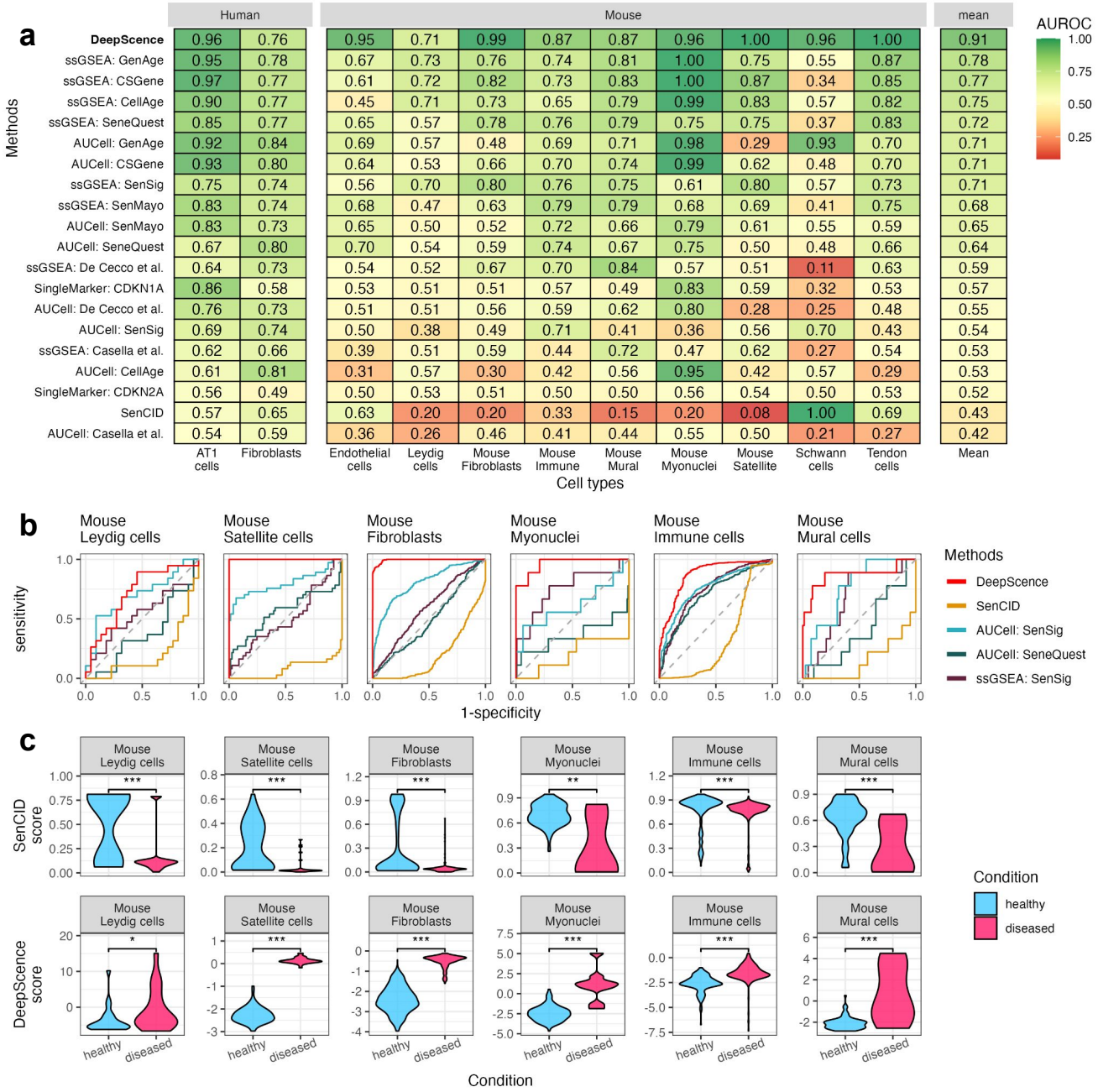

Figure 4. a, AUROCs for all methods across *in vivo* datasets. Methods are ordered in decreasing order by average AUROCs. **b,** ROC curves for six example cell types in *in vivo* datasets, showing methods with top performance in *in vitro* datasets. **c,** Distribution of SenCID and DeepScence scores, comparing cells from healthy and diseased conditions. Wilcoxon test was conducted to compare the two distributions in each case. '*' indicates p-value between 0.01 and 0.05, '**' indicates p-value between 0.001 and 0.01, and '***' indicates p-value < 0.001.

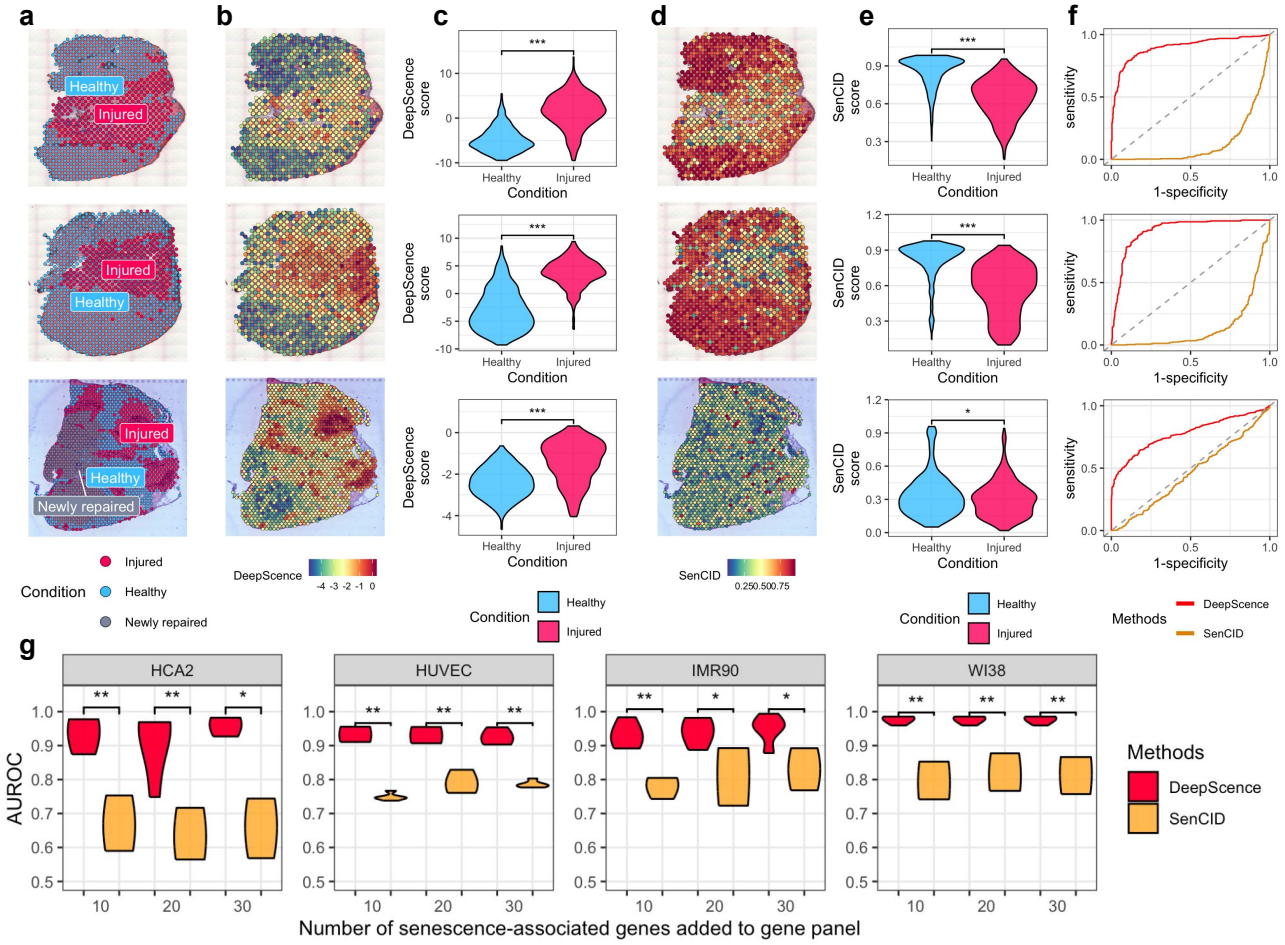

Figure 5. a, Spatial spots colored by healthy and injured regions in two Visium samples. **b,** DeepScence scores shown on spatial spots. **c,** Distribution of DeepScence scores, comparing spots from healthy and injured regions. Wilcoxon test was conducted to compare the two distributions in each sample. '***' indicates p-value < 0.001. **d,** SenCID scores shown on spatial spots. **e,** Distribution of SenCID scores, comparing spots from healthy and injured regions. Wilcoxon test was conducted to compare the two distributions in each sample. "*" indicates p-value between 0.05 and 0.01, and '***' indicates p-value < 0.001. **f,** ROC curves for DeepScence and SenCID. **g,** AUROC (y-axis) of DeepScence and SenCID in simulated Xenium datasets with different numbers of senescence-associated genes added to the default 10x gene panel (x-axis). Wilcoxon test was conducted to compare DeepScence and SenCID AUROCs. '*' indicates p-value between 0.01 and 0.05, '**' indicates p-value between 0.001 and 0.01, and '***' indicates p-value < 0.001.

References

1. Hernandez-Segura, A., Nehme, J. & Demaria, M. Hallmarks of Cellular Senescence. en. *Trends in Cell Biology* **28,** 436–453. ISSN: 09628924. https://linkinghub.elsevier.com/retrieve/pii/S0962892418300205 (2024) (June 2018).
2. Gorgoulis, V. *et al.* Cellular Senescence: Defining a Path Forward. English. *Cell* **179.** Publisher: Elsevier, 813–827. ISSN: 0092-8674, 1097-4172. https://www.cell.com/cell/abstract/S0092-8674(19)31121-3 (2024) (Oct. 2019).
3. López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The Hallmarks of Aging. English. *Cell* **153.** Publisher: Elsevier, 1194–1217. ISSN: 0092-8674, 1097-4172. https://www.cell.com/cell/abstract/S0092-8674(13)00645-4 (2024) (June 2013).
4. López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. Hallmarks of aging: An expanding universe. English. *Cell* **186.** Publisher: Elsevier, 243–278. ISSN: 0092-8674, 1097-4172. https://www.cell.com/cell/abstract/S0092-8674(22)01377-0 (2024) (Jan. 2023).
5. Coryell, P. R., Diekman, B. O. & Loeser, R. F. Mechanisms and therapeutic implications of cellular senescence in osteoarthritis. eng. *Nature Reviews. Rheumatology* **17,** 47–57. ISSN: 1759-4804 (Jan. 2021).
6. Martínez-Cué, C. & Rueda, N. Cellular Senescence in Neurodegenerative Diseases. eng. *Frontiers in Cellular Neuroscience* **14,** 16. ISSN: 1662-5102 (2020).
7. Saez-Atienzar, S. & Masliah, E. Cellular senescence and Alzheimer disease: the egg and the chicken scenario. en. *Nature Reviews Neuroscience* **21.** Publisher: Nature Publishing Group, 433–444. ISSN: 1471-0048. https://www.nature.com/articles/s41583-020-0325-z (2024) (Aug. 2020).
8. Schafer, M. J. *et al.* Cellular senescence mediates fibrotic pulmonary disease. en. *Nature Communications* **8.** Publisher: Nature Publishing Group, 14532. ISSN: 2041-1723. https://www.nature.com/articles/ncomms14532 (2024) (Feb. 2017).
9. Freund, A., Orjalo, A. V., Desprez, P.-Y. & Campisi, J. Inflammatory networks during cellular senescence: causes and consequences. English. *Trends in Molecular Medicine* **16.** Publisher: Elsevier, 238–246. ISSN: 1471-4914, 1471-499X. https://www.cell.com/trends/molecular-medicine/abstract/S1471-4914(10)00046-8 (2024) (May 2010).
10. Suryadevara, V. *et al.* SenNet recommendations for detecting senescent cells in different tissues. en. *Nature Reviews Molecular Cell Biology.* Publisher: Nature Publishing Group, 1–23. ISSN: 1471-0080. https://www.nature.com/articles/s41580-024-00738-8 (2024) (June 2024).
11. Kohli, J. *et al.* Algorithmic assessment of cellular senescence in experimental and clinical specimens. eng. *Nature Protocols* **16,** 2471–2498. ISSN: 1750-2799 (May 2021).
12. Williams, C. G., Lee, H. J., Asatsuma, T., Vento-Tormo, R. & Haque, A. An introduction to spatial transcriptomics for biomedical research. *Genome Medicine* **14,** 68. ISSN: 1756-994X. https://doi.org/10.1186/s13073-022-01075-1 (2024) (June 2022).
13. SenNet Consortium. NIH SenNet Consortium to map senescent cells throughout the human lifespan to understand physiological health. eng. *Nature Aging* **2,** 1090–1100. ISSN: 2662-8465 (Dec. 2022).
14. Gurkar, A. U. *et al.* Spatial mapping of cellular senescence: emerging challenges and opportunities. *Nature aging* **3.** Publisher: Nature Publishing Group US New York, 776–790. https://www.nature.com/articles/s43587-023-00446-6 (2024) (2023).
15. Hou, W., Ji, Z., Ji, H. & Hicks, S. C. A systematic evaluation of single-cell RNA-sequencing imputation methods. eng. *Genome Biology* **21,** 218. ISSN: 1474-760X (Aug. 2020).
16. Aibar, S. *et al.* SCENIC: Single-cell regulatory network inference and clustering. *Nature methods* **14,** 1083–1086. ISSN: 1548-7091. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5937676/ (2024) (Nov. 2017).
17. Yi, M., Nissley, D. V., McCormick, F. & Stephens, R. M. ssGSEA score-based Ras dependency indexes derived from gene expression data reveal potential Ras addiction mechanisms with possible clinical implications. en. *Scientific Reports* **10.** Publisher: Nature Publishing Group, 10258. ISSN: 2045-2322. https://www.nature.com/articles/s41598-020-66986-8 (2024) (June 2020).

18. Tao, W., Yu, Z. & Han, J.-D. J. Single-cell senescence identification reveals senescence heterogeneity, trajectory, and modulators. English. *Cell Metabolism* **36.** Publisher: Elsevier, 1126–1143.e5. ISSN: 1550-4131. https://www.cell.com/cell-metabolism/abstract/S1550-4131(24)00088-3 (2024) (May 2024).

19. Ogrodnik, M. Cellular aging beyond cellular senescence: Markers of senescence prior to cell cycle arrest in vitro and in vivo. eng. *Aging Cell* **20,** e13338. ISSN: 1474-9726 (Apr. 2021).

20. Eraslan, G., Simon, L. M., Mircea, M., Mueller, N. S. & Theis, F. J. Single-cell RNA-seq denoising using a deep count autoencoder. en. *Nature Communications* **10.** Publisher: Nature Publishing Group, 390. ISSN: 2041-1723. https://www.nature.com/articles/s41467-018-07931-2 (2024) (Jan. 2019).

21. Amodio, M. *et al.* Exploring single-cell data with deep multitasking neural networks. eng. *Nature Methods* **16,** 1139–1145. ISSN: 1548-7105 (Nov. 2019).

22. Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep generative modeling for single-cell transcriptomics. eng. *Nature Methods* **15,** 1053–1058. ISSN: 1548-7105 (Dec. 2018).

23. Saul, D. *et al.* A new gene set identifies senescent cells and predicts senescence-associated pathways across tissues. en. *Nature Communications* **13.** Publisher: Nature Publishing Group, 4827. ISSN: 2041-1723. https://www.nature.com/articles/s41467-022-32552-1 (2024) (Aug. 2022).

24. Cherry, C. *et al.* Transfer learning in a biomaterial fibrosis model identifies in vivo senescence heterogeneity and contributions to vascularization and matrix production across species and diverse pathologies. eng. *GeroScience* **45,** 2559–2587. ISSN: 2509-2723 (Aug. 2023).

25. 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. en. *Cell Death & Disease* **7.** Publisher: Nature Publishing Group, e2053–e2053. ISSN: 2041-4889. https://www.nature.com/articles/cddis2015414 (2024) (Jan. 2016).

26. Tacutu, R. *et al.* Human Ageing Genomic Resources: new and updated databases. eng. *Nucleic Acids Research* **46,** D1083–D1090. ISSN: 1362-4962 (Jan. 2018).

27. De Magalhães, J. P., Curado, J. & Church, G. M. Meta-analysis of age-related gene expression profiles identifies common signatures of aging. eng. *Bioinformatics (Oxford, England)* **25,** 875–881. ISSN: 1367-4811 (Apr. 2009).

28. De Cecco, M. *et al.* L1 drives IFN in senescent cells and promotes age-associated inflammation. eng. *Nature* **566,** 73–78. ISSN: 1476-4687 (Feb. 2019).

29. Casella, G. *et al.* Transcriptome signature of cellular senescence. eng. *Nucleic Acids Research* **47,** 7294–7305. ISSN: 1362-4962 (Aug. 2019).

30. M, C. *et al.* Novel insights from a multiomics dissection of the Hayflick limit. en. *eLife* **11.** Publisher: Elife. ISSN: 2050-084X. https://pubmed.ncbi.nlm.nih.gov/35119359/ (2024) (Feb. 2022).

31. Risso, D., Perraudeau, F., Gribkova, S., Dudoit, S. & Vert, J.-P. A general and flexible method for signal extraction from single-cell RNA-seq data. en. *Nature Communications* **9.** Publisher: Nature Publishing Group, 284. ISSN: 2041-1723. https://www.nature.com/articles/s41467-017-02554-5 (2024) (Jan. 2018).

32. Tang, H. *et al.* Single senescent cell sequencing reveals heterogeneity in senescent cells induced by telomere erosion. eng. *Protein & Cell* **10,** 370–375. ISSN: 1674-8018 (May 2019).

33. A, Z. *et al.* HMGB2 Loss upon Senescence Entry Disrupts Genomic Organization and Induces CTCF Clustering across Cell Types. en. *Molecular cell* **70.** Publisher: Mol Cell. ISSN: 1097-4164. https://pubmed.ncbi.nlm.nih.gov/29706538/ (2024) (May 2018).

34. Teo, Y. V. *et al.* Notch Signaling Mediates Secondary Senescence. eng. *Cell Reports* **27,** 997–1007.e5. ISSN: 2211-1247 (Apr. 2019).

35. Aarts, M. *et al.* Coupling shRNA screens with single-cell RNA-seq identifies a dual role for mTOR in reprogramming-induced senescence. eng. *Genes & Development* **31,** 2085–2098. ISSN: 1549-5477 (Oct. 2017).

36. Li, Y. *et al.* High Throughput scRNA-Seq Provides Insights Into Leydig Cell Senescence Induced by Experimental Autoimmune Orchitis: A Prominent Role of Interstitial Fibrosis and Complement Activation. eng. *Frontiers in Immunology* **12,** 771373. ISSN: 1664-3224 (2021).

37. Young, L. V. *et al.* Muscle injury induces a transient senescence-like state that is required for myofiber growth during muscle regeneration. eng. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology* **36,** e22587. ISSN: 1530-6860 (Nov. 2022).

38. Sui, J. *et al.* Loss of ANT1 Increases Fibrosis and Epithelial Cell Senescence in Idiopathic Pulmonary Fibrosis. *American Journal of Respiratory Cell and Molecular Biology* **69.** Publisher: American Thoracic Society - AJRCMB, 556–569. ISSN: 1044-1549. https://www.atsjournals.org/doi/abs/10.1165/rcmb.2022-0315OC (2024) (Nov. 2023).

39. Williams, D. W. *et al.* Human oral mucosa cell atlas reveals a stromal-neutrophil axis regulating tissue immunity. eng. *Cell* **184,** 4090–4104.e15. ISSN: 1097-4172 (July 2021).

40. Guo, S. *et al.* ROS-Induced Gingival Fibroblast Senescence: Implications in Exacerbating Inflammatory Responses in Periodontal Disease. eng. *Inflammation.* ISSN: 1573-2576 (Apr. 2024).

41. McKellar, D. W. *et al.* Large-scale integration of single-cell transcriptomic data captures transitional progenitor states in mouse skeletal muscle regeneration. eng. *Communications Biology* **4,** 1280. ISSN: 2399-3642 (Nov. 2021).

42. Janesick, A. *et al.* High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. en. *Nature Communications* **14.** Publisher: Nature Publishing Group, 8353. ISSN: 2041-1723. https://www.nature.com/articles/s41467-023-43458-x (2024) (Dec. 2023).

43. Kingma, D. P. & Ba, J. *Adam: A Method for Stochastic Optimization* arXiv:1412.6980. Jan. 2017. http://arxiv.org/abs/1412.6980 (2024).

44. Ashraf, H. M., Fernandez, B. & Spencer, S. L. The intensities of canonical senescence biomarkers integrate the duration of cell-cycle withdrawal. en. *Nature Communications* **14.** Publisher: Nature Publishing Group, 4527. ISSN: 2041-1723. https://www.nature.com/articles/s41467-023-40132-0 (2024) (July 2023).

45. Baldarelli, R. M. *et al.* Mouse Genome Informatics: an integrated knowledgebase system for the laboratory mouse. eng. *Genetics* **227,** iyae031. ISSN: 1943-2631 (May 2024).

46. Hou, W. & Ji, Z. *GPT-4V exhibits human-like performance in biomedical image classification* en. Jan. 2024. http://biorxiv.org/lookup/doi/10.1101/2023.12.31.573796 (2024).