Supplementary Information: Murine single cell RNA-seq reveals cell-identity- and tissue-specific trajectories of aging

Jacob C. Kimmel
Calico Life Sciences
South San Francisco, CA

Lolita Penland [*]
Calico Life Sciences
South San Francisco, CA

Nimrod D. Rubinstein [*]
Calico Life Sciences
South San Francisco, CA

David G. Hendrickson
Calico Life Sciences
South San Francisco, CA

David R. Kelley
Calico Life Sciences
South San Francisco, CA
`drk@calicolabs.com`

Adam Z. Rosenthal [†]
Calico Life Sciences
South San Francisco, CA
`adam.rosenthal@dupont.com`

Corresponding Author:

David R. Kelley
`drk@calicolabs.com`
1170 Veterans Blvd
Calico Life Sciences
South San Francisco, CA, USA, 94080

Adam Z. Rosenthal
`adam.rosenthal@dupont.com`
DuPont Nutrition and Biosciences
200 Powder Mill Rd
Wilmington, DE, 19803

---

[*] Equal Contribution

[†] Current Address: DuPont Nutrition and Biosciences, 200 Powder Mill Rd, Wilmington, DE, 19803

Supplementary Methods

Cell Isolation

We removed each tissue and washed each in HBSS then dissected the tissue into small pieces using a razor blade.

Kidney and Lung

We incubated dissected kidney tissue in 1.48 U/mL Liberase DL enzyme mixture (Roche Custom Biotech, Indianapolis, IN, USA) in a total volume of 7 mL DMEM at 37$^{\mathrm{O}}$C for 20 minutes in a 50 mL conical tube shaking at 200 rpm. We incubated dissected lung tissue in 1.48 U/mL Liberase TM (Roche) and 200 U/mL DNase I (Roche) in a total volume of 7 mL DMEM at 37$^{\mathrm{O}}$C for 30 minutes in a 50 mL conical tube shaking at 200 rpm. Tissue was further mixed using a 10 mL pipette tip and 40 mL of DMEM (2% FBS) was added to stop digestion. We sequentially pipetted cells through 100 μm, 70 μm, and 40 μm filters then pelleted cell suspensions by centrifugation at 500 x g for 10 minutes.

After centrifugation, we treated cells with ACK Lysing Buffer (Thermo Fischer, Waltham, MA, USA) for 5 minutes at room temperature. We centrifuged ACK treated cells, discarded the supernatant, and repeated ACK treatment. We subsequently washed cells in DMEM (2% FBS, GIBCO brand from Thermo Fischer) and pelleted by centrifugation. We removed debris from samples using the Miltenyi Debris Removal Solution (Miltenyi Biotec, Bergisch Gladbach, Germany). We added 6.2 mL cold PBS and 1.8 mL Debris Removal Solution to cell pellettes and resuspended by pipetting. We added 4 mL cold PBS to the top of the mixed solution and centrifuged samples at 4$^{\mathrm{O}}$C , 3000 x g for 10 minutes. We washed samples in 10 mL cold PBS and centrifuged at 4$^{\mathrm{O}}$C , 1000 x g for 10 minutes, then resuspended in PBS with 2% FBS. We counted cells using a TC20 Cell Counter (Bio-Rad, Hercules, CA, USA) and diluted cells to a concentration of 10^6 cells/mL. We then proceeded to single cell library preparation.

Spleen

We used a syringe plunger to further fragment the spleen tissue after razor blade dissection. We mechanically dissociated tissue fragments by pipetting in RPMI cell culture media (2% FBS). We subsequently forced tissue fragments through 100 and 40μm filters. After centrifugation at 500 x g for 10 minutes, we treated cells with ACK Lysing Buffer (Thermo Fischer) for 5 minutes at room temperature. We washed cells in RPMI (2% FBS) and pelleted by centrifugation. We counted cells using a TC20 Cell Counter (Bio-Rad) and diluted cells to a concentration of 10^6 cells/mL then proceeded to single cell library preparation.

Read Alignment and Gene Expression Quantification

We aligned reads to the `mm10` reference genome obtained from ENSEMBL. We used gene annotations from the GENCODE vM20 release with slight modification. We have replaced annotations for lincRNAs *Gm42418* and *AY036118* with a single contiguous gene annotation for the rRNA element *Rn45s*. This locus harbors an *Rn45s* repeat as reflected in RefSeq, such that contaminating 18S rRNA in our library preparations may lead to inflated expression counts for these lincRNAs.

We performed alignment to this amended reference using 10x `cellranger 3.0.2`, which employs the `STAR` sequence aligner [Dobin et al., 2013]. We determined gene expression counts using unique molecular identifiers (UMIs) for each cell barcode-gene combination. Following alignment, we filtered cell barcodes to identify those which contain cells using the approach implemented in `cellranger 3.0.2`, and only these barcodes were considered for downstream analysis. The output of this analysis is a Cells $\times$ Genes matrix, where each element i, j represents the number of UMIs mapping to gene j in cell i.

Dimensionality Reduction

To enable unsupervised clustering and cell type identification, we perform dimensionality reduction with principal component analysis (PCA) on the combined set of samples for each tissue. First, we identify a set of highly variable genes within the tissue based on overdispersion as described previously [Satija et al., 2015]. Briefly, we computed a "normalized dispersion" score for each gene by binning genes with similar mean expression levels, and subtracting the mean dispersion (variance/mean) within a bin from the dispersion score for each gene in the bin. The residual values represent the dispersion after accounting for the mean:variance relationship using this binning scheme. We set the following minimum criteria for the selection of genes as highly variable: minimum overdispersion of 0.5, minimum mean expression of $0.5 \ln(\text{CPM} + 1)$, maximum mean expression of $7 \ln(\text{CPM} + 1)$. After subsetting to this set of highly variable genes, we centered expression values to a mean of 0 and unit variance. We used these scaled genes as input features for PCA.

Cell Type Classification

Louvain community detection [Satija et al., 2015, Blondel et al., 2008] was applied to the nearest neighbor graph constructed in PCA space to define a cluster partition. To infer cell types, we trained a neural network classifier to predict cell ontology classes given single cell RNA-seq mRNA abundance profiles.

Our classifier was a fully connected neural network with four hidden layers each paired to a rectified linear unit activation, dropout layer, and batch normalization. Each hidden layer used 1024 units and the dropout probability on each layer was set to $p = 0.3$. We applied a softmax activation to the final layer and used cross entropy as an objective function for training. During training, we performed class balancing with a mixture of over- and under-sampling. Classes (cell types) with fewer than 128 examples were oversampled, while classes with more examples were undersampled. Optimization was performed using the Adagrad optimizer [Duchi et al., 2011].

As a training set, we used the *Tabula Muris* compendium which provides expert cell type annotations in the mouse [The Tabula Muris Consortium, 2018]. In addition to these annotations, we manually added cell state annotations to the Tabula Muris data to provide a level of granularity below cell ontology classes (Supp. Fig. S1). Example cell states include categorizing T cells into CD4 T cell and CD8 T cell subgroups, as well as the addition of subgroup labels to heterogeneous cell types such as lung stromal cells in the *Tabula Muris*. We named these cell states which do not have canonical names based on the expression of a prominent marker gene. For instance, we refer to a *Gucy1a1*+ subset of lung stromal cells as *Gucy1a1* lung stromal cells.

We first trained a classifier for each tissue individually, yielding a set of tissue-specific classifiers that have no knowledge of cell types outside the training tissue. We also trained a tissue independent classifier by training on all cell types present in the Tabula Muris simultaneously.

We first inferred cell types for the lung and spleen using a tissue-specific classification model trained to predict cell ontology classes. We classified kidney cells with a tissue independent model to provide further resolution of immune cell types not annotated in kidney cells of the training set. We subsequently inferred subtypes using a model trained to predict our added subtype annotations. We chose subtypes as the most likely subtype allowed within our defined hierarchy. For instance, we chose the subtype of a T cell as the most likely subtype among the "CD4", "CD8", and "memory" subtypes. We used subtype classification models trained on the spleen to predict T cell and macrophage subtypes in all tissues, as the training set contained insufficient cell numbers to perform subtype annotation in other tissues. After direct inference of types and subtypes, we refined cell type information by using a k-nearest neighbors smoothing approach. Here, we chose $k \in [30, 100]$ empirically depending on the tissue context.

We also classified cell types using the `scmap-cell` approach as an orthogonal cell type identification method [Kiselev et al., 2018]. We trained a k-nearest neighbors classification model using the cosine distance metric for kidney, lung, and spleen data in *Tabula Muris*. We selected the top 1000 genes with high residual dropout rate for training using the linear modeling approach of M3Drop [Kiselev et al., 2018]. As in the `scmap-cell` paper, we used only genes that are detected in both the *Tabula Muris* and the corresponding tissue of our data. We set $k = 10$ as in the `scmap-cell` paper. We assigned a cell type for all cells rather than using an "unassigned" label for low confidence predictions, as described in the original `scmap-cell` reference. This allowed us to interpret the classifier's "best guess," even when the classifications were uncertain, for better comparison with our method. We classified cells in our data using the relevant tissue classifier for each tissue. We also trained a classifier across all three tissues in the *Tabula Muris* and predicted cell types in the kidney data in an attempt to recall the B cell and T cell types not annotated in *Tabula Muris* kidney training data set. Matching the procedure for our neural network method, we used an `scmap-cell` classifier trained on spleen T cell and macrophage subtypes to predict T cell and macrophage subtypes in the kidney and lung. Overall, we found that `scmap-cell` classifications broadly agreed with our neural network predictions and subsequent manual validations, but failed to recover correct annotations for several cell types (Supp. Fig. S6).

Differential Variability Analysis

We measured differences in transcriptional variation between young and old animals in two distinct ways.

The first method evaluates changes in the variability of each gene between young and old animals, and attempts to identify a shift in the distribution of gene-wise variation. We assessed gene-specific variability by measuring the "overdispersion" of each gene. We defined overdispersion as the residual between a gene's observed dispersion and the expected dispersion based on the gene's mean expression value. We computed overdispersion values using the "difference from the median" (DM) method, as introduced previously [Kowalczyk et al., 2015]. We restricted DM calculation to genes with a mean expression value greater than 35 CPM to reduce the influence of poorly measured genes. To determine if the distribution of overdispersion values is significantly changed across ages, we employed the Wilcoxon Rank Sums test for a difference in means. We controlled the False Discovery Rate (FDR) to $\alpha = 0.05$

with the Benjamini-Hochberg procedure. We performed DM analysis for each cell state in each tissue in our data set separately.

The second method we employed evaluates cell-cell heterogeneity based on the Euclidean distance between cells in expression space, as introduced previously [Enge et al., 2017]. For each cell state in each tissue, we computed the centroid of the cell state in gene expression space. We computed the Euclidean distance from each cell to this centroid as a metric of cell-cell heterogeneity within each cell state. We employed the Wilcoxon Rank Sums test with the Benjamini-Hochberg procedure [Benjamini and Hochberg, 1995] as before to determine if this cell-cell heterogeneity metric is significantly different across ages.

As in previous reports [Enge et al., 2017, Angelidis et al., 2019], we also computed this distance to the centroid using a reduced set of genes as an alternative metric. These genes were chosen heuristically following previous methods [Enge et al., 2017]. We binned genes into 10 groups based on mean expression values, then selected genes within the lowest 10%-tile of coefficient of variation values.

As a second alternative, we computed the distance to the centroid using all measured genes but use the Manhattan distance rather than the Euclidean distance as a metric. The Manhattan distance implicitly models change in gene expression as independent across genes, in contrast to the Euclidean distance which implicitly models coordinated changes in gene expression. We found that results for both of these alternative cell-cell heterogeneity metrics are highly correlated with the primary analysis we present using the full set of measured genes and the Euclidean distance metric (Supp. Fig. S12).

Cell Cycle Scoring

Cell cycle activity was estimated by scoring the expression of a set of S phase associated and G_2/M phase associated genes, as shown previously [Tirosh et al., 2016] and as implement in `Seurat` [Satija et al., 2015]. Briefly, a set of genes associated with each of these phases was derived from single cell RNA-sequencing in 293T and 3T3 cell lines. For each cell, the sum of the expression of these phase-associated genes were computed. As a null distribution, a set of genes were selected from the set difference of the observed genes and the phase-associated genes. These null genes were selected by binning the sample genes into 24 equally sized bins by mean expression, then randomly sampling 100 null genes per bin per phase-associated gene with replacement. The difference between the mean of the phase-associated genes and the selected null genes was considered the module score.

For our analysis of cell cycle variation with scLVM [Buettner et al., 2015], we fit scLVM Gaussian process latent variable models with $k = 1$ latent variable to explain variation in genes contained in the cell cycle GO term (GO:0007049), as in the scLVM documentation. For spleen cells, we fit an scLVM model a randomly chosen subset of 10,000 cells to allow for convergence in a reasonable time frame.

Analysis of Variance in Transcriptional Space

To determine the proportion of variance in transcriptional space (gene-wise UMI counts, NMF embedding dimensions) explained by experimental factors in our data, we used the linear modeling approach of Robinson *et. al.* [Robinson et al., 2015]. Briefly, we fit a linear model for each dimension of the relevant transcriptional space (i.e. gene, NMF dimension) of the form

$$Y = \beta X^T + E$$

where Y is a Feature × Cell matrix of observed transcriptional features, X is a Samples × Parameters design matrix containing p experimental parameters, β is a Feature × Parameters coefficients matrix, and E is a Feature × Cell matrix of residuals. We calculated the proportion of variance explained by each experimental parameter for each gene by ANOVA. To determine the total proportion of variance explained by a factor, we summed the sum of squares for each parameter p across features for each parameter and divided by the summed total sum of squares:

$$P_{\text{var}}^{(p)} = \sum_{f}^{\mathcal{F}} \frac{\text{SS}_f^{(p)}}{\text{TSS}_f}$$

where f is a feature in the set of features $\mathcal{F}$.

NMF Embedding Interpretation

To assign semantic meaning to each dimension of the embedding, we first identified genes associated with each dimension by performing Otsu thresholding [Otsu, 1975] on log-transformed gene loadings. We used a scale factor of 2 on the Otsu threshold to generate more stringent gene associations. From the set of associated genes (loadings above the threshold), we performed Gene Ontology enrichment analysis with the Biological Process Gene Ontology database. We empirically named a consensus "program" for each dimension based on the enriched gene sets.

Aging Trajectory Permutation Testing and Bootstrap Sampling

We evaluated the statistical significance and robustness of aging trajectory results using permutation tests and bootstrap resampling. We performed permutation testing by randomizing tissue::cell type labels across all cells, followed by aging trajectory computation as above. We computed the adjusted rand index for lymphocyte clustering (B cell, T cell, NK cell classes) for each permuted set of labels compared against a common label where all lymphocytes share a cluster (as in Fig. 5A). For each permuted label, we also computed the proportion of variance explained by cell identity, tissue environment, and their interaction, as well as the distribution of cosine similarities. We compared the distribution of cosine similarities between permuted samples and our observation using the Kolmogorov–Smirnov (KS) test.

We performed bootstrap resampling across cells by randomly selecting 80% of the cells from each tissue::cell type combination for each of 100 iterations. At each iteration, we computed the cosine similarities between each tissue::cell type as described above. We also fit linear models to explain variation in these bootstrapped aging trajectories as described for the full data set above. These models estimate the proportion of variance explained by cell type, tissue environment, their interaction, and a residual amount of variation due to random sampling.

We also performed bootstrap resampling across animals using a "leave-one-out" strategy. For each iteration, we excluded one young and one old animal, for all possible combinations. We computed cosine similarities between tissue:cell type aging trajectories for each iteration as described above.

Supplementary Notes

Supplementary Note 1 – Tissue selection

We isolated single cells from the kidney, lung, and spleen. Each of these tissues exhibits age-related changes with functional implications for organismal health. For example, kidneys experience decreased glomerular filtration rate and elevated risk of kidney disease with age [Bolignano et al., 2014]. Lungs experience a decline in regenerative potential and a rise in pulmonary disorders [Navarro and Driscoll, 2017, Paxson et al., 2011, Bailey et al., 2014]. The spleen experiences structural changes with age with consequences for immune function [Turner and Mabbott, 2017, Aw et al., 2016]. The cellular etiology of these changes is not fully understood, motivating single cell analysis.

Supplementary Note 2 – Cell type and state proportion changes with age

One prospective way in which aging may influence tissue function is by altering the proportion of each cellular identity within the tissue. Lymphocytes were significantly more abundant in the kidneys and lungs of old animals (t-test on additive log-ratio (ALR) transformed proportions, $q < 0.05$, Fig. 1D). In old kidneys, we found a roughly 2 fold increase in T cells, classical monocytes, and non-classical monocytes. In old lungs, we found a corresponding 2 fold increase in classical and non-classical monocytes and a roughly 1.3 fold increase in T cells. This may reflect increasing immune infiltration of the the non-lymphoid tissues with age, as suggested in previous studies of kidney, lung, and other non-lymphoid tissues [O'Brown et al., 2015, Rodwell et al., 2004, Toapanta and Ross, 2009, Aoshiba and Nagai, 2007, Lumeng et al., 2011]. Both the kidney and lung experience elevated inflammation with age, and increased immune cell infiltration may be a factor [O'Sullivan et al., 2017, Kovacs et al., 2017, Kovacs et al., 2013].

However, we cannot rule out that recovery of specific cell types was confounded by an interaction of aging with our isolation procedures. The single cell isolation procedure required by scRNA-seq may lead to a biased recovery of some cell types relative to others, and this bias may change with age. Age-related changes in a tissue (i.e. altered extracellular matrix composition [Angelidis et al., 2019]) may lead to preferential isolation of one cell type relative to another, even if the underlying cell type proportions do not change. Therefore, this result raises the hypothesis that old kidneys and lungs have more immune cells than young counterparts, but *in situ* cell type quantification (e.g. by FISH or immunohistochemistry) is needed to confirm it.

For the spleen and only the non-immune cells in the kidney and lung, differences in cell type proportions between young and old animals were not statistically significant ($|\log_2 \frac{\text{Old}}{\text{Young}}| < 1$, $q > 0.05$, t-test of ALR-transformed proportions, Fig. 1D, Supp. Fig. S9B). However, we cannot rule out that even small, non-statistically significant changes may influence the aging process.

Shifting cell state proportions within a cell type may be an alternative mechanism by which aging phenotypes manifest. Investigating spleen T cells, the proportion of CD8 T cells decreased relative to other T cell subsets in older animals [Quinn et al., 2018, Quinn et al., 2016]. Circulating T cells may influence the spleen T cell compartment to affect this change. We also observed a decreased frequency of *Cald1*+ collecting duct cells relative to *Slc12a3*+ collecting duct cells in old kidneys (χ^2 contingency table, $q < 0.05$). These *Slc12a3*+ cells are likely principle cells of the collecting duct based on expression of *Scnn1a*. By contrast, *Cald1*+ collecting duct cells also express *Phgdh* and *Cryab*, which have been reported to mark a distinct collecting duct cell subpopulation [Takasaki et al., 2007, Han et al., 2018] (Supp. Fig. S9C). We are unaware of a defined role for this second cell population, making it difficult to speculate on the impact of this shift in cell state proportions. Other cell types with notable cell state substructure do not shift significantly with age, such as lung stromal cells (Supp. Fig. S9A).

Supplementary Note 3 – Cycling cells are similarly rare in young and old animals

Previous reports highlight cell cycle activity changes with age in multiple cell populations. In blood progenitors, cell cycle kinetics accelerate with age [Kowalczyk et al., 2015]. The frequency of cycling cells increases with age in muscle progenitors [Chakkalakal et al., 2012], but decreases in the intestinal crypt [Nalapareddy et al., 2017]. To investigate changes in cell cycle frequency in our data, we evaluated cell cycle activity by scoring the expression of S-phase associated genes and G_2M-associated genes [Tirosh et al., 2016] (Methods). We observed only small changes in either of these cell cycle module scores with age across cell identities, with only a handful of statistically significant differences (Supp. Fig. S10A). We did not find discrete populations of cells based on these cell cycle scores, and the standard cell cycle assignment scheme [Tirosh et al., 2016] assigned different phases to cells with similar scores (Supp. Fig. S10B). Cell cycle scores explained little variance in the data ($< 4\%$ in any cell identity, Supp. Fig. S10C). We also applied an alternative cell cycle scoring method, scLVM [Buettner et al., 2015]. However, we found that the scLVM model captures differences in the baseline expression of cell cycle-associated genes between cell types rather than variation within a cell type, as previously described [Barron and Li, 2016]. Thus, this method also suggests minimal cell cycle variation within cell types (Supp. Fig. S10D). Together, these results suggest that cell cycle rates change minimally with age, but we cannot eliminate the possibility that the cell cycle scoring methods are currently insufficient to detect differences.

Supplementary Note 4 – Senescent cells

Several previous studies reported accumulation of non-cycling senescent cells in aging tissues [Childs et al., 2017]. The reported magnitude of senescent cell accumulation varies between tissues. In aging mouse kidney, the proportion of cells with senescence-associated β-galactosidase activity increased from roughly 0.2% to 1.2%, whereas in epicardial cells the proportion increases from 2% to 10% (12 months to 18 months) [Baker et al., 2016]. Similar observations have been made using scRNA-seq in the human pancreas, where the proportion of cells expressing senescence marker gene *CDKN2A* increases from roughly 7% to 15+% between early- (21-22 years) and mid-adulthood (38-54 years) [Enge et al., 2017]. To investigate whether senescent cells are more prevalent in our old tissues, we similarly measured expression of *Cdkn2a*. The *Cdkn2a* locus (p16-Ink4a and p19-Arf) is not significantly upregulated with age in any of the cell identities (Supp. Fig. S11A). Due to the overlapping nature of the p16-Ink4a and p19-Arf reading frames, we note that we cannot distinguish transcripts from these two proteins using $3'$-end RNA-seq alone [Kamijo et al., 1997]. We also scored the activity of a curated set of senescence-associated genes using the AUCell approach [Aibar et al., 2017], but we did not find large differences in this score with age (median $|\log_2 \frac{\text{Old}}{\text{Young}}| = 0.061$, Supp. Fig. S11B). Only kidney activated macrophages and kidney non-classical monocytes showed a statistically significant increase with age ($q < 0.05$, $\log_2$ fold-change of 0.22 and 0.08 respectively).

Supplementary Note 5 – Changes in cell-cell variation with age depend on cell identity

Previous studies has reported that both gene expression variance and cell-cell heterogeneity increase with age using RNA-seq [Martinez-Jimenez et al., 2017, Enge et al., 2017, Angelidis et al., 2019] and quantitative PCR methods [Bahar et al., 2006]. These two types of variation differ in subtle but important ways. Gene expression variance quantifies the mean dispersion across genes in the transcriptome, such that each gene contributes equally. Because genes are equally weighted, changes in gene expression variance are unlikely to be driven by a small number of genes. Previous aging studies suggested that increased gene expression variance may reflect a global change in transcriptional noise, perhaps due to loss of regulatory control [Bahar et al., 2006, Todhunter et al., 2018].

By contrast, cell-cell heterogeneity measures the average distance in transcriptional space between cells in a population. These distances capture multivariate differences in gene expression. They also account for gene expression level, such that a small number of more highly expressed genes can drive changes in cell-cell heterogeneity. Increased cell-cell heterogeneity may reflect a diversification of cellular states within a population (Fig. 2A).

Both transcriptional variation and cell-cell heterogeneity have important implications for cell physiology and function, as explored in seminal studies of transcriptional noise in cell fate selection [Blake et al., 2006, Sueel et al., 2007] and bet hedging [Kussell et al., 2005, Altschuler and Wu, 2010, Cohen, 1966, Slatkin, 1974, Beaumont et al., 2009]. To determine if age-related changes in variation and heterogeneity depend on cell identity and tissue environment, we evaluated both properties across the many combinations that we observed.

We evaluated transcriptional variation using the difference from the median (DM) method [Kolodziejczyk et al., 2015] to estimate "overdispersion." Overdispersion refers to the residual variation in gene expression observed for a given gene after accounting for the mean:variance relationship in gene expression data (Methods). Across cell identity/environment combinations, we found that transcriptional variation experiences subtle changes (Fig. 2B, median Cohen's $d = 0.031$). Individual cell identities can be identified that exhibit either increased (lung leukocytes) or decreased (kidney mesangial cells) variance with age (Fig. 2D).

These results do not contradict previous observations of increased variation with age [Martinez-Jimenez et al., 2017, Enge et al., 2017]. Previous studies used a different method based on ERCC spike-in molecules to estimate gene-wise technical variation [Vallejos et al., 2015, Martinez-Jimenez et al., 2017, Enge et al., 2017]. In addition, we analyzed the whole transcriptome in unchallenged T calls, in contrast to previous studies that observed increased gene expression variance in a subset of T cell genes after stimulation. However, our results do suggest changes in gene expression variance with age are dependent on cell identity and cellular context.

We quantified cell heterogeneity in each cell identity/environment combination using the distance to the centroid method as previously introduced [Enge et al., 2017, Angelidis et al., 2019]. Cell-cell heterogeneity appears to increase for many cell identities (Wilcoxon Rank Sums, $q < 0.05$, median Cohen's $d = 0.473$), including B cells across all three tissues, and lung stromal cells. We also observed decreased heterogeneity with age in some cell identities, such as lung type II pneumocytes and kidney CD8 T cells (Fig. 2C). To directly compare with previous results in the literature, we also computed the distance to the centroid in a subspace of heuristically identified low-variance genes [Enge et al., 2017, Angelidis et al., 2019]. We find that the results from our full transcriptome analysis and this heuristically identified subset of genes are highly correlated ($\rho > 0.9$; Supp. Fig. S12A, C; Methods). We similarly observed highly similar changes in cell-cell heterogeneity when computing distance to the centroid using the Manhattan rather than Euclidean distance ($\rho > 0.9$; Supp. Fig. S12B, D; Methods). Taken together, these results indicate that changes in gene expression variance and cell-cell heterogeneity with age depend on cell identity. This suggests a nuanced view of changes in noise as a hallmark of aging at the single cell level [Todhunter et al., 2018].

Supplementary Note 6 - Inflammatory gene expression in non-immune cell identities

Increased inflammation-associated gene transcription has previously been reported in multiple tissues. However, previous investigations have relied on bulk transcriptional assays, making it difficult to determine if inflammatory genes signatures were upregulated in all cells within a tissue or if more immune cells had infiltrated the tissue [Rodwell et al., 2004, O'Brown et al., 2015, Benayoun et al., 2019]. To investigate the possibility that immune pathways are upregulated across many non-immune cell identities, we identified a set of genes that are differentially expressed in more than 3 non-immune cell identities (counted as above). Gene Ontology enrichment analysis on this gene set revealed that inflammatory gene sets (T cell activation, B cell activation, viral entry, cytokines) were upregulated, even in these non-immune cell identities. However, the enrichment was less significant than when immune cells are included ($q < 0.05$, rather than $q < 10^{-3}$, Supp. Fig. S15B). At the gene level, we found *Ikgc*, *Cd74*, and *B2m* are commonly upregulated with aging (Supp. Fig. S15A). *B2m* has previously been reported to increase in the aging systemic milieu, and reports have indicated it may play a causal role in aging brain pathology [Smith et al., 2015].

Supplementary Note 7 - Aging Trajectory Permutation Testing and Bootstrap Resampling

We performed permutation testing and bootstrap sampling to evaluate the robustness of our aging trajectory results. Permuted cell labels did not lead to clustering of lymphocyte trajectories, have roughly 2-fold less variance explained by cell identity in linear models, and show a significantly different distribution of cosine similarities (Supp. Fig. S19A, B, C). Using permuted samples as null distributions, our observations were statistically significant in each of these comparisons ($p < 0.01$). Bootstrap resampling likewise indicated that cosine similarities between aging vectors are robust to resampling noise (Supp. Fig. S19D, E). Bootstrap resampling by leaving out one animal of each age for each iteration yielded a similar result (Supp. Fig. S20). We also performed this analysis using a PCA embedding and achieved the same result, indicating robustness to the choice of an embedding space (Supp. Fig. S21).

Supplementary Figures

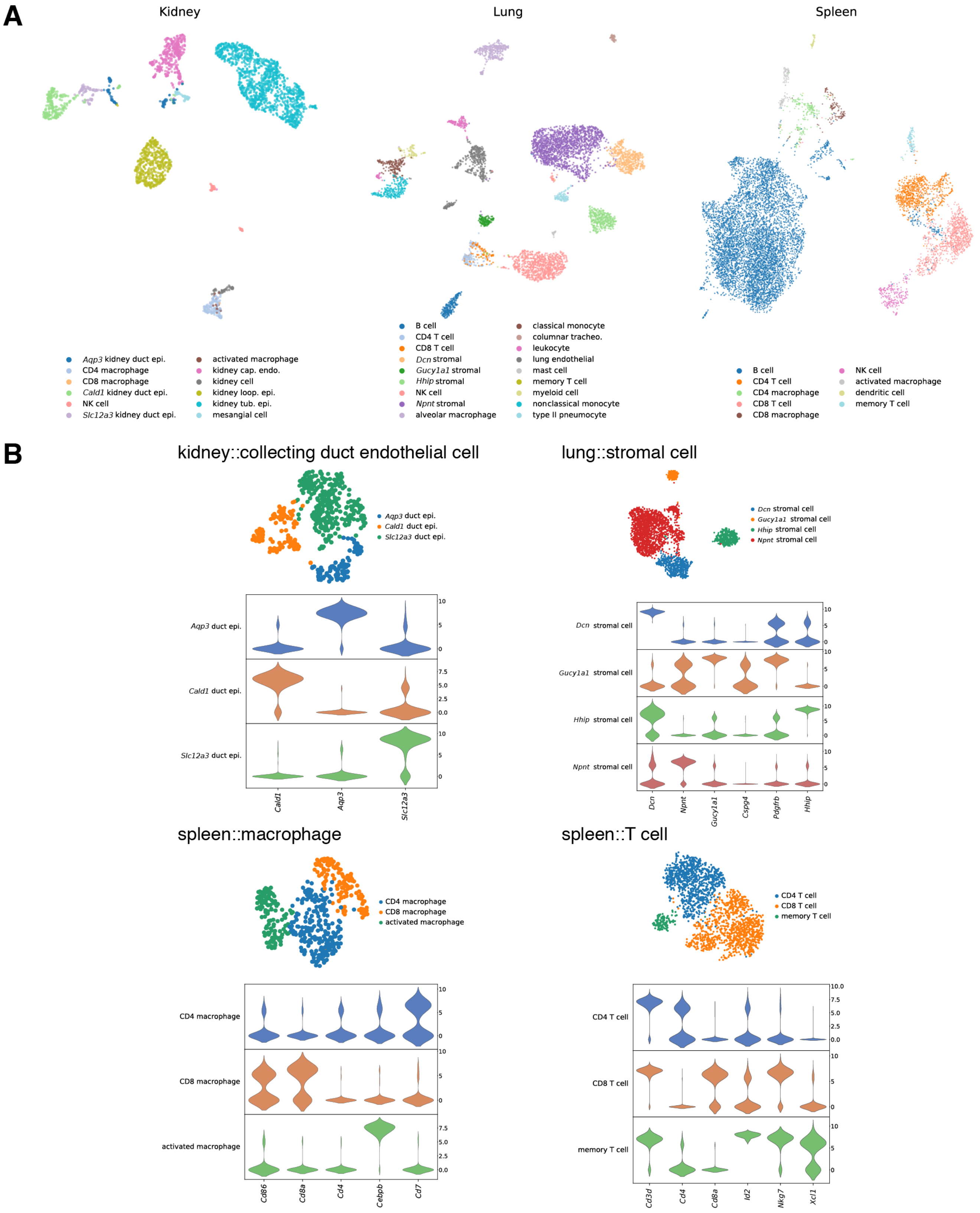
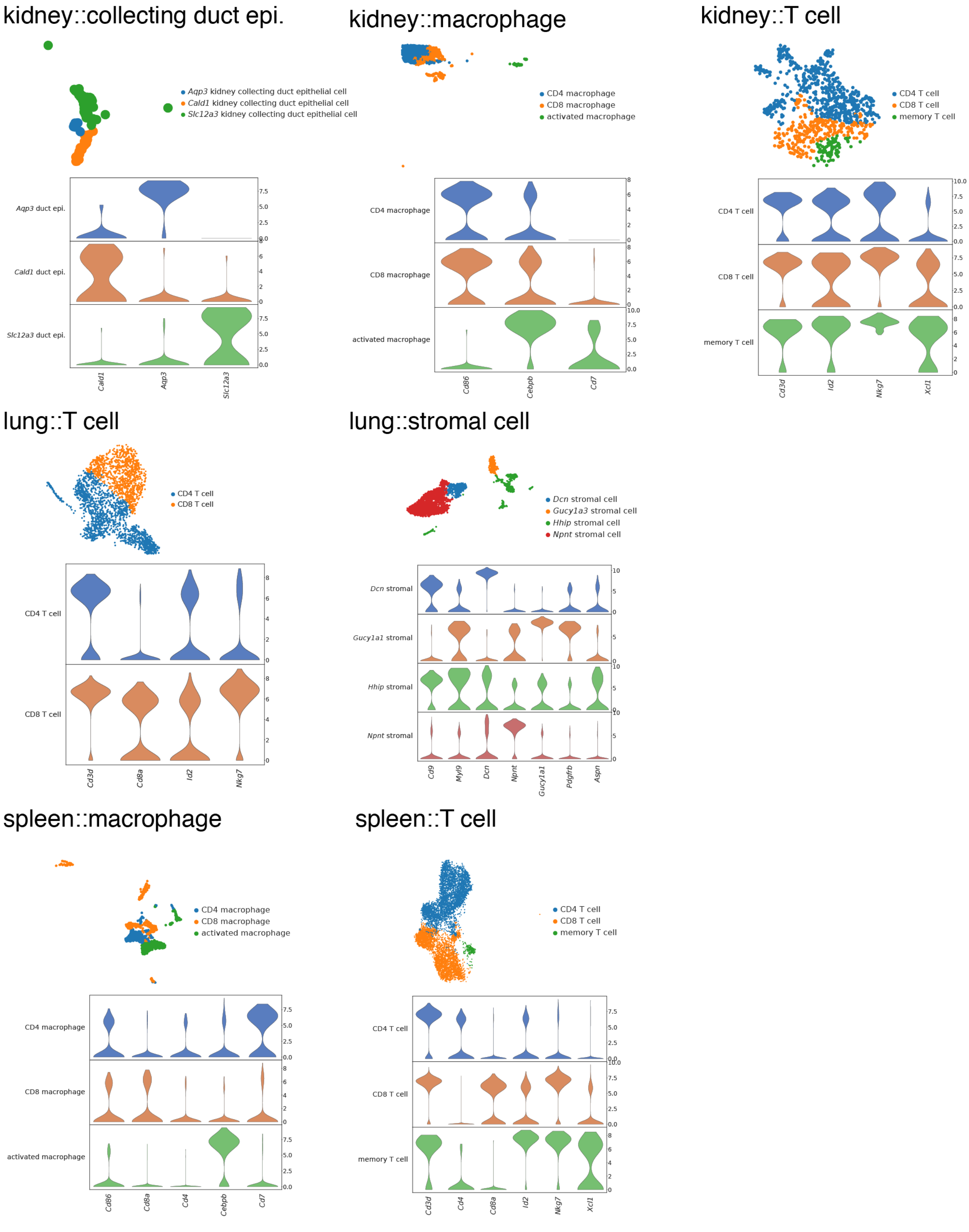


Figure S1: **Manual annotation of cell states in the Tabula Muris.** **(A)** UMAP projections of the Tabula Muris data with our manual cell state annotations overlaid as colors. We annotated cell states where there were known canonical states within a cell type (i.e. CD4 and CD8 T cells) and where we observed substructure within a cell type in our data and the *Tabula Muris* data (as in the lung stromal cells). **(B)** UMAP projections of each cell type, with manual cell state annotations. Below each UMAP projection, we show the expression of marker genes that guided our cell state annotations.



kidney::collecting duct epi.

kidney::macrophage

kidney::T cell

lung::T cell

lung::stromal cell

spleen::macrophage

spleen::T cell

Figure S2: **Cell state annotations in our data derived from neural networks.** We trained deep neural networks to classify cell states within individual cell types using our manual annotations of the *Tabula Muris* (Supp. Fig. 1). UMAP projections of each cell type are presented with cell state annotations overlaid as color labels. Cell states were enriched for corresponding marker genes, as in the *Tabula Muris*, presented as violins below each UMAP projection.

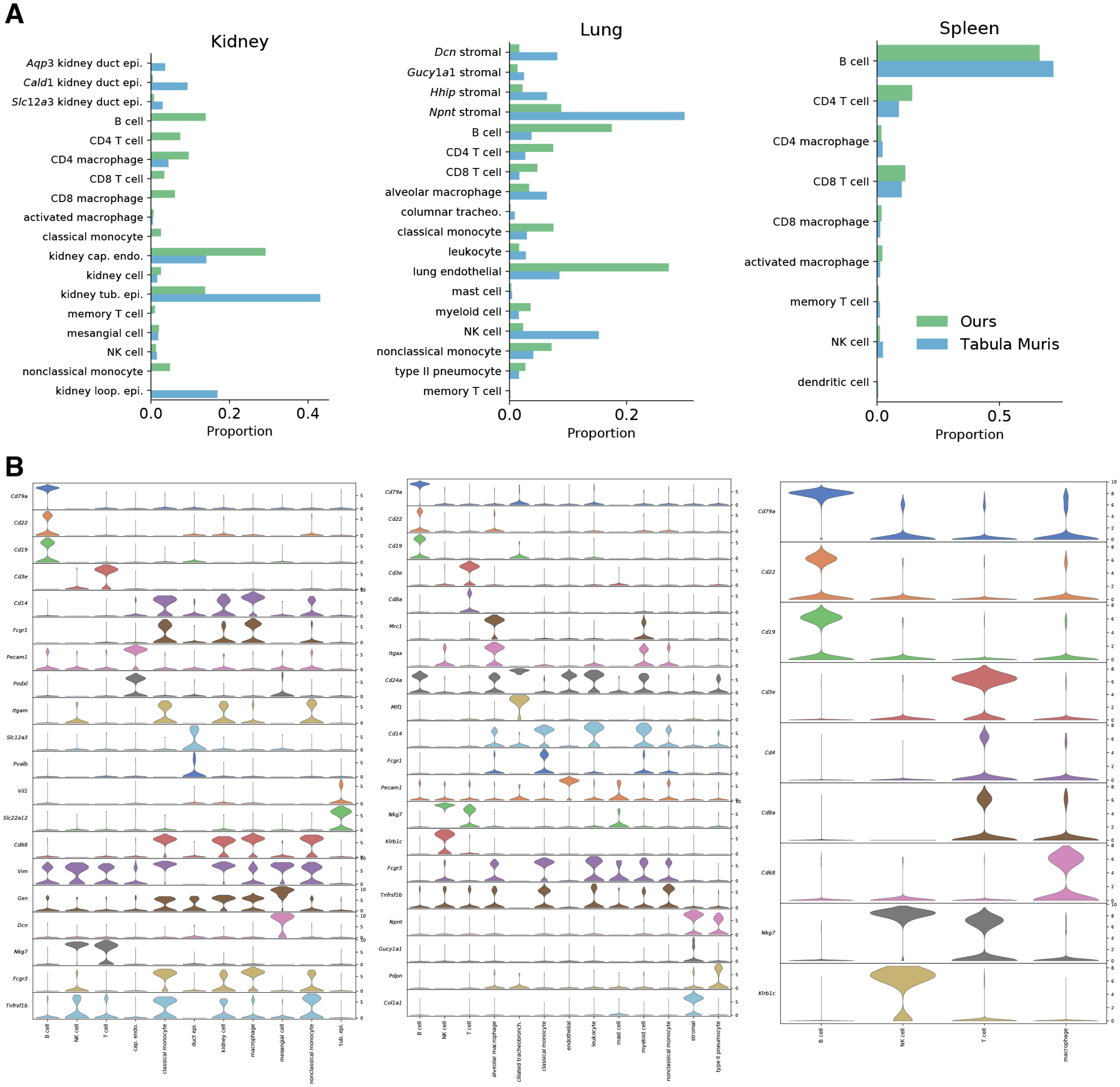


Figure S3: **Comparison of our cell state proportions to the *Tabula Muris* and cell type markers.** **(A)** Cell state proportions from our dataset (green) compared to proportions observed in the Tabula Muris (blue). Notable differences include a higher proportion of immune cells in our data set and the absence of kidney loop of Henle epithelial cells. These differences may be the result of intentional experimental differences (perfusion vs. no perfusion), animal ages (Tabula Muris animals are younger than our young animals), and laboratory-to-laboratory differences in isolation technique. **(B)** Expression of marker genes across cell states we identify in our data. Marker genes were either taken directly from those suggested by the Tabula Muris or from marker genes that we identified in their data using our differential expression procedure.

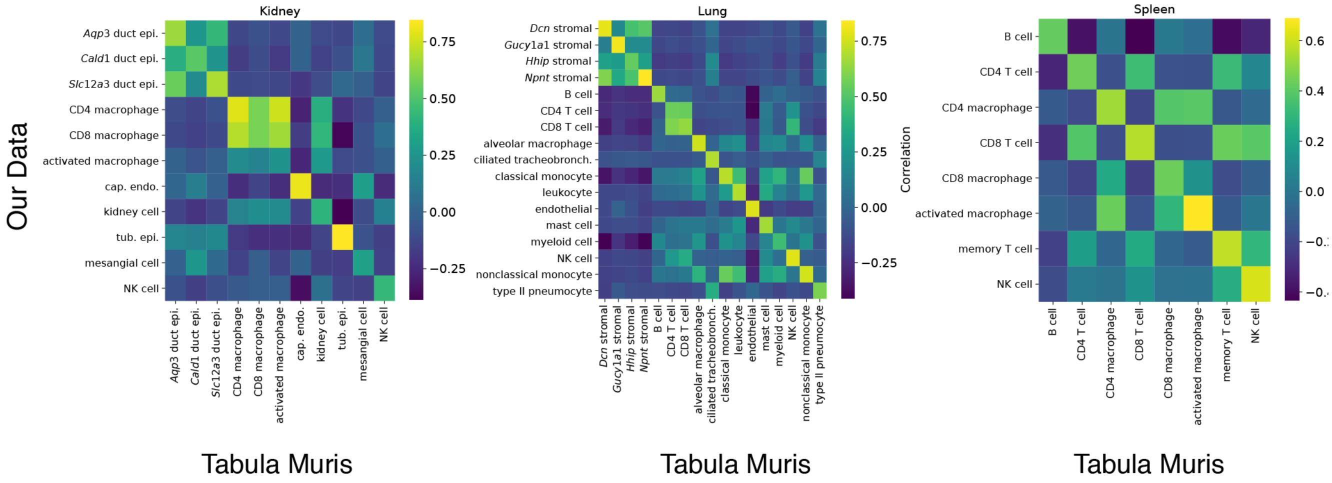


Figure S4: **Correlation of cell identities between our data and the *Tabula Muris*.** We computed mean expression vectors for each cell identity in our data and the *Tabula Muris* as the mean expression of each gene across cells in a given identity. We computed correlations between these mean expression vectors for all identities between the two data sets. Visualizing these correlations as a heatmap, we find that identities in our data are most similar to corresponding identities in the *Tabula Muris*.

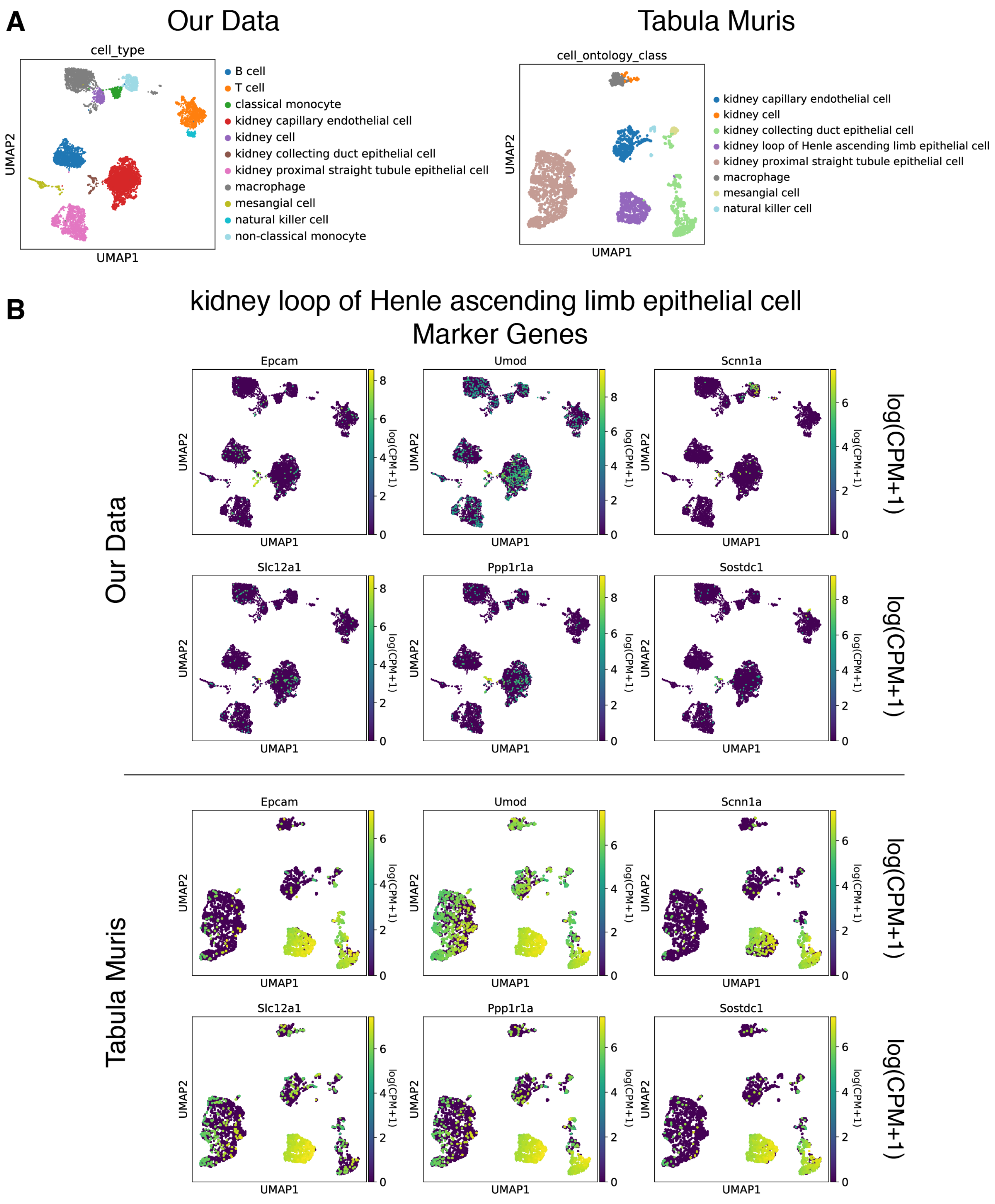


Figure S5: **Visualization of kidney loop of Henle epithelial cell markers from the Tabula Muris indicates these cells are not present in our data.** **(A)** Kidney UMAP projections of our data and Tabula Muris data. Note the absence of kidney loop of Henle epithelial cells in our data set. **(B)** Expression of marker genes for kidney loop of Henle epithelial cells in our dataset and the Tabula Muris. Colors indicate the expression of each gene after $\log(\mathrm{CPM}+1)$ transformation, with purple indicating low values while green and yellow colors indicate higher expression levels. We found no coherent group of kidney loop of Henle cells in our data based on these markers.

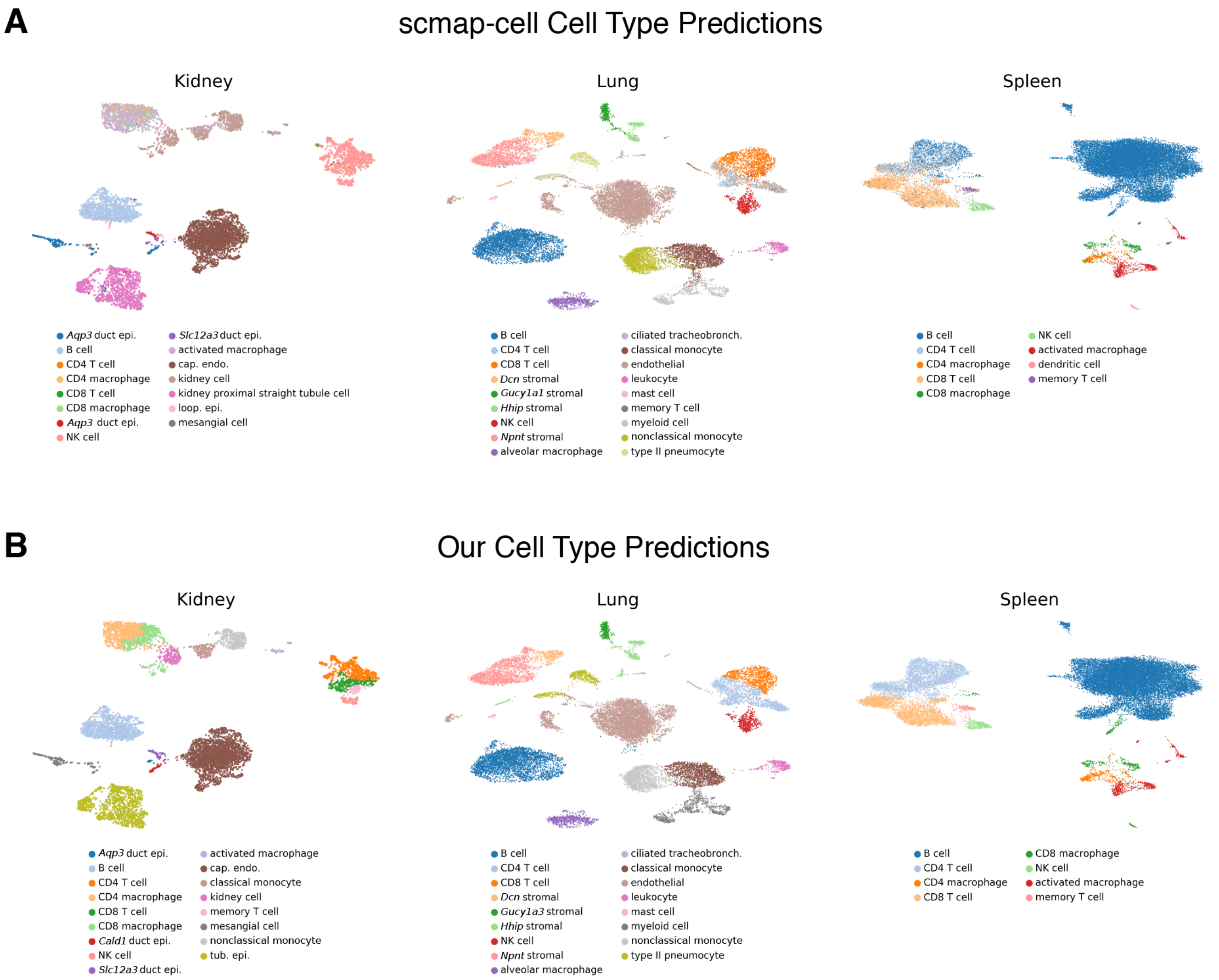


Figure S6: **scmap-cell cell type predictions are largely consistent with our annotations.** **(A)** Cell type predictions from `scmap-cell` after training on the *Tabula Muris*. `scmap-cell` predictions agreed with our annotations for most cell types, but failed to properly identify many rare cell types. For example, kidney mesangial cells were misclassified as kidney collecting duct epithelial cells and kidney classical and non-classical monocytes were misclassified as kidney cells. Activated macrophages are also misclassified as kidney cells. Many kidney T cells were also misclassified. In the spleen, CD4 T cells were likewise misclassified, despite clear marker gene expression. **(B)** Cell subtype labels derived from neural network classifiers and manual validation. We found that our labels are more consistent with marker gene expression than `scmap-cell` predictions.

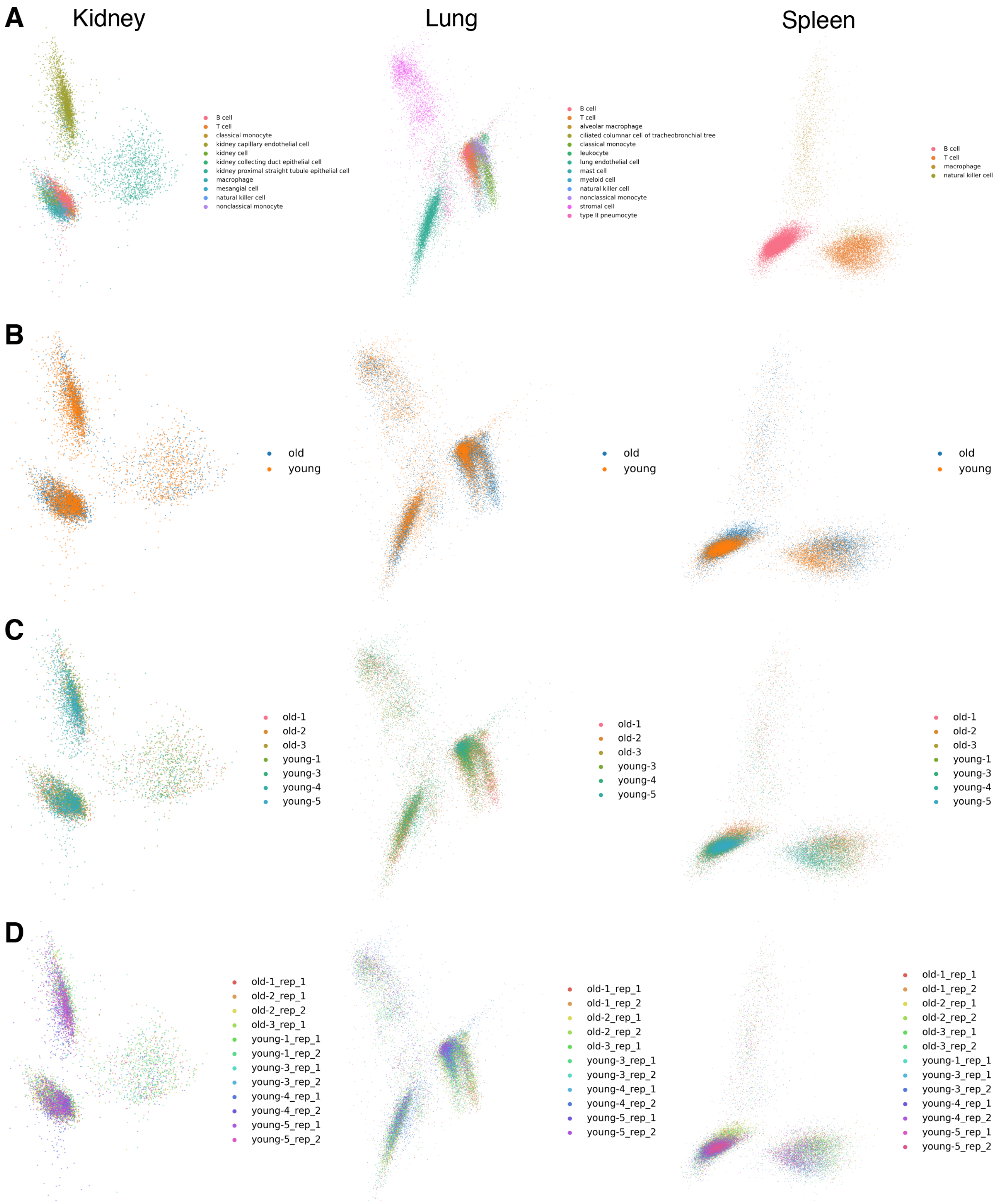


Figure S7: **Cells cluster by cell type and age, rather than batch.** **(A)** PCA projection of all cells, colored by cell type. We observed that cells segregate by type, but the PCA projection does not capture cell type variation as well as the UMAP. **(B)** PCA projection colored by cell age. We did not observe clustering by age in the PCA projection. **(C)** PCA projection of all cells, colored by the of the animal of origin. We observed that cells cluster primarily by cell type, although many cell types did not segregate in the linear projection. We did not observe clustering by age or batch. **(D)** PCA projection of all cells, colored by the of the animal of origin and technical replicate identifier. We likewise did not observe clustering by technical replicate.

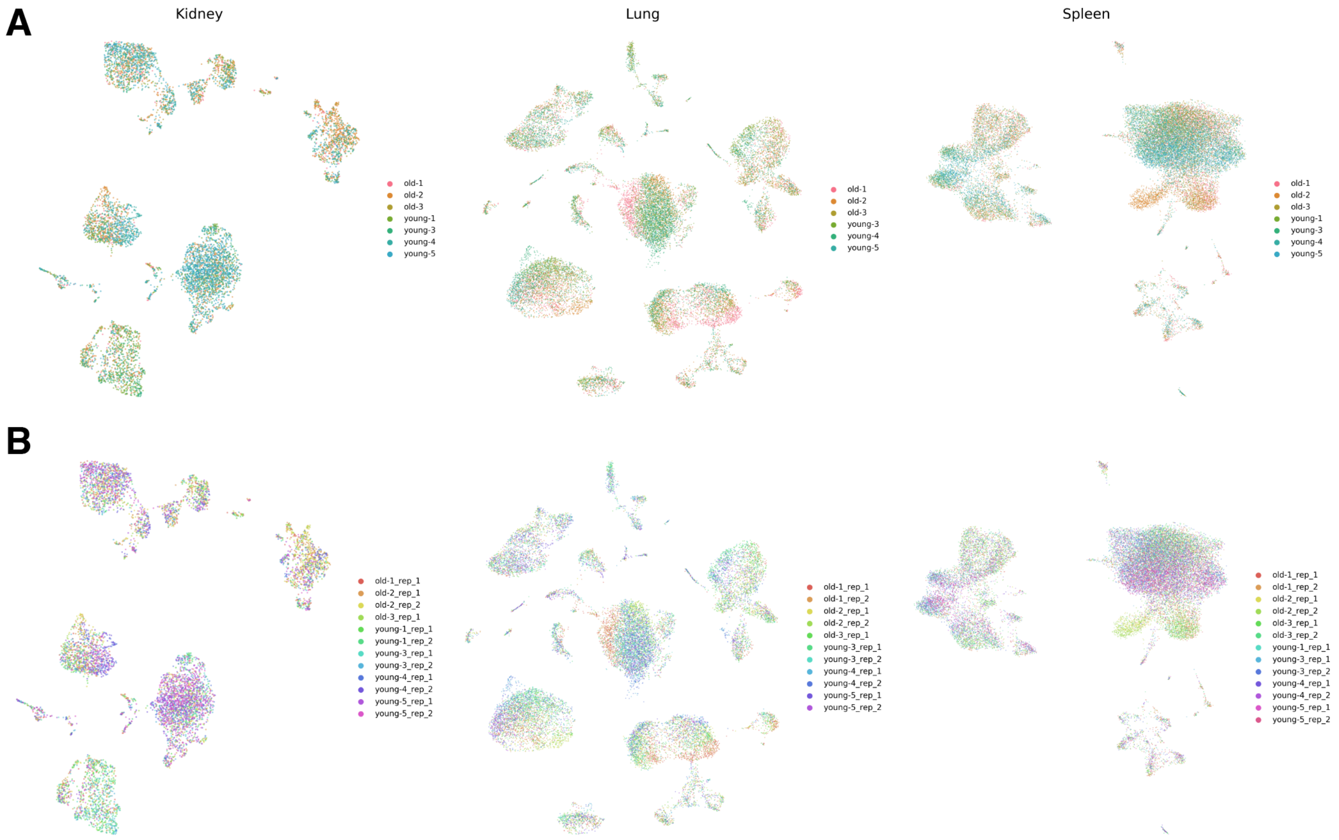


Figure S8: **Cells cluster by cell type and age, rather than batch.** **(A)** UMAP projection of all cells, colored by animal of origin. We observed that cells cluster primarily by cell type, to a lesser extent by age, and do not appear to cluster by replicate animals. **(B)** UMAP projection of all cells, colored by animal of origin and technical replicate identifier. We likewise find that cells do not cluster by technical replicates.

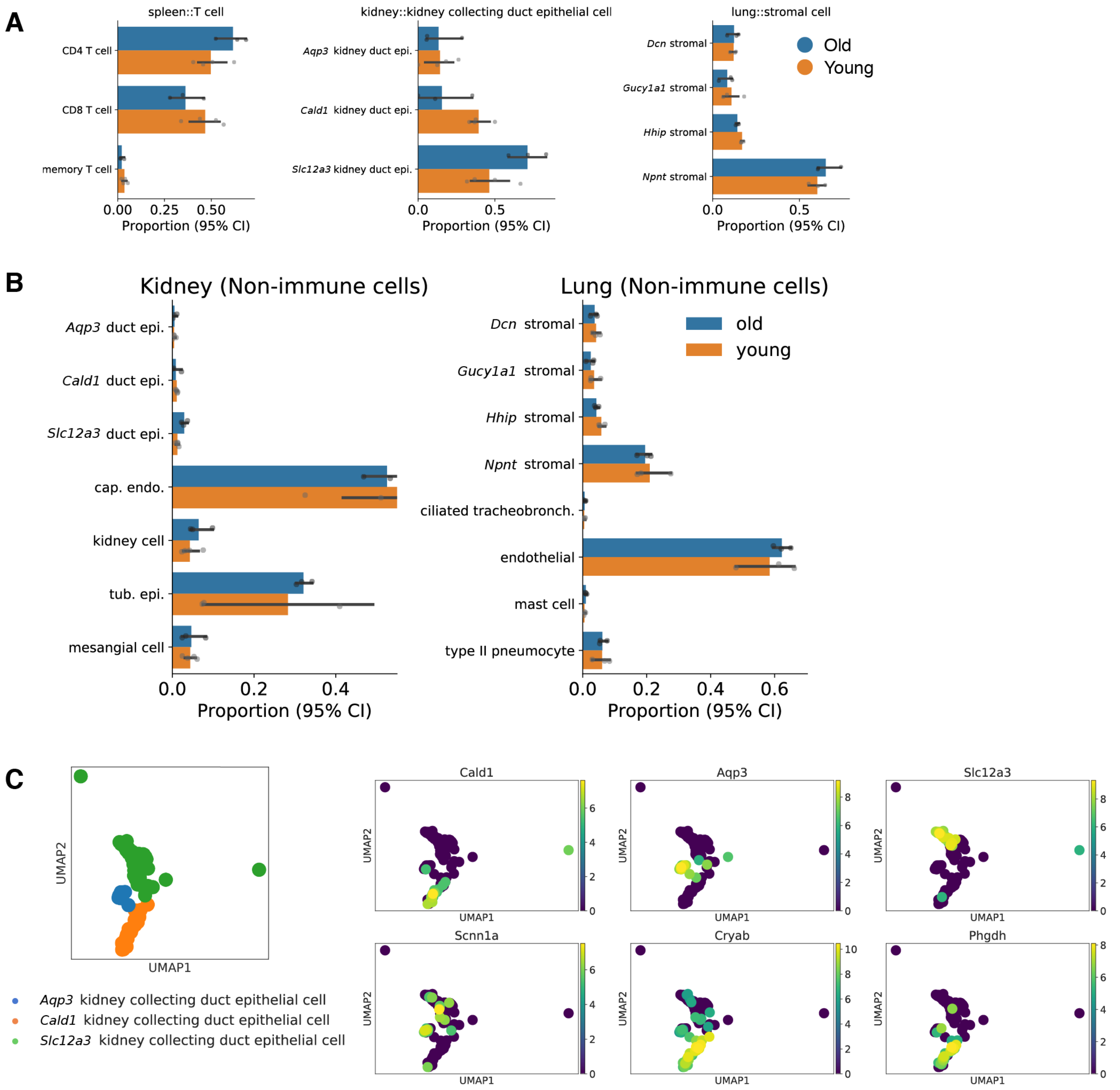


Figure S9: **Cell state proportions shift with age.** **(A)** Cell subtype proportions with age for individual cell types. Kidney collecting duct epithelial cells (left) showed a significant shift in the subtype distribution with age (Chi-squared contingency table, $p < 0.05$). By contrast, lung stromal cells, with a similarly complex cell type composition, did not change with age (center). In the spleen, we detected a well-known shift in T cell subtypes with age; old animals exhibited a lower proportion of CD8 T cells (Chi-squared contingency table, $p < 0.05$). **(B)** Cell subtype proportions with age for all non-immune cell types in the kidney and lung. After removing immune cells, we found that most cell types in both tissues have similar proportions between young and old animals. **(C)** *Aqp3* and *Slc12a3* kidney collecting duct epithelilial cells co-express *Scnn1a* (epi. sodium channel ENaC), suggesting they are principal cells of the collecting duct. *Cald1* duct epithelial cells also express *Phgdh* and *Cryab*, consistent with reports of a subpopulation in the literature with unknown function [Takasaki et al., 2007, Han et al., 2018].

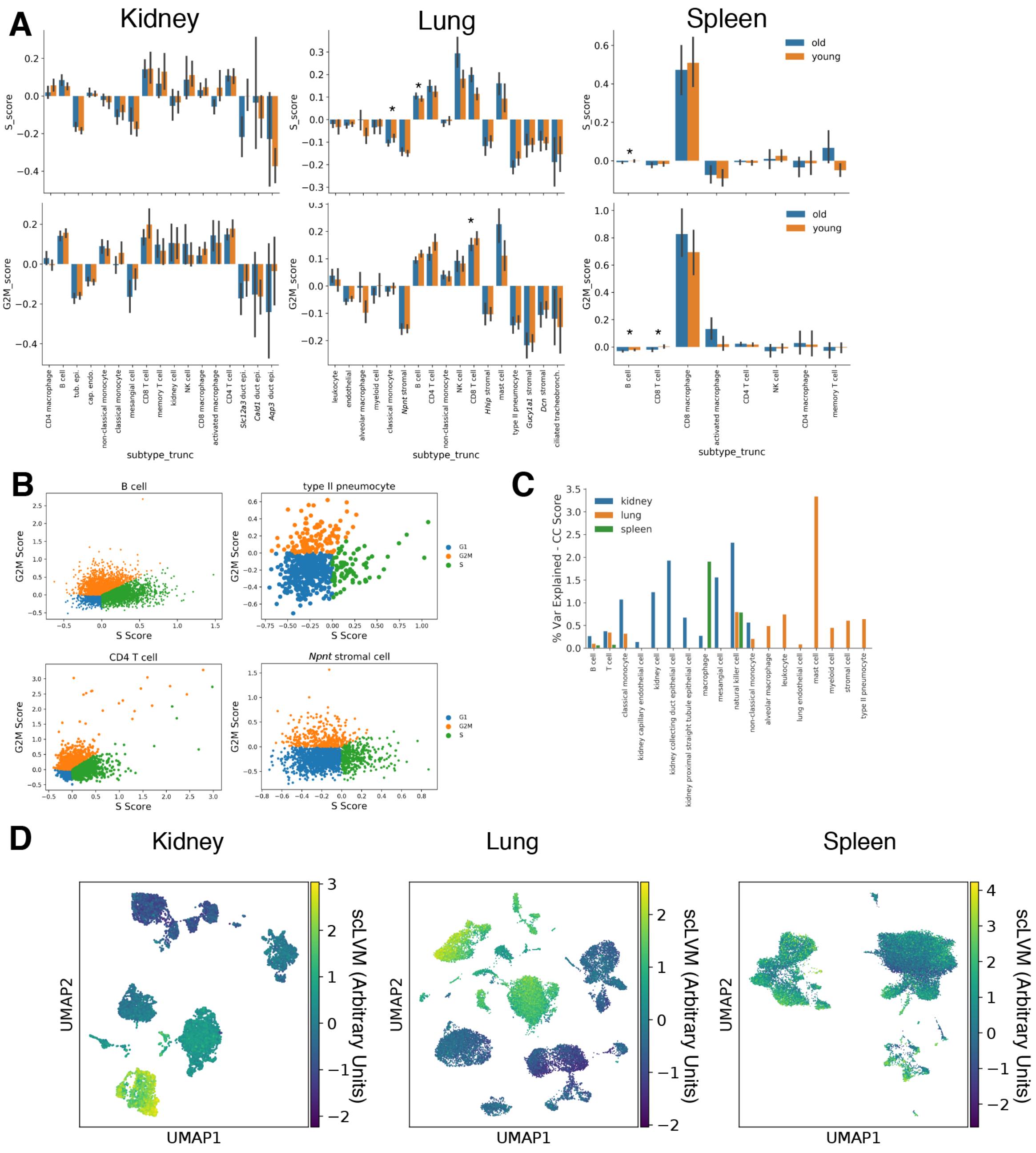


Figure S10: **Cell cycle activity is largely unchanged with age.** **(A)** S-phase gene module and G2/M-phase gene module scores each cell type in each tissue. We found little difference in cell cycle module scores across cell types, although a few changes are statistically significant (*: $q < 0.05$, Wilcoxon Rank Sums). **(B)** Representative cell cycle phase plots based on gene module scoring for S-phase and G2M-phase genes. Here, we show individual cell states from the lung. We found very few distinctly cycling cells across all those observed. **(C)** Proportion of variance explained by cell cycle module scores based on linear models. **(D)** We fit scLVM models to explain variance in cell cycle genes within each tissue. We found that the scLVM latent variable captures variation in the baseline expression of cell cycle genes across cell types, rather than variation within individual cell types due to the cell cycle. Here, we show the value of the scLVM latent variable pseudocolored on cells in the latent space. Differences in intensity mostly occur across cell type clusters.

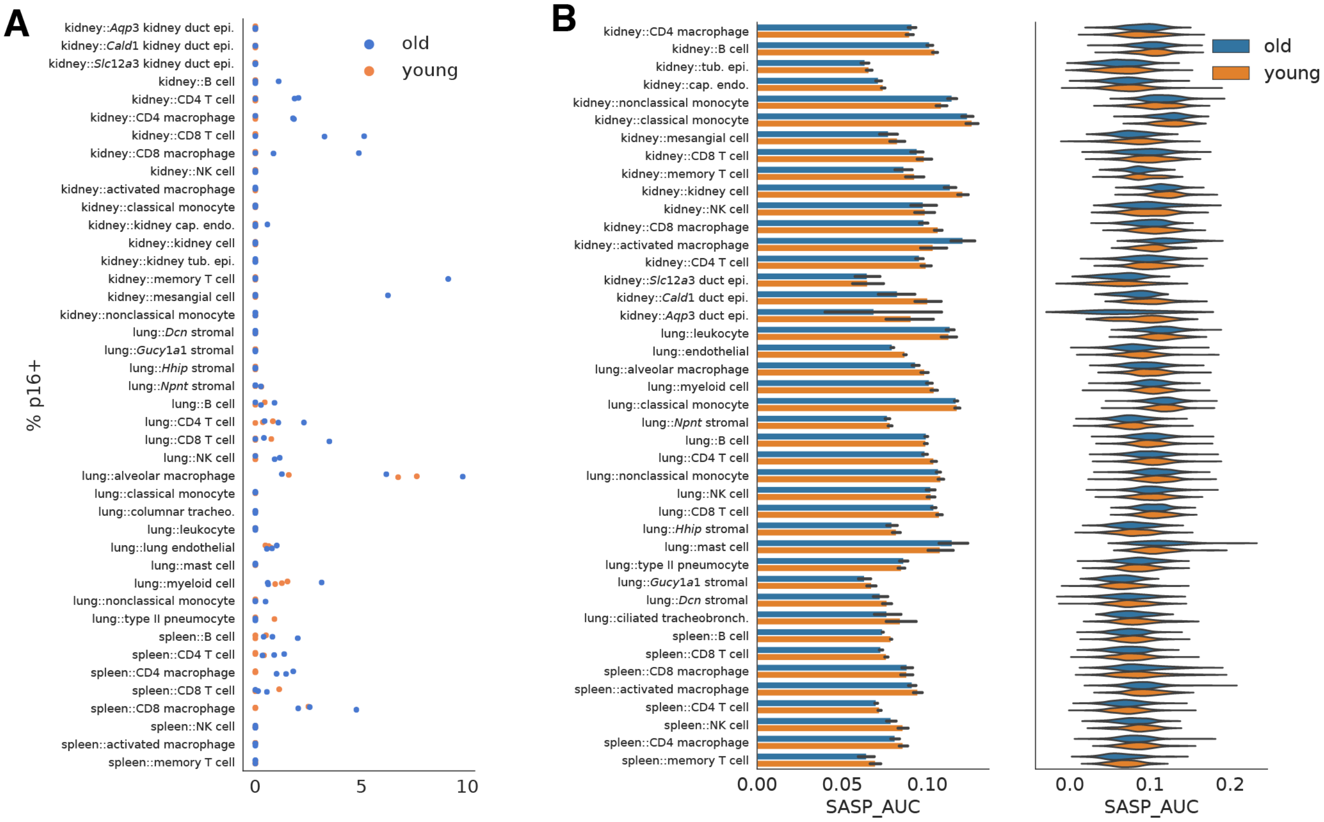


Figure S11: **Neither *Cdkn2a* (p16-Ink4a) nor senescence-associated gene activity is significantly upregulated in old cells.** **(A)** We quantified the percentage of cells in each identity that express *Cdkn2a* for each animal. No cell identity showed a significantly increased proportion of *Cdkn2a*+ cells. **(B)** We scored the activity of a manually curated set of senescence-associated secretory phenotype (SASP) genes using the AUCell approach for each cell identity and age. Greater AUCell score indicate greater activity of the gene program. SASP activity was not greater in old cells.

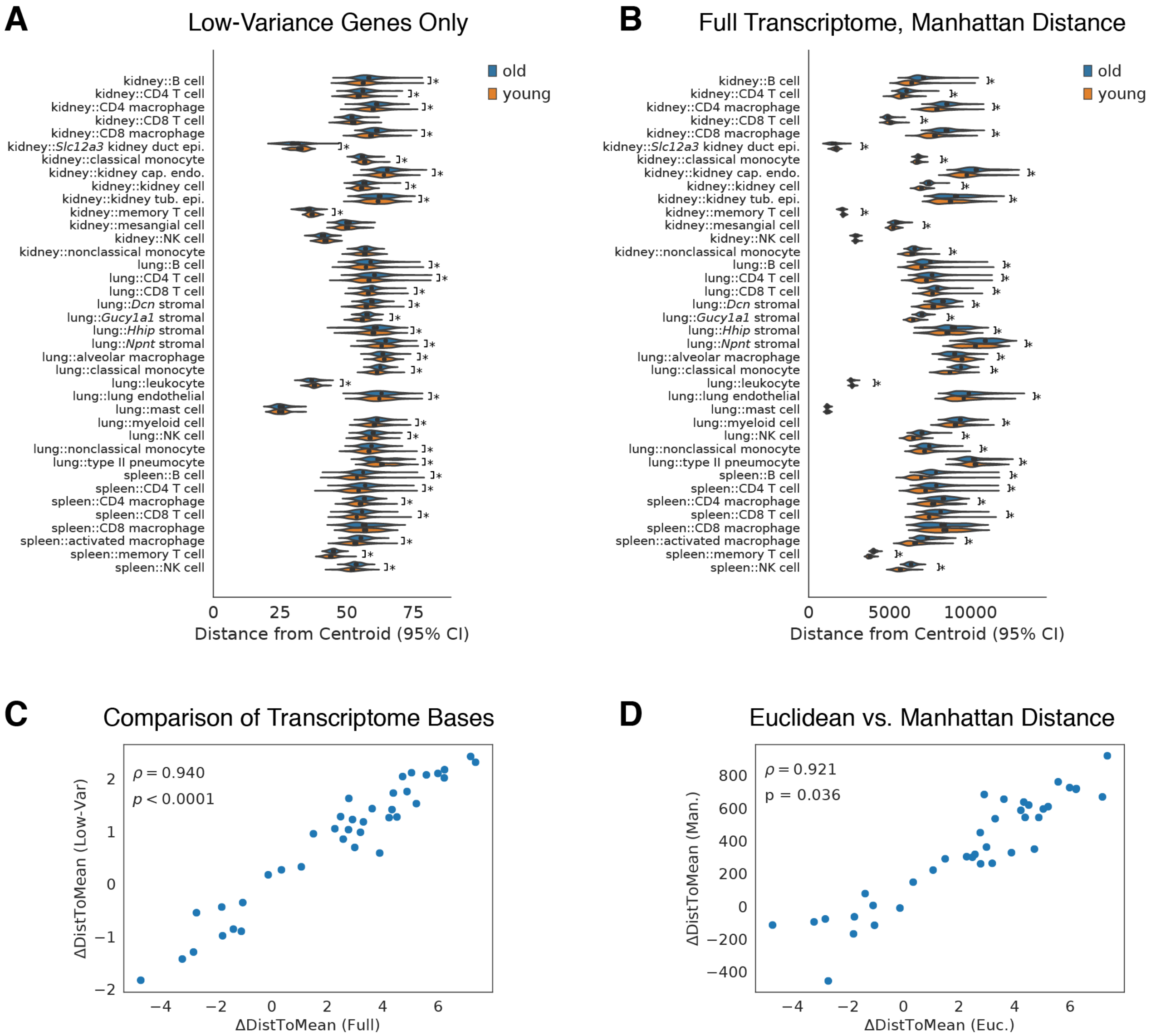


Figure S12: **Changes in gene expression variance and cell-cell heterogeneity are cell identity dependent.** **(A)** Cell-cell heterogeneity measurements based on Euclidean distances to the population centroid for each cell identity and age. We computed distances using only a subset of genes selected to have low-variance. We binned genes by mean expression into 10 bins and selected genes with a coefficient of variation at or below the 10%-tile for genes in the bin. Most cell identities exhibited increased cell-cell heterogeneity in old cells (*: Wilcoxon Rank Sums, $q < 0.05$). **(B)** Cell-cell heterogeneity measurements based on Manhattan distances to the population centroid for each cell identity and age. We used the full transcriptome for these comparisons, as in Fig. 2C. **(C)** Comparing the change in “distance to the centroid” (old - young) between our analysis using all transcripts (Fig. 2C, “Full”, x-axis) and a subset of low-variance genes (**A**, “Low-Var”, y-axis), measures of cell-cell heterogeneity were highly correlated between the two transcriptional spaces (Spearman’s $\rho > 0.9$). **(D)** Comparing the change in the “distance to the centroid” (old-young) computed using Euclidean and Manhattan distances across all transcripts. We find that the two distance metrics provide highly correlated results (Spearman’s $\rho > 0.9$).

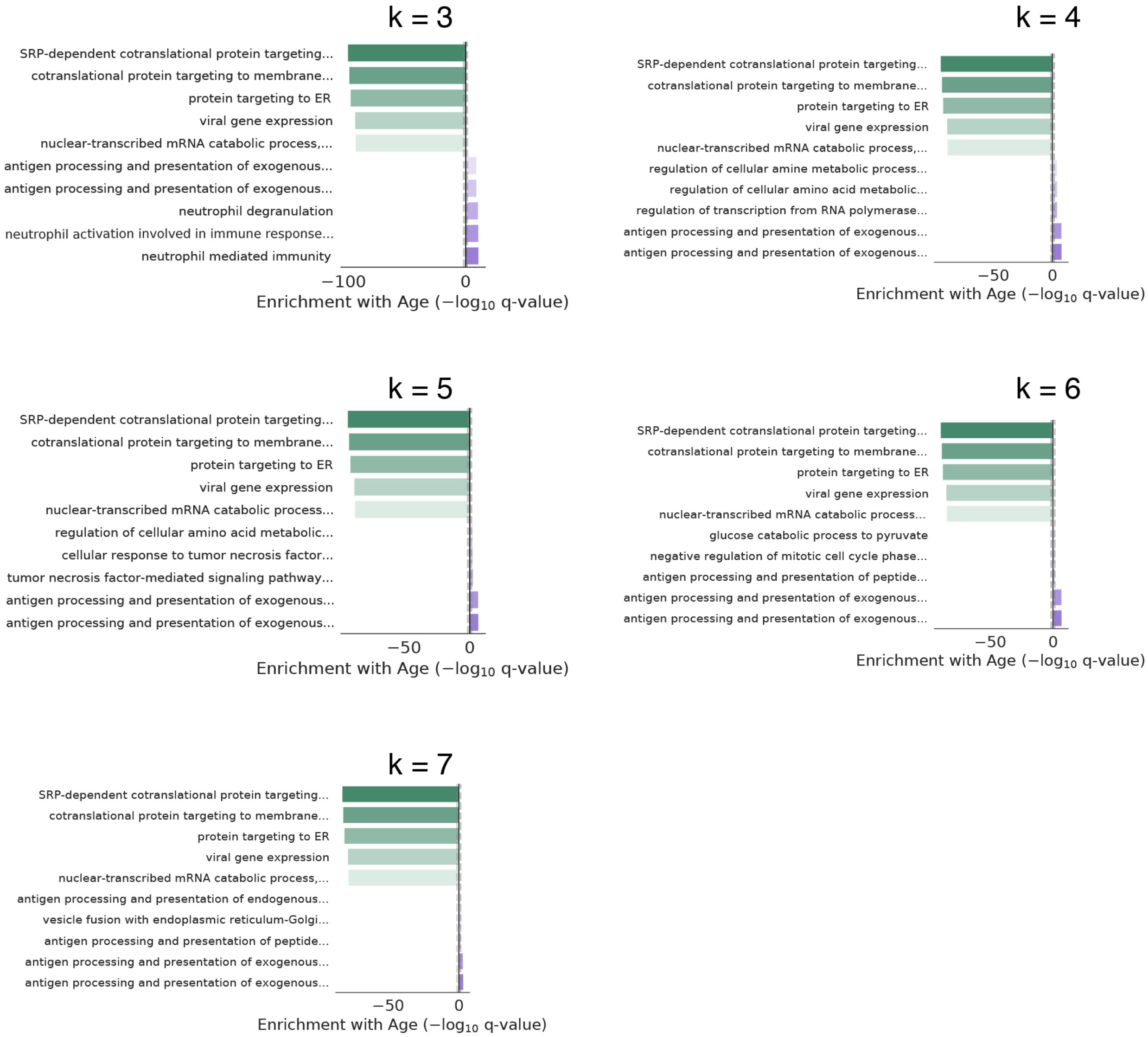


Figure S13: **Common differential expression gene sets are insensitive to threshold values.** We performed Gene Ontology enrichment analysis for common differentially expressed genes. We defined common differentially expressed genes as genes that change with age in the same direction in k or more cell identities. Gene Ontology enrichment results were not sensitive to a particular setting of k, providing similar results for values in the range $k \in [3, 7]$.

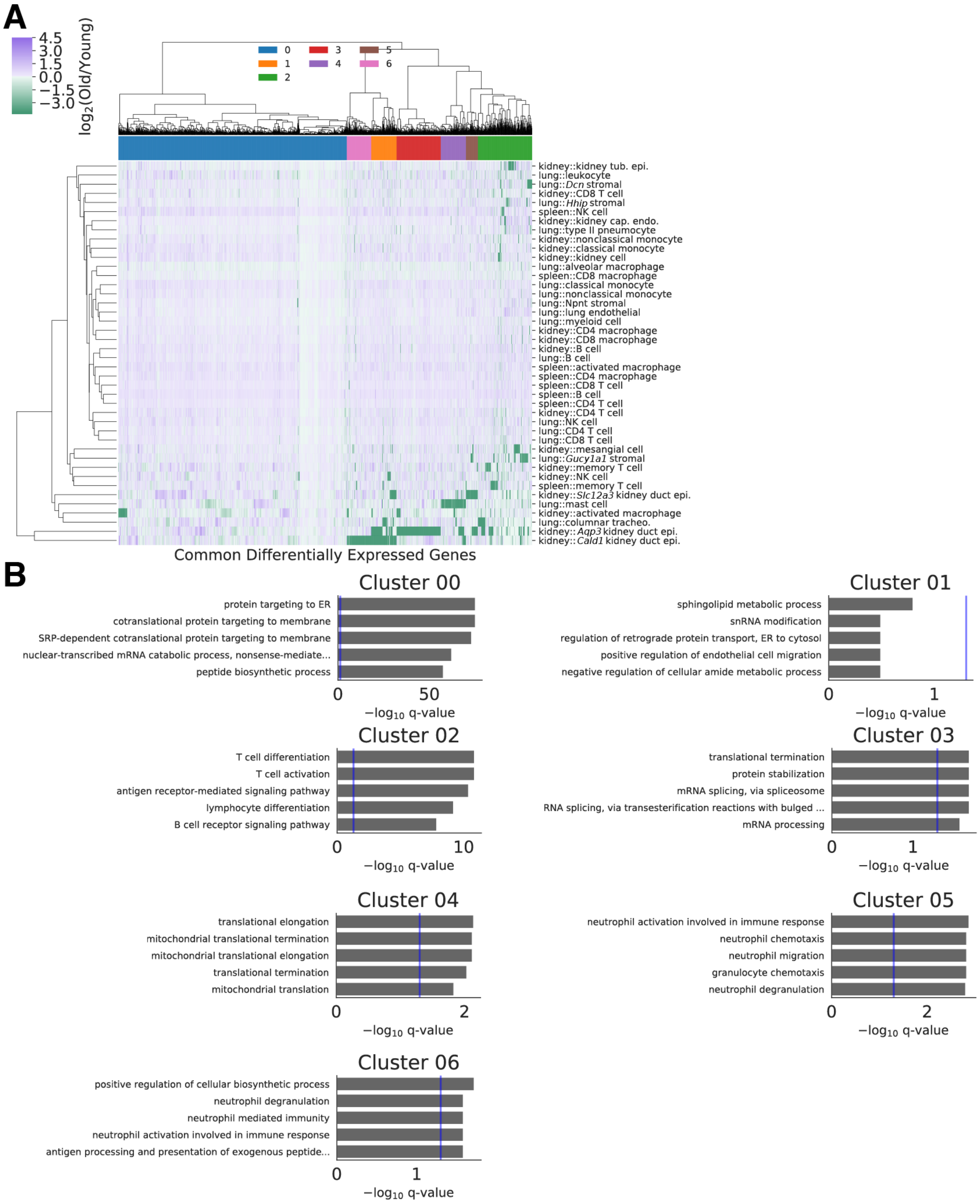


Figure S14: **Differential expression identifies a set of genes changed with age across many cell identities. (A)** We identified a set of 261 genes that change with age in at least $k = 5$ cell identities. Hierarchical clustering of the fold-changes within each cell identity identified subsets of these genes with similar behavior (color column labels). We used cosine similarity as an affinity metric for clustering. **(B)** Gene Ontology enrichment terms for genes within each gene cluster identified above.

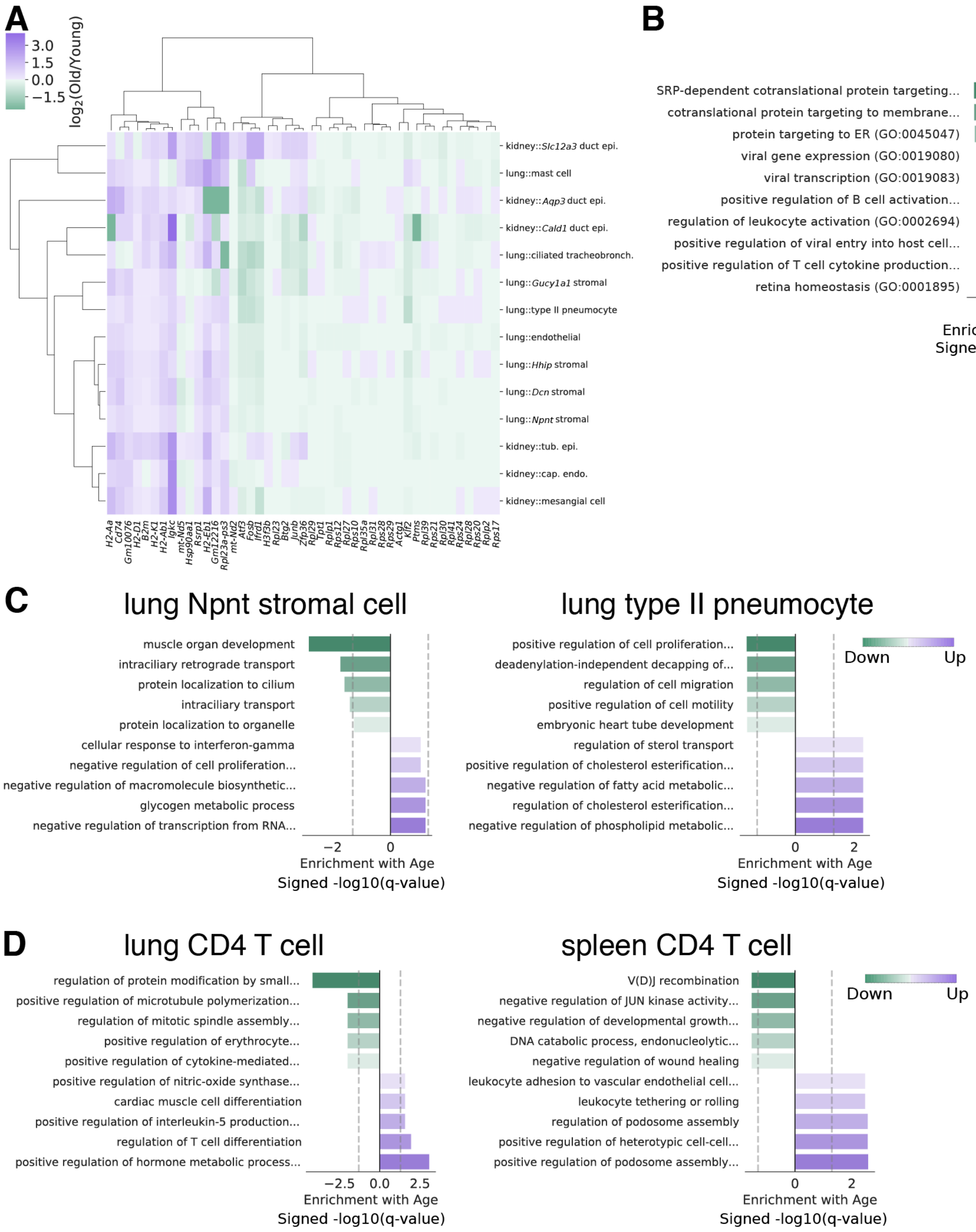


Figure S15: **Differential expression analysis identifies age-related changes unique to cell identity and tissue environment.** **(A)** Heatmap of common differentially expressed genes found in $k > 3$ non-immune cell states. Immune genes *B2m*, *Ikgc*, and *Cd74* were commonly upregulated with aging, even in these non-immune cells. **(B)** Gene Ontology enrichment terms for common differentially expressed genes in non-immune cell states. Dashed grey lines demarcate the $\alpha = 0.05$ significance threshold. Immunological activation pathways were upregulated even in these non-immune cell states, although the Gene Ontology enrichments were modest. SRP-dependent protein localization and ER targeting are again downregulated. **(C)** Gene Ontology enrichment analysis for genes up- and downregulated with age in type II pneumocytes, but not *Npnt* stromal cells and vice-versa. Dashed grey lines demarcate the $\alpha = 0.05$ significance threshold for enrichment. Genes downregulated with age reflect the mesenchymal nature of lung stroma and unique fluid shear stresses in type II pneumocytes. **(D)** Gene Ontology enrichment analysis for genes up- and downregulated with age in natural killer cells from the lung, but not natural killer cells from the spleen and vice-versa.

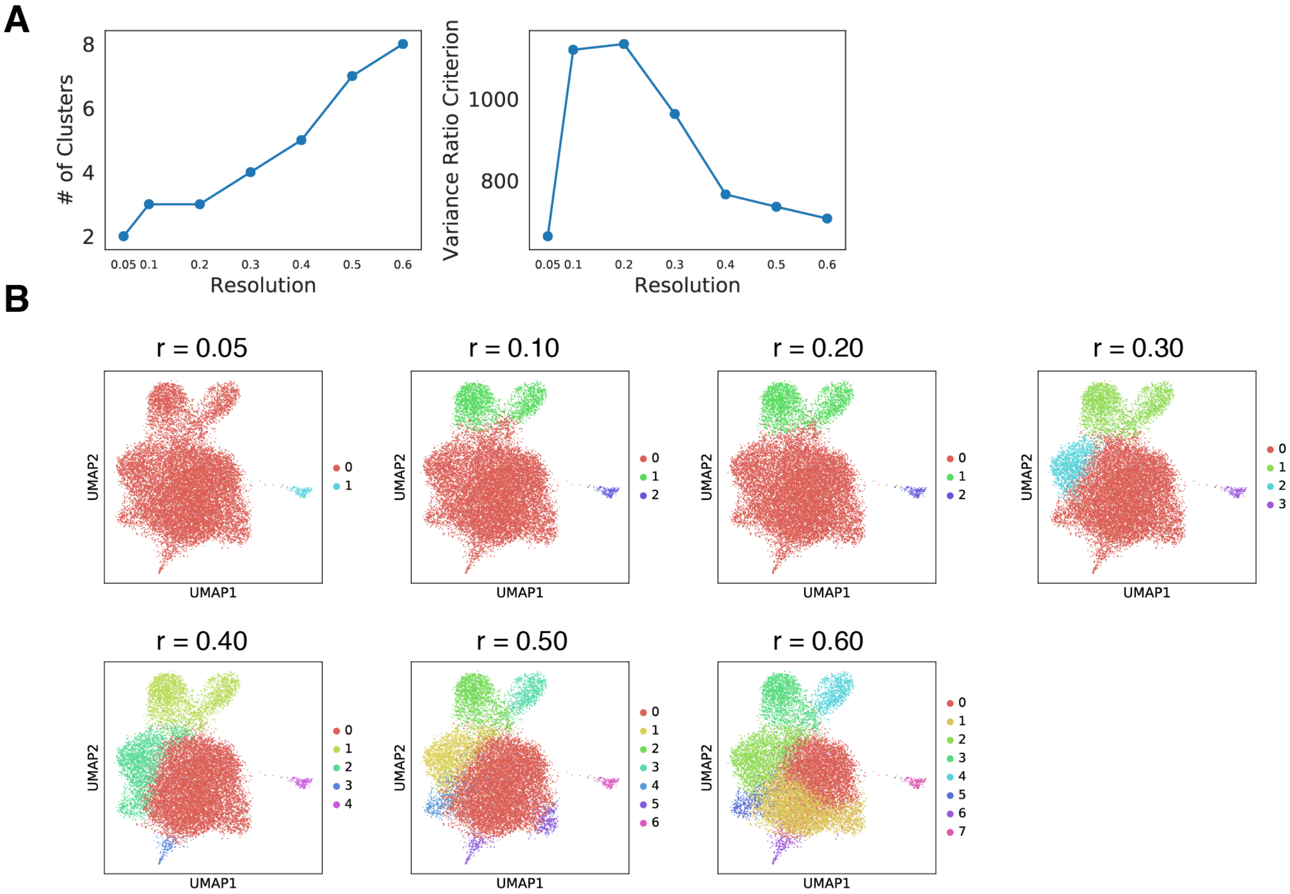


Figure S16: **Louvain community detection identifies three distinct B cell clusters.** **(A)** The number of clusters and Variance Ratio Criterion (VRC) for a range of resolution parameters to the Louvain community detection algorithm. Partitions with 3 clusters maximized the VRC, indicating that three clusters is optimal. **(B)** Louvain community detection partitions for a range of resolution parameters r visualized as colors on the UMAP projection.

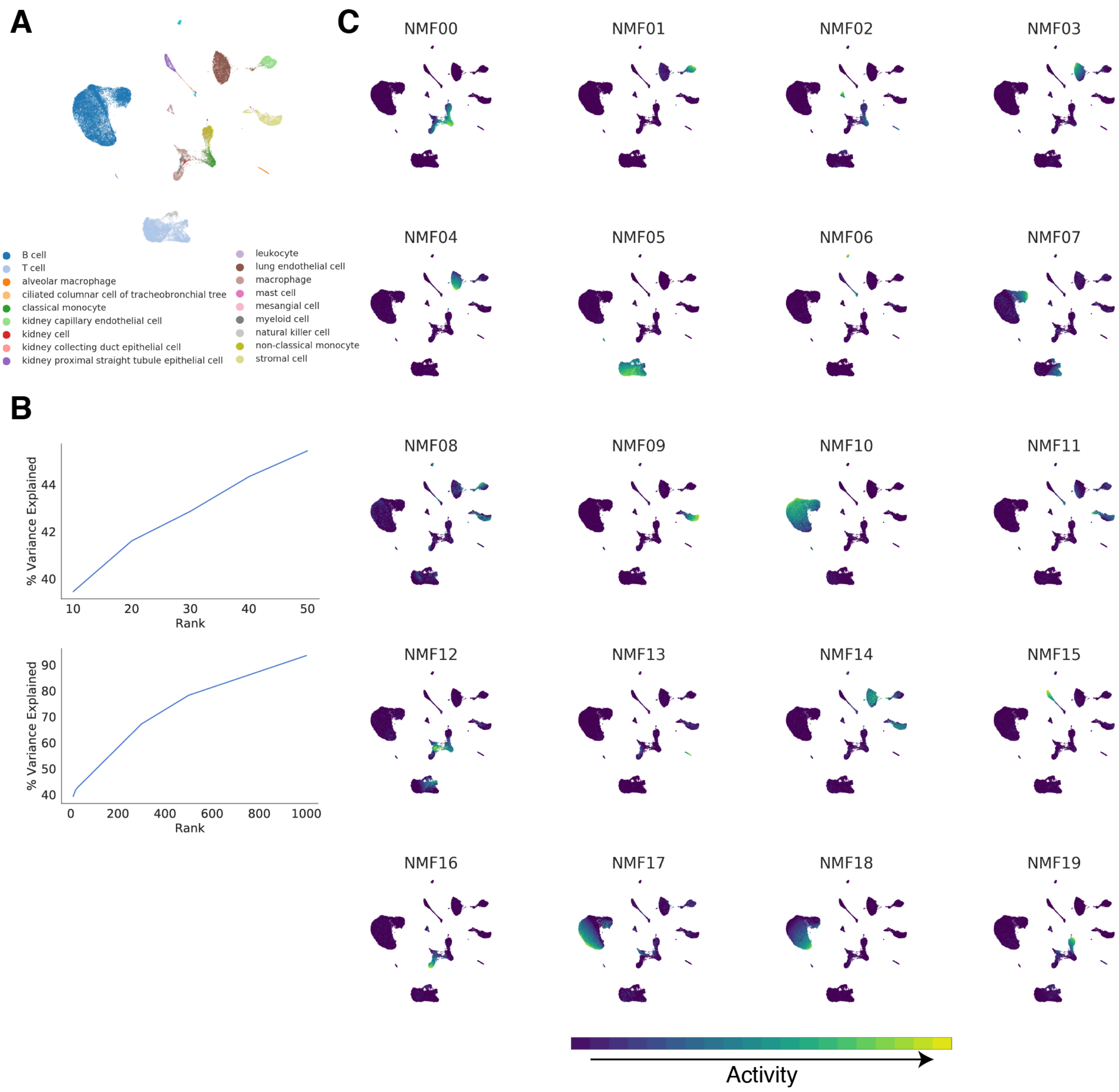


Figure S17: **Non-negative matrix factorization of all observed cell types.** **(A)** UMAP projection of our NMF embedding with rank $k = 20$. **(B)** Proportion of variance in count data explained by NMF embeddings as a function of rank and initialization procedure. A rank of $k = 20$ sits at the elbow of this relationship, capturing $\approx 42\%$ of the variance. **(C)** We visualized the activity level of each gene expression program captured by the NMF embedding across cell types in a UMAP projection. NMF dimensions were not explicitly ordered.

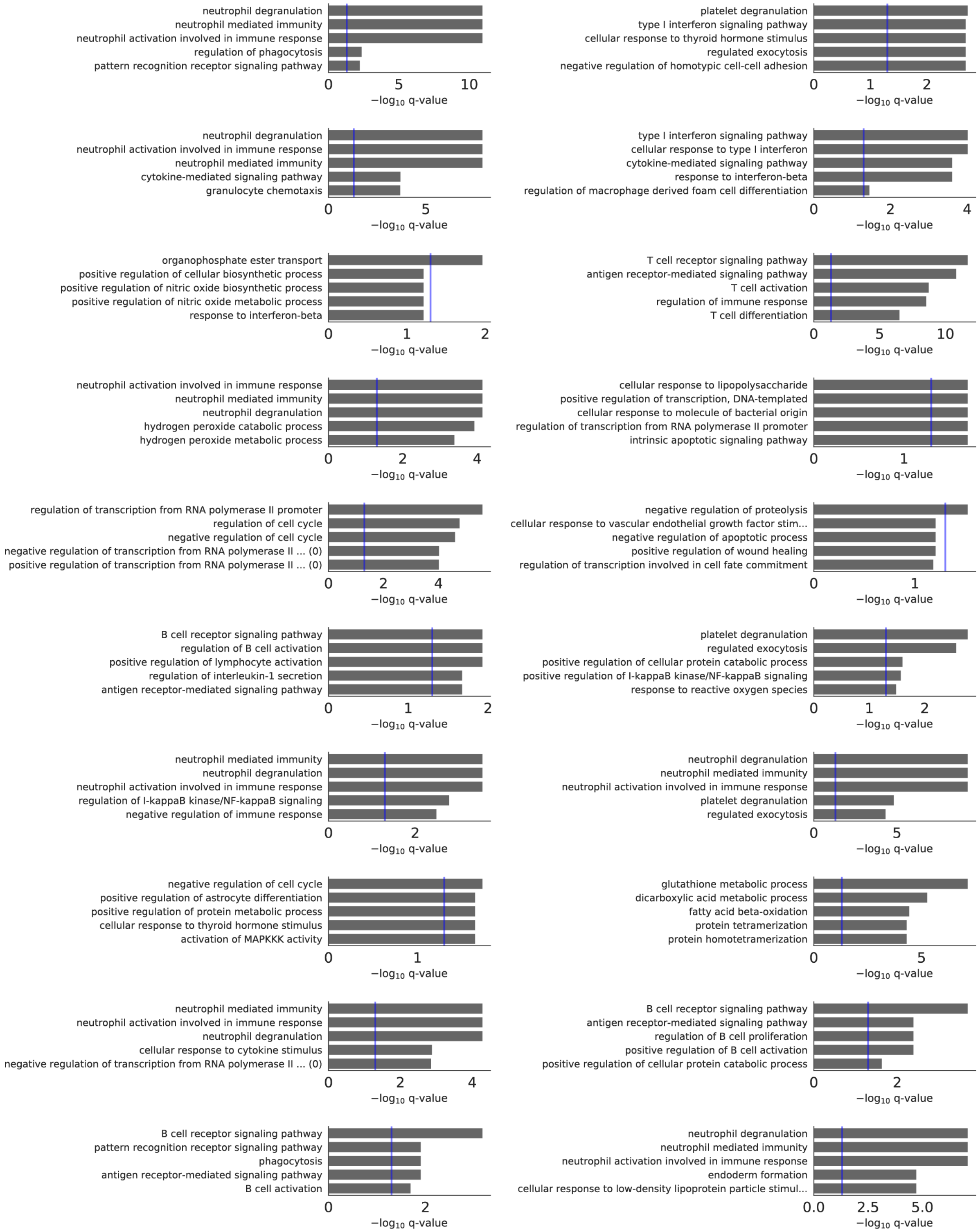


Figure S18: **Gene Ontology enrichment in non-negative matrix factorization embedding dimensions.** Top 10 enriched Gene Ontology Terms for each dimension of the NMF embedding. We performed Gene Ontology enrichment analysis on the genes with loadings above a threshold value for each dimension. Threshold values were chosen using Otsu's method [Otsu, 1975].

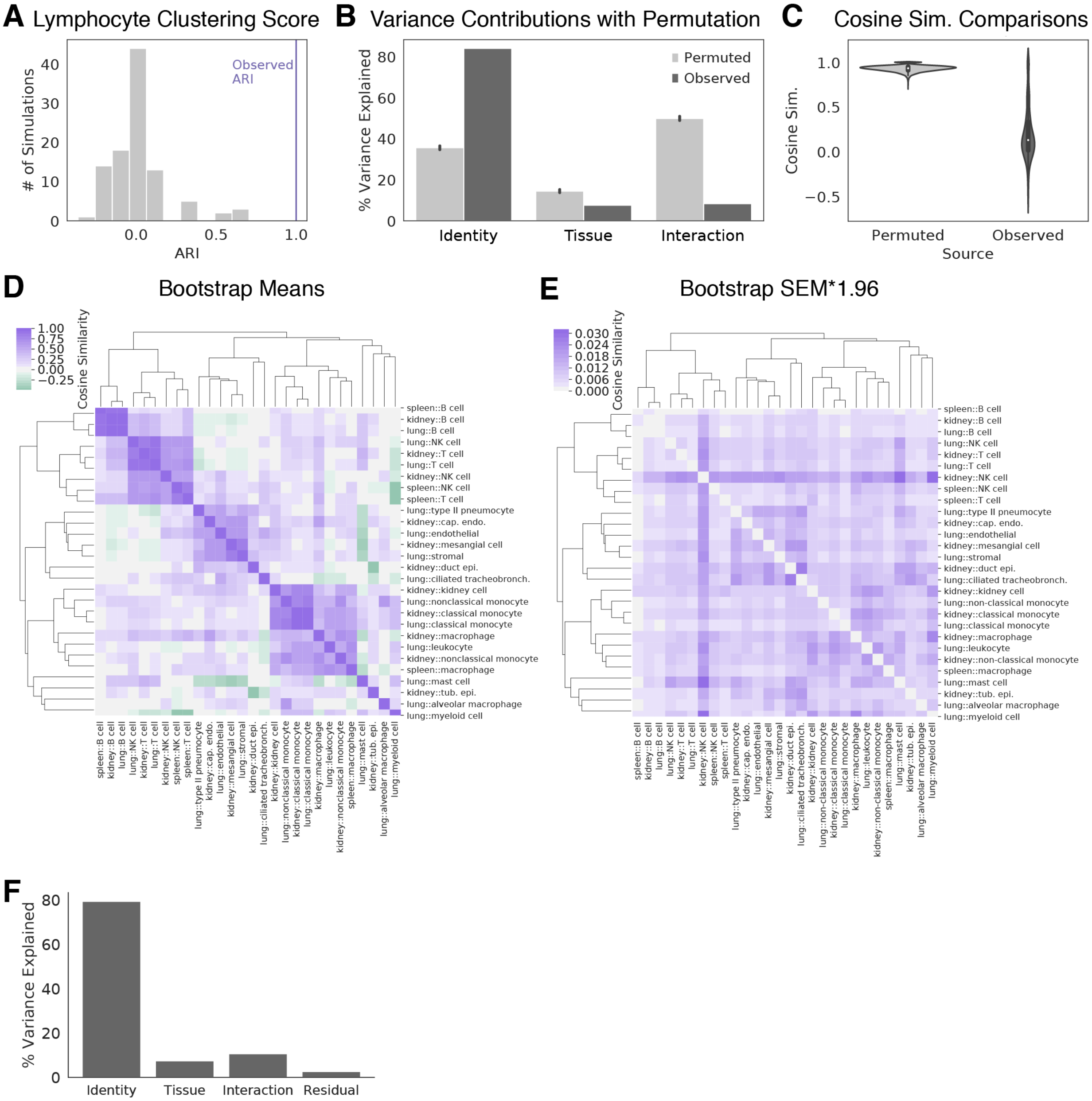


Figure S19: **Permutation and bootstrap evaluation of aging trajectory results.** We performed permutation testing on our aging vector clustering. For each iteration, we randomly permuted tissue::cell type labels, computed aging trajectories in the NMF embedding, and clustered the resulting cosine similarities. We performed this permutation procedure 100 times. **(A)** We evaluated the recovery of lymphocyte cell types in the same cluster using the Adjusted Rand Index for each of these permuted samples, as well as our observed sample. Complete recovery of lymphocytes in a single cluster (ARI $= 1.0$, our observation is marked with the purple line) did not occur in any permuted sample (empirical $p < 0.01$). **(B)** For each permuted sample, we assessed the variance explained by cell identity and tissue labels. Our observed values for the variance explained by cell identity are roughly 2-fold larger than results obtained for permuted samples (empirical $p < 0.01$ comparing variance explained by identity). **(C)** We computed cosine similarities between aging trajectories for each permuted sample. We expected highly similar aging trajectories for each tissue::cell type in the permuted sample, since random assignment of labels results in similar groups of cells within each label. We confirmed this expectation as permuted aging vectors all have very high cosine similarity. The distribution of similarities we observe is significantly different from that of permuted samples (KS-test, $p < 0.001$). **(D)** We performed bootstrap sampling on the observed cells and computed aging trajectories on each of 100 random samples of observed cells. 80% of cells from each type were randomly selected for each iteration. From bootstrap samples, we computed the mean cosine similarity and **(E)** standard error of the mean (SEM) for each 2-way aging vector comparison. Most positive and negative cosine similarities that we observed based on cosine similarity were an order of magnitude larger than the associated standard error. Standard errors were relatively small (< 0.017) for all comparisons, such that interpretation of the comparison between any two aging vectors is robust to sampling noise. **(F)** Variance in the aging vectors of immune cell types computed across bootstrap samples explained by cell type, tissue environment, and their interaction (ANOVA). Bootstrap sampling introduces residual variation into this model.

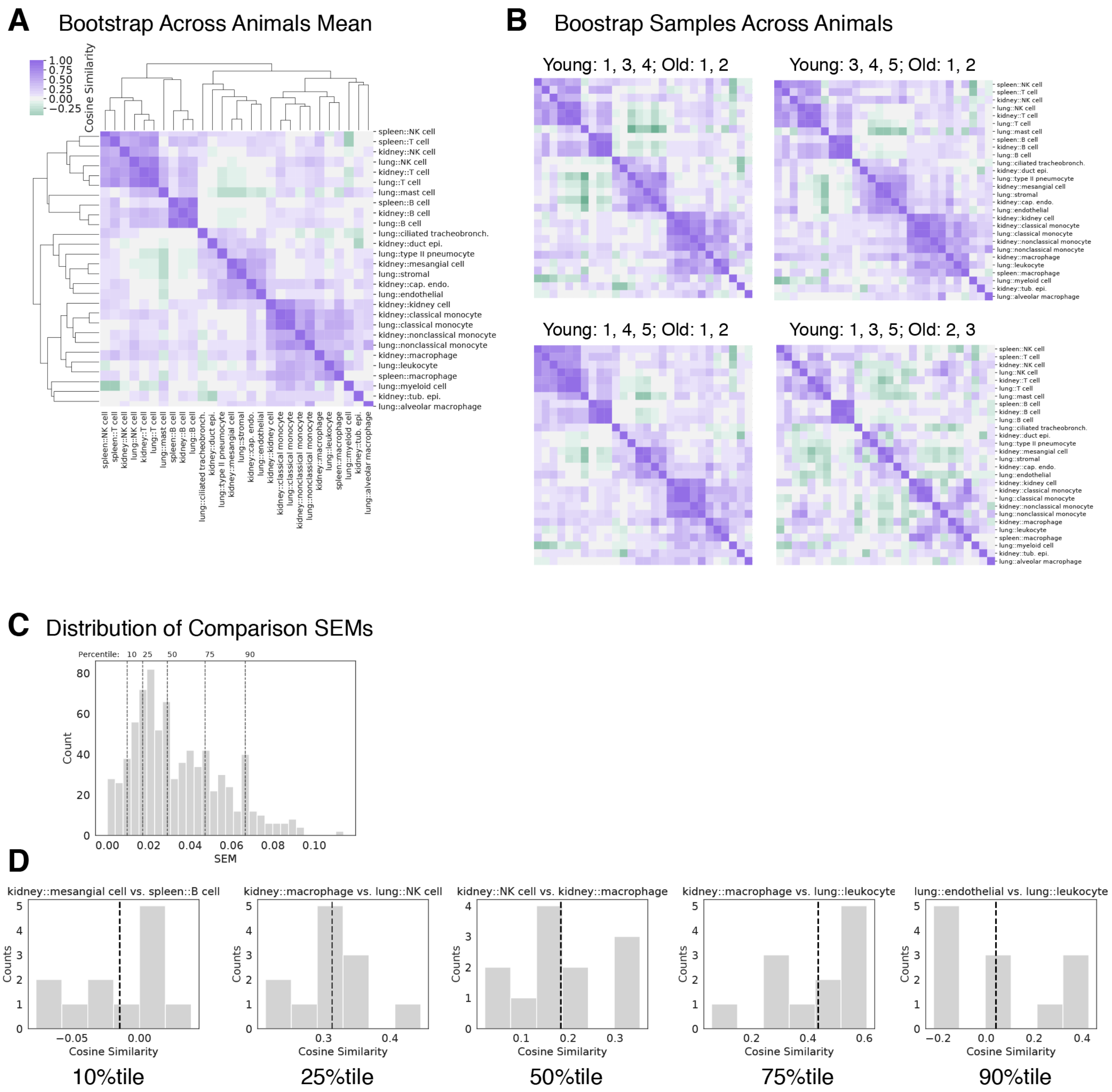


Figure S20: **Aging trajectory comparisons are robust to bootstrap sampling of animals.** We performed bootstrap sampling of cells *by animal of origin* to evaluate robustness of aging trajectory comparisons to resampling noise at the level of animals. This generates conservative confidence intervals. We sampled cells from all possible combinations of animals that leave out one young and one old animal ("leave-one-out sampling"). **(A)** We computed the mean cosine similarity and SEM across animal-wise bootstrap samples. SEMs were larger than than when we randomly sampled cells without regard for the animal of origin, but small relative to the magnitude of many comparisons (< 0.10). **(B)** Four example bootstraps across animals are shown. Labels indicate the animals included in each comparison. Matrices were qualitatively similar across bootstraps. **(C)** The distribution of standard error of the mean (SEM) values across all comparisons showed that SEMs are relatively small. **(D)** Distribution of cosine similarities across bootstrap samples for specific comparisons. Black lines mark the mean across samples. We chose comparisons based on their percentile rank in the SEM distribution. For each distribution, our interpretation of the cosine similarity result did not change across bootstraps. For example, all values of the kidney::macrophage vs. lung::NK cell comparison were positive but small (< 0.5), indicating weak similarity between trajectories. All values of the kidney::mesangial cell vs. spleen::B cell comparison were close to 0, indicating orthogonality.

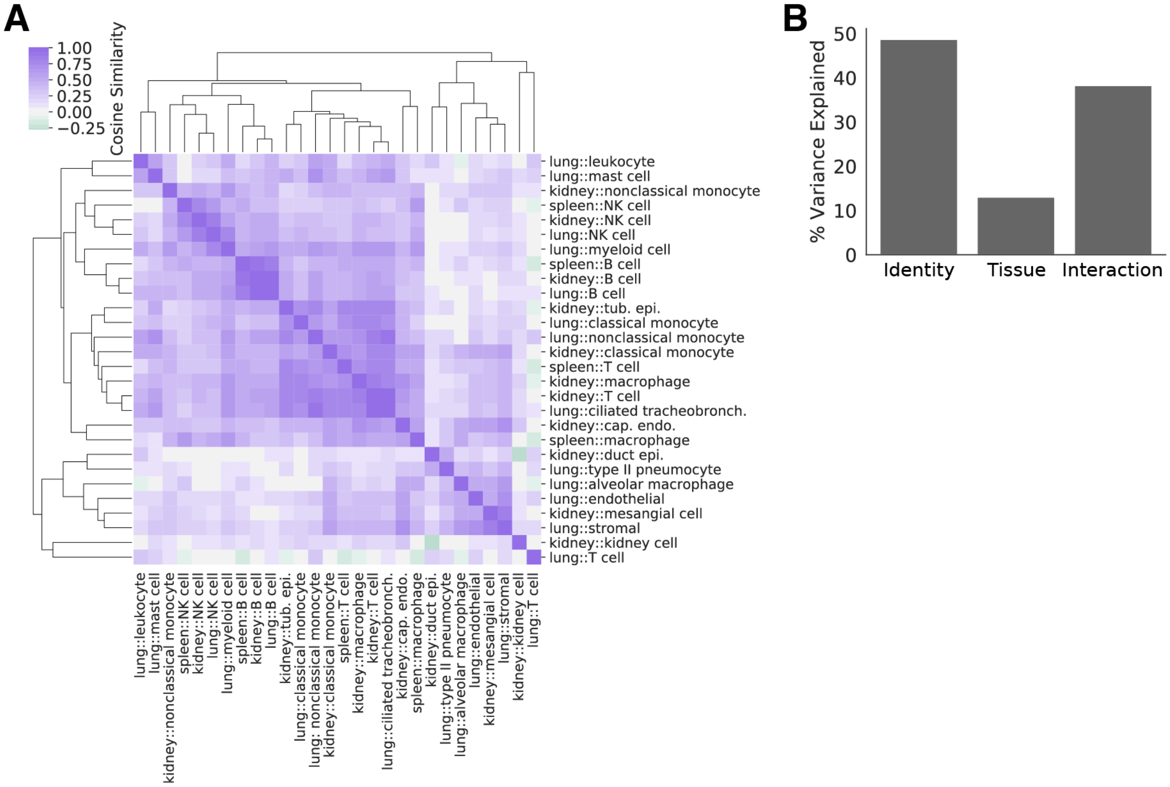


Figure S21: **Cell identity and tissue environment influence on aging trajectories in a PCA embedding. (A)** We computed aging trajectories between young and aged cell centroids in a PCA embedding. We compared these trajectories using cosine similarities. Heatmap values represent cosine similarities between the aging trajectories of each cell state in each tissue. **(B)** Variance in the aging vectors (computed in the PCA embedding) of immune cell types found in all three tissues explained by cell type, tissue environment, and their interaction (ANOVA).

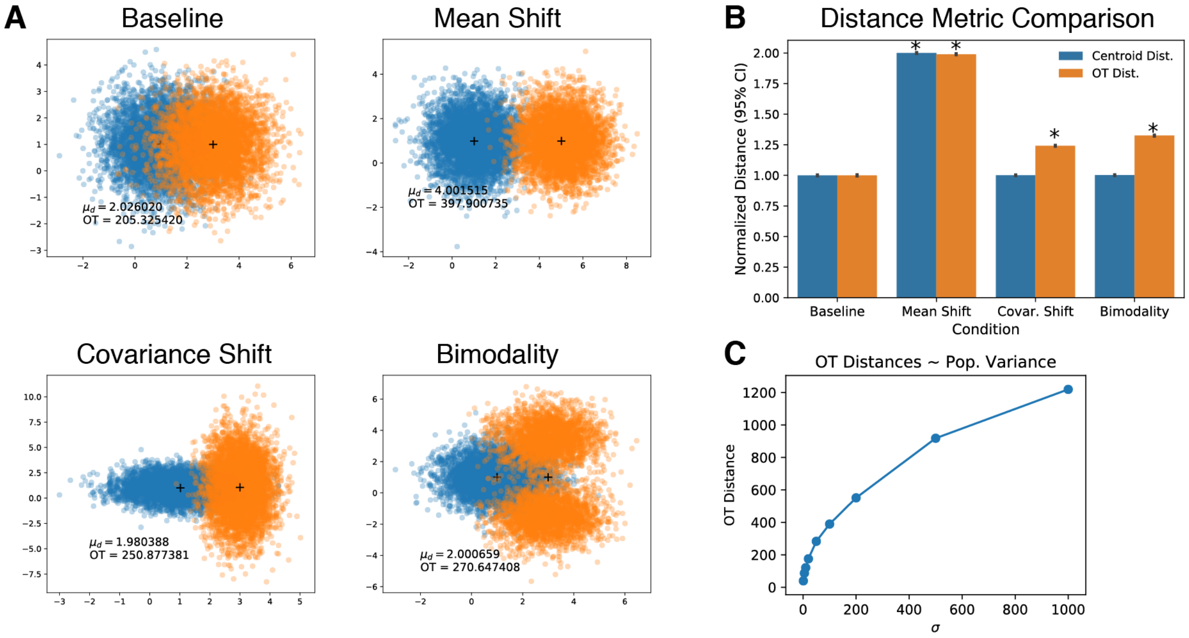


Figure S22: **Simulation experiments demonstrate that optimal transport distances capture differences between cell populations.** **(A)** We simulated two "cell populations" (blue, orange) as 2-dimensional Gaussian distributions. We computed the centroid distance and optimal transport (OT) distance between populations for a range of possible differences that may arise between cell populations (text insets). As a baseline, we simulated unit Gaussians with different means. When the difference in means is increased (Mean Shift), both the centroid distance and OT distance reflect the magnitude of change. However, when we shift the covariance matrix of one population or simulate a bimodal population with the same mean as the baseline unimodal population, only the OT distance reflects these differences. **(B)** Comparison of centroid distance and OT distance metrics for comparing simulated cell populations. Each simulated population contained $n = 5000$ cells. We performed each simulation 50 times to estimate confidence intervals. For each iteration of the simulation, we computed the OT distance as the mean OT distance across 30 random samples of $n = 100$ cells from each simulated population. (*: p-value < 0.05, t-test to baseline). **(C)** OT distances between random samples from the same population scale with population variance. Here, we drew random samples of simulated cells from a 2D Gaussian $\mathcal{N}(0, \sigma I)$ and computed OT distances between samples as in **(B)**. This simulated "isochronic" OT distance increased as we increased σ, demonstrating that isochronic OT distances reflects population heterogeneity.

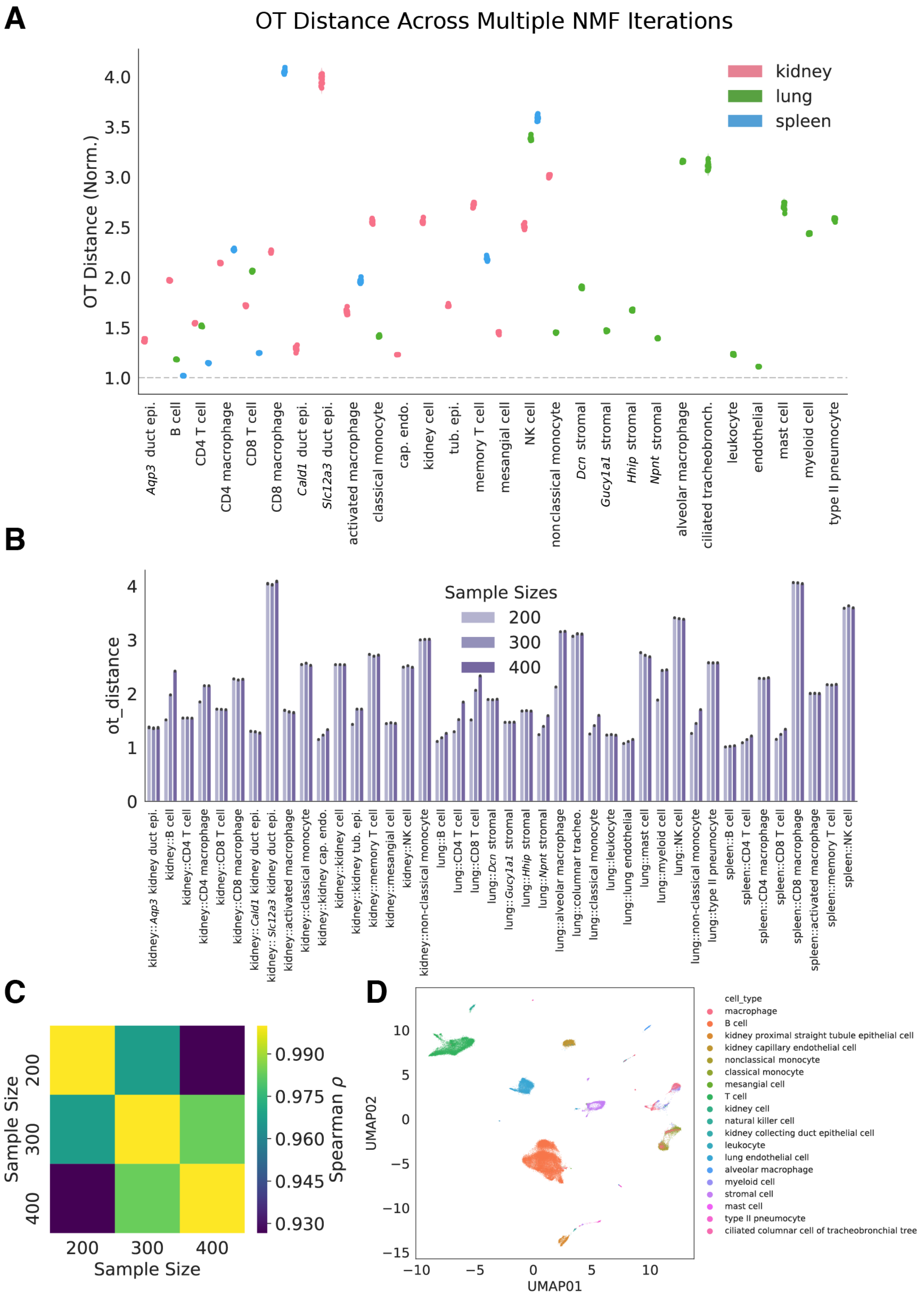


Figure S23: **Optimal transport estimates aging magnitude across cell identities.** **(A)** We computed optimal transport distances for each cell identity and tissue combination. We normalized our heterochronic comparison (Young-Old cells) by the larger mean of two null isochronic comparisons (Young-Young, Old-Old). We computed this normalized distance for 10 separate NMF optimizations. The mean normalized distance for a single NMF optimization is represented as a point, with violins outlining the distribution across 10 iterations. Relative distances between cell identity/tissue combinations were not changed across NMF optimization runs. **(B)** We computed optimal transport distances for a range of random sampling sizes. We present the normalized Old-Young distance values (normalized as in **(A)**) for each sample size. **(C)** Heatmap of Spearman correlations between the normalized Old-Young distances computed using different sample sizes. Old-Young distances have high correlation (Spearman's $\rho > 0.9$) across the range of sample sizes, suggesting the metric is robust to changes in sample size. **(D)** UMAP projection of the NMF (rank 500) embedding used for optimal transport distance calculation. Cell types are overlaid as colors.

Supplemental Tables

Table S1: Total cell counts for each tissue, stratified by age.

Age	Tissue	Cell Count
Old	Kidney	3,042
	Lung	10,460
	Spleen	13,815
Young	Kidney	4,653
	Lung	10,295
	Spleen	13,028

Table S2: Full cell type names corresponding to our abbreviations.

Abbreviation	Full Name
NK cell	natural killer cell
kidney duct epi.	kidney collecting duct epithelial cell
kidney cap. endo.	kidney capillary endothelial cell
kidney loop. epi.	kidney loop of Henle ascending limb epithelial cell
kidney tub. epi.	kidney proximal straight tubule epithelial cell
columnar tracheo.	ciliated columnar cell of tracheobronchial tree
stromal	stromal cell
lung endothelial	lung endothelial cell

References

Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, Rambow F, Marine JC, Geurts P, Aerts J, et al., 2017. SCENIC: single-cell regulatory network inference and clustering. *Nature methods* **14**: 1083–1086.

Altschuler SJ, Wu LF, 2010. Cellular Heterogeneity: Do Differences Make a Difference? *Cell* **141**: 559–563.

Angelidis I, Simon LM, Fernandez IE, Strunz M, Mayr CH, Greiffo FR, Tsitsiridis G, Ansari M, Graf E, Strom TM, et al., 2019. An atlas of the aging lung mapped by single cell transcriptomics and deep tissue proteomics. *Nat Commun* pp. 1–17.

Aoshiba K, Nagai A, 2007. Chronic lung inflammation in aging mice. *FEBS Letters* **581**: 3512–3516.

Aw D, Hilliard L, Nishikawa Y, Cadman ET, Lawrence RA, Palmer DB, 2016. Disorganization of the splenic microanatomy in ageing mice. *Immunology* **148**: 92–101.

Bahar R, Hartmann CH, Rodriguez KA, Denny AD, Busuttil RA, Dollé MET, Calder RB, Chisholm GB, Pollock BH, Klein CA, et al., 2006. Increased cell-to-cell variation in gene expression in ageing mouse heart. *Nature* **441**: 1011–1014.

Bailey KL, Bonasera SJ, Wilderdyke M, Hanisch BW, Pavlik JA, DeVasure J, Robinson JE, Sisson JH, Wyatt TA, 2014. Aging causes a slowing in ciliary beat frequency, mediated by PKCε. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **306**: L584–L589.

Baker DJ, Childs BG, Durik M, Wijers ME, Sieben CJ, Zhong J, Saltness RA, Jeganathan KB, Verzosa GC, Pezeshki A, et al., 2016. Naturally occurring p16(Ink4a)-positive cells shorten healthy lifespan. *Nature* **530**: 184–189.

Barron M, Li J, 2016. Identifying and removing the cell- cycle effect from single-cell RNA- Sequencing data. *Sci Rep* pp. 1–10.

Beaumont HJE, Gallie J, Kost C, Ferguson GC, Rainey PB, 2009. Experimental evolution of bet hedging . *Nature* **461**: 90–93.

Benayoun BA, Pollina EA, Singh PP, Mahmoudi S, Harel I, Casey KM, Dulken BW, Kundaje A, Brunet A, 2019. Remodeling of epigenome and transcriptome landscapes with aging in mice reveals widespread induction of inflammatory responses. *Genome Research* **29**: 697–709.

Benjamini Y, Hochberg Y, 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **57**: 289–300.

Blake WJ, Balázsi G, Kohanski MA, Isaacs FJ, Murphy KF, Kuang Y, Cantor CR, Walt DR, Collins JJ, 2006. Phenotypic consequences of promoter-mediated transcriptional noise. *Molecular Cell* **24**: 853–865.

Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E, 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* **2008**: P10008.

Bolignano D, Mattace-Raso F, Sijbrands EJG, Zoccali C, 2014. The aging kidney revisited: A systematic review. *Ageing Res Rev* **14**: 65–80.

Buettner F, Natarajan KN, Casale FP, Proserpio V, Scialdone A, Theis FJ, Teichmann SA, Marioni JC, Stegle O, 2015. Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nature Biotechnology* **33**: 155–160.

Chakkalakal JV, Jones KM, Basson MA, Brack AS, 2012. The aged niche disrupts muscle stem cell quiescence. *Nature* **490**: 355–360.

Childs BG, Gluscevic M, Baker DJ, Laberge RM, Marquess D, Dananberg J, van Deursen JM, 2017. Senescent cells: an emerging target for diseases of ageing. *Nat Rev Drug Discov* **16**: 718–735.

Cohen D, 1966. Optimizing reproduction in a randomly varying environment. *J Theor Biol* **12**: 119–129.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR, 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.

Duchi J, Hazan E, Singer Y, 2011. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. *Journal of Machine Learning Research* **12**: 2121–2159.

Enge M, Arda HE, Mignardi M, Beausang J, Bottino R, Kim SK, Quake SR, 2017. Single-Cell Analysis of Human Pancreas Reveals Transcriptional Signatures of Aging and Somatic Mutation Patterns. *Cell* **171**: 321–323.e14.

Han X, Wang R, Zhou Y, Fei L, Sun H, Lai S, Saadatpour A, Zhou Z, Chen H, Ye F, et al., 2018. Mapping the Mouse Cell Atlas by Microwell-Seq. *Cell* **173**: 1307.

Kamijo T, Zindy F, Roussel MF, Quelle DE, Downing JR, Ashmun RA, Grosveld G, Sherr CJ, 1997. Tumor suppression at the mouse INK4a locus mediated by the alternative reading frame product p19ARF. *Cell* **91**: 649–659.

Kiselev VY, Yiu A, Hemberg M, 2018. scmap: projection of single-cell RNA-seq data across data sets. *Nature methods* **15**: 359–362.

Kolodziejczyk AA, Kim JK, Tsang JCH, Ilicic T, Henriksson J, Natarajan KN, Tuck AC, Gao X, Bühler M, Liu P, et al., 2015. Single Cell RNA-Sequencing of Pluripotent States Unlocks Modular Transcriptional Variation. *Cell Stem Cell* **17**: 471–485.

Kovacs E, Lowery E, Kuhlmann E, Brubaker A, 2013. The aging lung. *Clinical Interventions in Aging* **8**: 1489–8.

Kovacs EJ, Boe DM, Boule LA, Curtis BJ, 2017. Inflammaging and the Lung. *Clinics in Geriatric Medicine* **33**: 459–471.

Kowalczyk MS, Tirosh I, Heckl D, Rao TN, Dixit A, Haas BJ, Schneider RK, Wagers AJ, Ebert BL, Regev A, 2015. Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. *Genome Research* **25**: 1860–1872.

Kussell E, Kishony R, Balaban NQ, Leibler S, 2005. Bacterial Persistence: A Model of Survival in Changing Environments. *Genetics* **169**: 1807–1814.

Lumeng CN, Liu J, Geletka L, Delaney C, Delproposto J, Desai A, Oatmen K, Martinez-Santibanez G, Julius A, Garg S, et al., 2011. Aging is associated with an increase in T cells and inflammatory macrophages in visceral adipose tissue. *Journal of immunology (Baltimore, Md. : 1950)* **187**: 6208–6216.

Martinez-Jimenez CP, Eling N, Chen HC, 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 (New York, N.Y.)* **355**: 1433–1436.

Nalapareddy K, Nattamai KJ, Kumar RS, Karns R, Wikenheiser-Brokamp KA, Sampson LL, Mahe MM, Sundaram N, Yacyshyn MB, Yacyshyn B, et al., 2017. Canonical Wnt Signaling Ameliorates Aging of Intestinal Stem Cells. *Cell Reports* **18**: 2608–2621.

Navarro S, Driscoll B, 2017. Regeneration of the Aging Lung: A Mini-Review. *Gerontology* **63**: 270–280.

O'Brown ZK, Van Nostrand EL, Higgins JP, Kim SK, 2015. The Inflammatory Transcription Factors NFκB, STAT1 and STAT3 Drive Age-Associated Transcriptional Changes in the Human Kidney. *PLoS Genet* **11**: e1005734–28.

O'Sullivan ED, Hughes J, Ferenbach DA, 2017. Renal Aging: Causes and Consequences. *Journal of the American Society of Nephrology : JASN* **28**: 407–420.

Otsu N, 1975. A threshold selection method from gray-level histograms. *Automatica* .

Paxson JA, Gruntman A, Parkin CD, Mazan MR, Davis A, Ingenito EP, Hoffman AM, 2011. Age-Dependent Decline in Mouse Lung Regeneration with Loss of Lung Fibroblast Clonogenicity and Increased Myofibroblastic Differentiation. *PLoS One* **6**: e23232–19.

Quinn KM, Fox A, Harland KL, Russ BE, Li J, Nguyen THO, Loh L, Olshanksy M, Naeem H, Tsyganov K, et al., 2018. Age-Related Decline in Primary CD8+ T Cell Responses Is Associated with the Development of Senescence in Virtual Memory CD8+ T Cells. *Cell Reports* **23**: 3512–3524.

Quinn KM, Zaloumis SG, Cukalac T, Kan WT, Sng XYX, Mirams M, Watson KA, McCaw JM, Doherty PC, Thomas PG, et al., 2016. Heightened self-reactivity associated with selective survival, but not expansion, of naïve virus-specific CD8 +T cells in aged mice. *Proc. Nat. Acad. Sci.* **113**: 1333–1338.

Robinson DG, Wang JY, Storey JD, 2015. A nested parallel experiment demonstrates differences in intensity-dependence between RNA-seq and microarrays. *Nucleic Acids Research* **4**: gkv636–8.

Rodwell GEJ, Sonu R, Zahn JM, Lund J, Wilhelmy J, Wang L, Xiao W, Mindrinos M, Crane E, Segal E, et al., 2004. A Transcriptional Profile of Aging in the Human Kidney. *PLOS Biology* **2**: e427–11.

Satija R, Farrell JA, Gennert D, Schier AF, Regev A, 2015. Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology* **33**: 495–502.

Slatkin M, 1974. Hedging one's evolutionary bets. *Nature* **250**: 704–705.

Smith LK, He Y, Park JS, Bieri G, Snethlage CE, Lin K, Gontier G, Wabl R, Plambeck KE, Udeochu J, et al., 2015. Beta 2-microglobulin is a systemic pro-aging factor that impairs cognitive function and neurogenesis. *Nat Med* pp. 1–8.

Sueel GM, Kulkarni RP, Dworkin J, Garcia-Ojalvo J, Elowitz MB, 2007. Tunability and noise dependence in differentiation dynamics. *Science* **315**: 1716–1719.

Takasaki C, Miura E, Watanabe M, 2007. Segmental and complementary expression of L-serine biosynthetic enzyme 3-phosphoglycerate dehydrogenase and neutral amino acid transporter ASCT1 in the mouse kidney. *Biomedical research (Tokyo, Japan)* **28**: 61–69.

The Tabula Muris Consortium, 2018. Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature* pp. 1–25.

Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, Rotem A, Rodman C, Lian C, Murphy G, et al., 2016. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science (New York, N.Y.)* **352**: 189–196.

Toapanta FR, Ross TM, 2009. Impaired immune responses in the lungs of aged mice following influenza infection. *Respiratory Research* **10**: 340–19.

Todhunter ME, Sayaman RW, Miyano M, LaBarge MA, 2018. Tissue aging: the integration of collective and variant responses of cells to entropic forces over time. *Curr Opin Cell Biol* **54**: 121–129.

Turner VM, Mabbott NA, 2017. Influence of ageing on the microarchitecture of the spleen and lymph nodes. *Biogerontology* **18**: 723–738.

Vallejos CA, Marioni JC, Richardson S, 2015. BASiCS: Bayesian Analysis of Single-Cell Sequencing Data. *PLoS computational biology* **11**: e1004333–18.