# Transcriptomic characterization of 20 organs and tissues from mouse at single cell resolution creates a *Tabula Muris*

The *Tabula Muris* Consortium

**We have created a compendium of single cell transcriptome data from the model organism *Mus musculus* comprising nearly 100,000 cells from 20 organs and tissues. These data reveal many new aspects of cell biology, including gene expression in poorly characterized cell populations and the characterization of new populations of cells in many tissues. Furthermore, they allow for direct and controlled comparison of gene expression in cell types shared between tissues, such as immune cells from distinct anatomical locations. Two distinct technical approaches were used for most organs: one approach, microfluidic droplet-based 3'-end counting[1,2], enabled the survey of thousands of cells at relatively low coverage, while the other, FACS-based full length transcript analysis[3], enabled characterization of cell types with high sensitivity and coverage. The cumulative resource provides the foundation for a comprehensive resource of transcriptomic cell biology.**

The cell is a fundamental unit of structure and function in biology, and multicellular organisms have evolved a wide variety of different cell types with specialized roles. Although cell types have historically been characterized on the basis of morphology and phenotype, the development of molecular methods has enabled ever more precise definition of their properties, typically by measuring protein or mRNA expression patterns[4]. Technological advances have enabled increasingly greater degrees of multiplexing of these measurements[5,6], and it is now possible to use highly parallel sequencing to enumerate nearly every mRNA molecule in a given single cell[3,7]. This approach has provided many novel insights into cell biology and the composition of organs and tissues from a variety of organisms[8–17]. However, while these reports provide valuable characterization of individual organs, it is challenging to compare data taken with varying experimental techniques in independent labs from different animals. It therefore remains an open question whether data from individual organs can be synthesized and used as a more general resource for biology.

Here we report a compendium of cell types from the mouse *Mus musculus*. The compendium is comprised of single cell transcriptome sequence data from 97,029 cells isolated from 20 organs and tissues (Fig. 1). These organs were collected from 3 female and 4 male, C57BL/6 NIA, 3 month old mice (10-15 weeks), whose developmental age is roughly analogous to humans at 20 years of age. We analyzed multiple organs and tissues from the same animal, thereby generating a data set controlled for age, environment and epigenetic effects. This enables the direct comparison of cell type composition between organs as well as comparison of shared cell types across the entire organism. All data, protocols, and analysis scripts from the compendium are shared as a public resource. Gene counts and metadata from all single cells are accessible via Digital Object Identifier (DOI) at 10.6084/m9.figshare.5715040 for the FACS data and 10.6084/m9.figshare.5715025 for the microfluidic droplet data. While these data are by no means a complete representation of all mouse organs and cell types, they provide a

first draft attempt to create an organism-wide representation of cellular diversity and a comparative framework for future studies using the large variety of murine disease models.

We developed a specific procedure to collect 20 organs and tissues from the same mouse (see Methods). Briefly, each mouse was anesthetized with 2.5% v/v Avertin, followed by transcardial perfusion with 20 ml PBS. Aorta, bladder, bone marrow, brain (cerebellum, cortex, hippocampus, striatum), colon, diaphragm, fat (brown, gonadal, mesenteric, subcutaneous), heart, kidney, liver, lung, mammary gland, pancreas, skin, skeletal muscle, spleen, thymus, tongue, and trachea were immediately dissected and processed into single cell suspensions (see Extended Data). Organs were dissociated into single cells and either sorted by FACS into 384-well plates (with the exception of liver hepatocytes that were sorted into 96 well plates and heart derived cardiomyocytes that were hand-picked into 96-well plates), and in many cases also loaded into a microfluidic droplet emulsion generating device for cell isolation followed by transcriptome capture (Fig. 1a). Once prepared into sequencing libraries, transcriptomes were sequenced to an average depth of 685,500 reads per cell for the plate data and 7,709 unique molecular identifiers (UMI) per cell for the microfluidic emulsion data. After quality control filtering, 44,879 FACS sorted cells and 55,656 microfluidic droplet processed cells were retained for further analysis. A comparison of the two methods showed differences for each organ in the number of cells analyzed (Fig. 1b,c), reads per cell (Fig. 1d,f) and genes per cell (Fig. 1e,g).

Using two distinct measurement approaches on the same samples yielded several insights. We compared the methods to understand their relative strengths and weaknesses. Importantly, our results here show that such comparisons can be misleading when performed on a single organ type, as there is substantial variation in performance across organs (Supp. Fig. 1). For example, the lung, trachea and thymus showed nearly twice as many genes detected per cell in the FACS data as compared to the microfluidic data, whereas heart, kidney and marrow show comparable numbers of expressed genes detected by both methods (Supp. Fig. 1-3). Number of genes detected is a fairly crude approach to sensitivity and one must also consider other metrics such as dynamic range for gene expression. The FACS based approach has much higher dynamic range per gene than the microfluidic droplets (Supp. Fig. 4) and enables the analysis of full length transcripts. While the FACS approach generally resulted in higher sensitivity and dynamic range, it also is more laborious and time consuming to perform. The microfluidic droplet approach is faster and enables analysis of larger numbers of cells, albeit at generally lower sensitivity and dynamic range, and with reads only from the 3' end of the transcript. We also expect this study to facilitate the development of computational tools to compare sequencing methods, and our data set will provide a training and validation set for such algorithms. The creation of these tools will crucial for comparing independent data generated by labs around the world as various tissue atlas studies begin generating results.

We performed unbiased graph-based clustering of the pooled set of transcriptomes across all organs, and visualized them using tSNE (Fig. 2a). The majority of clusters contain

cells from only one organ, but a substantial minority contained cells from many organs. To further dissect these clusters, we separately performed dimensional reduction and clustering on cellular transcriptomes from each individual organ (Fig. 2b). The resulting clusters were manually annotated using *a priori* biological knowledge about cell-type specific gene expression, which confirmed the presence of many specific cell types within organs (Fig. 3). Many of these cell types have not previously been obtained in pure populations and our data provide a wealth of new information on their characteristic gene expression profiles. Initial annotation of each organ and tissue can be found in the extended data, and a detailed discussion each cell types on an organ by organ basis can be found in the supplement. Some unexpected discoveries include: 1) novel immune cell types in the lung, 2) a putative neuroendocrine cell population in the trachea, 3) new differentiated keratinocyte cell types of the tongue, 4) strong sexual dimorphism in hepatocytes with differential gene expression of male pheromones, 5) the suggestions of new roles for genes such as *Neurog3*, *Hhex*, and *Prss53* in the adult pancreas, 6) sexual dimorphism in epithelial and mesenchymal cells of the bladder, and 7) a novel cell population expressing *Chodl* in skeletal muscle.

In order to better understand the relationships between cell types across organs, we mapped the biologically informed annotations of organ-specific cell types onto the clusters resulting from the unbiased analysis of the complete set of all cell transcriptomes. This analysis revealed that the clusters with cells from many organs generally represent shared cell types common to those organs (Figure 3). For example, B cells from fat, limb muscle, diaphragm, lung, spleen and marrow cluster together, as do T cells from spleen, marrow, lung, limb muscle, fat and thymus. Endothelial cells from fat, heart and lung cluster together, but form a distinct grouping from endothelial cells from mammary, kidney, trachea, limb muscle, aorta, diaphragm and pancreas. There are two clusters containing myeloid cells from limb muscle, brain, diaphragm, aorta, pancreas, kidney, trachea, heart, liver, fat and marrow. Lastly, a cluster containing mesenchymal stem cells from fat, diaphragm, and limb muscle suggests these cells are share similar transcriptomes with stromal cells in mammary, trachea, and lung. These findings show that gene-specific manual annotation of cell types is consistent with unbiased clustering based on whole transcriptome profiles, and that cell type identity is strong enough across tissues to enable their unbiased identification at the whole transcriptome level.

To further investigate a common cell type across organs, we pooled all T-cells and performed an unbiased clustering analysis, revealing 5 clusters (Figure 4) grouped into two distinct sub-groups: clusters 0,4 and clusters 1,2,3. Cluster 0 comprises cells from the Thymus that are undergoing VDJ recombination characterized by the expression of RAG (*Rag1* and *Rag2*) and TdT (*Dntt*), and includes uncommitted double positive T-cells (*Cd4* and *Cd8a* positive). Cluster 4 contains proliferating T-cells, predominantly from the thymus, and we hypothesize that these are pre-T cells that are expanding after having completed VDJ recombination. In contrast, Clusters 1,2,3 contain mature T-cells. The cells in Cluster 3 are also predominantly from the thymus and show high *Cd5* expression, suggesting that they are undergoing positive selection. Cluster 2 cells are mostly peripheral cells and are most likely activated T-cells expressing the high affinity IL2 receptor (*Il2ra* and *Il2rb)*. Interestingly, they also express MHC type II genes (*H2-Aa*

and *H2-Ab1*). Finally, Cluster 1 also represents mature T-cells, but primarily from the spleen.

A key challenge for many single cell studies to date is understanding the potential changes to the transcriptome caused by handling, dissociation and other experimental manipulation. A previous study in limb muscle showed that quiescent satellite cells tend to become activated by dissociation and consequently express immediate early genes among other genes.[18] We found that expression of these dissociation-related markers was also clearly observed in our limb muscle data, as well as in mammary and bladder (Supp. Fig. 6), but that many organs showed little evidence of cellular activation with this panel. Therefore the dissociation-related activation markers found in limb muscle are not universal across all organs. This is not to say that other tissues lack dissociation-related gene expression changes, but that some of the genes involved are specific to a given tissue. It is important to note that the presence of such gene expression changes does not prevent the identification of cell type or the comparison of cell types across organs.

To understand the extent to which transcription factor (TF) expression determines cell type, we performed a correlation analysis of TF expression[14] across the entire data set for TFs that were significantly enriched ($p < 10^{-5}$, log10 fold change > 0.3) in at least one unique combination of cell type and tissue (e.g., endothelial cells from liver or basal cell of the bladder). (Fig 5A). The off-diagonal elements of the correlation matrix indicate groups of transcription factors with correlated cell type expression. We measured co-expression of these groups of transcription factors and discovered that typically only one of a few cell types used any given combination of two to four TFs. (Supp. Fig. 7) The largest cluster of correlated genes enriched for a broad swath of immune cells across most tissues. Similarly, another cluster of genes enriched for endothelial cells. We also observed groups of TFs specific to several organs; for example within colon, goblet cells were highly enriched for a cluster of TFs that are either known or potential new targets for specification of goblet cells (Fig 5A, Supplementary Figure 7), and we also saw clusters specific to muscle cells, pancreatic cells, neurons, microglia and astrocytes. Another group of genes showed enrichment in several epithelial subtypes of tongue and bladder.

We further analyzed tissue-specific TFs within the endothelial, epithelial and stromal cell types found in several different organs (Fig 5B-D). Within epithelial cells, we found genes grouped by tissue: tongue (Pax9), bladder (Ets1), skin keratinocyte stem cells (Lhx2), and lung (Aebp1) (Fig 5B). Pax9 is enriched in tongue basal cells and is known to be necessary for formation of filiform papillae. Foxq1 is enriched in bladder basal cells, but its role in bladder basal cell differentiation is heretofore unexplored. Within fibroblast/stromal cells, we found TFs that separated heart from kidney fibroblasts, and specified the stromal cells of lung, trachea, or mammary gland (Fig 5C). Within endothelial cells, we found less strong signatures for tissue specificity; however, liver, brain, kidney, and lung-specific groups of TFs were distinguishable (Fig 5D). These results illustrate how single cell data taken across many organs and tissues can be used in

combination to identify which transcriptional regulatory programs are specific to cell types of interest.

In conclusion, we have created a compendium of single-cell transcriptional measurements across 20 organs and tissues of the mouse. This *Tabula Muris*, or "Mouse Atlas", has many uses, including the discovery of new putative cell types, the discovery of novel gene expression in known cell types, and the ability to compare cell types across organs. It will also serve as a reference of healthy young adult organs which can be used as a baseline for current and future mouse models of disease. While it is not an exhaustive characterization of all organs of the mouse, it does provide a rich data set of the most highly studied tissues in biology. The *Tabula Muris* provides a framework and description of many of the most populous and important cell populations within the mouse, and represents a foundation for future studies across a multitude of diverse physiological disciplines.

**Acknowledgements**

This work was funded by the Chan Zuckerberg Biohub, the National Institute on Aging (DP1AG053015, T.W.-C), the Veterans Administration (T.W.-C.), and the Stanford Diabetes Research Center (NIH DK116074, S.K.K.). We thank Sony Biotechnology for making an SH800S instrument available for this project. Some cell sorting/flow cytometry analysis for this project was done on a Sony SH800S instrument in the Stanford Shared FACS Facility. Some fluorescence activated cell sorting (FACS) was done with instruments in the VA Flow Cytometry Core, which is supported by the US Department of Veterans Affairs (VA), Palo Alto Veterans Institute for Research (PAVIR), and the National Institutes of Health (NIH).

**The *Tabula Muris* Consortium:**

**Overall Coordination:** Nicholas Schaum[b], Jim Karkanias[r], Norma F Neff[r], Andrew P. May[r], Stephen R. Quake[o,r]*, Tony Wyss-Coray[a,c,d]*, and Spyros Darmanis[r]*

* to whom correspondence should be addressed: quake@stanford.edu, twc@stanford.edu, spyros.darmanis@czbiohub.org

**Tissue collection and processing:** Ankit S. Baghel[b], Isaac Bakerman[b,h,i], Ishita Bansai[r], Daniela Berdnik[a], Biter Bilen[a], Douglas Brownfield[l], Corey Cain[s], Michelle B. Chen[o], Steven Chen[r] , Min Cho[r], Giana Cirolia[r], Stephanie G. Conley[b], Spyros Darmanis[r], Aaron Demers[r], Kubilay Demir[b,n], Tessa Divita[r], Haley du Bois[a], Laughing Bear Torrez Dulgeroff[b], Hamid Ebadi[r], Hernan Espinoza[l], Matt Fish[b,m,n], Benson M. George[b], Astrid Gillich[l], Foad Green[r], Geraldine Genetiano[r], Xueying Gu[m], Gunsagar S. Gulati[b], Yan Hang[m], Shayan Hosseinzadeh[r], Tal Iram[a], Taichi Isobe[b], Feather Ives[r] , Robert Jones[o], Kevin S. Kao[b], Guruswamy Karnam[q], Aaron M. Kershner[b], Bernhard Kiss[b,x], William Kong[b], Jonathan Lam[m], Davis P. Lee[d], Song E. Lee[a], Benoit Lehallier[a], Guang Li[j], Qingyun Li[w], Ling Liu[a,c], Annie Lo[r], Wan-Jin Lu[b,l], Anoop Manjunath[b], Andrew P. May[r],

Kaia L. May[r], Oliver L. May[r], Ashley Maynard[r], Marina McKay[r], Marco Mignardi[o], Dullei Min[k], Antoine de Morree[a], Ahmad N. Nabhan[l], Norma F Neff[r], Katharine M. Ng[o], Joseph Noh[b], Rasika Patkar[q], Weng Chuan Peng[m], Lolita Penland[r], Robert Puccinelli[r], Eric J. Rulifson[m], Nicholas Schaum[b], Shaheen S. Sikandar[b], Rahul Sinha[b,e,f,g], Rene V Sit[r], Krzysztof Szade[b,y], Weilun Tan[r], Cristina Tato[r], Krissie Tellez[m], Kyle J. Travaglini[l], Carolina Tropini[p], Lucas Waldburger[r], Linda J. van Weele[b], Michael N. Wosczyna[a], Jinyi Xiang[b], Soso Xue[o], Justin Youngyunpipatkul[r], Hanadie Yousef[a], Fabio Zanini[o], Macy E. Zardeneta[d], Lu Zhou[w]

**Library preparation and sequencing:** Ishita Bansai[r], Steven Chen[r], Min Cho[r], Giana Cirolia[r], Spyros Darmanis[r], Aaron Demers[r], Tessa Divita[r], Hamid Ebadi[r], Geraldine Genetiano[r], Foad Green[r], Shayan Hosseinzadeh[r], Feather Ives[r], Annie Lo[r], Andrew P. May[r], Ashley Maynard[r], Marina McKay[r], Norma F Neff[r], Lolita Penland[r], Rene V Sit[r], Weilun Tan[r], Lucas Waldburger[r], Justin Youngyunpipatkul[r]

**Computational Data Analysis:** Joshua Batson[r], Olga Botvinnik[r], Paola Castro[r], Spyros Darmanis[r], Joseph L. DeRisi[r,t], Jim Karkanias[r], Geoffrey M. Stanley[o], James T. Webber[r]

**Cell Type Annotation:** Ankit S. Baghel[b], Isaac Bakerman[b,h,i], Ben A. Barres[w], Joshua Batson[r], Phil Beachy[b,l,m,n], Biter Bilen[a], Olga Botvinnik[r], Douglas Brownfield[l], Charles K.F. Chan[v], Michelle B. Chen[o], Michael F. Clarke[b], Spyros Darmanis[r], Kubilay Demir[b,n], Hamid Ebadi[r], Hernan Espinoza[l], Matt Fish[b,m,n], Benson M. George[b], Astrid Gillich[l], Xueying Gu[m], Gunsagar S. Gulati[b], Yan Hang[m], Kerwyn Casey Huang[o,p,r], Tal Iram[a], Taichi Isobe[b], Guruswamy Karnam[q], Aaron M. Kershner[b], Seung K. Kim[m,u], Bernhard M. Kiss[b,x], William Kong[b], Mark A. Krasnow[l,n], Christin S. Kuo[k,l,n], Jonathan Lam[m], Benoit Lehallier[a], Guang Li[j], Qingyun Li[w], Ling Liu[a,c], Wan-Jin Lu[b,l], Dullei Min[k], Antoine de Morree[a], Ahmad N. Nabhan[l], Katharine M. Ng[o], Patricia K. Nguyen[b,h,i,j], Roel Nusse[b,m,n], Rasika Patkar[q], Weng Chuan Peng[m], Lolita Penland[r],Thomas A. Rando[a,c,d], Eric J. Rulifson[m], Nicholas Schaum[b], Shaheen S. Sikandar[b], Rahul Sinha[b,e,f,g], Krzysztof Szade[b,z], Krissie Tellez[m], Kyle J. Travaglini[l], Carolina Tropini[p], Linda J. van Weele[b], Bruce M. Wang[q], Kenneth Weinberg[k], Irving L. Weissman[b,e,f,g], Michael N. Wosczyna[a], Sean Wu[b,h,j], Tony Wyss-Coray[a,c,d] , Jinyi Xiang[b], Hanadie Yousef[a], Macy E. Zardeneta[d], Lu Zhou[w]

**Writing Group:** Joshua Batson[r], Olga Botvinnik[r], Steven Chen[r], Spyros Darmanis[r], Foad Green[r], Andrew P. May[r], Ashley Maynard[r], Stephen R. Quake[o,r], Nicholas Schaum[r], James T. Webber[r], Tony Wyss-Coray[a,c,d]

**Principal Investigators:** Ben A. Barres[w], Phil Beachy[b,l,m,n], Charles Chan[b], Michael F. Clarke[b], Spyros Darmanis[r], Kerwyn Casey Huang[o,p,r], Jim Karkanias[r], Seung K. Kim[m,u], Mark A. Krasnow[l,n], Christin S. Kuo[k,l,n], Andrew P. May[r], Norma Neff[r], Roel Nusse[b,m,n], Patricia K. Nguyen[b,h,i,j], Thomas A. Rando[a,c,d], Bruce M. Wang[q], Kenneth Weinberg[k], Irving L. Weissman[b,e,f,g], Sean Wu[b,h,j], Stephen R. Quake[o,r], Tony Wyss-Coray[a,c,d]

[a]Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, California, USA

[b]Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, California, USA
[c]Paul F. Glenn Center for the Biology of Aging, Stanford University School of Medicine, Stanford, California, USA
[d]Center for Tissue Regeneration, Repair, and Restoration, V.A. Palo Alto Healthcare System, Palo Alto, California, USA
[e]Department of Pathology, Stanford University School of Medicine, Stanford, California, USA
[f]Ludwig Center for Cancer Stem Cell Research and Medicine, Stanford University School of Medicine, Stanford, California, USA
[g]Stanford Cancer Institute, Stanford University School of Medicine, Stanford, California, USA.
[h]Stanford Cardiovascular Institute, Stanford University School of Medicine, Stanford, California, USA
[i]Department of Medicine, Division of Cardiology, Stanford University School of Medicine, Stanford, California, USA
[j]Department of Medicine, Division of Cardiovascular Medicine, Stanford University, Stanford, California, USA
[k]Department of Pediatrics, Stanford University school of Medicine, Stanford, California, USA
[l]Department of Biochemistry, Stanford University School of Medicine, Stanford, California, USA
[m]Department of Developmental Biology, Stanford University School of Medicine, Stanford, California, USA
[n]Howard Hughes Medical Institute, USA
[o]Department of Bioengineering, Stanford University, Stanford, California, USA
[p]Department of Microbiology & Immunology, Stanford University School of Medicine, Stanford, California, USA
[q] Department of Medicine and Liver Center, University of California San Francisco, San Francisco, California, USA
[r]Chan Zuckerburg Biohub, San Francisco, California, USA
[s]Flow Cytometry Core, V.A. Palo Alto Healthcare System, Palo Alto, California, USA
[t] Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, California USA
[u] Department of Medicine and Stanford Diabetes Research Center, Stanford University, Stanford, California USA
[v] Department of Surgery, Division of Plastic and Reconstructive Surgery, Stanford University, Stanford, California USA
[w] Department of Neurobiology, Stanford University School of Medicine, Stanford, CA USA
[x] Department of Urology, Stanford University School of Medicine, Stanford, California, USA
[y] Department of Medical Biotechnology, Faculty of Biophysics, Biochemistry and Biotechnology , Jagiellonian University, Poland

## Methods

### Mice and Tissue Collection

Four 10-15 week old male and four virgin female C57BL/6 mice were shipped from the National Institute on Aging colony at Charles River to the Veterinary Medical Unit (VMU) at the VA Palo Alto (VA). At both locations, mice were housed on a 12-h light/dark cycle, and provided food and water *ad libitum*. The diet at Charles River was NIH-31, and Teklad 2918 at the VA VMU. Littermates were not recorded or tracked, and mice were housed at the VA VMU for no longer than 2 weeks before euthanasia. Prior to tissue collection, mice were placed in sterile collection chambers for 15 minutes to collect fresh fecal pellets. Following anesthetization with 2.5% v/v Avertin, mice were weighed, shaved, and blood drawn via cardiac puncture before transcardial perfusion with 20 ml PBS. Mesenteric adipose tissue (MAT) was then immediately collected to avoid exposure to the liver and pancreas perfusate, which negatively impacts cell sorting. Isolating viable single cells from both pancreas and liver of the same mouse was not possible, therefore, 2 males and 2 females were used for each. Whole tissues were then dissected in the following order: colon, spleen, thymus, trachea, tongue, brain, heart, lung, kidney, gonadal adipose tissue (GAT), bladder, diaphragm, skeletal muscle (*tibialis anterior*), skin (dorsal), subcutaneous adipose tissue (SCAT, inguinal pad), mammary glands (fat pads 2, 3, and 4), brown adipose tissue (BAT, interscapular pad), aorta, and bone marrow (spine and limb bones). Following single cell dissociation as described below, cell suspensions were either used for FACS sorting of individual cells into 384-well plates, or for microfluidic droplet library preparation. All animal care and procedures were carried out in accordance with institutional guidelines approved by the VA Palo Alto Committee on Animal Research.

**Tissue dissociation and sample preparation:** Specific protocols for each tissue are described in the supplement.

## Single Cell Methods

### Lysis plate preparation

Lysis plates were created by dispensing 0.4 µl lysis buffer (0.5U Recombinant RNase Inhibitor (Takara Bio, 2313B), 0.0625% Triton™ X-100 (Sigma, 93443-100ML), 3.125 mM dNTP mix (Thermo Fisher, R0193), 3.125 µM Oligo-d$T_{30}$VN (IDT, 5'AAGCAGTGGTATCAACGCAGAGTACT$_{30}$VN-3') and 1:600,000 ERCC RNA spike-in mix (Thermo Fisher, 4456740)) into 384-well hard-shell PCR plates (Biorad HSP3901) using a Tempest or Mantis liquid handler (Formulatrix). 96 well lysis plates were also prepared with 4 µl lysis buffer. All plates were sealed with AlumaSeal CS Films (Sigma-Aldrich Z722634) or Microseal 'F' PCR plate seal (Biorad MSF 1001) and then spun down for 1 minute at 3220 xg and snap frozen on dry ice. Plates were stored at -80°C until used for sorting.

### FACS sorting

After dissociation, single cells from each tissue were isolated in 384 or 96-well plates via Fluorescence Activated Cell Sorting (FACS). Most tissues were sorted into 384-well plates using SH800S (Sony) sorters. Heart and liver were sorted into 96-well plates and cardiomyocytes were hand-picked into 96-well plates. Skeletal muscle and diaphragm were sorted into 384-well plates on an Aria III (Becton Dickinson) sorter. Most tissues used combinations of fluorescent antibodies to enrich for the presence of known rare cell populations (see tissue section below), but some tissues were simply sorted into plates after removal of dead cells, small and large debris. For all samples, an initial gate was set to exclude small debris, platelets and cell aggregates, and most tissues included a forward scattering gate to select for single cells. One color channel was used in most tissues to stain and exclude dead cells and high prevalence cells that would otherwise dominate the cell population. Color compensation was used whenever necessary. Cells were sorted using the highest purity setting on the sorter ("Single cell" in the case of the SH800s) for all but the rarest cell types, for which the "Ultrapure" setting was used. Sorters were calibrated using FACS buffer before collecting cells from any tissue and after every 8 sorted plates to ensure accurate dispensing into plate wells. For a typical sort, a tube containing 1-3ml the pre-stained cell suspension was filtered, vortexed gently and loaded onto the FACS machine. A small number of cells were flowed at low pressure to check cell concentration and amount of debris. Then the pressure was adjusted, flow was paused, the first destination plate was unsealed and loaded, and single-cell sorting started. If a cell suspension was too concentrated, it was diluted it using FACS buffer or 1X PBS. For some cell types (e.g. hepatocytes), 96 well plates were used when it was not possible to sort individual cells accurately into 384 well plates. Immediately after sorting, plates were sealed with a pre-labeled aluminium seal, centrifuged and flash frozen on dry ice. On average, each 384-well plate took around 8 minutes to sort.

**cDNA synthesis and library preparation**

cDNA synthesis was performed using the Smart-seq2 protocol [1,2]. Briefly, 384-well plates containing single-cell lysates were thawed on ice followed by first strand synthesis. 0.6 µl of reaction mix (16.7 U/µl SMARTScribe Reverse Transcriptase (Takara Bio, 639538), 1.67 U/µl Recombinant RNase Inhibitor (Takara Bio, 2313B), 1.67X First-Strand Buffer (Takara Bio, 639538), 1.67 µM TSO (Exiqon, 5'-AAGCAGTGGTATCAACGCAGACTACATrGrG+G-3'), 8.33 mM DTT (Bioworld, 40420001-1), 1.67 M Betaine (Sigma, B0300-5VL), and 10 mM $MgCl_2$ (Sigma, M1028-10X1ML)) was added to each well using a Tempest liquid handler. Reverse transcription was carried out by incubating wells on a ProFlex 2x384 thermal-cycler (Thermo Fisher) at 42°C for 90 min and stopped by heating at 70°C for 5 min.

Subsequently, 1.5 µl of PCR mix (1.67X KAPA HiFi HotStart ReadyMix (Kapa Biosystems, KK2602), 0.17 µM IS PCR primer (IDT, 5'-AAGCAGTGGTATCAACGCAGAGT-3'), and 0.038 U/µl Lambda Exonuclease (NEB, M0262L)) was added to each well with a Mantis liquid handler (Formulatrix), and second strand synthesis was performed on a ProFlex 2x384 thermal-cycler by using the following program: 1. 37°C for 30 minutes, 2. 95°C for 3 minutes, 3. 23 cycles of 98°C for 20 seconds, 67°C for 15 seconds, and 72°C for 4 minutes, and 4. 72°C for 5 minutes.

The amplified product was diluted with a ratio of 1 part cDNA to 10 parts 10mM Tris-HCl (Thermo Fisher, 15568025), and concentrations were measured with a dye-fluorescence assay (Quant-iT dsDNA High Sensitivity kit; Thermo Fisher, Q33120) on a SpectraMax i3x microplate reader (Molecular Devices). Sample plates were selected for downstream processing if the mean concentration of blanks - ERCC-containing, non-cell wells - was greater than 0 ng/µl, and, after linear regression of the values obtained from the Quant-iT dsDNA standard curve, the $R^2$ value was greater than 0.98. Sample wells were then selected if their cDNA concentrations were at least one standard deviation greater than the mean concentration of the blanks. These wells were reformatted to a new 384-well plate at a concentration of 0.3 ng/µl and final volume of 0.4 µl using an Echo 550 acoustic liquid dispenser (Labcyte).

Illumina sequencing libraries were prepared as described in Darmanis et al. 2015.[13] Briefly, tagmentation was carried out on double-stranded cDNA using the Nextera XT Library Sample Preparation kit (Illumina, FC-131-1096). Each well was mixed with 0.8 µl Nextera tagmentation DNA buffer (Illumina) and 0.4 µl Tn5 enzyme (Illumina), then incubated at 55°C for 10 min. The reaction was stopped by adding 0.4 µl "Neutralize Tagment Buffer" (Illumina) and spinning at room temperature in a centrifuge at 3220xg for 5 min. Indexing PCR reactions were performed by adding 0.4 µl of 5 µM i5 indexing primer, 0.4 µl of 5 µM i7 indexing primer, and 1.2 µl of Nextera NPM mix (Illumina). PCR amplification was carried out on a ProFlex 2x384 thermal cycler using the following program: 1. 72°C for 3 minutes, 2. 95°C for 30 seconds, 3. 12 cycles of 95°C for 10 seconds, 55°C for 30 seconds, and 72°C for 1 minute, and 4. 72°C for 5 minutes.

**Library pooling, quality control, and sequencing**

Following library preparation, wells of each library plate were pooled using a Mosquito liquid handler (TTP Labtech). Pooling was followed by two purifications using 0.7x AMPure beads (Fisher, A63881). Library quality was assessed using capillary electrophoresis on a Fragment Analyzer (AATI), and libraries were quantified by qPCR (Kapa Biosystems, KK4923) on a CFX96 Touch Real-Time PCR Detection System (Biorad). Plate pools were normalized to 2 nM and equal volumes from 10 or 20 plates were mixed together to make the sequencing sample pool. PhiX control library was spiked in at 0.2% before sequencing.

**Sequencing libraries from 384-well and 96-well plates**

Libraries were sequenced on the NovaSeq 6000 Sequencing System (Illumina) using 2 x 100bp paired-end reads and 2 x 8bp or 2 x 12bp index reads with either a 200 or 300-cycle kit (Illumina, 20012861 or 20012860).

**Microfluidic droplet single cell analysis**

Single cells were captured in droplet emulsions using GemCode Single-Cell Instrument (10x Genomics, Pleasanton, CA, USA) and SC RNA-seq libraries were constructed as per 10X Genomics protocol using GemCode Single-Cell 3′ Gel Bead and Library V2 Kit. Briefly, single cell suspensions were examined using an inverted microscope and if sample quality was deemed satisfactory, sample was diluted in PBS/2% FBS to achieve a

target concentration of 1000 cells/μl. If cell suspensions contained cell aggregate or debris, two additional washes in PBS/2% FBS at 300 x g for 5min at 4C were performed. Cell concentration was measured either with a Moxi GO II (Orflo Technologies) or a hemocytometer. Cells were loaded in each channel with a target output of 5,000 cells per sample. All reactions were performed in Biorad C1000 Touch Thermal cycler with 96-Deep Well Reaction Module. 12 cycles were used for cDNA amplification and sample index PCR. Amplified cDNA and final library were evaluated on a Fragment Analyzer using a High Sensitivity NGS Analysis Kit (Advanced Analytical). 10x cDNA libraries were quantitated for average fragment length using a 12 or 96 capillary Fragment Analyzer (Advanced Analytical) and by qPCR with Kapa Library Quantification kit for Illumina. Each library was diluted to 2nM and equal volumes of 16 libraries were pooled to fill each NovaSeq sequencing run. Pools were sequenced with 100 cycle run kits with 26 bases Read 1, 8 bases Index 1 and 90 bases Read 2 (Illumina 20012862). PhiX control library was spiked in at 0.2 to 1%. Libraries were sequenced on the NovaSeq 6000 Sequencing System (Illumina)

**Data Processing**

Sequences from the NovaSeq were demultiplexed using bcl2fastq version 2.19.0.316. Reads were aligned using to a copy of the mm10 genome augmented with ERCC sequences, using STAR version 2.5.2b with the following options:

```
--outFilterType BySJout --outFilterMultimapNmax 20 --alignSJoverhang
Min 8 --alignSJDBoverhangMin 1 --outFilterMismatchNmax 999 --outFilt
erMismatchNoverLmax 0.04 --alignIntronMin 20 --alignIntronMax 100000
0 --alignMatesGapMax 1000000 --outSAMstrandField intronMotif --outSA
Mtype BAM Unsorted --outSAMattributes NH HI NM MD --genomeLoad LoadA
ndKeep --outReadsUnmapped Fastx --readFilesCommand zcat
```

Gene counts were produced using HTSEQ version 0.6.1p1 with default parameters except stranded was set to false and mode was set to intersection-nonempty.

Sequences from the microfluidic platform were demultiplexed and aligned using CellRanger with default parameters.

**Clustering**

Standard procedures for filtering, variable gene selection, dimensionality reduction, and clustering were performed using the Seurat package. For each tissue and each sequencing method (FACS and microfluidic), the following steps were performed. (The parameters used were tuned on a per-tissue basis. Values appear in the supplement.)

1. Genes appearing in fewer than 5 cells were excluded.
2. Cells with fewer than 500 genes were excluded. Cells with fewer than 50,000 reads (FACS) or 1000 UMI (microfluidic) were excluded. (In some organs, cells with more than 2 million reads were also excluded as a conservative measure to avoid doublets.)
3. Counts were log-normalized (log(1 + counts per N)), then scaled by linear regression against the number of reads (or UMIs), the percent of reads mapping to Rn45s, and the percent of reads to ribosomal genes.

4. Variable genes were selected using a threshold for dispersion (log of variance/mean). The distribution of dispersion in each expression bin was computed, and variable genes were those with a z-score of at least 0.5-1..
5. The variable genes were projected onto a low-dimensional subspace using principal component analysis. The number of principal components was selected based on inspection of the plot of variance explained.
6. A shared-nearest-neighbors graph was constructed based on the Euclidean distance metric in the low-dimensional subspace. (The SNN is a weighted graph where w_ij is the Jaccard similarity of the k-neighborhood of i with the k-neighborhood of j, where k = 30.) Cells were clustered using the Louvain method to maximize modularity.
7. Cells were visualized using a 2-dimensional t-distributed Stochastic Neighbor Embedding on the same distance metric.
8. Cell types were assigned to each cluster using the abundance of known marker genes. (Plots showing the expression of the markers for each tissue appear in the supplement).
9. When clusters appeared to be mixtures of cell types, they were refined either by increasing the resolution parameter for Louvain clustering or subsetting the data and rerunning steps 3-7.

A similar analysis was done globally for all FACS processed cells and for all microfluidic processed cells.

To attempt to remove the effect of technical variation from the gene expression matrix, it is common practice to regress out statistics associated with quality, like the number of reads or the percent of ribosomal or mitochondrial RNA. If those statistics differ systematically between cell types or tissues, this can have the unwanted effect of compressing the true variation between those groups. For example, the percent of ribosomal RNA varies from 1% in the Liver and Pancreas to 5% in the Spleen and Tongue. Hence we regress out these factors separately in the analysis of each tissue.

For the "dynamic range" comparison in Supplementary Figure 4, the gene expression was scaled to log of reads per million for FACS and log of UMI per 10,000 for microfluidic emulsions. Genes with nonzero expression were rank ordered and plotted.

## References


1. Thorsen, T., Roberts, R. W., Arnold, F. H. & Quake, S. R. Dynamic pattern formation in a vesicle-generating microfluidic device. *Phys. Rev. Lett.* **86,** (2001).
2. Macosko, E. Z. *et al.* Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161,** 1202–1214 (2015).
3. Ramsköld, D. *et al.* Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat. Biotechnol.* **30,** 777–782 (2012).
4. Alberts, B. *et al*. *Molecular biology of the cell*. (Garland Science, 2002).
5. Guo, G. *et al.* Resolution of Cell Fate Decisions Revealed by Single-Cell Gene Expression Analysis from Zygote to Blastocyst. *Dev. Cell* **18,** 675–685 (2010).
6. Dalerba, P. *et al.* Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat. Biotechnol.* **29,** (2011).
7. Wu, A. R. *et al.* Quantitative assessment of single-cell RNA-sequencing methods. (2013). doi:10.1038/nmeth.2694
8. Treutlein, B. *et al.* Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* **509,** (2014).
9. Enge, M. *et al.* Single-Cell Analysis of Human Pancreas Reveals Transcriptional Signatures of Aging and Somatic Mutation Patterns. *Cell* **171,** (2017).
10. Halpern, K. B. *et al.* Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* **542,** 352–356 (2017).
11. Haber, A. L. *et al.* A single-cell survey of the small intestinal epithelium. *Nature* **551,** 333–339 (2017).
12. Villani, A.-C. *et al.* Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356,** eaah4573 (2017).
13. Darmanis, S. *et al.* A survey of human brain transcriptome diversity at the single cell level. *Proc. Natl. Acad. Sci. U. S. A.* **112,** (2015).
14. Gokce, O. *et al.* Cellular Taxonomy of the Mouse Striatum as Revealed by Single-Cell RNA-Seq. *Cell Rep.* **16,** (2016).
15. Usoskin, D. *et al.* Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. *Nat. Neurosci.* **18,** 145–153 (2014).
16. Zeisel, A. *et al.* Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347,** 1138–42 (2015).
17. Li, H. *et al.* Classifying Drosophila Olfactory Projection Neuron Subtypes by Single-Cell RNA Sequencing. *Cell* **171,** 1206–1220.e22 (2017).
18. van den Brink, S. C. *et al.* Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat. Methods* **14,** 935–936 (2017).

## Figure captions

**Figure 1.** a) 19 tissues from 4 male and 3 female mice were analyzed. After tissue dissociation, cells were either sorted by FACS or manipulated in microfluidic emulsions, after which they were lysed and their transcriptomes amplified, followed by sequencing, read mapping, and data analysis. b) Barplot showing number of sequenced cells prepared by FACS sorting from each tissue (n=20). c) Barplot showing number of sequenced cells prepared by microfluidic emulsion from each tissue (n=12). d) Histogram of number of reads per cell for each tissue from FACS sorted cells. e) Histogram of number of genes detected per cell for each tissue from FACS sorted cells. f) Histogram of number of unique molecular identifiers (UMI) sequenced per cell for each tissue from cells prepared by microfluidic emulsion. g) Histogram of number of genes detected per cell for each tissue for cells prepared by microfluidic emulsion.

**Figure 2**. a) Dimensionally reduced tSNE plot of all cells sorted by FACS color coded by tissue of origin. b) Dimensionally reduced tSNE plots for each tissue of cells sorted by FACS. Color coding indicates distinct clusters. c) Barplots of manually annotated cell types based on differential gene expression across all tissues.

**Figure 3**. Comparison of cell type determination as done by unbiased whole transcriptome comparison versus manual annotation by organ-specific experts. The x-axis represents clusters from Figure 2a with multiple organs contributing, while the y-axis represents manual expert annotation of cell types in an organ-specific fashion based on the data in Figure 2b and 2c. The unbiased method discovers relationships between similar cell types found in different organs (highlighted regions); in particular it groups T cells from different organs into a single cluster, B cells from different organs into a different single cluster, and endothelial cells from different organs into a single cluster.

**Figure 4**. Analysis of all T cells sorted by FACS. a) Dimensionally reduced tSNE plot of all T cells colored by cluster membership. Five clusters were identified. b) Dotchart showing level of expression (color scale) and number of expressing cells (point diameter) within each cluster of T cells. c) Barplot showing fraction of T cells originating from Fat, Lung, Marrow, Muscle, Spleen or Thymus for each of the T cell clusters. d) Barplot showing fraction of $Cd4^+$/ $Cd8^+$/ $Cd4^+Cd8^+$ / $Cd4^-Cd8^-$ T cells for each of the T cell clusters.

**Figure 5.** Transcription factor expression analysis. A. Gene-gene correlation (correlogram) of top cell type-tissue specific transcription factors (TFs) (selected by minimum differential expression of 0.3 and p-value of $10^{-6}$; immediate early genes excluded). Columns are colored by the cell ontology (left column) and tissue (right column) of the cell type+tissue with the highest expression of that gene. Insets show clusters of TFs enriched in selected cell types and tissues. B. Correlogram of top tissue-specific TFs for epithelial cells (cells ontologies containing “epithelial”, “basal”, “keratinocyte”, or “epidermis”). Row colors correspond to tissue (right) and cell ontology (left) of most-enriched cell type. tSNE plots are calculated using all variable genes and show tissue origin (right) and

gene expression of select TFs (bottom). C. Correlogram and tSNE of tissue-specific TFs within stromal cells (cells ontologies containing “stromal”, “fibroblast”, or “mesenchymal”). D. Correlogram and tSNE of tissue-specific TFs within endothelial cells.

**Supplementary Figure Captions**

**Supplementary Figure 1** Histograms of number of genes detected across tissues (Red = FACS sorted cells, Cyan = cells processed in microfluidic emulsion)

**Supplementary Figure 2** Plots across three cell types which compare number of cells expressing a given gene for measurements made by FACS (x axis) compared to microfluidic emulsion (y axis). Each data point is a gene and the axes represent fraction of cells which express that gene.

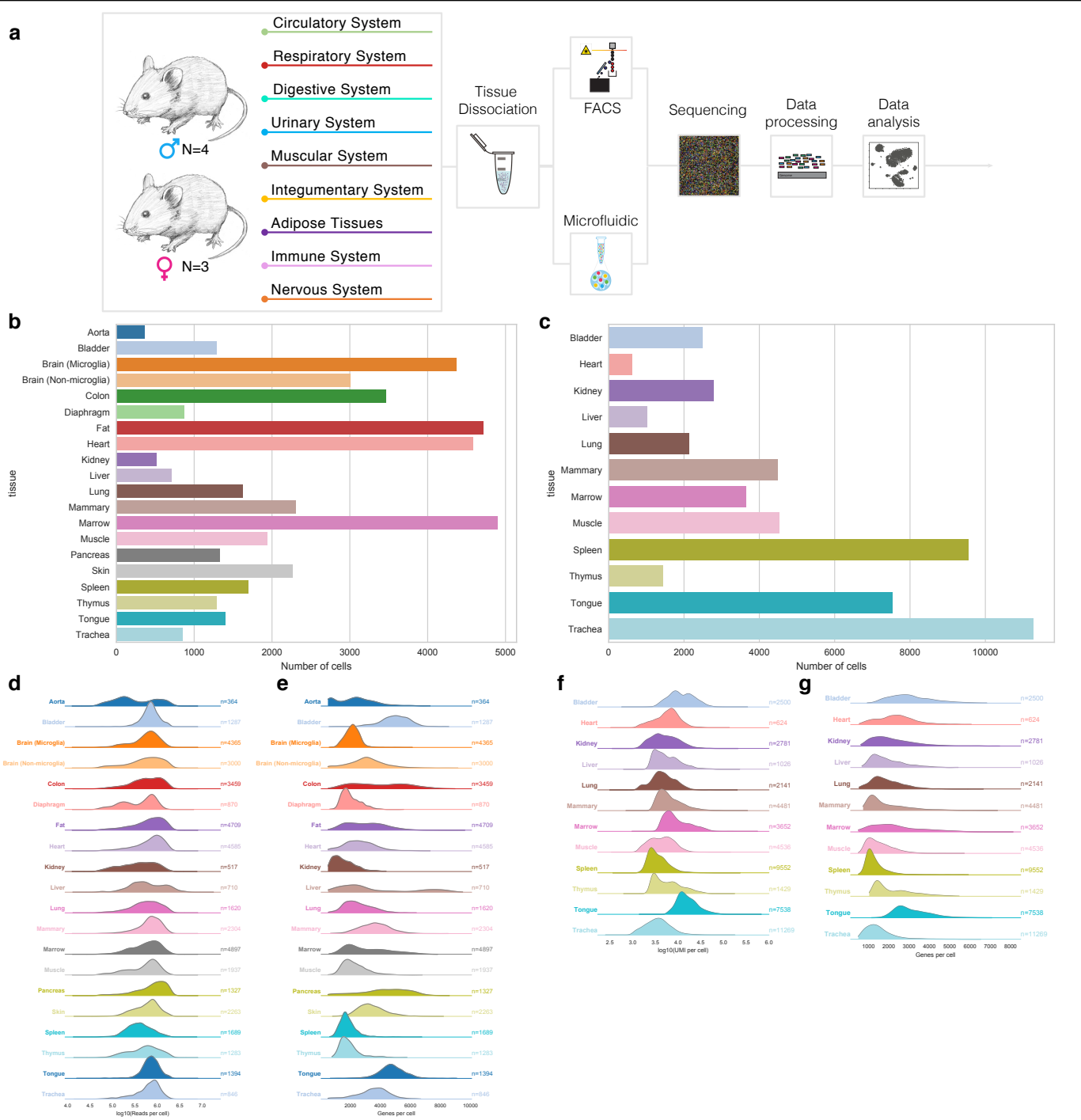
**Supplementary Figure 3** Correlation of number of genes detected in more cells across all common cell types. X-axis represents number of genes present in cell types measured by FACS, whereas Y-axis represents number of genes present in the same cell types measured by microfluidic emulsion. Most but not all of the cell types have more genes measured by FACS.

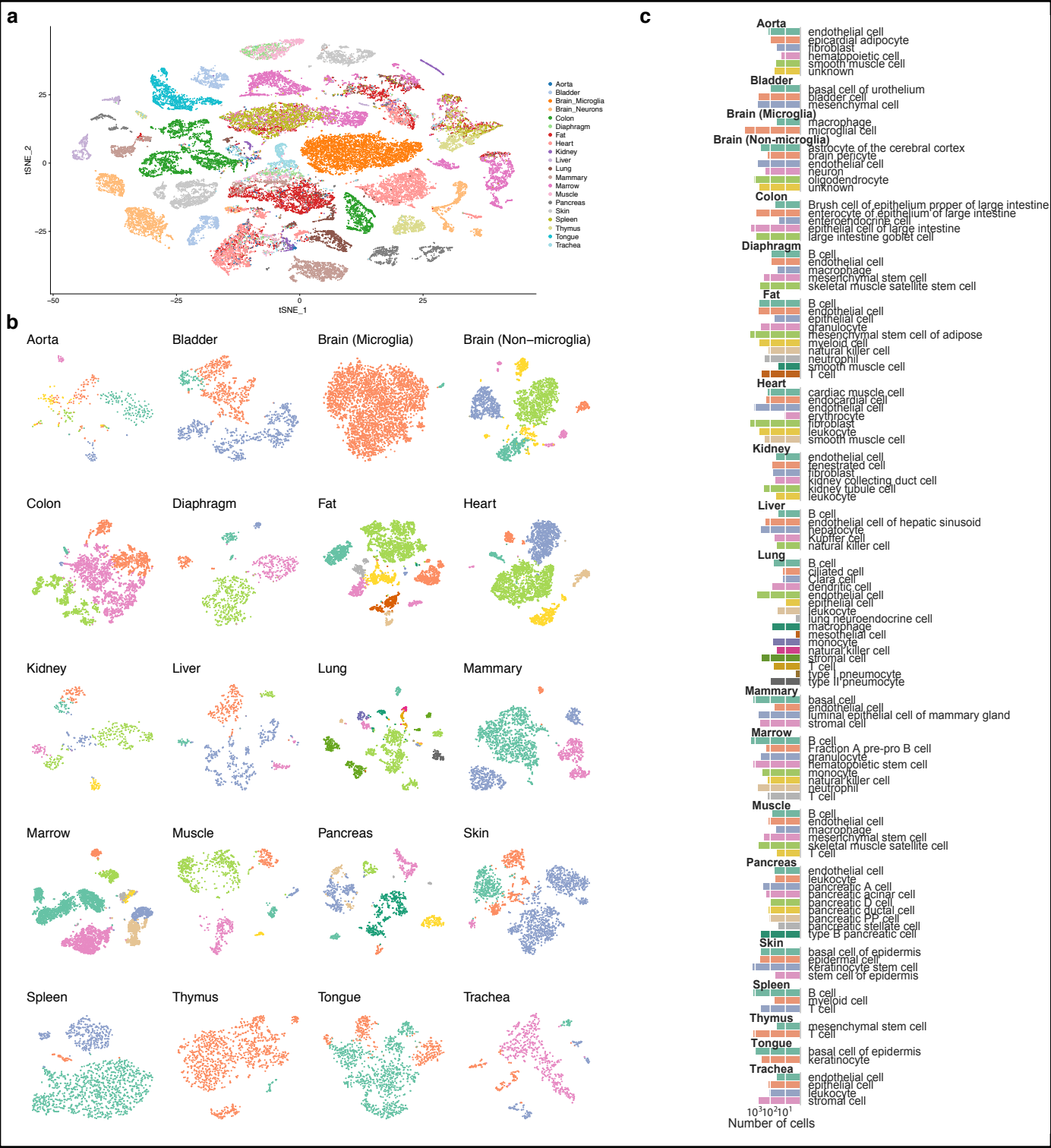
**Supplementary Figure 4** Plots showing examples of dynamic range for the two measurements approaches; FACS generally has many more reads per gene.

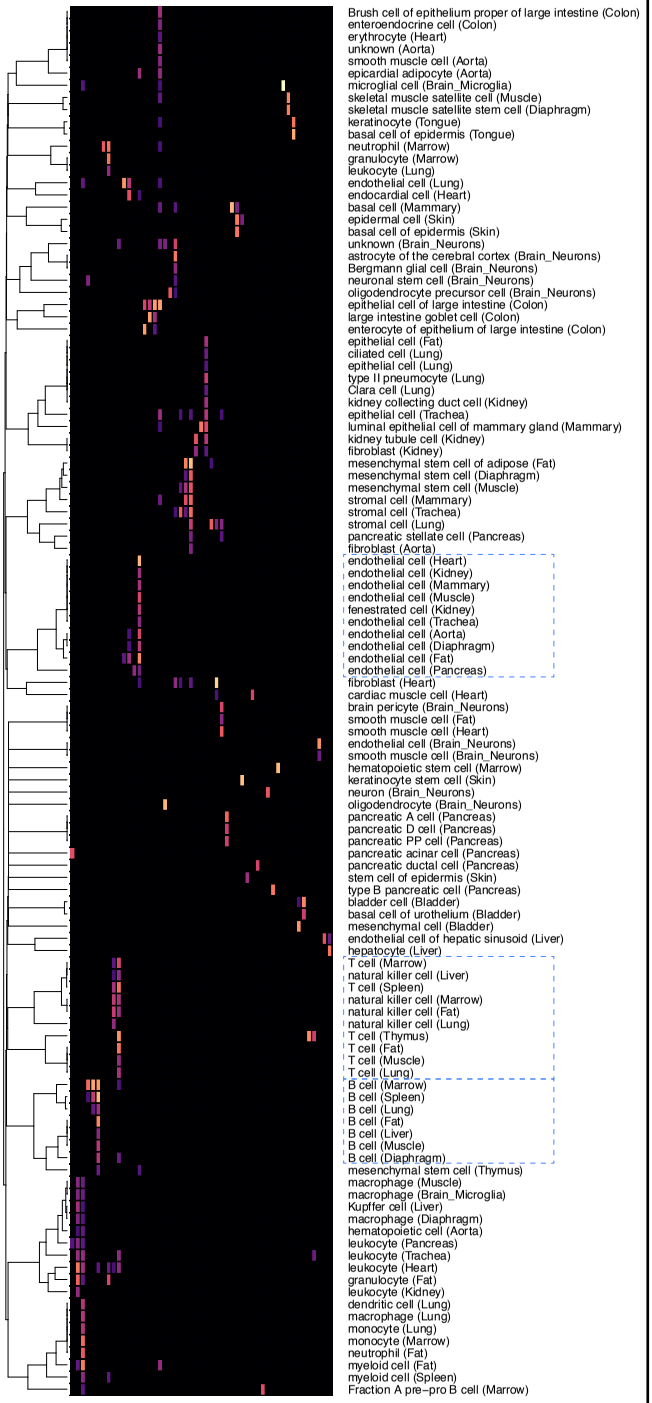
**Supplementary Figure 5** a) Dimensionally reduced tSNE plot of all cells manipulated by microfluidic emulsions color coded by tissue of origin. b) Dimensionally reduced tSNE plots for each tissue of cells sorted by microfluidic emulsions. Color coding indicates distinct clusters. c) Barplots of manually annotated cell types based on differential gene expression across all tissues.

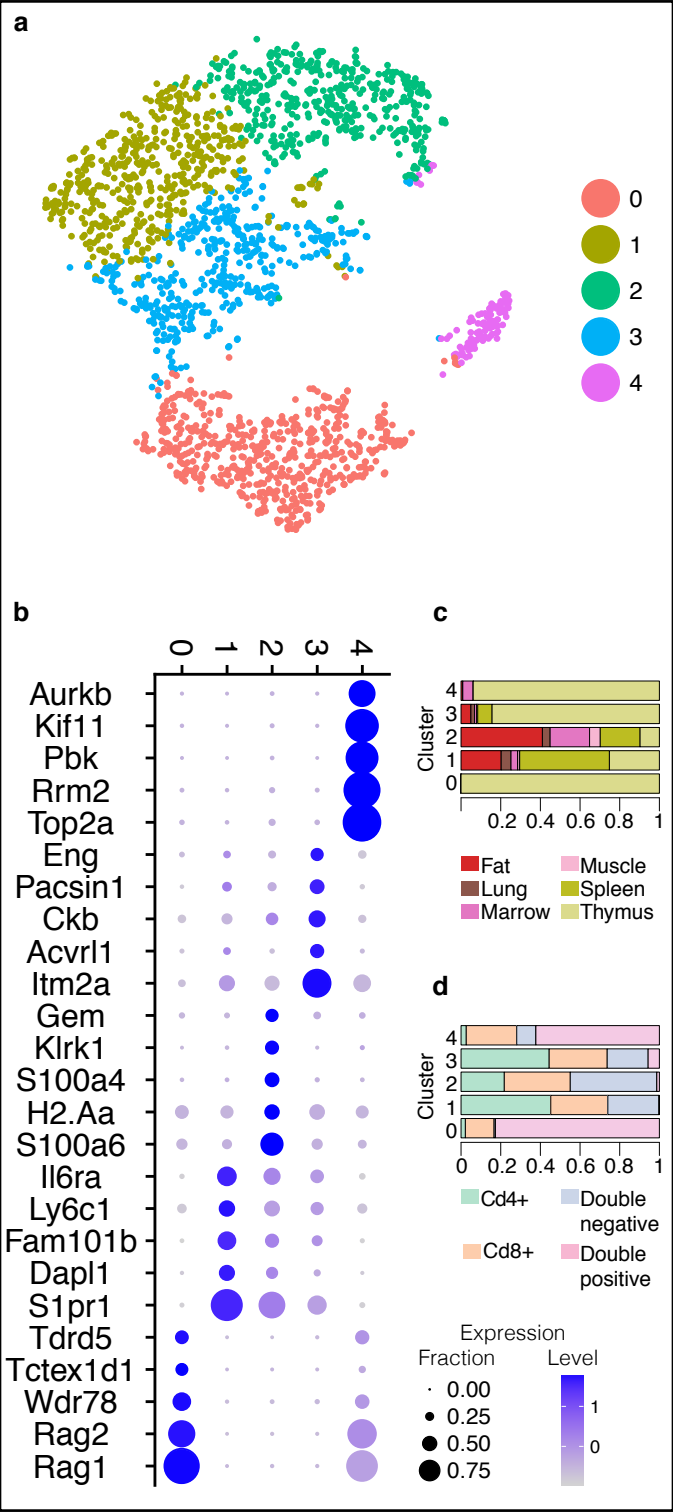
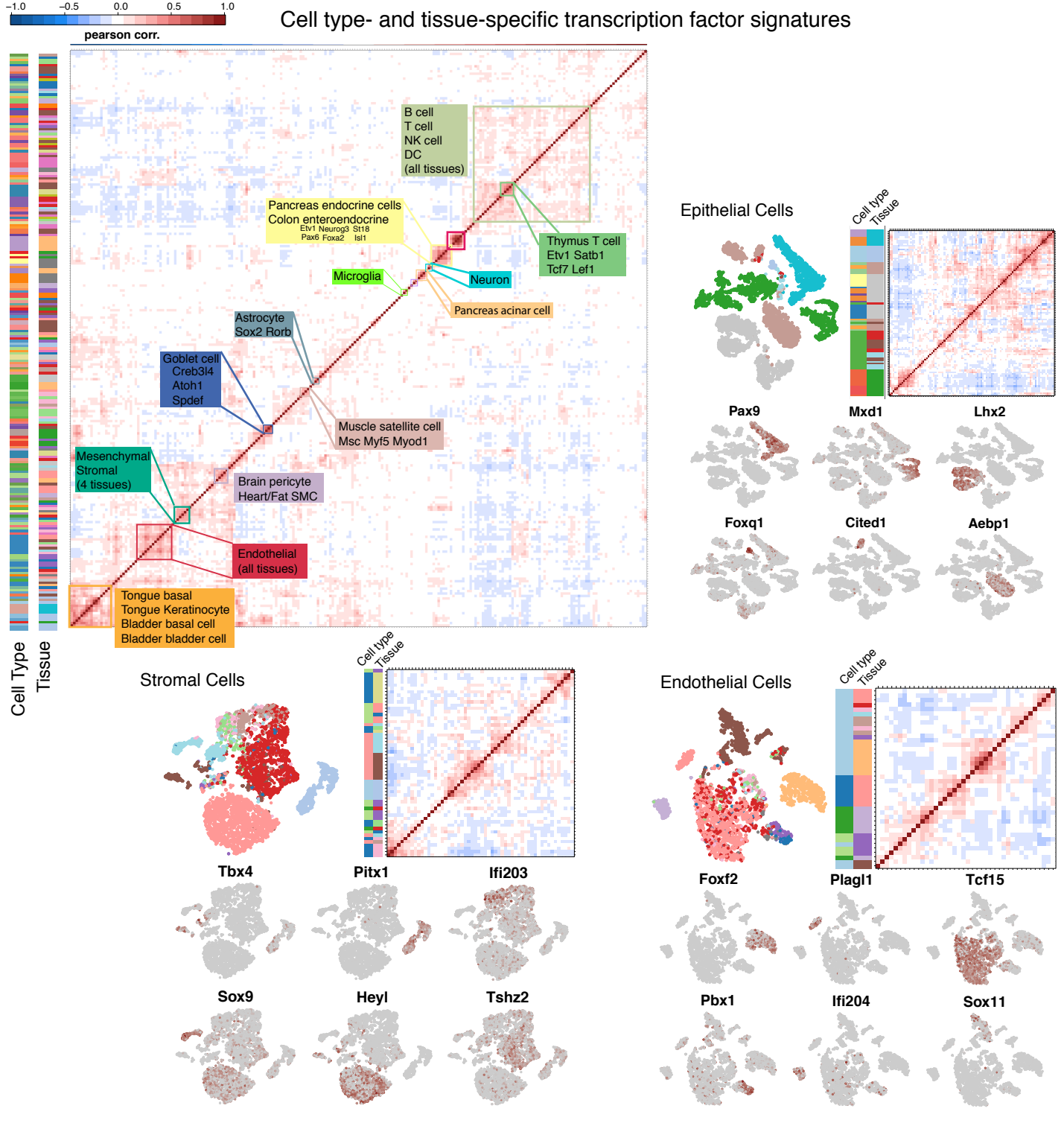
**Supplementary Figure 6** Analysis of dissociation induced gene expression scores across tissues.

**Supplementary Figure 7** Cell type- and tissue-specific enrichment of gene expression in TF gene clusters. Gene-gene correlations were calculated as in Figure 5 and clustered by cutting a hierarchical clustering tree at a height of 1.5 (Euclidean distance of correlations). To calculate scores, log10 cpm values were scaled to [0,1] for each gene. The score on the x-axis is the log of the geometric mean of all of the TFs in the cluster across all the cells of a cell type and tissue. The TFs in the cluster are indicated at the top of the plot; for clusters with more than 5 genes, only the first 5 are shown. The distribution of scores across all unique cell types and tissues is displayed, with the bar plots colored by tissue origin of the cell type (colors as in Fig. 5). Gene clusters that are highly specific show a strongly bimodal plot, with most cells expressing very low levels of the gene and one or a small number expressing high levels (for example, the “Spdef Atoh1 Creb3l4” cluster is highly specific to colon goblet cells). The names of the top three cell types are displayed as a subtitle below the genes, with the highest-scored at the top. Note that even though some gene clusters only separate one or two cell types, the top 3 by rank are still included in the subtitle.


a
Circulatory System
Respiratory System
Digestive System
Urinary System
Muscular System
Integumentary System
Adipose Tissues
Immune System
Nervous System
♂ N=4
♀ N=3
Tissue Dissociation
FACS
Microfluidic
Sequencing
Data processing
Data analysis
b
tissue
Aorta
Bladder
Brain (Microglia)
Brain (Non-microglia)
Colon
Diaphragm
Fat
Heart
Kidney
Liver
Lung
Mammary
Marrow
Muscle
Pancreas
Skin
Spleen
Thymus
Tongue
Trachea
0
1000
2000
3000
4000
5000
Number of cells
c
tissue
Bladder
Heart
Kidney
Liver
Lung
Mammary
Marrow
Muscle
Spleen
Thymus
Tongue
Trachea
0
2000
4000
6000
8000
10000
Number of cells
d
log10(Reads per cell)
e
Genes per cell
f
log10(UMI per cell)
g
Genes per cell


a
tSNE_2
tSNE_1
25
0
−25
−50
Aorta
Bladder
Brain_Microglia
Brain_Neurons
Colon
Diaphragm
Fat
Heart
Kidney
Liver
Lung
Mammary
Marrow
Muscle
Pancreas
Skin
Spleen
Thymus
Tongue
Trachea
b
Aorta
Bladder
Brain (Microglia)
Brain (Non−microglia)
Colon
Diaphragm
Fat
Heart
Kidney
Liver
Lung
Mammary
Marrow
Muscle
Pancreas
Skin
Spleen
Thymus
Tongue
Trachea
c
Aorta
endothelial cell
epicardial adipocyte
fibroblast
hematopoietic cell
smooth muscle cell
unknown
Bladder
basal cell of urothelium
bladder cell
mesenchymal cell
Brain (Microglia)
macrophage
microglial cell
Brain (Non-microglia)
astrocyte of the cerebral cortex
brain pericyte
endothelial cell
neuron
oligodendrocyte
unknown
Colon
Brush cell of epithelium proper of large intestine
enterocyte of epithelium of large intestine
enteroendocrine cell
epithelial cell of large intestine
large intestine goblet cell
Diaphragm
B cell
endothelial cell
macrophage
mesenchymal stem cell
skeletal muscle satellite stem cell
Fat
B cell
endothelial cell
epithelial cell
granulocyte
mesenchymal stem cell of adipose
myeloid cell
natural killer cell
neutrophil
smooth muscle cell
T cell
Heart
cardiac muscle cell
endocardial cell
endothelial cell
erythrocyte
fibroblast
leukocyte
smooth muscle cell
Kidney
endothelial cell
fenestrated cell
fibroblast
kidney collecting duct cell
kidney tubule cell
leukocyte
Liver
B cell
endothelial cell of hepatic sinusoid
hepatocyte
Kupffer cell
natural killer cell
Lung
B cell
ciliated cell
Clara cell
dendritic cell
endothelial cell
epithelial cell
leukocyte
lung neuroendocrine cell
macrophage
mesothelial cell
monocyte
natural killer cell
stromal cell
T cell
type I pneumocyte
type II pneumocyte
Mammary
basal cell
endothelial cell
luminal epithelial cell of mammary gland
stromal cell
Marrow
B cell
Fraction A pre-pro B cell
granulocyte
hematopoietic stem cell
monocyte
natural killer cell
neutrophil
T cell
Muscle
B cell
endothelial cell
macrophage
mesenchymal stem cell
skeletal muscle satellite cell
T cell
Pancreas
endothelial cell
leukocyte
pancreatic A cell
pancreatic acinar cell
pancreatic D cell
pancreatic ductal cell
pancreatic PP cell
pancreatic stellate cell
type B pancreatic cell
Skin
basal cell of epidermis
epidermal cell
keratinocyte stem cell
stem cell of epidermis
Spleen
B cell
myeloid cell
T cell
Thymus
mesenchymal stem cell
T cell
Tongue
basal cell of epidermis
keratinocyte
Trachea
endothelial cell
epithelial cell
leukocyte
stromal cell
$10^3$ $10^2$ $10^1$
Number of cells


Brush cell of epithelium proper of large intestine (Colon)
enteroendocrine cell (Colon)
erythrocyte (Heart)
unknown (Aorta)
smooth muscle cell (Aorta)
epicardial adipocyte (Aorta)
microglial cell (Brain_Microglia)
skeletal muscle satellite cell (Muscle)
skeletal muscle satellite stem cell (Diaphragm)
keratinocyte (Tongue)
basal cell of epidermis (Tongue)
neutrophil (Marrow)
granulocyte (Marrow)
leukocyte (Lung)
endothelial cell (Lung)
endocardial cell (Heart)
basal cell (Mammary)
epidermal cell (Skin)
basal cell of epidermis (Skin)
unknown (Brain_Neurons)
astrocyte of the cerebral cortex (Brain_Neurons)
Bergmann glial cell (Brain_Neurons)
neuronal stem cell (Brain_Neurons)
oligodendrocyte precursor cell (Brain_Neurons)
epithelial cell of large intestine (Colon)
large intestine goblet cell (Colon)
enterocyte of epithelium of large intestine (Colon)
epithelial cell (Fat)
ciliated cell (Lung)
epithelial cell (Lung)
type II pneumocyte (Lung)
Clara cell (Lung)
kidney collecting duct cell (Kidney)
epithelial cell (Trachea)
luminal epithelial cell of mammary gland (Mammary)
kidney tubule cell (Kidney)
fibroblast (Kidney)
mesenchymal stem cell of adipose (Fat)
mesenchymal stem cell (Diaphragm)
mesenchymal stem cell (Muscle)
stromal cell (Mammary)
stromal cell (Trachea)
stromal cell (Lung)
pancreatic stellate cell (Pancreas)
fibroblast (Aorta)
endothelial cell (Heart)
endothelial cell (Kidney)
endothelial cell (Mammary)
endothelial cell (Muscle)
fenestrated cell (Kidney)
endothelial cell (Trachea)
endothelial cell (Aorta)
endothelial cell (Diaphragm)
endothelial cell (Fat)
endothelial cell (Pancreas)
fibroblast (Heart)
cardiac muscle cell (Heart)
brain pericyte (Brain_Neurons)
smooth muscle cell (Fat)
smooth muscle cell (Heart)
endothelial cell (Brain_Neurons)
smooth muscle cell (Brain_Neurons)
hematopoietic stem cell (Marrow)
keratinocyte stem cell (Skin)
neuron (Brain_Neurons)
oligodendrocyte (Brain_Neurons)
pancreatic A cell (Pancreas)
pancreatic D cell (Pancreas)
pancreatic PP cell (Pancreas)
pancreatic acinar cell (Pancreas)
pancreatic ductal cell (Pancreas)
stem cell of epidermis (Skin)
type B pancreatic cell (Pancreas)
bladder cell (Bladder)
basal cell of urothelium (Bladder)
mesenchymal cell (Bladder)
endothelial cell of hepatic sinusoid (Liver)
hepatocyte (Liver)
T cell (Marrow)
natural killer cell (Liver)
T cell (Spleen)
natural killer cell (Marrow)
natural killer cell (Fat)
natural killer cell (Lung)
T cell (Thymus)
T cell (Fat)
T cell (Muscle)
T cell (Lung)
B cell (Marrow)
B cell (Spleen)
B cell (Lung)
B cell (Fat)
B cell (Liver)
B cell (Muscle)
B cell (Diaphragm)
mesenchymal stem cell (Thymus)
macrophage (Muscle)
macrophage (Brain_Microglia)
Kupffer cell (Liver)
macrophage (Diaphragm)
hematopoietic cell (Aorta)
leukocyte (Pancreas)
leukocyte (Trachea)
leukocyte (Heart)
granulocyte (Fat)
leukocyte (Kidney)
dendritic cell (Lung)
macrophage (Lung)
monocyte (Lung)
monocyte (Marrow)
neutrophil (Fat)
myeloid cell (Fat)
myeloid cell (Spleen)
Fraction A pre−pro B cell (Marrow)


a
0
1
2
3
4
b
0
1
2
3
4
Aurkb
Kif11
Pbk
Rrm2
Top2a
Eng
Pacsin1
Ckb
Acvrl1
Itm2a
Gem
Klrk1
S100a4
H2.Aa
S100a6
Il6ra
Ly6c1
Fam101b
Dapl1
S1pr1
Tdrd5
Tctex1d1
Wdr78
Rag2
Rag1
c
Cluster
4
3
2
1
0
0.2
0.4
0.6
0.8
1
Fat
Lung
Marrow
Muscle
Spleen
Thymus
d
Cluster
4
3
2
1
0
0
0.2
0.4
0.6
0.8
1
Cd4+
Cd8+
Double negative
Double positive
Expression
Fraction
0.00
0.25
0.50
0.75
Level
1
0

# Cell type- and tissue-specific transcription factor signatures