

Single-cell transcriptomics of 20 mouse organs creates a *Tabula Muris*

The Tabula Muris Consortium*

Here we present a compendium of single-cell transcriptomic data from the model organism *Mus musculus* that comprises more than 100,000 cells from 20 organs and tissues. These data represent a new resource for cell biology, reveal gene expression in poorly characterized cell populations and enable the direct and controlled comparison of gene expression in cell types that are shared between tissues, such as T lymphocytes and endothelial cells from different anatomical locations. Two distinct technical approaches were used for most organs: one approach, microfluidic droplet-based 3'-end counting, enabled the survey of thousands of cells at relatively low coverage, whereas the other, full-length transcript analysis based on fluorescence-activated cell sorting, enabled the characterization of cell types with high sensitivity and coverage. The cumulative data provide the foundation for an atlas of transcriptomic cell biology.

The cell is a fundamental unit of structure and function in biology, and multicellular organisms have evolved various cell types with specialized roles. Although cell types have historically been characterized by morphology and phenotype, the development of molecular methods has enabled increasingly precise descriptions of their properties, typically by measuring protein or mRNA expression patterns¹. Technological advances have also expanded measurement multiplexing such that highly parallel sequencing can now enumerate nearly every mRNA molecule in a single cell^{2–8}. This approach has provided insights into cell biology and organ composition from various organisms^{9–18}. However, although these reports provide valuable characterization of individual organs, it is challenging to compare data collected from different animals by independent labs with varying experimental techniques. It therefore remains unknown whether these data can be synthesized as a more general resource for biology.

Here we report a compendium of cell types from the mouse *Mus musculus*; we refer to this as a *Tabula Muris*, or 'Mouse Atlas'. We analysed several organs from the same mouse, generating a dataset controlled for age, environment and epigenetic effects. This enabled the direct comparison of cell-type composition between organs, and the comparison of shared cell types across organs. The compendium comprises single-cell transcriptomic data from 100,605 cells isolated from 20 organs from three female and four male, C57BL/6JN, three-month-old mice (10–15 weeks), analogous to 20-year-old humans (Fig. 1a). Aorta, bladder, bone marrow, brain (cerebellum, cortex, hippocampus and striatum), diaphragm, fat (brown, gonadal, mesenteric and subcutaneous), heart, kidney, large intestine, limb muscle, liver, lung, mammary gland, pancreas, skin, spleen, thymus, tongue and trachea from the same mouse were immediately processed into single-cell suspensions. All organs were single-cell-sorted into plates using fluorescence-activated cell sorting (FACS), and many were also loaded into microfluidic droplets (see Extended Data and Methods).

All data, protocols, analysis scripts and an interactive data browser are publicly available (for details, see 'Data availability'). This release enables the exact replication of all results, facilitates in-depth analyses not completed here, and provides a comparative framework for future studies using the large variety of murine disease models. Although 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.

Defining organ-specific cell types

To define cell types, we analysed each organ independently by performing principal component analysis (PCA) on the most variable genes between cells, followed by nearest-neighbour graph-based clustering. We then used cluster-specific gene expression of known markers and genes that are differentially expressed between clusters to assign cell-type annotations to each cluster (Extended Data Figs. 1, 2, Supplementary Table 1). We used a standard annotation method for all organs; step-by-step instructions to reproduce this method are provided in the supplemental Organ Annotation Vignette using the liver as an example. Cell type descriptions and defining genes for each organ are available in the Supplementary Information. For each cluster, we provide annotations in the controlled vocabulary of the cell ontology¹⁹ to facilitate inter-experiment comparisons. 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. Some unexpected discoveries include a potential new role for *Neurog3*, *Hhex* and *Prss53* in the adult pancreas, a cell population expressing *Chodl* in limb muscle, transcriptional heterogeneity of brain endothelial cells, the expression of MHC class II genes by adult mouse T cells, and sets of transcription factors that distinguish cell types across organs.

Methodological comparison

We performed single-cell RNA-sequencing with two methods: FACS-based cell capture in plates and microfluidic-droplet-based capture (hereafter denoted the FACS method and the microfluidic-droplet method, respectively). To understand the technical biases of each approach, we performed both methods on many organs. Overall, 44,949 cells from the FACS method and 55,656 cells from the microfluidic-droplet method were retained after quality control. Single-cell transcriptomes were sequenced to an average depth of 814,488 reads per cell (FACS) and 7,709 unique molecular identifiers (UMIs) per cell (microfluidic droplet). Comparing methods shows organ-specific differences in the number of cells analysed (Fig. 1b, c), reads per cell (Extended Data Fig. 3a, c) and genes per cell (Extended Data Fig. 3b, d). Furthermore, with both methods the most abundant cell types analysed were epithelial cells and leukocytes, although FACS captured a larger diversity of cell types (Extended Data Fig. 4).

Any individual single-cell sequencing experiment offers only a partial view of cell-type diversity within an organism and gene expression

*A list of authors and their affiliations is available online.

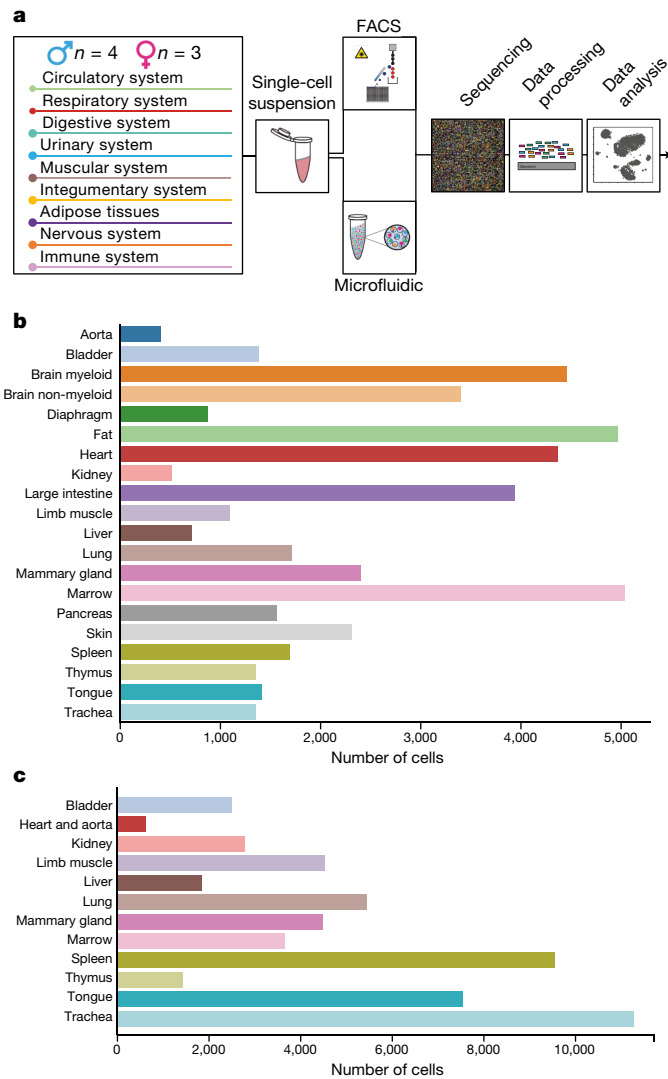


Fig. 1 | Overview of *Tabula Muris*. **a**, 20 organs from four male and three female mice were analysed. After dissociation, cells were sorted by FACS and, for some organs, captured in microfluidic oil droplets. Cells were lysed, transcriptomes amplified and sequenced, reads mapped, and data analysed. **b**, Bar plot showing the number of sequenced cells prepared by FACS from each organ ($n = 20$ organ types). **c**, Bar plot showing the number of sequenced cells prepared by microfluidic droplets from each organ ($n = 12$ organ types).

within each cell type. We illustrate the expected variability between methods and experiments by comparing our two measurement approaches to a third method, microwell-seq²⁰. One notable feature is the variability in the number of genes detected per cell between organs and methods. For example, the median number of genes detected in the bladder is around 4,900 (FACS), 2,900 (microfluidic droplet) and 900 (microwell-seq), whereas in the kidney it is around 1,400 (FACS), 1,900 (microfluidic droplet) and 500 (microwell-seq). In the bladder, liver, lung, mammary gland, trachea, tongue and spleen, nearly twice as many genes are detected per cell with the FACS method compared to the microfluidic-droplet method, whereas the heart and marrow show comparable numbers between the two methods (Extended Data Fig. 5a). This difference is probably not due to sequencing depth, as both FACS and microfluidic-droplet libraries are nearly saturated (Extended Data Fig. 5b). In these comparisons, a gene is considered detected if a single read maps to it, as that is the only value at which reads and UMIs can be treated equally. We also found that the number of detected genes decreases similarly across organs as the read or UMI threshold for a detectable gene is increased (Extended Data Fig. 6).

Next, we investigated whether the three methods agree on the genes defining each cell cluster (Methods). As expected, the FACS and microfluidic-droplet methods show the closest agreement, probably because they used the same biological samples. However, there are several dozen to several hundred genes common to all methods that define each cluster (Extended Data Fig. 7, Supplementary Table 2). This suggests that combining independent datasets can lead to more robust characterizations of gene expression.

Spleen and kidney are two organs for which FACS was performed without marker-based sorting, which enables us to compare the number and relative abundance of different cell types between methods. For those cell types that are captured by both methods, the proportion of each cell type is equivalent (Pearson correlation coefficient: spleen, 0.99; kidney, 0.99). Nonetheless, the microfluidic-droplet method identified cell types that were missed by the FACS method in both organs, for example kidney mesangial cells, and splenic dendritic and natural killer cells. This is partially explained by cellular abundance and sampling depth (12,333 microfluidic-droplet cells compared with 2,216 FACS cells, Supplementary Table 1), and possibly from cell capture and lysis biases between methods.

As the FACS method captures fewer cells but detects more molecules per cell than the microfluidic-droplet method, we asked whether the two methods agree in their 'bulk' gene-expression profiles for the 33 shared cell populations (Methods). Such gene-expression profiles largely correlate (Pearson correlation coefficient: 0.74–0.90), which suggests that although biases between methods exist, both accurately recapitulate average cell-type gene-expression profiles.

Global clustering across organs

To detect relationships between cells from different organs, we visualized all FACS cells with *t*-SNE and grouped them with unbiased, graph-based clustering (Fig. 2, Extended Data Fig. 8). As expected, cells from different organs often mixed, with 25 of 54 clusters containing (at least five) cells from distinct organs (Fig. 3). For example, clusters 3 and 48 each contain endothelial cells from five or more organs, and clusters 1 and 24 contain mesenchymal and stromal cells from four or more organs. Cluster 2 contains B cells from fat, limb muscle, lung, spleen, marrow and liver, but also cells annotated as leukocytes and lymphocytes from the thymus, heart and limb muscle. This suggests that the effect of cell type on measured gene expression is stronger than the effect of batch or dissociation protocol.

Cluster co-membership alone, however, is insufficient to conclude that two cell populations from different organs represent the same or similar cell types; at any given resolution, unbiased clustering that groups related cells may also group unrelated cells²¹. Therefore, to determine which clusters are composed of related or unrelated cell types, we computed a heterogeneity score for each cluster (Methods), and found low scores for the biologically sensible clusters discussed above (Extended Data Fig. 9). By contrast, the astrocytes and epithelial cells in cluster 53 are as different from one another as two random cells.

In addition to these heterogeneous groups, the clustering reveals small populations of potentially mislabelled cells inside homogenous populations. For example, ten thymus cells in cluster 3 (composed of 2,379 cells) are annotated as 'leukocytes', but they express *Pecam1*, which is an endothelial marker. This is a predictable artefact of the annotation scheme: because entire clusters, rather than individual cells, were annotated in each organ, a sufficiently rare cell type that was algorithmically grouped with a more populous cell type will be mis-annotated. This seems to occur only for populations smaller than about 30 cells, which comprise less than 4% of the overall dataset, and represents the lower limit of sensitivity in the current release of data interpretation.

The fact that most cells of similar cell types cluster together across organs and biological replicates shows that batch effects are not the main source of variance in the dataset. Our findings also show that manual annotation of cell types is consistent with unbiased transcriptomic clustering for sufficiently large populations. We expect that further development of multi-scale comparison algorithms will facilitate

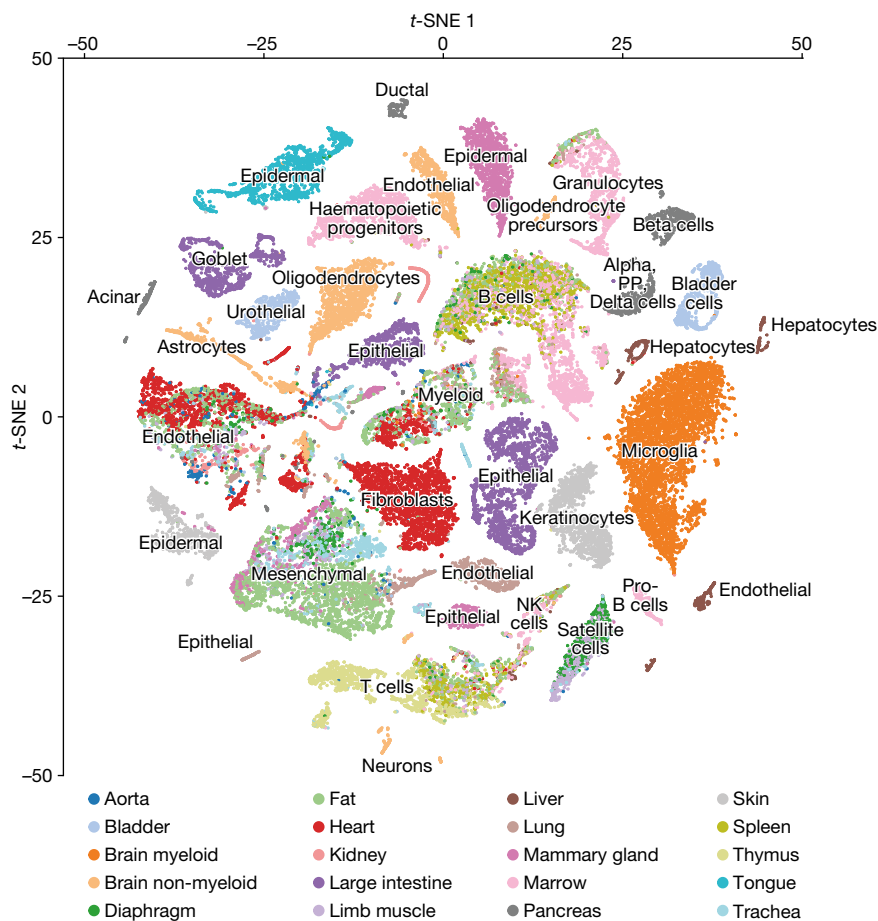


Fig. 2 | t-SNE visualization of all FACS cells. t-SNE plot of all cells collected by FACS, coloured by organ, overlaid with the predominant cell type composing each cluster; $n = 44,949$ individual cells.

the discovery of both universal and organ-specific gene modules within these shared cell types.

To demonstrate an example of investigating common cell types across organs, we collectively analysed all FACS cells annotated as T cells, which revealed five clusters (Fig. 4). Cluster 0 comprises thymic cells undergoing VDJ recombination characterized by the expression of *Rag1*, *Rag2* and *Dntt*, and includes uncommitted double-positive T cells (*Cd4*⁺ and *Cd8a*⁺). Cluster 4 contains predominantly proliferating thymic T cells, which may represent pre-T cells expanding after VDJ recombination. Clusters 1–3 contain mostly single-positive T cells (*Cd4*⁺ or *Cd8a*⁺). Cluster 3 contains *Cd5*^{hi} thymic T cells that are possibly undergoing positive selection, whereas Cluster 2 contains mostly non-thymic T cells expressing the high-affinity IL2 receptor (encoded by the genes *Il2ra* and *Il2rb*), which suggests that they are activated. Notably, they also express MHC class II genes (*H2-Aa* and *H2-Ab1*). Although this is known in human T cells, MHC class II was previously thought to be restricted to professional antigen-presenting cells in mice²². Finally, Cluster 1 also represents mature T cells, but primarily splenic.

Global transcription factor analysis

One major goal of defining cell identities is to understand the underlying regulatory networks. We investigated how transcription factors contribute to cell-type identity by clustering averaged gene-expression profiles for each cell type using only the 1,016 transcription factors expressed in our dataset (Fig. 5a). The resulting dendrogram closely resembles the dendrogram produced using all expressed genes, indicating that transcription factors can be used to reconstruct known cell-ontology relationships between bulk populations (entanglement = 0.11; Extended Data Fig. 10a). By contrast, when we repeated the analysis using cell-surface markers, RNA splicing factors, or the two groups

combined (equivalent to a random set of genes), the entanglement was 0.22, 0.25 and 0.34, respectively, which suggests that none of these molecular classes define cell type to the extent that transcription factors do.

We then analysed organ-specific transcription factors by performing correlation analysis on shared cell types between organs²³ (epithelial cells, endothelial cells, B cells and T cells; Fig. 5b–e, Extended Data Fig. 10b–i). To understand which transcription factors were most informative for specifying cell types, we performed variable selection using random forest models (Methods) and determined that 136 transcription factors are needed to simultaneously define all cell types across all organs (Fig. 5f, Supplementary Table 3). We then determined the transcription factor sets that distinguish each individual cell type from all other cells. These sets vary substantially in size (from 2 to 813 transcription factors) and are not necessarily unique to each cell type (Fig. 5g–i, Supplementary Table 4).

A possible application for such transcription factor networks is the design of reprogramming protocols. Indeed, the transcription factors used in published methods are found in the cell-type-specific transcription factors sets we discovered (Supplementary Table 5). For some cell types, such as hepatocytes, satellite cells and oligodendrocytes, those reprogramming factors are the top variables segregating cell types (Fig. 5g–i). In fact, for nearly all reprogramming protocols the transcription factors used also specified the targeted cell type in our data (Supplementary Table 5), which suggests that our data can inform novel reprogramming schemes.

Discussion

A key challenge for single-cell studies is to understand transcriptomic changes caused by dissociation. A previous study showed that quiescent limb-muscle satellite cells activate upon dissociation and consequently



Fig. 3 | Comparison of cell-type determination. Comparison of cell-type determination as performed by unbiased whole-transcriptome comparison versus manual annotation of clusters by organ-specific experts. The *x* axis represents clusters from Fig. 2, while the *y* axis represents manual expert annotation of clusters derived from individual organs analysed independently (Extended Data Fig. 1). The unbiased method discovers relationships between similar cell types found in different organs; in particular T cells from different organs are grouped into a single cluster, B cells from different organs into a different single cluster, and endothelial cells from different organs into a single cluster (regions outlined in blue boxes).

express immediate early genes and other dissociation-related markers²⁴. We clearly observed these markers in several organs including limb muscle (Extended Data Fig. 11), but many showed little evidence of cellular activation. Therefore, the dissociation-related satellite-cell markers are not universal, and organs probably display unique dissociation-related expression profiles. Importantly, the presence of such changes in gene expression does not prevent the identification of cell type or the comparison of cell types across organs.

Another challenge for single-cell studies is experimental design amid the choice of several technologies. Droplet-based technologies offer certain advantages in the discovery of rare cell types or states, for

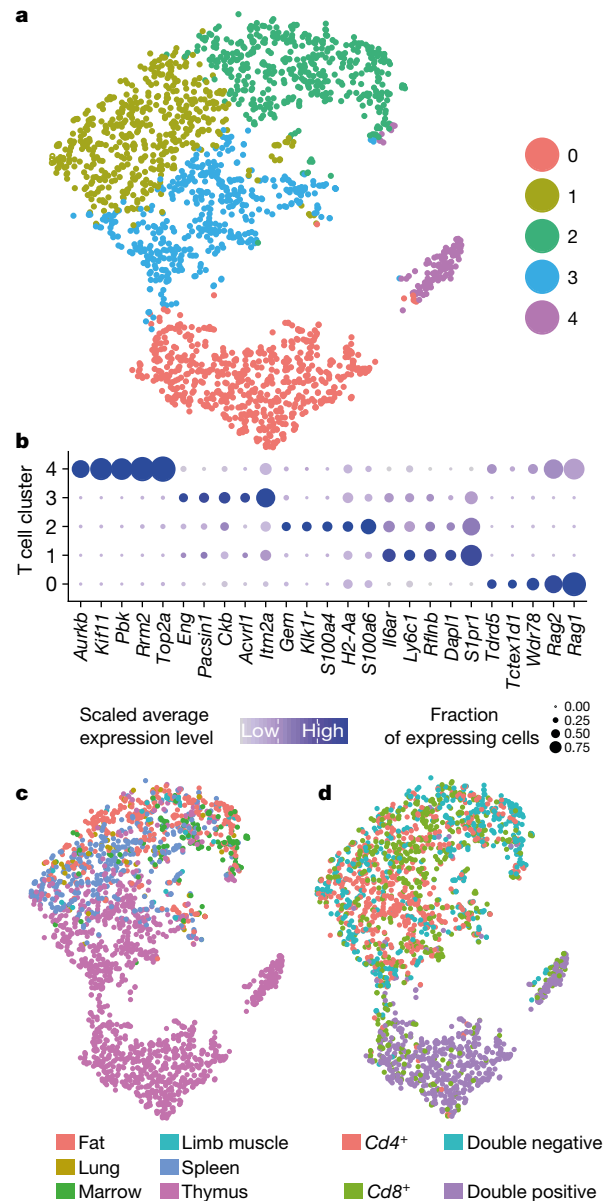


Fig. 4 | Analysis of all sorted T cells. **a**, *t*-SNE plot of all T cells coloured by cluster membership, highlighting the five identified clusters; *n* = 2,847 individual cells. **b**, Dot plot showing level of expression (colour scale) and number of expressing cells (point diameter) within each cluster of T cells. *Rflnb* is also known as *Fam101b*. **c**, *t*-SNE plot of all T cells coloured by organ of origin (fat, lung, marrow, limb muscle, spleen or thymus); *n* = 2,847 individual cells. **d**, *t*-SNE plot of all T cells coloured by classification of T cells to four categories based on expression of *Cd4* and *Cd8* (*Cd4*⁺, *Cd8*⁺, *Cd4*⁺*Cd8*⁺, *Cd4*⁻*Cd8*⁻); *n* = 2,847 individual cells.

example when many cells (tens of thousands) are required to reconstruct whole-organism architecture and developmental lineages^{25,26}. FACS-based methods generate high coverage over small cellular populations (tens to thousands), and are beneficial for enriching specific or rare cell types, and for studying subtle heterogeneity involving lowly expressed genes²⁷, alternative splicing¹⁵ and sequence variation analysis²⁸. There are opportunities to combine the two methods, such as by running sorted cells on a microfluidic-droplet platform, which could potentially accommodate both cell-type enrichment and cost factors.

Recently, a complementary scRNA-seq study across mouse organs was published²⁰. Those data contained four times as many cells and included several sample types not present in our data, such as neonatal and fetal organs, cell lines, and young adult ovary, peripheral blood,

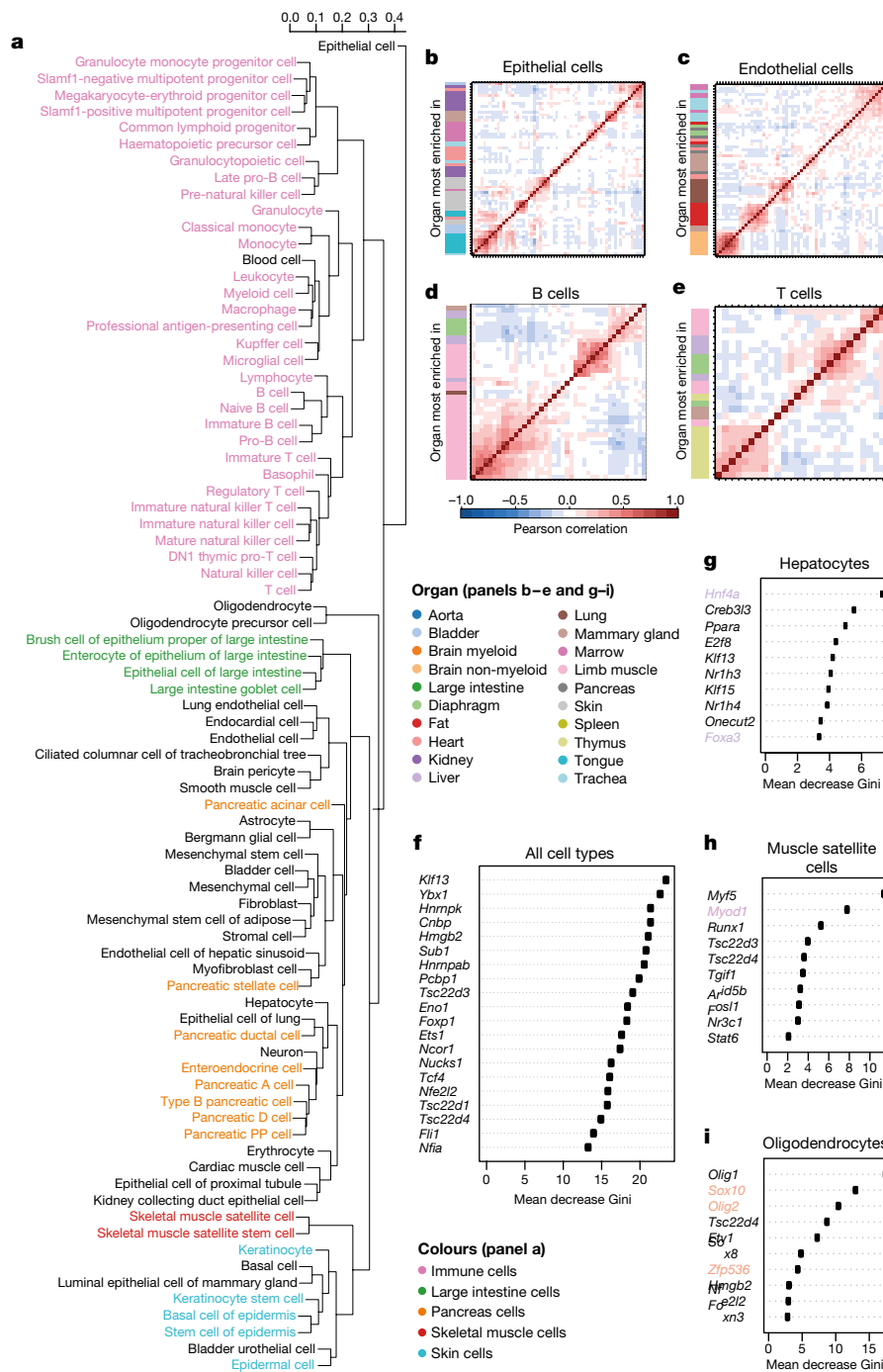


Fig. 5 | Transcription factor analysis. **a**, Dendrogram of cell types constructed with only transcription factors. **b–e**, Correlograms of top organ-specific transcription factors for epithelial cells (**b**), endothelial cells (**c**), B cells (**d**) and T cells (**e**). Row colours correspond to the organ of the most-enriched cell type; $n = 60$ randomly selected cells for each cell type.

f, Top 20 transcription factors (mean Gini importance) of the random-forest model when classifying all cell types. **g–i**, Top 10 transcription factors (mean Gini importance) of the random-forest model when classifying each cell type individually. The coloured genes correspond to transcription factors used in successful reprogramming protocols.

placenta, prostate, small intestine, stomach, testis and uterus. However, our FACS data contain four times as many genes per cell, and we analysed several organs not present in the other dataset²⁰, such as aorta, four brain regions, diaphragm, four fat types, four adult heart chambers, adult telogen and anagen skin, tongue and trachea. Additionally, several features of our study facilitate replication and cross-experiment analysis: all data, analysis and code are freely available; our web portal enables one to query gene expression in all organs simultaneously; we annotated cell types using standard cell ontology terms, thereby enabling cross-organ and cross-experiment analyses; age and sex are controlled in our data by collecting all organs from the same mice; both

sexes are represented for all organs in our data; organs were perfused, enabling the analysis of tissue-resident immune cells; and full-length transcript data make possible transcription factor, splice variant, and sequence variant analyses.

In conclusion, we have created a compendium of single-cell transcriptional measurements across 20 mouse organs. 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. Although

it is not an exhaustive characterization of all mouse organs, it does provide a rich dataset of the most highly studied organs 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.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at <https://doi.org/10.1038/s41586-018-0590-4>

Received: 20 December 2017; Accepted: 20 August 2018;

Published online: 03 October 2018

- Alberts, B. et al. *Essential Cell Biology* (W.W. Norton & Company, New York, 2016).
- 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).
- Dalerba, P. et al. Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat. Biotechnol.* **29**, 1120–1127 (2011).
- Thorsen, T., Roberts, R. W., Arnold, F. H. & Quake, S. R. Dynamic pattern formation in a vesicle-generating microfluidic device. *Phys. Rev. Lett.* **86**, 4163–4166 (2001).
- Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
- Klein, A. M. et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).
- 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).
- Wu, A. R. et al. Quantitative assessment of single-cell RNA-sequencing methods. *Nat. Methods* **11**, 41–46 (2014).
- Treutlein, B. et al. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* **509**, 371–375 (2014).
- Enge, M. et al. Single-cell analysis of human pancreas reveals transcriptional signatures of aging and somatic mutation patterns. *Cell* **171**, 321–330.e14 (2017).
- Halpern, K. B. et al. Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* **542**, 352–356 (2017).
- Haber, A. L. et al. A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333–339 (2017).
- Villani, A.-C. et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, eaah4573 (2017).
- Darmanis, S. et al. A survey of human brain transcriptome diversity at the single cell level. *Proc. Natl Acad. Sci. USA* **112**, 7285–7290 (2015).
- Gokce, O. et al. Cellular taxonomy of the mouse striatum as revealed by single-cell RNA-seq. *Cell Rep.* **16**, 1126–1137 (2016).
- Usoskin, D. et al. Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. *Nat. Neurosci.* **18**, 145–153 (2015).
- Zeisel, A. et al. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347**, 1138–1142 (2015).
- Li, H. et al. Classifying *Drosophila* olfactory projection neuron subtypes by single-cell RNA sequencing. *Cell* **171**, 1206–1220.e22 (2017).
- Bakken, T. et al. Cell type discovery and representation in the era of high-content single cell phenotyping. *BMC Bioinformatics* **18** (Suppl 17), 559 (2017).
- Han, X. et al. Mapping the mouse cell atlas by microwell-seq. *Cell* **172**, 1091–1107.e17 (2018).
- Freytag, S., Tian, L., Lonnstedt, I., Ng, M. & Bahlo, M. Comparison of clustering tools in R for medium-sized 10x Genomics single-cell RNA-sequencing data. *F1000Res* **7**, 1297 (2018).
- Holling, T. M., Schooten, E. & van Den Elsen, P. J. Function and regulation of MHC class II molecules in T-lymphocytes: of mice and men. *Hum. Immunol.* **65**, 282–290 (2004).
- Reichardt, J. & Bornholdt, S. Statistical mechanics of community detection. *Phys. Rev. E* **74**, 016110 (2006).
- van den Brink, S. C. et al. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat. Methods* **14**, 935–936 (2017).
- Alemany, A., Florescu, M., Baron, C. S., Peterson-Maduro, J. & van Oudenaarden, A. Whole-organism clone tracing using single-cell sequencing. *Nature* **556**, 108–112 (2018).
- Cao, J. et al. Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **357**, 661–667 (2017).
- Liu, Z. et al. Single-cell transcriptomics reconstructs fate conversion from fibroblast to cardiomyocyte. *Nature* **551**, 100–104 (2017).
- Darmanis, S. et al. Single-cell RNA-seq analysis of infiltrating neoplastic cells at the migrating front of human glioblastoma. *Cell Rep.* **21**, 1399–1410 (2017).

Acknowledgements We thank Sony Biotechnology for making an SH800S instrument available for this project. Some of the cell sorting/flow cytometry analysis for this project was performed using a Sony SH800S instrument in the Stanford Shared FACS Facility. Some FACS experiments used instruments in the VA Flow Cytometry Core, which is supported by the US Department of Veterans Affairs, Palo Alto Veterans Institute for Research and the National Institutes of Health. This work was supported by the Chan Zuckerberg Biohub, NIH Grant DP1 AG053015 and the NOMIS Foundation (T.W.-C.) as well as partly by the Stanford Islet Research Core in the Stanford Diabetes Research Center (P30 DK116074). We thank A. McGeever for contributions to the design of the *Tabula Muris* web portal.

Author contributions See author list for full contributions.

Competing interests The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41586-018-0590-4>.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-018-0590-4>.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

Correspondence and requests for materials should be addressed to S.R.Q., T.W.-C. or S.D.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

The Tabula Muris Consortium

Overall coordination: Nicholas Schaum¹, Jim Karkanias², Norma F. Neff², Andrew P. May², Stephen R. Quake^{2,3*}, Tony Wyss-Coray^{4,5,6*} & Spyros Darmanis^{2*}

Logistical coordination: Joshua Batson², Olga Botvinnik², Michelle B. Chen³, Steven Chen², Foad Green², Robert C. Jones³, Ashley Maynard², Lolita Penland², Angela Oliveira Pisco², Rene V. Sit², Geoffrey M. Stanley³, James T. Webber² & Fabio Zanini³

Organ collection and processing: Ankit S. Baghel¹, Isaac Bakerman^{1,7,8}, Ishita Bansal², Daniela Berdnik⁴, Biter Bilen⁴, Douglas Brownfield⁹, Corey Cain¹⁰, Michelle B. Chen³, Steven Chen², Min Cho², Giana Cirolia², Stephanie D. Conley¹, Spyros Darmanis², Aaron Demers², Kubilay Demir^{1,11}, Antoine de Morree⁴, Tessa Divita², Haley du Bois⁴, Laughing Bear Torrez Dulgeroff¹, Hamid Ebadi², F. Hernán Espinoza⁹, Matt Fish^{1,11,12}, Qiang Gan⁴, Benson M. George¹, Astrid Gillich⁹, Foad Green², Geraldine Genetiano², Xueying Gu¹², Gungsagar S. Gulati¹, Yan Hang¹², Shayan Hosseinzadeh², Albin Huang⁴, Tal Iram⁴, Taichi Isobe¹, Feather Ives², Robert C. Jones³, Kevin S. Kao¹, Guruswamy Karnam¹³, Aaron M. Kershner¹, Bernhard M. Kiss^{1,14}, William Kong¹, Maya E. Kumar^{15,16}, Jonathan Y. Lam¹², Davis P. Lee⁶, Song E. Lee⁴, Guang Li¹⁷, Qingyun Li¹⁸, Ling Liu⁴, Annie Lo², Wan-Jin Lu^{1,9}, Anoop Manjunath¹, Andrew P. May², Kaia L. May², Oliver L. May², Ashley Maynard², Marina McKay², Ross J. Metzger^{19,20}, Marco Mignardi³, Dullei Min²¹, Ahmad N. Nabhan⁹, Norma F. Neff², Katharine M. Ng³, Joseph Noh¹, Rasika Patkar¹³, Weng Chuan Peng¹², Lolita Penland², Robert Puccinelli², Eric J. Rulifson¹², Nicholas Schaum¹, Shaheen S. Sikandar¹, Rahul Sinha^{1,22,23,24}, Rene V. Sit², Krzysztof Szade^{1,25}, Weilun Tan², Cristina Tato², Krissie Tellez¹², Kyle J. Travaglini⁹, Carolina Tropini²⁶, Lucas Waldburger², Linda J. van Weele¹, Michael N. Wosczyzna⁴, Jinyi Xiang¹, Soso Xue³, Justin Youngyunpipatkul², Fabio Zanini³, Macy E. Zardeneta⁶, Fan Zhang^{19,20} & Lu Zhou¹⁸

Library preparation and sequencing: Ishita Bansal², Steven Chen², Min Cho², Giana Cirolia², Spyros Darmanis², Aaron Demers², Tessa Divita², Hamid Ebadi², Geraldine Genetiano², Foad Green², Shayan Hosseinzadeh², Feather Ives², Annie Lo², Andrew P. May², Ashley Maynard², Marina McKay², Norma F. Neff², Lolita Penland², Rene V. Sit², Weilun Tan², Lucas Waldburger² & Justin Youngyunpipatkul²

Computational data analysis: Joshua Batson², Olga Botvinnik², Paola Castro², Derek Croote³, Spyros Darmanis², Joseph L. DeRisi^{2,27}, Jim Karkanias², Angela Oliveira Pisco², Geoffrey M. Stanley³, James T. Webber² & Fabio Zanini³

Cell type annotation: Ankit S. Baghel¹, Isaac Bakerman^{1,7,8}, Joshua Batson², Biter Bilen⁴, Olga Botvinnik², Douglas Brownfield⁹, Michelle B. Chen³, Spyros Darmanis², Kubilay Demir^{1,11}, Antoine de Morree⁴, Hamid Ebadi², F. Hernán Espinoza⁹, Matt Fish^{1,11,12}, Qiang Gan⁴, Benson M. George¹, Astrid Gillich⁹, Xueying Gu¹², Gungsagar S. Gulati¹, Yan Hang¹², Albin Huang⁴, Tal Iram⁴, Taichi Isobe¹, Guruswamy Karnam¹³, Aaron M. Kershner¹, Bernhard M. Kiss^{1,14}, William Kong¹, Christin S. Kuo^{9,11,21}, Jonathan Y. Lam¹², Benoit Lehallier⁴, Guang Li¹⁷, Qingyun Li¹⁸, Ling Liu⁴, Wan-Jin Lu^{1,9}, Dullei Min²¹, Ahmad N. Nabhan⁹, Katharine M. Ng³, Patricia K. Nguyen^{1,7,8,17}, Rasika Patkar¹³, Weng Chuan Peng¹², Lolita Penland², Eric J. Rulifson¹², Nicholas Schaum¹, Shaheen S. Sikandar¹, Rahul Sinha^{1,22,23,24}, Krzysztof Szade^{1,25}, Serena Y. Tan²², Krissie Tellez¹², Kyle J. Travaglini⁹

Carolina Tropini²⁶, Linda J. van Weele¹, Bruce M. Wang¹³, Michael N. Wosczyzna⁴, Jinyi Xiang¹, Hanadie Yousef⁴ & Lu Zhou¹⁸

Writing group: Joshua Batson², Olga Botvinnik², Steven Chen², Spyros Darmanis², Foad Green², Andrew P. May², Ashley Maynard², Angela Oliveira Pisco², Stephen R. Quake^{2,3}, Nicholas Schaum¹, Geoffrey M. Stanley³, James T. Webber², Tony Wyss-Coray^{4,5,6} & Fabio Zanini³

Supplemental text writing group: Philip A. Beachy^{1,9,11,12}, Charles K. F. Chan²⁸, Antoine de Morree⁴, Benson M. George¹, Gungsagar S. Gulati¹, Yan Hang¹², Kerwyn Casey Huang^{2,3,26}, Tal Iram⁴, Taichi Isobe¹, Aaron M. Kershner¹, Bernhard M. Kiss^{1,14}, William Kong¹, Guang Li¹⁷, Qingyun Li¹⁸, Ling Liu⁴, Wan-Jin Lu^{1,9}, Ahmad N. Nabhan⁹, Katharine M. Ng³, Patricia K. Nguyen^{1,7,8,17}, Weng Chuan Peng¹², Eric J. Rulifson¹², Nicholas Schaum¹, Shaheen S. Sikandar¹, Rahul Sinha^{1,22,23,24}, Krzysztof Szade^{1,25}, Kyle J. Travaglini⁹, Carolina Tropini²⁶, Bruce M. Wang¹³, Kenneth Weinberg²¹, Michael N. Wosczyzna⁴, Sean M. Wu¹⁷ & Hanadie Yousef⁴

Principal investigators: Ben A. Barres¹⁸, Philip A. Beachy^{1,9,11,12}, Charles K. F. Chan²⁸, Michael F. Clarke¹, Spyros Darmanis², Kerwyn Casey Huang^{2,3,26}, Jim Karkanias², Seung K. Kim^{12,29}, Mark A. Krasnow^{9,11}, Maya E. Kumar^{15,16}, Christin S. Kuo^{9,11,21}, Andrew P. May², Ross J. Metzger^{19,20}, Norma F. Neff², Roel Nusse^{9,11,12}, Patricia K. Nguyen^{1,7,8,17}, Thomas A. Rando^{4,5,6}, Justin Sonnenburg^{2,26}, Bruce M. Wang¹³, Kenneth Weinberg²¹, Irving L. Weissman^{1,22,23,24}, Sean M. Wu^{1,7,17}, Stephen R. Quake^{2,3} & Tony Wyss-Coray^{4,5,6}

¹Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA, USA. ²Chan Zuckerberg Biohub, San Francisco, CA, USA. ³Department of Bioengineering, Stanford University, Stanford, CA, USA. ⁴Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, CA, USA. ⁵Paul F. Glenn Center for the Biology of Aging, Stanford University School of Medicine, Stanford, CA, USA. ⁶Center for Tissue Regeneration, Repair, and Restoration, VA Palo Alto Healthcare System, Palo Alto, CA, USA. ⁷Stanford Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA, USA. ⁸Department of Medicine, Division of Cardiology, Stanford University School of Medicine, Stanford, CA, USA. ⁹Department of Biochemistry, Stanford University School of Medicine, Stanford, CA, USA. ¹⁰Flow Cytometry Core, VA Palo Alto Healthcare System, Palo Alto, CA, USA. ¹¹Howard Hughes Medical Institute, Stanford University, Stanford, CA, USA. ¹²Department of Developmental Biology, Stanford University School of Medicine, Stanford, CA, USA. ¹³Department of Medicine and Liver Center, University of California San Francisco, San Francisco, CA, USA. ¹⁴Department of Urology, Stanford University School of Medicine, Stanford, CA, USA. ¹⁵Sean N. Parker Center for Asthma and Allergy Research, Stanford University School of Medicine, Stanford, CA, USA. ¹⁶Department of Medicine, Division of Pulmonary and Critical Care, Stanford University School of Medicine, Stanford, CA, USA. ¹⁷Department of Medicine, Division of Cardiovascular Medicine, Stanford University, Stanford, CA, USA. ¹⁸Department of Neurobiology, Stanford University School of Medicine, Stanford, CA, USA. ¹⁹Vera Moulton Wall Center for Pulmonary and Vascular Disease, Stanford University School of Medicine, Stanford, CA, USA. ²⁰Department of Pediatrics, Division of Cardiology, Stanford University School of Medicine, Stanford, CA, USA. ²¹Department of Pediatrics, Pulmonary Medicine, Stanford University School of Medicine, Stanford, CA, USA. ²²Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA. ²³Ludwig Center for Cancer Stem Cell Research and Medicine, Stanford University School of Medicine, Stanford, CA, USA. ²⁴Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA, USA. ²⁵Department of Medical Biotechnology, Faculty of Biochemistry, Biophysics and Biotechnology, Jagiellonian University, Kraków, Poland. ²⁶Department of Microbiology & Immunology, Stanford University School of Medicine, Stanford, CA, USA. ²⁷Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, CA, USA. ²⁸Department of Surgery, Division of Plastic and Reconstructive Surgery, Stanford University, Stanford, CA, USA. ²⁹Department of Medicine and Stanford Diabetes Research Center, Stanford University, Stanford, CA, USA. *e-mail: quake@stanford.edu; twc@stanford.edu; spyros.darmanis@czbiohub.org

METHODS

Mice and organ collection. Four 10–15 week old male and four virgin female C57BL/6JN mice were shipped from the National Institute on Aging colony at Charles River (housed at 67–73 °F) to the Veterinary Medical Unit (VMU; housed at 68–76 °F) 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. Before tissue collection, mice were placed in sterile collection chambers at 8 am for 15 min to collect fresh fecal pellets. After anaesthesia with 2.5% v/v Avertin, mice were weighed, shaved, and blood was drawn via cardiac puncture before transcardial perfusion with 20 ml PBS. Mesenteric adipose tissue was then immediately collected to avoid exposure to the liver and pancreas perfusate, which negatively affects cell sorting. Isolating viable single cells from both the pancreas and the liver of the same mouse was not possible; therefore, two males and two females were used for each. Whole organs were then dissected in the following order: large intestine, spleen, thymus, trachea, tongue, brain, heart, lung, kidney, gonadal adipose tissue, bladder, diaphragm, limb muscle (tibialis anterior), skin (dorsal), subcutaneous adipose tissue (inguinal pad), mammary glands (fat pads 2, 3 and 4), brown adipose tissue (interscapular pad), aorta and bone marrow (spine and limb bones). Organ collection concluded by 10 am. After single-cell dissociation as described below, cell suspensions were either used for FACS of individual cells into 384-well plates, or for preparation of the microfluidic droplet library. 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 Supplementary Information.

Sample size, randomization and blinding. No sample size choice was performed before the study. Randomization and blinding were not performed: the authors were aware of all data and metadata-related variables during the entire course of the study.

Single-cell methods. Lysis plate preparation. Lysis plates were created by dispensing 0.4 μ l lysis buffer (0.5 U Recombinant RNase Inhibitor (Takara Bio, 2313B), 0.0625% TritonTM X-100 (Sigma, 93443-100ML), 3.125 mM dNTP mix (Thermo Fisher, R0193), 3.125 μ M Oligo-dT₃₀VN (Integrated DNA Technologies, 5'AAGCAGTGGTATCAACGCAGAGTACT₃₀VN-3') and 1:600,000 ERCC RNA spike-in mix (Thermo Fisher, 4456740)) into 384-well hard-shell PCR plates (Bio-Rad HSP3901) using a Tempest liquid handler (Formulatrix). 96-well lysis plates were also prepared with 4 μ l lysis buffer. All plates were sealed with AlumaSeal CS Films (Sigma-Aldrich Z722634) and spun down (3,220g, 1 min) and snap-frozen on dry ice. Plates were stored at –80 °C until sorting.

FACS. After dissociation, single cells from each organ and tissue were isolated into 384- or 96-well plates via FACS. Most organs 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. Limb muscle and diaphragm were sorted into 384-well plates on an Aria III (Becton Dickinson) sorter. The last two columns of each 384 well plate were intentionally left as blanks. For most organs, single cells were selected with forward scatter, and dead cells and common cell types were excluded with a single colour channel. Combinations of fluorescent antibodies were used for most organs to enrich for rare cell populations (see Supplementary Information), but some were stained only for viable cells. Colour compensation was used whenever necessary. On the SH800, the highest purity setting ('Single cell') was used for all but the rarest cell types, for which the 'Ultrapur' setting was used. Sorters were calibrated using FACS buffer every day before collecting any cells, and also after every eight sorted plates. For a typical sort, 1–3 ml of 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 and debris concentrations. The pressure was then adjusted, flow paused, the first destination plate unsealed and loaded, and sorting started. If a cell suspension was too concentrated, it was diluted using FACS buffer or 1X PBS. For some cell types, such as hepatocytes, 96-well plates were used because it was not possible to sort individual cells accurately into 384-well plates. Immediately after sorting, plates were sealed with a pre-labelled aluminium seal, centrifuged, and flash frozen on dry ice. On average, each 384-well plate took 8 min to sort.

cDNA synthesis and library preparation. cDNA synthesis was performed using the Smart-seq2 protocol^{7,8}. In brief, 384-well plates containing single-cell lysates were thawed on ice followed by first-strand synthesis. 0.6 μ l of reaction mix (16.7 U μ l⁻¹ SMARTScribe Reverse Transcriptase (Takara Bio, 639538), 1.67 U μ l⁻¹ Recombinant RNase Inhibitor (Takara Bio, 2313B), 1.67X First-Strand Buffer (Takara Bio, 639538), 1.67 μ M TSO (Exiqon, 5'-AAGCAGTGGTATCAACGCAGAGTGAATrGrGrG-3'), 8.33 mM dithiothreitol (Bioworld, 40420001-1), 1.67 M Betaine (Sigma, B0300-5VL) and 10 mM MgCl₂ (Sigma, M1028-10X1ML)) was added to each well using a Tempest liquid handler. Reverse transcription was carried out by incubating wells on a ProFlex

2 × 384 thermal-cycler (Thermo Fisher) at 42 °C for 90 min, and stopped by heating at 70 °C for 5 min.

Subsequently, 1.5 μ l of PCR mix (1.67X KAPA HiFi HotStart ReadyMix (Kapa Biosystems, KK2602), 0.17 μ M IS PCR primer (IDT, 5'-AAGCAGTGGTATCAACGCAGAGT-3'), and 0.038 U μ l⁻¹ 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 min, 2) 95 °C for 3 min, 3) 23 cycles of 98 °C for 20 s, 67 °C for 15 s and 72 °C for 4 min, and 4) 72 °C for 5 min.

The amplified product was diluted with a ratio of 1 part cDNA to 10 parts 10 mM 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 μ l⁻¹, and, after linear regression of the values obtained from the Quant-iT dsDNA standard curve, the R² 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 μ l⁻¹ and a final volume of 0.4 μ l using an Echo 550 acoustic liquid dispenser (Labcyte).

Illumina sequencing libraries were prepared as described previously¹⁴. In brief, 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 μ l Nextera tagmentation DNA buffer (Illumina) and 0.4 μ l Tn5 enzyme (Illumina), then incubated at 55 °C for 10 min. The reaction was stopped by adding 0.4 μ l Neutralize Tagment Buffer (Illumina) and centrifuging at room temperature at 3,220g for 5 min. Indexing PCR reactions were performed by adding 0.4 μ l of 5 μ M i5 indexing primer, 0.4 μ l of 5 μ M i7 indexing primer, and 1.2 μ l 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 min, 2) 95 °C for 30 s, 3) 12 cycles of 95 °C for 10 s, 55 °C for 30 s and 72 °C for 1 min, and 4) 72 °C for 5 min.

Library pooling, quality control and sequencing. After 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. A 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 × 100-bp paired-end reads and 2 × 8-bp or 2 × 12-bp 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 the GemCode Single-Cell Instrument (10x Genomics), and scRNA-seq libraries were constructed as per the 10x Genomics protocol using GemCode Single-Cell 3' Gel Bead and Library V2 Kit. In brief, single cell suspensions were examined using an inverted microscope, and if sample quality was deemed satisfactory, the sample was diluted in PBS with 2% FBS to a concentration of 1000 cells per μ l. If cell suspensions contained cell aggregates or debris, two additional washes in PBS with 2% FBS at 300g for 5 min at 4 °C were performed. Cell concentration was measured either with a Moxi GO II (Orflo Technologies) or a haemocytometer. Cells were loaded in each channel with a target output of 5,000 cells per sample. All reactions were performed in the 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 libraries were evaluated on a Fragment Analyzer using a High Sensitivity NGS Analysis Kit (Advanced Analytical). The average fragment length of 10x cDNA libraries was quantitated on a Fragment Analyzer (AATI), and by qPCR with the Kapa Library Quantification kit for Illumina. Each library was diluted to 2 nM, and equal volumes of 16 libraries were pooled for each NovaSeq sequencing run. Pools were sequenced with 100 cycle run kits with 26 bases for Read 1, 8 bases for Index 1, and 90 bases for Read 2 (Illumina 20012862). A 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 de-multiplexed using bcl-2fastq version 2.19.0.316. Reads were aligned using the mm10plus genome using STAR version 2.5.2b with parameters TK. 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 droplet platform were de-multiplexed and aligned using Cell Ranger version 2.0.1, available from 10x Genomics with default parameters.

Clustering. Standard procedures for filtering, variable gene selection, dimensionality reduction and clustering were performed using the Seurat package version 2.2.1. A detailed worked example, including the mathematical formulae for each operation, is in the Organ Annotation Vignette. The parameters that were tuned on a per-tissue basis (resolution and number of principal components (PCs)) can be viewed in the tissue-specific Rmd files available on GitHub. For each tissue and each sequencing method (FACS and microfluidic droplet), the following steps were performed:

1. Cells were lexicographically sorted by cell ID to ensure reproducibility.
2. Cells with fewer than 500 detected genes were excluded. (A gene counts as detected if it has at least one read mapping to it). Cells with fewer than 50,000 reads (FACS) or 1,000 UMI (microfluidic droplet) were excluded.
3. Counts were log-normalized for each cell using the natural logarithm of $1 +$ counts per million (for FACS) or $1 +$ counts per ten thousand (for microfluidic droplet).
4. Variable genes were selected using a threshold (0.5) for the standardized log dispersion, in which the standardization was performed separately according to binned values of log mean expression.
5. The variable genes were projected onto a low-dimensional subspace using principal component analysis. The number of principal components was selected on the basis of inspection of the plot of variance explained.
6. A shared-nearest-neighbours graph was constructed on the basis of the Euclidean distance in the low-dimensional subspace spanned by the top principal components. Cells were clustered using a variant of the Louvain method that includes a resolution parameter in the modularity function¹³.
7. Cells were visualized using a 2-dimensional t-distributed Stochastic Neighbour Embedding of the PC-projected data.
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 Extended Data.
9. When clusters appeared to be mixtures of cell types, they were refined either by increasing the resolution parameter for 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-droplet-processed cells to produce an unbiased clustering.

Heterogeneity score. Let C be a cluster, decomposed into annotated cell types $C = T_1 \cup \dots \cup T_k$. For each pair of cell types T_i, T_j , we compute the average distance between their members: $d_{ij} = \frac{1}{|T_i| |T_j|} \sum_{x \in T_i, y \in T_j} |x - y|$. The heterogeneity score C is the maximum of those distances over cell types T with at least five cells. For the FACS data, the vector x for a cell is the PC-projection from step 5 above. Extended Data Fig. 9 contains heat maps of the cell-type distance matrix d_{ij} for select clusters and a bar plot of the heterogeneity scores for all clusters containing several cell types.

Differential expression overlap analysis. For FACS and microfluidic droplet data, differential expression analysis for each organ was performed using a Wilcoxon rank-sum test as implemented in the 'FindAllMarkers' function of the Seurat package. Differential expression was performed between cell ontology groups and resulted in a list of differentially expressed genes ($\ln(\text{FoldChange}) > 0.25$) between each cell ontology group and all other ontology groups of the same organ. For microwell-seq we used the corresponding published lists for each cell type and for every organ. We then assessed the overlap of those lists between the three methods. As the nomenclature is not identical, the analysis was performed between cell types that could be matched with a certain degree of confidence between the three methods (Supplementary Table 2).

Correlating bulk gene expression profiles. For the 33 cell populations shared between FACS and microfluidic droplets, the average gene-expression profile of each population was calculated. The quality of such a bulk gene-expression profile depends on the total number of detected molecules. FACS detects more molecules per cell, but fewer cells. Microfluidic droplets detect fewer molecules per cell, but more cells. To assess the agreement between methods on annotated cell types, Pearson correlation was used on the log expression profiles of each shared cell population. (Only genes present at 1 count per million or greater in at least one

of the datasets were considered. A pseudocount of 1 count per million was added before taking logarithms.)

Calculation of dissociation scores. For each organ, principal component analysis was performed on a subset of 140 dissociation-related genes²³. The first principal component was used as the 'dissociation score' as it corresponds to the variance within these genes.

Defining cell type-enriched transcription factors. Transcription factors were defined as the 1,140 genes annotated by the Gene Ontology term 'DNA binding transcription factor activity', downloaded from the Mouse Genome Informatics database (<http://www.informatics.jax.org/mgihome/GO/project.shtml>, accessed on 10 November 2017). Cell types were defined as unique combinations of cell ontology and organ annotation (for example, Lung_Endothelial_cell). All analyses were performed on the full dataset, except the correlograms for which the data was subsampled by randomly selecting 60 cells from each cell type. Enriched transcription factors were defined by the Seurat FindMarkers function with the Wilcoxon significance test for the target cell type against the all of the rest of the cell types combined. These were filtered by $p_val < 10^{-3}$, $avg_diff > 0.2$, $pct.1 - pct.2 > 0.1$ (per cent detected difference > 0.1), and $pct.1 > 0.3$ (detected in $> 30\%$ of target cells).

Cell-type comparisons between methods using cell ontology classes. We used the OntologyX R package family version 2.4 (libraries ontologyIndex, ontology-Plot, and ontologySimilarity) to draw the representative cell ontology dendrograms (function onto_plot). To compute the tanglegram (function tanglegram from dendextend R package version 1.8) we used the dendrogram created from all expressed genes as the reference for comparisons to the dendrograms produced using particular gene ontology cellular functions (transcription factors, cell surface markers, RNA splicing factors). The entanglement scores were calculated using the step2side method (function untangle from dendextend R package). Entanglement is a measure of alignment between two dendrograms. The entanglement score ranges from 0 (exact alignment) to 1 (no alignment)²⁹.

Defining transcription factor networks with random forests. We used random forests (a classifier that combines many single decision trees) to calculate the importance of each gene for defining cell types³⁰. The varSelRF R package version 0.7–8 uses the out-of-bag error as the minimization criterion and carries out variable elimination with random forests by successively eliminating the least important variables (with importance as returned from the random forest analysis). The algorithm iteratively fits random forests, at each iteration building a new forest after discarding those variables (genes) with the smallest variable importance; the selected set of genes is the one that yields the smallest out-of-bag error rate. This leads to the selection of small sets of non-redundant variables.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

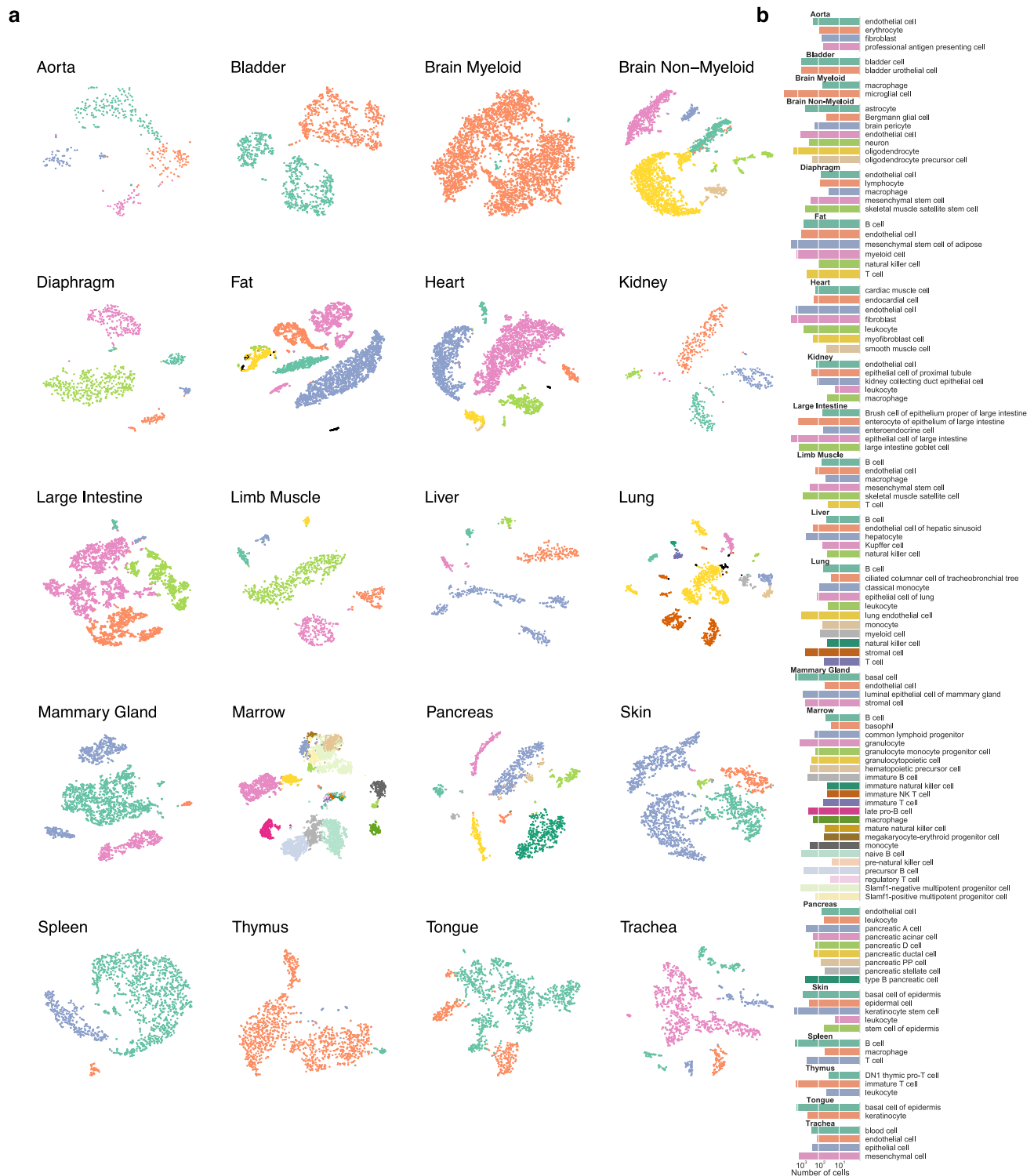
Code availability. All code used for analysis is available on GitHub (<https://github.com/czbiohub/tabula-muris>).

Data availability

All data, protocols and analysis scripts from the *Tabula Muris* are shared as a public resource (<http://tabula-muris.ds.czbiohub.org/>). Gene counts and metadata for FACS (<https://doi.org/10.6084/m9.figshare.5829687.v7>) and microfluidic droplets (<https://doi.org/10.6084/m9.figshare.5968960.v2>) from all single cells along with all produced R objects (<https://doi.org/10.6084/m9.figshare.5821263.v1>), as well as FACS Index data (<https://doi.org/10.6084/m9.figshare.5975392>) are accessible on Figshare (https://figshare.com/projects/Tabula_Muris_Transcriptomic_characterization_of_20_organ_and_tissues_from_Mus_musculus_at_single_cell_resolution/27733), and raw data are available from the Gene Expression Omnibus (GSE109774).

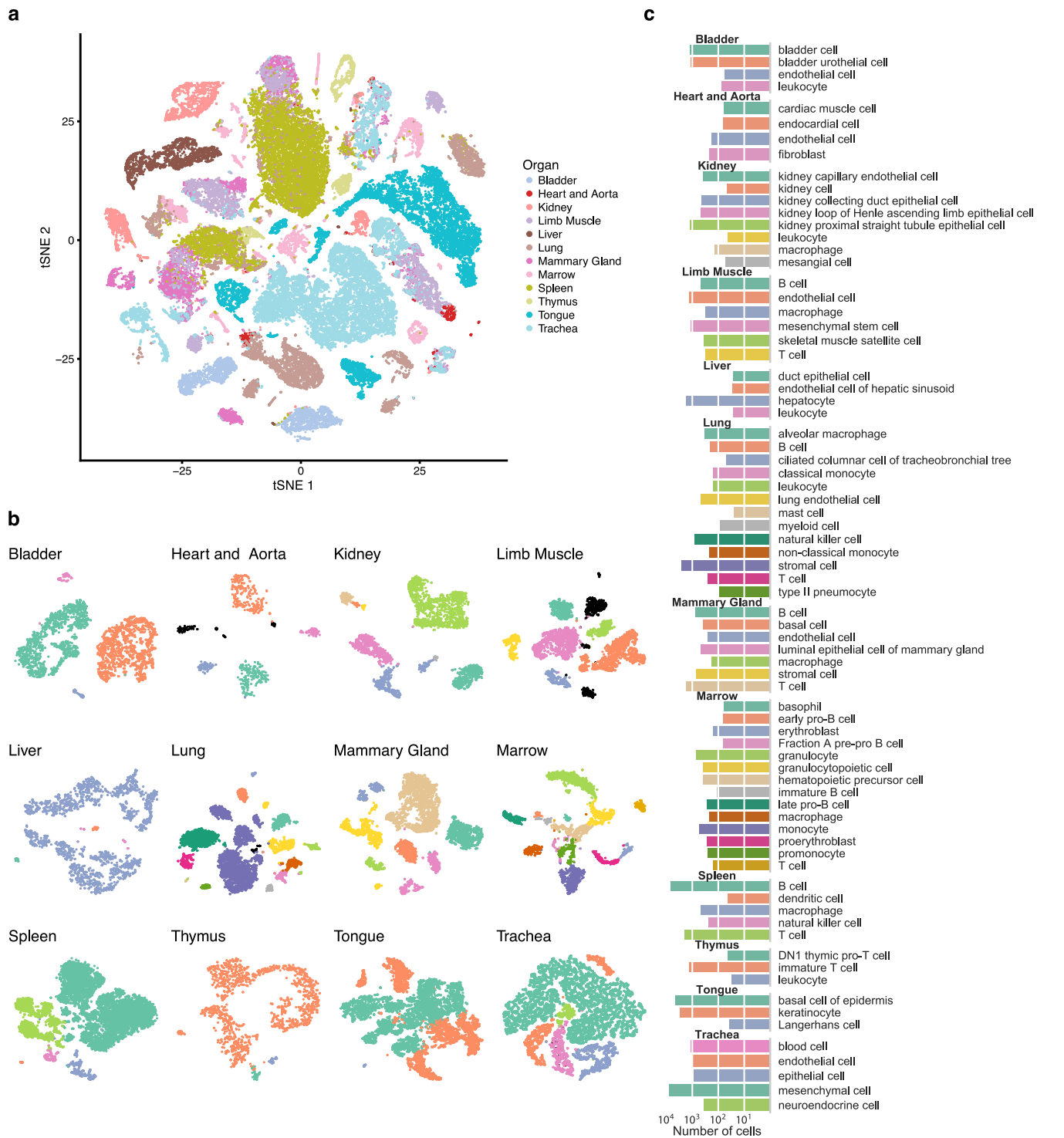
29. Kassambara, A. *Practical guide to cluster analysis in R: unsupervised machine learning* 1st edn (CreateSpace, North Charleston, 2017).

30. Díaz-Uriarte, R. & Alvarez de Andrés, S. Gene selection and classification of microarray data using random forest. *BMC Bioinformatics* **7**, 3 (2006).



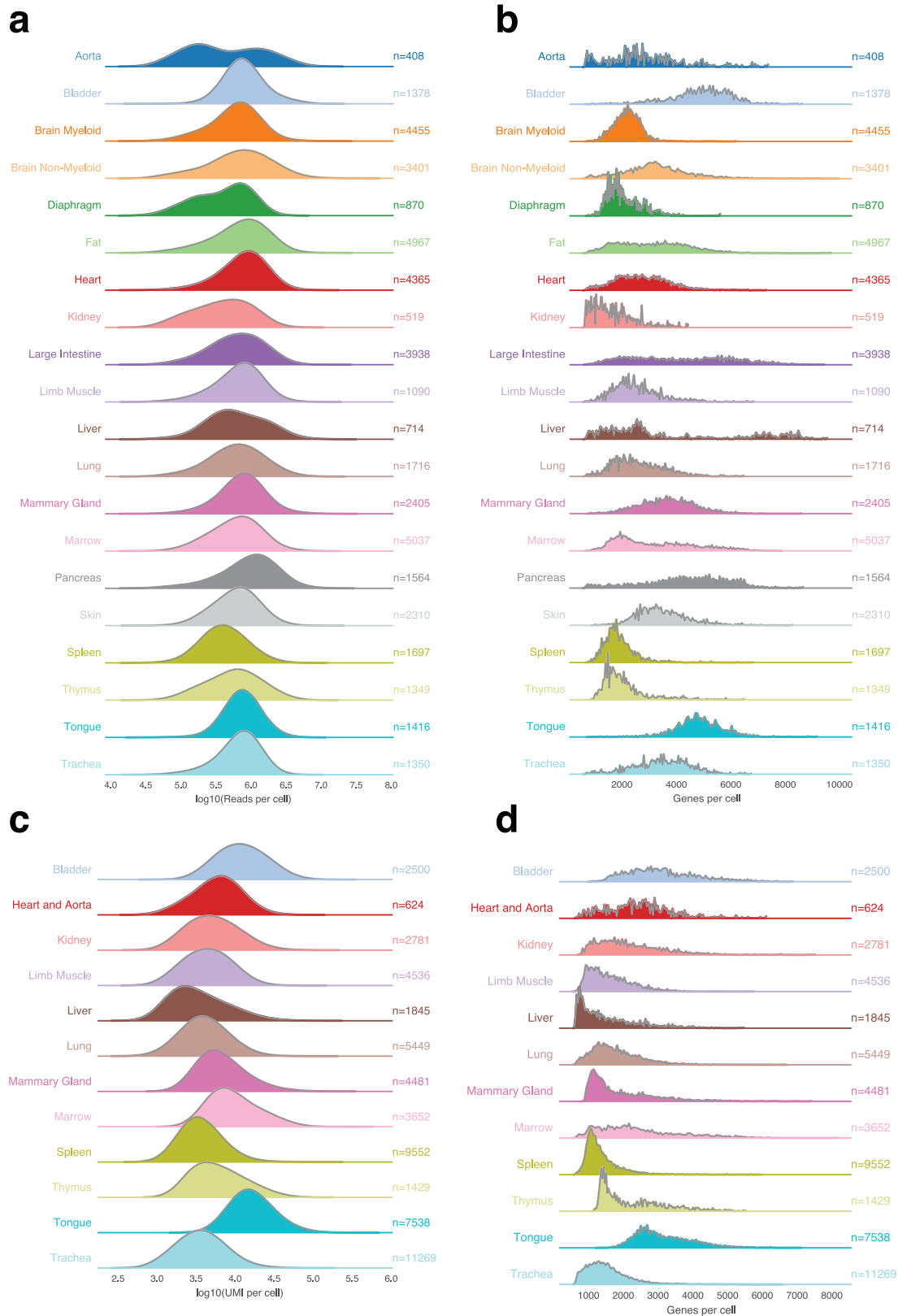
Extended Data Fig. 1 | The number and type of FACS cells that compose each organ. a, Cells for each organ visualized with *t*-SNE, coloured by cell type. Cell types were determined by differential gene expression of known

markers between clusters. **b**, Bar plots quantifying the number of each annotated cell type. Cell type colours match their respective *t*-SNE plot.



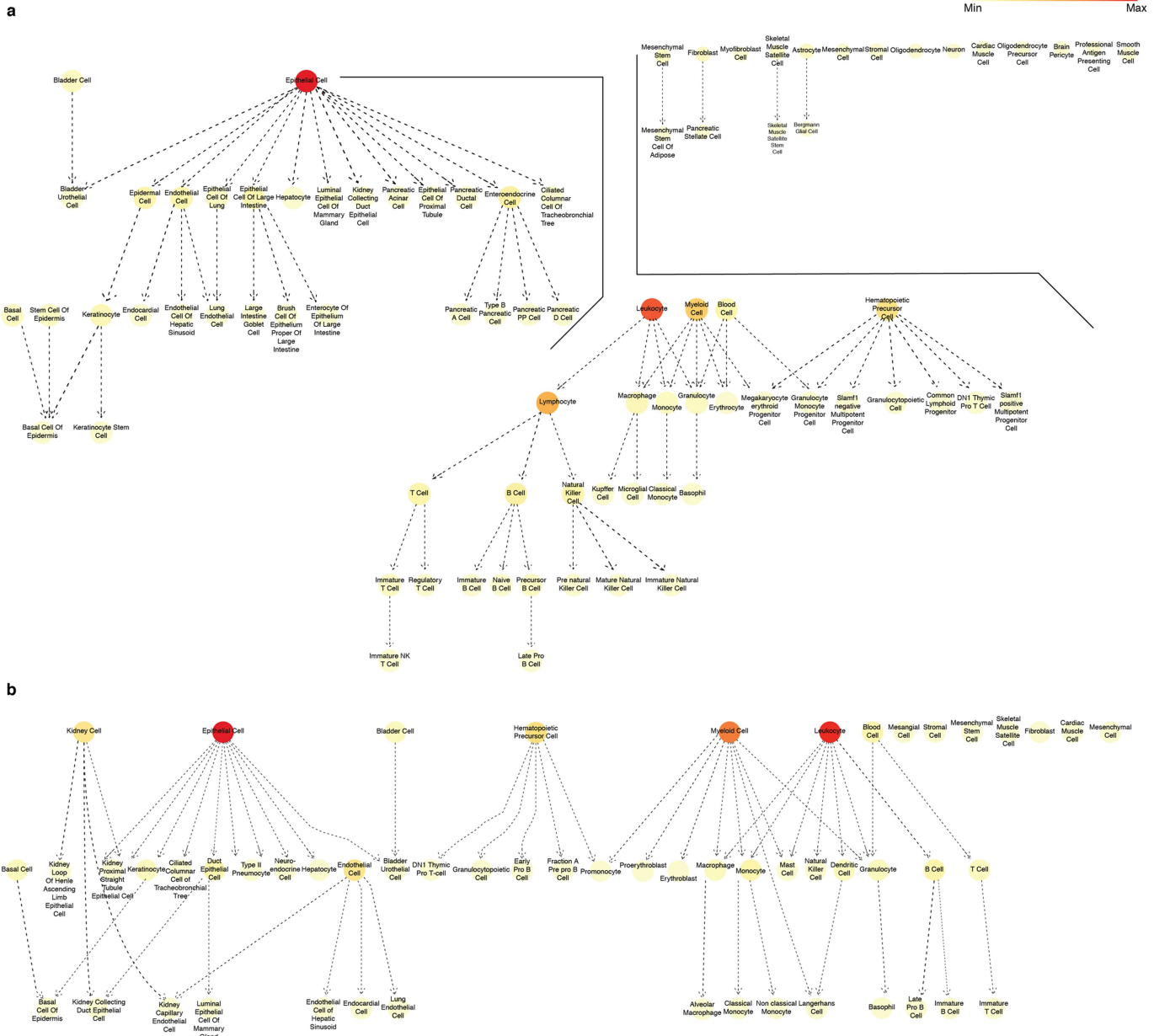
Extended Data Fig. 2 | The number and type of microfluidic cells that compose each organ. **a**, *t*-SNE plot of all cells collected by the microfluidic-droplet method, coloured by organ, overlaid with the predominant cell type that composes each cluster. **b**, Cells for each organ

visualized with *t*-SNE, coloured by cell type. Cell types were determined by differential gene expression of known markers between clusters. **c**, Bar plots quantifying the number of each annotated cell type. Cell type colours match their respective *t*-SNE plot.



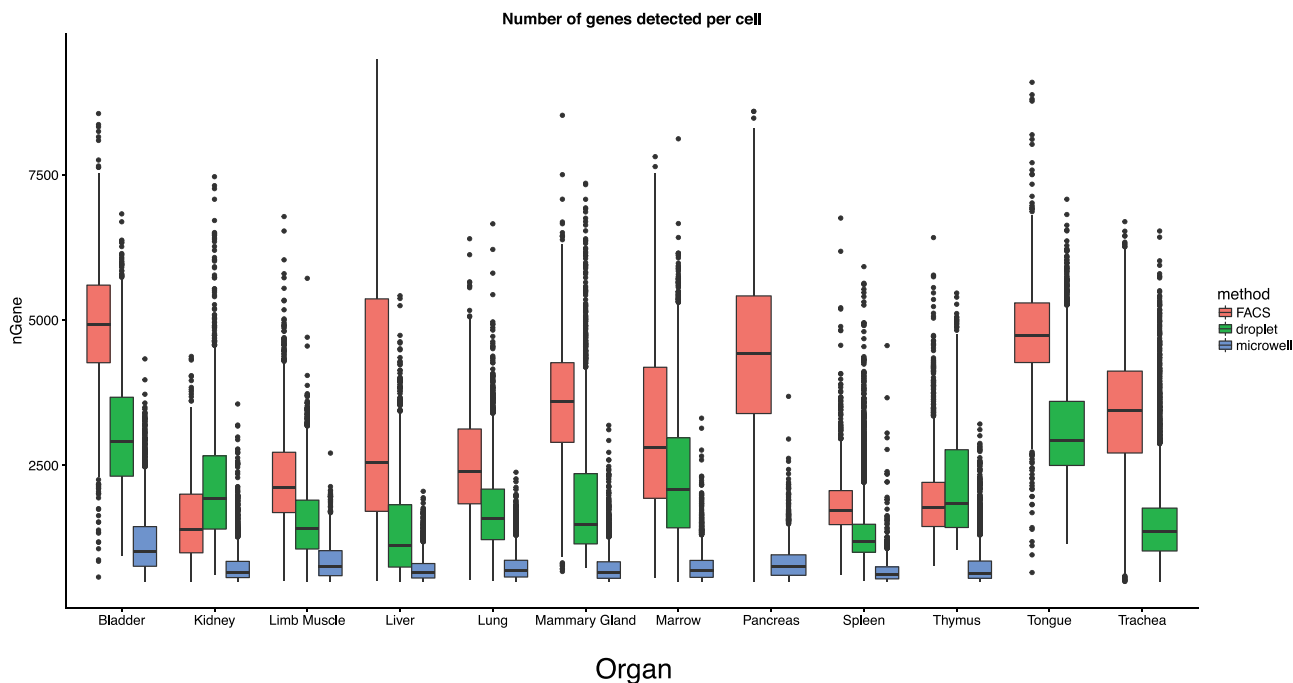
Extended Data Fig. 3 | The number of reads, UMIs and genes detected per cell for each organ. a, c, Histograms, for each organ, of the number of reads per cell (FACS) (a) and UMIs per cell (microfluidic droplet) (c).

b, d, Histogram of the number of genes detected per cell for each organ from the FACS method (b), and the microfluidic-droplet method (d).

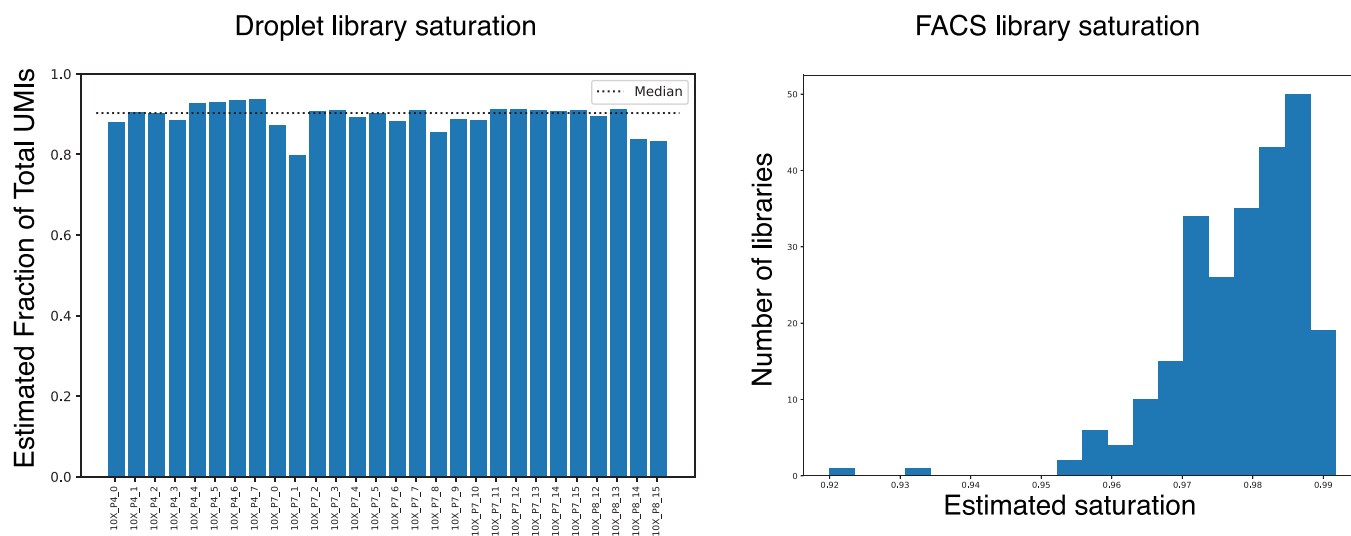


Extended Data Fig. 4 | Graphical representation of cell ontology class representation. a, b, Datasets from the FACS method (a) and the microfluidic-droplet method (b), coloured by the relative amount of each cell type in each dataset.

a

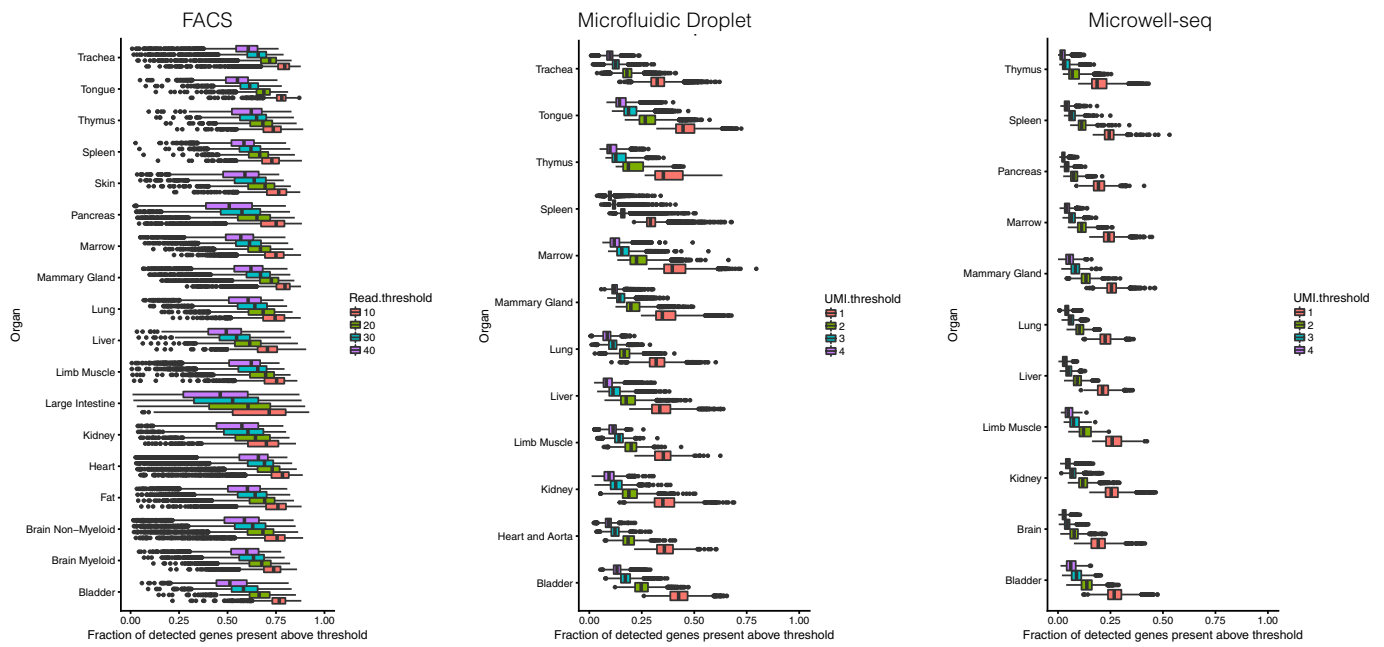


b



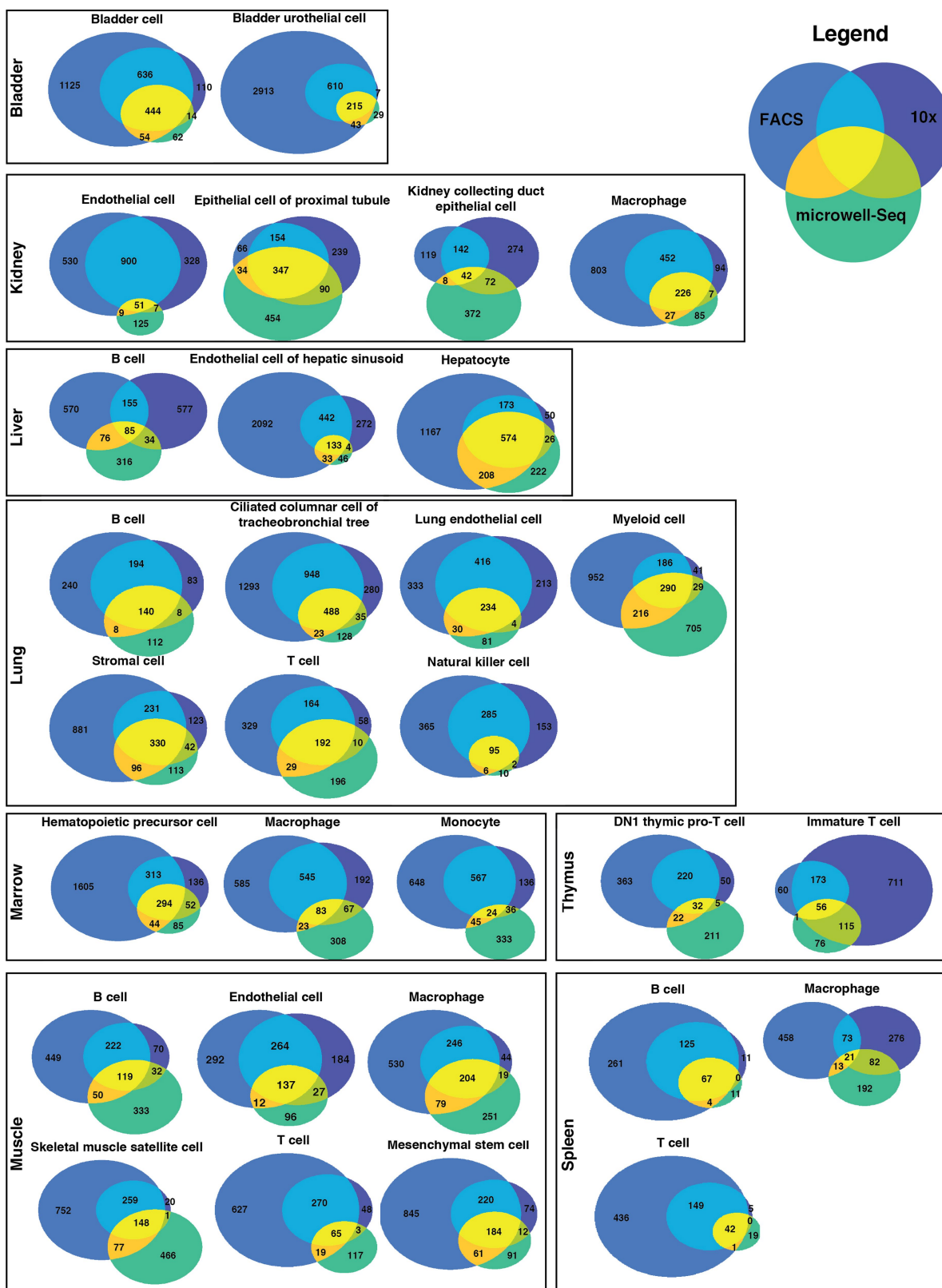
Extended Data Fig. 5 | Methodological comparison of detected genes and library saturation. **a.** The number of genes detected (threshold of >0 reads or UMIs per cell) by FACS (red; $n = 21,105$ individual cells), microfluidic-droplet (green; $n = 55,032$ individual cells) and microwell-seq (blue; $n = 25,891$ individual cells) methods²⁰. **b.** Library saturation

fraction for all microfluidic-droplet libraries. Dotted horizontal line demarcates the median saturation (around 0.9). **c.** Library saturation for all FACS libraries. Saturation was calculated using the number of detected genes while downsampling the number of reads per library. Summary statistics are contained in Supplementary Table 6.



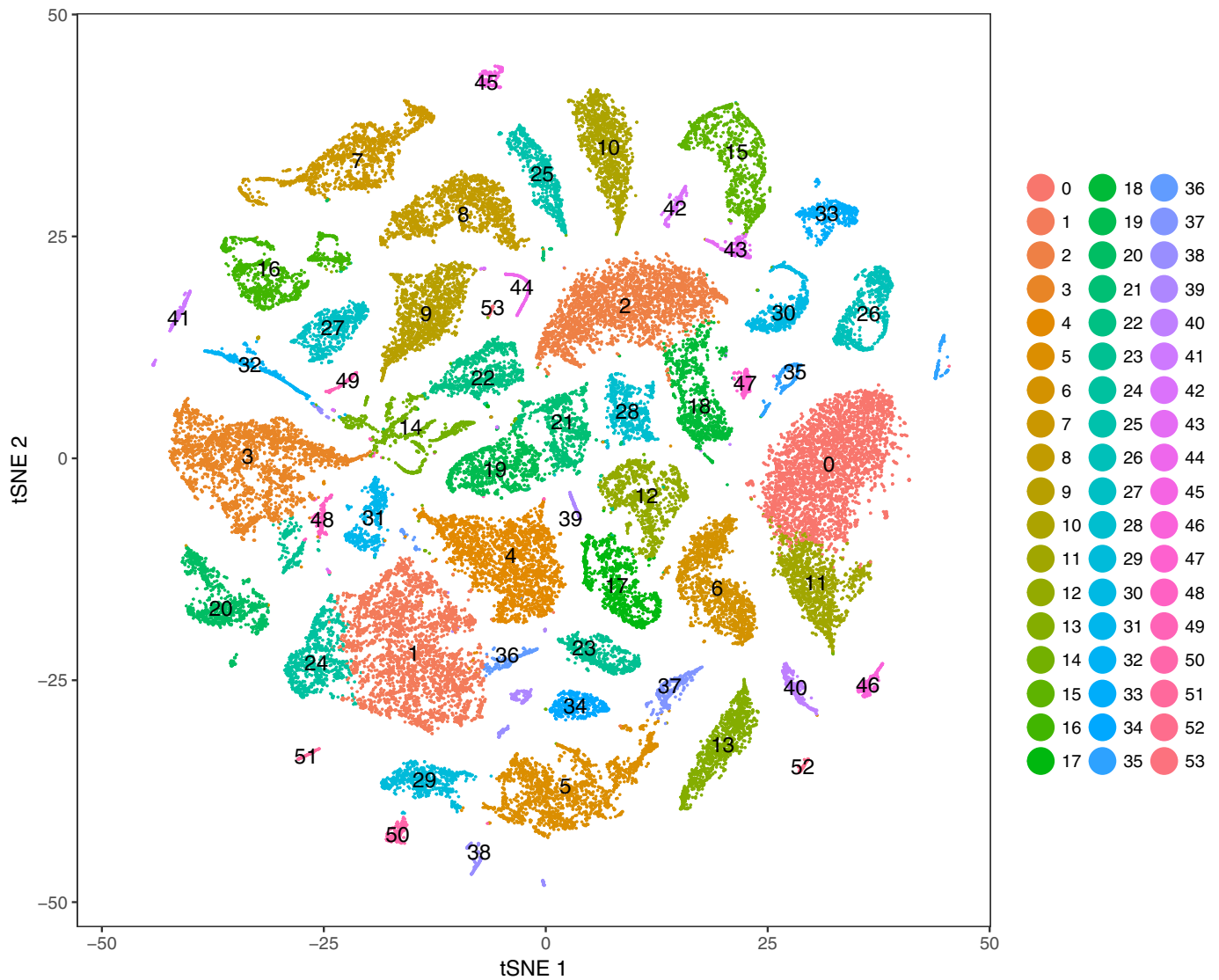
Extended Data Fig. 6 | The number of detected genes decreases similarly across organs as the read or UMI threshold is increased. Fraction of all detected genes (defined as >0 reads or UMIs) for each cell, across all organs, detected at increasing read or UMI thresholds for

FACS (left; $n = 44,949$ individual cells), microfluidic-droplet (middle; $n = 55,656$ individual cells), and microwell-seq (right; $n = 28,372$ individual cells) methods. Summary statistics are contained in Supplementary Table 6.

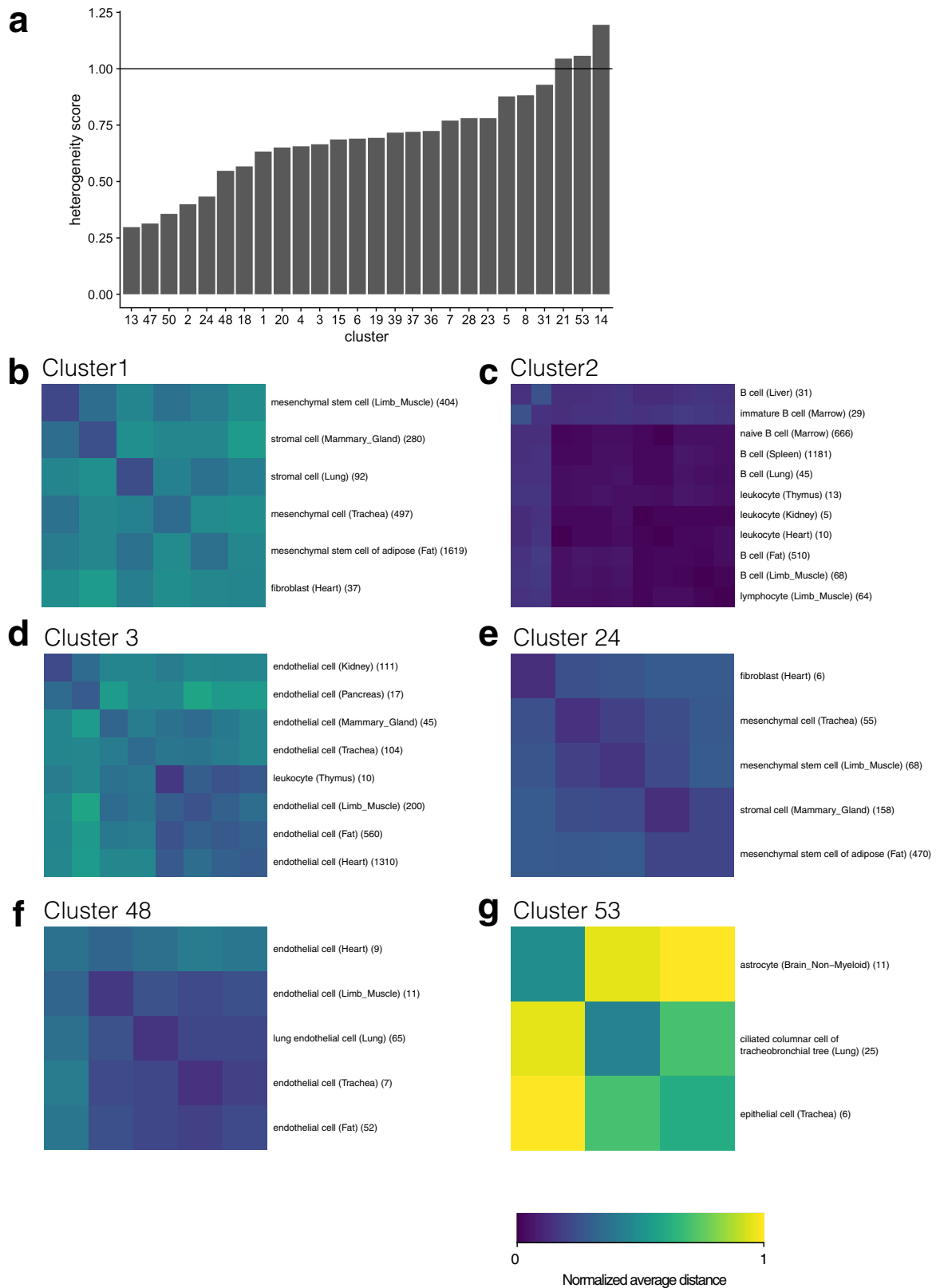


Extended Data Fig. 7 | The number of differentially expressed genes for each cell type that are common between methods. Venn diagrams showing the overlap between differentially expressed genes for each

common cell type across the three methods (FACS, microfluidic-droplet and microwell-seq). Plotted data are provided in tabular form in Supplementary Table 2.

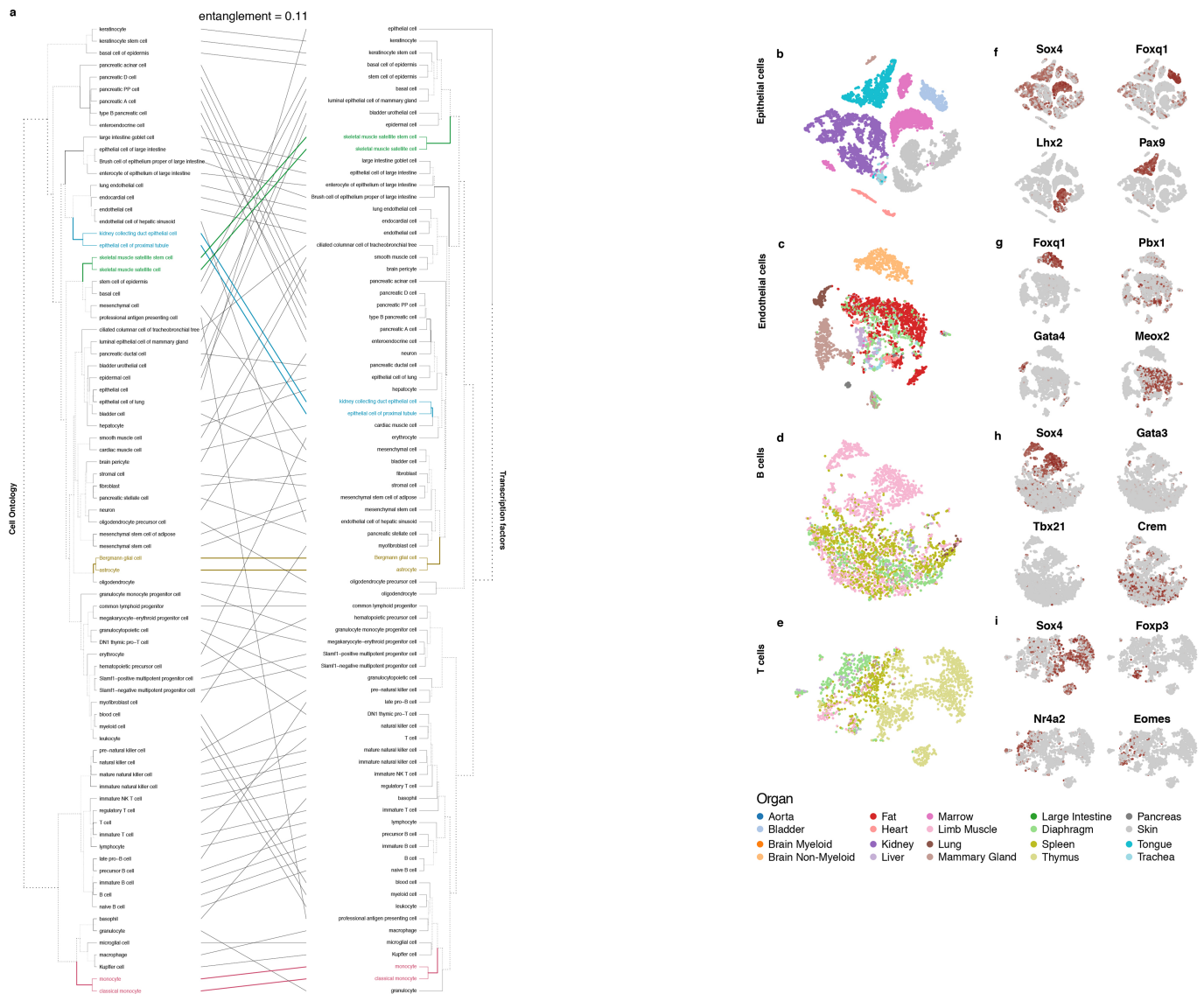


Extended Data Fig. 8 | *t*-SNE visualization of all FACS cells by cluster ID. $n = 44,949$ individual cells. Clusters are discussed in the text and further analysed in Fig. 3.



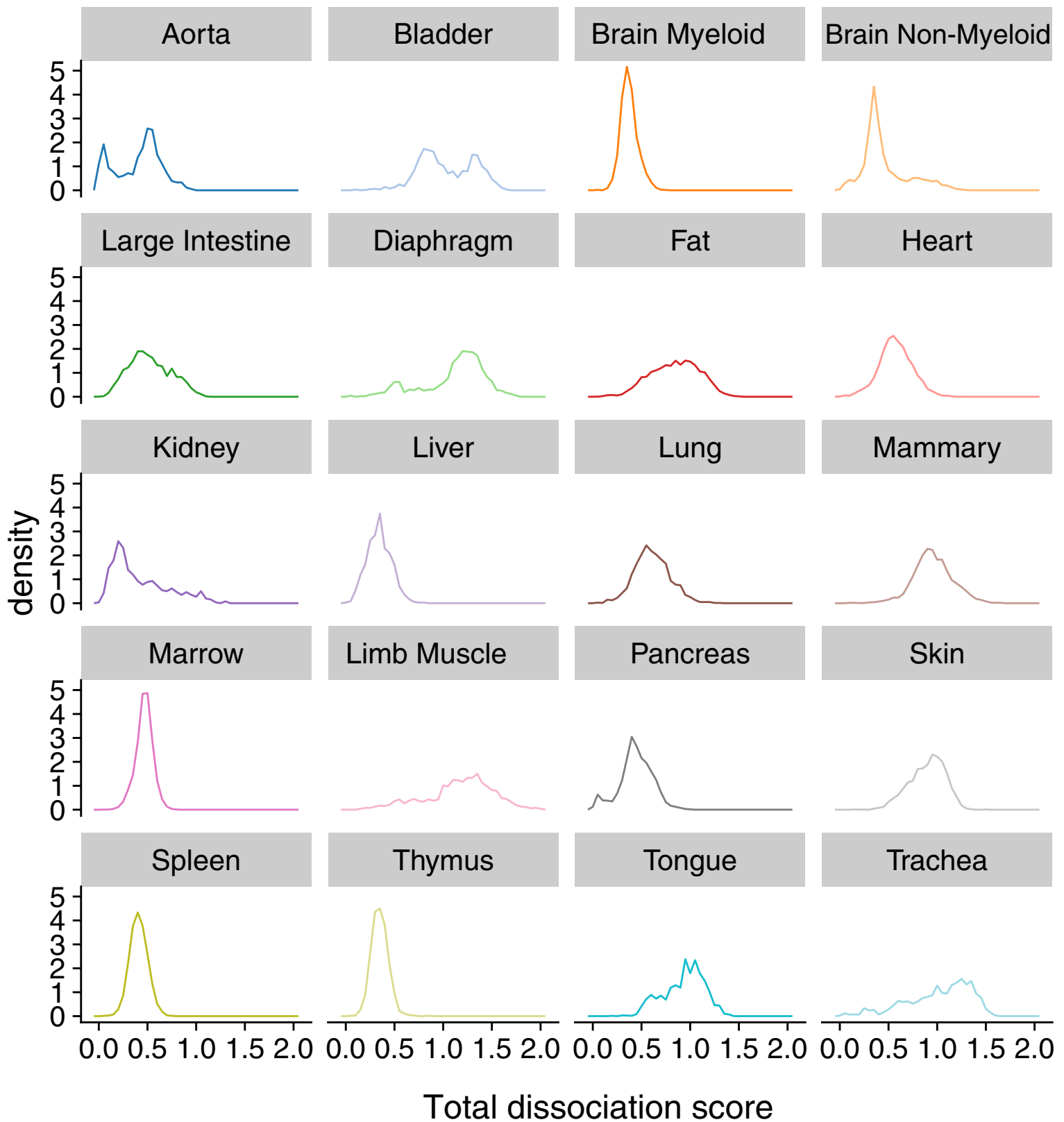
Extended Data Fig. 9 | Metrics of cluster heterogeneity. a, Bar plot showing the heterogeneity score for each cluster containing several cell types. **b–g**, Heat maps showing the average between-cell-type distances

within select clusters, normalized so that the average distance between pairs of FACS cells is 1, clipped to a max of 1, for clusters 1 (**b**), 2 (**c**), 3 (**d**), 24 (**e**), 48 (**f**) and 53 (**g**).



Extended Data Fig. 10 | Contribution of transcription factors to cell identity. **a**, Tanglegram contrasting the dendrogram obtained using all expressed genes with one obtained using only the expression of transcription factors. The solid lines indicate segments that did not change position during the alignment between the two trees, and the dotted lines correspond to dendrogram branches reordered during the entanglement

calculations. The colours indicate the branches for which identical leaves are aligned in both dendrograms. **b–e**, *t*-SNE visualization of epithelial (**b**), endothelial (**c**), B cells (**d**) and T cells (**e**), coloured by organ. **f–i**, *t*-SNE visualization of epithelial (**f**), endothelial (**g**) B cell (**h**) and T cell (**i**) expression of select transcription factors (from grey, low, to red, high). In **b–i**, $n = 60$ randomly selected cells for each cell type.



Extended Data Fig. 11 | Dissociation-induced gene-expression scores for each organ analysed with FACS. The dissociation score for each organ represents the magnitude of the first principal component of the

140 dissociation-associated genes from ref. ²⁴. The y axis shows the probability density of the normalized histogram.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated
- Clearly defined error bars
State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on [statistics for biologists](#) may be useful.

Software and code

Policy information about [availability of computer code](#)

Data collection

FACS data was collected with BD FACS Diva version 8 and the default SH800S software version 1.8.

Data analysis

Sequences from the Novaseq were de-multiplexed using bcl2fastq version 2.19.0.316. Reads were aligned using to the mm10plus genome using STAR version 2.5.2b with parameters TK. 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 droplet platform were de-multiplexed and aligned using CellRanger version 2.0.1, available from 10x Genomics with default parameters.

Standard procedures for filtering, variable gene selection, dimensionality reduction, and clustering were performed using the Seurat package version 2.2.1. A detailed worked example, including the mathematical formulae for each operation, is in the Tissue Annotation Vignette. The parameters that were tuned on a per-tissue basis (resolution and number of PCs can be viewed in the tissue-specific Rmd files available on GitHub).

Dendrograms were made with the OntologyX R package version 2.4, and tanglegrams with the dendextend R package version 1.8. The varselRF R package version 0.7-8 was used for random forest analysis.

Custom code was used for the transcription factor analysis and is available through GitHub.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data, protocols, and analysis scripts from the Tabula Muris are shared as a public resource (<http://tabula-muris.ds.czbiohub.org/>). Gene counts and metadata from all single cells are accessible on Figshare (https://figshare.com/projects/Tabula_Muris_Transcriptomic_characterization_of_20_organs_and_tissues_from_Mus_musculus_at_single_cell_resolution/27733), and raw data are available on GEO (GSE109774).

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to pre-determine sample size. Our sample size is sub-sampled for each organ from a cohort of 4 male and 4 female mice. The number of mice was restricted due to costs and practical constraints of collecting and processing organs. The sample size varies for each organ, as documented in the supplemental text. This was due to experimental constraints, for example: conflicting tissue dissociation protocols for liver and pancreas which limited our ability to process single cells from both of these organs from the same animal.
Data exclusions	Exclusion criteria were pre-establish: Cells with fewer than 500 detected genes were excluded. (A gene counts as detected if it has at least one read mapping to it). Cells with fewer than 50,000 reads (FACS) or 1000 UMI (microfluidic droplet) were excluded.
Replication	We have compared our results where possible, to those from Han et al., 2018. However, besides the biological replicates included in the study, we have not replicated any findings due to the financial constraints of single cell sequencing.
Randomization	Mice were not randomized due to practical constraints. Either 2 males or 2 females were collected on any particular day, spanning 4 days. 2 males: 6/15/2017 and 6/16/2017. 2 females: 7/7/2018 and 7/20/2017
Blinding	We did not have experimental groups in this study.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Unique biological materials
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used

Number;Antigen name;Fluorophore;Company;Cat No;Reactivity species;Origin species;Clonality;Clone;Application;Link
 1;ACSA2;PE;Miltenyi;130102365;mouse;rat;monoclonal; IH3-18A3;IF, FC;https://www.miltenyibiotec.com/US-en/products/macs-flow-cytometry/antibodies/primary-antibodies/anti-acsa-2-antibodies-mouse-ih3-18a3-1-10.html/
 2;rabbit IgG;AF488;Invitrogen;A11034;rabbit;goat;polyclonal;NA;IF, FC, IHC;https://www.thermofisher.com/antibody/product/Goat-anti-Rabbit-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-11034
 3;Ter119;Pacific Blue;BioLegend;116232;mouse;rat;monoclonal;Ter-119;FC;https://www.biolegend.com/en-us/products/pacific-blue-anti-mouse-ter-119-erythroid-cells-antibody-6137
 4;CD45R/B220;FITC;BioLegend;103206;mouse, human, cat;rat;monoclonal; RA3-6B2;FC;https://www.biolegend.com/en-us/products/fitc-anti-mouse-human-cd45r-b220-antibody-445
 5;CD106;PE/Cy7;BioLegend;105720;mouse;rat;monoclonal; 429 (MVCAM.A);FC;https://www.biolegend.com/en-us/products/pe-cy7-anti-mouse-cd106-antibody-6135
 6;CD11B/MAC1;FITC;BioLegend;101206;mouse, human, chimpanzee, baboon, cynomolgus, rhesus, rabbit;rat;monoclonal; M1/70;FC;https://www.biolegend.com/en-us/products/fitc-anti-mouse-human-cd11b-antibody-347
 7;anti-mouse Lineage Cocktail;FITC;BioLegend;133301;mouse;Armenian Hamster, rat;mixture of monoclonals;145-2C11, RB6-8C5, M1/70, RA3-6B2, Ter-119;FC;https://www.biolegend.com/de-de/products/fitc-anti-mouse-lineage-cocktail-with-isotype-ctrl-5803
 8;CD11b;BV421;BioLegend;101236;mouse, human, chimpanzee, baboon, cynomolgus, rhesus, rabbit;rat;monoclonal; M1/70;FC;https://www.biolegend.com/de-de/products/brilliant-violet-421-anti-mouse-human-cd11b-antibody-7163
 9;CD11b;PE/Cy7;BD Pharmingen;552850;mouse, human, chimpanzee, baboon, cynomolgus, rhesus, rabbit;rat;monoclonal; M1/70;FC;https://www.bdbiosciences.com/eu/applications/research/stem-cell-research/mesenchymal-stem-cell-markers-bone-marrow/mouse/negative-markers/pe-cy7-rat-anti-cd11b-m170/p/552850
 10;CD140a/PDGFRa;APC;R&D Systems;AF1062;mouse;mouse;polyclonal;NA;WB, IHC;https://www.rndsystems.com/products/mouse-pdgfr-alpha-antibody_af1062
 11;CD140b/PDGFRb;APC;eBioscience;17-1402-82;mouse;rat;monoclonal;APB5;FC;https://www.thermofisher.com/antibody/product/CD140b-PDGFRB-Antibody-clone-APB5-Monoclonal/17-1402-80
 12;CD171/L1CAM;PE-Vio770;Miltenyi;130-102-135;mouse;rat;monoclonal;555;FC, IF;https://www.miltenyibiotec.com/US-en/products/macs-flow-cytometry/antibodies/primary-antibodies/cd171-l1cam-antibodies-mouse-555-1-10.html/
 13;CD24 ;PE/Cy7;BioLegend;101822;mouse;rat;monoclonal; M1/69;FC;https://www.biolegend.com/en-us/products/pe-cy7-anti-mouse-cd24-antibody-3862
 14;CD2;APC;BioLegend;100112;mouse;rat;monoclonal;RM2-5;FC;https://www.biolegend.com/en-ie/products/apc-anti-mouse-cd2-antibody-9287
 15;CD2;PE/Cy7;BioLegend;100114;mouse;rat;monoclonal;RM2-5;FC;https://www.biolegend.com/de-de/products/pe-cy7-anti-mouse-cd2-antibody-9288
 16;CD31;Pacific Blue;BioLegend;102422;mouse;rat;monoclonal;390;FC;https://www.biolegend.com/de-de/products/pacific-blue-anti-mouse-cd31-antibody-6669
 17;CD31;APC;BD Biosciences;551262;mouse;rat;monoclonal;MEC 13.3;FC;http://www.bdbiosciences.com/eu/applications/research/stem-cell-research/cancer-research/mouse/apc-rat-anti-mouse-cd31-mec-133/p/551262
 18;CD31;APC;BioLegend;102510;mouse;rat;monoclonal;MEC 13.3;FC;https://www.biolegend.com/de-de/products/apc-anti-mouse-cd31-antibody-375
 19;CD31;BV421;BD;562939;mouse;rat;monoclonal;MEC 13.3;FC;http://www.bdbiosciences.com/us/applications/research/stem-cell-research/cancer-research/mouse/bv421-rat-anti-mouse-cd31-mec-133/p/562939
 20;CD31;FITC;BD Biosciences;561813;mouse;rat;monoclonal;MEC 13.3;FC;https://www.bdbiosciences.com/us/applications/research/stem-cell-research/cancer-research/mouse/fitc-rat-anti-mouse-cd31-mec-133/p/561813
 21;CD31;FITC;BioLegend;102506;mouse;rat;monoclonal;MEC 13.3;FC;https://www.biolegend.com/de-de/products/fitc-anti-mouse-cd31-antibody-377
 22;CD31;PE/Dazzle594;BioLegend;102526;mouse;rat;monoclonal;MEC 13.3;FC;https://www.biolegend.com/de-at/products/pe-dazzle-594-anti-mouse-cd31-antibody-14322
 23;CD326;APC/Fire750;BioLegend;118230;mouse;rat;monoclonal;G8.8;FC;https://www.biolegend.com/en-us/products/apc-fire-750-anti-mouse-cd326-ep-cam-antibody-13635
 24;CD326;PE/Cy7;BioLegend;118216;mouse;rat;monoclonal;G8.8;FC;https://www.biolegend.com/en-us/products/pe-cy7-anti-mouse-cd326-ep-cam-antibody-5303
 25;CD34;AF647;BD Pharmingen;560230;mouse;rat;monoclonal;RAM34;FC;http://www.bdbiosciences.com/us/applications/research/stem-cell-research/cancer-research/mouse/alexa-fluor-647-rat-anti-mouse-cd34-ram34/p/560230
 26;CD3;APC;BioLegend;100236;mouse;rat;monoclonal;17A2;FC;https://www.biolegend.com/en-us/products/apc-anti-mouse-cd3-antibody-8055
 27;CD3;FITC;BioLegend;100203;mouse;rat;monoclonal;17A2;FC;https://www.biolegend.com/de-de/products/fitc-anti-mouse-cd3-antibody-45
 28;CD44;APC;BioLegend;103012;mouse, human;rat;monoclonal;IM7;FC;https://www.biolegend.com/de-de/products/apc-anti-

mouse-human-cd44-antibody-312
 29;CD45;Pacific Blue;BioLegend;103126;mouse;rat;monoclonal;30-F11;FC;https://www.biolegend.com/de-de/products/pacific-blue-anti-mouse-cd45-antibody-3102
 30;CD45;APC;BioLegend;103112;mouse;rat;monoclonal;30-F11;FC;https://www.biolegend.com/de-de/products/apc-anti-mouse-cd45-antibody-97
 31;CD45;BV51;BioLegend;103138;mouse;rat;monoclonal;30-F11;FC;https://www.biolegend.com/de-de/products/brilliant-violet-510-anti-mouse-cd45-antibody-7995
 32;CD45;FITC;BioLegend;103108;mouse;rat;monoclonal;30-F11;FC;https://www.biolegend.com/de-de/products/fitc-anti-mouse-cd45-antibody-99
 33;CD45;PE;BioLegend;103106;mouse;rat;monoclonal;30-F11;FC;https://www.biolegend.com/de-de/products/pe-anti-mouse-cd45-antibody-100
 34;CD45;PE/Cy7;eBioscience;25-0451-82;mouse;rat;monoclonal;30-F11;FC;https://www.thermofisher.com/antibody/product/CD45-Antibody-clone-30-F11-Monoclonal/25-0451-81
 35;CD45R/B220;FITC;BioLegend;103206;mouse, human, cat; rat; monoclonal; RA3-6B2;FC;https://www.biolegend.com/en-us/products/fitc-anti-mouse-human-cd45r-b220-antibody-445
 36;CD49f;APC;BioLegend;313616;Human, African Green, Mouse, Baboon, Capuchin Monkey, Cat (Feline), Cattle (Bovine, Cow), Chimpanzee, Cynomolgus, Dog (Canine), Horse (Equine), Rabbit (Lapine), Rhesus, Sheep (Ovine), Swine (Pig, Porcine);rat;monoclonal;GoH3;FC;https://www.biolegend.com/de-de/products/apc-anti-human-mouse-cd49f-antibody-5617
 37;CD66a;PE;BioLegend;134506;mouse;mouse;monoclonal;Mab-CC1;FC;https://www.biolegend.com/de-de/products/pe-anti-mouse-cd66a-ceacam1a-antibody-5986
 38;CD90.2/Thy-1.2;APC/Cy7;BioLegend;105328;mouse;rat;monoclonal;30-H12;FC;https://www.biolegend.com/de-de/products/apc-cy7-anti-mouse-cd90-2-antibody-6671
 39;CD90.2/Thy-1.1;AF488;BioLegend;202506;Rat, Mouse (AKR/J, PL, and FVBIN mouse strains), Cross-Reactivity: Rabbit (Lapine), Guinea Pig;mouse;monoclonal; OX-7 ;FC;https://www.biolegend.com/de-de/products/alexa-fluor-488-anti-rat-cd90-mouse-cd90-1-thy-1-1-antibody-3126
 40;CD90.2/Thy-1.2;FITC;BioLegend;140304;mouse;rat;monoclonal; 53-2.1;FC;https://www.biolegend.com/de-de/products/fitc-anti-mouse-cd90-2-thy-1-2-antibody-6761
 41;C-FMS ;BV411;BioLegend;135513;mouse;rat;monoclonal;AFS98;FC;https://www.biolegend.com/de-de/products/brilliant-violet-421-anti-mouse-cd115-csf-1r-antibody-8971
 42;CKIT ;APC;BioLegend;105812;mouse;rat;monoclonal; 2B8;FC;https://www.biolegend.com/de-de/products/apc-anti-mouse-cd117-c-kit-antibody-72
 43;endomucin;FITC;eBioscience ;14-5851-82;mouse;rat;monoclonal;V.7C7;FC, IHC, IP, WB;https://www.thermofisher.com/antibody/product/Endomucin-Antibody-clone-eBioV-7C7-V-7C7-Monoclonal/14-5851-81
 44;EpCAM;APC;ThermoFisher;17-5791-82;mouse;rat;monoclonal;G8.8;FC, IHC, IP, WB;https://www.thermofisher.com/antibody/product/CD326-EpCAM-Antibody-clone-G8-8-Monoclonal/17-5791-80
 45;EpCAM;FITC;eBioscience;11-5791-82;mouse;rat;monoclonal;G8.8;FC, IHC, IP, WB;https://www.thermofisher.com/antibody/product/CD326-EpCAM-Antibody-clone-G8-8-Monoclonal/11-5791-80
 46;TER119;Pacific Blue;BioLegend;116232;mouse;rat;monoclonal;Ter-119;FC;https://www.biolegend.com/de-de/products/pacific-blue-anti-mouse-ter-119-erythroid-cells-antibody-6137
 47;IgM;PE/Cy7;BioLegend;406514;mouse;rat;monoclonal; RMM-1;FC;https://www.biolegend.com/de-de/products/pe-cy7-anti-mouse-igm-6935
 48;CD49f;FITC;BioLegend;313606;Human, African Green, Mouse, Baboon, Capuchin Monkey, Cat (Feline), Cattle (Bovine, Cow), Chimpanzee, Cynomolgus, Dog (Canine), Horse (Equine), Rabbit (Lapine), Rhesus, Sheep (Ovine), Swine (Pig, Porcine);rat;monoclonal;GoH3;FC;https://www.biolegend.com/de-de/products/fitc-anti-human-mouse-cd49f-antibody-2606
 49;LNGFR;Vio770;Miltenyi;103110079;mouse, monkey, goat, dog, pig, sheep;mouse;monoclonal;ME20.4-1.H4;FC, IF;https://www.miltenyibiotec.com/US-en/products/mac-flow-cytometry/antibodies/primary-antibodies/cd271-Ingfr-antibodies-human-me20-4-1-h4-1-11.html/
 50;Ly-6A/E;Pacific Blue;BioLegend;108120;mouse;rat;monoclonal;D7;FC;https://www.biolegend.com/de-de/products/pacific-blue-anti-mouse-ly-6a-e-sca-1-antibody-3140
 51;Ly-6G/Ly-6C/GR1;FITC;BioLegend;108405;mouse;rat;monoclonal; RB6-8C5;FC;https://www.biolegend.com/de-de/products/fitc-anti-mouse-ly-6g-ly-6c-gr-1-antibody-458
 52;Ly-6G/Ly-6C/GR1;PE/Cy7;BioLegend;108416;mouse;rat;monoclonal; RB6-8C5;FC;https://www.biolegend.com/de-de/products/pe-cy7-anti-mouse-ly-6g-ly-6c-gr-1-antibody-1931
 53;CD326;PE/Cy7;BioLegend;118216;mouse;rat;monoclonal;G8.8;FC;https://www.biolegend.com/de-de/products/pe-cy7-anti-mouse-cd326-ep-cam-antibody-5303
 54;CD44;APC;BioLegend;103012;mouse, human;rat;monoclonal;IM7;FC;https://www.biolegend.com/de-de/products/apc-anti-mouse-human-cd44-antibody-312
 55;CD45;Pacific Blue;BioLegend;103126;mouse;rat;monoclonal;30-F11;FC;https://www.biolegend.com/de-de/products/pacific-blue-anti-mouse-cd45-antibody-3102
 56;CD66a;PE;BioLegend;134506;mouse;mouse;monoclonal;Mab-CC1;FC;https://www.biolegend.com/de-de/products/pe-anti-mouse-cd66a-ceacam1a-antibody-5986
 57;O4;PE;Miltenyi;130-095-887;humam, mouse, rat;mouse;monoclonal;O4;FC, IF;https://www.miltenyibiotec.com/US-en/products/mac-flow-cytometry/antibodies/primary-antibodies/anti-o4-antibodies-human-mouse-rat-o4-1-11.html
 58;SCA1;APC;BioLegend;122512;mouse;rat;monoclonal;E13-161.7;FC;https://www.biolegend.com/en-us/products/apc-anti-mouse-ly-6a-e-sca-1-antibody-3897
 59;SCA1;FITC;BioLegend;122506;mouse;rat;monoclonal;E13-161.7;FC;https://www.biolegend.com/en-us/products/fitc-anti-mouse-ly-6a-e-sca-1-antibody-3894
 60;SCA1;Pacific Blue;BioLegend;108120;mouse;rat;monoclonal;E13-161.7;FC;https://www.biolegend.com/pacific-blue-anti-mouse-ly-6a-e-sca-1-antibody-3140.html
 61;SCA1;PE/Cy7;BioLegend;122514;mouse;rat;monoclonal;E13-161.7;FC;https://www.biolegend.com/pe-cy7-anti-mouse-ly-6a-e-sca-1-antibody-3898.html
 62;SDC4;APC;Miltenyi;130-109-831;mouse;rat;monoclonal;REA640;FC, IF;https://www.miltenyibiotec.com/US-en/products/mac-flow-cytometry/antibodies/primary-antibodies/anti-syndecan-4-antibodies-mouse-rea640-1-10.html/
 63;TER119;FITC;BioLegend;116205;mouse;rat;monoclonal;TER-119;FC;https://www.biolegend.com/en-us/products/fitc-anti-mouse-ter-119-erythroid-cells-antibody-1865

64;TER119;PE/Cy5;BioLegend;116210;mouse;rat;monoclonal;TER-119;FC;https://www.biolegend.com/en-us/products/pe-cy5-anti-mouse-ter-119-erythroid-cells-antibody-1868
 65;Thbs1;APC;Fisher;MA5-13398;Bovine, Dog, Horse, Human, Mouse, Sheep, Pig, Rat;mouse;monoclonal;A6.1;FC, IHC, IF, IP, WB;https://www.thermofisher.com/antibody/product/Thrombospondin-1-Antibody-clone-A6-1-Monoclonal/MA5-13398
 66;Tmem119 ;AF488;Abcam;ab210405;mouse;rabbit;monoclonal;106-6;FC;http://www.abcam.com/tmem119-antibody-106-6-microglial-marker-ab210405.html
 67;VCAM;PE/Cy7;Biolegend;105720;mouse;rat;monoclonal;MVCAM.A;FC;https://www.biolegend.com/en-us/products/pe-cy7-anti-mouse-cd106-antibody-6135

Validation

Please refer to references contained in the links in the "Antibodies Used" section above.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

Mus musculus, C57BL/6JN, male and female, 10-15 weeks

Wild animals

Study did not involve wild animals.

Field-collected samples

Study did not involve field samples.

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Please see the supplemental text for extensive details regarding each organ.

Instrument

Sony SH800S for all organs except limb muscle and diaphragm, which were sorted on a BD Aria III.

Software

BD FACS Diva and the default SH800S software were used to collect data.

Cell population abundance

Please see the supplemental text for extensive details regarding each organ.

Gating strategy

Please see the index data for each organ.

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.