Tabula Microcebus: A transcriptomic cell atlas of mouse lemur, an emerging primate model organism

The Tabula Microcebus Consortium*

The Tabula Microcebus Consortium

Lead authors

Camille Ezran[1, 2‡], Shixuan Liu[1, 2, 3‡], Stephen Chang[1, 2, 4‡], Jingsi Ming[5, 6], Olga Botvinnik[7], Lolita Penland[7], Alexander Tarashansky[7, 8], Antoine de Morree[9, 10], Kyle J. Travaglini[1, 2], Kazuteru Hasegawa[1, 11, 12], Hosu Sin[13], Rene Sit[7], Jennifer Okamoto[7], Rahul Sinha[14], Yue Zhang[2, 15], Caitlin J. Karanewsky[1, 2], Jozeph L. Pendleton[1, 2], Maurizio Morri[7], Martine Perret[16], Fabienne Aujard[16], Lubert Stryer[17], Steven Artandi[1, 11, 12], Margaret Fuller[13, 18], Irving L. Weissman[14], Thomas A. Rando[9], James E. Ferrell Jr.[1, 3], Bo Wang[8], Iwijn De Vlaminck[19], Can Yang[6], Kerriann M. Casey[20], Megan A. Albertelli[20], Angela Oliveira Pisco[7], Jim Karkanias[7], Norma Neff[7], Angela Wu[21, 22], Stephen R. Quake[7, 8, 23]†, Mark A. Krasnow[1, 2]†

‡ Co-first authors

† Corresponding authors: krasnow@stanford.edu (M.A.K.), steve@quake-lab.org (S.R.Q.)

Leadership and coordination

Trainees: Camille Ezran[1, 2], Shixuan Liu[1, 2, 3], Stephen Chang[1, 2, 4]

Principal investigators: Mark A. Krasnow[1, 2], Stephen R. Quake[7, 8, 23], Angela Wu[21, 22], Norma Neff[7], Jim Karkanias[7], Angela Oliveira Pisco[7]

Donor recruitment and animal husbandry

Megan A. Albertelli[20], Caitlin J. Karanewsky[1, 2], Jozeph L. Pendleton[1, 2], Fabienne Aujard[16], Martine Perret[16], Liza Shapiro[24], Andriamahery Razafindrakoto[25], Hajanirina Noëline Ravelonjanahary[25], Patricia Wright[26], Anne D. Yoder[27], Cathy V. Williams[28], Robert Schopler[28], Ute Radespiel[29], Jean-Michel Verdier[30], Corinne Lautier[30], E. Christopher Kirk[24], Rebecca Lewis[24]

Necropsy, sample collection, biobanking, histology, and pathology

Kerriann M. Casey[20], Megan A. Albertelli[20], Caitlin J. Karanewsky[1, 2], Jozeph L. Pendleton[1, 2], Camille Ezran[1, 2], Shixuan Liu[1, 2, 3], Kyle J. Travaglini[1, 2], Astrid Gillich[1, 2], Zicheng Zhao[2, 20], Stephen Chang[1, 2, 4], Elias Godoy[20], Jérémy Terrien[16], Jacques Epelbaum[16, 31], Dita Gratzinger[32], Katherine Lucot[32], Tom Montine[32]

Tissue processing, annotation, and analysis [Principal investigators]

Aorta: Jessica D'Addabbo[33, 34], Isaac Bakerman[4], [Patricia Nguyen[12, 33, 34]]

Bladder: Aaron Kershner[1, 35], Karim Mrouj[36], [Philip Beachy[3, 13, 36, 37]]

Blood: Rahul Sinha[14], Camille Ezran[1, 2], Yue Zhang[2, 15], Shixuan Liu[1, 2, 3], Caitlin J. Karanewsky[1, 2], [Irving L. Weissman[14], Mark A. Krasnow[1, 2]]

Bone (limb, spine): Thomas Ambrosi[36], Malachia Hoover[36], Alina Alam[36], [Charles Chan[36]]

Bone marrow: Rahul Sinha[14], SoRi Jang[1, 2], Camille Ezran[1, 2], [Irving L. Weissman[14], Mark A. Krasnow[1, 2]]

Brain (cortex, brainstem, cerebellum): Avin Veerakumar[1, 2, 8, 38], Peng Li[1, 2], Andrea R. Yung[1, 2], Connor V. Duffy[2, 18], Song-Lin Ding[39], [Ed S. Lein[39], Silvana Konermann[1,2], Liqun Luo[2, 15], Trygve E. Bakken[39], Justus Kebschull[40], Rebecca D. Hodge[39], Mark A. Krasnow[1, 2]]

Colon/Small intestine: Taichi Isobe[41], [Michael F. Clarke[36]]

Diaphragm/Limb muscle: Antoine de Morree[9, 10], Biter Bilen[9], Jean Farup[9, 10], Andoni Urtasun[9], Jengmin Kang[9], [Thomas A. Rando[9]]
Eye: Ming Chen[18], Bao Xiang[42], Varun Ramanan Subramaniam[42], Shravani Mukherjee[42], Aditi Swarup[42], Lily Kim[18], Bronwyn Scott[42], Ahmad Al-Moujahed[42], [Albert Y. Wu[42], Douglas Vollrath[18], Lubert Stryer[17]]
Fat: Nicholas Schaum[9], Jingsi Ming[5, 6], Shixuan Liu[1, 2, 3], Amanda L. Wiggenhorn[32, 43], [Tony Wyss-Coray[9, 44], Jonathan Z. Long[32, 45], Angela Wu[21, 22]]
Ganglion (vagus nerve): Yin Liu[1, 2], [Mark A. Krasnow[1, 2]]
Heart: Stephen Chang[1, 2, 4], [Mark A. Krasnow[1, 2]]
Hypothalamus & Pituitary: Andrea R. Yung[1, 2], Shixuan Liu[1, 2, 3], [Mark A. Krasnow[1, 2]]
Kidney: Shixuan Liu[1, 2, 3], Ahmad Nabhan[1, 2], Lolita Penland[7], Yue Zhang[2, 15], Andrea R. Yung[1, 2], [Gabriel Loeb[46], Mark A. Krasnow[1, 2]]
Liver: Shengda Lin[47], Shixuan Liu[1, 2, 3], Honor Paine[48], Deviana Burhan[48], Aris Taychameekiatchai[48], [Steven Artandi[1, 11, 12], Bruce Wang[48]]
Lung: Kyle J. Travaglini[1, 2], Ahmad Nabhan[1, 2], F. Hernán Espinoza[1, 2], Astrid Gillich[1, 2], [Christin Kuo[49], Ross Metzger[49, 50], Mark A. Krasnow[1, 2]]
Lymph node: SoRi Jang[1, 2], Lolita Penland[7], [Norma Neff[7], Mark A. Krasnow[1, 2]]
Mammary gland: Zhen Qi[36], [Michael F. Clarke[36]]
Microbiome: Rebecca Culver[18], [Kerwyn C. Huang[7, 8, 51]]
Pancreas: Patrick Neuhöfer[1, 11, 12], Charles A. Chang[13, 52], Yan Hang[13, 52], [Seung K. Kim[12, 13, 52, 53], Steven Artandi[1, 11, 12]]
Prostate: Hannah Weinstein[54, 55, 56], Paul Allegakoen[54], [Franklin W. Huang[54]]
Salivary gland: Sivakamasundari V[36], [Philip Beachy[3, 13, 36, 37]]
Skin: Song Eun Lee[9, 44], [Tony Wyss-Coray[9, 44]]
Spleen: Lolita Penland[7,] Hannah K. Frank[32], Dita Gratzinger[32], [Norma Neff7, Scott D. Boyd[32]]
Testes: Kazuteru Hasegawa[1, 11, 12], Hosu Sin[13], Jingsi Ming[5, 6], [Steven Artandi[1, 11, 12], Margaret Fuller[13, 18], Angela Wu[21, 22]]
Thymus: Isaac Bakerman[4], Jessica D'Addabbo[33, 34], [Patricia Nguyen[12, 33, 34]]
Tongue: Wan-Jin Lu[36], Ankit Baghel[15], [Philip Beachy[3, 13, 36, 37]]
Trachea/Thyroid/Parathyroid: William Kong[36], [Philip Beachy[3, 13, 36, 37]]
Uterus/Abdominal mass: Lolita Penland[7], [Norma Neff[7]]

Library processing, sorting, and sequencing
Stephen Chang[1, 2, 4], Lolita Penland[7], Carly Israel[7], Rene Sit[7], Jennifer Okamoto[7], Yue Zhang[2, 15], Rahul Sinha[14], Shixuan Liu[1, 2, 3], Camille Ezran[1, 2], Kyle J. Travaglini[1, 2], Antoine de Morree[9, 10], Ahmad Nabhan[1, 2], Youcef Ouadah[1, 58], Jalal Baruni[1, 2, 15, 59], Robert C. Jones[8], Norma Neff[7], Maurizio Morri[7], Spyros Darmanis[7], Sheela Crasta[7], Jia Yan[7], Aditi Agrawal[7], Shelly Huynh[7], Brian Yu[7], James Webber[7]

Global data analysis [Principal investigators]
General analysis: Camille Ezran[1, 2], Shixuan Liu[1, 2, 3], [Mark A. Krasnow[1, 2]]
Sequence alignment: Stephen Chang[1, 2, 4], Kyle J. Travaglini[1, 2], Rahul Sinha[14], Rene Sit[7], [Mark A. Krasnow[1, 2], Irving L. Weissman[14], Stephen R. Quake[7, 8, 23]]
FIRM integration: Jingsi Ming[5, 6], Jinxurong Yang[21], [Angela Wu[21, 22], Can Yang[6]]
Molecular trajectory: Shixuan Liu[1, 2, 3], Camille Ezran[1, 2], Weilun Tan[7], [Mark A. Krasnow[1, 2]]

Global cell type comparison: Shixuan Liu[1, 2, 3], [Mark A. Krasnow[1, 2]]
Cross-species comparison: Olga Botvinnik[7], Alexander Tarashansky[7, 8], Shixuan Liu[1, 2, 3], Camille Ezran[1, 2], Kyle J. Travaglini[1, 2], Antoine de Morree[9, 10], Jia Zhao[6], Gefei Wang[6], [Jim Karkanias[7], Bo Wang[8], Stephen R. Quake[7, 8, 23], Mark A. Krasnow[1, 2], Thomas A. Rando[9], Angela Wu[21, 22], Can Yang[6]]
Spermatogenesis: Jingsi Ming[5, 6], Hosu Sin[13], Kazuteru Hasegawa[1, 11, 12], Camille Ezran[1, 2], [Angela Wu[21, 22], Margaret Fuller[13, 18], Steven Artandi[1, 11, 12], Can Yang[6], Mark A. Krasnow[1, 2]]
uTAR: Michael F.Z. Wang[19, 60], [Iwijn De Vlaminck[19]]
SICILIAN: Roozbeh Dehghannasiri[1, 61], Julia Olivieri[62], [Julia Salzman[1, 61]]
SAMap: Alexander Tarashansky[7, 8], [Bo Wang[8]]
Major histocompatibility locus: Lisbeth A. Guethlein[51, 63], Camille Ezran[1, 2], [Peter Parham[51, 63], Mark A. Krasnow[1, 2]]
B cell receptor locus: Hannah K. Frank[32, 57], Camille Ezran[1, 2], [Scott D. Boyd[32], Mark A. Krasnow[1, 2]]
Cross-tissue neutrophil trajectory: Camille Ezran[1, 2], Shixuan Liu[1, 2, 3], [Mark A. Krasnow[1, 2]]
Cross-tissue monocyte and macrophage comparison: Qiuyu Jing[21], Camille Ezran[1, 2], Shixuan Liu[1, 2, 3], [Angela Wu[21, 22], Mark A. Krasnow[1, 2]]
Chemokine expression: Camille Ezran[1, 2], [Mark A. Krasnow[1, 2]]
Metastatic tumor: Jane Antony[36], Camille Ezran[1, 2], [Michael F. Clarke[36], Mark A. Krasnow[1, 2]]
Endothelium: Camille Ezran[1, 2], Astrid Gillich[1, 2], Geoff Stanley[8], [Ross Metzger[49, 50], Mark A. Krasnow[1, 2], Stephen R. Quake[7, 8, 23]]
Adipocytes: Jinxurong Yang[21], Jingsi Ming[5, 6], Camille Ezran[1, 2], Shixuan Liu[1, 2, 3], [Jonathan Z. Long[32, 45], Angela Wu[21, 22], Mark A. Krasnow[1, 2], Can Yang[6]]
Primate-specific genes: Winston Koh[64], Camille Ezran[1, 2], [Mark A. Krasnow[1, 2]]
Global hormone signaling: Shixuan Liu[1, 2, 3], [James E. Ferrell Jr.[1, 3], Mark A. Krasnow[1, 2]]
OnClass: [Sheng Wang[65]]
Data portal/Online histology atlas: Zicheng Zhao[2, 20], Snigdha Agarwal[7], Kyle Awayan[7], Aaron McGeever[7], Venkata Naga Pranathi Vemuri[7], Saba Nafees[7], [Stephen R. Quake[7, 8, 23], Angela Oliveira Pisco[7]]

Principal investigators (A-Z)
Megan A. Albertelli[20], Steven Artandi[1, 11, 12], Fabienne Aujard[16], Trygve E. Bakken[39], Philip Beachy[3, 13, 36, 37], Scott D. Boyd[32], Kerriann M. Casey[20], Charles Chan[36], Michael F. Clarke[36], Jacques Epelbaum[16, 31], James E. Ferrell Jr.[1, 3], Margaret Fuller[13, 18], Dita Gratzinger[32], Franklin W. Huang[54], Kerwyn C. Huang[7, 8, 51], Jim Karkanias[7], Justus Kebschull[40], Seung K. Kim[12, 13, 52, 53], Silvana Konermann[1,2], Mark A. Krasnow[1, 2], Christin Kuo[49], Ed S. Lein[39], Shengda Lin[47], Gabriel Loeb[46], Jonathan Z. Long[32, 45], Liqun Luo[2, 15], Ross Metzger[49, 50], Jingsi Ming[5, 6], Tom Montine[32], Norma Neff[7], Patricia Nguyen[12, 33, 34], Peter Parham[51, 63], Martine Perret[16], Angela Oliveira Pisco[7], Stephen R. Quake[7, 8, 23], Thomas A. Rando[9], Ute Radespiel[29], Julia Salzman[1, 61], Liza Shapiro[24], Lubert Stryer[17], Jérémy Terrien[16], Jean-Michel Verdier[30], Iwijn De Vlaminck[19], Douglas Vollrath[18], Bo Wang[8], Bruce Wang[48], Sheng Wang[65], Irving L. Weissman[14], Patricia Wright[26], Tony Wyss-Coray[9, 44], Albert Y. Wu[42], Angela Wu[21, 22], Can Yang[6], Anne D. Yoder[27]

Writing
Camille Ezran[1, 2,] Shixuan Liu[1, 2, 3], Stephen Chang[1, 2, 4], Jingsi Ming[5, 6], Antoine de Morree[9, 10], Lubert Stryer[17], Angela Wu[21, 22], Stephen R. Quake[7, 8, 23], Mark A. Krasnow[1, 2]

Affiliations
1. Department of Biochemistry, Stanford University School of Medicine, Stanford, CA, USA.
2. Howard Hughes Medical Institute, USA.
3. Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, CA, USA.
4. Department of Medicine, Division of Cardiovascular Medicine, Stanford University, Stanford, CA, USA.
5. Academy of Statistics and Interdisciplinary Sciences, KLATASDS-MOE, East China Normal University, Shanghai, China.
6. Department of Mathematics, Hong Kong University of Science and Technology, Hong Kong SAR, China.
7. Chan Zuckerberg Biohub, San Francisco, CA, USA.
8. Department of Bioengineering, Stanford University, Stanford, CA, USA.
9. Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, CA, USA.
10. Department of Biomedicine, Aarhus University, Aarhus, Denmark.
11. Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA, USA.
12. Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA.
13. Department of Developmental Biology, Stanford University School of Medicine, Stanford, CA, USA.
14. Stanford Institute for Stem Cell Biology and Regenerative Medicine, Stanford, CA, USA.
15. Department of Biology, Stanford University, Stanford, CA, USA.
16. Adaptive Mechanisms and Evolution (MECADEV), UMR 7179, National Center for Scientific Research, National Museum of Natural History, Brunoy, France.
17. Department of Neurobiology, Stanford University School of Medicine, Stanford, CA, USA.
18. Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA.
19. Nancy E. and Peter C. Meinig School of Biomedical Engineering, Cornell University, Ithaca, NY, USA.
20. Department of Comparative Medicine, Stanford University School of Medicine, Stanford, CA, USA.
21. Division of Life Science, Hong Kong University of Science and Technology, Hong Kong SAR, China.
22. Department of Chemical and Biological Engineering, Hong Kong University of Science and Technology, Hong Kong SAR, China.
23. Department of Applied Physics, Stanford University, Stanford, CA, USA.
24. Department of Anthropology, University of Texas at Austin, Austin, TX, USA.
25. Department of Animal Biology, Faculty of Science, University of Antananarivo, Antananarivo 101, BP 566, Madagascar.
26. Department of Anthropology, Stony Brook University, NY, USA.

27. Department of Biology, University Program in Genetics & Genomics, Duke University, Durham, NC, USA.
28. Duke Lemur Center, Durham, NC, USA.
29. Institute of Zoology, University of Veterinary Medicine Hannover, Germany.
30. MMDN, University of Montpellier, EPHE-PSL, INSERM, Montpellier, France.
31. Unité Mixte de Recherche en Santé 894 INSERM, Centre de Psychiatrie et Neurosciences, Université Paris Descartes, Sorbonne, Paris, France
32. Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA.
33. Division of Cardiovascular Medicine, Stanford University, Stanford, CA, USA.
34. Stanford Cardiovascular Institute, Stanford, CA, USA.
35. Institute of Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA, USA.
36. Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA, USA.
37. Department of Urology, Stanford University School of Medicine, Stanford, CA, USA.
38. Medical Scientist Training Program, Stanford University School of Medicine, Stanford, CA, USA.
39. Human Cell Types Department, Allen Institute for Brain Science, Seattle, WA, USA.
40. Department of Biomedical Engineering, Johns Hopkins School of Medicine, Baltimore, MD, USA.
41. Department of Oncology and Social Medicine, Kyushu University, Japan.
42. Department of Ophthalmology, Stanford University School of Medicine, Stanford, CA, USA.
43. Department of Chemistry, Stanford University, Stanford, CA, USA.
44. Wu Tsai Neurosciences Institute, Stanford, CA, USA.
45. Stanford ChEM-H, Stanford, CA, USA.
46. Division of Nephrology, Department of Medicine, University of California San Francisco, San Francisco, CA, USA.
47. Zhejiang Provincial Key Laboratory for Cancer Molecular Cell Biology, Life Sciences Institute, Zhejiang University, Hangzhou, Zhejiang, China.
48. Department of Medicine and Liver Center, University of California San Francisco, San Francisco, CA, USA.
49. Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, USA.
50. Division of Cardiology, Stanford University School of Medicine, Stanford, CA, USA.
51. Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA, USA.
52. Stanford Diabetes Research Center, Stanford, CA, USA.
53. JDRF Center of Excellence, Stanford, CA, USA.
54. Division of Hematology/Oncology, Department of Medicine, University of California San Francisco, San Francisco, CA, USA.
55. Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, CA, USA.
56. Bakar Computational Health Sciences Institute, University of California San Francisco, San Francisco, CA, USA.
57. Department of Ecology and Evolutionary Biology, Tulane University, New Orleans, LA, USA.

58. Program in Cancer Biology, Stanford University School of Medicine, Stanford, CA, USA.
59. Department of Anesthesiology, Perioperative and Pain Medicine, Stanford University School of Medicine, Stanford, CA, USA.
60. Department of Computational Biology, Cornell University, Ithaca, NY, USA.
61. Department of Biomedical Data Science, Stanford University, Stanford, CA, USA.
62. Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA, USA.
63. Department of Structural Biology, Stanford University School of Medicine, Stanford, CA, USA.
64. Institute of Molecular and Cell Biology, Agency of Science Technology and Research, Singapore.
65. Paul G. Allen School of Computer Science & Engineering, University of Washington, Seattle, WA, USA.

ABSTRACT

Mouse lemurs are the smallest, fastest reproducing, and among the most abundant primates, and an emerging model organism for primate biology, behavior, health and conservation. Although much has been learned about their physiology and their Madagascar ecology and phylogeny, little is known about their cellular and molecular biology. Here we used droplet- and plate-based single cell RNA-sequencing to profile 226,000 cells from 27 mouse lemur organs and tissues opportunistically procured from four donors clinically and histologically characterized. Using computational cell clustering, integration, and expert cell annotation, we defined and biologically organized over 750 mouse lemur molecular cell types and their full gene expression profiles. These include cognates of most classical human cell types, including stem and progenitor cells, and the developmental programs for spermatogenesis, hematopoiesis, and other adult tissues. We also described dozens of previously unidentified or sparsely characterized cell types and subtypes. We globally compared cell type expression profiles to define the molecular relationships of cell types across the body, and explored primate cell type evolution by comparing mouse lemur cell profiles to those of the homologous cells in human and mouse. This revealed cell type specific patterns of primate cell specialization even within a single tissue compartment, as well as many cell types for which lemur provides a better human model than mouse. The atlas provides a cellular and molecular foundation for studying this primate model organism, and establishes a general approach for other emerging model organisms.

INTRODUCTION

Systematic genetic and genomic studies of a handful of diverse organisms over the past half century have transformed our understanding of biology. But many of the most interesting and important aspects of primate biology, behavior, disease, and conservation are absent or poorly modeled in mice or any of the other established genetic model organisms (Elsea and Lucas 2002; Mestas and Hughes 2004; Watase and Zoghbi 2003). Mouse lemurs are the smallest (~50 g), fastest reproducing (2 month gestation, 8 month (minimum) generation, 1-4 offspring per pregnancy) and among the most abundant primates (millions to tens of millions), and an emerging primate model organism (Ezran et al. 2017). Although much has been learned from laboratory studies of their physiology and aging (Hozer et al. 2019; Perret and Aujard 2001; Génin, Schilling, and Perret 2005; Kraska et al. 2011; Terrien et al. 2018; Languille et al. 2012), and from field studies in Madagascar of their ecology, behavior, and phylogeny (Radespiel 2000; Yoder et al. 2000; Rina Evasoa et al. 2018; Atsalis 2015), little is known about their genetics or cellular and molecular biology.

To establish a new genetic model organism, the first step has traditionally been to characterize the wild type and then systematically screen for interesting phenotypes and map the affected genes and underlying mutations or, since the advent of gene targeting, to create targeted mutations and assess their phenotype, as is standard in mouse. Systematic screens are underway for mouse lemur, leveraging their standing genetic diversity and the large pool of naturally-occurring mutations (Karanewsky et al, in prep; Chang et al, in prep), as in human genetics research. The next step is to create a genetic map or reference genome sequence, which is already available for mouse lemur (Larsen et al. 2017) and is becoming increasingly affordable, accurate, and complete with powerful new genomic sequencing techniques (Jain et al. 2018).

With the accompanying development of single cell RNA-sequencing (scRNAseq) technologies (Wu et al. 2013), we reasoned that a reference molecular cell atlas would provide a cellular and molecular foundation that would aid definition and understanding of wild type organism, organ, cell and gene function, enable new types of cellular and molecular screens, and speed genetic mapping, while providing unprecedented resolution and insights into primate specializations and evolution.

Here we set out to create a comprehensive transcriptomic cell atlas of the mouse lemur *Microcebus murinus*, using a similar strategy to the one we used recently to construct the Tabula Muris mouse cell atlas (The Tabula Muris Consortium 2018, 2020) and the human lung cell atlas (Travaglini et al. 2020) (Fig. 1). We adapted the strategy to address several challenges for a new model organism. First, because there was no classical histological atlas, little molecular information, and few cell markers, we relied on the extensive knowledge of human and mouse cell markers (Table 1) from the long history of these fields plus the new atlases. Second, unlike classical model organisms but similar to human studies, donors are of different genetic backgrounds, ages, environments, and diseases. Hence we collected extensive clinical data and histopathology on every donor and organ (Casey et al. 2021), and as in the prior Tabula Muris atlases we procured multiple organs from each donor and processed them in parallel (Fig. 1 a-c, Table 2), thereby controlling for the many technical and biological variables, at least among cell profiles from the same donor. Finally, there is heightened sensitivity for primate studies, so our strategy was opportunistic and designed to maximize information from each donor. To achieve our goal, we brought together experts from diverse fields, including mouse lemur biologists, veterinarians, pathologists, tissue experts, single cell genomics specialists, and computational

leaders, to create the Tabula Microcebus Consortium - a team of over 150 collaborating scientists from over 50 laboratories at fifteen institutions worldwide.

RESULTS

1. Expression profiles of 226,000 cells from 27 mouse lemur organs

The approach used to create a mouse lemur molecular cell atlas is diagrammed in Fig. 1. Two male and two female aged laboratory *Microcebus murinus* mouse lemurs (L1, 9 y/o male; L2, 10 y/o female; L3 11 y/o female; L4, 11 y/o male) were euthanized for humane reasons over a 5 year period due to clinical conditions that failed to respond to therapy, as detailed in the companion manuscript (Casey et al, 2021). Samples were then opportunistically collected for analysis. At the time of euthanasia, blood was drawn, then fresh tissues were quickly isolated (<2 hours post-euthanasia) by a veterinary pathologist and divided into samples that were immediately fixed for pathology or rapidly transported (minutes) at 4ºC to the lab where organ-specific experts dissociated each tissue into single cell suspensions for expression profiling using protocols optimized for each organ (Fig. 1a and Supplementary Methods). Full veterinary evaluation, clinical pathology, and histopathological analysis are provided in a companion manuscript (Casey et al, 2021) as well as metadata for each individual, organ, and cell profiled in Supplementary Methods. This created a classical histological atlas of the mouse lemur (Fig. 1b, https://tabula-microcebus.ds.czbiohub.org/).

From each individual, 3 to 24 organs and tissues were profiled by scRNAseq, cumulatively totaling 27 different organs and tissues, most of which (20) were profiled in at least two subjects (Fig. 1c and Table 2). Beyond the 19 profiled in mouse for Tabula Muris Senis (aorta, bladder, bone marrow, brain, diaphragm, fat depot (mesenteric, subcutaneous, brown interscapular, and peri-gonadal), heart, kidney, colon, limb muscle, liver, lung, mammary gland, pancreas, skin, spleen, thymus, tongue, and trachea), we also profiled blood, bone, brainstem, pituitary gland, retina, small intestine, testis, and uterus (Fig. 1c).

RNA sequencing libraries were prepared directly from single cell suspensions of each organ from each individual using the droplet-based 10x Chromium (10x) protocol. For most organs, aliquots of the cell suspensions were also FACS-sorted for live cells, and for heart cardiomyocytes were hand-picked, and then RNA sequencing libraries were prepared robotically from individual cells using the plate-based SmartSeq2 (SS2) protocol. The 10x and SS2 libraries were sequenced to achieve saturation on an Illumina NovaSeq 6000 System (10x, 26 bp and 90 bp paired-end reads; SS2, 2 x 100 bp paired-end reads). The higher throughput and lower cost of 10x allowed profiling of more cells, whereas SS2 provided greater transcriptomic coverage that aided cell classification, detection of genes expressed at low levels, and gene discovery and structure characterization (accompanying manuscript, in prep). The 10x sequencing reads were demultiplexed, assigned to cells, and aligned to the *M. murinus* genome (assembly: Mmur_3.0, accession: GCF_000165445.2; annotation: NCBI Refseq Annotation Release 101) using Cell Ranger (10xGenomics). SS2 sequencing reads were aligned to the same genome annotation using STAR two-pass mapping, in which the first pass identified splice junctions that were added to the gene reference to improve second pass mapping. For both platforms, the aligned reads were counted (as unique molecular identifiers (UMIs) for 10x, and as reads for SS2) to determine the level of expression in each cell for each gene and scaled [10x: ln(UMI/N *1e4 +1); SS2: ln(reads/N *1e4 +1)] using Seurat v2 (Butler et al, 2018, Nature Biotech). After initial filtering of cell expression profiles for gene count (<500 expressed genes) and sequencing read/UMI parameters (<1000 UMIs for 10x and <5,000 reads for SS2, except for heart as described in Methods) and removing cells with profiles compromised by index switching (Larsson et al. 2018), we obtained 244,081 cell profiles. We then identified low quality profiles (predominantly tRNA, rRNAs, and/or high levels of immediate early genes) and putative cell doublets, leaving

226,701 high quality single cell transcriptomic profiles: 214,890 from 10x (16,682 - 88,910 per individual) and 11,811 from SS2 (394 - 6,723 per individual), distributed across four individuals and 27 organs (Fig. 1 c-d, Table 2).

To identify cells with similar gene expression profiles in each organ, the profiles obtained from 10x scRNAseq analysis of each organ of each individual were separately analyzed through dimensionality reduction by principal component analysis (PCA), visualization in 2-dimensional projections with t-Stochastic Neighbor Embedding (t-SNE) and Uniform Manifold Approximation and Projection (UMAP), and clustering by the Louvain method in Seurat (version 2). For each obtained cluster of cells with similar profiles, the tissue compartment (epithelial, endothelial, stromal, immune, neural, and germ) was assigned based on expression of the mouse lemur orthologs of compartment-specific marker genes from mouse and human (Table 1) (Travaglini et al. 2020). Cells from each compartment of each organ from each individual were then separately and iteratively clustered until the differentially-expressed genes that distinguished the resultant cell clusters were deemed not biologically relevant (e.g., stress or ribosomal genes) (Fig. 1e). Each resultant cluster was assigned a cell type designation, as detailed below.

We then integrated the cell expression profiles obtained from SS2 scRNAseq analysis with the 10x dataset from the same organ and individual (Fig. 1 e-f). For this cross-platform integration we used FIRM (Flexible Integration of single-cell RNA-sequencing data for large-scale Multi-tissue cell atlas datasets), an algorithm that accounts for differences in cell type composition between the two platforms by finding the nearest 10x cell cluster for each SS2 cluster in the same embedded space, then subsampling the data to ensure that the proportion of cells from each cluster match across both platforms, and finally rescaling the gene expression

values accordingly to re-calculate the embedding (Ming et al. 2021). Compared to existing batch correction algorithms, this method provided superior integration of shared cell types across platforms while preserving the original structure for each dataset (Fig. 1 f-g). The cell type designation of each SS2 cell profile was automatically assigned based on the designation of the neighboring 10x cells, and manually curated.

We next used FIRM to integrate the combined 10x/SS2 datasets for each organ across the 2-4 profiled individuals, and then to integrate the profiles across all 27 organs into a single embedded space (Fig. 1 e-g). At each integration step, cell type designations were manually verified by a single expert to ensure consistency of nomenclature throughout the atlas. In this way, cell profiles that co-clustered across organs were given the same cell type designation, and clusters distinguished only in the merged atlas (e.g., related cells too rare in an individual organ to separate from other clusters in that organ) were re-assigned their own cell type designation. This approach identified 768 molecularly-distinct cell populations ("molecular cell types") across the 27 profiled organs, with 28 $\pm$17 (mean $\pm$SD) cell populations per organ and 294 $\pm$1007 cells in each population (Fig. 1c-d, Table 2), which were given 256 different cell designations as detailed below (Fig. 2a).

2. Identification of hundreds of mouse lemur cell types and their expression profiles

To assign provisional cell identities and names to the 768 molecular cell types, we first compiled a list from the literature of canonical marker genes for all of the mouse and human cell types in each compartment of the 27 profiled organs, and found the orthologous mouse lemur genes (Table 1 and Methods). We then searched among the cell clusters of each organ compartment for clusters enriched in expression of each set of cell type marker genes and assigned cells in those clusters the name of the corresponding human and/or mouse cell type and

their corresponding cell ontology (Diehl et al. 2016); for cell types with small numbers, we used expert, biologically-guided manual curation. This allowed us to name almost all cell populations, although for many (34) cell designations there were multiple corresponding cell populations (Fig. 2 a-b, Fig. S1). This was resolved by adding a suffix to the cell designations recognizing a differentially-expressed gene or gene signature that distinguished them (see Section 4). For lung and muscle, we verified that these manual, expert-curated cell type assignments showed good agreement with automated cell identity assignments using Self-Assembling Manifold mapping (SAMap) (Tarashansky et al. 2021), an algorithm that computationally aligns the mouse lemur cell expression profiles to well curated and validated cell profiles we previously obtained for the corresponding mouse and human organs (see Section 6). We identified the differentially expressed genes that are enriched in each cell type relative to the entire atlas, to other cell types of the same tissue, and to the same compartment of that tissue (Table 3).

Examples of the identified and named cell types for six of the profiled tissues are shown in Fig. 3a (kidney) and Fig. S2 (brain cortex, brainstem, hypothalamus/pituitary, eye retina, limb muscle). For limb muscle (Fig. S2 e-f), we identified 31 molecular cell types distributed across the endothelial (7 types), stromal (10 types), and immune (14 types) compartments. We also profiled, though less deeply, another skeletal muscle (diaphragm) and identified 13 of the corresponding cell types. Stromal cells of limb muscle included putative tendon cells (express tenocyte markers tenomodulin *TNMD*+ and scleraxis *SCX*+) and adipocytes (adiponectin *ADIPOQ*+). Fatty infiltrates are rarely seen in aged murine skeletal muscle (Pisani et al. 2010) but are common in aged human muscle (Tuttle, Sinacore, and Mueller 2012), suggesting mouse lemur could provide a valuable model of fatty infiltration of muscle during human aging. We also identified vascular smooth muscle cells (*ACTA2+ MYH11+)* and pericytes (*ACTA2+*

HIGD1B+) as well as cells expressing mature myocyte markers (*ACTA1+, MYL1+*), presumably the self-sealed fragments of mature myofibers. Based on expression of troponin isoforms, myofibers were further subdivided into fast (*TNNT3+, TNNI2+, TNNC2+*) and slow (*TNNT1+, TNNI1+, TNNC1+*) myocytes.

We identified two putative stem/progenitor cell populations in limb muscle (Fig. S2 e-f). One was a cell cluster in the myocyte/myogenic compartment that expressed the human and mouse muscle stem cell (MuSC, or "satellite cell") transcription factors *MYF5* and *PAX7*. These cells also expressed *CD56 (NCAM1)* and *CD82*, cell surface markers used to purify human MuSCs (Alexander et al. 2016; Xu et al. 2015), as well as *VCAM1* or *ITGA7* used to purify mouse MuSCs (Liu et al. 2015; Sacco et al. 2008). A similar population of putative MuSCs (*CD56+ CD82+ VCAM1+ ITGA7+*) was found in the diaphragm. We also identified putative fibroadipogenic progenitors (FAPs), which in humans and mice give rise to fibroblasts and adipocytes (and perhaps chondrocytes and osteoblasts) and promote MuSC-mediated regeneration and sustain the MuSC pool (Joe et al. 2010; Uezumi et al. 2010; Wosczyna et al. 2019). The putative lemur FAPs were found in both limb muscle and diaphragm as stromal populations that selectively expressed *PDGFRA* and *THY1*, surface markers used to purify human FAPs (Uezumi et al. 2016) (note that *LY6A* (*SCA-1*), used to purify mouse FAPs (Joe et al. 2010) is not annotated in the lemur genome). In a companion paper (de Morree et al, in prep), we use the identified surface markers to purify and functionally characterize these putative stem/progenitor populations from lemur, and show they exhibit many characteristics more similar to their human counterparts than those of mouse.

From blood, we identified mouse lemur cognates of all the major human and mouse immune cell types including, in the lymphoid lineage, B cells, plasma cells, *CD4+* T cells,

CD8+ T cells, natural killer cells, natural killer T cells, and innate lymphoid cells; and in the myeloid lineage, erythrocytes, platelets, monocytes, macrophages, conventional and plasmacytoid dendritic cells, neutrophils, basophils, and even the rare and fragile eosinophils (Fig. 2). In the bone, bone marrow and other hematopoietic and lymphoid tissues, we identified presumptive progenitors including hematopoietic precursors, and progenitors of erythrocytes, megakaryocytes and granulocyte-monocytes and putative adipogenic and osteogenic progenitors. However, certain immune cell subtypes in human and/or mouse were not identified in lemur. For example, lemur monocytes formed a single cluster in most tissues that could not be resolved into classical and non-classical monocytes using the markers that distinguish the two cell types in human (*CD14, CD16*) or mouse (*LY6C1/2,* which has no primate ortholog*, CCR2, CX3CR1*) (Italiani and Boraschi 2014). Conventional dendritic cells (cDCs) detected in the lemur analysis could not be divided into type 1 and type 2 subtypes characteristic of humans and mice. Conversely, other dendritic molecular types were identified in lemur (e.g. *FLT3+ IGSF6+* DCs) that had no apparent human or mouse cognates. The full spectrum of developing and mature immune cells across the body allowed us to reconstruct the early stages of hematopoietic development (Section 3), as well as the subsequent dispersal of mouse lemur immune cells throughout the body and its alteration by disease (accompanying manuscript, in prep).

3. Molecular gradients of cell identity

Although the vast majority of profiled lemur cells could be computationally separated into discrete clusters of cells with similar expression profiles, we found many examples where cells instead formed a continuous gradient of gene expression profiles indicating a gradual transition from one molecular identity to another within the tissue. Some of these reflect a spatial

gradient of cell identity in the tissue, whereas others correspond to an ongoing developmental process or the induction of a physiological or pathological cell state.

The kidney provided a dramatic example of spatial gradients of cell types that are key to organ function. Coordinated gradients of renal epithelial and endothelial cell types define the position-specific molecular features and functions along the nephron that set fluid and electrolyte balance and other aspects of renal physiology. Among the ~14,800 cells profiled from lemur kidneys, we identified lemur cognates of almost every important cell type known from human and mouse in each of the major tissue compartments. These included cell types of the renal tubule and collecting duct, macula densa cells, and podocytes from the epithelium, and cells of the glomerulus and the vasa recta from the endothelium (Fig. 3 a-f). The only cell types we did not identify were parietal epithelial cells of the glomerulus, renin-secreting juxtaglomerular complex cells, and mesangial cells (see Section 4). However, most notable among the profiled kidney cells were the many epithelial cells that formed a long continuous gradient of molecular identity (Fig. 3 b-c, Fig. S3 a, g-h). With canonical renal markers, we determined that this gradient corresponds to the spatial gradient of epithelial cell types along the mouse lemur nephron, starting from proximal convoluted tubule cells, through the loop of Henle, and ending with principal cells of the collecting duct. Interestingly, macula densa cells, the sodium-sensing cells that regulate glomerular blood flow and filtration rate that are typically physically located at the distal end of the loop of Henle before the start of the distal convoluted tubule (Peti-Peterdi and Harris 2010), surprisingly localized in the lemur molecular cell gradient between cells of the thin and thick loop of Henle. We also found a prominent cell gradient in the endothelial compartment, with arterial markers (*GJA5+, BMX+*) expressed at one end of the gradient and venous markers (*ACKR1+, NR2F2+*) at the other, and capillary markers (*CA4+*) in between

(Fig. 3 d-e, Fig. S3 b, i-j). This appears to comprise the vasa recta, a network of arterioles and venules intermingled with the loop of Henle in the renal medulla, because it expressed some specific vasa recta markers (e.g., *AQP1, SLC14A1* for vasa recta descending arterioles) and was molecularly distinct from the clusters of glomerular endothelial cells (*EDH3*+), other capillary endothelial cells (possibly peritubular), and lymphatic endothelial cells (*CCL21*+). This deep molecular map of the lemur nephron also revealed region-specific hormonal regulation of nephron function (see accompanying hormone atlas manuscript, Liu et al, in prep).

Several other gradients of gradually changing gene expression profiles represent cells differentiating in adult stem cell lineages. Two such gradients observed among immune cell populations of bone marrow represent the ongoing development and maturation of hematopoietic progenitors (Fig. 3 i-j, Fig. S3 d-f, k-l). One gradient bifurcates into the monocyte/macrophage and the granulocyte/neutrophil lineages (accompanying paper, in prep), whereas the other represents the erythroid lineage. Some common but more subtle gradients marked differentiation of basal epithelial cells in the skin and at least four other organs (small intestine, colon, tongue, bladder) into their corresponding mature epithelial cell types along the suprabasal/luminal axis of each organ.

The most striking developmental gradient appeared in the male gonad. Among the ~6500 cells profiled from mouse lemur testes, we found on computational clustering that all except stromal and immune cells formed a single long continuous gradient, which the gene expression profiles indicated was a continuum of germline cells progressing through spermatogenesis (Fig. 3 g-h, Fig. S3c). The gradually changing expression levels of genes across the continuum allowed us to reconstruct the full gene expression program of mouse lemur spermatogenesis, assigning seven canonical stages from stem cells (spermatogonia) to mature spermatids using

orthologs of stage-specific markers from human and mouse (see Section 6). In addition to the essential role of male germ cell differentiation in reproduction, the mouse lemur spermatogenesis program is of special interest because such programs are rapidly evolving during primate speciation, with several notable evolutionary specializations already recognized for mouse lemur, including seasonal regulation and sperm competition (Andrès, Solignac, and Perret 2003).

4. Previously unknown or uncharacterized cell types and subtypes

Although we were able to assign a provisional identity to the vast majority of the cell populations based on expression of orthologs of canonical markers of human and mouse cell types, there were dozens of cases in which more than one cluster in a tissue expressed markers of the same cell type and their separation could not be attributed to technical differences (e.g., cell quality, batch effect, cell doublets) (Fig. 4a). In some cases these appear to be multiple states of the same cell type because the differentially-expressed genes that distinguished the clusters included proliferation markers (e.g., *MKI67, TOP2A)* indicating a proliferative state (hence PF was added to cell type name) or differentiation markers indicating a differentiating state. In most cases, however, the additional clusters appear to represent previously unknown or little characterized cell types or subtypes, and a distinguishing gene (or two) was added to the name of the minor cluster (e.g. B cell (*SOX5*+) in pancreas), or to both clusters if they were similar in abundance. Such clusters were uncovered in all compartments except the germline and in most profiled organs and tissues (Fig. 4a, Fig. 2b).

Fibroblast subtypes were particularly diverse (Fig. 4a). In the lung, we identified adventitial and alveolar fibroblasts, as well as two other clusters of fibroblasts. We also uncovered four molecular subtypes of fibroblasts in bladder, fat, and small intestine, three in kidney, and two in pancreas and tongue, most of which did not have known parallels in human or

mouse. Most appear to be organ-specific because there was little co-clustering across organs. Similarly, macrophages formed multiple distinct molecular clusters in many tissues, with substantial diversity within lung, liver, and kidney (Fig. 4a, and accompanying paper, in prep). There was also extensive molecular diversity among T cells that could not be readily harmonized with classical T cell subtypes defined in humans and mice (Martin and Badovinac 2018; Farber, Yudanin, and Restifo 2014).

We also identified interesting molecular subtypes of epithelial cells. These include three molecular types of pancreatic acinar cells (main, *FDX1+, PNISR+*) and two pancreatic ductal types (*SPOCK3+, UPK1A+*), three types of kidney collecting duct principal cells (main, *FXDY4+, KCNE1+*), and two types of small intestine enterocytes (main, *CA2+*) (Fig. 4a). We also found two molecular types of hepatocytes (Fig. 4 b-c): the *APOB+* population expressed higher levels of many classical liver proteins (e.g., apolipoprotein B (*APOB*), fibrinogen gene activator (*FGA*), histidine-rich glycoprotein (*HRG*)) and certain hormones (e.g. angiotensin (*AGT*), *FGF21*) and receptors (glucagon receptor (*GCGR*), hepcidin receptor (*SLC40A1*), androgen receptor (*AR*), growth hormone receptor (*GHR*), and prolactin receptor (*PRLR*)), whereas the *PHYH+* type expressed higher levels of some metabolic genes (e.g. *PHYH*, fatty acid degradation) and methylation regulator (*GNMT)* (Fig. 4c) (see also accompanying hormone atlas paper, Liu et al., in prep). The two types do not correspond to the known zonal heterogeneity of human and mouse hepatocytes (Ben-Moshe and Itzkovitz 2019), but notably the *APOB+* hepatocytes expressed far more transcripts and genes than *PHYH+* hepatocytes so the *APOB+* cells could correspond to the larger, polyploid hepatocytes and *PHYH+* cells to the smaller, diploid hepatocytes found in human and mouse (Celton-Morizur and Desdouets 2010) (Fig. S4). We uncovered a similar molecular distinction among hepatocytes in our mouse

scRNAseq dataset (Fig. S4) (The Tabula Muris Consortium 2020), indicating these are conserved but previously unrecognized molecular subtypes of hepatocytes.

In the endothelial compartment, there was substantial diversity among blood capillary cells including molecular subtypes found in multiple tissues. *FABP5+ RBP7+* capillary cells were found in 10 tissues, *FABP5- RBP7-* cells in seven, variants dubbed *FABP5+ RBP7-* and *FABP5+ RBP7*lo in five and four, respectively, and a *CXCL10+* population in four (Fig. 2b). In some tissues endothelial cells were comprised exclusively of one of these types (e.g. heart, *FABP5+ RBP7+*; brain cortex, *FABP5- RBP7-*; mammary gland, *FABP5+ RBP7*lo), but other tissues (e.g., kidney) contained a mix of two or more (Fig. 2b, Fig. 4a). *FABP5+ RBP7+* capillary cells appear to be specialized for energy storage because they are enriched in genes for fatty acid uptake and binding (e.g., retinol-binding protein 7 (*RBP7*), *FABP1*, *FABP4*, *FABP5*), as well as for transcriptional factors *MEOX2* and *TCF15* (Fig. 4d), all genes that are enriched in capillary cells of several high energy-demand tissues in human and mouse (Coppiello et al. 2015; Iso et al. 2013; Caprioli, Zhu, and Sato 2004; Hu et al. 2017). *CXCL10+* capillary cells express genes associated with interferon activation (*CXCL9, CXCL10, CX3CLl, GBP1, GBP2, IFIT3*) (Du et al. 2020) indicating an inflammatory state in the fat, kidney, and limb muscle of L2 and the lung of L4 (accompanying paper, in prep). We also identified molecular subtypes of lymphatic endothelial cells: *CCL21+* and *CCDC80+* in various tissues and *CCDC80+ ACKR4+* in pancreas, perhaps representing different lymphatic cell types in peripheral vessels and lining lymph node sinuses (Takeda et al. 2019; Fujimoto et al. 2020; Jalkanen and Salmi 2020) (Fig. 4a).

For eight of the 768 molecular cell types (~1%) we were unable to assign a specific identity, so they were designated “unknown” along with the organ they were isolated from and

the compartment inferred from expression of canonical compartment markers. These include mysterious stromal types in tongue (#142) and kidney (#143), and epithelial types in fat (#15) and blood (#16) (Fig. 4 a, e-f, Fig. S5 c-m). The remaining four stromal types, from bone, mammary gland, pancreas, and tongue, were given the same designation (#141, "unknown stromal *NGFR+ TNNT2+*") because they shared similar transcriptomic profiles (including high expression of *NGFR, IGFBP6, OGN, ALDH3A1, CLDN1, ITGB4*) and curiously also expressed high levels of cardiac troponin T (*TNNT2*), a component of the actomyosin contractile apparatus that is a sensitive and specific cardiac myocyte marker used clinically as a diagnostic marker for myocardial infarction (Neumann et al. 2019; Jarolim 2015; Zipes et al. 2018) (Fig. S5 a-e). Some of these unknown cell types (#15, #141, #142) could be specialized mesothelial cells, or at least cells that share features with them, because many of their differentially-expressed genes were enriched in mesothelial cells (Fig. S5 e, j). However, some of the differentially-expressed genes were also expressed by other cell types (e.g. leptomeningeal and Schwann cells for #141 and 142, urothelial cells for #15), and in the global comparison of cell types (Section 5) they did not closely localize with any of these cell types. The unknown kidney stromal type (#143) might be mesangial cells because they expressed high levels of *LMO7* and *ITGA8,* recently reported to be enriched in mouse mesangial cells (He et al. 2021) (Fig. 4 e-f, Fig. S5 f-g). The most perplexing unknown was the epithelial (*EPCAM+*) population from blood (#16) of one individual (L2). It had a unique gene signature including *PGAP1, PTPRG, FGF13, SOX2, CPNE6, CDH3*, many of which are also expressed by brain ependymal cells, astrocytes, and oligodendrocytes; however, this population did not express canonical markers of those cell types (Fig. S5 k-m).

Some of the molecular subtypes and unknown cell types described above may represent previously unrecognized or sparsely characterized cell states including pathological states

(accompanying paper, in prep). Others may be new cell types or subtypes, like the novel lung capillary (aerocyte) and epithelial (ionocyte) populations recently uncovered by scRNAseq (Gillich et al. 2020; Plasschaert et al. 2018; Montoro et al. 2018). It will be important to characterize each of these potentially novel molecular cell types and determine which are conserved across species but previously unrecognized, and which are mouse lemur or primate innovations.

5. Global molecular comparison of cell types across the body

To elucidate the molecular relationships of the over 750 defined lemur molecular cell types across the organism identified by our scRNAseq, we compared and visualized the cell type expression profiles in two ways. Typically such relationships are visualized as 2D representations (e.g., UMAP, tSNE) of expression profiles of all cells, in which the distance between cells (dots) indicates the relatedness of their expression profiles and all cells of a particular type are highlighted in their own color. However, such representations of massive datasets like ours are overly complicated and dominated by the most abundant cell types, obscuring all but the most superficial relationships (Fig. 1g, traditional UMAP of all 244,081 cells colored by tissue, see online portal for UMAP colored by molecular cell type). We found that simplified UMAPs, with each molecular cell type condensed into a single point representing the average expression value of all cells of that type ("pseudo-bulk" expression profile), allowed easy visualization of cell relationships across the entire body (Fig. 5 a-f, Fig. S6). Cell relationships could also be discerned in heatmaps showing pairwise correlation coefficients of all cell type "pseudo-bulk" expression profiles displayed as a large (~750x750) matrix, provided that the molecular cell types were arranged in a biologically sensible order, such as by tissue, compartment, or hierarchical clustering of similarity (Fig. 5 g-h, Fig. S7, https://tabula-

microcebus.ds.czbiohub.org/). Both plots revealed global patterns of similarity as well as unexpected cell relationships.

Molecular cell types within a tissue compartment generally showed more related expression profiles, even those from different organs. Endothelial cell types across the body formed the most coherent compartment. Next was the neural compartment including CNS glial cells, which the comparisons showed are surprisingly similar to neurons (Fig. 5 a, g, Fig. S6a). Immune compartment cell types were by far the most divergent, and particularly so between lymphoid and myeloid populations. But the global analyses also identified specific cell types that were more closely related to cell types in another compartment than to those in its own compartment. Some of these cross-compartment similarities were predictable. For example, neuroepithelial cells of the airway (neuroendocrine cells) and gut (enteroendocrine cells) were found to be more closely related to neurons and pituitary neuroendocrine cells than to most other epithelial cell types (Fig. 5 a, g, Fig. S7, https://tabula-microcebus.ds.czbiohub.org/). Likewise, proliferating skin intrafollicular cells were more similar to certain proliferative cells in other compartments (e.g., proliferating B cells, T cells, and neutrophils, and erythroid progenitors and skeletal muscle stem cells) than to other non-proliferating epithelial cell types (Fig. 5h); indeed such global comparisons of transcriptomic profiles are a facile way of identifying progenitors and proliferating cell populations.

Other identified cross-compartment similarities were surprising. The most striking was that male germ cells (spermatogonia) are more closely related to immune progenitor cells than to any other cells in the atlas, including progenitors and proliferating cells of other compartments (e.g., intestinal enterocyte progenitors, skin proliferating interfollicular cells, skeletal muscle stem cells) (Fig. 5 a, h). Maturing male germ line cells formed a developmental trajectory

emanating out from the immune progenitor cluster just like the various immune cell lineages (Fig. 5a). Differential gene expression analysis showed that this similarity included the common enrichment in both spermatogonia and hematopoietic progenitors of specific cell cycle genes, particularly genes involved in M phase (e.g., *CCNB1, VRK1, GINS2, CEP55, BUB1, TTK, FOXM1, CDCA3, DEPDC1B, SPAG5, BRCA1, NUDT1, CENPS*, *MKI67,* and *DBF4*) (Fig. 5b, Fig. S6b), indicating similarity in the mitotic machinery of proliferating spermatogonia and hematopoietic progenitors. Compared to progenitors and proliferating cells of other compartments, spermatogonia and immune progenitors also showed selective expression of non-cell-cycle genes (e.g., *TESMIN, RSPH14, ASB9, ARHGAP15, HEMGN, PLD6, ACAP1, NARF, PHACTR1, PLPPR3, EGFL7, DKC1, LOC105869343* (serine protease 40)*, LOC105861000*) (Fig. 5c, Fig. S6c).

Another surprising example of cross-compartment similarity was myelinating and non-myelinating Schwann cells of the peripheral tissues, which segregated with stromal cells and away from other neural compartment cells including non-myelinating Schwann cells of brain cortex (Fig. 5a). Differential expression analysis identified genes enriched in stromal and Schwann cells but not other neural compartment cells (e.g., *COL3A1, LAMC1, SOCS3, COL1A2, COL5A2, ID3, ID1, CDC42EP5, MMP2, TGFBR2, CCN1, TBX3, LOC105874070 (IFITM1-3), LOC105876248*) (Fig. 5d, Fig. S6d) and a complementary set enriched in neurons and CNS glial cells but not the PNS Schwann cells (e.g., *OMG, GPR137C, TCEAL3, MAPK10, ASTN1, NALCN, MAP7, CXXC4*). Many of the genes expressed in common by Schwann and stromal cells are components or regulators of extracellular matrix, suggesting an important and distinctive function of peripheral glia in collaborative remodeling of the extracellular matrix with surrounding stromal cells.

Within a compartment, cells of the same designated type or subtype also generally clustered closely despite their different tissue origins, most notably for endothelial, stromal, neural, and immune compartment cell types and subtypes (e.g., vascular smooth muscle cells, pericytes, mesothelial cells, five types of dendritic cells (conventional, plasmacytoid, mature, *FLT3*+ *IGSF6*+ dendritic cells, and a skin and tongue population), three types of neutrophils (main, *IL18BP*+, *CCL13*+), and two major types of lymphatic endothelial cells (*CCDC80*+ and *CCL21*+), and adipocytes (*UCP1*low, *UCP1*high) (Fig. 5 a, f, Fig. S6 e-f). In contrast, epithelial cell types were highly tissue-specific and generally clustered with other epithelial cells from the same organ, even those for which there were cell types of the same designation in another organ (e.g., skin and tongue basal and suprabasal cell types) (Fig. 5e). One example of initially perplexing cross-organ similarity was a distinctive population of epithelial cells from the lung of one individual (L2), which curiously clustered closely with uterine epithelial cells; these were later shown to be lung metastases of a uterine cancer (accompanying paper, in prep).

6. Evolution of gene expression in primate cell types

To elucidate how gene expression has changed during primate evolution, we compared the transcriptomic profiles of mouse lemur cell types to the corresponding cell types of human and mouse, using mouse as the non-primate outgroup. To ensure comparisons were made across truly homologous cell types and to minimize technical artifacts, we used our own human (Travaglini et al. 2020; The Tabula Sapiens Consortium 2021) and mouse datasets (The Tabula Muris Consortium 2018, 2020) obtained with the same scRNAseq method and clustered and annotated by canonical markers in a similar way by the same tissue experts. We focused our analysis on lung and on skeletal (limb) muscle cell types for which we had high quality profiles and annotations from all three species. Our assignments of homologous cell types were verified by

SAMap analysis (Fig. S8 a, c), which uses a self-assembling manifold algorithm and graph-based data integration to identify homologous (reciprocally-connected) cell types with shared expression programs across species (Tarashansky et al. 2021), and also by showing that the corresponding cell types from the three organisms co-clustered when their expression profiles were adjusted by BBKNN batch correction (Polański et al. 2020) and compared and displayed by dimensionality reduction (Fig. S8 b, d). We restricted the analysis at the gene level to the 13,302 one-to-one-to-one gene orthologs across all three species that we curated by assigned gene names from the ENSEMBL and NCBI databases (Table 4).

For each cell type, we computed the Xi correlation (Chatterjee 2019) of the expression profiles of 1000 trios of randomly-selected cells from human, lemur, and mouse, and then calculated the difference in correlation values between each pair of species. We found that the transcriptomic profiles of most lemur lung cell types analyzed (13 of 22, 59%) were significantly more similar to the homologous cell type of human than to that of mouse (e.g. ciliated cells, plasma cells, capillary aerocytes), although the degree of divergence from mouse varied by cell type (Fig. 6a). For five lung cell types (5 of 22, 23%), the lemur expression profile was equally similar to both. Surprisingly, the profiles of the other lemur lung cell types analyzed (alveolar macrophages and proliferating macrophages, NK T cells, CD4+ T cells) were significantly more similar to the corresponding mouse cell type than to the human, suggesting a high degree of evolutionary adaptation in gene expression for those cell types in the human lineage. Likewise, we found that the expression profiles of human and mouse lung lymphatic endothelial cells and CD8+ T cells were more similar to each other than to the corresponding cell types in lemur, suggesting a high degree of evolutionary adaptation in gene expression in those two cell types in the lemur lineage. The patterns of evolutionary adaptation in cell type gene expression varied

substantially even within a compartment. For example, the recently identified capillary endothelial "aerocytes" that mediate alveolar gas exchange (Gillich et al. 2020) showed among the highest degree of primate specialization of any lung cell type, whereas arterial and venous endothelial cells showed similar divergences across all three species and lymphatic endothelial cells showed the highest degree of lemur specialization. Likewise in the lymphoid compartment, plasma cells showed among the highest degree of primate specialization of any cell type, whereas B cells (from which plasma cells arise) showed only limited primate specialization, and CD4+ T cells showed the highest degree of human specialization and CD8+ T cells among the highest degree of lemur specialization. The same cross-species comparison of limb (skeletal) muscle cell types gave similar results, although 40% (4 of 10) of the analyzed lemur cell types were closer to the corresponding mouse cell type than to the human, and nearly all human cell types (80%, 8 of 10) were closer to the corresponding lemur cell type and none were closer to those of mouse (Fig. S8e). Interestingly, although human muscle *CD8+* T cells were more similar to those in lemur, human lung *CD8+* T cells were more similar to those in mouse, suggesting potential species differences in organ-specific influences on *CD8+* T cells. Thus, expression patterns of each cell type have apparently changed at different rates and to different degrees during primate evolution, including cell types in the same compartment as well as related cell types in different organs.

To provide molecular insight into these cell type specializations in primate evolution, we identified for each of the lung and muscle cell types the genes with expression levels in lemur and human substantially differ from those in mouse (>10-fold difference, $p < 1e-5$). These included general primate differences (both lemur and human vs mouse) as well as lemur-specific and human-specific differences (Table 5, Fig. S9). Dozens of genes showed such primate-

specific expression patterns for each cell type (58 ±31, mean ±SD; range 11 - 113 genes for each lung cell type; 57 ±30, range 11 - 103 genes for each muscle cell type), many of which are selectively expressed in the cell type or its compartment and serve canonical functions (Fig. 6b). Examples for lung include genes involved in immune (*PIGR, CXCL17*) and stem/progenitor (*RASGRF1, ERBB3*) functions of alveolar type 2 (AT2) cells; cilia structure (*SPATA4, WDR38, WDR54*) (Choksi et al. 2014; Albee et al. 2013)) and mucin production genes (*MUC20*) of ciliated cells; extracellular matrix genes (*LUM, DCN, SGCA*) in stromal compartment cell types (adventitial and alveolar fibroblasts, pericytes); pore forming cytolytic protein (*PRF1*), granzyme M (*GZMM*), and other immune-related genes (*PTGDR, KLF12, TIGIT*) in T and NK cells; and bacterial receptor *CD163*, *OSCAR*, and other immune-related genes (*CXCL16, IL7R, MSR1, PTAFR, RAB20, VSIG4*) in monocytes and macrophages.

Other genes enriched in primate cell types suggest primate-specific communication between cell types. For example, the classical vasopressor angiotensin II can target pericytes of lung and muscle via its receptor *AGTR1*, and vasoactive intestinal peptide (VIP) can target vascular endothelial cells of lung (but not those of muscle) via its receptor *VIPR1*. Also, insulin growth factor signaling can originate from lymphatic cells via the ligand *IGF1*, and Hedgehog signaling can be modulated in the lung by AT2 cell expression of the secreted inhibitor *HHIP*, implicated in chronic obstructive pulmonary disease (COPD) (Van Durme et al. 2010). We also identified many genes that were more broadly expressed in the lung and showed primate-specialization across many or all lung cell types (e.g., *ABHD2, BTG3, GNG7, KCNA3, TYMP*) (Fig. 6b, Fig. S9a, Table 5); many of these were also broadly expressed and primate enriched in muscle (Table 5, Fig. S9g). The complementary analysis of genes highly enriched in mouse cell types relative to the corresponding primate cell types identified hundreds of such genes (Fig.

S9b, Table 5), suggesting that cell type gain of expression in the mouse lineage (or loss in the primate lineage) also contributed substantially to primate-specific patterns of gene expression.

Comparison of the lemur spermatogenesis program (Section 3) with the corresponding expression programs recently elucidated in human and mouse (Hermann et al. 2018; M. Wang et al. 2018; Shami et al. 2020) Many of the marker genes that demarcate canonical stages of spermatogenesis are similarly expressed in all three species (e.g., *KIT, SOHLH1*, spermatogonia; *SYCP3*, early spermatocytes; *ACR*, pachytene/diplotene spermatocytes; *TNP1*, early/mid spermatids; *GAPDHS*, mid/late spermatids; Fig. 6c), defining a conserved core program of mammalian spermatogenesis. However, our analysis also identified primate-specific features of the program. Dozens of "primate-only" genes (orthologs identified in human and mouse lemur but missing in mouse) are selectively expressed during spermatogenesis, indeed these were the dominant class of "primate-only" genes (accompanying paper, in prep). We also uncovered genes that have undergone heterochronic changes in expression such that orthologous genes are expressed at different stages of the male germ cell differentiation program in different species, indicating re-wiring of aspects of the spermatogenesis program during evolution. Examples include *NME8*, which peaks in expression in pachytene spermatocytes of primates but in early spermatids of mice; *ID4*, which is expressed only in spermatogonia for primates but continues to be expressed in spermatocytes for mice; *GFRA1*, which is expressed in spermatogonia of all species but re-expressed in late spermatocytes and spermatids of primates (Fig. 6c); and the chromosome-wide down regulation of X-chromosome genes (meiotic sex chromosome inactivation, MSCI) ((Turner 2015) that occurs in pre-pachytene spermatocytes in primates (Fig. 6e) but pachytene spermatocytes in mouse (Ernst et al. 2019). One lemur-specific heterochronic change is in expression of *PIWIL1* (also known as MIWI in mice), a repressor of

retrotransposons during meiosis (Reuter et al. 2011). In both humans and mouse the gene is expressed throughout meiosis from early spermatocytes to diplotene and even in secondary spermatocytes with peak expression in pachytene spermatocytes (Jan et al. 2017; Watanabe et al. 2015). In contrast, *PIWIL1* expression in lemur was restricted to early spermatocytes, with reduced expression in pachytene spermatocytes (Fig. 6 c-d). Lemur-specific specializations like this in the spermatogenesis program are of particular interest because of their potential role in the remarkable radiation of the Lemuroidea clade (over 100 lemur species, nearly one-quarter of all primate species), and because several notable evolutionary specializations have already been recognized for mouse lemur, including the dramatic seasonal regulation of the testes and the role of sperm competition in reproduction (Andrès, Solignac, and Perret 2003).

One particularly prominent set of lemur specializations we uncovered in the spermatogenesis program was in the small subset of X-chromosome genes that become active after MSCI, ones that 'escape' postmeiotic sex chromatin silencing (Sin and Namekawa 2013; Sin et al. 2012; Namekawa et al. 2006). We identified 62 such 'escape' genes in mouse lemur by searching among X-chromosome genes for ones that are highly expressed selectively in the early spermatid stage (Table 6, Fig. 6f), as done for mouse and human spermatogenesis (Sin et al. 2012; Namekawa et al. 2006). Remarkably, 49 (79%) of the identified lemur 'escape' genes were unique to lemur, whereas only 3 (5%) were conserved across all three species (testes-specific histone variant *HYPM* (H2A.P), scaffold protein *AKAP4* required for sperm flagella function and fertility in mouse, and uncharacterized protein *CXHXorf65*), 8 (13%) were shared only with human, and 2 (3%) were shared only with mouse (Table 6, Fig. 6f).

Extending such comparisons to all homologous cell types and more broadly across phylogeny may provide insight into the selection pressures that underlie the evolutionary

differences among cell types. Even at this stage the three-way comparisons (lemur, human, mouse) described here suggest interesting biological hypotheses and identify many cell types and genes for which lemur provides a human modeling advantage over mice.

DISCUSSION

Our single cell RNA-sequencing and analytical pipeline defined over 750 mouse lemur molecular cell types and their expression profiles, comprising nearly all major cell types for most (27) tissues. The cell types include cognates of most canonical human cell types plus stem and progenitor cells and their developmental programs for spermatogenesis, hematopoiesis, and other adult tissues, many of which did not form discrete molecular types but merged into continuous gradients of cell intermediates along the developmental pathway. Biologically-guided expert curation also uncovered dozens of previously unidentified or little characterized cell types and subtypes, some of which appear to be conserved (e.g., two types of hepatocytes, several capillary types specialized for energy storage or activated by inflammation) but others may be primate or lemur-specific innovations.

By organizing the mouse lemur cell types by organ, compartment, and function, and then globally comparing their expression profiles in a simplified UMAP or matrix of all pairwise comparisons, we defined the molecular relationships of cell types across the body. This revealed global features like the high similarity of cell types within some compartments (e.g., endothelial) but marked divergence of cell types in others (e.g., immune), as well as the surprising similarity of a few cell types across compartments, such as spermatogonia to hematopoietic progenitors and peripheral glia to stromal populations.

The atlas provides a broad cellular and molecular foundation for studies of this emerging primate model organism. Although the first steps in establishing a new model organism have traditionally been mutant screens and generation of a genetic map or reference genome, with the technological advances and declining cost of scRNA-seq the creation of a reference transcriptomic cell atlas like ours can now be prioritized. It comprehensively defines cell types

and aids elucidation of their functions (muscle stem cell manuscript, de Morree et al, in prep), and allows molecular comparisons of mouse lemur cell types to each other and to their orthologs in human and mice, allowing exploration of primate biology and evolution at cellular resolution. This revealed cell type specific patterns of primate cell specialization even within a single organ and compartment, and identified the many cell types for which mouse lemur provides a human modeling advantage over mice, as well as cases like sperm with a multitude of primate- and lemur-specific expression innovations. But the atlas also provides a powerful new way of detecting genes, defining their structures and splicing, and assigning their function, as well as elucidating organism-wide processes such as hormonal signaling, immune cell activation and dysregulation, and primate-specific physiology, diseases, and genes, as we show in the accompanying papers (accompanying manuscript, in prep; hormone atlas, Liu et al, in prep).

Our cell atlas strategy can be adapted to other emerging model organisms. This strategy should include the opportunistic identification of donors and systematic procurement of tissues from each; application of the scRNA-seq technologies and analytical pipeline including computational cell clustering, integration, and expert cell annotation; and the biological organization of cell types and comparisons across the organism and between organisms. Importantly, extensive clinical and histopathological metadata should be collected from each donor and tissue to mitigate against complexities arising from genetic and environmental differences, as this can provide insight into individual-specific features of the atlas, as we exploit in the accompanying paper (in prep). Application of this strategy to a wide variety of organisms will rapidly expand our cellular, genetic, and molecular understanding of biology that has been dominated for a half century by a small number of model organisms.

ACKNOWLEDGMENTS

This work was supported by funding from The Chan-Zuckerberg Biohub, Howard Hughes Medical Institute and Vera Moulton Wall Center for Pulmonary Vascular Disease; unrestricted grant from Research to Prevent Blindness and NEI P30-EY026877 to the Stanford Department of Ophthalmology to Albert Y. Wu; The Hong Kong University of Science and Technology Start-up grant R9364, The Hong Kong University of Science and Technology Big Data for Bio Intelligence Laboratory (BDBI), and The Chau Hoi Shuen Foundation special research grant R9056 to Angela Wu; NSF-DBI-1701984 to Anne D. Yoder; funding from Developmental and Stem Cell Biology Graduate Program to Aris Taychameekiatchai; funding from Hong Kong Research Grant Council (16307818, 16301419, 16308120), the Hong Kong University of Science and Technology (startup grant R9405), and The Hong Kong University of Science and Technology Big Data for Bio Intelligence Laboratory (BDBI) to Can Yang; funding from Shanghai Sailing Program to Jingsi Ming; NIA 1K99AG066963 to Thomas Ambrosi; NIH DP2AI138242 to Iwijn De Vlaminck; National Sciences and Engineering Research Council of Canada fellowship PGS-D2 to Michael F.Z. Wang; NIH R35 GM139517, R01 GM116847, R35 GM139517, NSF MCB1552196 to Julia Salzman; NSF Graduate Research Fellowship DGE-1656518 and a Stanford Graduate Fellowship to Julia Olivieri; Cancer Systems Biology Scholars Fellowship (Grant R25 CA180993) and Clinical Data Science Fellowship (Grant T15 LM7033-36) to Roozbeh Dehghannasiri; Urology Care Foundation Research Scholar Award Program and AUA Western Section Research Scholar Fund II to Hosu Sin; Stanford Graduate Fellowship/HHMI/NIH CMB Traing Grant to Yue Zhang; American Cancer Society Postdoctoral Fellowship to SoRi Jang; Walter V. and Idun Berry Postdoctoral Fellowship to Andrea R. Yung; NSF Graduate Research Fellowship and Stanford Graduate Fellowship to Youcef Ouadah; the Wu Tsai Neurosciences Institute Interdisciplinary Scholar Award to Shixuan Liu; European Community's 7th Framework Programme (FP7/2007-2013) under grant agreement Nbr 278486 (DEVELAGE), from the Fonds Unique Interministériel and Région Languedoc-Roussillon under grant agreement Nbr 110284 (DiaTrAl), and from the Fondation Plan Alzheimer (PRADNET) to Jean-Michel Verdier and Corinne Lautier; R01 AI024258 to Peter Parham and Lisbeth A. Guethlein; NIH R01DC016892 to Wan-Jin Lu; NIH P30DK116074 to Yan Hang; NSF Graduate Research Fellowship to Connor V. Duffy; Postdoctoral Fellowships from the DFG (NE 2006/1-1) and California TRDRP (25FT-0011) to Patrick Neuhöfer; funding from Independent Research Fund Denmark (DFF–5053-00195) and Lundbeck Foundation (R232-2016-2459) to Jean Faru; NIH AG068667, AR073248 and AG036695 to Thomas A. Rando; funding from Wu Tsai Neurosciences Institute to Tony Wyss-Coray; NSF BCS 0647402 to Liza Shapiro and E. Christopher Kirk.

METHODS

Animal husbandry

Microcebus murinus gray mouse lemurs originated from the closed captive breeding colony at the Muséum National d'Histoire Naturelle in Brunoy, France, and transferred to the University of Texas (Austin) and then Stanford University and maintained for noninvasive phenotyping and genetic research as approved by the Stanford University Administrative Panel on Laboratory Animal Care (APLAC #27439) and in accordance with the *Guide for the Care and Use of Laboratory Animals*, as detailed in (Casey et al. 2021). Briefly, mouse lemurs were individually or group housed indoors in an AAALAC-accredited facility in a temperature (24°C) and light-controlled environment (daily 14:10 h and 10:14 h light:dark alternating every 6 months to stimulate photoperiod-dependent breeding behavior and metabolic changes) with perches and nest boxes, and were fed fresh fruits and vegetables, crushed primate chow plus live insect larvae as enrichment items. Health and welfare were routinely monitored and clinical care provided by the Veterinary Service Center.

Tissue procurement and processing

Animals in declining health that did not respond to standard therapy were euthanized by pentobarbital overdose under isoflurane anesthesia as described (Casey et al. 2021). Prior to euthanasia, a veterinary examination was performed, and animal body weight and electrocardiogram (ECG) were obtained (KardiaMobile 6L, AliveCor). Blood was immediately collected via cardiocentesis for serum chemistry, complete blood count, biobanking, and single cell RNA-sequencing. In three animals (L2, L3, and L4), transcardial perfusion of the lungs with phosphate buffered saline (PBS) was done to reduce circulating cells. Organs and tissues were sequentially removed and divided by a veterinary pathologist. One sample of each tissue was

immediately placed in formalin fixative for histopathology (Casey et al. 2021), and a second was embedded in Optimal Cutting Temperature compound (OCT) and then flash frozen on dry ice and stored at -80°C for biobanking. A third sample was placed directly in cold (4ºC) PBS pH 7.4 and immediately distributed to the tissue expert for cell dissociation and preparation for scRNAseq as detailed below. Additional diagnostics such as microbiological cultures were performed where clinically indicated. The entire necropsy was completed within 1-2 hrs, with ischemia-sensitive tissues prioritized as described in Supplementary Methods.

Histological and pathological analysis

Tissues were immersion-fixed in 10% neutral buffered formalin for 72 hours. Formalin-fixed tissues were processed routinely, embedded in paraffin, sectioned at 5 µm, and stained with hematoxylin and eosin (H&E). Tissues included the following: heart, aorta, lungs, trachea, thyroid gland, parathyroid gland, kidneys, urinary bladder, male reproductive tract (testicle, epididymis, seminal vesicle, prostate, and penile urethra), female reproductive tract (uterus, cervix, vagina, and ovaries), salivary glands, tongue, epiglottis, esophagus, stomach, small and large intestine, liver (with gallbladder), adrenal gland, spleen, lymph nodes, white adipose, brown adipose, bone, spinal cord, eyes, and bone marrow. Selected tissues stained with Von Kossa (for mineralization), Masson's trichrome (for collagen), Congo Red (for amyloid), and Gram stain (for bacteria) as part of the pathological analysis. H&E stained slides were scanned with a Leica Aperio AT2 High Volume Digital Whole Slide Scanner (40x objective), uploaded into Napari image viewer, software adapted by CZB (Sofroniew et al. 2021), and posted on the Tabula Microcebus portal https://tabula-microcebus.ds.czbiohub.org/.

Preparation of single cell suspensions and FACS-sorting for scRNA-seq

Fresh tissue samples procured as described above were placed on ice, received by organ/tissue experts and immediately dissociated and processed into single cell suspensions, except for samples from L3, which were kept cool overnight after necropsy and processed the next morning (see Supplementary Methods). For each solid tissue, this involved a standard combination of enzymatic digestion and mechanical disruption methods that were optimized for the specific tissue, many of which were adapted from procedures used for the corresponding mouse tissue (The Tabula Muris Consortium 2018, 2020). For blood, immune cells were isolated using a high density ficoll gradient (HISTOPAQUE®-1119, Sigma) to include peripheral blood mononuclear cells (PBMCs) and polymorphonuclear leukocytes (PMNs) (Travaglini et al. 2020).

The specific protocols for each of the 27 tissues are detailed in Supplementary Methods. Cell number and concentration for each single cell suspension were determined by manual counts using a hemocytometer, and then adjusted with 2% fetal bovine serum (FBS) in PBS to a target concentration of about 10^6 cells/mL. Samples were then used for droplet-based 10x library preparation and/or flow sorted for single live cells (Sytox blue negative; ThermoFisher S34857) for plate-based SmartSeq2 library preparation. To enrich for cardiomyocytes, the standard procedure for cardiac cell isolation was supplemented by hand-picking cardiomyocytes (Supplementary Methods). Residual cell suspensions were diluted 1:1 with serum-free Bambanker cell freezing media (GC Lymphotec #BB01) and cryopreserved at -80°C.

ScRNAseq library preparation, quality control, and sequencing

For 10x, single cells were profiled using the 10x Genomics single cell RNA-sequencing pipeline (Chromium Single Cell 3' Library and Gel Bead v2 Chemistry kit) and sequenced on a NovaSeq 6000 System as described (The Tabula Muris Consortium 2018, 2020; Travaglini et al.

2020) and detailed in Supplementary Methods. For SS2, single cells were sorted into 384- or 96-well lysis plates, reverse transcribed to complementary DNA (cDNA) and amplified, as described (The Tabula Muris Consortium 2018, 2020). cDNA libraries were prepared using the Nextera XT Library Sample Preparation kit (Illumina, FC-131-1096) or (for L4) an in-house protocol detailed in Supplementary Methods; no significant differences between protocols were observed in library read depth or quality. Pooling of individual libraries and subsequent quality control and DNA sequencing were done as described (The Tabula Muris Consortium 2018, 2020; Travaglini et al. 2020) with minor modifications (Supplementary Methods).

Genome alignment of scRNAseq sequencing reads and gene counts

Microcebus murinus genome assembly (Mmur 3.0) with NCBI annotation release 101 (date acquired, September 21, 2018) was used for downstream alignment and data analysis. A total of 31,509 genes were detected, including annotated genes and unannotated loci but excluding mitochondrial and Y-chromosome genes (unannotated at our acquisition date).

For 10x samples, downstream data was processed by standard methods using Cell Ranger (version 2.2, 10x Genomics). Raw base call (BCL) files directly generated by the NovaSeq were demultiplexed and converted to FASTQ files, and then aligned to the 10x genome index, with barcode and UMI counting performed to generate a gene counts table. Alignment files were outputted in standard BAM format.

For SS2 samples, demultiplexed fastq files were mapped to the genome using STAR aligner (version 2.6.1a). Briefly, the genome FASTA file was augmented with ERCC sequences to create a STAR genome index with 99 bp overhangs (optimized for Illumina 2 x 100 bp paired-end reads). Two-pass mapping was executed, with specific STAR options and parameters detailed in Supplementary Methods.

Contamination filtering of 10x data

We performed stringent contamination filtering to resolve cross-sample contamination in an Illumina sequencing run caused by cell barcode hopping among multiplexed 10x samples (Farouni et al. 2020). Such cross-sample contamination can occur when low levels of ambient mRNA containing the 10x cell barcode in one sample gets added onto the transcript of other samples during Illumina sequencing amplification, resulting in the incorrect assignment of a cell barcode to other samples; hence in subsequent analyses, a cell from one tissue could falsely appear as multiple cells from different tissues (samples). To exclude such artifacts, for each sequencing run we identified all cell barcodes that were assigned to multiple samples, and for each such barcode identified we compared the number of UMIs in each sample. If there was one dominant sample index (i.e., number of UMIs of the dominant sample was 10 times or more greater than that of the second most abundant sample), then the cell with the dominant sample index was kept (but labeled in its metadata as 'potentially contaminated') whereas all other instances of that "cell" were removed. If there was no dominant sample index, then all instances of the "cell" with that barcode were removed from the dataset. (This was not an issue for SS2 samples because they were sequenced using dual unique indices for each cell.)

Cell clustering, annotation, and cluster markers from scRNAseq profiles

Step 1 - Cell clustering and annotation of each tissue processed by 10x. Transcriptomic profiles of cells from each tissue from each individual lemur were clustered separately using Seurat software package (version 2.3.0) for R studio (version 3.6.1). We included in this step all cells with >100 genes or >1000 UMIs detected, a minimal threshold that was used to ensure inclusion of all cell types including ones in which the cells (or RNA) were unstable (see below for more stringent criteria used for final cell quality control). Data normalization, scaling,

dimensionality reduction (PCA), clustering, and visualization (t-SNE, UMAP) were performed as previously described (Travaglini et al. 2020) and detailed in Supplementary Methods. Resultant cell clusters were manually assigned to a compartment (endothelial, epithelial, stromal, lymphoid, myeloid, megakaryocyte-erythroid, neural, germ) based on expression of the mouse lemur orthologs of canonical marker genes for each compartment in human and mouse (see Table S2). Clusters expressing markers from more than one compartment were annotated as "doublets." Cells within each assigned compartment were then subclustered, repeating the data processing steps above, and then the clusters in each compartment were annotated separately. To annotate (determine) the cell type of each cluster, a list of canonical human and mouse gene markers for each cell type in each tissue was curated from the literature (Table S2), including genes previously validated by in situ hybridization and/or immunohistochemistry as well as differentially-expressed genes culled from recent scRNAseq studies, and the orthologous mouse lemur genes were identified and their expression visualized on the t-SNE plots. Based on the enriched expression of marker genes, each cluster of cells within a compartment was manually assigned a cell type identity. Clusters that contained more than one cell type were further subclustered to better resolve the cell types. Cell types represented by only a small number of cells that did not form a separate cluster were manually curated, aided by the cellxgene gene expression visualization tool (Megill et al. 2021) as detailed below.

Each cluster was assigned both a 'cell ontology' cell type designation using the standardized and structured nomenclature (Diehl et al. 2016), and a 'free annotation' that resolves biologically significant clusters not contained in the current cell ontology. Free annotations were assigned as follows. In cases where a smaller cluster stemmed off a larger ("main") cluster in t-SNE embedded space, the smaller cluster was distinguished with one or

more differentially-expressed genes added to the cell type name (e.g. B cell (*SOX5*+) clustered near the main population of B cells in the pancreas); differentially-expressed genes driving the subtype clustering were ascertained by Wilcoxon rank-sum tests. In cases where two approximately equal-sized clusters separated on the t-SNE, a marker gene was added to the cell type name for both clusters (e.g. hepatocyte (*APOB*+) and hepatocyte (*PHYH*+) in the liver). Clusters with a small number of cells that contained more than one cell type but could not be partitioned into separate clusters by subclustering with the Louvain algorithm or manually with cellxgene (see step 3) were labeled as a 'mix' cell type (e.g. the cluster labeled 'endothelial cell' in the uterus contains a mixture of artery, vein, and capillary cells). Clusters with cells that expressed markers for more than one cell type and it was biologically plausible they were not a technical artifact (e.g., doublet of two distinct cell types) were labeled as a 'hybrid' cell type (e.g. the cluster labeled as 'monocyte/macrophage' in the trachea contains cells that expressed markers of both cell types and could not be further distinguished based on current molecular definitions of these cell types). After examining the human/mouse markers for all known cell types in a tissue, clusters that could not be assigned a cell type were labeled 'unknown,' with the tissue, compartment, and one or more differentially expressed genes added to the cell type name (e.g. 'unknown_Bone_stromal_G1 (*NGFR*+ *TNNT2*+)' are bone stromal cells that do not correspond to any extant stromal cell type reported for human and mouse). To detect the differentially-expressed genes of an unknown cell type, we compared the unknown cell type to all other cells of the same compartment and tissue (see Fig. S5). Clusters containing a majority of cells that expressed cell proliferation markers (e.g. *TOP2A, MKI67, STMN1*) were appended the abbreviation "PF." Clusters that separated from a main cluster but did not express any distinguishing markers and differed only in parameters of technical quality (i.e. fewer genes and

counts detected per cell) were considered low quality and "LQ" was appended to the cell type name.

After annotations were assigned, the cutoff for the minimum number of genes per cell was increased from 100 to 500 and only the qualifying cells were further analyzed. For most tissues, this more stringent cutoff only resulted in removal of some erythrocytes and neutrophils; the only exception were cardiac cardiomyocytes, most of which expressed fewer than 500 genes per cell so separate filtering criteria were applied (see Supplementary Methods).

Step 2 - Annotation of each tissue processed by SS2. Cells processed by the SS2 protocol with <500 genes or <5000 reads were excluded from further analysis, and gene expression levels in the remaining cells was scaled and log transformed as described above for the 10x datasets. Cells from a particular tissue and individual were integrated with the 10x dataset of the same tissue and individual into the same UMAP embedded space using the FIRM algorithm (detailed below). Cells from SS2 were automatically annotated with the same label as the nearest neighboring 10x cell. Annotations were manually verified in Step 3 aided by cellxgene gene expression visualization. SS2 datasets for which there were no corresponding 10x dataset from the same individual/tissue were manually annotated using the method described in Step 1 for 10x datasets.

Step 3 – Integration of datasets across individuals. For each tissue, the combined 10x/SS2 datasets from each individual were further integrated into the same UMAP embedded space using the FIRM algorithm. This step resulted in 27 separate tissue UMAPs, each containing data from up to 4 individuals. To ensure consistency of cell type labeling across all individuals, annotations were verified and adjusted manually using cellxgene, an interactive tool

to visualize and annotate single cell RNA-sequencing data (https://chanzuckerberg.github.io/cellxgene/) (Megill et al. 2021).

Step 4 – Integration of datasets across tissues. All 27 tissue-level objects were integrated into a single UMAP embedded space using the FIRM algorithm. As above, annotations were verified/adjusted manually in cellxgene to ensure consistency of cell designations across all tissues. In most instances, each cell type clustered separately, irrespective of the tissue of origin, and the same designation was used across all tissues. Occasionally, similar cells types (e.g. fibroblasts, macrophages) clustered separately by tissue of origin, making it challenging to distinguish whether the separation was due to tissue-level batch effect or because of true biological differences. In these cases, the original tissue-level annotation label was kept for each cluster.

Step 5 – Detect differentially expressed genes for each cell type. We calculated the top 300 differentially-expressed genes for each cell type (represented by at least 5 individual cells) using Wilcoxon rank-sum test with Benjamini-Honchberg FDR values (Table 3). We compared each cell type to: (i) all other cell types from the same tissue (e.g., capillary cell type of the lung compared to all other lung cell types, 'organ-wide' comparison), (ii) all other cell types from the same compartment of that tissue (e.g., capillary cell type of the lung compared to all other lung endothelial cell types, 'organ-compartment-wide' comparison), (iii) all other cell types from the atlas (capillary cell type of the lung to all other cells in the atlas, 'atlas-wide' comparison), (iv) all other cell types from the same compartment across the atlas (capillary cell type of the lung to all other endothelial cell types in the atlas, 'atlas-compartment-wide' comparison).

FIRM integration

FIRM (Flexible Integration of single-cell RNA-sequencing data for large-scale Multi-tissue cell atlas datasets) is a newly developed algorithm that integrates multiple scRNAseq datasets (e.g., from different sequencing platforms, tissue types, and experimental batches) (Ming et al. 2021). In brief, FIRM optimizes dataset integration by computing the dataset-specific scale factors for gene-level normalization. Different datasets generally have varied cell-type compositions, resulting in dataset disparity when scaling the gene expression levels to unit variance for each dataset. Different from classical scaling procedures, FIRM computes the scale factors based on subsets of cells which have matched cell-type compositions between datasets. To construct these subsets, FIRM detects paired clusters between datasets based on similar overall gene expression and then samples the cells so that paired cell types have the same proportional representation in each dataset. Parameters used for integration are given in Supplementary Methods. The integrated datasets generated using FIRM show accurate mixing of shared cell-type identities and preserve the structure of the original datasets, as confirmed by expert manual inspection during cell annotation.

Trajectory analysis

We developed a custom in-house program in Matlab <give GitHub address?> to detect and characterize spatial and developmental pseudotime cell trajectories, which were similar but noisier when analyzed with existing programs such as Monocle 3 (https://github.com/cole-trapnell-lab/monocle3). For the mouse lemur kidney nephron spatial trajectory, all kidney epithelial cells were included in the analysis; podocytes, macula densa cells, intercalated cells, and urothelial cells were not part of nephron gradient and hence excluded from trajectory detection. For vasa recta endothelium spatial trajectory, all four vasa recta cell types were used.

For spermatogenesis pseudotime trajectory, all seven sperm and sperm progenitor cell types were used. For the myeloid cell developmental pseudotime trajectory, hematopoietic precursor cells and all myeloid cell types except dendritic cells (which did not form part of the continuum) were used. Analysis was performed independently for each trajectory using values from the 10x scRNAseq profiles of the indicated cells (low quality cells and technical doublets were excluded) that had been pre-processed (scaled: (ln(UP10K +1), normalized) as described above.

Principal component analysis (PCA) with highly-variable genes (dispersion > 0.5) was done with the PCA function of Matlab, and the high quality principal components (not driven by extreme outlier data points or immediate early genes) were selected from the top 20 principal components and used to generate a 2D UMAP using cell-cell Euclidean distances as input (https://www.mathworks.com/matlabcentral/fileexchange/71902). The trajectory of the cell continuum was detected as the probability density ridge of the data points in the UMAP, using automated image processing (Matlab Image Processing Toolbox™); any interruptions in the detected density ridge line were connected manually along the direction of the ridge line and guided by prior knowledge of the biological process, and the direction of the trajectory was assigned based on expression of marker genes. Individual cells were then aligned to the trajectory by the shortest connecting point to the trajectory; if the trajectory branched (e.g., in myeloid cell development), cells were assigned to the closest branch. Individual cells that were too distant from the trajectory (adaptive thresholding along the trajectory) were deemed outliers and removed from further analysis.

To detect genes whose expression followed the trajectory, we calculated the Spearman correlation coefficient and corresponding p-values (Bonferroni corrected) between the expression level of each gene and 20 preassigned unimodal patterns that smoothly change along

the trajectory (with their single peaks uniformly distributed from the beginning of the trajectory to its end point). Expression patterns of the top ranking (top 1000 with p-value<0.01) and highly variable (dispersion > 0.5) genes were smoothed with a moving average filter and clustered by k-means clustering to detect the major trajectory-dependent expression patterns. The trajectory differentially expressed genes were then ranked by the associated cluster (ranked by trajectory location of peak expression), and within the cluster by p-value from smallest to largest, and with the same p-value by mean expression level from highest to lowest.

For the myeloid cell analysis, four trajectories were independently detected: 1) from hematopoietic precursors to granulocyte monocyte progenitors, 2) from granulocyte monocyte progenitors, proliferating neutrophils, to neutrophils, 3) from granulocyte monocyte progenitors, proliferating macrophages/monocytes, proliferating monocytes, to monocytes and macrophages, and 4) from megakaryocyte progenitors, erythroid progenitor cells, proliferating erythroid lineage cells, to erythroid lineage cells. On the UMAP, trajectory 1 branched into trajectories 2 and 3, so two longer trajectories were generated (1+2, 1+3). Differential gene expression analysis was then independently performed for each of the constituent trajectories (1+2, neutrophil lineage; 1+3, monocyte/macrophage lineage; 4, erythrocyte lineage).

Comparison of expression profiles among mouse lemur cell types

UMAP of cell types

To visualize similarities among the mouse lemur cell type expression profiles, we applied uniform manifold approximation and projection (UMAP) and embedded the high dimensional scRNAseq expression data (~30,000 genes) to 2D. For this analysis, the 10x scRNAseq dataset was used and molecular cell types that were low quality (labeled with LQ in free annotation) or represented by less than 4 individual cells were excluded, resulting in a comparison of 681

molecular cell types. Molecular cell types were treated as pseudo bulk, with gene expression levels calculated by averaging the expression level of each gene for all cells of that type and then natural log transformed (ln(Avg Count Per 10K UMIs +1)). Expression levels were further normalized by the maximal value of each gene across all cell types, so that all ranged from 0 to 1. The cell type gene expression matrix was then projected onto a 2D space with cosine distances between pairs of cell types used in the UMAP function (https://www.mathworks.com/matlabcentral/fileexchange/71902). The Wilcoxon rank sum test was used to identify differentially-expressed genes that distinguished related molecular cell types identified in the cell type UMAP from other cell types (e.g., sperm/sperm progenitor cells and immune progenitor/proliferating cells vs. proliferating cells of other compartments) as described in Supplementary Methods.

Heat map of cell type pairwise correlation scores

To compare the overall gene expression profiles of molecular cell types, Pearson's correlation scores were calculated for every pair of molecular cell types. To compare data from different sequencing platforms (10x and SS2), we used the FIRM integrated dataset as described above which contains FIRM-generated principal component (PC) coefficients for each cell. Cell types were treated as pseudo-bulk and the cell type average PC coefficients were calculated and used to determine the correlation coefficients. The cell type pairwise correlation scores were plotted as heat map matrices with cell types arranged in three different orders in the matrices to facilitate comparison (Fig. 5 g-h and Fig. S7). Interactive forms of the heat map matrices are available online (Tabula Microcebus portal, https://tabula-microcebus.ds.czbiohub.org/).

SAMap analysis

SAMap (Self-Assembling Manifold mapping), a cross-species mapping method (Tarashansky et al. 2019, 2021), was used to calculate cell type similarity scores across species (Fig. S8 a, c). SAMap measures the similarity between cell types by calculating an alignment score (edge width in Fig. S8 a, c), which is defined as the average number of cross-species neighbors of each cell relative to the maximum possible number of neighbors in the combined manifold. For application of SAMap to lung and muscle cells, the 10x scRNAseq datasets for lemurs L1-L4 were used together with the corresponding mouse and human data from 10x scRNAseq datasets of Tabula Muris Senis (The Tabula Muris Consortium 2020) and Tabula Sapiens (The Tabula Sapiens Consortium 2021), respectively. The defaut SAMap parameters were used in the analysis, and similarity scores less than 0.1 were removed.

Comparison of mouse lemur, human, and mouse cell type expression profiles

For the comparisons, we used published mouse and human scRNAseq datasets obtained using approaches and standards similar to those described above for mouse lemur, and re-annotated where necessary for consistency with the mouse lemur annotations. Human and mouse lung cell 10x datasets (Travaglini et al. 2020) (annotations as published) were compared to lemur lung 10x datasets provided here. Human and mouse limb muscle cell 10x datasets (The Tabula Muris Consortium 2020; The Tabula Sapiens Consortium 2021) (re-annotated here) were compared to lemur limb muscle 10x datasets provided here. Mouse hepatocyte SS2 datasets (The Tabula Muris Consortium 2020) (re-annotated here) were compared to mouse lemur hepatocyte SS2 profiles provided here.

For each tissue, selection of equivalent cell types for comparison across the three species was done by grouping subtypes of the same cell type (with the exception of proliferating cell

types, which were kept separate) and reassigning them the same cell type designation, with the labeling convention kept consistent across all three species (column unified_annotation in h5ad files available on Figshare). Only cell types with at least 20 profiled cells in each of the three species were selected for use in the comparisons. To validate that the selected cell types were truly comparable across species, we used the SAMap algorithm (see above) to calculate cell type similarity scores across species, which confirmed that cells with the same cell type designation had the highest similarity score (Tarashansky et al. 2021). This was further corroborated by aligning tissue datasets across the three species (after removing cell types that were not selected for comparison) using the BBKNN (batch balanced k nearest neighbours) batch correction algorithm (Polański et al. 2020) on the 13,302 one-to-one-to-one genes orthologs described below. The resulting UMAP showed that cells with the same cell type designation across the three species almost always clustered nearest to each other (the exceptions were intermixing of lung monocytes, alveolar macrophages, and dendritic cells; lung and muscle CD4+ and CD8+ T cells; and muscle tendon and fibroadipoegnic progenitor cells).

For comparison of the transcriptomic profiles of homologous cell types, a list of human, mouse lemur, and mouse one-to-one-to-one gene orthologs was compiled from gene homology assignments in the National Center for Biotechnology Information (NCBI) and Ensembl databases. For all mouse lemur genes annotated in NCBI, the corresponding human and mouse orthologs were retrieved (gene_info.gz and gene_orthologs.gz from https://ftp.ncbi.nlm.nih.gov/gene/DATA/, February 2020). In addition, Ensembl gene ID numbers, gene symbols, and human/mouse ortholog assignments were retrieved from Ensembl Biomart (Ensembl Genes version 99, February 2020) and appended to the database by finding the corresponding Ensembl gene ID (variable 'Gene_stable_ID') attributed to each NCBI gene

ID (variable 'NCBI_gene_ID ') provided in Ensembl Biomart. Mouse lemur genes that did not have an assigned human and mouse ortholog in either Ensembl or NCBI were removed (15,297 out of 31,966 unique mouse lemur NCBI gene IDs), as were mouse lemur genes that had more than one human or mouse ortholog assigned, or that shared the same human or mouse ortholog with another mouse lemur gene. Note that unlike NCBI, Ensembl specifies the type of ortholog assignment (e.g., 'ortholog_one2one', 'ortholog_one2many'); however, we did not use the Ensembl specification to filter one-to-one-to-one orthologs because occasionally a mouse lemur gene symbol was assigned by homology to multiple currently unnamed loci in Ensembl and because of this imperfect genome annotation was labeled as sharing an 'ortholog_one2many' with human/mouse instead of 'ortholog_one2one'. A total of 15,515 one-to-one-to-one gene orthologs were thus uncovered, and of those, 13,302 were found in the human and mouse scRNAseq datasets described above (Table 5).

Similarity of the transcriptomic profiles between species for each homologous cell type was determined by randomly selecting (by bootstrapping) 1000 trios of cells of the same cell type designation from human, lemur, and mouse, and then calculating the pairwise similarities (human & lemur, lemur & mouse, and human & mouse) of the transcriptomic profiles (of the one-to-one-to-one gene orthologs) in each trio by Xi correlation (Chatterjee 2019) (https://github.com/czbiohub/xicor). Xi correlation is a rank-based algorithm that breaks ties randomly to handle the data sparsity of scRNAseq datasets. For each of the homologous cell types, differences in Xi correlation values (ΔXi) between each pair of species was used to determine the relative similarity between species (e.g., $Xi_{(lemur, human)} - Xi_{(lemur, mouse)} > 0$) indicates that the lemur cell type is more similar to the corresponding human cell type than to the

corresponding mouse cell type). p-values were calculated using a t-test under the null hypothesis $\Delta Xi=0$.

To detect gene orthologs with species-specific expression patterns, we performed differential gene expression analysis using the one-tailed Wilcoxon rank sum test of lemur vs human, lemur vs mouse, and human vs mouse, independently for each of the homologous cell types. Using a fold change threshold of 10 and a p-value threshold of 10^-5, we identified genes with species-selective expression patterns in at least one cell type, excluding genes that were lowly expressed (cell type ln(avgUP10K+1) < 0.5) in the corresponding cell type for all three species. We categorized 6 species-selective expression patterns: 1) high in human and lemur compared to mouse (HumanLemur), 2) high in human and mouse compared to lemur (HumanMouse), 3) high in lemur and mouse compared to human (LemurMouse), 4) high in human compared to lemur and mouse (Human), 5) high in lemur compared to human and mouse (Lemur), and 6) high in mouse compared human and lemur (Mouse). Note that these patterns are not mutually exclusive and a gene may appear within multiple categories for different cell types or identical cell types. Next, within a category, we grouped genes according to whether its species-selective expression pattern was specific to only one cell type or in multiple cell types. For genes with a species-selective pattern in multiple cell types, we further categorized them according to whether the cell types were from the same compartment (e.g., epithelial cells) or different compartments (cross-compartment). The full list of species-specific genes is provided in Table 5, and examples of the cell type-specific, compartment-specific, and cross-compartment expression patterns shown in Fig S9.

Evolutionary comparison of gene expression during spermatogenesis

To identify spermatid-specific genes that 'escape' post meiotic sex chromatin (PMSC) silencing, the expression level (10x dataset, L4) of the 1,453 X-chromosome genes (including non-protein coding genes) was compared across three stages of spermatogenesis: spermatogonia (SG), pachytene spermatocytes (PS), and early spermatids (ES). Expression patterns were categorized into 5 groups (A-E) as described for human and mouse (Namekawa et al. 2006; Sin et al. 2012). Group E is comprised of spermatogenesis-suppressed genes defined by low or undetectable expression (average expression <0.05) across all three stages. The remaining (spermatogenesis-expressed) genes were classified into four groups (A-D). Group A and B genes are repressed in PS (genes likely subject to MSCI), as defined by $E_{PS}<0.5E_{SG}$, with E representing the natural log transform of the average expression of a gene for the cell type denoted by the subscript. To determine if a pachytene-repressed gene was reactivated in ES, we calculated its recovery rate defined by $(E_{ES}-E_{PS})/(E_{SG}-E_{PS})$. PS-repressed genes that remained repressed in ES (recovery rate <0.5) were classified group A, and PS-repressed genes that were reactivated (recovery rate >0.5) were classified group B. Group C and D genes were not repressed in PS ($E_{PS}>0.5E_{SG}$). Genes whose expression increased in spermatid stages compared to pachytene stage ($E_{ES}>E_{PS}$) were classified group C (likely spermatid-specific genes that escape PMSC silencing) and the remaining spermatogenesis-expressed genes that did not follow one of these patterns were classified group D. Mouse lemur group C 'escape' genes were manually compared to the list of human and mouse group C genes obtained by similar classification methods using human and mouse microarray data (Table 6) (Namekawa et al. 2006; Sin et al. 2012). Note genes and unannotated loci from one species that lacked identified orthologs in either of the other two species were included as "species-specific expression

patterns" in our classification. Similar results obtained when this analysis was applied to the mouse lemur SS2 dataset.

Data and code availability

scRNAseq gene expression Counts/UMI tables, and cellular metadata are available on figshare (https://figshare.com/projects/Tabula_Microcebus/112227). Data can be explored interactively using cellxgene on the Tabula Microcebus portal: https://tabula-microcebus.ds.czbiohub.org/. Raw sequencing data and genome alignments are available on request.

FIGURE LEGENDS

Figure 1 Experimental scheme for constructing the mouse lemur cell atlas

a. Overview of four mouse lemurs profiled (name, age, sex), associated metadata collected, and uses of the procured tissues. L1, mouse lemur 1; yo, years old; M, male; F, female; CBC, complete blood count.

b. Representative tissue histology. Micrographs of hematoxylin-and-eosin-stained (H&E) lung section (left) and close-up (second from left) from L1 and small intestine (second from right) and close-up (right) from L3. Full histological atlas of all tissues analyzed from all individuals is available online (the Tabula Microcebus portal, https://tabula-microcebus.ds.czbiohub.org/), and identified histopathology is described in (Casey et al. 2021). Scale bars (left to right) 1 mm, 100 um, 100 um, 25 um.

c. The 27 mouse lemur tissues harvested, organized by system, showing for each tissue the number of high-quality cells obtained by 10x scRNAseq protocol, by SS2 scRNAseq protocol, number of biological replicates (individuals), and number of molecular cell types identified.

d. Flow diagram for obtaining and processing single cell RNA-sequencing data.

e. Flow diagram for clustering cells with related transcriptomic profiles and annotating the molecular cell types.

f. Representative tissue UMAP showing scRNAseq profiles of kidney cells (dots) integrated in the same embedded space via FIRM across 10x and SS2 datasets (top left, Step 2) and three individuals (top right, Step 3). Compartment identities of the cell clusters are shown (bottom left) along with heat maps of expression levels (ln(UP10K+1) for 10x data and ln(CP10K+1) for SS2 data) of the indicated compartment marker genes (bottom right; *EPCAM*, epithelial; *PTPRC*, immune lymphoid/myeloid; *PECAM*, endothelial; *COL1A1*, stromal).

g. UMAP of all 244,081 cells in the atlas integrated by FIRM algorithm across all 27 tissues analyzed from four individuals.

Figure 2 Taxonomy of identified mouse lemur molecular cell types

a. Dendrogram of the 256 assigned designations of molecular cell types across cell atlas. Designations are arranged by compartment (epithelial, endothelial, stromal, immune, neural, germ) and then ordered by organ system (epithelial compartment) or biological relatedness (other compartments). Designation number is given below the abbreviation and full names listed in box at right. Some closely related molecular cell types and states are grouped together but shown separately in panel b and Fig. S1, and described further in Section 4. PF, group includes a proliferative cell state; H1-12, hybrid cell types with symbol placed between the two types whose expression signatures the hybrid type shares; M1-7, mixed clusters of distinct cell types too few to assign separately; *, pathological cell state (Path; metastatic tumor found in lung of L2 and described in accompanying paper, in prep).

b. Dot plot showing number of profiled cells (dot intensity shown by heat map scale, small red dot indicates <10 cells) for each of the 768 identified molecular cell types (including 38 hybrid types) plus 24 mixed clusters isolated from the tissues indicated at left. Molecular cell types in each tissue (rows) are arranged (columns) by cell type number/designation and separated by compartment as in panel a. Black bars, closely related molecular types/subtypes/states (see Fig. 4). +, unknown molecular cell type (Section 4).

Figure 3 Cell types with gene expression gradients

a. Dendrogram of the 71 molecular cell types identified by scRNAseq of kidney, arranged as in Fig. 2a.

b. UMAP of kidney epithelial cells (L4, 10x scRNAseq dataset) showing a detected molecular trajectory. Dots, individual cells colored by molecular cell type as indicated; thick black line, cell density ridge (trajectory line); thin grey lines, shortest connecting point of cell to trajectory line. Trajectory shows spatial continuum of molecular cell identities along nephron, beginning with proximal convoluted tubule (PCT, #55; see panel a) and ending with principal cells of collecting duct (CDp, #64-66). Macula densa cells (MD, #62) cluster between thin and thick ascending loop of Henle (LoH) cell types (#57-60), and urothelial cells (#69) cluster near CDp cells. Intercalated cells of collecting duct (CDi, #67, 68) and podocytes (#54) cluster separately from trajectory. PCT, proximal convoluted tubule; PST, proximal straight tubule; LoH thin D, loop of Henle thin descending limb; LoH thin A, loop of Henle thin ascending limb; LoH thick A, loop of Henle thick ascending limb; DCT, distal convoluted tubule; CDp, principal cell; CDi alpha, alpha intercalated cell; CDi beta, beta intercalated cell; MD, macula densa.

c. Heat map of expression of indicated canonical nephron cell marker genes in cells along nephron trajectory in panel b. Colored bar at top shows cell type designations (colors as in panel **b**). Gene expression values (ln(UP10K+1) are normalized to maximal value for each gene across all cells in trajectory. See panel e for heat map scale (for panels c, e, and j).

d. UMAP of kidney vasa recta endothelial cells (L4, 10x scRNAseq dataset) showing another detected molecular trajectory as in panel b. Trajectory shows spatial continuum of molecular cell identities along vasa recta. VR D, vasa recta descending limb; VR A, vasa recta ascending limb.

e. Heat map of expression of indicated canonical marker genes in cells along vasa recta trajectory in panel d. Bar (at top), cell type designations (colors as in panel d). Note transitions along

trajectory (left to right) from artery/arteriole (*GJA5*+) to capillary (*CA4*+) to vein (*ACKR1*+) markers.

f. Micrographs of H&E-stained sections of kidney medulla and cortex analyzed in panels b-e. Scale bars: 100 um (left), 50 um (right).

g. UMAP of germ cells from testis (L4, 10x scRNAseq dataset) showing detected molecular trajectory as in panel b. Trajectory shows developmental continuum (developmental pseudotime) of molecular cell identities during spermatogenesis, beginning with stem cells (spermatogonium, #250) and progressing to late spermatids (#256). See Fig. 6c for dot plot of expression of canonical marker genes along trajectory.

h. Micrographs of H&E-stained sections of testis seminiferous tubules analyzed in panel g. Scale bars: 200 um (left), 50 um (right).

i. UMAP of myeloid cells from bone and bone marrow (L2, 10x scRNAseq dataset) showing two detected molecular trajectories. One (at left) shows developmental continuum (developmental pseudotime) of molecular cell identities beginning with hematopoietic precursor cells (HPC, #169) and bifurcates at granulocyte-monocyte progenitor cells (GMP, #181) into neutrophil lineage (#175-176) and monocyte/macrophage lineage (#182-184,187). Another trajectory (at right) connects the erythroid progenitor and lineage cells (#172-174), with megakaryocyte progenitor cells (#170) nearby.

j. Heat map of expression of indicated canonical marker genes in cells (uniformly subsampled for neutrophils) along developmental pseudotime trajectories in panel i (top left, neutrophil trajectory; top right, monocyte/macrophage trajectory; bottom: erythroid trajectory). [], description of gene identified by NCBI as a gene locus (LOC105862649 [CD14], LOC105883507 [HBB], LOC105856254 [HBA2]).

Heat maps of genes differentially expressed along each of the six trajectories are provided in Fig. S3.

Figure 4 Previously unknown and understudied molecular cell types

a. Expansion of molecular cell type dendrogram in Fig. 2a showing individually each of the molecular cell types that are grouped in Fig. 2a. These include molecular types related to a known cell type but distinguished from the major population ("main") or each other indicated by a distinguishing marker gene (e.g., B cell (*SOX5*+), hepatocytes (*APOB*+), hepatocytes (*PHYH*+)) or by a proliferative gene signature (PF), and unidentified ("unknown") molecular cell types. *, molecular types found in only one individual. (), tissue abbreviation indicated for molecular types found in <3 tissues. Bla, Bladder; Blo, Blood; Bon, Bone; BM, Bone marrow; Col, Colon; Cor, Brain cortex; Fat, Fat; Pit, Pituitary; Kid, Kidney; Liv, Liver; Lun, Lung; MG, Mammary gland; Pan, Pancreas; Ski, Skin; SI, Small intestine; Spl, Spleen; Ste, Brainstem; Ton, Tongue; Tra, Trachea; Ute, Uterus.

b. UMAP of liver hepatocytes and cholangiocytes from three individuals (L2, L3, L4, 10x scRNAseq datasets) integrated by FIRM.

c. Dot plot of expression of indicated cell type markers and other differentially-expressed genes in liver *APOB*+ and *PHYH*+ hepatocytes and cholangiocytes of panel b. +, hormone receptor, ligand, and synthase genes. [], description of gene identified by NCBI as a gene locus (LOC105859005 [HP], LOC105861535 [A2M]).

d. Dot plot of expression of indicated endothelial and capillary markers and differentially-expressed genes for all *FABP5*+ *RBP7*+ and *FABP5*- *RBP7*- capillary molecular types in atlas (L1-L4, 10x scRNAseq dataset). Note the two sets of differentially-expressed genes (CNS,

Peripheral) that distinguish *FABP5- RBP7-* capillary cells from CNS (Cortex (Cor), Brainstem (Ste), Pituitary (Pit)) and peripheral tissues (Blood (Blo), Lung (Lun), Kidney (Kid)), c. [], description of gene designated in NCBI as a gene locus (LOC105857591 [FABP4], LOC105879342 [CD36]).

e. UMAP of kidney stromal cells (L2, L3, L4; 10x scRNAseq dataset) integrated by FIRM. Dashed circle, unknown molecular cell type (#143, *ST6GAL2+*; potentially mesangial cells).

f. Dot plot of expression of indicated cell type markers and differentially-expressed genes for kidney stromal cells (L2, L3, L4, 10x scRNAseq dataset). Note unknown molecular cell type (#143) and the set of differentially-expressed genes(Unkn) that distinguish it from the other molecular types.

Figure 5 Relationships of molecular cell types across the mouse lemur cell atlas

a. UMAP of all molecular cell types (dots) in the atlas based on their transcriptomic profiles (mean expression level of each gene for all cells of that type in atlas (L1-4, 10x scRNAseq data)). Dot fill color indicates tissue compartment; black outline, progenitor cell types; grey outline, proliferating cell types; +, unknown cell types. Dashed grey contours, groups of the related cell types indicated.

b-d. UMAP of panel a showing mean expression level of indicated gene in each cell type as heat map, with expression normalized across all cell types. Shown here are examples of genes with expression pattern specific to immune progenitor/proliferating cells and germ cells compared to cell types in other compartments (*CCNB1*, panel b), or to proliferating cells in other compartment (*TESMIN*, panel c), and genes with expression pattern specific to Schwann and

stromal cells, compared to non-Schwann cell types in the neural compartment (*COL3A1*, panel d).

e-f. Close up of portions of UMAP in panel a showing segregation of the two types of lymphatic cells indicated (e), and segregation of skin and tongue epithelial cell types including those with similar locations/functions in the two tissues (e.g., basal cells) (f).

g. Heat maps of pairwise Pearson correlation coefficients between the transcriptomic profiles of each of the 749 molecular cell type in atlas (10x and SS2 datasets, excluding all cardiac cells), calculated from principal component values of FIRM-integrated UMAP (Fig. 1g) averaged across all cells of each type. Cell types ordered by compartment, then cell type designation/number, and then tissue.

h. Close up of heat map from panel d showing pairwise correlations between skin epithelial cells and all other cell types (top), and between testis germ cells and all other cell types (bottom). Note proliferating skin interfollicular suprabasal cells show high correlation with proliferating and progenitor cell types across all compartments in atlas. Testis spermatogonia also show high correlation with proliferating cell types in the atlas, but especially with hematopoietic stem cells.

Figure 6 Comparison of the transcriptomic profiles of lemur, human, and mouse molecular cell types

a. Comparison of the transcriptomic profiles (10x scRNAseq datasets generated and annotated in same way for all three species) of orthologous lemur (Tabula Microcebus, this paper), human (Travaglini et al. 2020), and mouse (Tabula Muris Senis, (The Tabula Muris Consortium 2020)) lung cell types. Pairwise (human-lemur, lemur-mouse, human-mouse) Xi correlations of the profiles of 1000 randomly-selected trios of the indicated cell types were computed using the

13,302 one-to-one-to-one gene orthologs in the three species, and ΔXi values with 95% confidence intervals are shown (left plot, (lemur, human) - (lemur, mouse) ΔXi values, with positive values indicating lemur cell type is more similar to human cell type than to mouse cell type; right plot, (human, lemur) - (human, mouse) ΔXi values, with positive values indicating human cell type is more similar to lemur cell type than to mouse cell type). P-values, t-test with null hypothesis ΔXi=0. Similar comparisons of orthologous muscle cell types are provided in Fig. S9. AT2, alveolar type 2 cells; Adv: adventitial; Alv: alveolar; NK, natural killer cells; NKT, natural killer T cells; NK/T(PF): proliferating NK and T cell hybrid; cDC, conventional dendritic cells; pDC: plasmacytoid dendritic cells.

b. Dot plots of expression levels (datasets as in panel a) in indicated lung cell types from human (H), lemur (L), and mouse (M) of selected one-to-one-to-one gene orthologs (gene symbol is that of human ortholog) that show primate-selective expression (>10-fold higher expression in human and lemur than in mouse, p <10^-5) in one or more lung cell types. Cell types are colored by compartment as in Fig. 2, and genes are labeled with the same compartment color if it is selectively expressed in cell type(s) of that compartment, or black for the cross-compartment genes. See Fig. S9 for similar analyses of human-, lemur-, and mouse-selective genes.

c. Dot plot of expression in testis sperm cells (10x scRNAseq dataset, L4) of the indicated lemur gene orthologs of regulators and marker genes of spermatogenesis previously identified in human (Sin et al. 2012) and mouse (Namekawa et al. 2006). Filled arrowhead, temporal pattern of expression similar for human and lemur; open arrowhead, pattern similar for lemur and mouse; bold font, pattern specific to lemur; all other genes shown, pattern similar for all three species (comparison data for human and mouse spermatogenesis were curated from published

studies (Hermann et al. 2018; Shami et al. 2020; M. Wang et al. 2018; Jung et al. 2019; Ernst et al. 2019; Thul and Lindskog 2018; Y. H. Sun, Jiang, and Li 2018; F. Sun et al. 2015; Oatley et al. 2014; Ravnik and Wolgemuth 1996; Miranda-Vizuete et al. 2003; Desvignes et al. 2009; Watanabe et al. 2015; Jan et al. 2017)). [], description of gene identified by NCBI as a gene locus (e.g., LOC105858542 [DMC1], LOC105858168 [HSPA1L]).

d. UMAP of testis germ cells (10x scRNAseq dataset, L4) colored by cell type designation (top) and as heatmap showing expression level of *PIWIL1* (bottom). Note *PIWIL1* expression is restricted to early and pachytene spermatocytes.

e. Box (mean and interquartile range) and whisker (outlier cells) plot showing relative expression of all genes on the X-chromosome normalized to all genes on a control autosome (chromosome 2) (mean(ChrX genes)/mean(Chr2 genes)), calculated for each cell during spermatogenesis (10x and SS2 datasets included in plot, L4), with stages as indicated in panel d. Note sharp diminution in expression of X-chromosome genes in early spermatocytes (#251, corresponding to “meiotic sex chromosome inactivation”, MSCI) followed by gradual restoration to about one-third the original value in maturing spermatids during postmeiotic sex chromatin silencing (#254-256).

f. Gene expression levels (ln(UP10K+1) during mouse lemur spermatogenesis for each identified X-chromosome gene that “escapes” postmeiotic silencing during spermatogenesis. Mouse lemur escape gene profiles are separated into four panels according to whether the orthologous gene in the species indicated (human, mouse) is also an escape gene (Sin et al. 2012; Namekawa et al. 2006). Note most escape genes in lemur are unique to lemur (49 genes) or primate-specific (8 genes). [], description of gene identified by NCBI as a gene locus (e.g., LOC105868203 [TCP11X2], LOC105855602 [H2AB1], LOC105862836 [uncharacterized 1], LOC105855134 [TNP3]).

TABLES

Table 1 Marker genes in human and mouse for each cell type (and lemur ortholog) as well as compartment markers

Table 2 Number of cells and list of cell types sequenced per individual and tissue (separated by 10x and SS2)

Table 3 Differentially expressed genes for each mouse lemur molecular cell type

Table 4 One-to-one-to-one orthologs across lemur, human, and mouse

Table 5 Ortholog genes with species expression patterns in lung and muscle cell types

Table 6 Expression patterns of X-chromosome genes during spermatogenesis in lemur, human, and mouse

SUPPLEMENTAL MATERIAL

Legends for Supplementary Figures

Figure S1 Full dendrogram of cell type designations in mouse lemur atlas

Expansion of Fig. 2a showing full dendrogram of the 256 assigned designations of molecular cell type across atlas. Designations are arranged by compartment (epithelial, endothelial, stromal, immune-lymphoid/hematopoietic/myeloid, neural, germ) and then ordered by tissue system (epithelial compartment) or biological relatedness (other compartments). PF, proliferative cell state; H1-12, hybrid cell types (cell clusters with gene signature of two cell types); M1-7, mixed clusters of distinct cell types too few to assign separately; (), tissue abbreviation indicated for cell types present in <4 tissues with full tissue name in black box; *, pathological cell state (metastatic tumor found in lung of L2 and described in accompanying paper, in prep).

Figure S2 Dendrogram of cell types in brain cortex, brainstem, retina, hypothalamus/pituitary, and limb muscle.

a, b. Dendrograms of the 23 molecular cell types identified from scRNAseq of brain cortex (3120 10x cells, 292 SS2 cells; L1, L2, L4) (a), and the 25 molecular cell types from brainstem (1988 10x cells, 122 SS2 cells; L2, L4) (b). The cortex and brainstem samples captured 1099 neurons including excitatory (*SLC17A6+, SLC17A7+*) and inhibitory (*GAD1+, GAD2+, SLC32A1+*) neurons that with further resolved by subclustering into at least 10 major molecular types including four cortex-specific excitatory, one brainstem-specific excitatory, one cortex-specific inhibitory, three brainstem-specific inhibitory, and one hybrid population expressing both excitatory and inhibitory markers. Different glial cell types were also identified including three molecular types of astrocytes, oligodendrocytes and their precursors (OPCs), choroid

plexus cells, ependymal cells, as well as surprisingly Schwann cells (myelinating and non-myelinating) which are generally associated with peripheral nervous system. In the endothelial compartment we identified vein cells as well as capillary cells (*FABP5- RBP7-*) that showed distinct gene signatures compared to capillary cells of peripheral tissues (Fig. 4d). In the stromal compartment, vascular smooth muscle cells, pericytes, and leptomeningeal cells were identified. For the immune compartment, major classes of lymphoid and myeloid cells (B, NK/T, macrophages/microglia, neutrophils, and erythroid cells) were identified.

c. Dendrogram of the 11 molecular cell types identified from retina scRNAseq (938 10x cell, 120 SS2 cells, L4). Identified cells include two types of photoreceptors (dim-light monochromatic rod cells and bright-light multichromatic cones), ON- and OFF- bipolar cells (Shekhar et al. 2016) that rods or cones synapse on plus a hybrid bipolar cell type, a type of horizontal neurons, two types of glial cells (Muller and non-myelinating Schwann cells), and three immune populations (macrophages/microglia, neutrophils, erythroid cells). We found that the photoreceptor cells of the mouse lemur retina is rod-dominated (99.4% rods vs. 0.6% cones in the 10x scRNAseq unsorted sample), which fits the nocturnal behavior of the mouse lemur and is consistent with reported photoreceptor cell densities (Peichl et al. 2019). Lemurs (and new world monkeys) have only two opsins (unlike old world monkeys and apes which have three (Peichl et al. 2019)), and two of the six sequenced cone cells expressed the long-wave-sensitive cone opsin (*OPN1LW*) but none expressed the short/medium-wave-opsin (*OPN1SW*). Two cell types not obtained are the ganglion cells on which bipolar cells synapse and amacrine cells which serve as inhibitory interneurons that modulate and shape signaling.

d. Dendrogram of the 34 molecular cell types identified from hypothalamus/pituitary scRNAseq (2431 10x cells, L4). Identified cells include all five canonical neuroendocrine cell types of the

anterior pituitary (based on expression of their corresponding hormones and other classical markers, see Table 2); interestingly, a few cells (labeled "hybrid") expressed markers for more than one neuroendocrine cell type. We also identified a GABAergic neuron and multiple glial molecular types including oligodendrocytes, OPCs, ependymal cells, and astrocytes. In the endothelial compartment, two capillary types were identified, one (capillary *FABP5- RBP7-*) shared with brain cortex and brainstem and another (sinusoid *MAFB+*) likely the fenestrated sinusoidal capillaries of the pituitary gland. Stromal compartment cell types identified include vascular smooth muscle cells, pericytes, leptomeningeal cells, and fibroblasts, and immune cell types spanned diverse lymphoid and myeloid cells including several types of immune progenitors and proliferating cells.

e. Dendrogram of the 31 molecular cell types identified from limb muscle scRNAseq (2968 10x cells, 994 SS2 cells; L2, L4). Identified cells include 10 distinct stromal molecular types as described in the main text, as well as diverse endothelial, lymphoid, and myeloid cells. One significant cell type that was not identified is Schwann cells, which cover the neuromuscular junction and the terminal axon. This might be because, as in humans, the neuromuscular junctions in mouse lemurs may be small (Boehm et al. 2020) and have few associated Schwann cells.

f. UMAPs of limb muscle cells (10x scRNAseq dataset, L2) separated by compartment. Left to right: endothelial, stromal, lymphoid, and myeloid compartments.

Figure S3 Expression profiles of selected differentially-expressed genes along cell trajectories in Fig. 3.

a-f. Heat map showing normalized expression (10x scRNAseq datasets from lemur, tissue, and compartment indicated at top of panel) of the indicated differentially-expressed genes along trajectories of Fig. 3. Genes shown are top three genes from each of the detected trajectory-dependent expression patterns described in Methods. Cells are ordered left to right by their trajectory coordinates, and their cell type designations are indicated by color in top bar (color key as in UMAPs of Fig. 3). Trajectories in panels d and e include bone and bone marrow hematopoietic precursor cells (HPC) and granulocyte/monocyte progenitors (GMP), and heat map in panel f includes bone and bone marrow erythroid progenitors and lineage cells. [], description of gene identified by NCBI as a gene locus (LOC105884179 [BEX1], LOC105877478 [CCSAP], LOC105862715 [DCUN1D1], LOC105881500 [DEFA1], LOC105865511 [FCGR1A], LOC105867419 [GSTP1], LOC105856255 [HBA1], LOC105883507 [HBB], LOC105859005 [HP], LOC105872012 [HLA-DRB1-1], LOC105859819 [HRNR], LOC105874071 [IFITM3], LOC105882927 [LRRC9], LOC105884612 [NAT8B], LOC105880511 [PTMA], LOC105883741 [SERPINB3], LOC105876721 [SERPINB3], LOC105859377 [TMEM45A], LOC105865212 [TMEM14C], LOC105871594 [RGCC], LOC105865610 [RDH7], LOC105865617 [RDH16], LOC105864771 [RNASE2], LOC105876678 [uncharacterized 1], LOC105873147 [uncharacterized 2], LOC105862290 [uncharacterized 3], LOC105858108 [uncharacterized 4], LOC105880776 [uncharacterized 5], LOC105881161 [uncharacterized 6], LOC105871650 [uncharacterized 7]). **g-l**. UMAPs and detected molecular trajectories of cells from the indicated tissues and compartments as in Fig. 3, but from other lemur individuals as indicated at bottom left of each UMAP.

Figure S4 Hepatocyte molecular subtypes in mouse and lemur

a. UMAP of kidney hepatocytes and cholangiocytes (10x scRNAseq datasets; L2, L3, and L4) integrated by FIRM. Top to bottom: cells colored by molecular cell type annotation, by individual mouse lemur from which each was isolated, by heatmap showing normalized expression of indicated marker genes: *SPP1*, cholangiocyte marker; *APOB*, hepatocyte (*APOB*+) marker; *PHYH*, hepatocyte (*PHYH*+) enriched gene.

b. UMAP of mouse kidney hepatocytes (Tabula Muris Senis SS2 dataset; 9 individuals including 3 females and 6 males, aged 3-24 months) (The Tabula Muris Consortium 2020). Note two main clusters that we designate hepatocyte (*Apob*+) and hepatocyte (*Phyh*+), corresponding to differentially-expressed genes and molecular type designations in lemur. Top to bottom: cells colored by molecular cell type annotation, by sex and by age of animal cell was isolated from, and by heatmap showing normalized expression of indicated marker genes (*Apob, Phyh*).

c. Dot plot of mean level of expression in the indicated lemur (10x dataset as above) and mouse (SS2 dataset as above) hepatocyte molecular types of the indicated hepatocyte markers, differentially-expressed genes between the two molecular types, liver zonation markers, and cell stress markers including immediate-early genes and heat shock proteins. Note that in both species, hepatocyte molecular types do not correspond to expression of liver zonation markers and do not differ significantly in expression of cell stress markers. +, known hormone receptor, ligand, or synthase gene; [], description of gene identified by NCBI as a gene locus (LOC105859005 [HP], LOC105861535 [A2M], LOC105854822 [CYP1A2], LOC105861495 [CYP2E1]).

d, e. Box and whisker plots of relative number of UMIs (lemur, 10x dataset as above) or transcripts (mouse, SS2 dataset as above) (d) and number of genes (e) detected per cell for each

hepatocyte molecular type indicated. Note that in both species, more transcripts/UMIs and genes were detected in *APOB*+ hepatocytes compared to *PHYH*+ hepatocytes. The greater number of genes detected in mouse hepatoctyes relative to that of lemur is due to the greater sensitivity of the SS2 scRNAseq method used for mouse samples.

Figure S5 UMAP and marker genes of previously unknown cell types

a, b. UMAP (a) of the bone stromal and neural cells (10x scRNAseq dataset; L2, L4), integrated by FIRM and colored according to molecular cell types. Dashed circle, unknown (*NGFR*+ *TNNT2*+) stromal cell type (#141). Dot plot (b) showing fraction of expressing cells and mean level of expression in the molecular cell types in panel a for the indicated compartment/cell type marker genes and genes differentially expressed in the unknown stromal cell type (#141) relative to all other cells from bone.

c, d. UMAP (c) of the tongue stromal and neural cells (10x scRNAseq dataset; L2, L4) integrated by FIRM and colored according to molecular cell types. Dashed circles, unknown stromal (*NGFR*+ *TNNT2*+) cell type (#141) and (*COL15A1*+ *PTGDS*+) cell type (#142). Dot plot (d) as above showing expression in the molecular cell types in panel c of compartment/cell type marker genes and genes differentially expressed in the unknown stromal cell types (#141, 142) relative to all other cells from tongue.

e. Dot plot as above showing expression (10x scRNAseq datasets, L1-L4) of indicated marker genes and differentially-expressed genes in the unknown stromal (*NGFR*+ *TNNT2*+) populations (#141) from mammary gland, bone, tongue, and pancreas, and the (*COL15A1*+ *PTGDS*+) population (#142) found in tongue. Also included are cardiomyocytes, mesothelial cells,

leptomeningeal cells, and Schwann cells from all tissues, which express high levels of the two unknown populations' tissue-specific, differentially-expressed genes.

f-g. UMAP (f) of kidney stromal cells (10x scRNAseq dataset, L2, L3, L4), integrated by FIRM and colored according to molecular cell types. Dashed circle, unknown stromal (*ST6GAL2*+) cell type (#143). Dot plot (g) as above showing expression in kidney epithelial, endothelial, and stromal cell types of indicated compartment/cell type markers and genes differentially expressed in the unknown stromal cell type (#143) relative to all other cells from kidney.

h, i. UMAP (h) of epithelial and stromal cells from gonadal fat adipocyte tissue (GAT) (10x scRNAseq dataset; L2), integrated by FIRM and colored according to molecular cell types. Dashed circle, unknown epithelial (*CRISP3*+) cell type (#15). Dot plot (i) as above showing expression in the molecular cell types in panel h of indicated compartment/cell type marker genes and genes differentially expressed in the unknown epithelial molecular cell type (#15) relative to all other cells in the GAT.

j. Dot plot as above showing expression of indicated compartment/cell type marker genes (and genes differentially expressed in epithelial (*CRISP3*+) cell type) in the epithelial (*CRISP3*+) population (#15) from GAT, as well as urothelial cells and mesothelial cells from all tissues (10x scRNAseq dataset, L1-L4), which share expression of some of the genes differentially expressed in the unknown epithelial (*CRISP3*+) population (#15). Fibroblasts are also included for comparison. [], description of gene identified by NCBI as a gene locus (LOC109730246 [uncharacterized 1]).

k, l. UMAP (k) of the blood cells (SS2 scRNAseq dataset, L2, L4), integrated by FIRM and colored according to molecular cell types. Dashed circle, unknown epithelial (*PGAP1*+) cell type (#16). Dot plot (l) as above showing expression of the molecular cell types in panel k of

indicated compartment/cell type marker genes and genes differentially expressed in the unknown epithelial molecular cell type (#16) relative to all other cells from blood. [], description of gene identified by NCBI as a gene locus (LOC105861777 [uncharacterized 2], LOC105884518 [TTC39B]).

m. Dot plot as above showing expression of indicated compartment/cell type marker genes (and genes differentially expressed in epithelial (*PGAP1+*) cell type) in the epithelial (*PGAP1+*) cell type (#16) from blood, as well as astrocytes, oligodendrocyte precursor cells (OPC), oligodendrocytes, and ependymal cells from all tissues (SS2 scRNAseq dataset, L1-L4), which express high levels of some of the epithelial (*PGAP1+*) tissue-specific, differentially-expressed genes.

Figure S6 UMAP of lemur molecular cell types showing expression of compartment genes and differentially-expressed genes, and clustering of neutrophil and adipocyte molecular types

a. UMAP of all molecular cell types identified from 10x scRNAseq showing mean expression level of compartment (e.g., epithelial, endothelial, stromal) markers as heat map, with expression values normalized across all cell types. Colors below data points indicate tissue compartment (see Fig. 5a for compartment color legend and more detailed cell type labels).

b-d. UMAP as in panel a showing expression level of indicated differentially-expressed genes. Examples shown include genes enriched in immune progenitor/proliferating cells and germ cells relative to cell types in other compartments (*VRK1, CINS2*, b) or to proliferating cells in other compartments (*RSPH14, ASB9*, c), and genes enriched in Schwann and stromal cells relative to non-Schwann cell types in neural compartment (*LAMC1, SOCS3*, d).

e, f. Close up of portions of UMAP in panel a showing segregation of three types of neutrophils (e) and of two types of adipocytes (f).

Figure S7 Pairwise comparisons of the relationships of the transcriptomic profiles of all molecular cell types across the mouse lemur atlas

Heat maps as in Fig. 5g of pairwise Pearson correlation coefficients between the transcriptomic profiles of each of the 749 molecular cell type in atlas (10x and SS2 datasets, excluding all cardiac cells because of low quality), calculated from principal component values of FIRM-integrated UMAP (Fig. 1g) averaged across all cells of each type. Cell types ordered by compartment, then cell type annotation number, then tissue (a); by tissue, then compartment, then cell type annotation number (b); and by hierarchical clustering of cell type expression profiles (c). Interactive forms of the heat maps are available at the Tabula Microcebus portal https://tabula-microcebus.ds.czbiohub.org/.

Figure S8 Molecular relationship of human, lemur, and mouse lung and limb muscle cell types

a, c. Sankey plot showing the relationship between human, mouse lemur, and mouse cell types (10x scRNAseq dataset; L) for lung (a) and for limb muscle (c) as determined by SAMap algorithm (Tarashansky et al. 2021) (see Methods). Each cell type in lemur is connected (grey line) to the cell types it maps to in human and mouse datasets; thickness of line indicates similarity score (0-1) between connected cell types. A cell type with no connecting lines indicates it did not map with similarity score > 0.1 to any cell type in the other species.

b, d. UMAP of the indicated homologous cell types in human, lemur, and mouse from lung (b) and from limb muscle (d) corrected for batch effect by BBKNN integration algorithm (Polański et al. 2020) based on the 13,302 one-to-one-to-one gene orthologs. UMAP colored by compartment (left panel), cell type designation (center), and species (right).

e. Comparison of the transcriptomic profiles (10x scRNAseq datasets generated and annotated in the same way for all three species) of orthologous lemur (Tabula Microcebus, this paper), human (Tabula Sapiens, (The Tabula Sapiens Consortium 2021)), and mouse (Tabula Muris Senis, (The Tabula Muris Consortium 2020)) muscle cell types, as in Fig. 6a for lung cell types. ΔXi values and 95% confidence intervals are shown, with (lemur, human) - (lemur, mouse) ΔXi values in left plot (positive values indicate lemur cell type is more similar to human cell type than to mouse cell type), and (human, lemur) - (human, mouse) ΔXi values in right plot (positive values indicate human cell type is more similar to lemur cell type than to mouse cell type). P-values are from t-test with null hypothesis ΔXi=0.

Figure S9 Orthologous genes with species-selective expression patterns in lung and muscle cell types

Dot plots showing expression of the indicated one-to-one-to-one gene orthologs (symbol is for human ortholog) in human (H), lemur (L), and mouse (L) lung (a-f) and muscle (g-l) cell types showing human/lemur-enriched (a, g), mouse-enriched (b, h), human/mouse-enriched (c, i), lemur-enriched (d, j), lemur/mouse-enriched (e, k), and human-enriched (f, l) expression while also displaying selective expression in the corresponding cell type and/or compartment (genes and cell types labeled in compartment colors of Fig. 2). See Table 5 for full list of species-

enriched genes for each lung and muscle cell type. In addition to the primate-enriched genes of lung cell types described in Section 6, note in panel g primate-enriched expression in muscle stromal cells (including MuSCs, FAPs, and tenocytes) of splicing factor *NOVA1*, a proposed driver of human divergence in neuronal cells (Trujillo et al. 2021)). Note in panel h the primate-reduced expression in tenocytes of ECM proteins *FBLN7*, *OGN*, and *FMOD,* and matrix processing enzyme *HAS1*, suggesting evolutionarily divergent remodeling mechanisms of tenocyte ECM. Primate MuSCs show reduced expression of *VCAM1*, *SDC4*, and *CD34*, known cell surface markers for mouse MuSCs (Liu et al. 2015; Cornelison et al. 2001; Sacco et al. 2008), and enriched expression of *MYF6* and *DLK1*. Note also primate enrichment of type 1 angiotensin II receptor (*AGTR1*) in muscle pericytes, as in lung pericytes (Fig. 6b), and of genes that affect tumor angiogenesis including *IGFBP2 (Azar et al. 2011)*, *NOSTRIN (J. Wang et al. 2016)*, and *HEG1 (Zhao et al. 2019)* in vein and capillary endothelial cells. Primate muscle macrophages are enriched for *VSIG4*, an inflammatory response suppressor (Li et al. 2017).

REFERENCES

Albee, Alison J., Alan L. Kwan, Huawen Lin, David Granas, Gary D. Stormo, and Susan K. Dutcher. 2013. "Identification of Cilia Genes That Affect Cell-Cycle Progression Using Whole-Genome Transcriptome Analysis in Chlamydomonas Reinhardtti." *G3* 3 (6): 979–91.

Alexander, Matthew S., Anete Rozkalne, Alessandro Colletta, Janelle M. Spinazzola, Samuel Johnson, Fedik Rahimov, Hui Meng, et al. 2016. "CD82 Is a Marker for Prospective Isolation of Human Muscle Satellite Cells and Is Linked to Muscular Dystrophies." *Cell Stem Cell* 19 (6): 800–807.

Andrès, M., M. Solignac, and M. Perret. 2003. "Mating System in Mouse Lemurs: Theories and Facts, Using Analysis of Paternity." *Folia Primatologica; International Journal of Primatology* 74 (5-6). https://doi.org/10.1159/000073319.

Atsalis, Sylvia. 2015. *A Natural History of the Brown Mouse Lemur*. Vol. 6. Primate Field Studies Series. Prentice Hall.

Azar, Walid J., Sheena H. X. Azar, Sandra Higgins, Ji-Fan Hu, Andrew R. Hoffman, Donald F. Newgreen, George A. Werther, and Vincenzo C. Russo. 2011. "IGFBP-2 Enhances VEGF Gene Promoter Activity and Consequent Promotion of Angiogenesis by Neuroblastoma Cells." *Endocrinology* 152 (9): 3332–42.

Ben-Moshe, Shani, and Shalev Itzkovitz. 2019. "Spatial Heterogeneity in the Mammalian Liver." *Nature Reviews. Gastroenterology & Hepatology* 16 (7): 395–410.

Boehm, Ines, Abrar Alhindi, Ana S. Leite, Chandra Logie, Alyssa Gibbs, Olivia Murray, Rizwan Farrukh, et al. 2020. "Comparative Anatomy of the Mammalian Neuromuscular Junction." *Journal of Anatomy* 237 (5): 827–36.

Caprioli, Arianna, Hongxin Zhu, and Thomas N. Sato. 2004. "CRBP-III:lacZ Expression Pattern Reveals a Novel Heterogeneity of Vascular Endothelial Cells." *Genesis* 40 (3): 139–45.

Casey, Kerriann M., Caitlin J. Karanewsky, Jozeph L. Pendleton, Mark R. Krasnow, and Megan A. Albertelli. 2021. "Fibrous Osteodystrophy, Chronic Renal Disease, and Uterine Adenocarcinoma in Aged Gray Mouse Lemurs ()." *Comparative Medicine* 71 (3): 256–66.

Celton-Morizur, S., and C. Desdouets. 2010. "Polyploidization of Liver Cells." *Advances in Experimental Medicine and Biology* 676. https://doi.org/10.1007/978-1-4419-6199-0_8.

Chatterjee, Sourav. 2019. "A New Coefficient of Correlation." http://arxiv.org/abs/1909.10140.

Choksi, Semil P., Gilbert Lauter, Peter Swoboda, and Sudipto Roy. 2014. "Switching on Cilia: Transcriptional Networks Regulating Ciliogenesis." *Development* 141 (7): 1427–41.

Coppiello, Giulia, Maria Collantes, María Salomé Sirerol-Piquer, Sara Vandenwijngaert, Sandra Schoors, Melissa Swinnen, Ine Vandersmissen, et al. 2015. "Meox2/Tcf15 Heterodimers Program the Heart Capillary Endothelium for Cardiac Fatty Acid Uptake." *Circulation* 131 (9): 815–26.

Cornelison, D. D., M. S. Filla, H. M. Stanley, A. C. Rapraeger, and B. B. Olwin. 2001. "Syndecan-3 and Syndecan-4 Specifically Mark Skeletal Muscle Satellite Cells and Are Implicated in Satellite Cell Maintenance and Muscle Regeneration." *Developmental Biology* 239 (1): 79–94.

Desvignes, Thomas, Pierre Pontarotti, Christian Fauvel, and Julien Bobe. 2009. "Nme Protein Family Evolutionary History, a Vertebrate Perspective." *BMC Evolutionary Biology*. https://doi.org/10.1186/1471-2148-9-256.

Diehl, A. D., T. F. Meehan, Y. M. Bradford, M. H. Brush, W. M. Dahdul, D. S. Dougall, Y. He, et al. 2016. "The Cell Ontology 2016: Enhanced Content, Modularization, and Ontology Interoperability." *Journal of Biomedical Semantics* 7 (1). https://doi.org/10.1186/s13326-016-0088-7.

Du, Y., R. A. Fenton, L. Schoonjans, M. Dewerchin, G. Eelen, B. Thienpont, L. Lin, et al. 2020. "Single-Cell Transcriptome Atlas of Murine Endothelial Cells." *Cell* 180 (4): 764–79.e20.

Elsea, S. H., and R. E. Lucas. 2002. "The Mousetrap: What We Can Learn When the Mouse Model Does Not Mimic the Human Disease." *ILAR Journal / National Research Council, Institute of Laboratory Animal Resources* 43 (2). https://doi.org/10.1093/ilar.43.2.66.

Ernst, Christina, Nils Eling, Celia P. Martinez-Jimenez, John C. Marioni, and Duncan T. Odom. 2019. "Staged Developmental Mapping and X Chromosome Transcriptional Dynamics during Mouse Spermatogenesis." *Nature Communications* 10 (1): 1–20.

Ezran, Camille, Caitlin J. Karanewsky, Jozeph L. Pendleton, Alex Sholtz, Maya R. Krasnow, Jason Willick, Andriamahery Razafindrakoto, Sarah Zohdy, Megan A. Albertelli, and Mark A. Krasnow. 2017. "The Mouse Lemur, a Genetic Model Organism for Primate Biology, Behavior, and Health." *Genetics* 206 (2): 651–64.

Farber, Donna L., Naomi A. Yudanin, and Nicholas P. Restifo. 2014. "Human Memory T Cells: Generation, Compartmentalization and Homeostasis." *Nature Reviews. Immunology* 14 (1): 24–35.

Farouni, Rick, Haig Djambazian, Lorenzo E. Ferri, Jiannis Ragoussis, and Hamed S. Najafabadi. 2020. "Model-Based Analysis of Sample Index Hopping Reveals Its Widespread Artifacts in Multiplexed Single-Cell RNA-Sequencing." *Nature Communications* 11 (1): 2704.

Fujimoto, Noriki, Yuliang He, Marco D'Addio, Carlotta Tacconi, Michael Detmar, and Lothar C. Dieterich. 2020. "Single-Cell Mapping Reveals New Markers and Functions of Lymphatic Endothelial Cells in Lymph Nodes." *PLoS Biology* 18 (4): e3000704.

Génin, F., A. Schilling, and M. Perret. 2005. "Social Inhibition of Seasonal Fattening in Wild and Captive Gray Mouse Lemurs." *Physiology & Behavior* 86 (1-2): 185–94.

Gillich, A., F. Zhang, C. G. Farmer, K. J. Travaglini, S. Y. Tan, M. Gu, B. Zhou, J. A. Feinstein, M. A. Krasnow, and R. J. Metzger. 2020. "Capillary Cell-Type Specialization in the Alveolus." *Nature* 586 (7831). https://doi.org/10.1038/s41586-020-2822-7.

He, Bing, Ping Chen, Sonia Zambrano, Dina Dabaghie, Yizhou Hu, Katja Möller-Hackbarth, David Unnersjö-Jess, et al. 2021. "Single-Cell RNA Sequencing Reveals the Mesangial Identity and Species Diversity of Glomerular Cell Transcriptomes." *Nature Communications* 12 (1): 2141.

Hermann, B. P., K. Cheng, A. Singh, Roa-De La Cruz L, K. N. Mutoji, I. C. Chen, H. Gildersleeve, et al. 2018. "The Mammalian Spermatogenesis Single-Cell Transcriptome, from Spermatogonial Stem Cells to Spermatids." *Cell Reports* 25 (6). https://doi.org/10.1016/j.celrep.2018.10.026.

Hozer, Clara, Fabien Pifferi, Fabienne Aujard, and Martine Perret. 2019. "The Biological Clock in Gray Mouse Lemur: Adaptive, Evolutionary and Aging Considerations in an Emerging Non-Human Primate Model." *Frontiers in Physiology* 10 (August): 1033.

Hu, Chunyan, Henry L. Keen, Ko-Ting Lu, Xuebo Liu, Jing Wu, Deborah R. Davis, Stella-Rita C. Ibeawuchi, Silke Vogel, Frederick W. Quelle, and Curt D. Sigmund. 2017. "Retinol-Binding Protein 7 Is an Endothelium-Specific PPAR Cofactor Mediating an Antioxidant Response through Adiponectin." *JCI Insight* 2 (6): e91738.

Iso, Tatsuya, Kazuhisa Maeda, Hirofumi Hanaoka, Toshihiro Suga, Kosaku Goto, Mas Rizky A. A. Syamsunarno, Takako Hishiki, et al. 2013. "Capillary Endothelial Fatty Acid Binding Proteins 4 and 5 Play a Critical Role in Fatty Acid Uptake in Heart and Skeletal Muscle." *Arteriosclerosis, Thrombosis, and Vascular Biology* 33 (11): 2549–57.

Italiani, Paola, and Diana Boraschi. 2014. "From Monocytes to M1/M2 Macrophages: Phenotypical vs. Functional Differentiation." *Frontiers in Immunology* 5 (October): 514.

Jain, Miten, Sergey Koren, Karen H. Miga, Josh Quick, Arthur C. Rand, Thomas A. Sasani, John R. Tyson, et al. 2018. "Nanopore Sequencing and Assembly of a Human Genome with Ultra-Long Reads." *Nature Biotechnology* 36 (4): 338–45.

Jalkanen, Sirpa, and Marko Salmi. 2020. "Lymphatic Endothelial Cells of the Lymph Node." *Nature Reviews. Immunology* 20 (9): 566–78.

Jan, Sabrina Z., Tinke L. Vormer, Aldo Jongejan, Michael D. Röling, Sherman J. Silber, Dirk G. de Rooij, Geert Hamer, Sjoerd Repping, and Ans M. M. van Pelt. 2017. "Unraveling Transcriptome Dynamics in Human Spermatogenesis." *Development* 144 (20): 3659–73.

Jarolim, Petr. 2015. "High Sensitivity Cardiac Troponin Assays in the Clinical Laboratories." *Clinical Chemistry and Laboratory Medicine: CCLM / FESCC* 53 (5): 635–52.

Joe, Aaron W. B., Lin Yi, Anuradha Natarajan, Fabien Le Grand, Leslie So, Joy Wang, Michael A.

Rudnicki, and Fabio M. V. Rossi. 2010. "Muscle Injury Activates Resident Fibro/adipogenic Progenitors That Facilitate Myogenesis." *Nature Cell Biology* 12 (2): 153–63.

Jung, Min, Daniel Wells, Jannette Rusch, Suhaira Ahmad, Jonathan Marchini, Simon R. Myers, and Donald F. Conrad. 2019. "Unified Single-Cell Analysis of Testis Gene Regulation and Pathology in Five Mouse Strains." *eLife* 8 (June). https://doi.org/10.7554/eLife.43966.

Kraska, Audrey, Olene Dorieux, Jean-Luc Picq, Fanny Petit, Emmanuel Bourrin, Evelyne Chenu, Andreas Volk, et al. 2011. "Age-Associated Cerebral Atrophy in Mouse Lemur Primates." *Neurobiology of Aging* 32 (5): 894–906.

Languille, S., S. Blanc, O. Blin, C. I. Canale, A. Dal-Pan, G. Devau, M. Dhenain, et al. 2012. "The Grey Mouse Lemur: A Non-Human Primate Model for Ageing Studies." *Ageing Research Reviews* 11 (1): 150–62.

Larsen, P. A., R. A. Harris, Y. Liu, S. C. Murali, C. R. Campbell, A. D. Brown, B. A. Sullivan, et al. 2017. "Hybrid de Novo Genome Assembly and Centromere Characterization of the Gray Mouse Lemur (Microcebus Murinus)." *BMC Biology* 15 (1). https://doi.org/10.1186/s12915-017-0439-6.

Larsson, Anton J. M., Geoff Stanley, Rahul Sinha, Irving L. Weissman, and Rickard Sandberg. 2018. "Computational Correction of Index Switching in Multiplexed Sequencing Libraries." *Nature Methods* 15 (5): 305–7.

Li, Jialin, Bo Diao, Sheng Guo, Xiaoyong Huang, Chengying Yang, Zeqing Feng, Weiming Yan, et al. 2017. "VSIG4 Inhibits Proinflammatory Macrophage Activation by Reprogramming Mitochondrial Pyruvate Metabolism." *Nature Communications* 8 (1): 1322.

Liu, Ling, Tom H. Cheung, Gregory W. Charville, and Thomas A. Rando. 2015. "Isolation of Skeletal Muscle Stem Cells by Fluorescence-Activated Cell Sorting." *Nature Protocols* 10 (10): 1612–24.

Martin, Matthew D., and Vladimir P. Badovinac. 2018. "Defining Memory CD8 T Cell." *Frontiers in Immunology* 9 (November): 2692.

Megill, Colin, Bruce Martin, Charlotte Weaver, Sidney Bell, Lia Prins, Seve Badajoz, Brian McCandless, et al. 2021. "Cellxgene: A Performant, Scalable Exploration Platform for High Dimensional Sparse Matrices." *bioRxiv*. https://doi.org/10.1101/2021.04.05.438318.

Mestas, J., and C. C. Hughes. 2004. "Of Mice and Not Men: Differences between Mouse and Human Immunology." *Journal of Immunology* 172 (5). https://doi.org/10.4049/jimmunol.172.5.2731.

Ming, J., Z. Lin, J. Zhao, X. Wan, C. Yang, and A. R. Wu. 2021. "FIRM: Flexible Integration of Single-Cell RNA-Sequencing Data for Large-Scale Multi-Tissue Cell Atlas Datasets." *bioRxiv*. https://www.biorxiv.org/content/10.1101/2020.06.02.129031.abstract.

Miranda-Vizuete, Antonio, Katie Tsang, Yang Yu, Alberto Jiménez, Markku Pelto-Huikko, Charles J. Flickinger, Peter Sutovsky, and Richard Oko. 2003. "Cloning and Developmental Analysis of Murid Spermatid-Specific Thioredoxin-2 (SPTRX-2), a Novel Sperm Fibrous Sheath Protein and Autoantigen." *The Journal of Biological Chemistry* 278 (45): 44874–85.

Montoro, Daniel T., Adam L. Haber, Moshe Biton, Vladimir Vinarsky, Brian Lin, Susan E. Birket, Feng Yuan, et al. 2018. "A Revised Airway Epithelial Hierarchy Includes CFTR-Expressing Ionocytes." *Nature* 560 (7718): 319–24.

Namekawa, Satoshi H., Peter J. Park, Li-Feng Zhang, James E. Shima, John R. McCarrey, Michael D. Griswold, and Jeannie T. Lee. 2006. "Postmeiotic Sex Chromatin in the Male Germline of Mice." *Current Biology: CB* 16 (7): 660–67.

Neumann, Johannes T., Raphael Twerenbold, Francisco Ojeda, Nils A. Sörensen, Andrew R. Chapman, Anoop S. V. Shah, Atul Anand, et al. 2019. "Application of High-Sensitivity Troponin in Suspected Myocardial Infarction." *The New England Journal of Medicine* 380 (26): 2529–40.

Oatley, Melissa, Frieda Chan, Amy Kaucher, and Jon Oatley. 2014. "Functional and Molecular Features of the ID4 Germline Stem Cell Population in Mouse Testes." *Reproduction Abstracts*. https://doi.org/10.1530/repabs.1.p177.

Peichl, L., A. Kaiser, F. Rakotondraparany, R. R. Dubielzig, S. M. Goodman, and P. M. Kappeler. 2019. "Diversity of Photoreceptor Arrangements in Nocturnal, Cathemeral and Diurnal Malagasy Lemurs."

The Journal of Comparative Neurology 527 (1). https://doi.org/10.1002/cne.24167.
Perret, M., and F. Aujard. 2001. "Daily Hypothermia and Torpor in a Tropical Primate: Synchronization by 24-H Light-Dark Cycle." *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology* 281 (6). https://doi.org/10.1152/ajpregu.2001.281.6.R1925.
Peti-Peterdi, János, and Raymond C. Harris. 2010. "Macula Densa Sensing and Signaling Mechanisms of Renin Release." *Journal of the American Society of Nephrology: JASN* 21 (7): 1093–96.
Pisani, Didier F., Cynthia D. K. Bottema, Catherine Butori, Christian Dani, and Claude A. Dechesne. 2010. "Mouse Model of Skeletal Muscle Adiposity: A Glycerol Treatment Approach." *Biochemical and Biophysical Research Communications* 396 (3): 767–73.
Plasschaert, Lindsey W., Rapolas Žilionis, Rayman Choo-Wing, Virginia Savova, Judith Knehr, Guglielmo Roma, Allon M. Klein, and Aron B. Jaffe. 2018. "A Single-Cell Atlas of the Airway Epithelium Reveals the CFTR-Rich Pulmonary Ionocyte." *Nature* 560 (7718): 377–81.
Polański, Krzysztof, Matthew D. Young, Zhichao Miao, Kerstin B. Meyer, Sarah A. Teichmann, and Jong-Eun Park. 2020. "BBKNN: Fast Batch Alignment of Single Cell Transcriptomes." *Bioinformatics* 36 (3): 964–65.
Radespiel, U. 2000. "Sociality in the Gray Mouse Lemur (Microcebus Murinus) in Northwestern Madagascar." *American Journal of Primatology* 51 (1). https://doi.org/10.1002/(SICI)1098-2345(200005)51:1<21::AID-AJP3>3.0.CO;2-C.
Ravnik, Stuart E., and Debra J. Wolgemuth. 1996. "The Developmentally Restricted Pattern of Expression in the Male Germ Line of a MurineCyclin A, Cyclin A2,Suggests Roles in Both Mitotic and Meiotic Cell Cycles." *Developmental Biology* 173 (1): 69–78.
Reuter, Michael, Philipp Berninger, Shinichiro Chuma, Hardik Shah, Mihoko Hosokawa, Charlotta Funaya, Claude Antony, Ravi Sachidanandam, and Ramesh S. Pillai. 2011. "Miwi Catalysis Is Required for piRNA Amplification-Independent LINE1 Transposon Silencing." *Nature* 480 (7376): 264–67.
Rina Evasoa, Mamy, Ute Radespiel, Alida F. Hasiniaina, Solofonirina Rasoloharijaona, Blanchard Randrianambinina, Romule Rakotondravony, and Elke Zimmermann. 2018. "Variation in Reproduction of the Smallest-Bodied Primate Radiation, the Mouse Lemurs (Microcebus Spp.): A Synopsis." *American Journal of Primatology* 80 (7): e22874.
Sacco, Alessandra, Regis Doyonnas, Peggy Kraft, Stefan Vitorovic, and Helen M. Blau. 2008. "Self-Renewal and Expansion of Single Transplanted Muscle Stem Cells." *Nature* 456 (7221): 502–6.
Shami, Adrienne Niederriter, Xianing Zheng, Sarah K. Munyoki, Qianyi Ma, Gabriel L. Manske, Christopher D. Green, Meena Sukhwani, Kyle E. Orwig, Jun Z. Li, and Saher Sue Hammoud. 2020. "Single-Cell RNA Sequencing of Human, Macaque, and Mouse Testes Uncovers Conserved and Divergent Features of Mammalian Spermatogenesis." *Developmental Cell* 54 (4): 529–47.e12.
Shekhar, K., S. W. Lapan, I. E. Whitney, N. M. Tran, E. Z. Macosko, M. Kowalczyk, X. Adiconis, et al. 2016. "Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics." *Cell* 166 (5). https://doi.org/10.1016/j.cell.2016.07.054.
Sin, Ho-Su, Yosuke Ichijima, Eitetsu Koh, Mikio Namiki, and Satoshi H. Namekawa. 2012. "Human Postmeiotic Sex Chromatin and Its Impact on Sex Chromosome Evolution." *Genome Research* 22 (5): 827–36.
Sin, Ho-Su, and Satoshi H. Namekawa. 2013. "The Great Escape: Active Genes on Inactive Sex Chromosomes and Their Evolutionary Implications." *Epigenetics: Official Journal of the DNA Methylation Society* 8 (9): 887–92.
Sofroniew, Nicholas, Talley Lambert, Kira Evans, Juan Nunez-Iglesias, Grzegorz Bokota, Gonzalo Peña-Castellanos, Philip Winston, et al. 2021. *Napari/napari: 0.4.12rc2*. Zenodo. https://doi.org/10.5281/ZENODO.3555620.
Sun, Feng, Qing Xu, Danfeng Zhao, and Charlie Degui Chen. 2015. "Id4 Marks Spermatogonial Stem Cells in the Mouse Testis." *Scientific Reports* 5 (December): 17594.
Sun, Yu H., Feng Jiang, and Xin Zhiguo Li. 2018. "Disruption of Tdrd5 Decouples the Stepwise

Processing of Long Precursor Transcripts during Pachytene PIWI-Interacting RNA Biogenesis." *Biology of Reproduction* 99 (4): 684–85.

Takeda, Akira, Maija Hollmén, Denis Dermadi, Junliang Pan, Kevin Francis Brulois, Riina Kaukonen, Tapio Lönnberg, et al. 2019. "Single-Cell Survey of Human Lymphatics Unveils Marked Endothelial Cell Heterogeneity and Mechanisms of Homing for Neutrophils." *Immunity* 51 (3): 561–72.e5.

Tarashansky, Alexander J., Jacob M. Musser, Margarita Khariton, Pengyang Li, Detlev Arendt, Stephen R. Quake, and Bo Wang. 2021. "Mapping Single-Cell Atlases throughout Metazoa Unravels Cell Type Evolution." *bioRxiv*. https://doi.org/10.1101/2020.09.28.317784.

Tarashansky, Alexander J., Yuan Xue, Pengyang Li, Stephen R. Quake, and Bo Wang. 2019. "Self-Assembling Manifolds in Single-Cell RNA Sequencing Data," September. https://doi.org/10.7554/eLife.48994.

Terrien, J., M. Gaudubois, D. Champeval, V. Zaninotto, L. Roger, J. F. Riou, and F. Aujard. 2018. "Metabolic and Genomic Adaptations to Winter Fattening in a Primate Species, the Grey Mouse Lemur (Microcebus Murinus)." *International Journal of Obesity* 42 (2): 221–30.

The Tabula Muris Consortium. 2018. "Single-Cell Transcriptomics of 20 Mouse Organs Creates a Tabula Muris." *Nature*. https://doi.org/10.1038/s41586-018-0590-4.

The Tabula Muris Consortium. 2020. "A Single-Cell Transcriptomic Atlas Characterizes Ageing Tissues in the Mouse." *Nature* 583 (7817): 590–95.

The Tabula Sapiens Consortium. 2021. "The Tabula Sapiens: A Single Cell Transcriptomic Atlas of Multiple Organs from Individual Human Donors." *bioRxiv*. https://doi.org/10.1101/2021.07.19.452956.

Thul, Peter J., and Cecilia Lindskog. 2018. "The Human Protein Atlas: A Spatial Map of the Human Proteome." *Protein Science*. https://doi.org/10.1002/pro.3307.

Travaglini, Kyle J., Ahmad N. Nabhan, Lolita Penland, Rahul Sinha, Astrid Gillich, Rene V. Sit, Stephen Chang, et al. 2020. "A Molecular Cell Atlas of the Human Lung from Single-Cell RNA Sequencing." *Nature* 587 (7835): 619–25.

Trujillo, Cleber A., Edward S. Rice, Nathan K. Schaefer, Isaac A. Chaim, Emily C. Wheeler, Assael A. Madrigal, Justin Buchanan, et al. 2021. "Reintroduction of the Archaic Variant of in Cortical Organoids Alters Neurodevelopment." *Science* 371 (6530). https://doi.org/10.1126/science.aax2537.

Turner, James M. A. 2015. "Meiotic Silencing in Mammals." *Annual Review of Genetics* 49: 395–412.

Tuttle, Lori J., David R. Sinacore, and Michael J. Mueller. 2012. "Intermuscular Adipose Tissue Is Muscle Specific and Associated with Poor Functional Performance." *Journal of Aging Research* 2012 (May): 172957.

Uezumi, Akiyoshi, So-Ichiro Fukada, Naoki Yamamoto, Shin 'ichi Takeda, and Kunihiro Tsuchida. 2010. "Mesenchymal Progenitors Distinct from Satellite Cells Contribute to Ectopic Fat Cell Formation in Skeletal Muscle." *Nature Cell Biology* 12 (2): 143–52.

Uezumi, Akiyoshi, Masashi Nakatani, Madoka Ikemoto-Uezumi, Naoki Yamamoto, Mitsuhiro Morita, Asami Yamaguchi, Harumoto Yamada, et al. 2016. "Cell-Surface Protein Profiling Identifies Distinctive Markers of Progenitor Cells in Human Skeletal Muscle." *Stem Cell Reports* 7 (2): 263–78.

Van Durme, Y. M. T. A., M. Eijgelsheim, G. F. Joos, A. Hofman, A. G. Uitterlinden, G. G. Brusselle, and B. H. Ch Stricker. 2010. "Hedgehog-Interacting Protein Is a COPD Susceptibility Gene: The Rotterdam Study." *The European Respiratory Journal: Official Journal of the European Society for Clinical Respiratory Physiology* 36 (1): 89–95.

Wang, Jian, Shouhui Yang, Peijun He, Aaron J. Schetter, Jochen Gaedcke, B. Michael Ghadimi, Thomas Ried, et al. 2016. "Endothelial Nitric Oxide Synthase Traffic Inducer (NOSTRIN) Is a Negative Regulator of Disease Aggressiveness in Pancreatic Cancer." *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 22 (24): 5992–6001.

Wang, M., X. Liu, G. Chang, Y. Chen, G. An, L. Yan, S. Gao, et al. 2018. "Single-Cell RNA Sequencing

Analysis Reveals Sequential Cell Fate Transition during Human Spermatogenesis." *Cell Stem Cell* 23 (4). https://doi.org/10.1016/j.stem.2018.08.007.

Watanabe, Toshiaki, Ee-Chun Cheng, Mei Zhong, and Haifan Lin. 2015. "Retrotransposons and Pseudogenes Regulate mRNAs and lncRNAs via the piRNA Pathway in the Germline." *Genome Research* 25 (3): 368–80.

Watase, Kei, and Huda Y. Zoghbi. 2003. "Modelling Brain Diseases in Mice: The Challenges of Design and Analysis." *Nature Reviews. Genetics* 4 (4): 296–307.

Wosczyna, Michael N., Colin T. Konishi, Edgar E. Perez Carbajal, Theodore T. Wang, Rachel A. Walsh, Qiang Gan, Mark W. Wagner, and Thomas A. Rando. 2019. "Mesenchymal Stromal Cells Are Required for Regeneration and Homeostatic Maintenance of Skeletal Muscle." *Cell Reports* 27 (7): 2029–35.e5.

Wu, Angela R., Norma F. Neff, Tomer Kalisky, Piero Dalerba, Barbara Treutlein, Michael E. Rothenberg, Francis M. Mburu, et al. 2013. "Quantitative Assessment of Single-Cell RNA-Sequencing Methods." *Nature Methods* 11 (1): 41–46.

Xu, Xiaoti, Karlijn J. Wilschut, Gayle Kouklis, Hua Tian, Robert Hesse, Catharine Garland, Hani Sbitany, et al. 2015. "Human Satellite Cell Transplantation and Regeneration from Diverse Skeletal Muscles." *Stem Cell Reports* 5 (3): 419–34.

Yoder, A. D., R. M. Rasoloarison, S. M. Goodman, J. A. Irwin, S. Atsalis, M. J. Ravosa, and J. U. Ganzhorn. 2000. "Remarkable Species Diversity in Malagasy Mouse Lemurs (primates, Microcebus)." *Proceedings of the National Academy of Sciences of the United States of America* 97 (21): 11325–30.

Zhao, Yan-Rong, Ji-Long Wang, Cong Xu, Yi-Ming Li, Bo Sun, and Lian-Yue Yang. 2019. "HEG1 Indicates Poor Prognosis and Promotes Hepatocellular Carcinoma Invasion, Metastasis, and EMT by Activating Wnt/β-Catenin Signaling." *Clinical Science* 133 (14): 1645–62.

Zipes, Douglas P., Peter Libby, Robert O. Bonow, Douglas L. Mann, and Gordon F. Tomaselli. 2018. *Braunwald's Heart Disease E-Book: A Textbook of Cardiovascular Medicine*. Elsevier Health Sciences.

Fig. 1

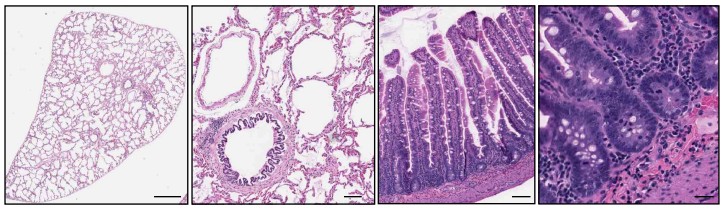
a

Mouse lemur (M. *murinus*)

L1: 9 yo, M
L2: 10 yo, F
L3: 11 yo, F
L4: 11 yo, M

- Clinical history & physical exam
- Blood
  - Metabolic panel & CBC
  - Single cell RNAseq
  - Biobank (cells)
- 26 other organs
  - Histology & pathology
  - Single cell RNAseq
  - Biobank (tissue, cells)
  - Culture (when indicated)

b



c

Nervous
1. Brain cortex (3120 10x cells, 292 SS2 cells, 3 individuals, 23 molecular types)
2. Brainstem (1988, 122, 2, 25)
3. Hypothalamus/Pituitary (2431, 0, 1, 34)
4. Retina (938, 120, 1, 11)

Reproductive
26. Uterus (1672, 0, 1, 18)
27. Testis (5639, 705, 1, 12)

Integumentary
5. Skin (23480, 676, 2, 19)
6. Fat (12601, 0, 2, 50)
7. Mammary gland (5396, 0, 1, 25)

Urinary
24. Kidney (13404, 1120, 3, 71)
25. Bladder (6708, 354, 2, 35)

Musculoskeletal
8. Limb muscle (2968, 994, 2, 31)
9. Diaphragm (76, 150, 2, 12)
10. Bone (14585, 1037, 2, 36)

Gastrointestinal
19. Tongue (5952, 449, 2, 27)
20. Small intestine (5825, 302, 2, 46)
21. Colon (1362, 194, 2, 10)
22. Liver (8024, 62, 3, 31)
23. Pancreas (7448, 324, 2, 54)

Respiratory
17. Lung (31527, 1014, 4, 64)
18. Trachea (489, 0, 3, 13)

Circulatory
11. Bone marrow (29948, 1434, 3, 33)
12. Blood (17607, 569, 4, 23)
13. Spleen (8806, 1292, 3, 31)
14. Thymus (0, 66, 1, 8)
15. Heart (2785, 535, 3, 19)
16. Aorta (111, 0, 1, 7)

d

Single Cell RNA sequencing pipeline

Tissue single cell suspension → Droplet-based cDNA library (10x protocol) / FACS sorting for live cells → Plate-based cDNA library (SS2 protocol) → Illumina sequencing → Genome alignment (NCBI Mmur3) 10x - cellranger SS2 - STAR → Quality filtering → 244,081 total cells; 226,701 high quality cells; 10x: 214,890; SS2: 11,811 → Cell annotation → 256 cell names; 768 molecular types

e

Cell annotation pipeline

Step 1 Manual annotation of each 10x object: cells from one tissue of one individual

2 FIRM integration of 10x & SS2 within same individual & tissue → Automatically annotate SS2 cells based on manual 10x annotation

3 FIRM integration of all individuals within same tissue → Review annotations for consistency across individuals

4 FIRM integration of all tissues → Review annotations for consistency across tissues

Visualize and cluster cells based on molecular expression (2-D tSNE/UMAP) → Compartment separation
- *Epithelial*
- *Endothelial*
- *Stromal*
- *Immune*
- *Neural*
- *Germ*

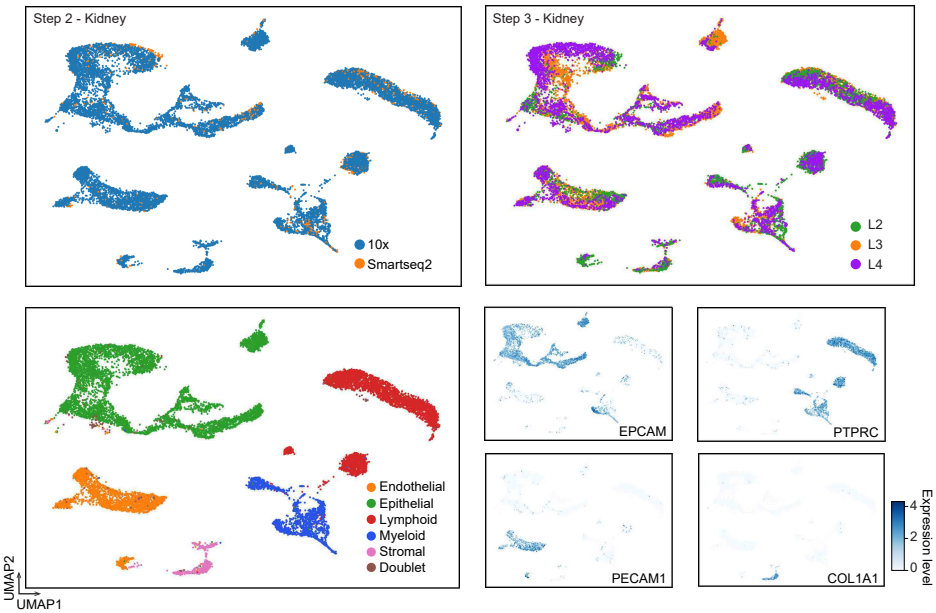
→ Cluster cells in compartment (Louvain) → *Unique molecular type in each cluster + expert curation* → Assign cell type name and corresponding cell ontology based on cannonical human and mouse marker genes → Add qualifier

Molecular diversity within cluster — Iteratively subcluster

- *Cell type*
- *Unknown*
- *Hybrid*
- *Mix*
- *Doublet*

Marker gene if >1 molecular cluster

f



g

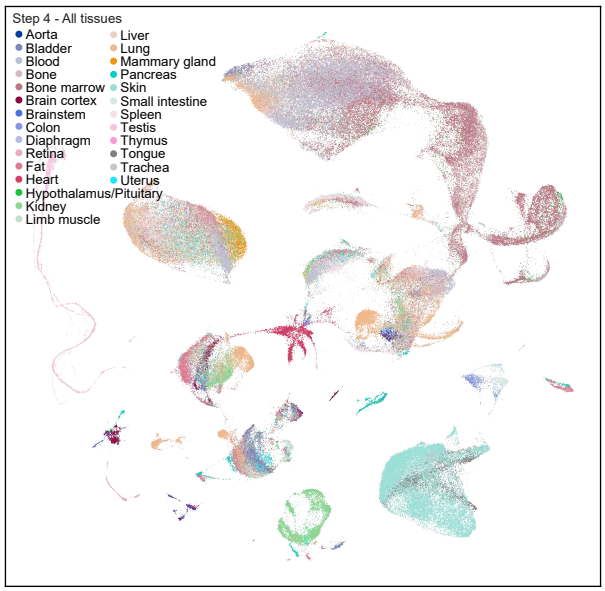
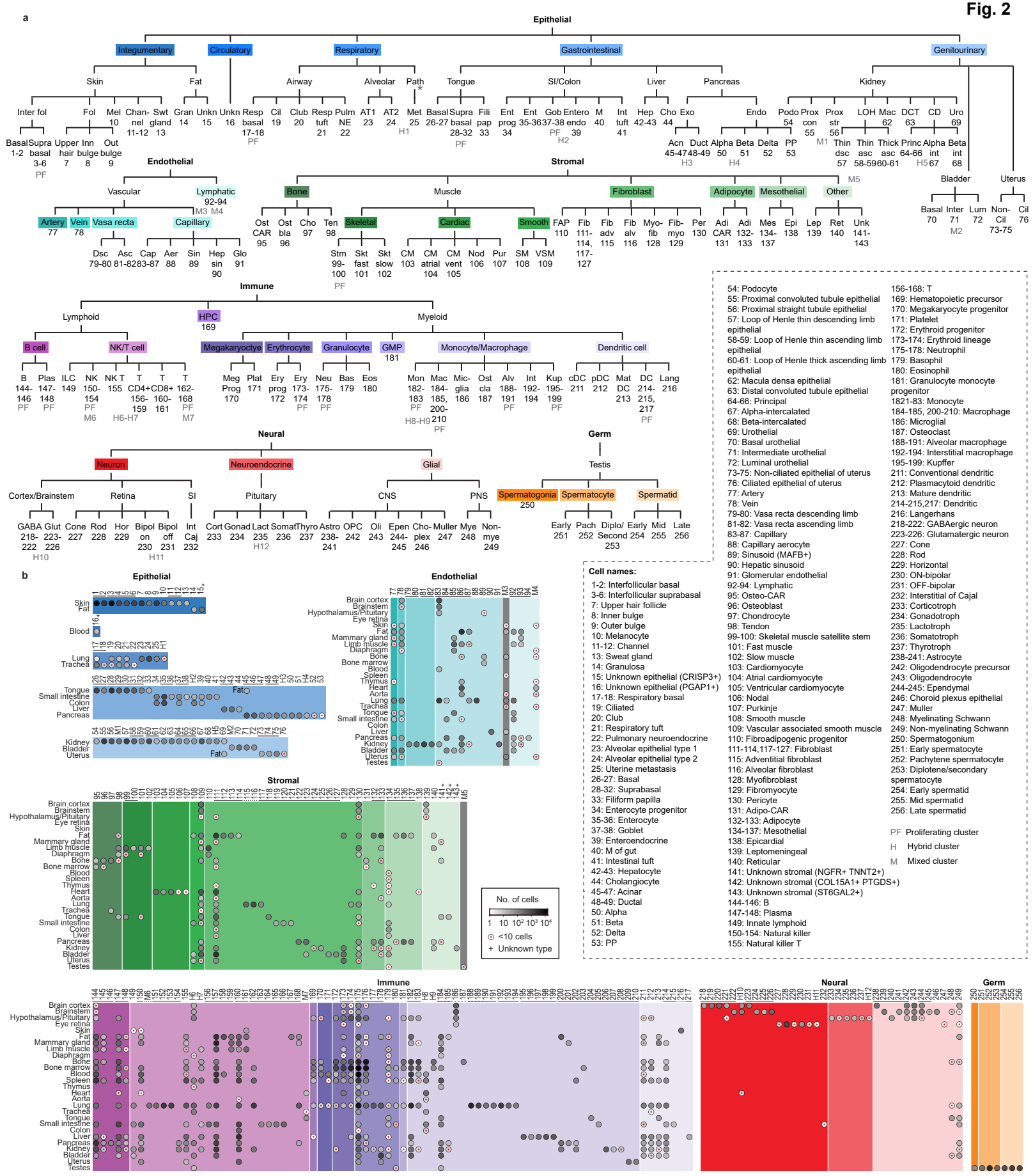


Fig. 2



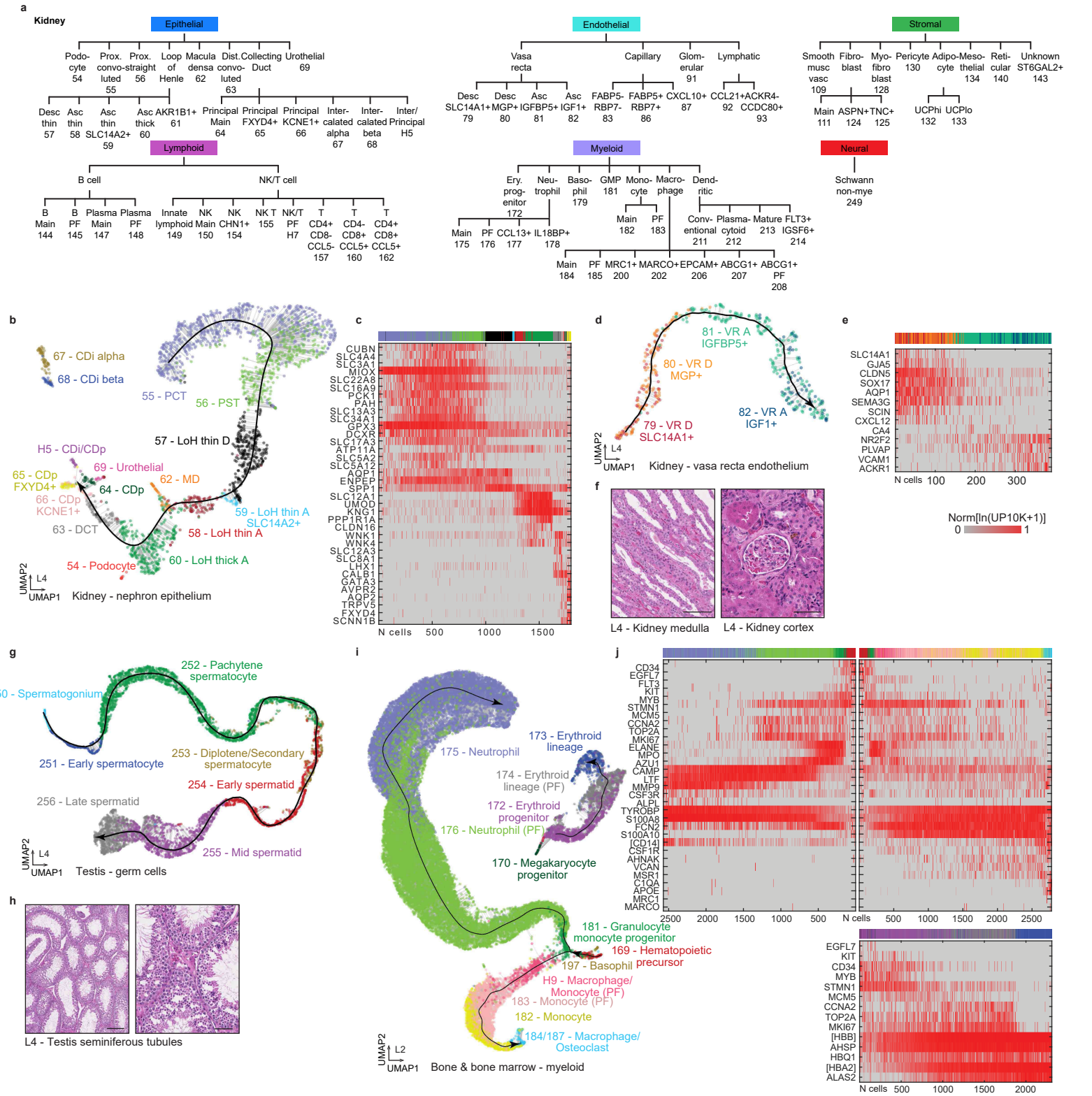
54: Podocyte
55: Proximal convoluted tubule epithelial
56: Proximal straight tubule epithelial
57: Loop of Henle thin descending limb epithelial
58-59: Loop of Henle thin ascending limb epithelial
60-61: Loop of Henle thick ascending limb epithelial
62: Macula densa epithelial
63: Distal convoluted tubule epithelial
64-66: Principal
67: Alpha-intercalated
68: Beta-intercalated
69: Urothelial
70: Basal urothelial
71: Intermediate urothelial
72: Luminal urothelial
73-75: Non-ciliated epithelial of uterus
76: Ciliated epithelial of uterus
77: Artery
78: Vein
79-80: Vasa recta descending limb
81-82: Vasa recta ascending limb
83-87: Capillary
88: Capillary aerocyte
89: Sinusoid (MAFB+)
90: Hepatic sinusoid
91: Glomerular endothelial
92-94: Lymphatic
95: Osteo-CAR
96: Osteoblast
97: Chondrocyte
98: Tendon
99-100: Skeletal muscle satellite stem
101: Fast muscle
102: Slow muscle
103: Cardiomyocyte
104: Atrial cardiomyocyte
105: Ventricular cardiomyocyte
106: Nodal
107: Purkinje
108: Smooth muscle
109: Vascular associated smooth muscle
110: Fibroadipogenic progenitor
111-114,117-127: Fibroblast
115: Adventitial fibroblast
116: Alveolar fibroblast
128: Myofibroblast
129: Fibromyocyte
130: Pericyte
131: Adipo-CAR
132-133: Adipocyte
134-137: Mesothelial
138: Epicardial
139: Leptomeningeal
140: Reticular
141: Unknown stromal (NGFR+ TNNT2+)
142: Unknown stromal (COL15A1+ PTGDS+)
143: Unknown stromal (ST6GAL2+)
144-146: B
147-148: Plasma
149: Innate lymphoid
150-154: Natural killer
155: Natural killer T
156-168: T
169: Hematopoietic precursor
170: Megakaryocyte progenitor
171: Platelet
172: Erythroid progenitor
173-174: Erythroid lineage
175-178: Neutrophil
179: Basophil
180: Eosinophil
181: Granulocyte monocyte progenitor
1821-83: Monocyte
184-185, 200-210: Macrophage
186: Microglial
187: Osteoclast
188-191: Alveolar macrophage
192-194: Interstitial macrophage
195-199: Kupffer
211: Conventional dendritic
212: Plasmacytoid dendritic
213: Mature dendritic
214-215,217: Dendritic
216: Langerhans
218-222: GABAergic neuron
223-226: Glutamatergic neuron
227: Cone
228: Rod
229: Horizontal
230: ON-bipolar
231: OFF-bipolar
232: Interstitial of Cajal
233: Corticotroph
234: Gonadotroph
235: Lactotroph
236: Somatotroph
237: Thyrotroph
238-241: Astrocyte
242: Oligodendrocyte precursor
243: Oligodendrocyte
244-245: Ependymal
246: Choroid plexus epithelial
247: Muller
248: Myelinating Schwann
249: Non-myelinating Schwann
250: Spermatogonium
251: Early spermatocyte
252: Pachytene spermatocyte
253: Diplotene/secondary spermatocyte
254: Early spermatid
255: Mid spermatid
256: Late spermatid

PF Proliferating cluster
H Hybrid cluster
M Mixed cluster

Cell names:

1-2: Interfollicular basal
3-6: Interfollicular suprabasal
7: Upper hair follicle
8: Inner bulge
9: Outer bulge
10: Melanocyte
11-12: Channel
13: Sweat gland
14: Granulosa
15: Unknown epithelial (CRISP3+)
16: Unknown epithelial (PGAP1+)
17-18: Respiratory basal
19: Ciliated
20: Club
21: Respiratory tuft
22: Pulmonary neuroendocrine
23: Alveolar epithelial type 1
24: Alveolar epithelial type 2
25: Uterine metastasis
26-27: Basal
28-32: Suprabasal
33: Filiform papilla
34: Enterocyte progenitor
35-36: Enterocyte
37-38: Goblet
39: Enteroendocrine
40: M of gut
41: Intestinal tuft
42-43: Hepatocyte
44: Cholangiocyte
45-47: Acinar
48-49: Ductal
50: Alpha
51: Beta
52: Delta
53: PP

Fig. 3


a
Kidney
Epithelial
Podo-cyte 54
Prox. convo-luted 55
Prox. straight 56
Loop of Henle
Macula densa 62
Dist. convo-luted 63
Collecting Duct
Urothelial 69
Desc thin 57
Asc thin 58
Asc thin SLC14A2+ 59
Asc thick 60
AKR1B1+ 61
Principal Main 64
Principal FXYD4+ 65
Principal KCNE1+ 66
Inter-calated alpha 67
Inter-calated beta 68
Inter/ Principal H5
Lymphoid
B cell
NK/T cell
B Main 144
B PF 145
Plasma Main 147
Plasma PF 148
Innate lymphoid 149
NK Main 150
NK CHN1+ 154
NK T 155
NK/T PF H7
T CD4+ CD8- CCL5- 157
T CD4- CD8+ CCL5+ 160
T CD4+ CD8+ CCL5+ 162
Endothelial
Vasa recta
Capillary
Glom-erular 91
Lymphatic
Desc SLC14A1+ 79
Desc MGP+ 80
Asc IGFBP5+ 81
Asc IGF1+ 82
FABP5- RBP7- 83
FABP5+ RBP7+ 86
CXCL10+ 87
CCL21+ 92
ACKR4- CCDC80+ 93
Stromal
Smooth musc vasc 109
Fibro-blast
Myo-fibro blast 128
Pericyte 130
Adipo-cyte
Meso-thelial 134
Reti-cular 140
Unknown ST6GAL2+ 143
Main 111
ASPN+ 124
TNC+ 125
UCPhi 132
UCPlo 133
Myeloid
Ery. prog-enitor 172
Neu-trophil
Baso-phil 179
GMP 181
Mono-cyte
Macro-phage
Dend-ritic
Main 175
PF 176
CCL13+ 177
IL18BP+ 178
Main 182
PF 183
Conv-entional 211
Plasma-cytoid 212
Mature 213
FLT3+ IGSF6+ 214
Main 184
PF 185
MRC1+ 200
MARCO+ 202
EPCAM+ 206
ABCG1+ 207
ABCG1+ PF 208
Neural
Schwann non-mye 249
b
67 - CDi alpha
68 - CDi beta
55 - PCT
56 - PST
57 - LoH thin D
H5 - CDi/CDp
69 - Urothelial
65 - CDp FXYD4+
64 - CDp
62 - MD
66 - CDp KCNE1+
59 - LoH thin A SLC14A2+
63 - DCT
58 - LoH thin A
54 - Podocyte
60 - LoH thick A
UMAP2
L4
UMAP1
Kidney - nephron epithelium
c
CUBN
SLC4A4
SLC3A1
MIOX
SLC22A8
SLC16A9
PCK1
PAH
SLC13A3
SLC34A1
GPX3
DCXR
SLC17A3
ATP11A
SLC5A2
SLC5A12
AQP1
ENPEP
SPP1
SLC12A1
UMOD
KNG1
PPP1R1A
CLDN16
WNK1
WNK4
SLC12A3
SLC8A1
LHX1
CALB1
GATA3
AVPR2
AQP2
TRPV5
FXYD4
SCNN1B
N cells 500 1000 1500
d
81 - VR A IGFBP5+
80 - VR D MGP+
82 - VR A IGF1+
79 - VR D SLC14A1+
UMAP2
L4
UMAP1
Kidney - vasa recta endothelium
e
SLC14A1
GJA5
CLDN5
SOX17
AQP1
SEMA3G
SCIN
CXCL12
CA4
NR2F2
PLVAP
VCAM1
ACKR1
N cells 100 200 300
Norm[ln(UP10K+1)]
0 1
f
L4 - Kidney medulla
L4 - Kidney cortex
g
252 - Pachytene spermatocyte
250 - Spermatogonium
251 - Early spermatocyte
253 - Diplotene/Secondary spermatocyte
254 - Early spermatid
256 - Late spermatid
255 - Mid spermatid
UMAP2
L4
UMAP1
Testis - germ cells
h
L4 - Testis seminiferous tubules
i
175 - Neutrophil
173 - Erythroid lineage
174 - Erythroid lineage (PF)
172 - Erythroid progenitor
176 - Neutrophil (PF)
170 - Megakaryocyte progenitor
181 - Granulocyte monocyte progenitor
169 - Hematopoietic precursor
197 - Basophil
H9 - Macrophage/ Monocyte (PF)
183 - Monocyte (PF)
182 - Monocyte
184/187 - Macrophage/ Osteoclast
UMAP2
L2
UMAP1
Bone & bone marrow - myeloid
j
CD34
EGFL7
FLT3
KIT
MYB
STMN1
MCM5
CCNA2
TOP2A
MKI67
ELANE
MPO
AZU1
CAMP
LTF
MMP9
CSF3R
ALPL
TYROBP
S100A8
FCN2
S100A10
[CD14]
CSF1R
AHNAK
VCAN
MSR1
C1QA
APOE
MRC1
MARCO
2500 2000 1500 1000 500 N cells 500 1000 1500 2000 2500
EGFL7
KIT
CD34
MYB
STMN1
MCM5
CCNA2
TOP2A
MKI67
[HBB]
AHSP
HBQ1
[HBA2]
ALAS2
N cells 500 1000 1500 2000

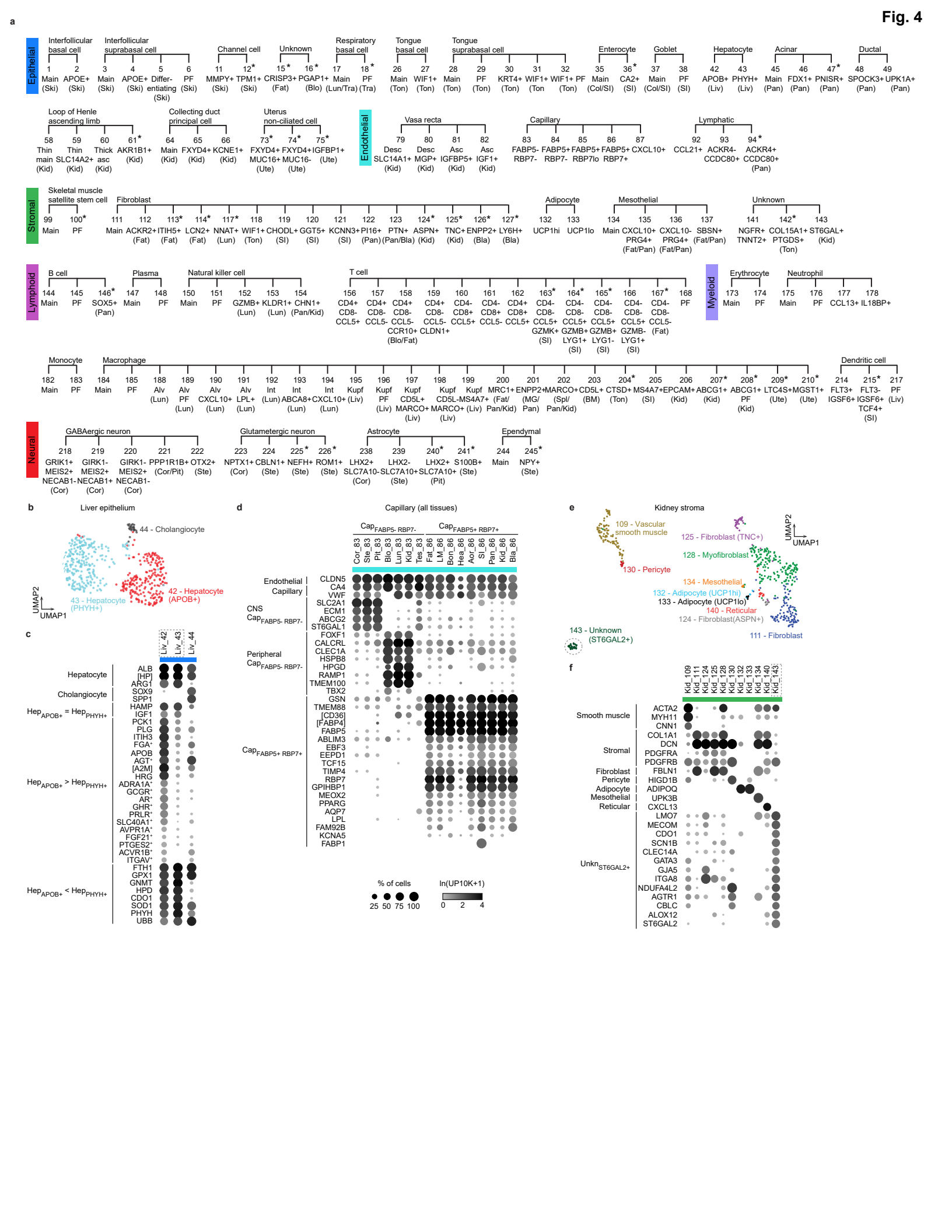


Fig. 4
a
Epithelial
Interfollicular basal cell
1 Main (Ski)
2 APOE+ (Ski)
Interfollicular suprabasal cell
3 Main (Ski)
4 APOE+ (Ski)
5 Differentiating (Ski)
6 PF (Ski)
Channel cell
11 MMPY+ (Ski)
12* TPM1+ (Ski)
Unknown
15* CRISP3+ (Fat)
16* PGAP1+ (Blo)
Respiratory basal cell
17 Main (Lun/Tra)
18* PF (Tra)
Tongue basal cell
26 Main (Ton)
27 WIF1+ (Ton)
Tongue suprabasal cell
28 Main (Ton)
29 PF (Ton)
30 KRT4+ (Ton)
31 WIF1+ (Ton
32 WIF1+ PF (Ton)
Enterocyte
35 Main (Col/SI)
36* CA2+ (SI)
Goblet
37 Main (Col/SI)
38 PF (SI)
Hepatocyte
42 APOB+ (Liv)
43 PHYH+ (Liv)
Acinar
45 Main (Pan)
46 FDX1+ (Pan)
47* PNISR+ (Pan)
Ductal
48 SPOCK3+ (Pan)
49 UPK1A+ (Pan)
Loop of Henle ascending limb
58 Thin main (Kid)
59 Thin SLC14A2+ (Kid)
60 Thick asc (Kid)
61* AKR1B1+ (Kid)
Collecting duct principal cell
64 Main (Kid)
65 FXYD4+ (Kid)
66 KCNE1+ (Kid)
Uterus non-ciliated cell
73* FXYD4+ MUC16+ (Ute)
74* FXYD4+ MUC16- (Ute)
75* IGFBP1+ (Ute)
Endothelial
Vasa recta
79 Desc SLC14A1+ (Kid)
80 Desc MGP+ (Kid)
81 Asc IGFBP5+ (Kid)
82 Asc IGF1+ (Kid)
Capillary
83 FABP5- RBP7-
84 FABP5+ RBP7-
85 FABP5+ RBP7lo
86 FABP5+ RBP7+
87 CXCL10+
Lymphatic
92 CCL21+
93 ACKR4- CCDC80+
94* ACKR4+ CCDC80+ (Pan)
Stromal
Skeletal muscle satellite stem cell
99 Main
100* PF
Fibroblast
111 Main
112 ACKR2+ (Fat)
113* ITIH5+ (Fat)
114* LCN2+ (Fat)
117* NNAT+ (Lun)
118 WIF1+ (Ton)
119 CHODL+ (SI)
120 GGT5+ (SI)
121 KCNN3+ (SI)
122 PI16+ (Pan)
123 PTN+ (Pan/Bla)
124* ASPN+ (Kid)
125* TNC+ (Kid)
126* ENPP2+ (Bla)
127* LY6H+ (Bla)
Adipocyte
132 UCP1hi
133 UCP1lo
Mesothelial
134 Main
135 CXCL10+ PRG4+ (Fat/Pan)
136 CXCL10- PRG4+ (Fat/Pan)
137 SBSN+ (Fat/Pan)
Unknown
141 NGFR+ TNNT2+
142* COL15A1+ PTGDS+ (Ton)
143 ST6GAL+ (Kid)
Lymphoid
B cell
144 Main
145 PF
146* SOX5+ (Pan)
Plasma
147 Main
148 PF
Natural killer cell
150 Main
151 PF
152 GZMB+ (Lun)
153 KLDR1+ (Lun)
154 CHN1+ (Pan/Kid)
T cell
156 CD4+ CD8- CCL5+
157 CD4+ CD8- CCL5-
158 CD4+ CD8- CCL5- CCR10+ (Blo/Fat)
159 CD4+ CD8- CCL5- CLDN1+
160 CD4- CD8+ CCL5+
161 CD4- CD8+ CCL5-
162 CD4+ CD8+ CCL5+
163* CD4- CD8- CCL5+ GZMK+ (SI)
164* CD4- CD8- CCL5+ GZMB+ LYG1+ (SI)
165* CD4- CD8- CCL5+ GZMB+ LYG1- (SI)
166 CD4- CD8- CCL5+ GZMB- LYG1+ (SI)
167* CD4- CD8- CCL5- (Fat)
168 PF
Myeloid
Erythrocyte
173 Main
174 PF
Neutrophil
175 Main
176 PF
177 CCL13+
178 IL18BP+
Monocyte
182 Main
183 PF
Macrophage
184 Main
185 PF
188 Alv (Lun)
189 Alv PF (Lun)
190 Alv CXCL10+ (Lun)
191 Alv LPL+ (Lun)
192 Int (Lun)
193 Int ABCA8+ (Lun)
194 Int CXCL10+ (Lun)
195 Kupf (Liv)
196 Kupf PF (Liv)
197 Kupf CD5L+ MARCO+ (Liv)
198 Kupf CD5L- MARCO+ (Liv)
199 Kupf MS4A7+ (Liv)
200 MRC1+ (Fat/ Pan/Kid)
201 ENPP2+ (MG/ Pan)
202 MARCO+ (Spl/ Pan/Kid)
203 CD5L+ (BM)
204* CTSD+ (Ton)
205 MS4A7+ (SI)
206 EPCAM+ (Kid)
207* ABCG1+ (Kid)
208* ABCG1+ PF (Kid)
209* LTC4S+ (Ute)
210* MGST1+ (Ute)
Dendritic cell
214 FLT3+ IGSF6+
215* FLT3- IGSF6+ TCF4+ (SI)
217 PF (Liv)
Neural
GABAergic neuron
218 GRIK1+ MEIS2+ NECAB1- (Cor)
219 GIRK1- MEIS2+ NECAB1+ (Cor)
220 GIRK1- MEIS2+ NECAB1- (Cor)
221 PPP1R1B+ (Cor/Pit)
222 OTX2+ (Ste)
Glutamatergic neuron
223 NPTX1+ (Cor)
224 CBLN1+ (Ste)
225* NEFH+ (Ste)
226* ROM1+ (Ste)
Astrocyte
238 LHX2+ SLC7A10- (Cor)
239 LHX2- SLC7A10+ (Ste)
240* LHX2+ SLC7A10+ (Pit)
241* S100B+ (Ste)
Ependymal
244 Main
245* NPY+ (Ste)
b
Liver epithelium
44 - Cholangiocyte
42 - Hepatocyte (APOB+)
43 - Hepatocyte (PHYH+)
UMAP2
UMAP1
c
Liv_42
Liv_43
Liv_44
Hepatocyte
Cholangiocyte
Hep APOB+ = Hep PHYH+
Hep APOB+ > Hep PHYH+
Hep APOB+ < Hep PHYH+
ALB
[HP]
ARG1
SOX9
SPP1
HAMP
IGF1
PCK1
PLG
ITIH3
FGA+
APOB
AGT+
[A2M]
HRG
ADRA1A+
GCGR+
AR+
GHR+
PRLR+
SLC40A1+
AVPR1A+
FGF21+
PTGES2+
ACVR1B+
ITGAV+
FTH1
GPX1
GNMT
HPD
CDO1
SOD1
PHYH
UBB
d
Capillary (all tissues)
Cap FABP5- RBP7-
Cap FABP5+ RBP7+
Cor_83
Ste_83
Pit_83
Blo_83
Lun_83
Kid_83
Tes_83
Fat_86
LM_86
Bon_86
Hea_86
Aor_86
SI_86
Pan_86
Kid_86
Bla_86
Endothelial
Capillary
CNS Cap FABP5- RBP7-
Peripheral Cap FABP5- RBP7-
Cap FABP5+ RBP7+
CLDN5
CA4
VWF
SLC2A1
ECM1
ABCG2
ST6GAL1
FOXF1
CALCRL
CLEC1A
HSPB8
HPGD
RAMP1
TMEM100
TBX2
GSN
TMEM88
[CD36]
[FABP4]
FABP5
ABLIM3
EBF3
EEPD1
TCF15
TIMP4
RBP7
GPIHBP1
MEOX2
PPARG
AQP7
LPL
FAM92B
KCNA5
FABP1
% of cells
25 50 75 100
ln(UP10K+1)
0 2 4
e
Kidney stroma
109 - Vascular smooth muscle
125 - Fibroblast (TNC+)
128 - Myofibroblast
130 - Pericyte
134 - Mesothelial
132 - Adipocyte (UCP1hi)
133 - Adipocyte (UCP1lo)
140 - Reticular
124 - Fibroblast(ASPN+)
111 - Fibroblast
143 - Unknown (ST6GAL2+)
UMAP2
UMAP1
f
Kid_109
Kid_111
Kid_124
Kid_125
Kid_128
Kid_130
Kid_132
Kid_133
Kid_134
Kid_140
Kid_143
Smooth muscle
Stromal
Fibroblast
Pericyte
Adipocyte
Mesothelial
Reticular
Unkn ST6GAL2+
ACTA2
MYH11
CNN1
COL1A1
DCN
PDGFRA
PDGFRB
FBLN1
HIGD1B
ADIPOQ
UPK3B
CXCL13
LMO7
MECOM
CDO1
SCN1B
CLEC14A
GATA3
GJA5
ITGA8
NDUFA4L2
AGTR1
CBLC
ALOX12
ST6GAL2

Fig. 5

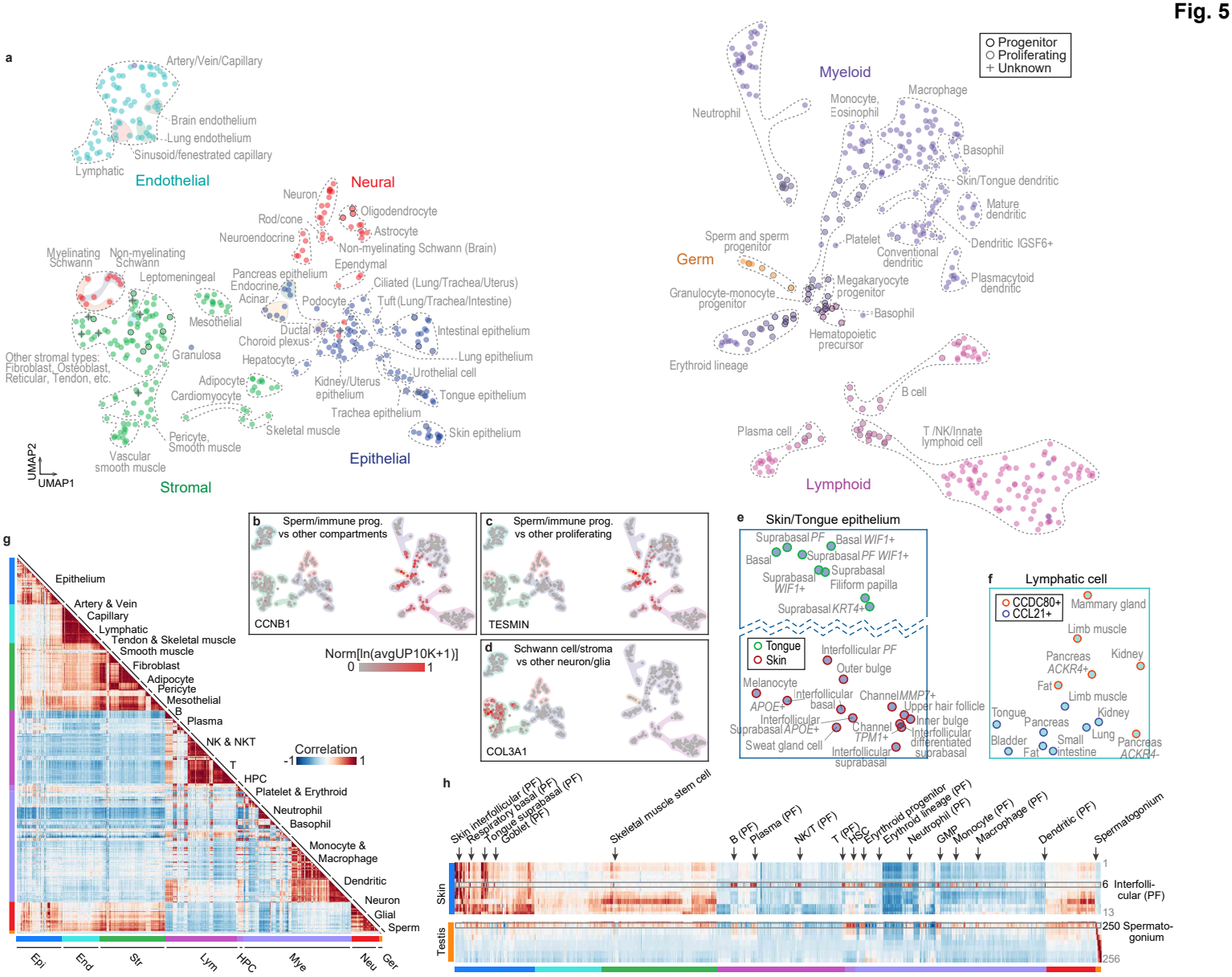
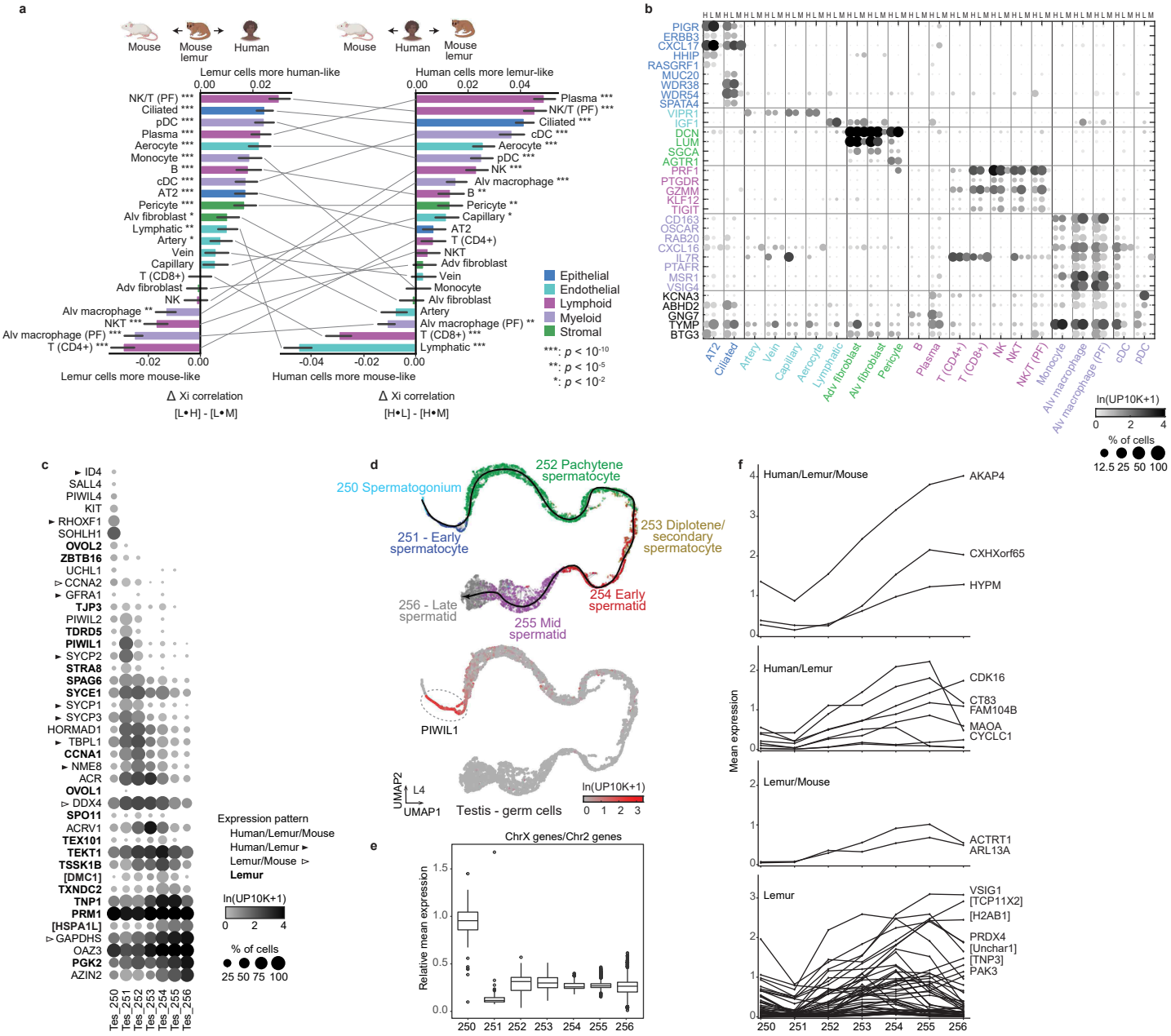

a
Artery/Vein/Capillary
Brain endothelium
Lung endothelium
Sinusoid/fenestrated capillary
Lymphatic
Endothelial
Neural
Neuron
Rod/cone
Neuroendocrine
Oligodendrocyte
Astrocyte
Non-myelinating Schwann (Brain)
Ependymal
Ciliated (Lung/Trachea/Uterus)
Tuft (Lung/Trachea/Intestine)
Myelinating Schwann
Non-myelinating Schwann
Leptomeningeal
Pancreas epithelium
Endocrine
Acinar
Podocyte
Mesothelial
Ductal
Choroid plexus
Intestinal epithelium
Lung epithelium
Urothelial cell
Granulosa
Hepatocyte
Other stromal types: Fibroblast, Osteoblast, Reticular, Tendon, etc.
Adipocyte
Cardiomyocyte
Kidney/Uterus epithelium
Tongue epithelium
Trachea epithelium
Skeletal muscle
Skin epithelium
Pericyte, Smooth muscle
Vascular smooth muscle
UMAP2
UMAP1
Stromal
Epithelial
Progenitor
Proliferating
Unknown
Myeloid
Neutrophil
Monocyte, Eosinophil
Macrophage
Basophil
Skin/Tongue dendritic
Mature dendritic
Dendritic IGSF6+
Platelet
Conventional dendritic
Plasmacytoid dendritic
Sperm and sperm progenitor
Germ
Megakaryocyte progenitor
Granulocyte-monocyte progenitor
Basophil
Hematopoietic precursor
Erythroid lineage
B cell
Plasma cell
T /NK/Innate lymphoid cell
Lymphoid
b Sperm/immune prog. vs other compartments
CCNB1
c Sperm/immune prog. vs other proliferating
TESMIN
d Schwann cell/stroma vs other neuron/glia
COL3A1
Norm[ln(avgUP10K+1)]
0
1
e Skin/Tongue epithelium
Suprabasal PF
Basal WIF1+
Basal
Suprabasal PF WIF1+
Suprabasal WIF1+
Suprabasal
Filiform papilla
Suprabasal KRT4+
Tongue
Skin
Interfollicular PF
Outer bulge
Melanocyte
Interfollicular basal
Channel MMP7+
APOE+
Upper hair follicle
Interfollicular
Inner bulge
Suprabasal APOE+
Channel TPM1+
Interfollicular differentiated suprabasal
Sweat gland cell
Interfollicular suprabasal
f Lymphatic cell
CCDC80+
CCL21+
Mammary gland
Limb muscle
Pancreas ACKR4+
Kidney
Fat
Limb muscle
Tongue
Pancreas
Kidney
Lung
Bladder
Small intestine
Fat
Pancreas ACKR4-
g
Epithelium
Artery & Vein
Capillary
Lymphatic
Tendon & Skeletal muscle
Smooth muscle
Fibroblast
Adipocyte
Pericyte
Mesothelial
B
Plasma
NK & NKT
T
HPC
Platelet & Erythroid
Neutrophil
Basophil
Monocyte & Macrophage
Dendritic
Neuron
Glial
Sperm
Correlation
-1
1
Epi
End
Str
Lym
HPC
Mye
Neu
Ger
h
Skin interfollicular (PF)
Respiratory basal (PF)
Tongue suprabasal (PF)
Goblet (PF)
Skeletal muscle stem cell
B (PF)
Plasma (PF)
NK/T (PF)
T (PF)
HSC
Erythroid progenitor
Erythroid lineage (PF)
Neutrophil (PF)
GMP
Monocyte (PF)
Macrophage (PF)
Dendritic (PF)
Spermatogonium
Skin
Testis
1
6
13
Interfollicular (PF)
250
256
Spermatogonium

Fig. 6


a
Mouse
Mouse lemur
Human
Lemur cells more human-like
Human cells more lemur-like
NK/T (PF) ***
Ciliated ***
pDC ***
Plasma ***
Aerocyte ***
Monocyte ***
B ***
cDC ***
AT2 ***
Pericyte ***
Alv fibroblast *
Lymphatic **
Artery *
Vein
Capillary
T (CD8+)
Adv fibroblast
NK
Alv macrophage **
NKT ***
Alv macrophage (PF) ***
T (CD4+) ***
Plasma ***
NK/T (PF) ***
Ciliated ***
cDC ***
Aerocyte ***
pDC ***
NK ***
Alv macrophage ***
B **
Pericyte **
Capillary *
AT2
T (CD4+)
NKT
Adv fibroblast
Vein
Monocyte
Alv fibroblast
Artery
Alv macrophage (PF) **
T (CD8+) ***
Lymphatic ***
Epithelial
Endothelial
Lymphoid
Myeloid
Stromal
***: p < 10-10
**: p < 10-5
*: p < 10-2
Lemur cells more mouse-like
Human cells more mouse-like
Δ Xi correlation
[L•H] - [L•M]
[H•L] - [H•M]
b
PIGR
ERBB3
CXCL17
HHIP
RASGRF1
MUC20
WDR38
WDR54
SPATA4
VIPR1
IGF1
DCN
LUM
SGCA
AGTR1
PRF1
PTGDR
GZMM
KLF12
TIGIT
CD163
OSCAR
RAB20
CXCL16
IL7R
PTAFR
MSR1
VSIG4
KCNA3
ABHD2
GNG7
TYMP
BTG3
AT2
Ciliated
Artery
Vein
Capillary
Aerocyte
Lymphatic
Adv fibroblast
Alv fibroblast
Pericyte
B
Plasma
T (CD4+)
T (CD8+)
NK
NKT
NK/T (PF)
Monocyte
Alv macrophage
Alv macrophage (PF)
cDC
pDC
ln(UP10K+1)
% of cells
12.5 25 50 100
c
ID4
SALL4
PIWIL4
KIT
RHOXF1
SOHLH1
OVOL2
ZBTB16
UCHL1
CCNA2
GFRA1
TJP3
PIWIL2
TDRD5
PIWIL1
SYCP2
STRA8
SPAG6
SYCE1
SYCP1
SYCP3
HORMAD1
TBPL1
CCNA1
NME8
ACR
OVOL1
DDX4
SPO11
ACRV1
TEX101
TEKT1
TSSK1B
[DMC1]
TXNDC2
TNP1
PRM1
[HSPA1L]
GAPDHS
OAZ3
PGK2
AZIN2
Tes_250
Tes_251
Tes_252
Tes_253
Tes_254
Tes_255
Tes_256
Expression pattern
Human/Lemur/Mouse
Human/Lemur
Lemur/Mouse
Lemur
ln(UP10K+1)
% of cells
25 50 75 100
d
250 Spermatogonium
252 Pachytene spermatocyte
251 - Early spermatocyte
253 Diplotene/ secondary spermatocyte
256 - Late spermatid
254 Early spermatid
255 Mid spermatid
PIWIL1
UMAP2
UMAP1
L4
Testis - germ cells
ln(UP10K+1)
e
ChrX genes/Chr2 genes
Relative mean expression
250 251 252 253 254 255 256
f
Human/Lemur/Mouse
AKAP4
CXHXorf65
HYPM
Human/Lemur
CDK16
CT83
FAM104B
MAOA
CYCLC1
Lemur/Mouse
ACTRT1
ARL13A
Lemur
VSIG1
[TCP11X2]
[H2AB1]
PRDX4
[Unchar1]
[TNP3]
PAK3
Mean expression
250 251 252 253 254 255 256

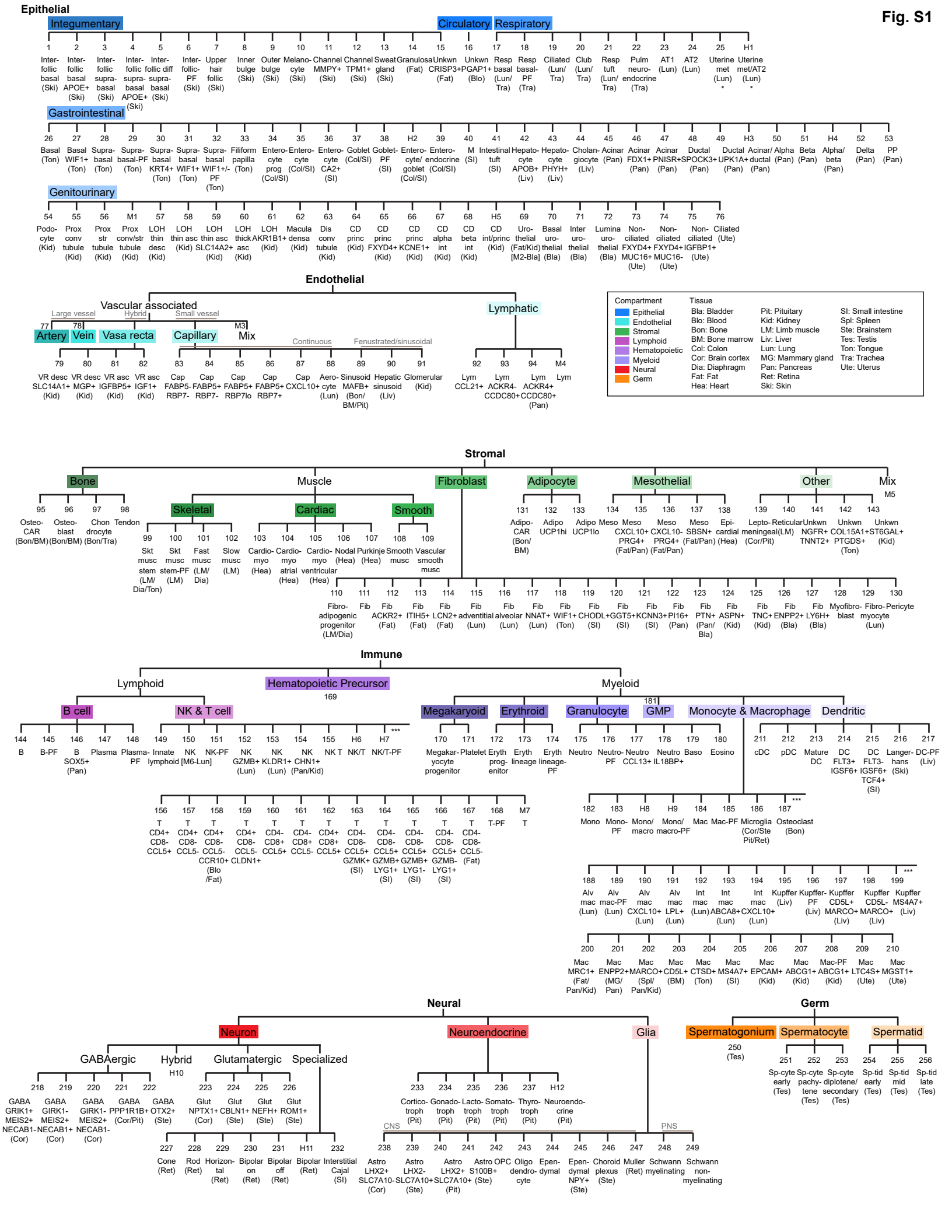


Fig. S1
Epithelial
Integumentary
1 Inter-follic basal (Ski)
2 Inter-follic basal APOE+ (Ski)
3 Inter-follic supra-basal (Ski)
4 Inter-follic supra-basal APOE+ (Ski)
5 Inter-follic diff supra-basal (Ski)
6 Inter-follic-PF (Ski)
7 Upper hair follic (Ski)
8 Inner bulge (Ski)
9 Outer bulge (Ski)
10 Melano-cyte (Ski)
11 Channel MMPY+ (Ski)
12 Channel TPM1+ (Ski)
13 Sweat gland (Ski)
14 Granulosa (Fat)
Circulatory
15 Unkwn CRISP3+ (Fat)
16 Unkwn PGAP1+ (Blo)
Respiratory
17 Resp basal (Lun/Tra)
18 Resp basal-PF (Tra)
19 Ciliated (Lun/Tra)
20 Club (Lun/Tra)
21 Resp tuft (Lun/Tra)
22 Pulm neuro-endocrine (Tra)
23 AT1 (Lun)
24 AT2 (Lun)
25 Uterine met (Lun) *
H1 Uterine met/AT2 (Lun) *
Gastrointestinal
26 Basal (Ton)
27 Basal WIF1+ (Ton)
28 Supra-basal (Ton)
29 Supra-basal-PF (Ton)
30 Supra-basal KRT4+ (Ton)
31 Supra-basal WIF1+ (Ton)
32 Supra-basal WIF1+/- PF (Ton)
33 Filiform papilla (Ton)
34 Entero-cyte prog (Col/SI)
35 Entero-cyte (Col/SI)
36 Entero-cyte CA2+ (SI)
37 Goblet (Col/SI)
38 Goblet-PF (SI)
H2 Entero-cyte/ goblet (Col/SI)
39 Entero-endocrine (Col/SI)
40 M (SI)
41 Intestinal tuft (SI)
42 Hepato-cyte APOB+ (Liv)
43 Hepato-cyte PHYH+ (Liv)
44 Cholan-giocyte (Liv)
45 Acinar (Pan)
46 Acinar FDX1+ (Pan)
47 Acinar PNISR+ (Pan)
48 Ductal SPOCK3+ (Pan)
49 Ductal UPK1A+ (Pan)
H3 Acinar/ ductal (Pan)
50 Alpha (Pan)
51 Beta (Pan)
H4 Alpha/ beta (Pan)
52 Delta (Pan)
53 PP (Pan)
Genitourinary
54 Podo-cyte (Kid)
55 Prox conv tubule (Kid)
56 Prox str tubule (Kid)
M1 Prox conv/str tubule (Kid)
57 LOH thin desc (Kid)
58 LOH thin asc (Kid)
59 LOH thin asc SLC14A2+ (Kid)
60 LOH thick asc (Kid)
61 LOH AKR1B1+ (Kid)
62 Macula densa (Kid)
63 Dis conv tubule (Kid)
64 CD princ (Kid)
65 CD princ FXYD4+ (Kid)
66 CD princ KCNE1+ (Kid)
67 CD alpha int (Kid)
68 CD beta int (Kid)
H5 CD int/princ (Kid)
69 Uro-thelial (Fat/Kid) [M2-Bla]
70 Basal uro-thelial (Bla)
71 Inter uro-thelial (Bla)
72 Lumina uro-thelial (Bla)
73 Non-ciliated FXYD4+ MUC16+ (Ute)
74 Non-ciliated FXYD4+ MUC16- (Ute)
75 Non-ciliated IGFBP1+ (Ute)
76 Ciliated (Ute)
Endothelial
Vascular associated
Large vessel
Hybrid
Small vessel
77 Artery
78 Vein
Vasa recta
Capillary
M3 Mix
Continuous
Fenustrated/sinusoidal
79 VR desc SLC14A1+ (Kid)
80 VR desc MGP+ (Kid)
81 VR asc IGFBP5+ (Kid)
82 VR asc IGF1+ (Kid)
83 Cap FABP5- RBP7-
84 Cap FABP5+ RBP7-
85 Cap FABP5+ RBP7lo
86 Cap FABP5+ RBP7+
87 Cap CXCL10+
88 Aero-cyte (Lun)
89 Sinusoid MAFB+ (Bon/ BM/Pit)
90 Hepatic sinusoid (Liv)
91 Glomerular (Kid)
Lymphatic
92 Lym CCL21+
93 Lym ACKR4- CCDC80+
94 Lym ACKR4+ CCDC80+ (Pan)
M4 Lym
Compartment
Epithelial
Endothelial
Stromal
Lymphoid
Hematopoietic
Myeloid
Neural
Germ
Tissue
Bla: Bladder
Blo: Blood
Bon: Bone
BM: Bone marrow
Col: Colon
Cor: Brain cortex
Dia: Diaphragm
Fat: Fat
Hea: Heart
Pit: Pituitary
Kid: Kidney
LM: Limb muscle
Liv: Liver
Lun: Lung
MG: Mammary gland
Pan: Pancreas
Ret: Retina
Ski: Skin
SI: Small intestine
Spl: Spleen
Ste: Brainstem
Tes: Testis
Ton: Tongue
Tra: Trachea
Ute: Uterus
Stromal
Bone
95 Osteo-CAR (Bon/BM)
96 Osteo-blast (Bon/BM)
97 Chon drocyte (Bon/Tra)
98 Tendon
Muscle
Skeletal
99 Skt musc stem (LM/ Dia/Ton)
100 Skt musc stem-PF (LM)
101 Fast musc (LM/ Dia)
102 Slow musc (LM)
Cardiac
103 Cardio-myo (Hea)
104 Cardio-myo atrial (Hea)
105 Cardio-myo ventricular (Hea)
106 Nodal (Hea)
107 Purkinje (Hea)
Smooth
108 Smooth musc
109 Vascular smooth musc
Fibroblast
110 Fibro-adipogenic progenitor (LM/Dia)
111 Fib
112 Fib ACKR2+ (Fat)
113 Fib ITIH5+ (Fat)
114 Fib LCN2+ (Fat)
115 Fib adventitial (Lun)
116 Fib alveolar (Lun)
117 Fib NNAT+ (Lun)
118 Fib WIF1+ (Ton)
119 Fib CHODL+ (SI)
120 Fib GGT5+ (SI)
121 Fib KCNN3+ (SI)
122 Fib PI16+ (Pan)
123 Fib PTN+ (Pan/ Bla)
124 Fib ASPN+ (Kid)
125 Fib TNC+ (Kid)
126 Fib ENPP2+ (Bla)
127 Fib LY6H+ (Bla)
128 Myofibro-blast
129 Fibro-myocyte (Lun)
130 Pericyte
Adipocyte
131 Adipo-CAR (Bon/ BM)
132 Adipo UCP1hi
133 Adipo UCP1lo
Mesothelial
134 Meso
135 Meso CXCL10+ PRG4+ (Fat/Pan)
136 Meso CXCL10- PRG4+ (Fat/Pan)
137 Meso SBSN+ (Fat/Pan)
138 Epi-cardial (Hea)
Other
139 Lepto-meningeal (Cor/Pit)
140 Reticular (LM)
141 Unkwn NGFR+ TNNT2+
142 Unkwn COL15A1+ PTGDS+ (Ton)
143 Unkwn ST6GAL+ (Kid)
Mix
M5
Immune
Lymphoid
B cell
144 B
145 B-PF
146 B SOX5+ (Pan)
147 Plasma
148 Plasma-PF
NK & T cell
149 Innate lymphoid
150 NK [M6-Lun]
151 NK-PF
152 NK GZMB+ (Lun)
153 NK KLDR1+ (Lun)
154 NK CHN1+ (Pan/Kid)
155 NK T
H6 NK/T
H7 NK/T-PF
***
Hematopoietic Precursor
169
Myeloid
Megakaryoid
170 Megakar-yocyte progenitor
171 Platelet
Erythroid
172 Eryth prog-enitor
173 Eryth lineage
174 Eryth lineage-PF
Granulocyte
175 Neutro
176 Neutro-PF
177 Neutro CCL13+
178 Neutro IL18BP+
179 Baso
180 Eosino
181 GMP
Monocyte & Macrophage
182 Mono
183 Mono-PF
H8 Mono/ macro
H9 Mono/ macro-PF
184 Mac
185 Mac-PF
186 Microglia (Cor/Ste Pit/Ret)
187 Osteoclast (Bon)
***
Dendritic
211 cDC
212 pDC
213 Mature DC
214 DC FLT3+ IGSF6+
215 DC FLT3- IGSF6+ TCF4+ (SI)
216 Langer-hans (Ski)
217 DC-PF (Liv)
156 T CD4+ CD8- CCL5+
157 T CD4+ CD8- CCL5-
158 T CD4+ CD8- CCL5- CCR10+ (Blo /Fat)
159 T CD4+ CD8- CCL5- CLDN1+
160 T CD4- CD8+ CCL5+
161 T CD4- CD8+ CCL5-
162 T CD4+ CD8+ CCL5+
163 T CD4- CD8- CCL5+ GZMK+ (SI)
164 T CD4- CD8- CCL5+ GZMB+ LYG1+ (SI)
165 T CD4- CD8- CCL5+ GZMB+ LYG1- (SI)
166 T CD4- CD8- CCL5+ GZMB- LYG1+ (SI)
167 T CD4- CD8- CCL5- (Fat)
168 T-PF
M7 T
188 Alv mac (Lun)
189 Alv mac-PF (Lun)
190 Alv mac CXCL10+ (Lun)
191 Alv mac LPL+ (Lun)
192 Int mac (Lun)
193 Int mac ABCA8+ (Lun)
194 Int mac CXCL10+ (Lun)
195 Kupffer (Liv)
196 Kupffer-PF (Liv)
197 Kupffer CD5L+ MARCO+ (Liv)
198 Kupffer CD5L- MARCO+ (Liv)
199 Kupffer MS4A7+ (Liv)
***
200 Mac MRC1+ (Fat/ Pan/Kid)
201 Mac ENPP2+ (MG/ Pan)
202 Mac MARCO+ (Spl/ Pan/Kid)
203 Mac CD5L+ (BM)
204 Mac CTSD+ (Ton)
205 Mac MS4A7+ (SI)
206 Mac EPCAM+ (Kid)
207 Mac ABCG1+ (Kid)
208 Mac-PF ABCG1+ (Kid)
209 Mac LTC4S+ (Ute)
210 Mac MGST1+ (Ute)
Neural
Neuron
GABAergic
218 GABA GRIK1+ MEIS2+ NECAB1- (Cor)
219 GABA GIRK1- MEIS2+ NECAB1+ (Cor)
220 GABA GIRK1- MEIS2+ NECAB1- (Cor)
221 GABA PPP1R1B+ (Cor/Pit)
222 GABA OTX2+ (Ste)
Hybrid
H10
Glutamatergic
223 Glut NPTX1+ (Cor)
224 Glut CBLN1+ (Ste)
225 Glut NEFH+ (Ste)
226 Glut ROM1+ (Ste)
Specialized
227 Cone (Ret)
228 Rod (Ret)
229 Horizon-tal (Ret)
230 Bipolar on (Ret)
231 Bipolar off (Ret)
H11 Bipolar (Ret)
232 Interstitial Cajal (SI)
Neuroendocrine
233 Cortico-troph (Pit)
234 Gonado-troph (Pit)
235 Lacto-troph (Pit)
236 Somato-troph (Pit)
237 Thyro-troph (Pit)
H12 Neuroendo-crine (Pit)
Glia
CNS
PNS
238 Astro LHX2+ SLC7A10- (Cor)
239 Astro LHX2- SLC7A10+ (Ste)
240 Astro LHX2+ SLC7A10+ (Pit)
241 Astro S100B+ (Ste)
242 OPC
243 Oligo dendro-cyte
244 Epen-dymal
245 Epen-dymal NPY+ (Ste)
246 Choroid plexus (Ste)
247 Muller (Ret)
248 Schwann myelinating
249 Schwann non-myelinating
Germ
Spermatogonium
250 (Tes)
Spermatocyte
251 Sp-cyte early (Tes)
252 Sp-cyte pachy-tene (Tes)
253 Sp-cyte diplotene/ secondary (Tes)
Spermatid
254 Sp-tid early (Tes)
255 Sp-tid mid (Tes)
256 Sp-tid late (Tes)

Fig. S2

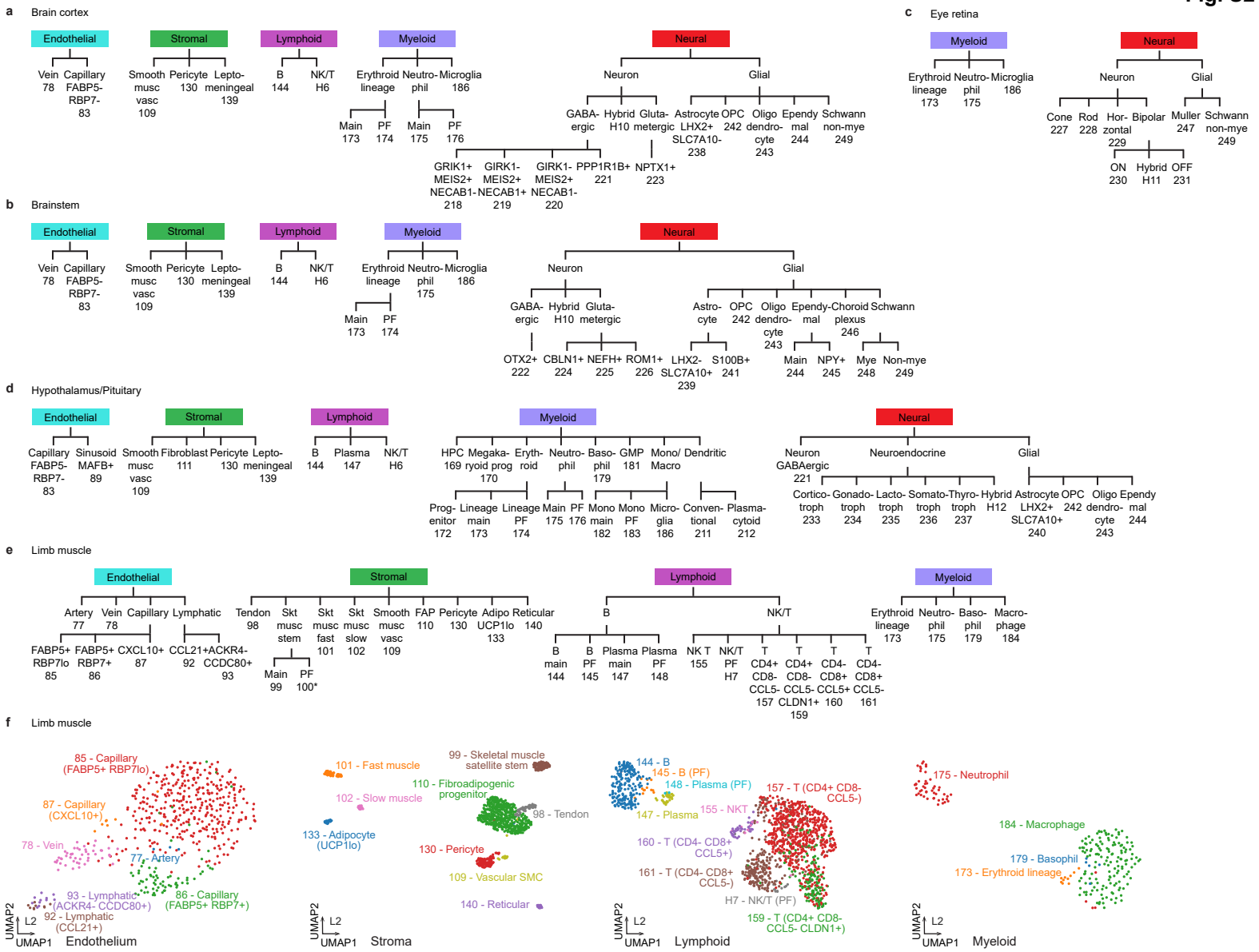

a
Brain cortex
Endothelial
Vein 78
Capillary FABP5- RBP7- 83
Stromal
Smooth musc vasc 109
Pericyte 130
Lepto-meningeal 139
Lymphoid
B 144
NK/T H6
Myeloid
Erythroid lineage
Main 173
PF 174
Neutro-phil
Main 175
PF 176
Microglia 186
Neural
Neuron
GABA-ergic
GRIK1+ MEIS2+ NECAB1- 218
GIRK1- MEIS2+ NECAB1+ 219
GIRK1- MEIS2+ NECAB1- 220
Hybrid H10
Gluta-metergic
PPP1R1B+ 221
NPTX1+ 223
Glial
Astrocyte LHX2+ SLC7A10- 238
OPC 242
Oligo dendro-cyte 243
Ependy-mal 244
Schwann non-mye 249
b
Brainstem
Endothelial
Vein 78
Capillary FABP5- RBP7- 83
Stromal
Smooth musc vasc 109
Pericyte 130
Lepto-meningeal 139
Lymphoid
B 144
NK/T H6
Myeloid
Erythroid lineage
Main 173
PF 174
Neutro-phil 175
Microglia 186
Neural
Neuron
GABA-ergic
OTX2+ 222
Hybrid H10
Gluta-metergic
CBLN1+ 224
NEFH+ 225
ROM1+ 226
Glial
Astro-cyte
LHX2- SLC7A10+ 239
S100B+ 241
OPC 242
Oligo dendro-cyte 243
Ependy-mal
Main 244
NPY+ 245
Choroid plexus 246
Mye 248
Non-mye 249
Schwann
c
Eye retina
Myeloid
Erythroid lineage 173
Neutro-phil 175
Microglia 186
Neural
Neuron
Cone 227
Rod 228
Hor-zontal 229
Bipolar
ON 230
Hybrid H11
OFF 231
Glial
Muller 247
Schwann non-mye 249
d
Hypothalamus/Pituitary
Endothelial
Capillary FABP5- RBP7- 83
Sinusoid MAFB+ 89
Stromal
Smooth musc vasc 109
Fibroblast 111
Pericyte 130
Lepto-meningeal 139
Lymphoid
B 144
Plasma 147
NK/T H6
Myeloid
HPC 169
Megaka-ryoid prog 170
Eryth-roid
Neutro-phil
Main 175
PF 176
Baso-phil 179
GMP 181
Mono/ Macro
Mono main 182
Mono PF 183
Dendritic
Prog-enitor 172
Lineage main 173
Lineage PF 174
Micro-glia 186
Conven-tional 211
Plasma-cytoid 212
Neural
Neuron GABAergic 221
Neuroendocrine
Cortico-troph 233
Gonado-troph 234
Lacto-troph 235
Somato-troph 236
Thyro-troph 237
Hybrid H12
Glial
Astrocyte LHX2+ SLC7A10+ 240
OPC 242
Oligo dendro-cyte 243
Ependy-mal 244
e
Limb muscle
Endothelial
Artery 77
Vein 78
Capillary
FABP5+ RBP7lo 85
FABP5+ RBP7+ 86
CXCL10+ 87
Lymphatic
CCL21+ 92
ACKR4- CCDC80+ 93
Stromal
Tendon 98
Skt musc stem
Main 99
PF 100*
Skt musc fast 101
Skt musc slow 102
Smooth musc vasc 109
FAP 110
Pericyte 130
Adipo UCP1lo 133
Reticular 140
Lymphoid
B
B main 144
B PF 145
Plasma main 147
Plasma PF 148
NK/T
NK T 155
NK/T PF H7
T CD4+ CD8- CCL5- 157
T CD4+ CD8- CCL5- CLDN1+ 159
T CD4- CD8+ CCL5+ 160
T CD4- CD8+ CCL5- 161
Myeloid
Erythroid lineage 173
Neutro-phil 175
Baso-phil 179
Macro-phage 184
f
Limb muscle
85 - Capillary (FABP5+ RBP7lo)
87 - Capillary (CXCL10+)
78 - Vein
77 - Artery
93 - Lymphatic (ACKR4- CCDC80+)
92 - Lymphatic (CCL21+)
86 - Capillary (FABP5+ RBP7+)
UMAP2
L2
UMAP1
Endothelium
101 - Fast muscle
102 - Slow muscle
133 - Adipocyte (UCP1lo)
99 - Skeletal muscle satellite stem
110 - Fibroadipogenic progenitor
98 - Tendon
130 - Pericyte
109 - Vascular SMC
140 - Reticular
UMAP2
L2
UMAP1
Stroma
144 - B
145 - B (PF)
148 - Plasma (PF)
147 - Plasma
155 - NKT
157 - T (CD4+ CD8- CCL5-)
160 - T (CD4- CD8+ CCL5+)
161 - T (CD4- CD8+ CCL5-)
H7 - NK/T (PF)
159 - T (CD4+ CD8- CCL5- CLDN1+)
UMAP2
L2
UMAP1
Lymphoid
175 - Neutrophil
184 - Macrophage
179 - Basophil
173 - Erythroid lineage
UMAP2
L2
UMAP1
Myeloid

Fig. S3

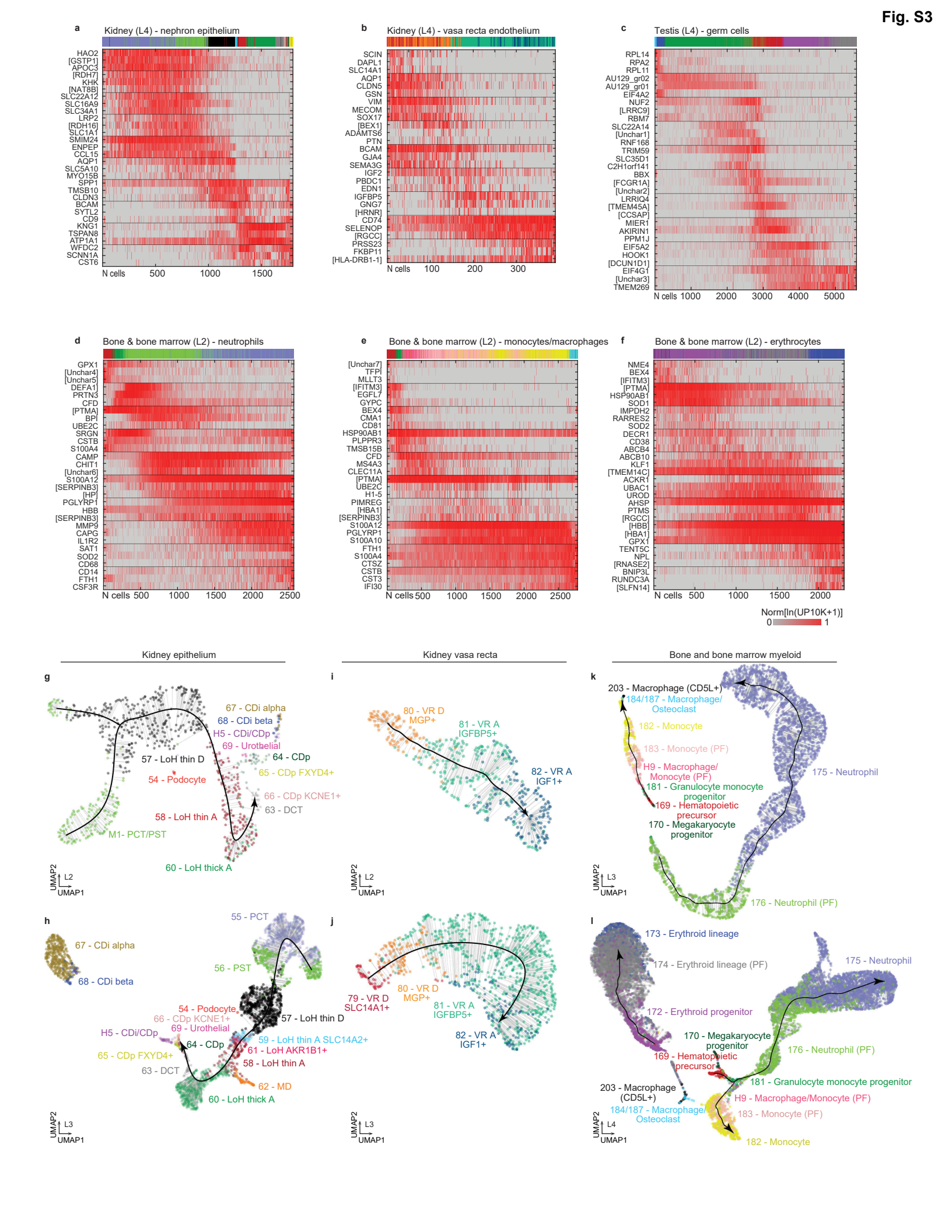
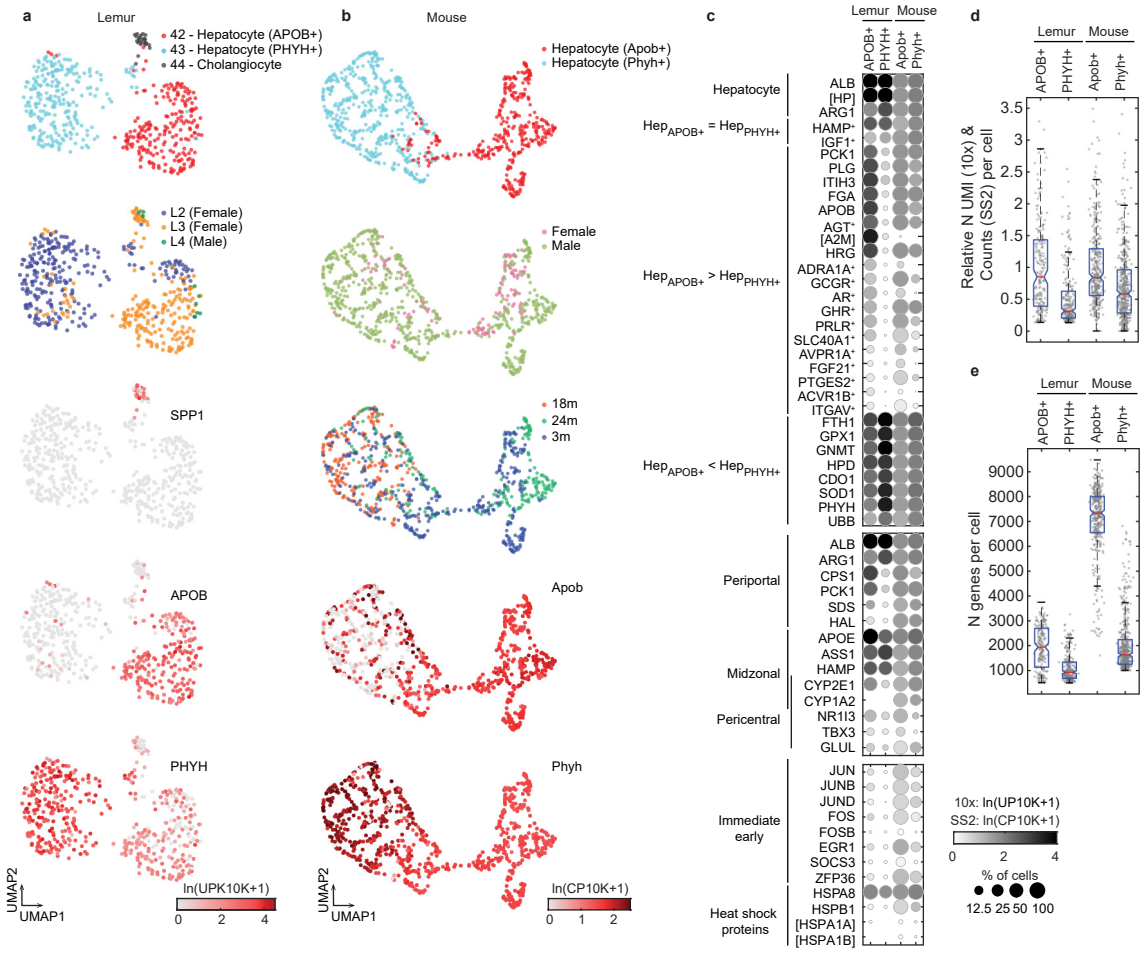

a
Kidney (L4) - nephron epithelium
HAO2
[GSTP1]
APOC3
[RDH7]
KHK
[NAT8B]
SLC22A12
SLC16A9
SLC34A1
LRP2
[RDH16]
SLC1A1
SMIM24
ENPEP
CCL15
AQP1
SLC5A10
MYO15B
SPP1
TMSB10
CLDN3
BCAM
SYTL2
CD9
KNG1
TSPAN8
ATP1A1
WFDC2
SCNN1A
CST6
N cells 500 1000 1500
b
Kidney (L4) - vasa recta endothelium
SCIN
DAPL1
SLC14A1
AQP1
CLDN5
GSN
VIM
MECOM
SOX17
[BEX1]
ADAMTS6
PTN
BCAM
GJA4
SEMA3G
IGF2
PBDC1
EDN1
IGFBP5
GNG7
[HRNR]
CD74
SELENOP
[RGCC]
PRSS23
FKBP11
[HLA-DRB1-1]
N cells 100 200 300
c
Testis (L4) - germ cells
RPL14
RPA2
RPL11
AU129_gr02
AU129_gr01
EIF4A2
NUF2
[LRRC9]
RBM7
SLC22A14
[Unchar1]
RNF168
TRIM59
SLC35D1
C2H1orf141
BBX
[FCGR1A]
[Unchar2]
LRRIQ4
[TMEM45A]
[CCSAP]
MIER1
AKIRIN1
PPM1J
EIF5A2
HOOK1
[DCUN1D1]
EIF4G1
[Unchar3]
TMEM269
N cells 1000 2000 3000 4000 5000
d
Bone & bone marrow (L2) - neutrophils
GPX1
[Unchar4]
[Unchar5]
DEFA1
PRTN3
CFD
[PTMA]
BPI
UBE2C
SRGN
CSTB
S100A4
CAMP
CHIT1
[Unchar6]
S100A12
[SERPINB3]
[HP]
PGLYRP1
HBB
[SERPINB3]
MMP9
CAPG
IL1R2
SAT1
SOD2
CD68
CD14
FTH1
CSF3R
N cells 500 1000 1500 2000 2500
e
Bone & bone marrow (L2) - monocytes/macrophages
[Unchar7]
TFPI
MLLT3
[IFITM3]
EGFL7
GYPC
BEX4
CMA1
CD81
HSP90AB1
PLPPR3
TMSB15B
CFD
MS4A3
CLEC11A
[PTMA]
UBE2C
H1-5
PIMREG
[HBA1]
[SERPINB3]
S100A12
PGLYRP1
S100A10
FTH1
S100A4
CTSZ
CSTB
CST3
IFI30
N cells 500 1000 1500 2000 2500
f
Bone & bone marrow (L2) - erythrocytes
NME4
BEX4
[IFITM3]
[PTMA]
HSP90AB1
SOD1
IMPDH2
RARRES2
SOD2
DECR1
CD38
ABCB4
ABCB10
KLF1
[TMEM14C]
ACKR1
UBAC1
UROD
AHSP
PTMS
[RGCC]
[HBB]
[HBA1]
GPX1
TENT5C
NPL
[RNASE2]
BNIP3L
RUNDC3A
[SLFN14]
N cells 500 1000 1500 2000
Norm[ln(UP10K+1)]
0 1
Kidney epithelium
g
67 - CDi alpha
68 - CDi beta
H5 - CDi/CDp
69 - Urothelial
64 - CDp
65 - CDp FXYD4+
66 - CDp KCNE1+
63 - DCT
57 - LoH thin D
54 - Podocyte
58 - LoH thin A
M1- PCT/PST
60 - LoH thick A
UMAP2
L2
UMAP1
Kidney vasa recta
i
80 - VR D MGP+
81 - VR A IGFBP5+
82 - VR A IGF1+
UMAP2
L2
UMAP1
Bone and bone marrow myeloid
k
203 - Macrophage (CD5L+)
184/187 - Macrophage/ Osteoclast
182 - Monocyte
183 - Monocyte (PF)
H9 - Macrophage/ Monocyte (PF)
181 - Granulocyte monocyte progenitor
169 - Hematopoietic precursor
170 - Megakaryocyte progenitor
175 - Neutrophil
176 - Neutrophil (PF)
UMAP2
L3
UMAP1
h
55 - PCT
67 - CDi alpha
56 - PST
68 - CDi beta
54 - Podocyte
66 - CDp KCNE1+
69 - Urothelial
57 - LoH thin D
H5 - CDi/CDp
64 - CDp
59 - LoH thin A SLC14A2+
65 - CDp FXYD4+
61 - LoH AKR1B1+
58 - LoH thin A
63 - DCT
62 - MD
60 - LoH thick A
UMAP2
L3
UMAP1
j
79 - VR D SLC14A1+
80 - VR D MGP+
81 - VR A IGFBP5+
82 - VR A IGF1+
UMAP2
L3
UMAP1
l
173 - Erythroid lineage
174 - Erythroid lineage (PF)
175 - Neutrophil
172 - Erythroid progenitor
170 - Megakaryocyte progenitor
176 - Neutrophil (PF)
169 - Hematopoietic precursor
181 - Granulocyte monocyte progenitor
203 - Macrophage (CD5L+)
H9 - Macrophage/Monocyte (PF)
184/187 - Macrophage/ Osteoclast
183 - Monocyte (PF)
182 - Monocyte
UMAP2
L4
UMAP1

Fig. S4


a
Lemur
42 - Hepatocyte (APOB+)
43 - Hepatocyte (PHYH+)
44 - Cholangiocyte
L2 (Female)
L3 (Female)
L4 (Male)
SPP1
APOB
PHYH
UMAP2
UMAP1
ln(UPK10K+1)
0 2 4
b
Mouse
Hepatocyte (Apob+)
Hepatocyte (Phyh+)
Female
Male
18m
24m
3m
Apob
Phyh
ln(CP10K+1)
0 1 2
c
Lemur Mouse
APOB+ PHYH+ Apob+ Phyh+
Hepatocyte
Hep$_{APOB+}$ = Hep$_{PHYH+}$
Hep$_{APOB+}$ > Hep$_{PHYH+}$
Hep$_{APOB+}$ < Hep$_{PHYH+}$
ALB
[HP]
ARG1
HAMP
IGF1
PCK1
PLG
ITIH3
FGA
APOB
AGT
[A2M]
HRG
ADRA1A
GCGR
AR
GHR
PRLR
SLC40A1
AVPR1A
FGF21
PTGES2
ACVR1B
ITGAV
FTH1
GPX1
GNMT
HPD
CDO1
SOD1
PHYH
UBB
Periportal
ALB
ARG1
CPS1
PCK1
SDS
HAL
Midzonal
APOE
ASS1
HAMP
CYP2E1
Pericentral
CYP1A2
NR1I3
TBX3
GLUL
Immediate early
JUN
JUNB
JUND
FOS
FOSB
EGR1
SOCS3
ZFP36
Heat shock proteins
HSPA8
HSPB1
[HSPA1A]
[HSPA1B]
d
Lemur Mouse
APOB+ PHYH+ Apob+ Phyh+
Relative N UMI (10x) & Counts (SS2) per cell
0 0.5 1 1.5 2 2.5 3 3.5
e
Lemur Mouse
APOB+ PHYH+ Apob+ Phyh+
N genes per cell
1000 2000 3000 4000 5000 6000 7000 8000 9000
10x: ln(UP10K+1)
SS2: ln(CP10K+1)
0 2 4
% of cells
12.5 25 50 100

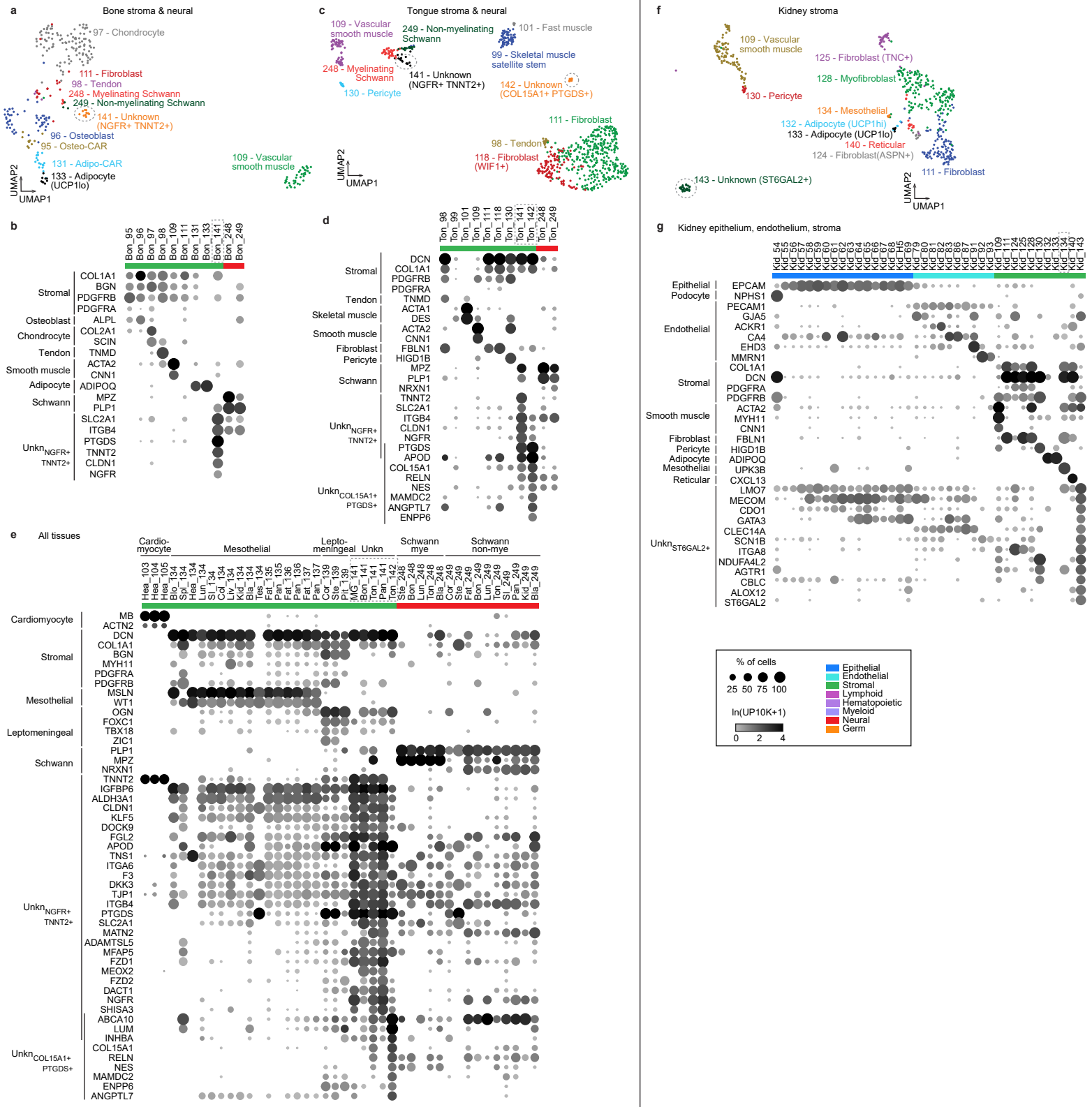

a
Bone stroma & neural
97 - Chondrocyte
111 - Fibroblast
98 - Tendon
248 - Myelinating Schwann
249 - Non-myelinating Schwann
141 - Unknown (NGFR+ TNNT2+)
96 - Osteoblast
95 - Osteo-CAR
131 - Adipo-CAR
133 - Adipocyte (UCP1lo)
109 - Vascular smooth muscle
UMAP2
UMAP1
c
Tongue stroma & neural
109 - Vascular smooth muscle
249 - Non-myelinating Schwann
101 - Fast muscle
99 - Skeletal muscle satellite stem
248 - Myelinating Schwann
141 - Unknown (NGFR+ TNNT2+)
142 - Unknown (COL15A1+ PTGDS+)
130 - Pericyte
111 - Fibroblast
98 - Tendon
118 - Fibroblast (WIF1+)
UMAP2
UMAP1
f
Kidney stroma
109 - Vascular smooth muscle
125 - Fibroblast (TNC+)
128 - Myofibroblast
130 - Pericyte
134 - Mesothelial
132 - Adipocyte (UCP1hi)
133 - Adipocyte (UCP1lo)
140 - Reticular
124 - Fibroblast(ASPN+)
111 - Fibroblast
143 - Unknown (ST6GAL2+)
UMAP2
UMAP1
b
Bon_95 Bon_96 Bon_97 Bon_98 Bon_109 Bon_111 Bon_131 Bon_133 Bon_141 Bon_248 Bon_249
Stromal: COL1A1, BGN, PDGFRB, PDGFRA
Osteoblast: ALPL
Chondrocyte: COL2A1, SCIN
Tendon: TNMD
Smooth muscle: ACTA2, CNN1
Adipocyte: ADIPOQ
Schwann: MPZ, PLP1
Unkn NGFR+ TNNT2+: SLC2A1, ITGB4, PTGDS, TNNT2, CLDN1, NGFR
d
Ton_98 Ton_99 Ton_101 Ton_109 Ton_111 Ton_118 Ton_130 Ton_141 Ton_142 Ton_248 Ton_249
Stromal: DCN, COL1A1, PDGFRB, PDGFRA
Tendon: TNMD
Skeletal muscle: ACTA1, DES
Smooth muscle: ACTA2, CNN1
Fibroblast: FBLN1
Pericyte: HIGD1B
Schwann: MPZ, PLP1, NRXN1
Unkn NGFR+ TNNT2+: TNNT2, SLC2A1, ITGB4, CLDN1, NGFR, PTGDS
Unkn COL15A1+ PTGDS+: APOD, COL15A1, RELN, NES, MAMDC2, ANGPTL7, ENPP6
e
All tissues
Cardio-myocyte
Mesothelial
Lepto-meningeal
Unkn
Schwann mye
Schwann non-mye
Hea_103 Hea_104 Hea_105 Blo_134 Spl_134 Hea_134 Lun_134 Sl_134 Col_134 Liv_134 Kid_134 Bla_134 Tes_134 Fat_135 Fat_136 Fat_137 Pan_136 Pan_137 Cor_139 Ste_139 Pit_139 MG_141 Bon_141 Ton_141 Pan_141 Ton_142 Ste_248 Bon_248 Lun_248 Ton_248 Bla_248 Cor_249 Ste_249 Fat_249 Bon_249 Lun_249 Ton_249 Sl_249 Pan_249 Kid_249 Bla_249
Cardiomyocyte: MB, ACTN2
Stromal: DCN, COL1A1, BGN, MYH11, PDGFRA, PDGFRB
Mesothelial: MSLN, WT1
Leptomeningeal: OGN, FOXC1, TBX18, ZIC1
Schwann: PLP1, MPZ, NRXN1
Unkn NGFR+ TNNT2+: TNNT2, IGFBP6, ALDH3A1, CLDN1, KLF5, DOCK9, FGL2, APOD, TNS1, ITGA6, F3, DKK3, TJP1, ITGB4, PTGDS, SLC2A1, MATN2, ADAMTSL5, MFAP5, FZD1, MEOX2, FZD2, DACT1, NGFR, SHISA3
Unkn COL15A1+ PTGDS+: ABCA10, LUM, INHBA, COL15A1, RELN, NES, MAMDC2, ENPP6, ANGPTL7
g
Kidney epithelium, endothelium, stroma
Kid_54 Kid_55 Kid_56 Kid_57 Kid_58 Kid_59 Kid_60 Kid_61 Kid_62 Kid_63 Kid_64 Kid_65 Kid_66 Kid_67 Kid_68 Kid_H5 Kid_69 Kid_79 Kid_80 Kid_81 Kid_82 Kid_83 Kid_86 Kid_87 Kid_91 Kid_92 Kid_93 Kid_109 Kid_111 Kid_124 Kid_125 Kid_128 Kid_130 Kid_132 Kid_133 Kid_134 Kid_140 Kid_143
Epithelial Podocyte: EPCAM, NPHS1
Endothelial: PECAM1, GJA5, ACKR1, CA4, EHD3, MMRN1
Stromal: COL1A1, DCN, PDGFRA, PDGFRB
Smooth muscle: ACTA2, MYH11, CNN1
Fibroblast: FBLN1
Pericyte: HIGD1B
Adipocyte: ADIPOQ
Mesothelial: UPK3B
Reticular: CXCL13
Unkn ST6GAL2+: LMO7, MECOM, CDO1, GATA3, CLEC14A, SCN1B, ITGA8, NDUFA4L2, AGTR1, CBLC, ALOX12, ST6GAL2
% of cells
25 50 75 100
ln(UP10K+1)
0 2 4
Epithelial
Endothelial
Stromal
Lymphoid
Hematopoietic
Myeloid
Neural
Germ

Fig. S6

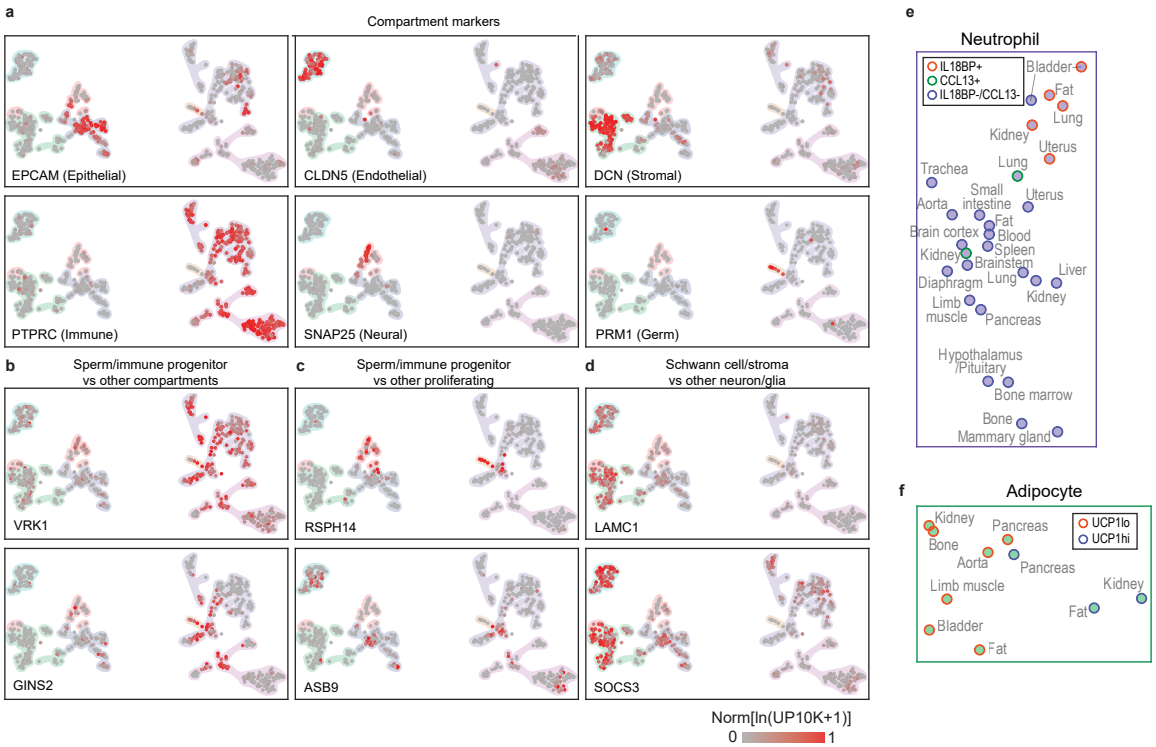
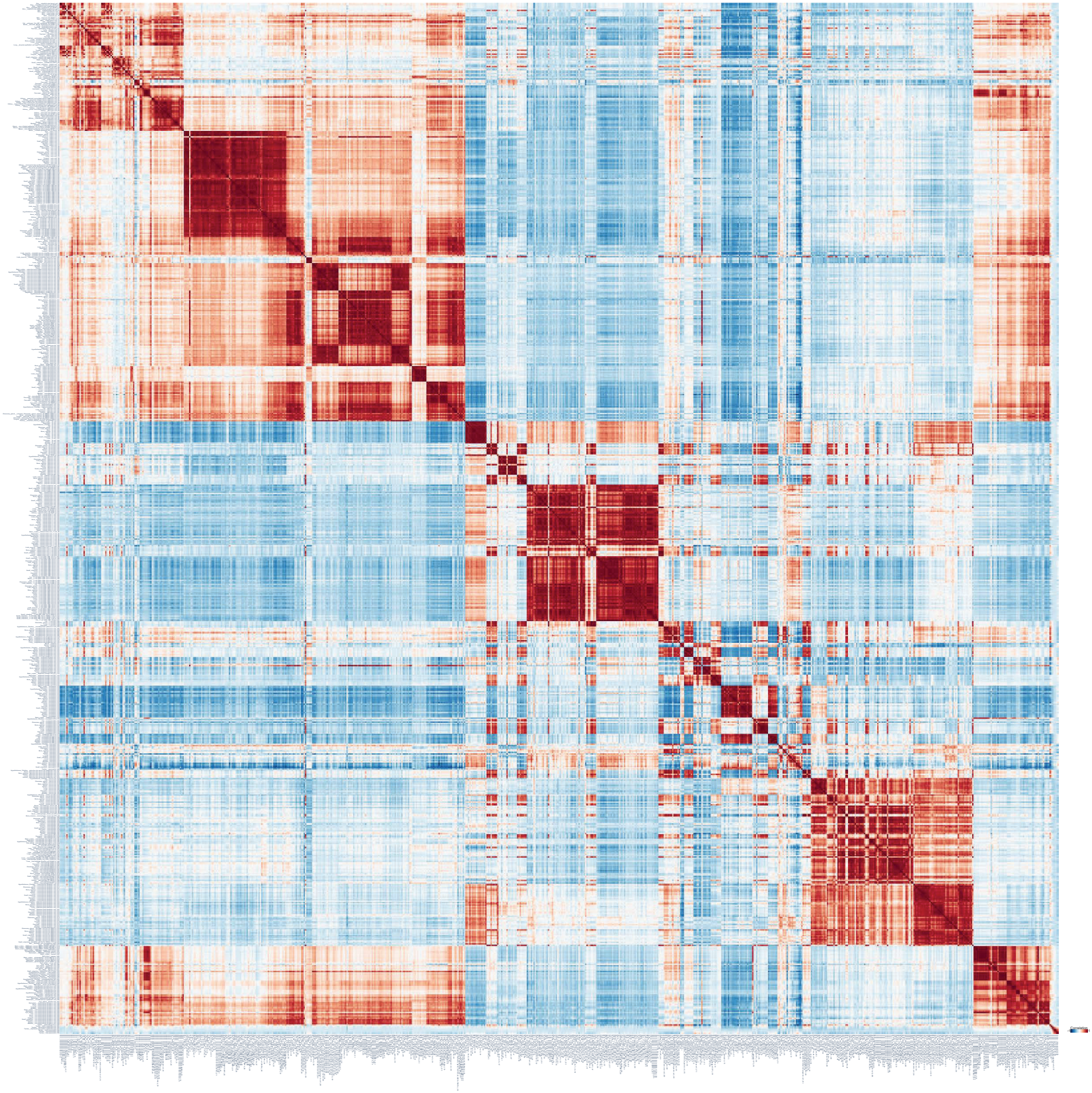

a
Compartment markers
EPCAM (Epithelial)
CLDN5 (Endothelial)
DCN (Stromal)
PTPRC (Immune)
SNAP25 (Neural)
PRM1 (Germ)
b
Sperm/immune progenitor vs other compartments
VRK1
GINS2
c
Sperm/immune progenitor vs other proliferating
RSPH14
ASB9
d
Schwann cell/stroma vs other neuron/glia
LAMC1
SOCS3
Norm[ln(UP10K+1)]
0
1
e
Neutrophil
IL18BP+
CCL13+
IL18BP-/CCL13-
Bladder
Fat
Lung
Kidney
Uterus
Lung
Trachea
Small intestine
Uterus
Aorta
Fat
Brain cortex
Blood
Spleen
Kidney
Brainstem
Liver
Diaphragm
Lung
Kidney
Limb muscle
Pancreas
Hypothalamus /Pituitary
Bone marrow
Bone
Mammary gland
f
Adipocyte
UCP1lo
UCP1hi
Kidney
Pancreas
Bone
Aorta
Pancreas
Limb muscle
Kidney
Fat
Bladder
Fat

Fig. S7

a. Correlation heatmap A - ordered by compartment, cell type, tissue



b. Correlation heatmap B - ordered by tissue, compartment, cell type

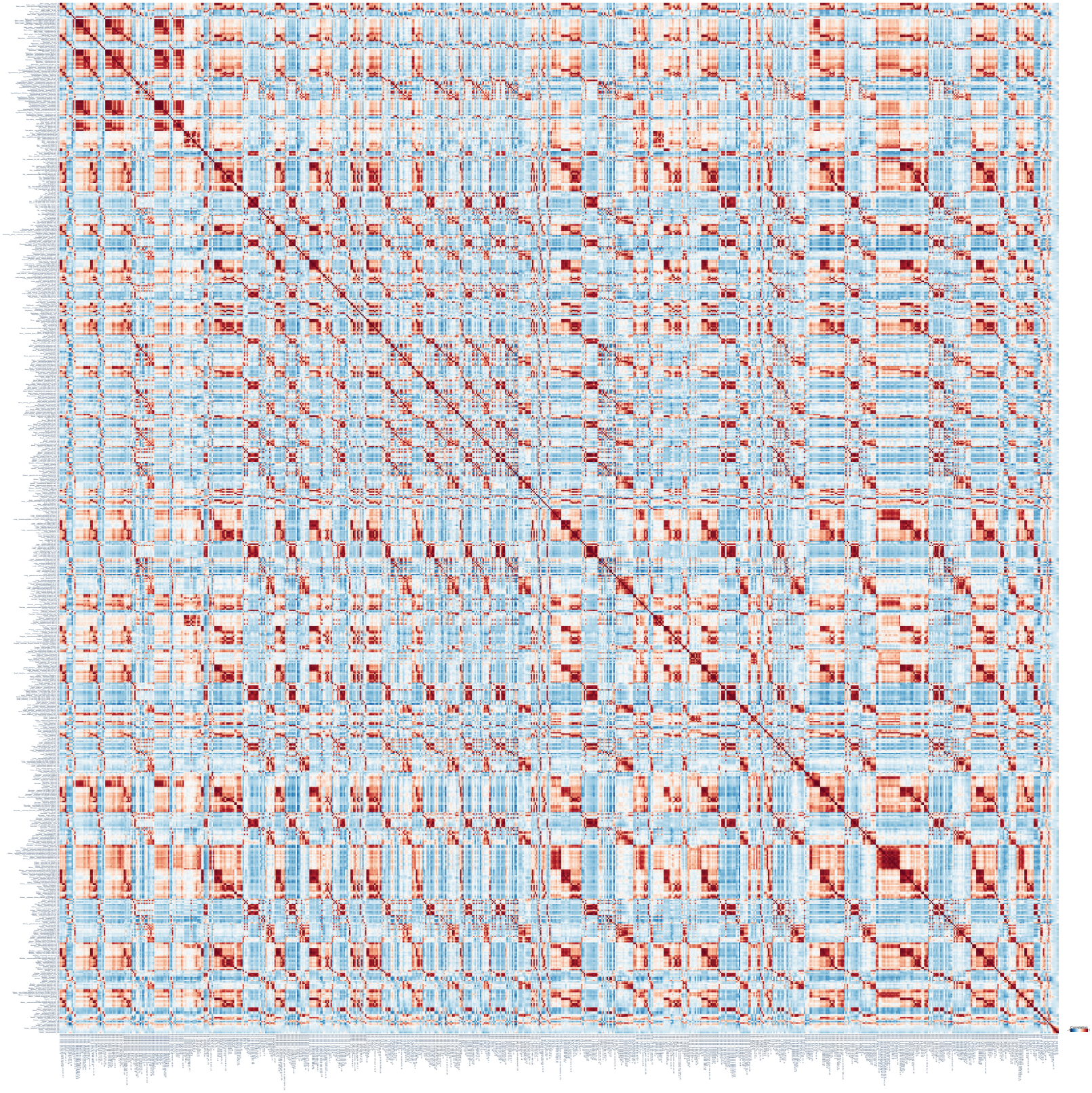


Fig. S7 - continued

c. Correlation heatmap C - ordered by hierarchical clustering

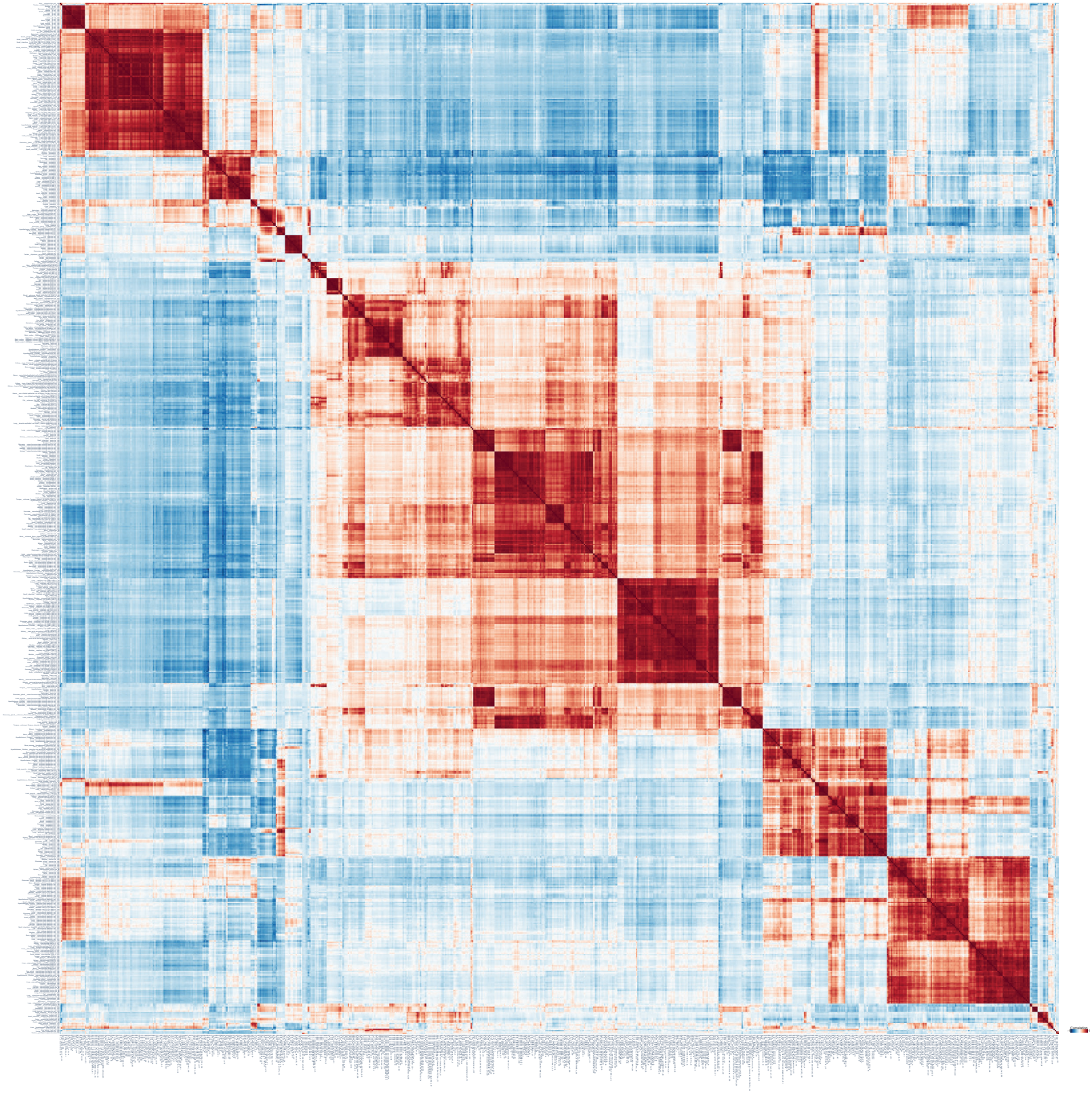
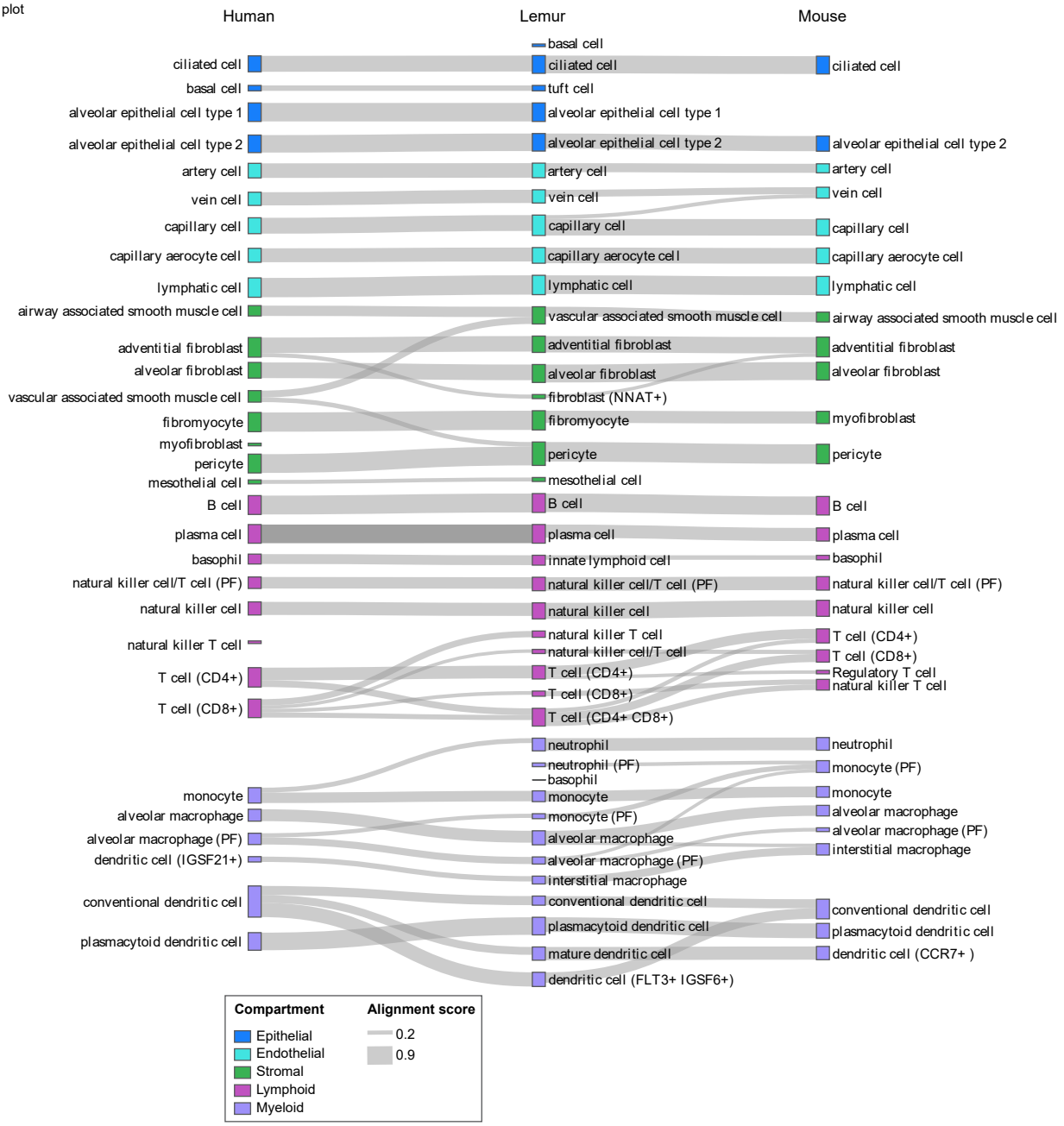


Fig. S8

a Lung Sankey plot



b Lung BBKNN-corrected UMAP

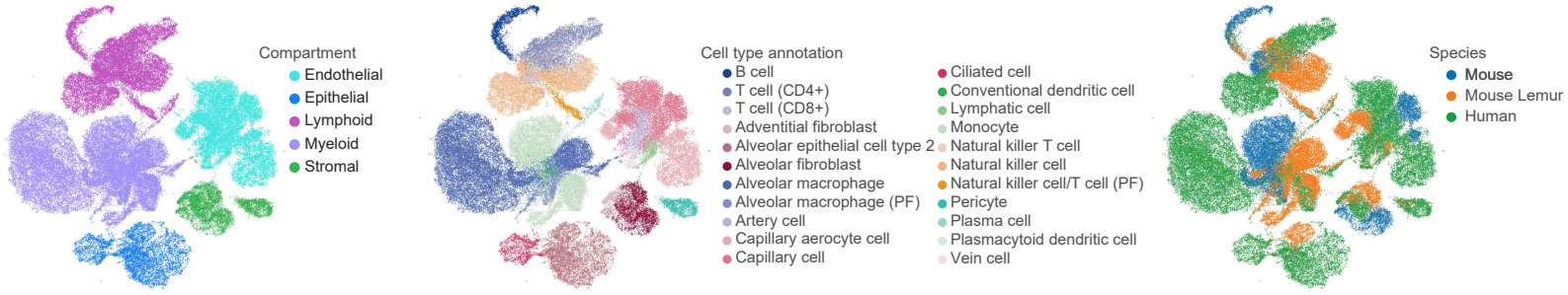
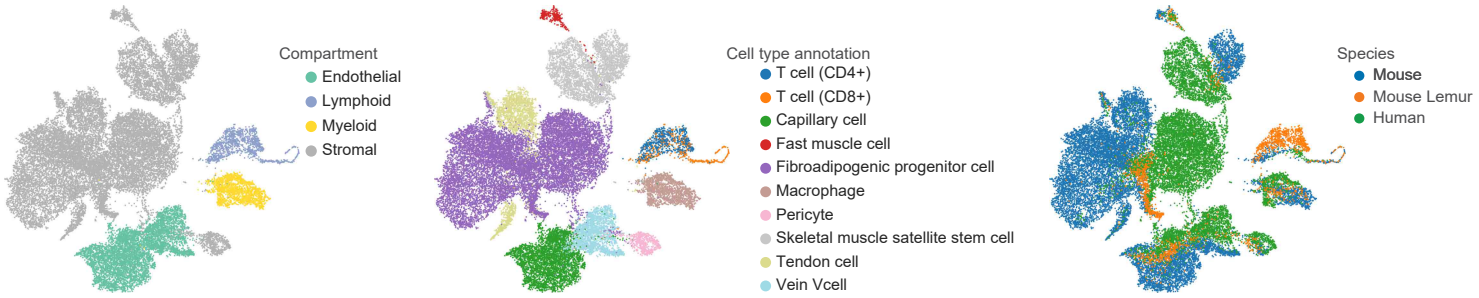


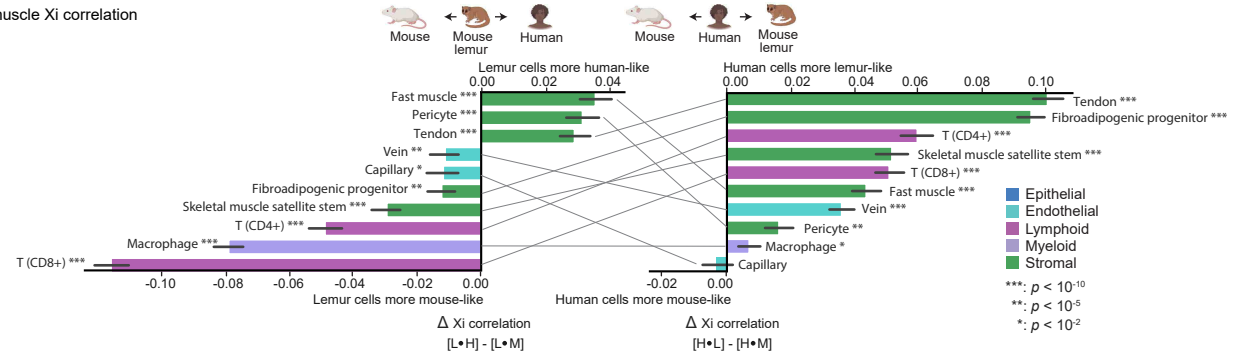
Fig. S8 - continued

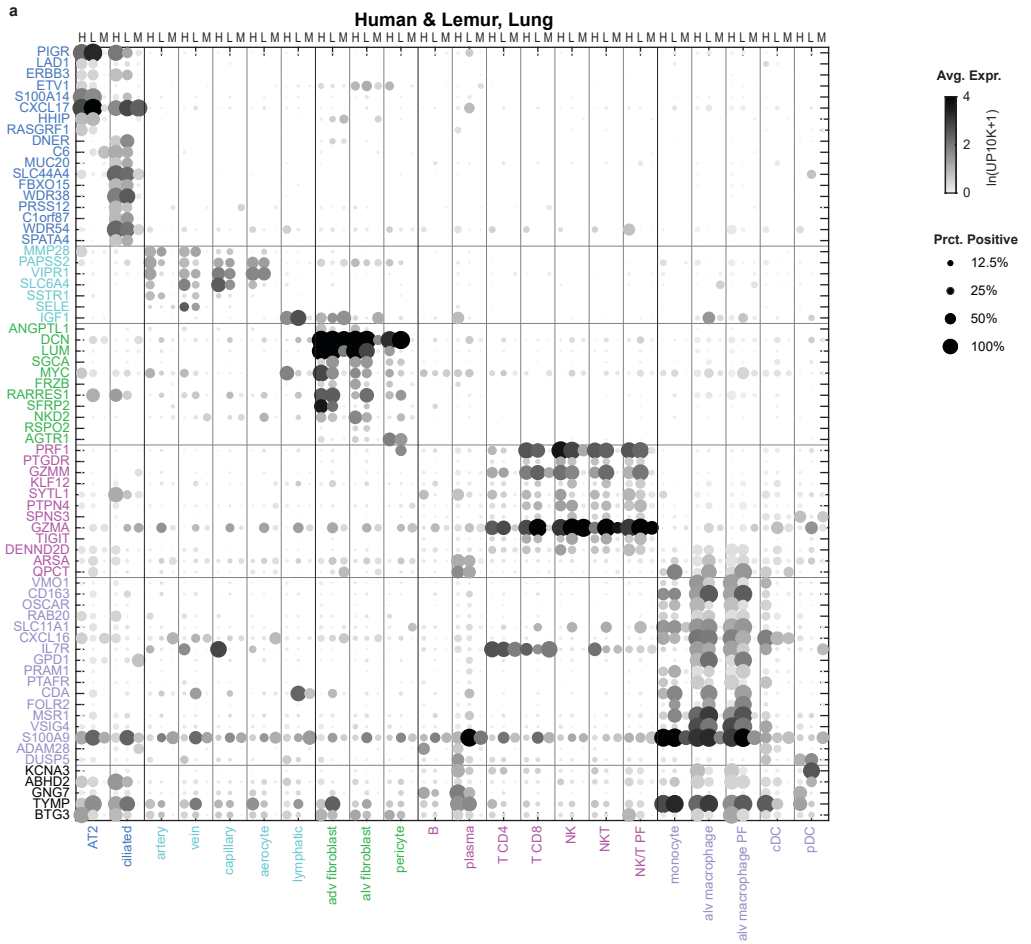
c Limb muscle Sankey plot

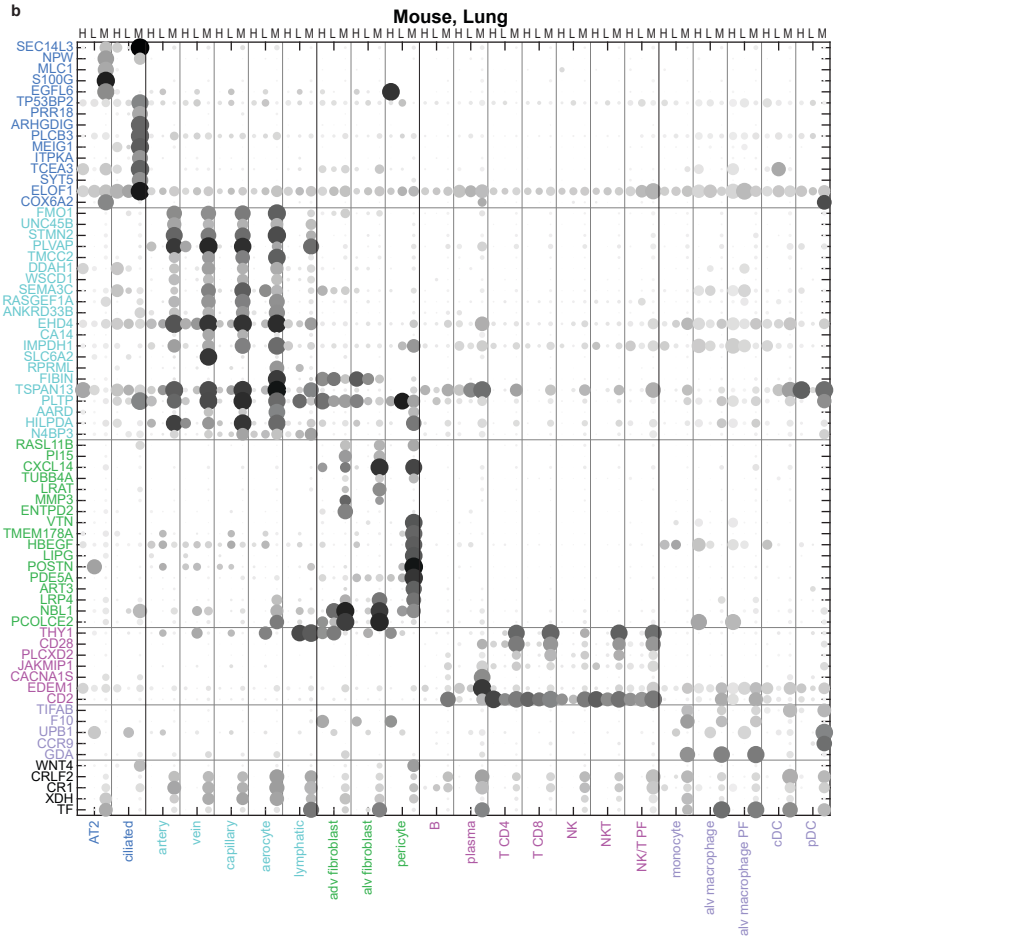
d Limb muscle BBKNN-corrected UMAP

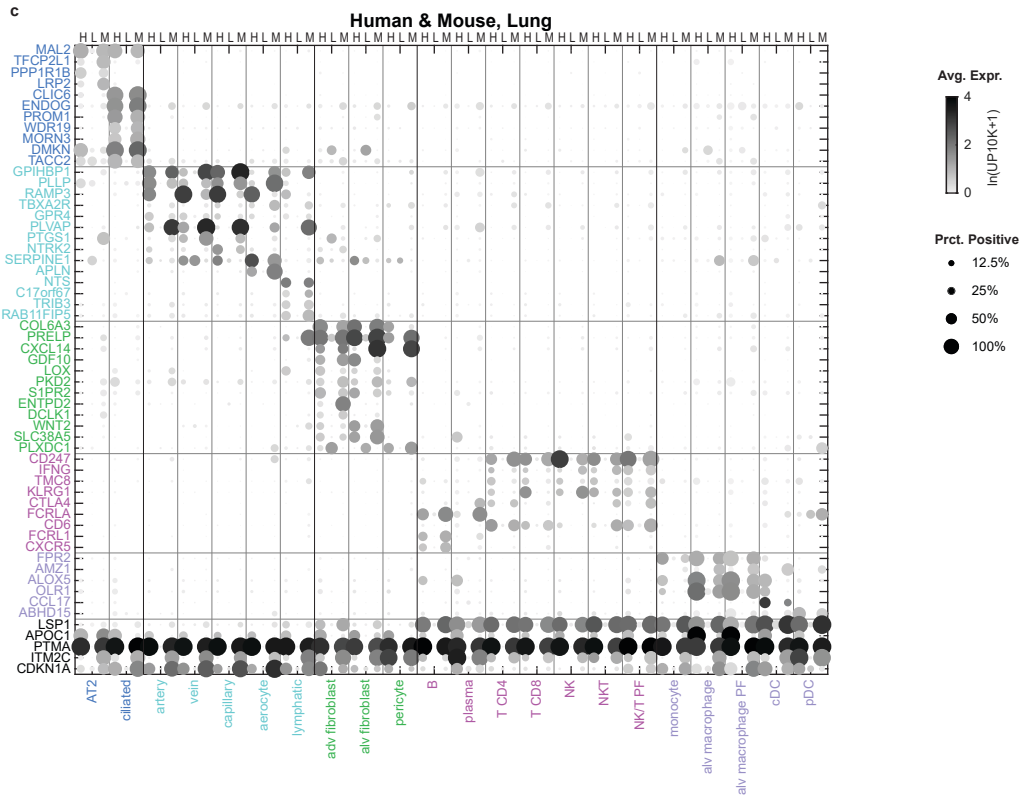


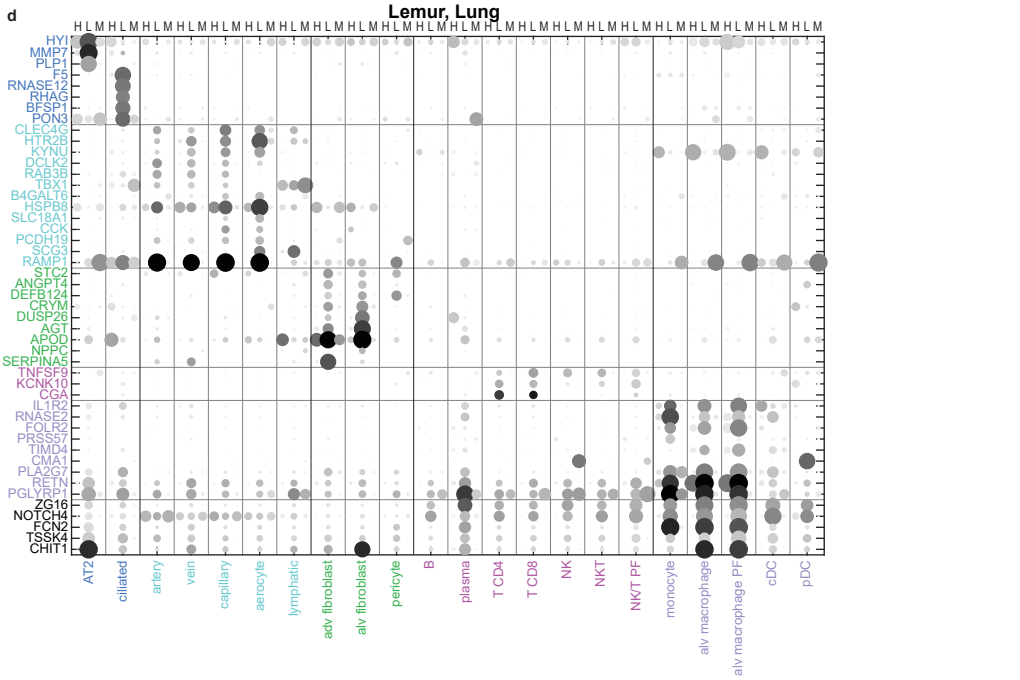
e Limb muscle Xi correlation

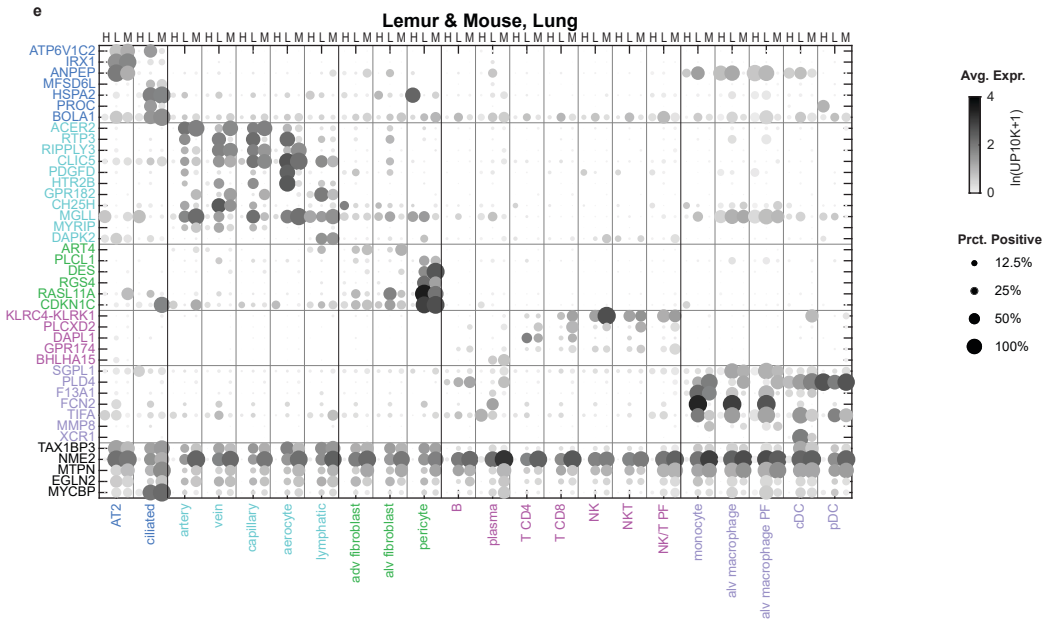


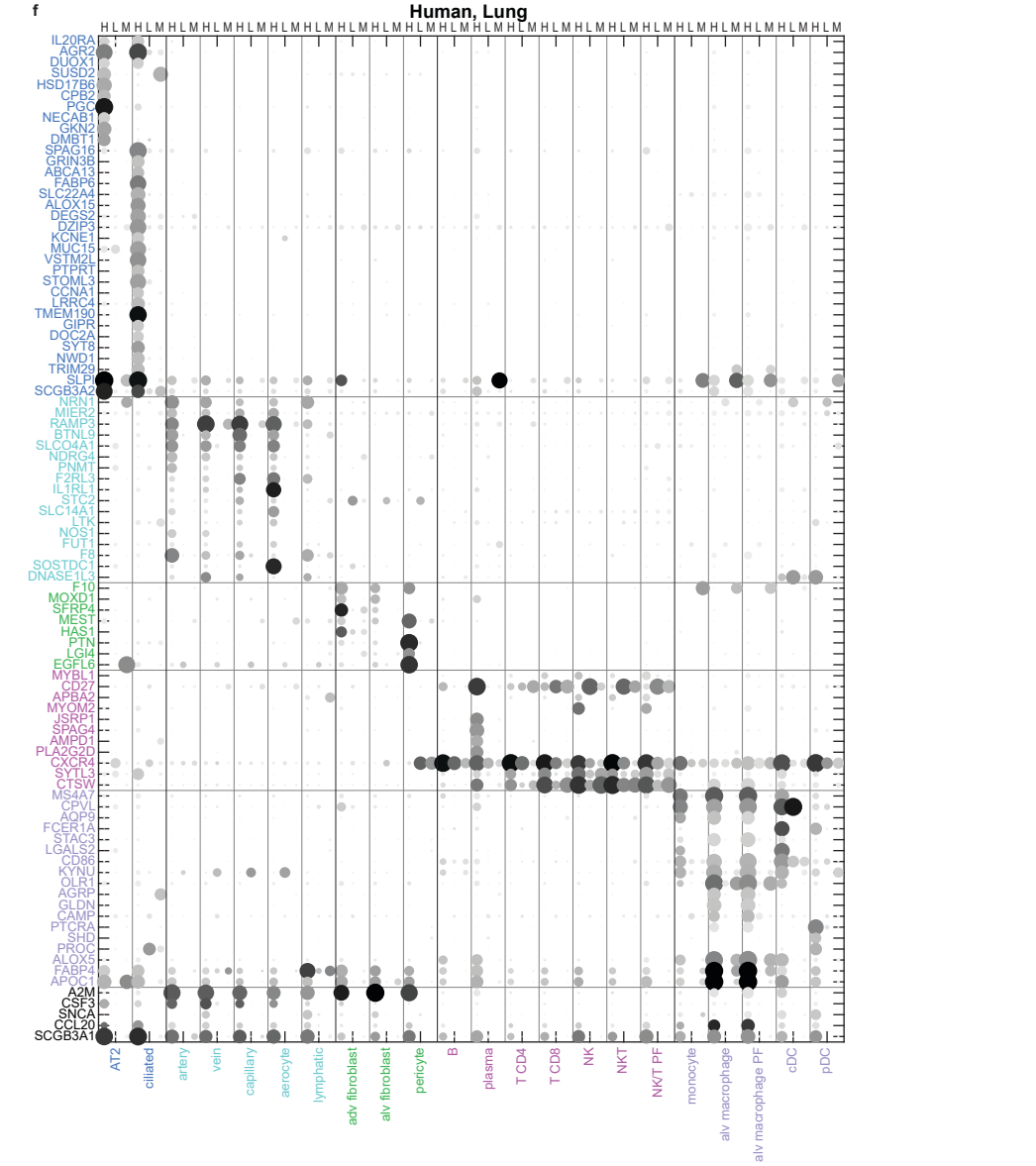

a
Human & Lemur, Lung
H L M
PIGR
LAD1
ERBB3
ETV1
S100A14
CXCL17
HHIP
RASGRF1
DNER
C6
MUC20
SLC44A4
FBXO15
WDR38
PRSS12
C1orf87
WDR54
SPATA4
MMP28
PAPSS2
VIPR1
SLC6A4
SSTR1
SELE
IGF1
ANGPTL1
DCN
LUM
SGCA
MYC
FRZB
RARRES1
SFRP2
NKD2
RSPO2
AGTR1
PRF1
PTGDR
GZMM
KLF12
SYTL1
PTPN4
SPNS3
GZMA
TIGIT
DENND2D
ARSA
QPCT
VMO1
CD163
OSCAR
RAB20
SLC11A1
CXCL16
IL7R
GPD1
PRAM1
PTAFR
CDA
FOLR2
MSR1
VSIG4
S100A9
ADAM28
DUSP5
KCNA3
ABHD2
GNG7
TYMP
BTG3
AT2
ciliated
artery
vein
capillary
aerocyte
lymphatic
adv fibroblast
alv fibroblast
pericyte
B
plasma
T CD4
T CD8
NK
NKT
NK/T PF
monocyte
alv macrophage
alv macrophage PF
cDC
pDC
Avg. Expr.
4
2
0
ln(UP10K+1)
Prct. Positive
12.5%
25%
50%
100%

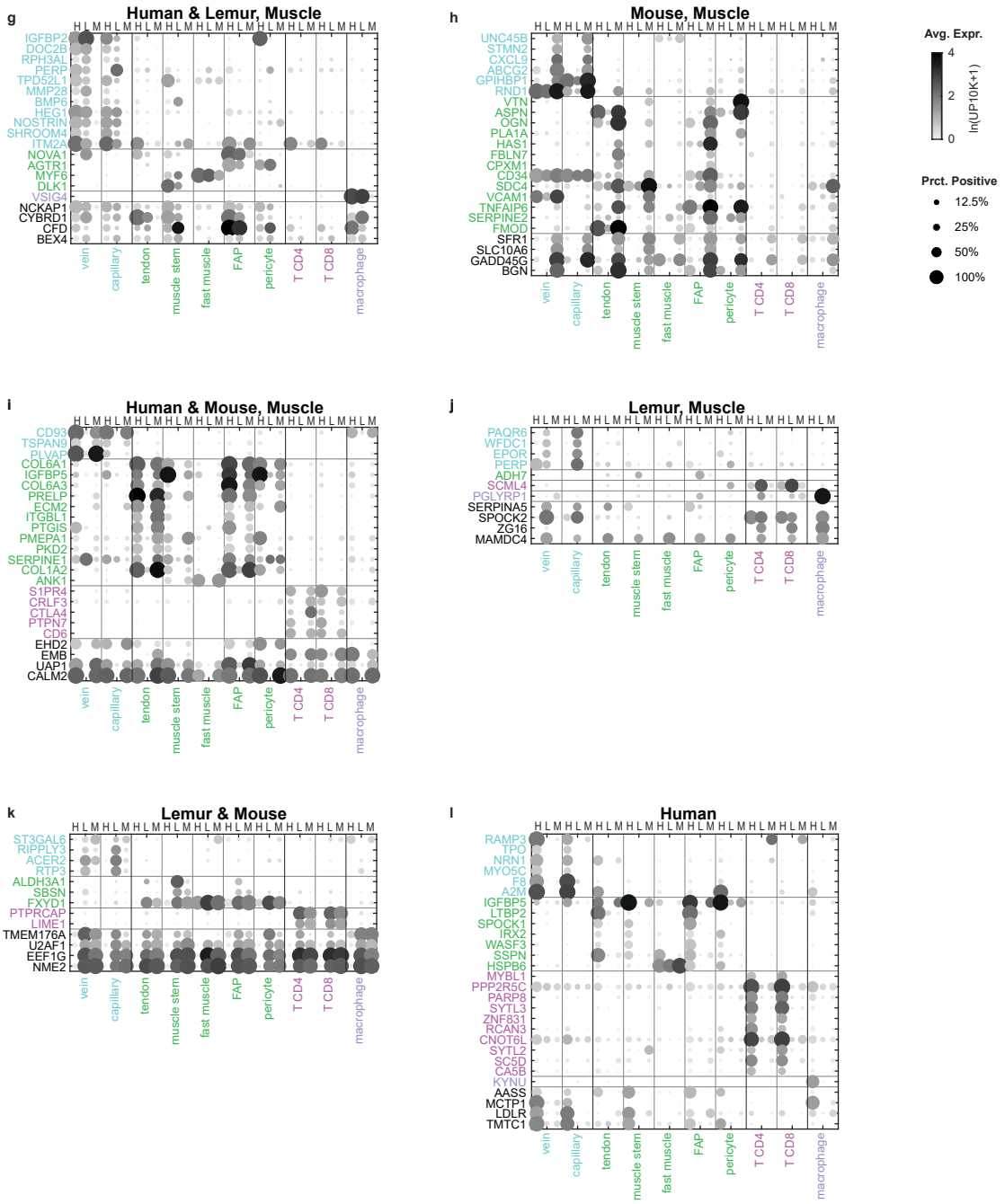

b
Mouse, Lung
H L M
SEC14L3
NPW
MLC1
S100G
EGFL6
TP53BP2
PRR18
ARHGDIG
PLCB3
MEIG1
ITPKA
TCEA3
SYT5
ELOF1
COX6A2
FMO1
UNC45B
STMN2
PLVAP
TMCC2
DDAH1
WSCD1
SEMA3C
RASGEF1A
ANKRD33B
EHD4
CA14
IMPDH1
SLC6A2
RPRML
FIBIN
TSPAN13
PLTP
AARD
HILPDA
N4BP3
RASL11B
PI15
CXCL14
TUBB4A
LRAT
MMP3
ENTPD2
VTN
TMEM178A
HBEGF
LIPG
POSTN
PDE5A
ART3
LRP4
NBL1
PCOLCE2
THY1
CD28
PLCXD2
JAKMIP1
CACNA1S
EDEM1
CD2
TIFAB
F10
UPB1
CCR9
GDA
WNT4
CRLF2
CR1
XDH
TF
AT2
ciliated
artery
vein
capillary
aerocyte
lymphatic
adv fibroblast
alv fibroblast
pericyte
B
plasma
T CD4
T CD8
NK
NKT
NK/T PF
monocyte
alv macrophage
alv macrophage PF
cDC
pDC


c
Human & Mouse, Lung
H L M
MAL2
TFCP2L1
PPP1R1B
LRP2
CLIC6
ENDOG
PROM1
WDR19
MORN3
DMKN
TACC2
GPIHBP1
PLLP
RAMP3
TBXA2R
GPR4
PLVAP
PTGS1
NTRK2
SERPINE1
APLN
NTS
C17orf67
TRIB3
RAB11FIP5
COL6A3
PRELP
CXCL14
GDF10
LOX
PKD2
S1PR2
ENTPD2
DCLK1
WNT2
SLC38A5
PLXDC1
CD247
IFNG
TMC8
KLRG1
CTLA4
FCRLA
CD6
FCRL1
CXCR5
FPR2
AMZ1
ALOX5
OLR1
CCL17
ABHD15
LSP1
APOC1
PTMA
ITM2C
CDKN1A
AT2
ciliated
artery
vein
capillary
aerocyte
lymphatic
adv fibroblast
alv fibroblast
pericyte
B
plasma
T CD4
T CD8
NK
NKT
NK/T PF
monocyte
alv macrophage
alv macrophage PF
cDC
pDC
Avg. Expr.
4
2
0
ln(UP10K+1)
Prct. Positive
12.5%
25%
50%
100%


d
Lemur, Lung
H L M
HYI
MMP7
PLP1
F5
RNASE12
RHAG
BFSP1
PON3
CLEC4G
HTR2B
KYNU
DCLK2
RAB3B
TBX1
B4GALT6
HSPB8
SLC18A1
CCK
PCDH19
SCG3
RAMP1
STC2
ANGPT4
DEFB124
CRYM
DUSP26
AGT
APOD
NPPC
SERPINA5
TNFSF9
KCNK10
CGA
IL1R2
RNASE2
FOLR2
PRSS57
TIMD4
CMA1
PLA2G7
RETN
PGLYRP1
ZG16
NOTCH4
FCN2
TSSK4
CHIT1
AT2
ciliated
artery
vein
capillary
aerocyte
lymphatic
adv fibroblast
alv fibroblast
pericyte
B
plasma
T CD4
T CD8
NK
NKT
NK/T PF
monocyte
alv macrophage
alv macrophage PF
cDC
pDC


e
Lemur & Mouse, Lung
H L M
ATP6V1C2
IRX1
ANPEP
MFSD6L
HSPA2
PROC
BOLA1
ACER2
RTP3
RIPPLY3
CLIC5
PDGFD
HTR2B
GPR182
CH25H
MGLL
MYRIP
DAPK2
ART4
PLCL1
DES
RGS4
RASL11A
CDKN1C
KLRC4-KLRK1
PLCXD2
DAPL1
GPR174
BHLHA15
SGPL1
PLD4
F13A1
FCN2
TIFA
MMP6
XCR1
TAX1BP3
NME2
MTPN
EGLN2
MYCBP
AT2
ciliated
artery
vein
capillary
aerocyte
lymphatic
adv fibroblast
alv fibroblast
pericyte
B
plasma
T CD4
T CD8
NK
NKT
NK/T PF
monocyte
alv macrophage
alv macrophage PF
cDC
pDC
Avg. Expr.
4
2
0
ln(UP10K+1)
Prct. Positive
12.5%
25%
50%
100%


f
Human, Lung
H L M
IL20RA
AGR2
DUOX1
SUSD2
HSD17B6
CPB2
PGC
NECAB1
GKN2
DMBT1
SPAG16
GRIN3B
ABCA13
FABP6
SLC22A4
ALOX15
DEGS2
DZIP3
KCNE1
MUC15
VSTM2L
PTPRT
STOML3
CCNA1
LRRC4
TMEM190
GIPR
DOC2A
SYT8
NWD1
TRIM29
SLPI
SCGB3A2
NRN1
MIER2
RAMP3
BTNL9
SLCO4A1
NDRG4
PNMT
F2RL3
IL1RL1
STC2
SLC14A1
LTK
NOS1
FUT1
F8
SOSTDC1
DNASE1L3
F10
MOXD1
SFRP4
MEST
HAS1
PTN
LGI4
EGFL6
MYBL1
CD27
APBA2
MYOM2
JSRP1
SPAG4
AMPD1
PLA2G2D
CXCR4
SYTL3
CTSW
MS4A7
CPVL
AQP9
FCER1A
STAC3
LGALS2
CD86
KYNU
OLR1
AGRP
GLDN
CAMP
PTCRA
SHD
PROC
ALOX5
FABP4
APOC1
A2M
CSF3
SNCA
CCL20
SCGB3A1
AT2
ciliated
artery
vein
capillary
aerocyte
lymphatic
adv fibroblast
alv fibroblast
pericyte
B
plasma
T CD4
T CD8
NK
NKT
NK/T PF
monocyte
alv macrophage
alv macrophage PF
cDC
pDC


g
Human & Lemur, Muscle
IGFBP2
DOC2B
RPH3AL
PERP
TPD52L1
MMP28
BMP6
HEG1
NOSTRIN
SHROOM4
ITM2A
NOVA1
AGTR1
MYF6
DLK1
VSIG4
NCKAP1
CYBRD1
CFD
BEX4
h
Mouse, Muscle
UNC45B
STMN2
CXCL9
ABCG2
GPIHBP1
RND1
VTN
ASPN
OGN
PLA1A
HAS1
FBLN7
CPXM1
CD34
SDC4
VCAM1
TNFAIP6
SERPINE2
FMOD
SFR1
SLC10A6
GADD45G
BGN
Avg. Expr.
ln(UP10K+1)
Prct. Positive
12.5%
25%
50%
100%
i
Human & Mouse, Muscle
CD93
TSPAN9
PLVAP
COL6A1
IGFBP5
COL6A3
PRELP
ECM2
ITGBL1
PTGIS
PMEPA1
PKD2
SERPINE1
COL1A2
ANK1
S1PR4
CRLF3
CTLA4
PTPN7
CD6
EHD2
EMB
UAP1
CALM2
j
Lemur, Muscle
PAQR6
WFDC1
EPOR
PERP
ADH7
SCML4
PGLYRP1
SERPINA5
SPOCK2
ZG16
MAMDC4
k
Lemur & Mouse
ST3GAL6
RIPPLY3
ACER2
RTP3
ALDH3A1
SBSN
FXYD1
PTPRCAP
LIME1
TMEM176A
U2AF1
EEF1G
NME2
l
Human
RAMP3
TPO
NRN1
MYO5C
F8
A2M
IGFBP5
LTBP2
SPOCK1
IRX2
WASF3
SSPN
HSPB6
MYBL1
PPP2R5C
PARP8
SYTL3
ZNF831
RCAN3
CNOT6L
SYTL2
SC5D
CA5B
KYNU
AASS
MCTP1
LDLR
TMTC1
H L M
vein
capillary
tendon
muscle stem
fast muscle
FAP
pericyte
T CD4
T CD8
macrophage