Supplement – Detailed Discussion of Tissue Cell Types

Table of Contents

Circulatory System

Heart and Aorta

Figures are located in the Tabula Muris Tissue Supplement under section(s):
Aorta FACS
Heart FACS
Heart and Aorta Droplet

The adult heart has four main chambers including the right and left atria, and right and left ventricles. The heart pumps blood through the aorta, the body's main vascular conduit that carries oxygenated blood. The aorta emerges from the left ventricle and extends down into the abdomen where it splits into two main branches. Like the heart, the aorta is mainly an organ of mesodermal origin with contributions from migratory neural crest cells (e.g., giving rise to parts of the muscle and connective tissue walls). The elastic vascular wall is composed of three main layers, the tunica intima, the tunica media, and the tunica externa. Resident immune cells including macrophages, lymphocytes, and dendritic cells, can be found in the aortic wall. Perivascular adipose tissue containing adipocytes is contiguous with the tunica adventitia, providing additional support for the vascular wall.

Tissue Processing

Hearts were dissected into four chambers (left and right atria, left and right ventricles) based on anatomical landmarks such as the atrial and ventricular canal and the septal groove. Dissected tissue pieces were minced in PBS on ice, and digested sequentially with 0.25% trypsin-EDTA (Gibco 25200056) and collagenase A/B mixture (Sigma 10103586001, 11088815001) for 10 and 20 minutes respectively at 37ºC. Cells were triturated and filtered through a 100 µm (Falcon 352360) strainer and stained with 1:500 propidium iodide (Life Tech P3655). Cardiomyocytes were then manually picked from half of each single cell solution, and the remaining cells were run unbiasedly on both the FACS and microfluidic platforms.

Aortas were minced with scissors and pelleted (270 x g, 5°C, 5 minutes) before digestion with 2.2 mg/ml Collagenase II (Sigma C6885) for 10 minutes at 37°C, and incubation at 37°C for 30 minutes with agitation. Digestion was quenched with FACS buffer (2% FBS, 1% Antibiotics (Gibco 15240-062), and 10% Pluronics (ThermoFisher 24040032) in PBS), cells were pelleted (270 x g, 5°C, 5 minutes), and stained using 1:50 dilutions of antibodies for TER119-PB (Biolegend 116232), CD45-Pe-Cy7 (eBioscience 25-0451- 82), EpCAM-G8.8 FITC (eBioscience 11-5791-82), and CD31-APC (BD Pharmingen 551262), and stored on ice for 20 minutes. Cells were washed and resuspended in FACS buffer, and stained with 1:1000 Sytox Blue (ThermoFisher S34857). Aortic cells were run together with the heart cells on the microfluidic droplet platform, or sorted separately from the heart cells by FACS. Viable heart cells were FACS sorted whereas cells from the aorta were sorted into 2 bins: endothelial ($CD45^-$, $TER119^-$, $CD31^+$, $EPCAM^-$) and smooth muscle cells ($CD45^-$, $TER119^-$, $CD31^-$, $EPCAM^-$).

Data analysis

From the cardiac and aortic tissue of 3 female and 2 male mice, 624 single cells were profiled with the microfluidic droplet platform and grouped into 5 clusters: vascular endothelial cells (*Fabp4*$^+$, *Cdh5*$^+$, *Cav1*$^+$)[1], fibroblasts (*Ddr2*$^+$, *Tcf21*$^+$, *Col3a1*$^+$, *Col1a2*$^+$, *Col1a1*$^+$)[2], atrial cardiomyocytes (*Nppa*$^+$, *Myl7*$^+$, *Sln*$^+$)[1], endocardial cells (*Npr3*$^+$, *Pecam1*$^+$)[3], immune cells (*C1qa*$^+$, *H2-Eb1*$^+$), and myofibroblasts/smooth muscle cells that can be separated by *Myh11* and *Tcf21*

expression. The presence of atrial cardiomyocytes is somewhat surprising given that these cells were thought to be too large for the microfluidic channel.

FACS sorted cardiac cells and picked cardiomyocytes formed 12 clusters totaling 4,365 cells. We observe cardiomyocytes (*Tnni3*$^{+}$), vascular endothelial cells (*Cdh5*$^{+}$, *Pecam1*$^{+}$, *Fabp4*$^{+}$, *Cav1*$^{+}$), aortic endothelial cells (*Ehd3*$^{+}$), endocardial cells (*Npr3*$^{+}$), fibroblasts (*Dcn*$^{+}$, *Gsn*$^{+}$), smooth muscle cells (*Myh11*$^{+}$), 2 clusters of immune cells (*C1qa*$^{+}$), and red blood cells (*Hbb-b1*$^{+}$). We also tracked the chamber of origin for all cells, with cells from each chamber present in all clusters.

Subanalysis of the 408 FACS-sorted aortic cells generated 5 clusters and better resolved specific cell types. For example, a cluster of mesenchymal cells that expresses vimentin[4] and genes encoding secreted extracellular matrix proteins (*Dcn, Col1a1, Ddr2, Col3a1, Eln, and Fn1*) is present. The majority of these cells are fibroblasts, while a few are likely smooth muscle cells that express *Acta2* and *Tagln*[5,6]. Three endothelial cell clusters are also present (*Cdh5*$^{+}$, *Pecam1*$^{+}$), which are similar to the vascular endothelial cells (*Cav1*$^{+}$, *Fabp4*$^{+}$) in the heart. Two of the clusters also express high levels of angiopoietin 1 receptor (*Tek)*[7], endothelial cell specific molecular 1 (*Esm1*)[8], and *Ehd3*[9]. In addition to the cluster of red blood cells (*Hbb*$^{+}$), a cluster of antigen presenting cells is present (*Ptprc*$^{+}$, *H2-Eb1*$^{+}$, *Cd86*$^{+}$). The majority are most likely macrophages (*Selplg*$^{+}$, *Cd14*$^{+}$, *Cd3e*$^{-}$, *Cd19*$^{-}$).

Respiratory System

Lung

Figures are located in the Tabula Muris Tissue Supplement under section(s):
Lung Droplet
Lung FACS

The lung carries out many physiologic functions, with gas exchange the most critical. To function effectively while constantly exposed to airborne pathogens and potentially noxious substances, the branched network of tubes comprising the lung is lined by highly specialized epithelial cells. This epithelial monolayer forms a tight mechanical barrier against inhaled pathogens and particulates. The most abundant cells in the airway epithelium are secretory cells (termed club cells and goblet cells), and multiciliated cells that generate a directional flow of mucus by coordinated beating of their cilia. Rare neuroendocrine cells form neuroepithelial bodies at airway branch points and may function as chemoreceptors and oxygen sensors. The airway epithelium is surrounded by different types of supporting cells including contractile smooth muscle, matrix producing stromal cells, immune cells, and innervating neurons. The lung also contains a highly branched vasculature composed of endothelial cells and mural cells. At the terminal airways, millions of tiny air sacs called alveoli interface with this vasculature to form the gas exchange surface. This surface is formed by two types of epithelial cells, the squamous alveolar type 1 cells that are closely apposed to endothelial cells of capillary tubes, and the cuboidal type 2 cells that supply alveoli with surfactant to prevent their collapse during breathing[10].

Tissue Processing

The lung was dissected, minced, and digestion buffer (2 U/ml liberase, Sigma 5401127001, in RPMI) media was added before placing the tissue in gentleMACS c-tubes (Miltenyi 130-096-334). The sample was then incubated on a nutator at 37°C for 30 minutes, run again on the gentleMACS, and placed on ice for the remainder of the protocol. After adding FACS buffer (5% FBS in PBS), cells were pelleted (300 x g, 5 minutes, 4°C), resuspended in FACS buffer, filtered through a 70 µm strainer (Fisherbrand 22363548), pelleted again (300 x g, 5 minutes, 4°C), and resuspended in FACS buffer with 1:100 Fc block (BD 553141). Cells were then stained with one

of three panels, depending on instrument availability, to improve representation of low abundance cell populations. For sorts done on a BD FACS Aria II at the Stanford Shared FACS Facility (annotated 'Aria' in the metadata), cells were stained with CD326-APC/Fire750 (Biolegend 118230, 1:50), CD31-PE/Dazzle594 (Biolegend 102526, 1:100), CD45-BV510 (Biolegend 103138, 1:100), endomucin-FITC, (eBioscience 14-5851-82 1:30, conjugated with Abcam ab102884), CD140a/PDGFRa-APC (R&D Systems AF1062 1:100, conjugated with Abcam ab201807), CD140b/PDGFRb-APC (eBioscience 17-1402-82, 1:30), Thbs1-APC (Fisher MA5-13398, 1:50, conjugated with Abcam ab201807), Sdc4-APC (Miltenyi 130-109-831 1:20), LNGFR-PE/Vio770 (Miltenyi 103110079 1:20), C-FMS-BV411 (Biolegend 135513 1:30) for 30 minutes at 4°C shaking. For sorts done on a Sony SH800S, cells were split and stained with CD31-APC (BD Biosciences 551262), CD45-PE (Biolegend 103106), and endomucin-FITC (annotated 'Endomucin' in the metadata) or CD140a/PDGFRa-APC, CD140b/PDGFRb-APC, Thbs1-APC, Sdc4-APC, CD31-FITC (BD Biosciences 561813), and CD326-APC/Fire750 (annotated as 'Epcam' in the metadata). Cells were then pelleted (300 x g, 5 minutes, 4°C), washed with 5% FBS in PBS, spun again, and resuspended in 2% FBS in PBS before being passed through a 35 µm FACS tube (Falcon 352235). Propidium iodide (AnaSpec A-83215, 1:1000, Aria sorts) or Sytox Blue (ThermoFisher S34857, Sony sorts) was added immediately prior to sorting. Cells were sorted into 5 bins on the Sony instrument: 2 endothelial clusters (Endomucin$^-$/CD31$^+$/CD45$^-$, and Endomucin$^+$/CD31$^+$/CD45$^-$) both termed "Endomucin" in the metadata, epithelial (CD31$^-$, EpCAM$^+$), mesenchymal (CD31$^-$, PDGFRa$^+$, PDGFRb$^+$, Thbs1$^+$, syndecan4$^+$), and immune (CD31$^-$, EpCAM$^-$, PDGFRa$^-$, PDGFRb$^-$, Thbs1$^-$ syndecan4$^-$), each termed "Epcam" in the metadata. On the BD FACS Aria I, cells were sorted into 10 bins (figure reference).

Data Analysis

From 5,449 and 1,716 cells isolated using microfluidic droplets and FACS from 2 males and 2 females, we identified 25 transcriptionally distinct populations from nearly every tissue compartment in the lung. Specifically, we identified 6 epithelial (basal, club, multiciliated, neuroendocrine, alveolar type 1, and alveolar type 2 cells), 4 endothelial, 4 stromal, and 11 immune cell types (alveolar macrophages, interstitial macrophages, dendritic cells, invading monocytes, circulating monocytes, mast, B, T, natural killer cells, and two non-canonical populations). All cell types were shared between datasets except basal cells (only in FACS data), one endothelial cell type (only in FACS data), and one unknown immune population (only in droplet data). Epithelial and immune cell identities were assigned using established marker genes (such as surfactant proteins for alveolar type 2 cells, the transcription factor *Foxj1* for ciliated cells, or *Marco* for alveolar macrophages) or correlations to expression signatures from previous work[11]. Markers for specific endothelial and stromal cell types are less characterized, and we grouped these clusters together using genes expressed broadly within each population (such as *Pecam1* and *Col1a1*, respectively). While the low number of epithelial cells and macrophages prevented separating these populations into distinct clusters, we were able to use the most sensitive and specific canonical markers to identify most of the known cell types.

Lung resident and itinerant immune cells are of particular interest because they are not well characterized, yet they are the first responders to damage caused by inhaled pathogens, toxins, and other irritants associated with lung diseases including cancer and emphysema[12–15]. Genetic lineage tracing and parabiosis experiments have identified three distinct macrophage lineages that localize to different lung regions: primitive macrophages (localized to the peripheral interstitial space), interstitial macrophages (broadly distributed throughout the interstitium), and alveolar macrophages (localized to the alveolar lumen)[16]. Alveolar macrophages (*Marco*$^+$, *Itgax*$^+$, *Mrc1*$^+$) clustered separately from other immune populations and, although interstitial macrophages clustered with other myeloid cells, we could identify them based on marker expression (*Csf1r*$^+$, MHCII$^+$, *Mrc1*^{low}, Itgax$^-$, *Ly6c2*$^-$). We did not identify any primitive macrophages.

Comparison of the different macrophage population's full gene expression profiles shows that interstitial macrophages are enriched for expression of genes involved in antigen processing and presentation and interferon responses, indicating a dominant role among lung macrophages in canonical immune functions and inflammation. By contrast, alveolar macrophages are enriched for expression of genes involved in catalysis of lipids, iron sequestration, and the transport of long-chain fatty acids, suggesting functional specialization for airway clearance.

We also identified lung dendritic cells based on their expression of dendritic cell markers *Itgax*, *Cd24a*, and *Cd68,* and low expression of monocyte and macrophage markers *Ly6c2*, *Cx3cr1*, *Cd14*, *Csf1r*, and *Mrc1*. Their full gene expression profile shows enrichment for genes involved in regulating T cell differentiation and the helper T cell response, consistent with their known function[17]. A cluster of invading monocytes distinct from the circulating monocyte cluster is also evident, based on expression of *Itgam*, *Csf1r*, *Ly6c2* and *Cd14*. Interestingly, invading monocytes express a unique set of pro-inflammatory chemokines and receptors that recruit macrophages and other monocytes.

The other two unknown immune clusters have expression signatures most similar to dendritic cells. Both clusters express some neutrophil markers such as *Itgam* and *Cxcr4*, but are missing other canonical markers like *Ly6g*. Both also express the dendritic marker *Cd24a*, and one cluster expresses additional dendritic, monocyte, and macrophage markers *Itgax*, *Cd68*, *Ly6c2*, *Cx3cr1*, and *Csf1r*. The lack or contradictory expression of canonical cell type marker genes means these populations could represent distinct subtypes (or cell states) of myeloid cells, neutrophils (or other granulocytes), hematopoietic intermediates, or other rare or uncharacterized immune cell types. The markers identified here will facilitate purifying these cells, localizing them in vivo, and defining their biological functions and fates.

Trachea

Figures are located in the Tabula Muris Tissue Supplement under section(s):
Trachea Droplet
Trachea FACS

The trachea, like other gas-conducting components of the respiratory system, consists of mucosal elements including an epithelium with ciliated, secretory, and neuroendocrine cells. It differs from more distal portions of the mouse airway due to the presence of an extensive system of epithelial basal cells, as well as a set of proximally located submucosal glands, whose primary function is to secrete mucus. The cells of the tracheal epithelium are generally characterized by expression of *Epcam* and *Cdh1,* and more specifically by expression of *Foxj1* within ciliated cells, *Scgb1a1* within secretory cells, and *Krt15* and/or *Krt14* within basal cells. Subjacent to the basement membrane, the tracheal epithelium is supported by *Pdgfrb*-expressing mesenchymal cells interspersed with cartilaginous elements[18–21].

Tissue Processing

Trachea, excluding thyroid and parathyroids but including submucosal glands, was dissected from below the cricoid cartilage. Dissected tissue was rinsed in ice cold PBS (Gibco, 10010049). The trachea was then cut longitudinally to expose the epithelium and directly placed in 1 ml of digest buffer (5 U/ml dispase (Gibco, 17105041) and 40 U/ml collagenase I (Gibco, 17018029) in HBSS (Gibco, 14175095) with 2% FBS and 1X Penicillin-Streptomycin (Gibco, 15140-122)) for 30-45 minutes on ice. The tissue was then minced in digest buffer and incubated at 37°C for 45-60 minutes, with pipetting throughout to further dissociate the tissue. After filtering through a 40 μm strainer (BD, 08-771-1), the cells were pelleted (2,000 x g, 5 minutes, 21°C), and washed twice with PBS. They were then resuspended in ACK lysis buffer for 1 minute at 21°C,

pelleted (2,000 x g, 5 minutes, 21°C), and washed twice with PBS. After resuspension in 2% FBS with 1X Penicillin-Streptomycin in PBS, the cells were filtered through a 40 μm strainer, and 1:1000 Sytox Blue (Invitrogen S34857) was added immediately prior to sorting.

Data Analysis

Our dissection and dissociation strategy captured viable epithelial and mesenchymal cells, including cells from the submucosal glands. Cells were analyzed with microfluidic droplets (11,269 cells) or were FACS-sorted for viability and sequenced (1,350 cells). We identified with both platforms epithelial, endothelial, and immune cell clusters, as well as three corresponding mesenchymal cell populations. Both platforms also revealed a small cluster of possible neuroendocrine cells.

Analysis of gene expression within the cells of the FACS clusters identified three mesenchymal populations, each showing high expression of *Pdgfrb*. These populations can be differentiated by their expression of 3 genes: *Col1a1*^{high} *Col8a1*^{low} *Pdgfra*^{high}; *Col1a1*^{high} *Col8a1*^{high} *Pdgfra*^{low}; *Col1a1*^{low} *Col8a1*^{high} *Pdgfra*^{low}. An epithelial population in the FACS analysis is defined by high *Epcam* and *Cdh1* expression; cells within this cluster express *Krt5, Scgb1a1,* and *Foxj1*, suggestive of basal, secretory, and ciliated cell types, respectively. Subclustering of this population identified and separated basal from secretory types. Ciliated cells did not form a distinct cluster, probably due to the low number of this cell type. Two distinct cell clusters identified by expression of *Ptprc* (CD45) seem likely to represent immune cells. Finally, an endothelial population is identified by expression of *Pecam1* (CD31).

Analysis of gene expression within the microfluidic data also identified three mesenchymal clusters corresponding to those identified in the FACS cells on the basis of their expression of *Pdgfrb*, *Pdgfra, Col1a1*, and *Col8a1*. Similar to the FACS analysis, a single immune cell cluster was identified by *Ptprc* expression, and an endothelial cell cluster was identified by *Pecam1* expression. *Epcam* and *Cdh1* expression identified an epithelial cluster which contains cells expressing *Krt5*, *Foxj1* and *Scgb1a1*, suggestive of basal, ciliated, and secretory cell identities; sub-clustering of this population indeed appeared to separate these three cell types. Our analysis also revealed a small but distinct cluster with cells expressing cholecystokinin (*Cck*) and synaptophysin (*Syn*), consistent with properties of a neuroendocrine cell type[22,23]. In summary, we have for the first time described a variety of tracheal cell types based on unbiased clustering of single-cell transcriptomic data.

Digestive System

Tongue

Figures are located in the Tabula Muris Tissue Supplement under section(s):
Tongue Droplet
Tongue FACS

The tongue mediates mechanical processing and chemosensory discrimination of ingested substances. The tongue epithelium is organized into different types of specialized, protruding papillae, including filiform, fungiform, foliate, and circumvallate. The filiform papillae comprise heavily keratinized cells and cover the majority of the tongue's dorsal surface, whereas the latter three papillae contain taste receptor cells, which are responsible for chemoreception and comprise less than 1% of the lingual epithelial cells. Undifferentiated basal cells are highly proliferative and give rise to all mature cells of the lingual epithelium[24].

Tissue Processing

Tongues excised at the level of the anterior intermolar eminence, removing the circumvallate and foliate papillae, were injected at several points along the sub-epithelial lingual tissue with 5 units/ml dispase (Corning 354235) until fully distended, and then incubated in an external solution of dispase for 10 minutes at 37°C. The tongue epithelium was then peeled away from the sub-epithelial lingual tissue, rinsed with PBS, minced, and digested at 37°C on an orbital shaker for 1 hour in Collagenase Type IV (Worthington LS004188) and DNase I (Worthington LS006343), followed by 30 minutes in 1x TrypLE (ThermoFisher A1217701) and DNase I. Cells were then filtered through a 40 µm strainer (Falcon 352340), pelleted (500 x g, 4°C, 5 minutes), and resuspended in FACS buffer (1X Penicillin-Streptomycin (ThermoFisher 15140122), 1X Pluronic F-68 (ThermoFisher 24040032), and 2% FBS (Atlanta Biologicals S11550H) in PBS pH 7.4 (ThermoFisher 100100-23)). Cells were stained with 1:50 anti-CD45-Pacific Blue (Biolegend 103126), 1:50 anti-TER119-Pacific Blue (Biolegend 116232), 1:100 Sca1-FITC (Biolegend 122506) and 1:100 EpCAM-APC (ThermoFisher 17-5791-82). Live/Dead stain was performed with 1:1000 SYTOX Blue (ThermoFisher S34857) immediately prior to sorting. Cells ($CD45^-$, $TER119^-$) were sorted into 2 bins: $EpCAM^{hi}$ and $EpCAM^{low}$.

Data Analysis

Our single-cell tongue preparations enriched for epithelial cells and omitted the underlying mesenchymal tissues, thereby excluding muscle and blood vessels. We analyzed 1,416 cells from 3 males and 2 females using FACS, and 7,538 cells from 2 males and 1 female using microfluidic droplets. Through unbiased clustering analysis, both platforms identified a cluster containing highly proliferative cells, as indicated by the expression of cell cycle related genes, *Top2a*, *Cdc20,* and *Mki67*. Similar percentages of proliferating cells are present in the two datasets (13.9% and 14.5% in FACS and microfluidic droplet, respectively). Cells in this proliferative cluster are basal in character, as indicated by expression *Krt14*, *Krt5,* and *Krt15*. We also noted basal character in cells of other non-proliferative clusters. The proportion of total cells expressing the three basal markers is 65.0% and 58.9 % in the FACS and microfluidic droplet platforms, respectively.

The remaining clusters show features of more differentiated keratinocytes at distinct stages of differentiation. One group of clusters shows enriched expression of *Krt10* and *Sbsn*, previously associated with the suprabasal spinous layer of skin epidermis[25,26], but not previously characterized in the tongue. Interestingly, one cluster shows co-expression of genes indicating both basal (*Krt14*) and suprabasal (*Krt10* and *Sbsn*) character, suggesting a distinct cluster of cells in an intermediate state of differentiation. The last small but distinct cluster has high levels of *Krt84* and *Krt36*, which are specifically associated with keratinocytes of the filiform papillae[27]. This cluster also shows high *Hoxc13*, a gene often associated with specification of the filiform papilla cell fate.

Our study presents the first unbiased single-cell analysis of the entire tongue epithelium. We note the expression of genes previously reported in isolated type II and type III taste receptor cells (*Krt8*, *Krt19*, *Tas1r1*, *Tas1r2*, *Tas1r3*, *P2rx7*)[28], but in our analysis this expression occurs in rare cells widely distributed among all clusters. This may occur because insufficient cells are captured in our preparations to define clusters comprised of taste receptor cell types. Perhaps the most striking feature to emerge from our unbiased single-cell study of tongue epithelium is the degree of enrichment for basal cell types, including a substantial fraction of proliferating cells. The abundance of such proliferative basal cells is consistent with the 4-8 day turnover rate of the lingual epithelium, one of the fastest epithelial turnover rates for any mouse organs[29].

Liver

Figures are located in the Tabula Muris Tissue Supplement under section(s):
Liver Droplet

Liver FACS

As the largest internal organ, the liver has several essential functions including blood detoxification, clotting factor and albumin synthesis, glycogen storage, and bile production, which aids digestion in the small intestine. The principle anatomical unit of the liver is the lobule, comprised primarily of hepatocytes grouped in interconnected plates between the afferent portal veins and the efferent central veins. Hepatocytes form bile canaliculi on their basolateral surface, which drain into bile ducts lined by biliary epithelial cells (BECs). Along the central vein to portal vein axis, hepatocytes show both functional and gene expression differences[30–33]. For example, periportal (PP) hepatocytes are known to express higher levels of gluconeogenesis and ureagenesis enzymes while pericentral (PC) hepatocytes express higher levels of glycolysis and xenobiotic metabolism enzymes.

The lobule is also traversed by sinusoidal endothelia composed of fenestrated endothelial cells[34]. Lipid-storing pericytes called stellate cells reside in the space between these cells and hepatocytes, and can become activated fibroblasts after injury[35]. Moreover, the liver is rich with resident immune cells[36], including macrophages known as Kupffer cells, natural killer (NK) cells, and natural killer T (NKT) cells.

Tissue Processing

Hepatocytes and non-parenchymal cells (NPCs) were isolated by a two-step collagenase perfusion technique with modifications[37]. Briefly, after the inferior vena cava was cannulated with a 24 gauge catheter and the portal vein was cut, the liver was perfused at 10 ml/minute through the inferior vena cava with Liver Perfusion Medium (Invitrogen 17701-038) at 37 °C for 5 minutes, followed by perfusion with collagenase type IV (Wellington LS004188) in HBSS (GIBCO 14025126) for 5 minutes. The liver was dissected out, incubated on ice for 30 minutes, and passed through a 100 μm filter (Falcon 352360). Hepatocytes were separated from NPCs by low-speed centrifugation (50 x g, 5 minutes, 3X, brake = 2), and further purified by Percoll gradient centrifugation (50% v/v), as described previously[38]. NPCs were pelleted from supernatant by centrifugation (300 x g, 5 minutes). Hepatocytes and NPC were either stained with Hoechst 33342 (1 μg/ml, Invitrogen, H3570) with 5 μM Reserpine (Sigma R0875) and propidium iodide (2 μg/ml, Life Tech P3655) for FACS sorting, or resuspended in 2% FBS (v/v) in PBS for the microfluidic droplet platform.

Data Analysis

Following retrograde perfusion via the inferior vena cava (IVC) with liver perfusion medium and collagenase, hepatocytes were isolated by centrifugation. A total of 1,845 single hepatocytes from 1 male and 2 females were successfully analyzed on the microfluidic platform. Approximately 5% of the captured cells are non-parenchymal cells, which is in line with the reported efficiency of the isolation procedure. Hepatocyte-specific genes are broadly expressed, including *Alb*, *Ttr*, *Apoa1*, and *Serpina1c*. Using known zonated metabolic genes, we could identify within this population PC hepatocytes (*Cyp2e1*$^{+}$, *Glul*$^{+}$, *Oat*$^{+}$, *Gulo*$^{+}$), periportal PP hepatocytes (*Cyp2f2*$^{+}$, *Pck1*$^{+}$, *Hal*$^{+}$, *Cdh1*$^{+}$), and midlobular (ML) hepatocytes (*Cyp2e1*^{low} or *Cyp2f2*^{low}, *Ass1*^{high}, *Hamp*^{high}, *Gstp1*^{high}, *Ubb*^{high})[33,39].

In addition to using the microfluidic droplet platform, we also analyzed hepatocytes and non-parenchymal cells sorted by FACS based on viability. Hepatocytes were purified as before and non-parenchymal cells were separately isolated using density centrifugation. Of 714 viable cells, we identified hepatocytes (*Alb*$^{+}$, *Ttr*$^{+}$, *Apoa1*$^{+}$, *Serpina1c*$^{+}$), endothelial cells (*Pecam*1^{+}, *Nrp1*$^{+}$, *Kdr*$^{+}$ and *Oit3*$^{+}$)[40], and Kupffer cells (*Emr1*$^{+}$, *Clec4f*$^{+}$, *Cd68*$^{+}$, *Irf7*$^{+}$)[41–43].

We also identified minor populations[44] that are most likely NK/NKT cells (*Zap70*$^{+}$, *Il2rb*$^{+}$, *Nkg7*$^{+}$, *Cxcr6*$^{+}$, *Gzma*$^{+}$) and circulating B cells (*Cd79a*$^{+}$, *Cd79b*$^{+}$, *Cd74*$^{+}$ and *Cd19*$^{+}$). We did not,

however, detect BEC or stellate cells in our cluster analysis, possibly due to low cell viability as a result of perfusion or lack of enrichment with our purification/sorting strategy.

Pancreas

Figures are located in the Tabula Muris Tissue Supplement under section(s):
Pancreas FACS

Located in the upper abdominal cavity, the pancreas is a dual-functioning organ that includes both exocrine and endocrine compartments. The exocrine pancreas, comprising ~95% of organ mass, contains acinar and ductal cells. Acinar cells produce and secrete digestive enzymes (e.g. amylases, proteases) into pancreatic ducts which drain into the duodenum to play an essential role in food digestion. The endocrine pancreas is composed of the islets of Langerhans. Interspersed among acinar clusters, they account for ~5% of pancreas mass. Islets consist of 5 types of cells which each produce a unique hormone, namely glucagon-producing alpha, insulin-producing beta, somatostatin-producing delta, pancreatic peptide-producing PP (also called gamma), and a transient fetal population of ghrelin-producing epsilon cells. These hormones are secreted in response to metabolic cues and act coordinately to maintain blood glucose homeostasis. Less abundant vascular, neural, stromal, and immune cells are also present in the pancreas.

Loss or damage of insulin-producing islet beta cells can lead to diabetes, and therefore a considerable amount of single cell transcriptomic research has been conducted on islets, especially in humans. However, only a few studies have investigated the mouse pancreas, with an inevitable focus on endocrine cells[45,46].

Tissue Processing

Immediately after cardiac perfusion with PBS, the pancreas of each mouse was inflated through the bile duct with 5 ml of cold 1 mg/ml collagenase type XI solution (Sigma-Aldrich, C7657), and then excised. The pancreas was then digested in another 5 ml of collagenase solution at 37°C for 10 minutes, and subjected to Ficoll (GE Healthcare, 17-1440-02) gradient centrifugation as described previously[47]. The resultant pellet contained acinar tissue while the supernatant contained islets, which were further purified by handpicking. Separated acinar cells and islets were dispersed into single cells by enzymatic digestion using Accumax (Thermo Fisher Scientific, 00-4666-56) and 1 U/ml Dispase solution (Thermo Fisher Scientific, 17105041) as previously described[48,49]. After each enzymatic digestion step, cells were washed with 1X PBS. After centrifugation, the cell pellets were resuspended in PBS and passed through a 70 µm cell strainers (BD Biosciences, 352350). The still separated endocrine and exocrine cells were stained with LIVE/DEAD Fixable Near-IR Dead Cell Dye following manufacturer's instructions (Thermo Fisher Scientific, L34975) to assess viability.

Data Analysis

We first separated islets from exocrine cells and subjected both fractions independently to FACS sorting based on live/dead staining. Of the resulting 1,564 cells from 2 male and 2 female mice, we identified 10 distinct populations. This includes many previously characterized pancreatic cell types based on known cell-specific transcripts: beta (*Ins1*$^+$, *Ins2*$^+$, *Slc2a2*$^+$, *Nkx6-1*$^+$, *Pdx1*$^+$, and *Mafa*$^+$), alpha (*Gcg*$^+$, *Mafb*$^+$, and *Arx*$^+$), delta (*Sst*$^+$, *Hhex*$^+$, and *Pdx1*$^+$), and PP cells (*Ppy*$^+$, *Mafb*^{neg}, and *Spp1*$^+$) of islets, as well as acinar (*Amy2b*$^+$, *Cpa1*$^+$, and *Ptf1a*$^+$), ductal (*Krt19*$^+$ and *Hnf1b*$^+$), endothelial (*Pecam1*$^+$, *Cdh5*$^+$, and *Kdr*$^+$), immune (*Ptprc*$^+$), and pancreatic stellate cells (*Pdgfra*$^+$ and *Pdgfrb*$^+$).

All four types of endocrine cells are represented in our dataset, accounting for 5 of 10 cell clusters. Consistent with immunolabeling studies, we did not detect *Ghrl*-expressing epsilon cells that are normally absent in the adult mouse pancreas. As previously reported[45,46,50,51], we also

observed a group of endocrine cells with high expression of multiple hormones. In addition to the previously identified *Gcg*$^{+}$*Ppy*$^{+}$ population[46], *Ppy* transcripts were detected at high levels (>10,000 reads/cell) in more than half of delta cells (n=71). These cells have a similar number of expressed genes and total reads as single-hormonal endocrine cells, indicating that they are likely not artifacts caused by doublets during FACS. Further analysis will be needed to understand the origin and functional significance of this unique population.

In addition, we detected the endocrine progenitor transcription factor *Neurog3* in a small number of endocrine cells (n=78, >100 reads/cell). Essential for islet cell differentiation, *Neuorg3* is believed to be expressed transiently in islet cell progenitors during mouse development and does not co-express with hormones[52]. However, our data reveal *Neurog3* expression mainly in somatostatin-producing delta cells in all mice tested at a level comparable to another well-established delta cell specific transcription factor *Hhex*, suggesting a potentially overlooked function of *Neurog3* in adult islets, consistent with a prior study[53]. Interestingly, *Neurog3*^{neg}*Hhex*$^{+}$, *Neurog3*$^{+}$*Hhex*^{neg}, and *Neurog3*$^{+}$*Hhex*$^{+}$ mark three distinct subsets of delta cells, which to our knowledge has not been reported before. Whether the heterogeneous gene profiles lead to functional heterogeneity awaits further analysis.

Furthermore, our data enables the discovery of novel cell-specific genes. For instance, *Prss53* encodes a protein with tandem serine-protease domains. Multiple microarray and bulk RNA-seq studies have shown the mRNA level of *Prss53* changed in islets of mice lacking important islet transcription factors such as *Pax6*[54,55]. However, the cellular distribution and function of PRSS53 is unexplored in the pancreas[56]. Our single-cell transcriptome analysis shows *Prss53* is exclusively detected in islet beta cells as one of the most significantly differentially expressed genes between beta and the other pancreatic cell clusters, consistent with the notion that beta cell specific transcription factor *Mafa* directly regulates *Prss53*[55].

Taken together, this inclusive pancreatic cell-sorting approach has generated a mouse single-cell transcriptome library representing almost all known pancreatic cell types, abundant or rare. This analysis suggests potential new roles for genes such as *Neurog3*, *Hhex*, and *Prss53* in the adult mouse. Future studies will delineate the functional consequences of these discoveries.

Large Intestine

Figures are located in the Tabula Muris Tissue Supplement under section(s):
Large Intestine FACS

As the central location of digestion, the intestines regulate nutrient and water uptake. Partially digested food called chyme enters the small intestine from the stomach, where it is exposed to digestive enzymes, many from the pancreas and liver. The majority of nutrients are absorbed here through densely packed villi, microscopic protrusions which greatly increase the intestinal surface area. The remaining chyme is then transported to the large intestine, consisting of the cecum, colon, and rectum. The colon is the major center of water resorption. Although it also intakes nutrients, it lacks villi. The colon also houses a diverse microbiota, microbes which assist digestion. These organisms are common in the colon and distal small intestine, but rare in the acidic, proximal small intestine. The proximal and distal colon also differ developmentally, with the midgut forming the proximal region, and the hindgut forming the distal region. These regions also display distinct histology and function, with the proximal colon serving more to absorb water, and distal to store feces. Colorectal cancers also differ between these regions, as do patient outcomes.

The large intestinal tract is divided into 4 layers: 1) the mucosa, which absorbs water and is composed of epithelium, lamina propria, and muscularis mucosa, 2) the submucosa, made of connective tissue that supports the mucosa, 3) the muscular layer, containing smooth muscle responsible for gut movement, and 4) the serosa, the outermost layer which secretes serous fluid, reducing friction with other organs. Although only the small intestines have villi, both the small

and large intestines contain crypts. The bottoms of these invaginations house intestinal stem cells which are critical for replenishing the constantly shed mucosal cells. As these stem cells differentiate, they emigrate to the top of the crypt to create several cell types: water absorbing enterocytes, mucus secreting goblet cells, hormone secreting chromaffin cells (enteroendocrine cells), and microvilli containing tuft cells which initiate an immune response after sensing pathogens[57]. Importantly, colonic crypts are also the source of colorectal carcinomas.

Tissue Processing

The large intestine was dissected from cecum to rectum and cut longitudinally in PBS to remove feces. After shaking with 5 mM EDTA in PBS briefly, the tissue was placed in fresh 5 mM EDTA in PBS and incubated at 37˚C for 5-10 minutes shaking to isolate crypts. Residual tissue was removed, and crypts were pelleted (200 x g, 2 minutes, 4˚C). Crypts were resuspended with serum-free medium (Advanced RPMI/F12, 10 mM HEPES, penicillin/streptomycin), then incubated for 60 minutes at 37˚C with trituration every 15 minutes. After centrifugation (500 x g, 5 minutes, 4˚C), cells were resuspended in FACS buffer (HBSS, 2% FBS, 10 mM HEPES, 1 mM sodium pyruvate, penicillin/streptomycin) with 100 U/mL DNase I (Worthington DP), and passed through a 40 µm strainer (Falcon 352340). Cells were then counted, pelleted (500 x g, 5 minutes, 4˚C), resuspended at 10^7 cells/mL in FACS buffer, and stained with 1:50 mCD45-Pacific Blue (Biolegend 103126), 1:50 mCD66a-PE (Biolegend 134506), 1:50 mCD326-PE-Cy7 (Biolegend 118216), and 1:20 mCD44-APC (Biolegend 103012) for 10 minutes on ice. Cells were washed and the pellet was resuspended at 10^6 cells/mL in FACS buffer with 1 µg/mL DAPI (Sigma 32670). We excluded immune cells and enriched for cells of the epithelial crypts ($EpCAM^+$, $CD45^-$), further gating on crypt-bottom cells ($CD44^+$), mid-crypt cells ($CD44^-CD66a^{low}$), and crypt-top cells ($CD44^-CD66a^{high}$), as previously shown[58]

Data analysis

The proximal colon and distal cecum was analyzed from 4 males and 3 females, and distal colon from 3 males and 3 females. In total, 3,938 cells formed 15 distinct clusters. As expected, cells of the distal and proximal colon separate distinctly, with distal cells lacking *Hoxb6* expression[59]. Three clusters contain enterocytes (*Krt20*$^+$, *Slc26a3*$^+$)[60,61], 1 from the distal colon, 2 from the proximal colon. While 5 clusters express goblet cell markers (*Atoh1*$^+$, *Spdef*$^+$)[62,63], 1 out of 3 distal clusters also expresses *Krt20,* indicating these goblet cells are derived from the crypt-top. Clusters of largely proximal tuft cells (*Dclk1*$^+$)[64], and chromaffin cells (*Chga*$^+$, *Chgb*$^+$)[65], are also apparent. Proliferating stem cells (*Lgr5*$^+$, *Mki67*$^+$)[66] segregate into 2 clusters: one composed mainly of proximal cells, and one composed mainly of distal cells which also express goblet cell markers (*Atoh1*$^+$, *Spdef*$^+$). These proliferating cells are distinct from the cluster of non-proliferating stem cells (*Lgr5*$^+$, *Mki67*$^-$). And although proliferating (*Lgr5*$^-$, *Mki67*$^+$) and non-proliferating (*Lgr5*$^-$, *Mki67*$^-$) progenitor populations are evident, the vast majority of these cells originate from the proximal colon.

In summary, we sorted and captured 3,938 cells of the murine large intestine, comprising enterocytes, goblet cells, tuft cells, chromaffin cells, and undifferentiated stem and progenitor cells. While tuft cells are rare, they form a distinct cluster, consistent with previous results[67]. We also detected the rare chromaffin cells. Furthermore, we identified *Hoxb13* as a new gene marker of the distal colon.

Microbiome

Figures are located in the Tabula Muris Tissue Supplement under section(s):
Microbiome

Stool was collected immediately prior to euthanasia and frozen at -80ºC. DNA extraction and library preparation was carried out by Microbiome Insights. DNA was extracted from stool using the MO Bio PowerSoil DNA Kit optimized for the KingFisher robot, with the addition of a bead-beating step. PCR amplification was performed using Phusion polymerase and dual-barcoded primers[68] targeting the V4 region of the 16S rRNA gene (Bacteria). PCR samples were cleaned and normalized using the high-throughput SequalPrep 96-well Plate kit (Thermo Fisher). Pooled libraries were quantified using the KAPA qPCR Library Quant kit (KAPA Biosystems).

QIIME v. 1.9.1[69] was used to perform quality filtering and demultiplexing. Reads were assigned to open-reference OTUs using UCLUST[70] with a threshold of 97% identity and seeded with Greengenes v. 13.8 database[71] sequences. The resulting OTU counts per sample were rarefied to a depth of 15956. Diversity was analyzed with QIIME using UniFrac[72] (PyNAST alignment[69]) and with Shannon and Chao1 metrics.

The V4 region of the 16S ribosomal rRNA gene was sequenced to assess gut microbiota composition within each mouse. Analysis of these communities revealed a predominance of two main phyla, the Bacteroidetes and Firmicutes, and to a lesser extent, a third phylum, Verrucomicrobia. The community compositions observed in the mice in this study are in line with other C57BL/6 mouse studies[73,74]. PERMANOVA analysis was conducted on calculated distance matrices and did not reveal significant differences between male and female mice ($p = 0.48$). This finding was corroborated by principal coordinate analysis of weighted Unifrac distances (a metric that incorporates relative abundances as well as phylogenetic distance[69]), which did not demonstrate clear separation of communities by sex. Alpha diversity (a measure of the diversity of microbes within a community) was calculated, and while there was greater variation in female mice, the mean alpha diversity as measured by Faith's PD did not differ significantly between male and female mice ($p=0.52$, two-tailed t-test).

Urinary System

Kidney

<u>Figures are located in the Tabula Muris Tissue Supplement under section(s):</u>
Kidney Droplet
Kidney FACS

The kidney regulates blood volume, pH, and salt concentration, and removes waste while reabsorbing nutrients. These waste products are then sent to the bladder as urine. These functions are enabled through the complex physiology of the nephron, the basic repeating unit of the kidney. The nephron consists of the renal corpuscle (composed of the glomerulus and Bowman's capsule), where blood is filtered, and tubules which use a variety of transporters to reabsorb water and nutrients while sending waste to the collecting duct[75]. The renal corpuscle is made up of three major cell types: fenestrated capillary endothelial cells composing the glomerulus, epithelial podocytes that filter the blood into the renal tubules, and mesangial support cells, which help clear debris and modulate flow rate through their contractile properties. The tubules can be subdivided into the proximal convoluted tubules, the loop of Henle, the distal convoluted tubules, and the collecting tubules. Although it is well established that these areas serve different physiological functions and show unique gene expression, little is known about the different cell types responsible for these functions.

<u>Tissue Processing</u>

Kidneys were minced with razors and dissociated in RPMI with 10 U Liberase TM enzyme (Roche 5401119001), 2% FBS, and 1X Antibiotic-Antimycotic (Gibco 15240-062), agitating for

30 minutes at 37°C. Following trituration with a 5 ml serological pipette, cells were ground through 100 µm, 70 µm, and 40 µm filters (Falcon 352340, 352350, 352360) with a syringe plunger. Cells were pelleted (541 x g, 10 minutes, 4°C), treated with ACK (Gibco A10492-01) for 5 minutes at 21°C, washed in 2% FBS and 1X Antibiotic-Antimycotic in RPMI, and debris was removed using a cell debris removal kit (Miltenyi 130-109-398). Cells were resuspended in 2% FBS in PBS and filtered into 35 µm FACS tubes (Falcon 352235) before staining with 1:500 PI (Life Tech P3655) immediately prior to sorting.

Data analysis

We profiled 2,781 kidney cells from 2 males and 1 female on the microfluidic droplet platform, and 519 cells from 4 males and 2 females following FACS for viability only. This approach captured cells from all tissue compartments, and clustering produced similar populations for both techniques. In the renal corpuscle for example, we identified clusters of fenestrated endothelial cells (*Plvap*$^+$) and mesangial cells (*Des*$^+$, *Vim*$^+$, *Acta2* $^+$). Although podocytes are present (*Podxl*$^+$, *Wt1*$^+$, *Mme*$^+$), they do not segregate from the other cell types. Several tubule cell clusters are also evident: proximal brush border cells (*Vil1*$^+$), and distal collecting duct principal cells (*Kcne1*$^+$, *Scnn1a*$^+$, *Aqp2*$^+$*)*. In both microfluidic and FACS approaches, epithelial cells are the most highly represented cells, followed by endothelial cells. Immune cells are the least numerous. Nominal cell types that were identified include proximal tubule cells, fenestrated endothelial cells, mesangial cells, distal convoluted tubule, collecting duct, thick ascending tubule, fibroblasts, and immune cells (NK cluster and myeloid cluster).

In conclusion, after analyzing 2,781 cell libraries prepared by microfluidic droplets, and 519 FACS-sorted cells, we were able to identify 11 cell types which have been previously described but not collectively studied in pure populations. Reassuringly, both methods produced similar clusters with relatively equal proportions of the different cell types. Interestingly, our data show possible sex differences, but only in a particular population, the proximal tubule cells. While sex-specific gene expression differences have been reported in the proximal and distal tubules previously[76], our data show much more robust differences in the proximal tubules, identifying dozen of differentially expressed genes. Because this was the only population to cluster by sex, these cells may be responsible for the observed ischemic tolerance and resistance to renal disease in female rodents and humans[76–78]. Possible sex differences should be interpreted with caution, however, due to the small sample size.

Bladder

Figures are located in the Tabula Muris Tissue Supplement under section(s):
Bladder Droplet
Bladder FACS

The urinary bladder is an organ of endodermal origin[79] lined by a specialized epithelium called urothelium, comprising umbrella, intermediate, and basal cell layers, with extracellular matrix-producing mesenchymal cells in the stromal compartment subjacent to the basal layer[80]. The umbrella cells of the luminal lining are often polyploid and express the transcription factor *Grhl3*[81] as well as uroplakins, which help form a water-tight barrier at the luminal surface[82]. The basal layer consists of smaller, undifferentiated cells characterized by expression of *Krt5*; a subset of basal cells also express *Krt14*[83]. Additionally, the basal cell layer harbors the urothelial stem cells that give rise to all urothelial cell types[83,84]. Intermediate cells between the basal and umbrella cell layers can express both basal and umbrella cell markers[83–85]. Importantly, the vast majority of bladder cancers are of epithelial origin[86], and mesenchymal cells play essential roles in regulating both urothelial repair and cancer progression[84,87,88]. Large scale single cell RNA sequencing of the urothelium and the underlying mesenchyme has not yet been described.

Tissue Processing

Bladders were cut above the bladder neck and inverted. The epithelial layer and associated stroma were mechanically teased away from the bladder muscle using forceps, and the epithelial/stromal sheet was then minced with butterfly shears. The minced bladder tissue was digested sequentially at 37ºC on an orbital shaker for 1 hour in Collagenase Type IV (Worthington LS004188) and DNase I (Worthington LS006343), followed by 30 minutes with 1X TrypLE (ThermoFisher A1217701) and DNase I. Cells were then filtered through a 40 µm strainer (Falcon 352340), pelleted (500 x g, 4ºC, 5 minutes), and resuspended in FACS buffer (1X Penicillin-Streptomycin (ThermoFisher 15140122), 1X Pluronic F-68 (ThermoFisher 24040032), and 2% FBS (Atlanta Biologicals S11550H) in PBS pH 7.4 (ThermoFisher 100100-23)). Cells were stained with 1:50 anti-CD45-Pacific Blue (Biolegend 103126), 1:50 anti-TER119-Pacific Blue (Biolegend 116232), 1:100 SCA1-FITC (Biolegend 122506) and 1:100 EpCAM-APC (ThermoFisher 17-5791-82). Cells were stained with 1:1000 Sytox Blue (ThermoFisher S34857) to identify live/dead cells immediately prior to sorting. Cells were sorted into 3 bins: $EpCAM^+$ $SCA\text{-}1^-$ epithelial cells, $EpCAM^+$ $SCA\text{-}1^+$ epithelial cells, and $EpCAM^-$ $SCA\text{-}1^+$ mesenchymal cells.

Data Analysis

Bladder single cell preparations enriched for epithelial and stromal cells and omitted the bladder smooth muscle. We captured and sequenced a total of 3,878 cells, 1,378 from 3 females and 3 males with FACS, and 2,500 from 1 female and 2 males with microfluidics. Both platforms identified the same number of urothelial and mesenchymal cell populations with reasonably good correspondence of specific epithelial and mesenchymal clusters between platforms. Because the microfluidic cell preparations were non pre-selected, they contained minor subpopulations of *Pecam1*$^+$ ($CD31^+$) endothelial cells and *Cd14*$^+$ immune cells.

Of the 3 bladder epithelial clusters characterized by high expression of *Epcam*, 2 express higher levels of uroplakin transcripts (e.g., *Upk3a* and *Upk1b*), and show higher expression of *Grhl3*, thus likely representing umbrella or intermediate cells. The third *Epcam*$^+$ cluster expresses basal markers such as *Krt5* and *Krt14*, thus suggesting basal cell identity. These results indicate the presence within the urothelium of a continuum of molecular phenotypes ranging from undifferentiated progenitors to differentiated umbrella cells, a result suggested previously using a limited panel of marker genes[85].

Three clusters of mesenchymal cells identified by high *Dcn* expression include 2 clusters expressing the marker gene *Car3*, and one cluster expressing *Scara5*. We noted that one mesenchymal cluster appears to comprise mostly female cells in both the microfluidic droplet and FACS data. Within epithelium, one luminal cluster and the basal cluster are largely male in the FACS data, but more mixed in the microfluidic data. Animal sex thus may contribute to differences between mesenchymal as well as epithelial cell clusters, perhaps consistent with androgen sensitivity of bladder tissues[89], but additional work will be required to confirm these sex differences and to determine whether they relate to the sexually dimorphic bladder cancer rates, which are 3-4 fold higher in males[90].

In summary, we have described the bladder urothelium and mesenchyme at the single-cell transcriptomic level. This has enabled a detailed characterization of the urothelial differentiation continuum, and has uncovered possible sex differences that may help explain sexually dimorphic cancer rates in humans.

Skeletal Muscle

Limb muscle

Figures are located in the Tabula Muris Tissue Supplement under section(s):
Limb Muscle Droplet
Limb Muscle FACS

Skeletal muscle is the largest organ of the body. Individual skeletal muscles are attached to bones by tendons and produce skeletal movement by the coordination of contraction and relaxation between muscle groups. Each muscle is composed of bundles of myofibers, and each myofiber is a multinucleated muscle cell derived from the fusion of myogenic progenitors during development. In adult animals, myofibers are post-mitotic and form the basic machinery for muscle contraction.

Despite the extremely low turnover rate of myofibers, skeletal muscle possesses very effective regenerative potential owning to the presence of resident muscle stem cells. The primary muscle stem cells, or satellite cells, reside underneath the basal lamina of individual myofibers in a quiescent state[91,92]. Acute muscle injury or chronic disease conditions that cause damage to the myofibers trigger the satellite cells to divide. These activated satellite cells give rise to myogenic progenitors which then fuse with and repair damaged myofibers. Other types of muscle-resident cells that have been documented to orchestrate the regeneration process include mesenchymal stem cells[93], endothelial cells[94], and resident macrophages and other immune cells[95]. These cells secrete factors that regulate the proliferation and differentiation of myogenic progenitors, and participate in the repair of vasculature and other connective tissues within the muscle.

Tissue Processing

Muscle cells collected from hindlimbs and forelimbs were prepared for FACS isolation as described[96]. In brief, muscles from each mouse were minced and digested in 10 ml 2 mg/ml collagenase II (Worthington LS004179) at 37°C for 1 hour with agitation followed by a second digestion with 1,000 Units of collagenase II and 11 Units of dispase (Thermo Fisher 17105-041) for 30 minutes. The digested tissues were then passed through a 20-gauge needle to release mononucleated cells. The resulting cell suspension was filtered with a 40 μm strainer (Falcon cat. 352340) and pelleted. Cells were then resuspended in 0.5 ml wash medium (Ham's F-10 (Thermo Fisher SH30025.01) supplemented with 10% (vol/vol) horse serum (Invitrogen 16050-122) and 1× penicillin-streptomycin (Omega Scientific PS-20)) and stained with 1:1000 CD31-APC (BioLegend 102510), 1:1000 CD45-FITC (BioLegend 103108), 1:250 Ly-6A/E-PB (BioLegend 108120), and 1:100 CD106-PE/Cy7 (BioLegend 105720) at 4°C for 30 minutes. Cells were washed once in wash medium before FACS isolation. Cells were sorted into 4 bins: satellite cells ($Vcam1^+$, $Sca1^-$, $CD31^-$, $CD45^-$), mesenchymal progenitors ($Sca1^+$, $CD31^-$, $CD45^-$), endothelial cells ($CD31^+$, $CD45^-$), and immune cells ($CD45^+$).

Data Analysis

Muscle single cell preparation enriched for satellite cells and other types of resident mono-nucleated cells, and depleted multinucleated muscle cells. 1,090 total mono-nucleated cells were sorted by FACS accordingly: satellite cells (*Vcam1*$^+$ [CD106]), mesenchymal stem cells (*Atxn1*$^+$ [Sca1]), and endothelial cells (*Pecam1*$^+$ [CD31]). As expected, these cells formed distinct clusters. Immune cells were isolated by *Ptprc* (CD45) expression and generated 3 clusters with differential expression of the monocyte/macrophage marker *Itgam*, the T cell marker *Cd3g*, and the B-cell marker *Cd19*.

Total mono-nucleated cells were also run on the microfluidic droplet platform without prior purification, revealing the relative ratios between muscle cell types. For instance, 1136 cells were identified as mesenchymal stem cells with adipogenic and fibrogenic potential, and 354 cells were identified as satellite cells. This is consistent with the 2-3:1 ratio between these two types of cells revealed by surface marker staining and FACS[96]. While the presence and function of macrophages in muscle have been previously described[97], T cells and B cells have not been identified locally in

the muscle. It is therefore surprising to identify comparable numbers of macrophages, T cells, and B cells by the microfluidic analysis. Further characterization is needed to understand the location and the role of these lymphocytes in muscle.

In addition to the cell clusters identified with both methods, the microfluidic droplet analysis revealed two additional clusters signified by their expression of *Chodl* and *Acta2*. *Chodl* is a well-known chondrocyte marker and cells expressing this marker have not been previously identified in muscle. It will be interesting to further characterize these cells to understand whether these are bona-fide chondrocytes or a subset of mesenchymal progenitors with chondrogenic potential.

Consistent with previously published single cell analysis[98], satellite cells can be sub-divided into two clusters in both our FACS and microfluidic analyses based on their differential expression of *Myod1* and *Calcr*. Given that *Calcr* expression has been shown to decrease in satellite cells when they activate in response to injury[99], and *Myod1* expression is generally believed to increase with cell activation, it is likely that the 2 sub-clusters represent satellite cells at different stages of the cell cycle.

In summary, we have generated a dataset of single cell transcriptomic data from skeletal muscle consistent with previous studies, but which has also identified novel muscle genes such as the chondrocyte marker *Chodl*. Future studies will elucidate the role of this gene and characterize the cells expressing it.

Diaphragm

Figures are located in the Tabula Muris Tissue Supplement under section(s):
Diaphragm FACS

The diaphragm is an essential skeletal muscle of mesodermal origin which separates the chest and abdominal cavities and enables respiration. It is comprised of three distinct domains: the crural muscle, the costal muscle, and the central tendon[100]. The muscle domains consist of myofibers that are surrounded by connective tissue and connect the central tendon to the ribs. These myofibers, which contain hundreds of post-mitotic nuclei with a common cytoplasm, contract and relax to control breathing. Although the myofibers are terminally differentiated, adult muscle stem cells give the diaphragm regenerative potential. Muscle stem cells, also called satellite cells due to their juxtaposed position on the myofibers, exist in a quiescent state in healthy tissue. Upon injury, however, these cells activate, enter the cell cycle, and give rise to myoblasts that fuse to form new myofibers[101]. Satellite cells are molecularly defined by the expression of the transcription factor *Pax7*. The diaphragm contains several other populations of mononuclear cells that are thought to aid in the regenerative process, including mesenchymal stem cells, also called fibro-adipogenic progenitors, endothelial cells, and resident immune cells[93–95].

Tissue Processing

The diaphragm muscle was isolated in one piece by cutting along the inner side of the ribcage. It was then washed in ice cold wash medium (10% horse serum (Invitrogen 16050114) and 1% pen/strep (Omega Scientific PS-20) in Ham's F-10 medium (HyClone SH30026.FS)), dried on tissue, minced with scissors, and digested in dissociation buffer (0.2% w/v Collagenase II (Worthington LS004177) in wash medium) at 37˚C for 35 minutes with shaking. After digestion, cells were pelleted (1,600 x g, 5 minutes, 4˚C) and resuspended in PBS with 0.1% Dispase (Gibco 17105-041) and 0.05% Collagenase II, and shaken at 37˚C for 20 minutes. Samples were then passed through a 20' needle (Fisher 305175) 5 times, filtered through a 40 μm strainer (Falcon 352340), pelleted (1,600 x g, 5 minutes, 4˚C), and resuspended in wash medium for staining. Antibodies were added at 1:100 - Sca1-PB (Biolegend 108120), CD31-FITC (Biolegend 102506), CD45-APC (Biolegend 103112), VCAM-PECy7 (Biolegend 105720), and shaken at 4˚C for 25

minutes. Following centrifugation (1,600 x g, 5 minutes, 4˚C), cells were resuspended in wash medium and filtered into 35 µm FACS tubes (Falcon 352235). Cells were sorted into 4 bins: skeletal muscle satellite cells (Sca1$^-$, CD31$^-$, CD45$^-$, VCAM$^+$), mesenchymal progenitors (Sca-1$^+$, CD31$^-$, CD45$^-$), immune cells (CD45$^+$, CD31$^-$), and endothelial cells (CD31$^+$, CD45$^-$).

Data Analysis

We isolated diaphragms from 2 male and 2 female mice and dissociated them into single cell suspensions depleted of multinucleated myofibers[96]. The cells were stained to identify 4 major mononucleated cell populations, and we analyzed 870 total cells including satellite cells, mesenchymal stem cells, endothelial cells, and immune cells. Following clustering, these populations separated distinctly. Two clusters of satellite cells defined by expression of known muscle stem cell genes (*Pax7*, *Pax3*, *Myod1*, *Myf5*) are present, as are clusters of fibro-adipogenic progenitors (*Pdgfra*$^+$), endothelial cells (*Pecam1*$^+$), B and T cells (*Cd19*$^+$, *Cd3d*$^+$), and macrophages (*Itgam*$^+$, *Fcer1g*$^+$, *C1qa*$^+$).

To our knowledge, this is the first published scRNA-seq dataset from the diaphragm. Consistent with published scRNA-seq datasets of satellite cells isolated from hind limb muscles[98,102], these cells separate into two clusters that show overlap of many classical markers, including all surface markers that can be used to purify satellite cells: *Vcam1*, *Itga7*, *Sdc4*, *Cd34*, *Itgb1*, and *Cxcr4*[96]. Not all cells that cluster among the satellite cells have detectable *Pax7* expression (27% in our dataset), which is consistent with published scRNA-seq data sets and prior observations that *Pax7* is expressed at low levels[98,102–104]. This is therefore likely an underestimation of the actual number of *Pax7*$^+$ cells.

Integumentary System

Skin

Figures are located in the Tabula Muris Tissue Supplement under section(s):
Skin FACS

Human and murine skin is composed of three distinct layers: the hypodermis, or subcutaneous fat, which contains adipocytes and connective tissue; the dermis, a highly vascularized layer chiefly composed of collagen and elastin, but also containing fibroblasts, adipocytes, immune cells, and nerve endings; and the outermost layer, the epidermis, which contains not only keratinocytes forming a physical barrier to external pathogens, but also epidermal appendages which extend into the dermis: sweat glands, and hair follicles containing oil-producing sebaceous glands.

The epidermis is typically categorized into two parts, the interfollicular epidermis (IFE), and the hair follicles (HFs). The IFE comprises the majority of the epidermis, and contains 5 distinct layers. The *stratum basale*, adjacent to the dermis, is typically a single cell layer of keratinocytes, stem cells which continuously self-renew and generate more committed progeny which migrate distally to the *stratum spinosum*. This thicker layer is characterized by a high concentration of keratin, as well as cell-anchoring desmosomes. Upon further differentiation, cells form the *stratum granulosum*, which produces lipids that act as a water sealant. In thicker regions of skin, the *stratum lucidum* containing dead cells is evident. Finally, the most superficial layer, the *stratum corneum*, also contains dead keratinocytes providing a physical barrier to the external environment. These cells are continually shed and replaced by cells migrating from the more basal layers.

In addition to the IFE, the epidermis also contains well-defined specialized niches: the hair follicles. HF stem cells, housed in the bulge region, differentiate as they migrate proximally, and are even capable of reconstituting the IFE upon injury. Importantly, the HF morphology and gene

expression changes with the hair growth cycle. During telogen, the resting phase, the HFs are contracted, and the skin of C57BL/6 appears grey. During active hair growth, anagen, the follicles are enlarged and the skin appears black. In adult mice, anagen occurs in distinct dark patches.

Tissue Processing

After scraping off the subcutaneous fat layer with a scalpel, we incubated mouse back skin in 0.25% trypsin (Thermo 15050057) for 30 minutes at 37°C with gentle agitation. The epidermis was then scraped off the dermis, releasing keratinocytes into a single cell suspension. After 5 minutes of additional trypsin digestion, cells were filtered through 70 μm and 40 μm cell strainers (Falcon 352350, 352340) on ice and washed with cold PBS with 5% FBS. Keratinocytes were then stained with 1:50 CD34-AlexaFluor 647 (BD Pharmingen 560230), 1:200 Integrin 6 [CD49f]-FITC (BioLegend 313606) for 30 minutes at 4°C in PBS with 5% FBS and washed. 1:1000 SytoxBlue (Invitrogen S34857) was added to exclude dead cells prior to sorting. Cells were sorted into 3 bins: interfollicular epidermis ($CD34^{low}$, $ITGA6^{high}$), HF inner bulge stem cells ($CD34^{high}$, $ITGA6^{low}$), and HF outer bulge stem cells ($CD34^{high}$, $ITGA6^{high}$).

Data Analysis

Here, the epidermis was separated into telogen and anagen prior to dissociation into single cell suspensions. Telogen skin was analyzed from 3 males and 2 females, and anagen from 4 males. From the FACS data, 2,310 cells were analyzed: 891 cells of the IFE and upper follicle, 831 HF outer bulge cells, and 573 HF inner bulge cells. After clustering and t-SNE, the HF bulge populations separated distinctly from IFE clusters, as evidenced by high *Cd34* expression. Outer bulge cells ($Dkk3^{high}$) clustered separately from inner bulge cells ($Fgf18^{high}$). Multiple IFE clusters are also evident, consistent with stem cells ($Krt14^{high}$, $Krt10^{low}$ basal IFE), actively dividing stem cells ($Top2a^{high}$ replicating basal IFE), and more differentiated progeny ($Krt10^{high}$ intermediate IFE).

We sorted and captured 2,310 cells of the murine epidermis, comprising predominantly cells of the hair follicle bulge and interfollicular epidermis. Consistent with previous publications, many classical spatial markers of the epidermis showed overlap between clusters[25]. And while this study was unable to resolve the subtle heterogeneity of bulge stem cells and transit amplifying cells characterized elsewhere[26], it does allow for intra- and inter-animal comparisons of telogen and anagen epidermis.

Mammary gland

Figures are located in the Tabula Muris Tissue Supplement under section(s):
Mammary Gland Droplet
Mammary Gland FACS

The mammary gland is a hormone responsive organ that exists as a rudimentary structure during birth. During puberty, hormones signal breast development, and the stroma signals primordial epithelial cells to form a mature ductal network[105]. The mammary epithelial tree consists of two main cell lineages – luminal cells, and basal myoepithelial cells adjacent to the basement membrane. The basal layer also contains the putative mammary epithelial stem cell that is capable of giving rise to both lineages upon transplantation[106,107]. During the estrus cycle, the epithelial tree, and the resident stem cells, respond to estrogen and progesterone changes[108,109]. Furthermore, the mammary epithelium undergoes extensive proliferation and remodeling during pregnancy to give rise to the mature milk producing cells during lactation. Following lactation, the epithelial cells again undergo remodeling and apoptosis to return to the pre-pregnancy state. Deciphering the composition of the mammary fat pad is therefore crucial to our understanding of mammary epithelial biology, including breast cancer progression.

Tissue Processing

Fat pads 2, 3 and 4 were surgically resected, mechanically dissociated, and digested in DMEM/F12 for 2 hours using collagenase and hyaluronidase (StemCell Technologies 07912). Red blood cells were lysed in ACK lysis buffer (Lonza 10-548E) for 5 minutes, washed with FACS buffer (2% FBS and Penicillin-Streptomycin (100 U/ml) in HBSS), treated with pre-warmed 0.25% trypsin-EDTA (Invitrogen 25200114) for 1-2 minutes, washed, and treated with a pre-warmed mixture of Dispase (StemCell Technologies 07913) and DNase I (Worthington LS002139) for 2 minutes. Cells were then filtered through a 40 μm cell strainer (Falcon 352340) and washed with FACS buffer. Cells were counted and stained with CD45 (Biolegend 103126, Clone 30-F11), CD31 (Biolegend 102422, Clone 390), Ter-119 (Biolegend 116232, Clone TER119), CD24 (Biolegend 101822, Clone M1/69), and CD49f (Biolegend 313616, Clone GoH3), all at 1:100. Cells were subsequently washed with FACS buffer and resuspended in FACS buffer containing DAPI (Sigma 39542, 1:10,000). Cells were sorted into 4 bins: basal cells ($CD45^-$, $CD31^-$, $TER119^-$, $CD49f^{high\text{-}med}$, $CD24^{med\text{-}low}$), luminal cells ($CD45^-$, $CD31^-$, $TER119^-$, $CD49f^{med\text{-}low}$, $CD24^{high\text{-}med}$), mammary repopulating cells ($CD45^-$, $CD31^-$, $TER119^-$, $CD49f^{high}$, $CD24^{med}$), stromal cells ($CD45^-$, $CD31^-$, $TER119^-$, $CD49f^-$, $CD24^-$).

Data Analysis

Here, we processed mammary fat pad cells with microfluidic droplets from 2 mice, and we sorted epithelial cells with FACS from 4 mice. The microfluidic droplets yielded 4,481 cells, the majority of which are immune cells (~61%). This includes T-cells (*Cd3*$^+$), B-cells (*Cd74*$^+$), and macrophages (*Cd74*$^+$, *Cd14*$^+$, *Csf1r*$^+$). We also detected endothelial cells (*Esam*$^+$, *Pecam1*$^+$) and stromal cells (*Vim*$^+$, *Fn1*$^+$). Mammary epithelial populations comprising 16% of the total cells are also present. We classified the clusters using well established markers of mammary epithelial cells[110–112]. These include basal cells (*Krt14*$^+$, *Krt5*$^+$, *Krt17*$^+$) and luminal cells (*Krt8*$^+$, *Krt18*$^+$, *Krt19*$^+$). The luminal cells can be further separated into luminal progenitors (*Krt8*$^+$, *Cd14*$^+$, *Aldh1a3*$^+$) and hormone responsive luminal cells (*Krt8*$^+$, *Esr1*$^+$, *Pgr*$^+$), as previously reported[110,111,113].

To specifically study the transcriptional heterogeneity and sub-structure of the epithelial compartment within the mammary gland, we sorted luminal ($CD24^+$, $CD49f^{med\text{-}low}$), basal ($CD24^{med\text{-}low}$, $CD49f^+$), and stromal cells ($CD49f^-$, $CD24^-$). We then analyzed 2,405 cells, half of which are basal cells that contain the mammary stem cells and can be further subdivided[106,107]. Consistent with our microfluidic droplet data, we identified the hormone responsive luminal cells and luminal progenitors. We also observed a small population of *Cd55*$^+$ luminal progenitors as recently reported[114].

We then subclustered the FACS basal cells, which formed 4 groups based on differentially expressed genes: *Vcam1*$^+$/*Procr*$^+$[115], *Id1*$^+$/*Id2*$^+$, *Fos*$^+$, *Igfbp2*$^+$/*Igfpb4*$^+$. Although we attempted to identify specific basal cell clusters expressing luminal genes based on previously described markers[114], *Prlr*, *Csn3*, *Cited1*, and *Areg* were expressed sporadically.

In summary, we analyzed 6,886 single cells of the mammary fat pad, both in an unbiased fashion with microfluidic droplets, and enriching for mammary gland epithelium with FACS. Consistent with previously published single cell data, we identified the major epithelial compartments using classical epithelial markers[114]. In addition, we captured a wide variety of immune cells and heterogeneity within the stromal compartment of the mammary fat pad that has not been characterized previously at the single cell level. We also captured a large number of basal cells in order to study the stem cell compartment and can identify novel subpopulations including a population expressing the stem cell marker *Procr*[115]. Further functional analysis of these basal cell populations will help to understand the heterogeneity of mammary epithelial stem cells, and provide novel insights into mammary gland biology.

Adipose Tissues

Fat

Figures are located in the Tabula Muris Tissue Supplement under section(s):
Fat FACS

Fat is situated in specific depots throughout the body, and is classified as white adipose tissue (WAT) or brown adipose tissue (BAT). WAT is further divided into 2 types: subcutaneous (e.g. inguinal) and visceral (e.g. gonadal and mesenteric). These depots function not only as energy reservoirs, but they maintain organismal metabolic homeostasis by responding to insulin, releasing leptin, and modulating systemic inflammation[116]. Interestingly, these different depots play very different roles, both in healthy and in diseased or obese individuals. For example, BAT is primarily located in the interscapular space and is extremely metabolically active. With densely packed mitochondria that produce heat instead of ATP, brown fat regulates thermogenesis. WAT types also play distinct roles, with inflammatory visceral depots promoting obesity, and protective subcutaneous depots sustaining metabolic health.

Although adipocytes store fat and provide many of the aforementioned functions[117], adipose tissue contains many more resident cell types termed the stromal vascular fraction (SVF). These include mesenchymal progenitors (MPs), immune cells (ICs), and endothelial cells (ECs). MPs have been intensely investigated over the past decade and they include a subpopulation of $CD24^-$ adipocyte precursors. These adipogenic lineage preadipocytes are thought to arise from $CD24^+$ adipogenic progenitors of the MP population[118]. The ICs have also attracted attention due to their potential role in the inflammatory response related to obesity. Both innate and adaptive immune cells are present in fat depots and have an integral role in regulating adipocyte functions[119]. Similarly, the vascular ECs have a multitude of functions in fat depots. ECs are central to communication between local and distant cells. They also secrete growth factors to induce adipocyte hyperplasia and hypertrophy, while also producing cytokines to influence other non-adipogenic constituents. Additionally, ECs in fat are also potentially capable of differentiating into preadipocytes[120].

Tissue Processing

Fat tissues including inguinal subcutaneous adipose tissue (SCAT), gonadal adipose tissue (GAT), mesenteric adipose tissue (MAT), and interscapular brown adipose tissue (BAT), were dissected out and minced before digestion in 760 U/mL Collagenase II (Worthington LS004177) and Dispase II (Gibco 17105-041, 1U/ml), shaking for 30 minutes at 37°C. After trituration, cells were filtered consecutively through 100 µm (Falcon 352360) and 40 µm (Falcon 352340) strainers on ice and washed with cold F-10/Ham's medium containing 10% horse serum (Invitrogen 16050114). Cells were then stained with 1:100 SCA1-APC (Biolegend 122512), 1:100 CD31-FITC (Biolegend 102506), 1:500 CD45-PE/Cy7 (eBioscience 25-0451-82) for 30 minutes at 4°C in F-10/Ham's with 10% horse serum and washed. 1:1000 SytoxBlue (Invitrogen S34857) was added immediately prior to sorting. Cells were sorted into 3 bins: MPs ($Sca\text{-}1^+$, $CD31^-$, $CD45^-$), ICs ($CD45^+$, $CD31^-$), and ECs ($CD31^+$, $CD45^-$),

Data Analysis

SCAT, GAT, MAT, and BAT were isolated from 4 males and 3 females, resulting in 4,967 single cell libraries. Following clustering of all fat types combined, these populations separated distinctly: MPs (*Pdgfra*$^+$, *Cd34*$^+$)[116,121], ECs (*Pecam1*$^+$, *Cdh5*$^+$, *Cd34*$^+$)[122], and ICs (*Ptprc*$^+$). ICs further separated into myeloid cells (*Itgam*$^+$, *Cd68*$^+$, *C1qa*$^+$, *Lyz2*$^+$, *Cd14*$^+$), B cells (*Cd19*$^+$), T cells (*Cd3g*$^+$, *Cd8a*$^+$, *Cd4*$^+$) and NK cells (*Irf8*$^+$, *Gzma*$^+$, *Nkg7*$^+$)[119].

Interestingly, cells from BAT are largely absent within the MP population, and within one population of myeloid cells. We were also surprised to find no appreciable *Cd24* expression in MPs, as this marker typically labels adipogenic progenitors in this population[118]. However, we did detect Tie2 (*Tek*), a marker previously used to identify multipotent MPs in other organs, in a subpopulation of MPs. This suggests that these cells represent the multipotent population, and possibly provide a marker for further fractionation[123]. Furthermore, *Fabp4* expression in SVF MPs has been a matter of debate; our results confirm the presence of this gene in a subpopulation of MPs[124]. We also detected *Pecam1*$^{+}$*/Ptprc*$^{+}$ cells, suggesting these are endothelial progenitor cells of hematopoietic origin[125].

In summary, we have sequenced the stromal vascular fraction originating from both brown and white, and both visceral and subcutaneous fat depots. These data provide insight into the gene expression of known and purported markers, and will also allow the comparison of immune, endothelial, and multipotent progenitors between fat tissues and other tissues throughout the body.

Immune System

Bone marrow

Figures are located in the Tabula Muris Tissue Supplement under section(s):
Bone Marrow Droplet
Bone Marrow FACS

The bone marrow (BM) is the major site of hematopoiesis in adult vertebrates. Here, the hematopoietic system is organized in a hierarchical fashion with self-renewing, multipotent hematopoietic stem cells (HSCs) at the top of the hierarchy[126]. These HSCs are capable of giving rise to all mature blood cell types, including red blood cells (RBCs), platelets, and all innate (granulocytes and monocytes) and adaptive immune cells (B, T, and NK cells), throughout life[127,128]. HSCs give rise to these mature functional blood cells by differentiating into increasingly specialized and lineage restricted progenitors, which then progressively branch out and terminally differentiate into multiple effector cell types[126,129]. The resulting progenitor and effector cells are classified into two major lineage branches: myeloid and lymphoid[130,131].

Downstream of HSCs, the myeloid branch in the hematopoietic tree begins with the common myeloid progenitor (CMP), which can bifurcate into two major types of myeloid progenitors. The first of these progenitors is the megakaryocyte-erythroid progenitors (MEPs). MEPs further differentiate into (1) megakaryocyte progenitors (MkPs) that give rise to platelets, and (2) erythrocyte progenitors (EPs) that differentiate into erythroblasts that are nucleated precursors of RBCs[126]. The second major myeloid progenitor subset is the granulocyte-macrophage progenitors (GMPs) that differentiate into (1) monocytes, which are the precursors to BM-derived macrophages, (2) dendritic cells, and (3) granulocytes, which include neutrophils, basophils, and eosinophils[130].

Similar to the myeloid branch, HSCs also give rise to the lymphoid branch where the earliest progenitors are the common lymphoid progenitors (CLPs), which give rise to the three major classes of adaptive immune cells: (1) T cells, (2) B cells, and (3) natural killer (NK) cells[131]. Each of these classes of hematopoietic cells are seemingly homogeneous as per their immunophenotype—i.e. proteins expressed on their cell-surface. However, we and others have shown that these cell-types are functionally heterogeneous through cell transplantation experiments[132,133]. These experiments have illustrated that the hematopoietic subsets, particularly within the stem and progenitor population, are heterogeneous and dynamic, occupying distinct

states. This heterogeneity was previously indiscernible since these cell-types were isolated and studied in bulk. The advent of high-resolution techniques, such as scRNA-seq, allows studying cellular heterogeneity within these seemingly homogenous cell types at the single-cell level. Our scRNA-seq analysis of the BM aims to provide a roadmap and a resource where researchers can query molecular details that may explain the inherent single cell diversity related to cellular identity, fate-bias during differentiation, and possibly, multi-lineage reconstitution potential, which could prove useful for developing cell-transplantation based therapies.

Tissue Processing

Isolated bones from forelimbs, hindlimbs, hips, and vertebrae were crushed in FACS buffer (2% sterile FBS in PBS) on ice, filtered through a 100 µm strainer (Falcon 352360), pelleted (300 x g, 4°C, 5 minutes), and resuspended in FACS buffer with 100 U/ml DNase I (Worthington LS006344) and 20 ug/ml Rat-IgG (Abcam ab37361) for 10 minutes on ice. Cells were incubated with anti-CD117 MicroBeads (Miltenyi Biotec 130-091-224) on ice for 20 minutes, pelleted, resuspended in FACS buffer, and filtered through a 40 µm strainer (Falcon 352340) into an LS column (Miltenyi Biotec 130-042-401). Following magnetic purification, c-Kit$^+$ cells were eluted, and the flow through was depleted of RBCs with Histopaque-1119 (Sigma-Aldrich RNBF0417). The c-Kit$^+$ cells and RBC-depleted flow through were resuspended in FACS buffer with Rat-IgG (20 ug/mL) and incubated on ice for 10 minutes. The flow-through was stained for B-cells, T-cells, and granulocytes. All antibodies were used at 1:50.

c-Kit$^+$ cells were stained with CD3-FITC (BioLegend 133301, clone 17A2), Ly-6G/Ly-6C/GR1-FITC (BioLegend 133301, clone RB6-8C5), CD11B/MAC1-FITC (BioLegend 133301, clone M1/70), CD45R/B220-FITC (BioLegend 133301, clone RA3-6B2), TER119-FITC (BioLegend 133301), CKIT-APC (BioLegend 105812, clone 2B8), SCA-1-PECy7 (BioLegend 122514, clone E13-161.7).

B-cells were stained with TER119-PECy5 (BioLegend 116210), B220-FITC (BioLegend 103206, clone RA3-6B2), IgM-PECy7 (BioLegend 406514, clone RMM-1).

T-cells were stained with TER119-PECy5 (BioLegend 116210), CD90/Thy1.1-AF488 (BioLegend 202506, clone OX-7), CD90/Thy1.2-FITC (BioLegend 140304, clone 53-2.1), CD2-PECy7 (BioLegend 100114, clone RM2-5).

Granulocytes were stained with TER119-PECy5 (BioLegend 116210), CD11B/MAC1-FITC (BioLegend 101206, clone M1/70), Ly-6G/Ly-6C/GR1-PECy7 (BioLegend 108416, clone RB6-8C5).

Cells were incubated in antibody for 20 minutes, washed with FACS buffer, and pelleted (300 x g, 4°C, 5 minutes). After resuspension in FACS buffer, cells were filtered into 35 µm FACS tubes (Falcon 352235) and stained with 1:1000 SYTOX Blue (ThermoFisher S34857) immediately prior to sorting.

Data Analysis

Bone marrow samples were analyzed as fractionated and unfractionated pools. We used microfluidic droplets to sequence whole bone marrow cells (depleted of RBCs through Ficoll separation). For a higher resolution view of the various populations in the bone marrow, we subfractionated the BM using FACS into five different populations: (1) Lin$^-$Kit$^+$Sca1$^+$ HPCs that are known to include rarer long-term HSCs and much more abundant short-term multipotent progenitors (MPPs), (2) Ter119$^-$B220$^+$ cells that include both immature and mature B cells, (3) Ter119$^-$Cd90$^+$Cd2$^+$ cells that include immature T cells and NK cells, (4) Ter119$^-$Mac1$^+$Gr1^{Hi} cells that include predominantly granulocytes and (5) TER119$^-$Mac1$^+$Gr1$^{Mid/Lo}$ cells enriched in monocytes. From both types of analysis, we found the following similarities:

1. From the whole bone marrow, we identified T cells (*Ahnak*$^+$, *Thy1*$^+$, *Cd3e*$^+$, *Cd8*$^+$), NK cells (*Kirb1a*$^+$, *Kirb1b*$^+$, *Kirb1c*$^+$, *Ncr1*$^+$), and NKT cells that express both T and NK gene expression programs. The in-depth view of Ter119$^-$Cd90$^+$Cd2$^+$ cells further subdivided

these populations into immature T cells (*Cd3e*$^{+}$, *Cd4*$^{+}$, *Cd8a*$^{+}$, *and Cd6*$^{+}$), immature NK cells (*Kirb1a*$^{+}$, *Kirb1b*$^{+}$, *Kirb1c*$^{+}$, *Ncr1*$^{+}$), and immature NKT cells that express both T and NK cell genes. We also identified a proliferating pre-NK cell (*Mki67*$^{+}$, *Stmn1*$^{+}$), which may be a precursor to immature NK cells.

2. Both the unfractionated marrow and fractionated marrow yielded four unique populations. In the whole bone marrow, we find populations similar to early pro-B cells (*Dntt*$^{+}$, *Pax5*$^{-}$, *Rag1*$^{+}$, *Rag2*$^{+}$) and late pro-B cells (*Dntt*$^{-}$, *Pax5*$^{+}$, *Rag1*$^{+}$, *Rag2*$^{+}$). More downstream B cells were found that can be further subclustered into immature B cells (*Chchd10*$^{+}$, *Cd79a*$^{+}$, *Cd79b*$^{+}$, *Cd19*$^{+}$, *Ms4a1*$^{-/lo}$, *Cd74*$^{-}$, *Mki67*$^{+}$, *Stmn1*$^{+}$), and mature (naïve) B cells (*Chchd10*$^{-}$, *Cd79a*$^{+}$, *Cd79b*$^{+}$, *Cd19*$^{+}$, *Ms4a1*$^{+}$, *Cd74*$^{+}$, *Mki67*$^{-}$, *Stmn1*$^{-}$). In the sorted Ter119-B220+ fraction, we identified a late pro-B cell (*Dntt*$^{+/-}$, *Vpreb1*$^{+}$, *Pax5*$^{+}$, *Rag1*$^{+}$, *Rag2*$^{+}$, *Cd19*$^{+}$, *Cd20*$^{-}$, *Cd22*$^{-}$), a more downstream intermediate pre-B cell (*Dntt*$^{-}$, *Vpreb1*$^{-}$, *Pax5*$^{+}$, *Rag1*$^{+}$, *Rag2*$^{+}$, *Cd19*$^{+}$, *Chchd10*^{hi}, *Cd20*$^{-}$, *Cd22*$^{-}$, *Cd74*^{lo}), an immature B cell subset (*Dntt*$^{-}$, *Vpreb1*$^{-}$, *Pax5*$^{+}$, *Rag1*$^{-}$, *Rag2*$^{-}$, *Cd19*$^{+}$, *Chchd10*^{hi}, *Cd20*$^{+}$, *Cd22*$^{+}$, *Cd74*^{hi}) and a more mature (naïve) B cell subset (*Dntt*$^{-}$, *Vpreb1*$^{-}$, *Pax5*$^{+}$, *Rag1*$^{-}$, *Rag2*$^{-}$, *Cd19*$^{+}$, *Chchd10*^{lo}, *Cd20*$^{+}$, *Cd22*$^{+}$, *Cd74*^{hi}). The late-pro B cell could be further subdivided into *Dntt*$^{+}$, and *Dntt*$^{-}$ late pro-B cells.

Each approach also generated more nuanced data. From the unfractionated marrow, our analysis shows the hematopoietic precursor cells (HPCs) (*Kit*$^{+}$, *Stmn1*$^{+}$, *Mki67*$^{+}$) that primarily includes the HSCs and the early hematopoietic progenitors[134].

1. We identified distinct clusters of proerythroblasts (*Beta-s*$^{+}$, *Hbb-b2*$^{+}$, *Tfrc*$^{+}$, *Mki67*$^{+}$, *Stmn1*$^{+}$, *Bpgm*$^{-}$) and erythroblasts (*Beta-s*$^{+}$, *Hbb-b2*$^{+}$, *Tfrc*$^{-}$, *Mki67*$^{-}$, *Stmn1*$^{-}$, *Bpgm*$^{+}$) subsets. The proerythroblasts (*Beta-s*$^{+}$, *Hbb-b2*$^{+}$, *Tfrc*$^{+}$, *Mki67*$^{+}$, *Stmn1*$^{+}$, *Bpgm*$^{-}$) cluster can be seen diffusing into a subpopulation of downstream, more differentiated erythroblasts (*Beta-s*$^{+}$, *Hbb-b2*$^{+}$, *Tfrc*$^{-}$, *Mki67*$^{-}$, *Stmn1*$^{-}$, *Bpgm*$^{+}$).
2. We find proliferating granulopoietic cells (*Ltf*$^{+}$, *Pglyrp1*$^{+}$, *Lcn2*$^{+}$, *Camp*$^{+}$, *Mki67*$^{+}$, *Stmn1*$^{+}$) downstream of GMPs that are most likely present in the HPC cluster, and granulocytes (*Ltf*$^{+}$, *Pglyrp1*$^{+}$, *Lcn2*$^{+}$, *Camp*$^{+}$, *Mki67*$^{-}$, *Stmn1*$^{-}$). We also identified a distinct small cluster of basophils (*Ccl3*$^{+}$, *Fcer1a*$^{+}$, *Mcpt8*$^{+}$) based on our previous results[134].
3. Proliferating promonocytes (*Ahnak*$^{+}$, *Mpeg1*$^{+}$, *Emr1*$^{-}$, *Cd68*$^{-}$, *Mki67*$^{+}$, *Stmn1*$^{+}$) are also present, again, downstream of the GMPs that are most likely present in the HPC cluster. These are transitioning into more mature monocytes (*Ahnak*$^{+}$, *Mpeg1*$^{+}$, *Emr1*$^{+}$, *Cd68*$^{+}$, *Mki67*$^{-}$, *Stmn1*$^{-}$, *Ly6d*$^{-}$, *Irf8*^{lo}, *Cd74*^{lo}) and BM macrophages (*Ahnak*$^{+}$, *Mpeg1*$^{+}$, *Emr1*$^{+/-}$, *Cd68*$^{+}$, *Mki67*$^{-}$, *Stmn1*$^{-}$, *Ly6d*$^{+}$, *Irf8*^{hi}, *Cd74*^{hi}). We can see a subcluster within the monocyte cluster that are *Cd74*^{hi} (indicating the presence of major histocompatibility complex, class II antigen [MHC-II]), most likely in the process of transitioning to a more mature bone marrow resident macrophage population.

From subfractionation using FACS, we identified 13 major clusters in our single-cell dataset.

1. The largest group of cells are the HPC clusters (*Kit*$^{+}$, *Gpr56*$^{+}$, *Cd34*$^{+}$) that separated into multiple subpopulations.
    a. <u>Subclustering of HPC cluster</u>: To further analyze transcriptional profile of HPCs, we gated out clusters characterized by *Kit* expression. Then, we analyzed cellular heterogeneity among these cells and distinguished progenitor populations: a quiescent HSC-enriched fraction (*Stmn1*$^{-}$, *Mki67*$^{-}$, *Cd48*^{lo}, *Flt3*^{lo}, *Pdzk1ip1*^{hi}, *Ly6a*^{hi}), and cells that are immediately downstream of HSCs—the MPPs (*Stmn1*$^{+}$, *Mki67*$^{+}$, *Cd48*^{hi}, *Flt3*^{hi}, *Cd34*^{hi}). The MPPs could be divided into *Slamf1*^{hi} and *Slamf1*^{lo} populations. Next, we detected more committed fractions of progenitor cells that express markers corresponding to megakaryocyte erythrocyte progenitor

(MEP) (*Tgm2*$^{+}$, *Klf1*$^{+}$, *Treml1*$^{+}$), and Common lymphocyte progenitor lineage (CLP) (*Dntt*$^{+}$, *Ly6c2*$^{+}$, *Ctsg*$^{+}$, *Tyrobp*$^{+}$, *Cd69*$^{+}$).

2. Amongst the myeloid cells, we identified one population enriched for granulocyte-monocyte progenitors (GMP) (*Flt3*$^{+}$, *Kit*$^{+}$, *Mpeg1*$^{+}$, *Itgb2*$^{+}$, *Ahnak*$^{+}$, *Pld4*$^{+}$, *Cd68*$^{+}$, *Hp*$^{-}$), and several more mature (*Kit*$^{-}$) downstream populations—monocytes (*Cd68*$^{+}$, *Mpeg1*$^{+}$, *Itgam*$^{+}$, *Emr1*$^{+}$, *Fcgr3*$^{+}$) and macrophages, which in addition to all monocyte markers also expressed MHC-II (*Cd74*^{hi}). Surprisingly, we found the expression of *Cd4* on the cluster identified as macrophages—we have recently shown that infiltrating macrophages in a tumor tissue can express CD4[135]. In addition, we identified two different subsets of granulocytes (*Ltf*$^{+}$, *Camp*$^{+}$, *Ngp*$^{+}$, *S100a11*^{Mid}, *Pglyrp1*$^{+}$). One of the granulocyte clusters is distinct from the majority of other granulocytes and was classified as basophils (*Ccl3*$^{+}$, *Fcer1a*$^{+}$, *Mcpt8*$^{+}$) based on our previous results[134].

It is important to note for both unfractionated and fractionated analysis, that the HPC cluster likely contains very few (1-5) true, long-term reconstituting HSCs (LT-HSCs), although they are not immediately distinguishable[136–139]. This is likely because (1) cell-surface protein levels and mRNA expression do not perfectly correlate with the surface markers that are classically used to identify phenotypic HSCs (e.g. *Cd150*$^{+}$, *Cd48*$^{-}$, *Flt3*$^{-}$, *Cd34*$^{-}$) and therefore, these markers cannot be reliably used at the transcriptional level to distinguish LT-HSCs from downstream progenitors; (2) there is a chance of drop out of lowly expressed HSC-specific reporter genes such as *Hoxb5*, *Fgf5*, and *Ctnnal1*[136–138]. Additionally, in these datasets, the level of resolution needed to distinguish LT-HSCs from progenitors may be obscured by the index switching[140] of key negative marker genes across the sequenced pool. It also seems likely that many of the early B, T, NK, granulocyte, and monocyte progenitors that were captured during the four different sorting schemes (B, T, Granulocytes, and Monocytes; see above) also clustered with the primary HPC cluster as this resolution and quality of data was not sufficient to distinguish them for reasons mentioned earlier. This would explain the presence of MEPs and CLPs in this cluster.

In summary, we were able to generate single cell transcriptomic data for all major cell populations of bone marrow using two different sequencing approaches. Given the depth of sequencing and the quality of our data, we were able to identify many immature cell-types that are in the process of transitioning to a more mature state. We expect these rich datasets will be valuable in investigating cell-fate decisions during both early myelopoiesis and early lymphopoiesis.

Thymus

Figures are located in the Tabula Muris Tissue Supplement under section(s):
Thymus Droplet
Thymus FACS

The thymus is a bi-lobed organ of epithelial origin that contains developing T cells (thymocytes) and is surrounded by a mesenchymal capsule[141]. The capsule is composed of loose connective tissue that extends into the gland and forms septae. The septae, which contain blood vessels, nerves, and efferent lymphatics, further divide the thymus into lobules ranging in size from 0.5 to 2.0 μm. These lobules are permeated by two networks: 1) a supportive framework of reticular fibers, and 2) a mesh-like network of thymic epithelial cells (TECs). The thymocytes migrate through the interstices in this 3-dimensional mesh, where they interact with TECs expressing trophic and inhibitory cytokines and growth factors, major histocompatibility complex (MHC) proteins, and self-antigens. In addition to the resident TECs, the thymic microenvironment is also comprised of hematopoietic macrophages and dendritic cells derived from circulating progenitors in the blood, and mesenchymal cells which may be of neural crest origin.

Histologically, each lobe of the thymus is comprised of cortical and medullary regions, which play different roles in T cell development. Each contains distinct populations of stromal cells including TECs, mesenchymal (e.g., fibroblasts), endothelial, and dendritic cells[142], as well as thymocytes in different stages of development. The differentiation of thymocytes has been largely defined on the basis of surface antigen (protein) expression, and selected transcripts for gene expression analysis, but global transcriptomic analyses have been lacking. Hematopoietic progenitors enter the thymus at the cortical-medullary junction and migrate to the cortex. Cortical thymocytes are a tightly packed group of immature cells expressing the surface protein CD2, and absent or low levels of the invariant CD3 subunits of the T-cell receptor (TCR). The most immature thymocytes, which do not express either of the co-receptors CD4 and CD8, are termed double negative (DN) cells. The DN thymocytes can be further divided into four successive developmental stages based on their expression of the proteins CD44 and Il2ra (CD25): 1) DN1, $CD44^{+}CD25^{-}$; 2) DN2, $CD44^{+}CD25^{+}$; 3) DN3, $CD44^{-}CD25^{+}$; and 4) DN4, $CD44^{-}CD25^{-}$. Formation of the TCR repertoire occurs in immature thymocytes by productive V(D)J recombination of the TCR and then TCR loci mediated by the *Rag1* and *Rag2* genes. Further diversity of the rearranged TCR loci can occur by trimming of nucleotides at V(D)J junctions and random addition of junctional nucleotides at 3' ends of V, D, or J segments by the DNTT DNA polymerase. Selection of cells that have generated a productive TCR rearrangement ("β-selection") is mediated by expression of the pre-TCR, which is composed of a non-rearranged (invariant) pre-TCRα chain and a rearranged TCR β-chain, is followed by V(D)J recombination at the TCR locus. The end result is a population of immature cells expressing a clonotypic TCR comprised of a TCR and TCR subunit, which with the CD3 subunits comprises a signaling receptor for MHC-peptide complexes. After the DN stages, immature thymocytes transiently express CD8 alone, and become immature single positive (ISP) cells. Replacement of the pre-TCR with a rearranged TCRα and TCRβ, in addition to expression of CD4, marks the differentiation into $CD8^{+}CD4^{+}$ double positive (DP) thymocytes. DPs in the cortex are selected for survival and proliferation based on the affinity of interactions between the clonotypic TCR with MHC and self-peptides expressed by cortical TECs. Thymocytes with either too much or too little MHC affinity are negatively selected and undergo apoptosis, while those with intermediate affinity are positively selected and proliferate. Positively selected DP thymocytes lose expression of either CD4 or CD8 and differentiate into single-positive CD4 (SP4) or CD8 (SP8) thymocytes and migrate into the medulla. Negative selection of SP thymocytes based on affinity for MHC and self-peptides expressed by medullary TECs then prunes self-reactive mature thymocytes. The surviving SP thymocytes undergo final maturation before export into the periphery as CD4 or CD8 T lymphocytes.

Tissue Processing

Thymi were crushed on a 70 μm strainer (Falcon 352350), and centrifuged (270 x g, 5°C, 5 minutes). The cell pellet was digested with 2.2 mg/ml Collagenase II (Sigma C6885) for 10 minutes at 37°C, and incubated at 37°C for 30 minutes with agitation. The digestion was quenched with FACS buffer (2% FBS, 1% Antibiotics (Gibco 15240-062), and 10% Pluronics (ThermoFisher 24040032) in PBS), cells were pelleted (270 x g, 5°C, 5 minutes), and stained using 1:50 dilutions of antibodies for TER119-PB (Biolegend 116232), CD3-APC (Biolegend 100236), CD2-APC (Biolegend 100112), and CD45-Pe-Cy7 (eBioscience 25-0451-82, clone 30-F11) on ice for 20 minutes. Cells were washed and resuspended in FACS buffer, and stained with 1:1000 Sytox Blue (ThermoFisher S34857) immediately prior to sorting. Cells were sorted into two bins: thymocytes ($TER119^{-}$, $CD45^{+}$, $CD2^{+}$ or $CD3^{+}$) and stromal cells ($TER119^{-}$, $CD45^{-}$).

Data Analysis

Here, thymi from 2 females and 3 males were dissociated and 1,349 and 1,421 cells were analyzed by FACS and microfluidic droplet, respectively. For the FACS analysis, 6 clusters emerged, 5 of which represent immature thymocytes of different differentiation stages. The most

immature cells, the DN1 thymocytes, lack expression of *Cd8*, *Cd4,* and *Il2ra (Cd25)*, but express *Cd44*. They also lack expression of genes involved in V(D)J recombination, including the recombination activation genes *Rag1* and *Rag2*, and *Dntt*, the enzyme that mediates non-template based nucleotide addition at TCR junctions. Another cluster representing more mature DN2, DN3, and DN4 thymocytes (*Cd44*$^{-}$, *Il2ra*$^{-}$) also has a subset of DP cells. Although some of these cells express *Dntt*, they have either not yet expressed or have down-regulated expression of the *Rag1* and *Rag2* genes.

The remaining clusters represent the intermediate stages between DN and DP, showing *Cd8* expression but variable expression of *Cd4*, *Rag1, Rag2,* and *Dntt*. Of these cells, one cluster represents proliferating thymocytes (*Top2a*$^{+}$*, Mki67*$^{+}$) with low *Rag1* and *Rag2*, yet expressing *Ptprca,* suggesting that these cells are undergoing TCRβ selection.

Markers for lymphocyte subset specific differentiation (e.g., THPOK (*Zbtb7b*) for CD4 differentiation[143] and *Runx3* for CD8 differentiation[144]) are present in all the clusters but mainly in the DN2-4 population. They appear to be mutually exclusive, consistent with their predefined roles (*Zbtb7b*$^{+}$ favoring CD8 differentiation, and *Runx3*$^{+}$ favoring CD3 differentiation).

The final, non-thymocyte cluster is likely composed of stromal cells (*Cd3*^{low}, *Cd4*$^{-}$, *Cd8*$^{-}$). These cells express MHC II genes (e.g., *H2-Aa* and *H2-Ab1*), suggesting they are involved in antigen presentation. Some of these cells express *Ptprc* and *Cd86*, indicating that they are of hematopoietic origin and may be dendritic cells, which have the ability to present both tissue specific and circulating antigens. These cells lack markers of mesenchymal cells (e.g., *Igfbp7*, *C3*, and *Pdgfrb*), and TECs, such as cytokeratins (e.g., *Krt5*, *Krt8*)[142].

In summary, the majority of cells were immature thymocytes in different stages of differentiation with a small percentage of cells representing stromal cells. Given that the majority of thymic cells are thymocytes, the low stromal yield is not unexpected. While other studies have reported single transcriptional analysis of isolated thymic epithelial cells[142,145–147], to our knowledge, an analysis of whole thymus tissue using single cell sequencing that includes thymocytes has not yet been performed.

Spleen

<u>Figures are located in the Tabula Muris Tissue Supplement under section(s):</u>
Spleen Droplet
Spleen FACS

The spleen is a major site of blood filtration, iron metabolism, and pathogen detection. Anatomically, it is divided into the vascular tree, the red and white pulp, and the perifollicular zone (PFZ), a specialized compartment of the red pulp. These regions serve different physiological functions and are known to consist of distinct cell populations.

Constituting 75% of splenic volume (ref), the preponderant function of red pulp is blood filtration. However, a fair amount of the red pulp does not include capillary endings and is instead surrounded primarily by sinuses. In these regions, aggregates of lymphocytes and mononuclear phagocytes are present. These areas are thus regarded as part of the splenic lymphoid compartment, similar to the white pulp. The perifollicular zone is a specialized compartment of the red pulp containing its own reticular stroma. It also contains a mixture of blood cells comparable to that of peripheral blood and has been suggested to be responsible for the passage of about 10% of the splenic blood.

The splenic white pulp consists of both B and T cell lymphoid compartments. The B cell compartment mainly contains splenic lymphoid follicles, similar to lymph nodes found throughout the body. However, the splenic lymphoid follicles have a unique structure called the splenic marginal zone, which contains a population of macrophages functionally distinct from those of the

red pulp. These macrophages are responsible for maintaining the anatomic structure of the marginal zone.

Tissue Processing

Spleens were minced with razors, ground with a syringe plunger, and further dissociated in RPMI with 2% FBS and 1X Antibiotic-Antimycotic (Gibco 15240-062). Following trituration with a 5 ml serological pipette, cells were ground through 100 μm, 70 μm, and 40 μm filters (Falcon 352340, 352350, 352360) with a syringe plunger. Cells were pelleted (541 x g, 10 minutes, 4°C), treated with ACK (A10492-01) for 5 minutes at 21°C, washed in 2% FBS and 1X Antibiotic-Antimycotic in RPMI, and debris was removed using a cell debris removal kit (Miltenyi 130-109-398). Cells were resuspended in 2% FBS in PBS and filtered into 35 μm FACS tubes (Falcon 352235) before staining with 1:500 PI (Life Tech P3655) immediately prior to sorting.

Data Analysis

Here, we profiled 9,552 spleen cells from 2 male mice with microfluidic droplets, and 1,697 cells from 4 males and 2 females following FACS for viability only. We captured all major cell types, including the most abundant B cells (*Cd79a*$^+$), consisting of a marginal zone B cell cluster (*Ccr2*$^+$), and a cluster of T1/T2/follicular B cells. We also observe populations of *Cd4*$^+$ and *Cd8*$^+$ T cells, monocytes, and natural killer (NK) cells.

The majority of cells were different subpopulations of B cells, which is not unexpected as the spleen is a major cite of B cell development. While other studies have reported single transcriptional analysis of the spleen[148], this is the largest data set to date and the large number of B cells should help elucidate the mechanisms in the development of this important cell type.

Nervous System

Brain

Figures are located in the Tabula Muris Tissue Supplement under section(s):
Brain Myeloid FACS
Brain non-Myeloid FACS

The brain is arguably the most complex organ of the body, consisting of neuronal and non-neuronal cell types. Neurons display an extraordinary level of diversity, as demonstrated by their functional properties, gene expression, and especially connectivity, which is the foundation of normal brain functionality. Synapsing with hundreds to thousands of other cells, neurons conduct the electrical signals which govern thought and behavior. Non-neuronal cell types including microglia, astrocytes, oligodendrocyte lineage cells, brain endothelial cells (BECs), and pericytes are also increasingly recognized as heterogeneous[149,150], and play critical functions in the brain.

For example, oligodendrocytes facilitate neuronal signal transmission by forming an insulating myelin sheath around axons[151]. Interestingly, oligodendrocyte progenitor cells (OPCs) remain present in the adult brain, constantly surveilling the environment and contributing to remyelination during injury[150]. The brain also has specialized innate immune cells residing in the parenchyma, macrophage-like microglia. Microglia play important developmental functions, maintain tissue homeostasis, and aid in tissue repair[152]. Also responsible for a variety of functions are astrocytes. Though classically defined as only supportive cells, astrocytes conduct critical functions such as regulating neuronal synapses, recycling neurotransmitters, and maintaining the blood-brain barrier (BBB) of the brain's vasculature. Formed by BECs linked with tight junctions, the BBB closely regulates transport between the brain and the periphery (discussed in detail below). Finally, surrounding BECs are pericytes, contractile cells that also maintain the BBB. These

different cell types and subtypes interact extensively with one another, forming an intricate network that is essential for maintaining brain homeostasis and function.

Tissue Processing – Myeloid Cells

Cortex (CTX), cerebellum (CB), hippocampus (HIP), and striatum (STR) were dissected from one hemisphere into cold medium A (15 mM HEPES, 0.5% glucose in HBSS without phenol red), where white matter was manually removed. The remaining tissue was minced and dounced in medium A (2ml) with 1:25 12,500 unit/mL DNase (Worthington Bio LS002007) and 1:400 RNase inhibitor (Clontech 2313B), prior to filtering in a 70 µm strainer (Falcon 352350) and pelleting (400 x g, 5 minutes, Brake=5, 4˚C).

CTX and CB were resuspended in 1 ml MACS buffer (0.5% BSA and 2 mM EDTA in PBS) with 1:500 RNase inhibitor and 1:10 myelin removal beads (Miltenyi Biotec 130-096-433). HIP and STR were resuspended in 500 ul MACS buffer with 1:500 RNase inhibitor and 1:10 myelin removal beads. All samples were incubated on ice for 10 minutes. The volume of CTX was adjusted to 2 ml with MACS buffer and the volumes of all other samples were adjusted to 1 ml, before loading CTX onto an LD column (Miltenyi Biotec 130-042-901), and CB, HIP, and STR onto LS columns (Miltenyi Biotec 130-042-401). Cells were then filtered into 35 µm FACS tubes (Falcon 352235) and pelleted (400 x g, 5 minutes, Brake=5, 4˚C) before being resuspended in FACS buffer (1% FCS, 2 mM EDTA, and 25mM HEPES (pH 7.2-7.5) in PBS). Cells were then incubated on ice for 5 minutes in 1:60 Fc block (BD Pharmingen 553142), followed by 1:400 rabbit anti-mouse Tmem119 (Abcam ab210405) agitating for 10 minutes at 21˚C. After pelleting cells (400 x g, 5 minutes, Brake=5, 4˚C) and resuspending in FACS buffer, cells were stained with 1:300 CD45-PE-Cy7 (eBioscience 25-0451-82), 1:300 CD11b-BV421 (BioLegend 101236), and 1:300 Alexa 488 goat anti-rabbit (Invitrogen A11034) at 21˚C for 10 minutes. The cells were then pelleted (400 x g, 5 minutes, Brake=5, 4˚C) and resuspended in FACS buffer with 1:500 RNase inhibitor and 1:1000 propidium iodide (Invitrogen P3566) immediately before sorting.

Tissue Processing – Non-Myeloid Cells (Neurons, Glia, and Endothelial Cells)

Cortex (CTX), cerebellum (CB), hippocampus (HIP), and striatum (STR) were dissected from one hemisphere and dissociated with the neural tissue dissociation kit (P) using manual dissociation (Miltenyi Biotec 130-092-628). Briefly, tissue was minced in buffer X prior to transferring to preheated 37ºC buffer X containing enzyme P. The mixture was triturated and incubated at 37ºC for 10 minutes, and triturated again before adding enzyme 2 mix. Following a 10 minute incubation at 37ºC with agitation, samples were triturated and 10 ml modified DPBS buffer (Gibco 14287080) was added. The tissue settled for 2 minutes, the supernatant was filtered through a 70 µm strainer (Falcon 352350), and cells pelleted (300 x g, 10 minutes, 21ºC).

HIP was resuspended in FACS buffer (modified DPBS + 0.5% BSA) and passed through a FACS tube with 35 µm cap (Falcon 352235) prior to centrifugation (200 x g, 5 minutes, 21ºC).

For CTX/CB/STR, myelin was removed by resuspending cell pellets in 0.9 M sucrose in pure DPBS with calcium and magnesium (Gibco 14040) and centrifuging (850 x g, 15 minutes, 21ºC). This was repeated on the pellets from CTX/CB. CTX/CB/STR were then resuspended in FACS buffer and passed through 35 µm FACS tubes before centrifugation (200 x g, 5 minutes, 21ºC).

CTX/CB/STR/HIP were then blocked with 1:65 Fc block (BD Pharmingen 553142) for 5 minutes shaking before being stained with 1:10 CD90.2/Thy-1.2-APC-Cy7 (Biolegend 105328), 1:10 CD171/L1CAM-PE Vio 770 (Miltenyi 130-102-135, clone 555), 1:100 CD31-BV421 (BD 562939), 1:10 ACSA-2-PE (Miltenyi 130102365, clone IH3-18A3), and 1:10 O4-PE (Miltenyi 130-095-887, clone O4) for 10 minutes shaking.

Cells were then washed with DPBS, pelleted (200 x g, 5 minutes, 21ºC), and resuspended in FACS buffer with 1:10,000 SytoxBlue (ThermoFisher S34857), and 1:500 RNase inhibitor (Clontech 2313B), immediately prior to sorting.

To maintain microglial quiescence, each right hemisphere from 3 males and 2 females was designated for microglial isolation, which required a specialized, low temperature protocol. Microglia ($CD11b^+$, $CD45^{low}$) and macrophages ($CD11b^+$, $CD45^{hi}$) were collected. The remaining cell types were separated by FACS from the left hemispheres of 4 males and 3 females. Astrocytes ($ACSA\text{-}2^+$), neurons ($L1CAM^+$ or $Thy1.2^+$), and oligodendrocyte lineage cells ($O4^+$) were positively selected as previously reported[153–157]. For enrichment of brain endothelial cells, pericytes, and other cell types, cells were collected in the negative bin ($ACSA\text{-}2^-$, $L1CAM^-$, $Thy1.2^-$, $O4^-$), as well as by unbiased sorting of all viable $CD11b^+$ cells.

Data Analysis

Overall, 4,455 myeloid cells formed 6 distinct clusters, and 3,401 non-myeloid cells formed 10 distinct clusters. All major central nervous system (CNS) cell types are represented, but clusters did not segregate by brain region. Microglia and non-parenchymal macrophages constitute the biggest cluster (*Cx3cr1*$^+$, *P2ry12*$^+$, *Tmem119*$^+$), and subclustering separated non-parenchymal macrophages from microglia by their expression of MHC-II class genes and *Cd163*. We also detected oligodendrocyte lineage cells ranging from OPCs (*Pdgfra*$^+$, *Susd5*$^+$, *Cspg4*$^+$), to mature myelinating oligodendrocytes (*Mog*$^+$, *Mag*$^+$, *Gjc2*$^+$). BECs were identified with pan-endothelial markers (*Pecam1*$^+$, *Cldn5*$^+$) and BBB-specific markers (*Slco1c1*$^+$, *Ocln*$^+$)[158–160]. A cluster expressing astrocyte-specific markers[158] is also evident (*Aldh1l1*$^+$, *Slc1a3*$^+$, *Aqp4*$^+$). Interestingly, subclustering revealed 5 astrocyte populations, one uniquely derived from the cerebellum. Indeed, this cluster expresses genes specific for Bergman glia (*Gdf10*$^+$, *Vim*$^+$, *Nbl1*$^+$, *A2m*$^+$), a specialized cerebellar astroglial cell[160]. Characterization of the neuronal cluster allowed us to identify excitatory (*Slc17a7*$^+$, *Neurod6*$^+$, *Mab21l1*$^+$) and inhibitory (*Gad1*$^+$, *Reln*$^+$, *Calb1*$^+$) neurons, as well as neuroprogenitor cells (*Dcx*$^+$, *Dlx2*$^+$, *Ascl1*$^+$, *Hes5*$^+$). Lastly, the smallest cluster consists of pericytes (*Des*$^+$, *Mcam*$^+$, *Pdgfrb*$^+$).

Of the 715 BECs analyzed, unsupervised clustering revealed a number transcriptionally distinct subpopulations. As observed previously[161,162], several clusters reflect the presence of vessel segmental heterogeneity, as seen by cluster enrichment of distinct markers for arterioles (*Bmx*$^+$, *Jag1*$^+$, *Efnb2*$^+$), post-capillary venules (*Nr2f2*$^+$, *Flt4*$^+$, *Vwf*$^+$) and capillaries (*Tfrc*$^+$, *Car4*$^+$, *Slc16a1*$^+$)[163–165]. One particularly distinct population is composed of post-capillary venules and capillaries enriched with inflammatory genes including *Vcam1, Icam1, Lcn2, Hif1a, Vwf*, and *Csf1*. A cluster of pro-angiogenic arterioles is also present, with distinctly high expression of *Jag1, Notch1, Hey1, Vegfc*, *Edn1,* and *Tmem100*, key players in the Notch-signaling pathway.

The brain is the most commonly studied organ using single cell technologies both in humans[166,167] and mice[168,169], and many studies have sampled more cells than we did. However, our dataset adds to existing datasets by i) including cell types that are underrepresented in most other studies (as most focus on neurons) and ii) sampling cells across four distinct brain regions of multiple male and female mice.

References

1. Li, G. *et al.* Transcriptomic Profiling Maps Anatomically Patterned Subpopulations among Single Embryonic Cardiac Cells. *Dev. Cell* **39,** 491–507 (2016).
2. Doppler, S. A. *et al.* Cardiac fibroblasts: more than mechanical support. *J. Thorac. Dis.* **9,** S36–S51 (2017).
3. Zhang, H. *et al.* Endocardium Minimally Contributes to Coronary Endothelium in the Embryonic Ventricular Free Walls. *Circ. Res.* **118,** 1880–1893 (2016).
4. Frank, E. D. & Warren, L. Aortic smooth muscle cells contain vimentin instead of desmin. *Proc. Natl. Acad. Sci. U. S. A.* **78,** 3020–4 (1981).
5. Sauteur, L. *et al.* Cdh5/VE-cadherin promotes endothelial cell interface elongation via cortical actin polymerization during angiogenic sprouting. *Cell Rep.* **9,** 504–513 (2014).

6. Gomez, D., Swiatlowska, P. & Owens, G. K. Epigenetic Control of Smooth Muscle Cell Identity and Lineage Memory. *Arterioscler. Thromb. Vasc. Biol.* **35,** 2508–2516 (2015).
7. Davis, S. *et al.* Isolation of angiopoietin-1, a ligand for the TIE2 receptor, by secretion-trap expression cloning. *Cell* **87,** 1161–1169 (1996).
8. Brooks, A. R., Lelkes, P. I. & Rubanyi, G. M. Gene expression profiling of human aortic endothelial cells exposed to disturbed flow and steady laminar flow. *Physiol. Genomics* **9,** 27–41 (2002).
9. Liu, D., Jia, H., Holmes, D. I. R., Stannard, A. & Zachary, I. Vascular Endothelial Growth Factor-Regulated Gene Expression in Endothelial Cells: KDR-Mediated Induction of Egr3 and the Related Nuclear Receptors Nur77, Nurr1, and Nor1. *Arterioscler. Thromb. Vasc. Biol.* **23,** 2002–2007 (2003).
10. Desai, T. J., Brownfield, D. G. & Krasnow, M. A. Alveolar progenitor and stem cells in lung development, renewal and cancer. *Nature* **507,** 190–194 (2014).
11. Treutlein, B. *et al.* Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* **509,** 371–375 (2014).
12. Herold, S., Mayer, K. & Lohmeyer, J. Acute lung injury: How macrophages orchestrate resolution of inflammation and tissue repair. *Frontiers in Immunology* **2,** 65 (2011).
13. Stämpfli, M. R. & Anderson, G. P. How cigarette smoke skews immune responses to promote infection, lung disease and cancer. *Nature Reviews Immunology* **9,** 377–384 (2009).
14. Vlahos, R. & Bozinovski, S. Role of alveolar macrophages in chronic obstructive pulmonary disease. *Frontiers in Immunology* **5,** 435 (2014).
15. Holt, P. G. Antigen presentation in the lung. in *American Journal of Respiratory and Critical Care Medicine* **162,** S151–S156 (2000).
16. Tan, S. Y. S. & Krasnow, M. A. Developmental origin of lung macrophage diversity. *Development* **143,** 1318–1327 (2016).
17. Mahnke, K., Schmitt, E., Bonifaz, L., Enk, A. H. & Jonuleit, H. Immature, but not inactive: The tolerogenic function of immature dendritic cells. *Immunology and Cell Biology* **80,** 477–483 (2002).
18. Rock, J. R., Futtner, C. R. & Harfe, B. D. The transmembrane protein TMEM16A is required for normal development of the murine trachea. *Dev. Biol.* **321,** 141–149 (2008).
19. Rock, J. R., Randell, S. H. & Hogan, B. L. M. Airway basal stem cells: a perspective on their roles in epithelial homeostasis and remodeling. *Dis. Model. Mech.* **3,** 545–556 (2010).
20. Tadokoro, T. *et al.* IL-6/STAT3 promotes regeneration of airway ciliated cells from basal stem cells. *Proc. Natl. Acad. Sci.* **111,** E3641–E3649 (2014).
21. Tata, P. R. *et al.* Dedifferentiation of committed epithelial cells into stem cells in vivo. *Nature* **503,** 218–223 (2013).
22. Branchfield, K. *et al.* Pulmonary neuroendocrine cells function as airway sensors to control lung immune response. *Science (80-.).* **351,** 707–710 (2016).
23. Kotton, D. N. & Morrisey, E. E. Lung regeneration: Mechanisms, applications and emerging stem cell populations. *Nature Medicine* **20,** 822–832 (2014).
24. Okubo, T., Clark, C. & Hogan, B. L. M. Cell Lineage Mapping of Taste Bud Cells and Keratinocytes in the Mouse Tongue and Soft Palate. *Stem Cells* **27,** 442–450 (2009).
25. Joost, S. *et al.* Single-Cell Transcriptomics Reveals that Differentiation and Spatial Signatures Shape Epidermal and Hair Follicle Heterogeneity. *Cell Syst.* **3,** 221–237.e9 (2016).
26. Yang, H., Adam, R. C., Ge, Y., Hua, Z. L. & Fuchs, E. Epithelial-Mesenchymal Micro-niches Govern Stem Cell Lineage Choices. *Cell* **169,** 483–496.e13 (2017).
27. Xu, M. *et al.* WNT10A mutation causes ectodermal dysplasia by impairing progenitor cell proliferation and KLF4-mediated differentiation. *Nat. Commun.* **8,** 15397 (2017).

28. Sukumaran, S. K. *et al.* Whole transcriptome profiling of taste bud cells. *Sci. Rep.* **7,** 7595 (2017).
29. Perea-Martinez, I., Nagai, T. & Chaudhari, N. Functional Cell Types in Taste Buds Have Distinct Longevities. *PLoS One* **8,** e53399 (2013).
30. Jungermann, K. & Keitzmann, T. Zonation of Parenchymal and Nonparenchymal Metabolism in Liver. *Annu. Rev. Nutr.* **16,** 179–203 (1996).
31. Torre, C., Perret, C. & Colnot, S. Molecular determinants of liver zonation. *Prog. Mol. Biol. Transl. Sci.* **97,** 127–150 (2010).
32. Wang, B., Zhao, L., Fish, M., Logan, C. Y. & Nusse, R. Self-renewing diploid Axin2+ cells fuel homeostatic renewal of the liver. *Nature* **524,** 180–185 (2015).
33. Halpern, K. B. *et al.* Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* **542,** 1–5 (2017).
34. Poisson, J. *et al.* Liver sinusoidal endothelial cells: Physiology and role in liver diseases. *J. Hepatol.* **66,** 212–227 (2017).
35. Tsuchida, T. & Friedman, S. L. Mechanisms of hepatic stellate cell activation. *Nat. Rev. Gastroenterol. Hepatol.* **14,** 397–411 (2017).
36. Jenne, C. N. & Kubes, P. Immune surveillance by the liver. *Nature Immunology* **14,** 996–1006 (2013).
37. Berry, M. N. & Friend, D. S. High-yield preparation of isolated rat liver parenchymal cells: a biochemical and fine structural study. *J. Cell Biol.* **43,** 506–520 (1969).
38. Wang, B., Zhao, L., Fish, M., Logan, C. Y. & Nusse, R. Self-renewing diploid Axin2 + cells fuel homeostatic renewal of the liver. *Nature* **524,** 180–185 (2015).
39. Braeuning, A. *et al.* Differential gene expression in periportal and perivenous mouse hepatocytes. *FEBS J.* **273,** 5051–5061 (2006).
40. Nolan, D. J. *et al.* Molecular Signatures of Tissue-Specific Microvascular Endothelial Cell Heterogeneity in Organ Maintenance and Regeneration. *Dev. Cell* **26,** 204–219 (2013).
41. Yang, H. *et al.* Intravenous Administration of Human Umbilical Cord Mesenchymal Stem Cells Improves Cognitive Impairments and Reduces Amyloid-Beta Deposition in an AβPP/PS1 Transgenic Mouse Model. *Neurochem. Res.* **38,** 2474–2482 (2013).
42. Beattie, L. *et al.* Interferon regulatory factor 7 contributes to the control of Leishmania donovani in the mouse liver. *Infect. Immun.* **79,** 1057–1066 (2011).
43. Cha, S.-J. *et al.* CD68 acts as a major gateway for malaria sporozoite liver infection. *J. Exp. Med.* **212,** 1391–1403 (2015).
44. Jenne, C. N. & Kubes, P. Immune surveillance by the liver. *Nat. Immunol.* **14,** 996–1006 (2013).
45. Baron, M. *et al.* A Single-Cell Transcriptomic Map of the Human and Mouse Pancreas Reveals Inter- and Intra-cell Population Structure. *Cell Syst.* **3,** 346–360 (2016).
46. Xin, Y. *et al.* Use of the Fluidigm C1 platform for RNA sequencing of single mouse pancreatic islet cells. *Proc. Natl. Acad. Sci.* **113,** 3293–3298 (2016).
47. Chen, H. *et al.* Polycomb protein Ezh2 regulates pancreatic β-cell Ink4a/Arf expression and regeneration in diabetes mellitus. *Genes Dev.* **23,** 975–985 (2009).
48. Arda, H. E. *et al.* Age-dependent pancreatic gene regulation reveals mechanisms governing human β cell function. *Cell Metab.* **23,** 909–920 (2016).
49. Lee, J. *et al.* Reconstituting development of pancreatic intraepithelial neoplasia from primary human pancreas duct cells. *Nat. Commun.* **8,** 14686 (2017).
50. Blodgett, D. M. *et al.* Novel observations from next-generation RNA sequencing of highly purified human adult and fetal islet cell subsets. *Diabetes* **64,** 3172–3181 (2015).
51. Enge, M. *et al.* Single-Cell Analysis of Human Pancreas Reveals Transcriptional Signatures of Aging and Somatic Mutation Patterns. *Cell* **171,** 321–330.e14 (2017).
52. Schwitzgebel, V. M. *et al.* Expression of neurogenin3 reveals an islet cell precursor population in the pancreas. *Development* **127,** 3533–3542 (2000).

53. Wang, S. *et al.* Sustained Neurog3 expression in hormone-expressing islet cells is required for endocrine maturation and function. *Proc. Natl. Acad. Sci.* **106,** 9715–9720 (2009).
54. Mitchell, R. K. *et al.* The transcription factor Pax6 is required for pancreatic β cell identity, glucose-regulated ATP synthesis, and Ca2+ dynamics in adult mice. *J. Biol. Chem.* **292,** 8892–8906 (2017).
55. Scoville, D. W. *et al.* MLL3 and MLL4 methyltransferases bind to the MafA and MAFB transcription factors to regulate islet β-cell function. *Diabetes* **64,** 3772–3783 (2015).
56. Xin, Y. *et al.* RNA Sequencing of Single Human Islet Cells Reveals Type 2 Diabetes Genes. *Cell Metab.* **24,** 608–615 (2016).
57. Barker, N. Adult intestinal stem cells: Critical drivers of epithelial homeostasis and regeneration. *Nature Reviews Molecular Cell Biology* **15,** 19–33 (2014).
58. Rothenberg, M. E. *et al.* Identification of a cKit+ colonic crypt base secretory cell that supports Lgr5+ stem cells in mice. *Gastroenterology* **142,** 1195–1205.e6 (2012).
59. Zorn, A. M. & Wells, J. M. Vertebrate Endoderm Development and Organ Formation. *Annu. Rev. Cell Dev. Biol.* **25,** 221–251 (2009).
60. Byeon, M. K. *et al.* The down-regulated in adenoma (DRA) gene encodes an intestine-specific membrane glycoprotein. *Oncogene* **12,** 387–396 (1996).
61. Dalerba, P. *et al.* Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat. Biotechnol.* **29,** 1120–1127 (2011).
62. Yang, Q., Bermingham, N. A., Finegold, M. J. & Zoghbi, H. Y. Requirement of Math1 for Secretory Cell Lineage Commitment in the Mouse Intestine. *Science (80-.).* **294,** 2155–2158 (2001).
63. Gregorieff, A. *et al.* The Ets-Domain Transcription Factor Spdef Promotes Maturation of Goblet and Paneth Cells in the Intestinal Epithelium. *Gastroenterology* **137,** 1333–1345.e3 (2009).
64. Gerbe, F., Brulin, B., Makrini, L., Legraverend, C. & Jay, P. DCAMKL-1 Expression Identifies Tuft Cells Rather Than Stem Cells in the Adult Mouse Intestinal Epithelium. *Gastroenterology* **137,** 2179–2180 (2009).
65. Facer, P. *et al.* Chromogranin: a newly recognized marker for endocrine cells of the human gastrointestinal tract. *Gastroenterology* **89,** 1366–73 (1985).
66. Barker, N. *et al.* Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* **449,** 1003–1007 (2007).
67. Herring, C. A. *et al.* Unsupervised Trajectory Analysis of Single-Cell RNA-Seq and Imaging Data Reveals Alternative Tuft Cell Origins in the Gut. *Cell Syst.* (2017). doi:10.1016/j.cels.2017.10.012
68. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. *Appl. Environ. Microbiol.* **79,** 5112–5120 (2013).
69. Caporaso, J. G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci.* **108,** 4516–4522 (2011).
70. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26,** 2460–2461 (2010).
71. DeSantis, T. Z. *et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72,** 5069–5072 (2006).
72. Lozupone, C. & Knight, R. UniFrac: A new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* **71,** 8228–8235 (2005).
73. Stecher, B. *et al.* Salmonella enterica serovar typhimurium exploits inflammation to compete with the intestinal microbiota. *PLoS Biol.* **5,** 2177–2189 (2007).
74. Thompson, J. A., Oliveira, R. A., Djukovic, A., Ubeda, C. & Xavier, K. B. Manipulation of the quorum sensing signal AI-2 affects the antibiotic-treated gut microbiota. *Cell Rep.*

10, 1861–1871 (2015).
75. Dressler, G. R. The Cellular Basis of Kidney Development. *Annu. Rev. Cell Dev. Biol.* **22,** 509–529 (2006).
76. Veiras, L. C. *et al.* Sexual Dimorphic Pattern of Renal Transporters and Electrolyte Homeostasis. *J. Am. Soc. Nephrol.* **28,** 1–14 (2017).
77. Baylis, C. Changes in renal hemodynamics and structure in the aging kidney; sexual dimorphism and the nitric oxide system. *Experimental Gerontology* **40,** 271–278 (2005).
78. Tanaka, R. *et al.* Sex differences in ischaemia/reperfusion-induced acute kidney injury depends on the degradation of noradrenaline by monoamine oxidase. *Clin. Exp. Pharmacol. Physiol.* **44,** 371–377 (2017).
79. Wells, J. M. & Melton, D. A. Vertebrate Endoderm Development. *Annu. Rev. Cell Dev. Biol.* **15,** 393–410 (1999).
80. Khandelwal, P., Abraham, S. N. & Apodaca, G. Cell biology and physiology of the uroepithelium. *AJP Ren. Physiol.* **297,** F1477–F1501 (2009).
81. Yu, Z., Mannik, J., Soto, A., Lin, K. K. & Andersen, B. The epidermal differentiation-associated Grainyhead gene Get1/Grhl3 also regulates urothelial differentiation. *EMBO J.* **28,** 1890–1903 (2009).
82. HICKS, R. M. THE MAMMALIAN URINARY BLADDERAN ACCOMMODATING ORGAN. *Biol. Rev.* **50,** 215–246 (1975).
83. Papafotiou, G. *et al.* KRT14 marks a subpopulation of bladder basal cells with pivotal role in regeneration and tumorigenesis. *Nat. Commun.* **7,** 11914 (2016).
84. Shin, K. *et al.* Hedgehog/Wnt feedback supports regenerative proliferation of epithelial stem cells in bladder. *Nature* **472,** 110–116 (2011).
85. Gandhi, D. *et al.* Retinoid signaling in progenitors controls specification and regeneration of the urothelium. *Dev. Cell* **26,** 469–482 (2013).
86. Lopez-Beltran, A. Bladder cancer: Clinical and pathological profile. in *Scandinavian Journal of Urology and Nephrology* **42,** 95–109 (2008).
87. Mysorekar, I. U., Isaacson-Schmid, M., Walker, J. N., Mills, J. C. & Hultgren, S. J. Bone Morphogenetic Protein 4 Signaling Regulates Epithelial Renewal in the Urinary Tract in Response to Uropathogenic Infection. *Cell Host Microbe* **5,** 463–475 (2009).
88. Shin, K. *et al.* Hedgehog signaling restrains bladder cancer progression by eliciting stromal production of urothelial differentiation factors. *Cancer Cell* **26,** 521–533 (2014).
89. Miyamoto, H. *et al.* Promotion of Bladder Cancer Development and Progression by Androgen Receptor Signals. *JNCI J. Natl. Cancer Inst.* **99,** 558–568 (2007).
90. Society, A. C. Cancer Facts & Figures 2017. **Atlanta,** American Cancer Society (2017).
91. Mauro, A. Satellite cell of skeletal muscle fiber. *J. Cell Biol.* **9,** 493–495 (1961).
92. Morgan, J. E. & Partridge, T. A. Muscle satellite cells. *Int. J. Biochem. Cell Biol.* **35,** 1151–1156 (2003).
93. Joe, A. W. B. *et al.* Muscle injury activates resident fibro/adipogenic progenitors that facilitate myogenesis. *Nat. Cell Biol.* **12,** 153–163 (2010).
94. Mounier, R., Chrétien, F. & Chazaud, B. *Blood vessels and the satellite cell niche*. *Current Topics in Developmental Biology* **96,** (2011).
95. Honda, H., Kimura, H. & Rostami, A. Demonstration and phenotypic characterization of resident macrophages in rat skeletal muscle. *Immunology* **70,** 272–7 (1990).
96. Liu, L., Cheung, T. H., Charville, G. W. & Rando, T. A. Isolation of skeletal muscle stem cells by fluorescence-activated cell sorting. *Nat. Protoc.* **10,** 1612–1624 (2015).
97. Brigitte, M. *et al.* Muscle resident macrophages control the immune cell reaction in a mouse model of notexin-induced myoinjury. *Arthritis Rheum.* **62,** 268–279 (2010).
98. Van Den Brink, S. C. *et al.* Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nature Methods* **14,** 935–936 (2017).
99. Cheung, T. H. *et al.* Maintenance of muscle stem-cell quiescence by microRNA-489.

Nature **482,** 524–528 (2012).
100. Randolph, M. E. & Pavlath, G. K. A muscle stem cell for every muscle: Variability of satellite cell biology among different muscle groups. *Frontiers in Aging Neuroscience* **7,** 190 (2015).
101. Cheung, T. H. & Rando, T. A. Molecular regulation of stem cell quiescence. *Nature Reviews Molecular Cell Biology* **14,** 329–340 (2013).
102. Cho, D. S. & Doles, J. D. Single cell transcriptome analysis of muscle satellite cells reveals widespread transcriptional heterogeneity. *Gene* **636,** 54–63 (2017).
103. de Morrée, A. *et al.* Staufen1 inhibits MyoD translation to actively maintain muscle stem cell quiescence. *Proc. Natl. Acad. Sci.* **114,** 201708725 (2017).
104. van Velthoven, C. T. J., de Morree, A., Egner, I. M., Brett, J. O. & Rando, T. A. Transcriptional Profiling of Quiescent Muscle Stem Cells In Vivo. *Cell Rep.* **21,** 1994–2004 (2017).
105. Visvader, J. E. Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis. *Genes and Development* **23,** 2563–2577 (2009).
106. Shackleton, M. *et al.* Generation of a functional mammary gland from a single stem cell. *Nature* **439,** 84–88 (2006).
107. Stingl, J. *et al.* Purification and unique properties of mammary epithelial stem cells. *Nature* **439,** 993–7 (2006).
108. Joshi, P. A. *et al.* Progesterone induces adult mammary stem cell expansion. *Nature* **465,** 803–807 (2010).
109. Asselin-Labat, M. L. *et al.* Control of mammary stem cell function by steroid hormone signalling. *Nature* **465,** 798–802 (2010).
110. Asselin-Labat, M. L. *et al.* Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. *Nat. Cell Biol.* **9,** 201–209 (2007).
111. Asselin-Labat, M.-L. *et al.* Gata-3 Negatively Regulates the Tumor-Initiating Capacity of Mammary Luminal Progenitor Cells and Targets the Putative Tumor Suppressor Caspase-14. *Mol. Cell. Biol.* **31,** 4609–4622 (2011).
112. Visvader, J. E. & Stingl, J. Mammary stem cells and the differentiation hierarchy: Current status and perspectives. *Genes and Development* **28,** 1143–1158 (2014).
113. Bach, K. *et al.* Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. *Nat. Commun.* **8,** 2128 (2017).
114. Pal, B. *et al.* Construction of developmental lineage relationships in the mouse mammary gland by single-cell RNA profiling. *Nat. Commun.* **8,** 1627 (2017).
115. Wang, D. *et al.* Identification of multipotent mammary stem cells by protein C receptor expression. *Nature* **517,** 81–84 (2014).
116. Majka, S. M., Barak, Y. & Klemm, D. J. Concise review: Adipocyte origins: Weighing the possibilities. *Stem Cells* **29,** 1034–1040 (2011).
117. Betz, M. J. & Enerbäck, S. Targeting thermogenesis in brown fat and muscle to treat obesity and metabolic disease. *Nat. Rev. Endocrinol.* (2017). doi:10.1038/nrendo.2017.132
118. Berry, R. & Rodeheffer, M. S. Characterization of the adipocyte cellular lineage in vivo. *Nat. Cell Biol.* **15,** 302–308 (2013).
119. Kumari, M., Heeren, J. & Scheja, L. Regulation of immunometabolism in adipose tissue. *Semin. Immunopathol.* (2017). doi:10.1007/s00281-017-0668-3
120. Cao, Y. Angiogenesis and vascular functions in modulation of obesity, adipose metabolism, and insulin sensitivity. *Cell Metabolism* **18,** 478–489 (2013).
121. Traktuev, D. O. *et al.* A population of multipotent CD34-positive adipose stromal cells share pericyte and mesenchymal surface markers, reside in a periendothelial location, and stabilize endothelial networks. *Circ. Res.* **102,** 77–85 (2008).
122. Lu, W. & Li, X. Vascular stem/progenitor cells: functions and signaling pathways.

Cellular and Molecular Life Sciences 1–11 (2017). doi:10.1007/s00018-017-2662-2
123. Wosczyna, M. N., Biswas, A. A., Cogswell, C. A. & Goldhamer, D. J. Multipotent progenitors resident in the skeletal muscle interstitium exhibit robust BMP-dependent osteogenic activity and mediate heterotopic ossification. *J. Bone Miner. Res.* **27,** 1004–1017 (2012).
124. Shan, T., Liu, W. & Kuang, S. Fatty acid binding protein 4 expression marks a population of adipocyte progenitors in white and brown adipose tissues. *FASEB J.* **27,** 277–287 (2013).
125. Baumann, C. I. *et al.* PECAM-1 is expressed on hematopoietic stem cells throughout ontogeny and identifies a population of erythroid progenitors. *Blood* **104,** 1010–1016 (2004).
126. Bryder, D., Rossi, D. J. & Weissman, I. L. Hematopoietic Stem Cells. *Am. J. Pathol.* **169,** 338–346 (2006).
127. Baum, C. M., Weissman, I. L., Tsukamoto, A. S., Buckle, A. M. & Peault, B. Isolation of a candidate human hematopoietic stem-cell population. *Proc. Natl. Acad. Sci.* **89,** 2804–2808 (1992).
128. Spangrude, G. J., Heimfeld, S. & Weissman, I. L. Purification and characterization of mouse hematopoietic stem cells. *Science* **241,** 58–62 (1988).
129. Christensen, J. L. & Weissman, I. L. Flk-2 is a marker in hematopoietic stem cell differentiation: a simple method to isolate long-term stem cells. *Proc Natl Acad Sci U S A* **98,** 14541–14546 (2001).
130. Akashi, K., Traver, D., Miyamoto, T. & Weissman, I. L. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* **404,** 193–197 (2000).
131. Kondo, M. *et al.* Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell* **91,** 661–72 (1997).
132. Beerman, I. *et al.* Functionally distinct hematopoietic stem cells modulate hematopoietic lineage potential during aging by a mechanism of clonal expansion. *Proc. Natl. Acad. Sci.* **107,** 5465–5470 (2010).
133. Yamamoto, R. *et al.* Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell* **154,** 1112–26 (2013).
134. Seita, J. *et al.* Gene expression commons: An open platform for absolute gene expression profiling. *PLoS One* **7,** e40321 (2012).
135. Gordon, S. R. *et al.* PD-1 expression by tumour-associated macrophages inhibits phagocytosis and tumour immunity. *Nature* **545,** 495–499 (2017).
136. Acar, M. *et al.* Deep imaging of bone marrow shows non-dividing stem cells are mainly perisinusoidal. *Nature* **526,** 126–130 (2015).
137. Chen, J. Y. *et al.* Hoxb5 marks long-term haematopoietic stem cells and reveals a homogenous perivascular niche. *Nature* **530,** 223–227 (2016).
138. Gazit, R. *et al.* Fgd5 identifies hematopoietic stem cells in the murine bone marrow. *J. Exp. Med.* **211,** 1315–1331 (2014).
139. Seita, J. & Weissman, I. L. Hematopoietic stem cell: Self-renewal versus differentiation. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* **2,** 640–653 (2010).
140. Sinha, R. *et al.* Index Switching Causes 'Spreading-Of-Signal' Among Multiplexed Samples In Illumina HiSeq 4000 DNA Sequencing. *bioRxiv* 125724 (2017). doi:10.1101/125724
141. Farley, A. M. *et al.* Dynamics of thymus organogenesis and colonization in early human development. *Development* **140,** 2015–2026 (2013).
142. Ki, S. *et al.* Global transcriptional profiling reveals distinct functions of thymic stromal subsets and age-related changes during thymic involution. *Cell Rep.* **9,** 402–415 (2014).
143. Carpenter, A. C. *et al.* Control of Regulatory T Cell Differentiation by the Transcription Factors Thpok and LRF. *J. Immunol.* **199,** ji1700181 (2017).

144. Park, J. H. *et al.* Signaling by intrathymic cytokines, not T cell antigen receptors, specifies CD8 lineage choice and promotes the differentiation of cytotoxic-lineage T cells. *Nat. Immunol.* **11,** 257–264 (2010).
145. St-Pierre, C. *et al.* Transcriptome sequencing of neonatal thymic epithelial cells. *Sci. Rep.* **3,** 1860 (2013).
146. Brennecke, P. *et al.* Single-cell transcriptome analysis reveals coordinated ectopic gene-expression patterns in medullary thymic epithelial cells. *Nat. Immunol.* **16,** 933–941 (2015).
147. Sansom, S. N. *et al.* Population and single-cell genomics reveal the Aire dependency, relief from Polycomb silencing, and distribution of self-antigen expression in thymic epithelia. *Genome Res.* **24,** 1918–1931 (2014).
148. Jaitin, D. A. *et al.* Massively Parallel Single-Cell RNA-Seq for Marker-Free Decomposition of Tissues into Cell Types. *Science (80-.).* **343,** 776–779 (2014).
149. Ben Haim, L. & Rowitch, D. H. Functional diversity of astrocytes in neural circuit regulation. *Nature Reviews Neuroscience* **18,** 31–41 (2016).
150. Marques, S. *et al.* Oligodendrocyte heterogeneity in the mouse juvenile and adult central nervous system. *Science (80-.).* **352,** 1326–1329 (2016).
151. Miron, V. E., Kuhlmann, T. & Antel Jack P., J. P. Cells of the oligodendroglial lineage, myelination, and remyelination. *Biochimica et Biophysica Acta - Molecular Basis of Disease* **1812,** 184–193 (2011).
152. Li, Q. & Barres, B. A. Microglia and macrophages in brain homeostasis and disease. *Nat. Rev. Immunol.* (2017). doi:10.1038/nri.2017.125
153. Batiuk, M. Y. *et al.* An immunoaffinity-based method for isolating ultrapure adult astrocytes based on ATP1B2 targeting by the ACSA-2 antibody. *J. Biol. Chem.* **292,** 8874–8891 (2017).
154. Horinouchi, K., Nakamura, Y., Yamanaka, H., Watabe, T. & Shiosaka, S. Distribution of L1cam mRNA in the adult mouse brain: In situ hybridization and northern blot analyses. *J. Comp. Neurol.* **482,** 386–404 (2005).
155. G. Kantzer, C. *et al.* Anti-ACSA-2 defines a novel monoclonal antibody for prospective isolation of living neonatal and adult astrocytes. *Glia* **65,** 990–1004 (2017).
156. Lin, G., Mela, A., Guilfoyle, E. M. & Goldman, J. E. Neonatal and adult O4+ oligodendrocyte lineage cells display different growth factor responses and different gene expression patterns. *J. Neurosci. Res.* **87,** 3390–3402 (2009).
157. Taylor-Clark, T. E. *et al.* Thy1.2 YFP-16 transgenic mouse labels a subset of large-diameter sensory neurons that lack TRPV1 expression. *PLoS One* **10,** e0119538 (2015).
158. Cahoy, J. D. *et al.* A Transcriptome Database for Astrocytes, Neurons, and Oligodendrocytes: A New Resource for Understanding Brain Development and Function. *J. Neurosci.* **28,** 264–278 (2008).
159. Obermeier, B., Daneman, R. & Ransohoff, R. M. Development, maintenance and disruption of the blood-brain barrier. *Nature Medicine* **19,** 1584–1596 (2013).
160. Koirala, S. & Corfas, G. Identification of novel glial genes by single-cell transcriptional profiling of Bergmann glial cells from mouse cerebellum. *PLoS One* **5,** e9198 (2010).
161. Yousef, H. *et al.* Aged blood inhibits hippocampal neurogenesis and activates microglia through VCAM1 at the blood-brain barrier. *bioRxiv* (2018). doi:10.1101/242198
162. Vanlandewijck, M. *et al.* A molecular atlas of cell types and zonation in the brain vasculature. *Nature* **554,** 475–480 (2018).
163. Murakami, M. Signaling required for blood vessel maintenance: Molecular basis and pathological manifestations. *International Journal of Vascular Medicine* **2012,** 293641 (2012).
164. Macdonald, J. A., Murugesan, N. & Pachter, J. S. Endothelial cell heterogeneity of blood-brain barrier gene expression along the cerebral microvasculature. *J. Neurosci. Res.* **88,**

1457–1474 (2010).
165. Daneman, R. *et al.* The mouse blood-brain barrier transcriptome: A new resource for understanding the development and function of brain endothelial cells. *PLoS One* **5,** e13741 (2010).
166. Lake, B. B. *et al.* Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science (80-.).* **352,** 1586–1590 (2016).
167. Darmanis, S. *et al.* A survey of human brain transcriptome diversity at the single cell level. *Proc. Natl. Acad. Sci.* **112,** 7285–7290 (2015).
168. Zeisel, A. *et al.* Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science (80-.).* **347,** 1138–1142 (2015).
169. Tasic, B. *et al.* Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* **19,** 335–346 (2016).