

## RESEARCH ARTICLE SUMMARY

## HUMAN CELL ATLAS

# Cross-tissue immune cell analysis reveals tissue-specific features in humans

C. Domínguez Conde<sup>†</sup>, C. Xu<sup>†</sup>, L. B. Jarvis<sup>†</sup>, D. B. Rainbow<sup>†</sup>, S. B. Wells<sup>†</sup>, T. Gomes, S. K. Howlett, O. Suchanek, K. Polanski, H. W. King, L. Mamanova, N. Huang, P. A. Szabo, L. Richardson, L. Bolt, E. S. Fasouli, K. T. Mahbubani, M. Prete, L. Tuck, N. Richoz, Z. K. Tuong, L. Campos, H. S. Mousa, E. J. Needham, S. Pritchard, T. Li, R. Elmentaita, J. Park, E. Rahmani, D. Chen, D. K. Menon, O. A. Bayraktar, L. K. James, K. B. Meyer, N. Yosef, M. R. Clatworthy, P. A. Sims, D. L. Farber, K. Saeb-Parsy<sup>\*‡</sup>, J. L. Jones<sup>\*‡</sup>, S. A. Teichmann<sup>\*‡</sup>

**INTRODUCTION:** Immune cells that seed peripheral tissues play a vital role in health and disease, yet most studies of human immunity focus on blood-derived cells. Immune cells adapt to local microenvironments, acquiring distinct features and functional specialization. Dissecting these molecular adaptations through the systematic assessment of cells across the human body promises to transform our understanding of the immune system at the organismal level.

**RATIONALE:** To comprehensively assess immune cell types, we collected donor-matched tissues from 12 deceased organ donors. We isolated immune cells and performed single-cell RNA sequencing and paired VDJ sequencing for T cell and B cell receptors, resulting in high-quality data for ~330,000 immune cells. To resolve cell identities, we developed CellTypist, a logistic regression-based framework using stochastic gradient descent learning. This cross-tissue annotation enabled interrogation of shared and

tissue-specific expression modules and cell states within myeloid and lymphoid cell lineages.

**RESULTS:** We developed CellTypist by curating and harmonizing public datasets to assemble a comprehensive immune cell type reference database (<https://www.celltypist.org>). CellTypist was then applied to our data, generated across multiple tissues and individuals (<https://www.tissueimmunecellatlas.org/>). Altogether, we detected 101 immune populations and performed extensive cross-tissue comparisons for each subset. Although macrophages displayed prominent tissue-restricted features, some convergent features were also detected. For example, macrophages expressing erytrophagocytosis-related genes were widely found in spleen, liver, bone marrow, and lymph nodes. Heterogeneity within defined subpopulations was also observed, such as migratory dendritic cell adaptations. Within adaptive immune lineages, we identified tissue-specific distributions of memory popula-

tions. Plasma cells showed a restricted tissue distribution, whereas memory B cells were more widely distributed. Similarly, tissue-resident memory T ( $T_{RM}$ ) cells were more restricted in distribution compared with central and effector memory T cells. Notably,  $T_{RM}$  cells harbored significant diversity, including  $\alpha\beta$  and  $\gamma\delta$  lineages, ascertained by VDJ sequencing. Assessment of clonal dynamics revealed the highest clonal expansions in  $T_{RM}$  cells and the most frequent clonal sharing between resident and effector memory populations.

**CONCLUSION:** Here we present an immune cell atlas of myeloid and lymphoid lineages across adult human tissues. We developed CellTypist for automated immune cell annotation and performed an in-depth dissection of cell populations, identifying 101 cell types or states from more than one million cells, including previously underappreciated cell states. We also uncovered convergent phenotypes across tissues within given lineages and described tissue adaptation signatures for a number of cell types, including macrophages and resident memory T cells. Together, we have extended our understanding of how human immunity functions as an integrated, cross-tissue network, and we provide the scientific community with several key new resources. ■

The list of author affiliations is available in the full article online.

\*Corresponding author. Email: st9@sanger.ac.uk (S.A.T.); jls53@medsch.cam.ac.uk (J.L.J.); ks10014@cam.ac.uk (K.S.-P.)

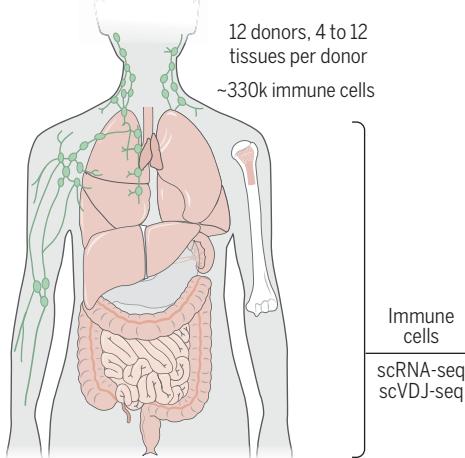
†These authors contributed equally to this work.

‡These authors contributed equally to this work.

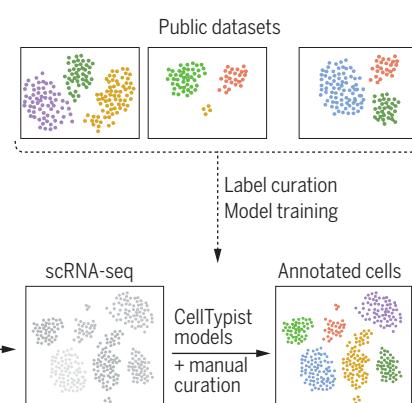
Cite this article as C. Domínguez Conde et al., *Science* **376**, eabl5197 (2022). DOI: 10.1126/science.abl5197

**S READ THE FULL ARTICLE AT**  
<https://doi.org/10.1126/science.abl5197>

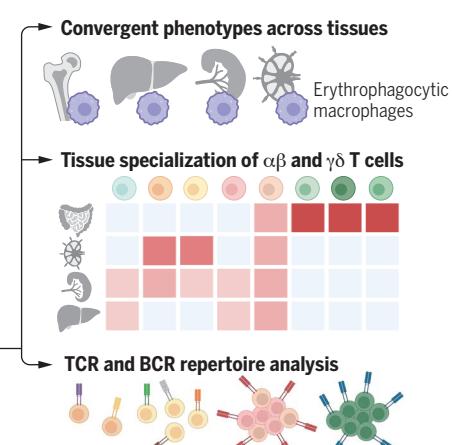
## Human immune cells across tissues



## CellTypist: automated cell type annotation



## Emerging cross-tissue features



**Cross-tissue atlas of human immune cells.** Donor-matched tissues from 12 deceased organ donors were profiled using single-cell RNA sequencing [including single-cell VDJ sequencing (scVDJ-seq)]. Cell type annotation was achieved with CellTypist, an automated cell type annotation pipeline, followed by manual curation.

In-depth analysis of transcriptional features revealed convergent and divergent gene expression programs for each cell lineage across lymphoid and nonlymphoid tissues. scVDJ-seq enabled repertoire analysis across cell subsets and tissues. scRNA-seq, single-cell RNA sequencing; TCR, T cell receptor; BCR, B cell receptor.

## RESEARCH ARTICLE

## HUMAN CELL ATLAS

# Cross-tissue immune cell analysis reveals tissue-specific features in humans

C. Domínguez Conde<sup>1†</sup>, C. Xu<sup>1†</sup>, L. B. Jarvis<sup>2†</sup>, D. B. Rainbow<sup>2†</sup>, S. B. Wells<sup>3†</sup>, T. Gomes<sup>1</sup>, S. K. Howlett<sup>2</sup>, O. Suchanek<sup>4</sup>, K. Polanski<sup>1</sup>, H. W. King<sup>5</sup>, L. Mamanova<sup>1</sup>, N. Huang<sup>1</sup>, P. A. Szabo<sup>6</sup>, L. Richardson<sup>1</sup>, L. Bolt<sup>1</sup>, E. S. Fasouli<sup>1</sup>, K. T. Mahbubani<sup>7</sup>, M. Prete<sup>1</sup>, L. Tuck<sup>1</sup>, N. Richoz<sup>4</sup>, Z. K. Tuong<sup>1,4</sup>, L. Campos<sup>1,8</sup>, H. S. Mousa<sup>2</sup>, E. J. Needham<sup>2</sup>, S. Pritchard<sup>1</sup>, T. Li<sup>1</sup>, R. Elmentaita<sup>1</sup>, J. Park<sup>1</sup>, E. Rahmani<sup>9,10</sup>, D. Chen<sup>3</sup>, D. K. Menon<sup>11</sup>, O. A. Bayraktar<sup>1</sup>, L. K. James<sup>5</sup>, K. B. Meyer<sup>1</sup>, N. Yossef<sup>9,10,12,13</sup>, M. R. Clatworthy<sup>1,4</sup>, P. A. Sims<sup>3</sup>, D. L. Farber<sup>6</sup>, K. Saeb-Parsy<sup>7\*†</sup>, J. L. Jones<sup>2\*†</sup>, S. A. Teichmann<sup>1,14\*†</sup>

Despite their crucial role in health and disease, our knowledge of immune cells within human tissues remains limited. We surveyed the immune compartment of 16 tissues from 12 adult donors by single-cell RNA sequencing and VDJ sequencing generating a dataset of ~360,000 cells. To systematically resolve immune cell heterogeneity across tissues, we developed CellTypist, a machine learning tool for rapid and precise cell type annotation. Using this approach, combined with detailed curation, we determined the tissue distribution of finely phenotyped immune cell types, revealing hitherto unappreciated tissue-specific features and clonal architecture of T and B cells. Our multitissue approach lays the foundation for identifying highly resolved immune cell types by leveraging a common reference dataset, tissue-integrated expression analysis, and antigen receptor sequencing.

The immune system is a dynamic and integrated network made up of many different cell types distributed across the body that act together to maintain tissue homeostasis and mediate protective immunity. In recent years, a growing appreciation of immune ontogeny and diversity across tissues has emerged. For example, we have gained insights into how macrophages derived in embryogenesis contribute to the distinctive adaptation of tissue-resident myeloid cells, such as Langerhans cells in the skin, microglia in the brain, and Kupffer cells in the liver (1–3). Other pop-

ulations, such as innate lymphoid cells (ILCs), including natural killer (NK) cells and non-conventional T cells [natural killer T (NKT) cells, mucosal-associated invariant T (MAIT) cells, and gamma-delta ( $\gamma\delta$ ) T cells], have circulating counterparts but are highly enriched at barrier sites where they mediate tissue defense and repair (4). In addition, following resolution of an immune response, antigen-specific, long-lived tissue-resident memory T ( $T_{RM}$ ) cells persist in diverse sites, where they provide optimal protection against secondary infections [reviewed in (5–7)].

Studies in mice have revealed the central role of tissue immune responses in protective immunity, antitumor immunity, and tissue repair; however, human studies have largely focused on peripheral blood as the most accessible site. Given that circulating immune cells make up only a subset of the entire immune cell landscape, understanding human immunity requires a comprehensive assessment of the properties and features of immune cells within and across tissues. Recent progress in the analysis of tissue immune cells has implemented organ-focused approaches (8–12), whereas use of tissues obtained from organ donors allows for analysis of immune cells across multiple sites across an individual (13–19). We previously reported single-cell RNA sequencing (scRNA-seq) analysis of T cells in three tissues from two donors (20), identifying tissue-specific signatures. However, despite the effort to assemble murine (21) and human (22, 23) multitissue atlases, large-scale cross-tissue scRNA-seq studies that investigate tissue-specific features of innate

and adaptive immune compartments have not been reported.

Furthermore, a fundamental challenge of increasingly large single-cell transcriptomics datasets is cell type annotation, including identifying rare cell subsets and distinguishing novel discoveries from previously described cell populations. Currently available automated annotation workflows leverage organ-focused studies and lack a comprehensive catalog of all cell types found across tissues. Therefore, a single unified approach is required to provide an in-depth dissection of immune cell type and immune state heterogeneity across tissues.

To address these needs, we comprehensively profiled immune cell populations isolated from a wide range of donor-matched tissues from 12 deceased individuals, generating nearly 360,000 single-cell transcriptomes. To systematically annotate multitissue immune cells, we developed CellTypist, a machine learning framework for cell type prediction initially trained on data from studies across 20 human tissues (see supplementary text in the supplementary materials) and then updated and extended by integrating immune cells from our dataset.

## Results

### *CellTypist: A pan-tissue immune database and a tool for automated cell type annotation*

To systematically assess immune cell type heterogeneity across human tissues, we performed scRNA-seq on 16 different tissues from 12 deceased organ donors (Fig. 1, A and B, and table S1). The tissues studied included primary (bone marrow) and secondary (spleen and lung-draining and mesenteric lymph nodes) lymphoid organs, mucosal tissues (gut and lung), as well as blood and liver. When available, we also included additional donor-matched samples from tissues such as thymus, skeletal muscle, and omentum. Immune cells were isolated from tissues as detailed in the methods. After stringent quality control, we obtained a total of 357,211 high-quality cells, of which 329,762 belonged to the immune compartment (fig. S1, A and B).

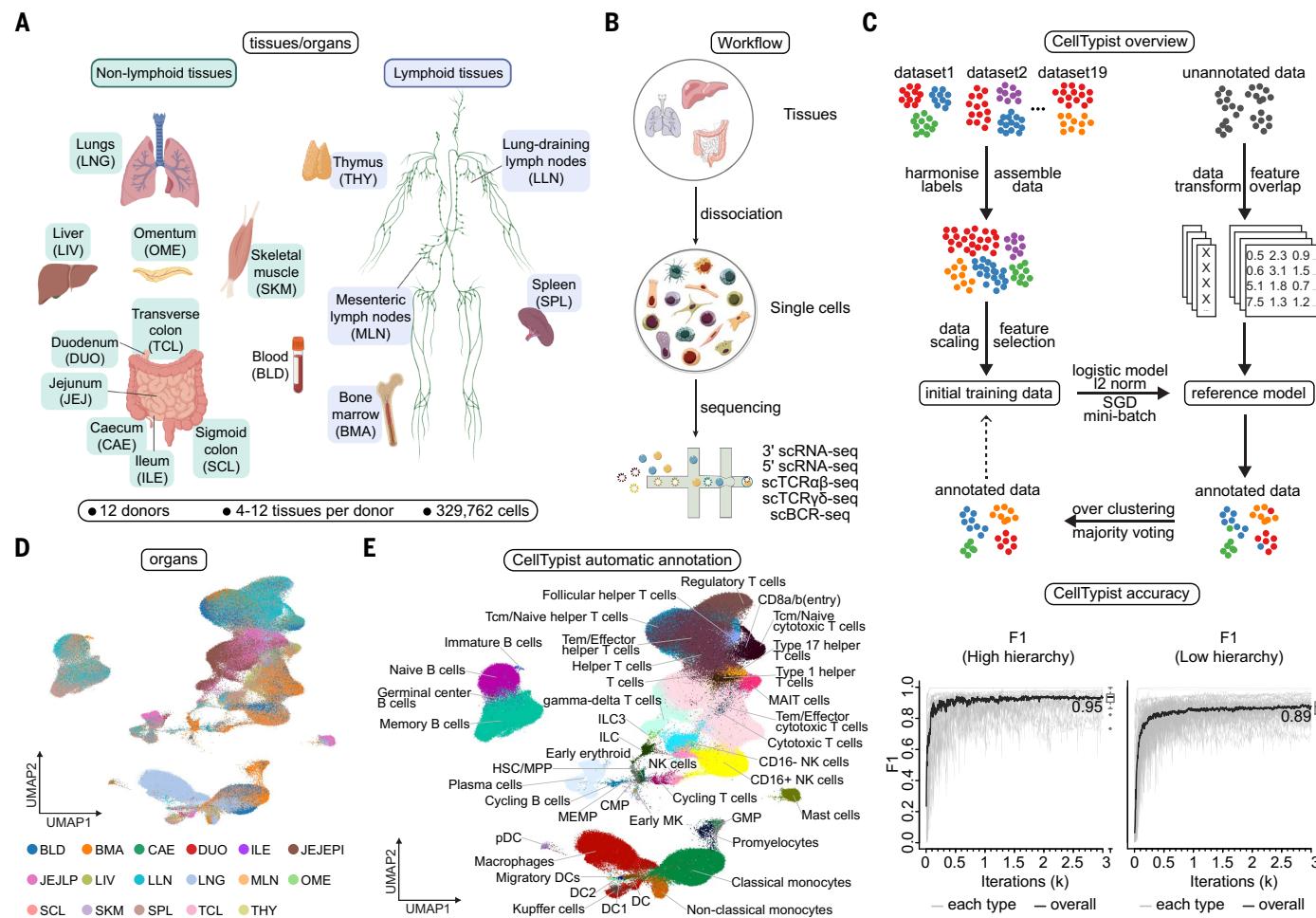
Robust cell type annotation remains a major challenge in single-cell transcriptomics. To computationally predict cellular heterogeneity in our dataset, we developed CellTypist (24), a cell type database, which in its current form is focused on immune cells, that provides a directly interpretable pipeline for the automatic annotation of scRNA-seq data (Fig. 1C). One of the distinctive and valuable aspects of CellTypist is that its training set encompasses a wide range of immune cell types across tissues. This breadth is of critical importance given that immune compartments are shared across tissues, warranting accurate and automated cell annotation in an organ-agnostic manner. In brief, to develop CellTypist,

<sup>1</sup>Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK. <sup>2</sup>Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK.

<sup>3</sup>Department of Systems Biology, Columbia University Irving Medical Center, New York, NY, USA. <sup>4</sup>Molecular Immunology Unit, Department of Medicine, University of Cambridge, Cambridge, UK. <sup>5</sup>Centre for Immunobiology, Blizard Institute, Queen Mary University of London, London, UK.

<sup>6</sup>Department of Microbiology and Immunology, Columbia University Irving Medical Center, New York, NY, USA. <sup>7</sup>Department of Surgery, University of Cambridge and NIHR Cambridge Biomedical Research Centre, Cambridge, UK. <sup>8</sup>West Suffolk Hospital NHS Trust, Bury Saint Edmunds, UK. <sup>9</sup>Center for Computational Biology, University of California, Berkeley, Berkeley, CA, USA. <sup>10</sup>Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA, USA. <sup>11</sup>Department of Anaesthesia, University of Cambridge, Cambridge, UK. <sup>12</sup>Chan Zuckerberg Biohub, San Francisco, CA, USA. <sup>13</sup>Ragon Institute of MGH, MIT and Harvard, Cambridge, MA, USA. <sup>14</sup>Theory of Condensed Matter, Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge, UK.

\*Corresponding author. Email: st9@sanger.ac.uk (S.A.T.); jls53@medsch.cam.ac.uk (J.L.J.); ks10014@cam.ac.uk (K.S.-P.)  
†These authors contributed equally to this work.  
‡These authors contributed equally to this work.



**Fig. 1. Automated annotation of immune cells across human tissues using CellTypist.** (A) Schematic showing sample collections of human lymphoid and nonlymphoid tissues and their assigned tissue name acronyms. (B) Schematic of single-cell transcriptome profiling and paired sequencing of  $\alpha\beta$  TCR,  $\gamma\delta$  TCR, and BCR variable regions. (C) Workflow of CellTypist including data collection, processing, model training, and cell type prediction (upper panel). Performance curves showing the F1 score at each iteration of training with mini-batch

stochastic gradient descent for high- and low-hierarchy CellTypist models, respectively (lower panel). The black curve represents the median F1 score averaged across the individual F1 scores of all predicted cell types. (D) Uniform manifold approximation and projection (UMAP) visualization of the immune cell compartment colored by tissues. Note that jejunum samples in (A) were further split into epithelial (JEJ/EP) and lamina propria (JEJLP) fractions. (E) As in (D), but colored by predicted cell types using CellTypist.

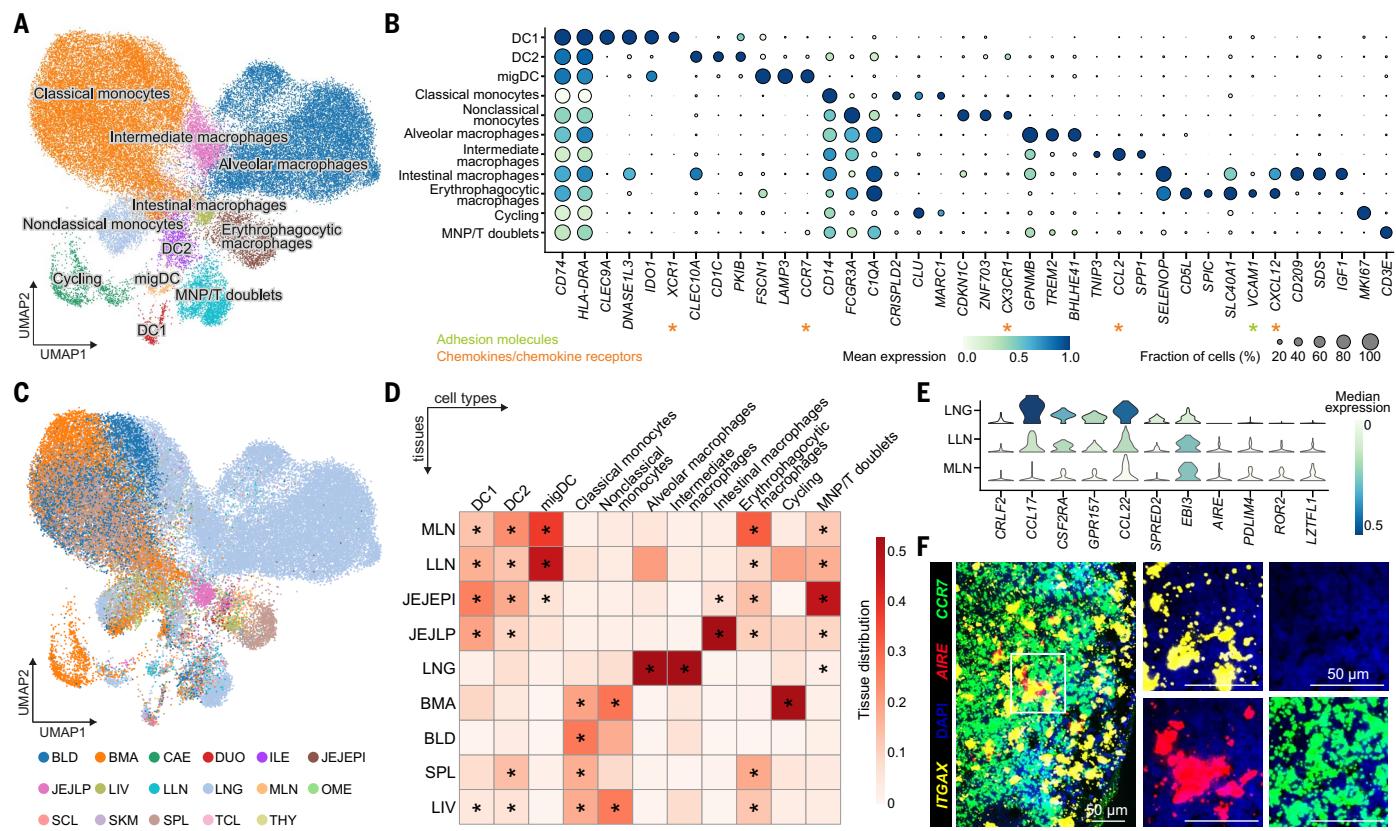
we integrated cells from 20 different tissues from 19 reference datasets (fig. S2) where we had deeply curated and harmonized cell types at two knowledge-driven hierarchical levels (figs. S3 to S8). This was followed by a machine learning framework to train a model using logistic regression with stochastic gradient descent learning (see methods). Performance of the derived models, as measured by the precision, recall, and global F1 score, reached ~0.9 for cell type classification at both the high- and low-hierarchy levels (Fig. 1C and fig. S9, A and B). Notably, representation of a given cell type in the training data was a major determinant of its prediction accuracy (fig. S9C), implying that higher model performance can be achieved by incorporating additional datasets. Moreover, CellTypist prediction was robust to differences between training and query datasets including gene

expression sparseness (fig. S10) and batch effects (fig. S11).

Next, we applied CellTypist's high-hierarchy (i.e., low-resolution, 32 cell types) classifier to our cross-tissue dataset (Fig. 1D) and found 15 major cell populations (fig. S1C). At this level of resolution, clear compositional patterns emerged in lymphoid versus nonlymphoid tissues, and within the lymphoid tissues between spleen and lymph nodes, and appeared not to be driven by differences in dissociation protocols (fig. S12). As the training datasets of CellTypist contained hematopoietic tissues with definitive annotations for progenitor populations, the classifier was able to resolve several progenitors including hematopoietic stem cells (HSCs), multipotent progenitors (MPPs), promyelocytes, early megakaryocytes, and pre-B and pro-B cells. Furthermore, CellTypist clearly resolved monocytes and

macrophages, which often form a transcriptomic continuum in scRNA-seq datasets owing to their functional plasticity. Thus, CellTypist was successfully able to identify major groups of cell populations with different abundances in our dataset (fig. S1C).

To automatically annotate fine-grained immune subpopulations, we next applied the low-hierarchy (high-resolution, 91 cell types and subtypes) classifier, which was able to classify cells into 43 specific subtypes, including subsets of T cells, B cells, ILCs, and mononuclear phagocytes (Fig. 1E). This revealed a high degree of heterogeneity within the T cell compartment, not only distinguishing between  $\alpha\beta$  and  $\gamma\delta$  T cells but also between CD4 $^{+}$  and CD8 $^{+}$  T cell subtypes and their more detailed effector and functional phenotypes. Specifically, the CD4 $^{+}$  T cell cluster was classified into helper, regulatory, and cytotoxic subsets,



**Fig. 2. Myeloid compartment across tissues.** (A) UMAP visualization of the cell populations in the myeloid compartment. (B) Dot plot for expression of marker genes of the identified myeloid populations. Color represents maximum-normalized mean expression of cells expressing marker genes, and size represents the percentage of cells expressing these genes. (C) UMAP visualization of the tissue distribution in the myeloid compartment. (D) Heatmap showing the distribution of each myeloid cell population across different tissues. Cell numbers are normalized within each tissue and later calculated as

proportions across tissues. Only tissues containing >50 myeloid cells in at least two donors were included. Asterisks mark significant enrichment in a given tissue relative to the remaining tissues [Poisson regression stratified by donors,  $P < 0.05$  after Benjamini-Hochberg (BH) correction]. (E) Violin plot for genes differentially expressed in migratory dendritic cells across tissues. Color represents maximum-normalized mean expression of cells expressing marker genes. (F) smFISH visualization of *ITGAX*, *CCR7*, and *AIRE* transcripts, validating the *AIRE*<sup>+</sup> migratory dendritic cells in lung-draining lymph nodes.

and the CD8<sup>+</sup> T cell clusters contained unconventional T cell subpopulations such as MAIT cells. In the B cell compartment, a clear distinction was observed between naïve and memory B cells. Moreover, CellTypist revealed three distinct subsets of dendritic cells (DCs)—DC1, DC2, and migratory DCs (migDCs) (25, 26)—again highlighting the granularity that CellTypist can achieve. This fine-grained dissection of each compartment also allowed for the cross-validation and consolidation of cell types by examining the expression of marker genes derived from CellTypist models in cells from our dataset (fig. S1D).

To summarize, we generated an in-depth map of immune cell populations across human tissues and developed a framework for automated annotation of immune cell types and subtypes. CellTypist produced fine-grained annotations on our multitissue and multilineage dataset, and its performance, as assessed on multiple metrics, was comparable to or better than other label-transfer methods, with minimal computational cost (figs.

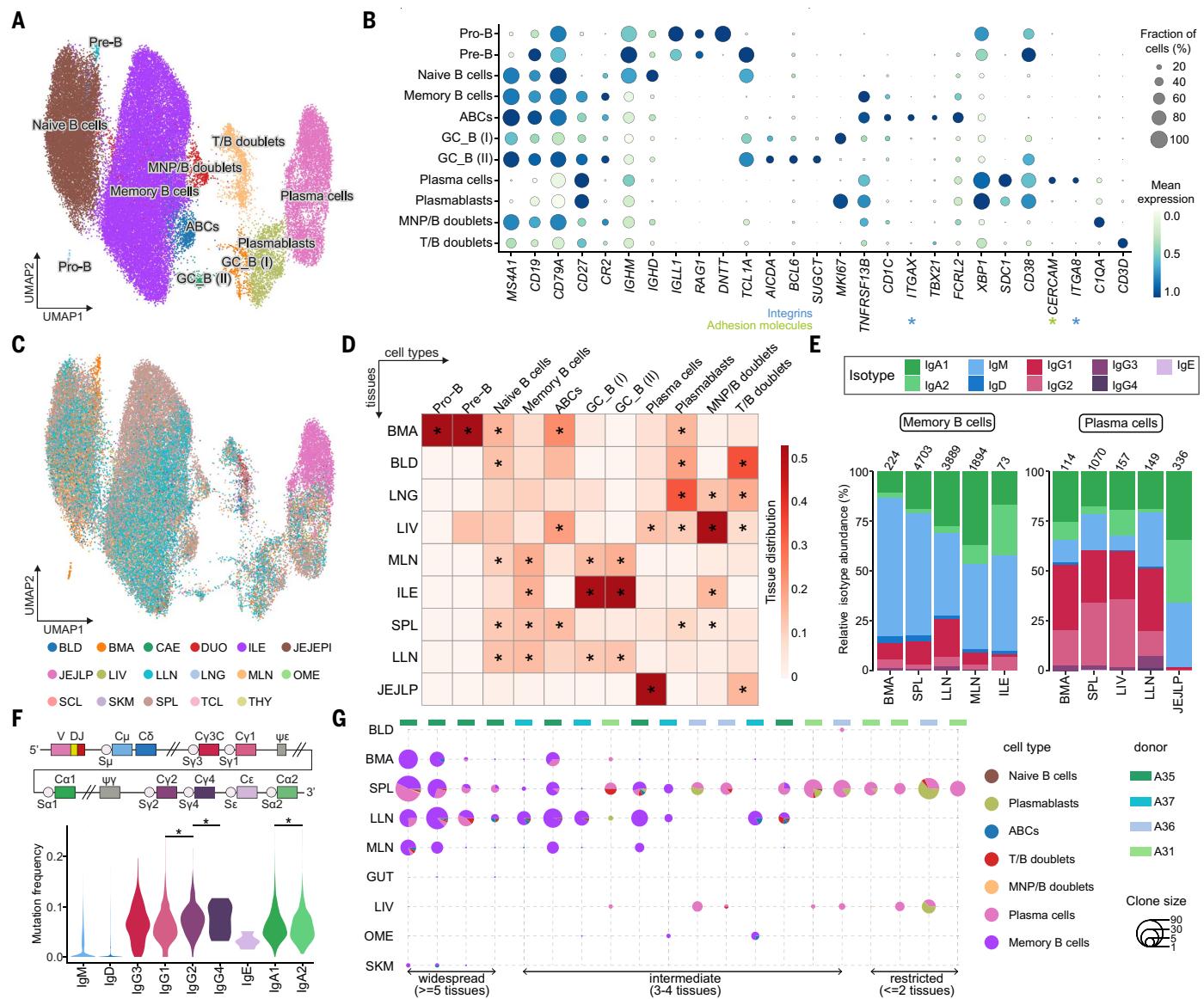
S13 and S14). This approach allowed us to identify fine-grained cell subtypes, such as the progenitors and dendritic cell compartments at full transcriptomic breadth, resolving 43 cell states in total across our dataset. This automated annotation forms the basis for further cross-tissue comparisons of cell compartments in the sections below.

#### Tissue-restricted features of mononuclear phagocytes

Mononuclear phagocytes, including monocytes, macrophages, and dendritic cells, are critical for immune surveillance and tissue homeostasis. Automatic annotation by CellTypist identified eight populations in this compartment (fig. S15A). To explore macrophage heterogeneity further, we built upon CellTypist's annotation by performing additional manual curation, which revealed further heterogeneity within the macrophages (Fig. 2A and fig. S15B). The identities of these cells were supported by expression of well-established marker genes (Fig. 2B) and by markers in-

dependently identified from CellTypist models (fig. S15C). Moreover, existence of these cell types was cross-validated, and thus consolidated, in the training datasets of CellTypist (fig. S15D), as well as in myeloid cells from two additional studies of the human intestinal tract (27) (fig. S15E) and lung (28) (fig. S15F).

Among macrophages, lung-resident cells constituted the majority and were classified into two major clusters: (i) alveolar macrophages expressing *GPNMB* and *TREM2*, markers that have been related to alveolar macrophages (29) and disease-associated macrophages (30), respectively; and (ii) intermediate macrophages with distinctive expression of *TNIP3* (Fig. 2, B to D). *TNIP3* (TNFAIP3-interacting protein 3) binds to A20, encoded by *TNFAIP3*, and inhibits tumor necrosis factor, interleukin-1 (IL-1), and lipopolysaccharide-induced nuclear factor κB (NF-κB) activation (31). Its expression in lung macrophages may be related to underlying pathology, as it was primarily detected in one donor (A29), a multitrauma donor with



**Fig. 3. B cell compartment across tissues.** (A) UMAP visualization of the cell populations in the B cell compartment. (B) Dot plot for expression of marker genes of the identified B cell populations. Color represents maximum normalized mean expression of cells expressing marker genes, and size represents the percentage of cells expressing these genes. (C) UMAP visualization of the tissue distribution in the B cell compartment. (D) Heatmap showing the distribution of each B cell population across different tissues. Cell numbers are normalized within each tissue and later calculated as proportions across tissues. Only tissues containing >50 B cells

in at least two donors were included. Asterisks mark significant enrichment in a given tissue relative to the remaining tissues (Poisson regression stratified by donors,  $P < 0.05$  after BH correction). (E) Stacked bar plots showing the isotype distribution per tissue within memory B cells and the plasma cells. (F) Violin plot of the hypermutation frequency on the IgH chain across isotypes. Significant difference among IgG4, IgG2, and IgG1 and between IgA2 and IgA1 is marked by asterisks (Wilcoxon rank sum test,  $P < 0.05$ ). (G) Scatterpie plot showing the tissue distribution and B cell subsets of expanded clonotypes (>10 cells). Each vertical line represents one clonotype.

substantial lung contusions. Notably, this population also expressed the monocyte-recruiting chemokine CCL2 (Fig. 2B), providing a means of replenishing the lung macrophage pool.

Other macrophage subsets in our data also showed a high degree of tissue restriction (Fig. 2D). Erythrophagocytic macrophages, including red pulp macrophages and Kupffer cells, mainly populated the spleen and liver, as expected, and shared high expression of

*CD5L*, *SCL40A1*, and the transcription factor *SPIC* (32). Notably, a number of macrophages from lymph nodes clustered together with erythrophagocytic macrophages, pointing to the presence of a small number of iron-recycling macrophages in lymph nodes (Fig. 2D). Another macrophage population specifically present in the gut expressed *CD209* (encoding DC-SIGN) and *IGFL1*, markers that have been previously reported in mature intestinal macrophages and M2-like macrophages,

respectively (33, 34). Lastly, monocytes were clearly grouped in two major subgroups, classical and nonclassical monocytes, which differed in the expression of *CD14*, *FCGR3A*, and *CX3CR1* and in their tissue distribution (Fig. 2, A to D).

Among dendritic cells, DC1 expressed *XCR1* and *CLEC9A*, consistent with their identity as conventional DCs (DC1), specialized for cross-presentation of antigens (Fig. 2B). Conventional DC2s expressed *CD1C* and *CLEC10A*, and migDCs were *CCR7*<sup>+</sup>*LAMP3*<sup>+</sup>. *CCR7* is up-regulated in

tissue DCs following Toll-like receptor or Fc-gamma receptor ligation (35, 36), enabling migration toward CCL19/21-expressing lymphatic endothelium and stromal cells in the T cell zone of lymph nodes (37, 38). Consistent with this, we observed a marked enrichment of *CCR7<sup>+</sup>* migDCs in lung-draining and mesenteric lymph nodes (Fig. 2D). Notably, migDCs showed up-regulation of *AIRE*, *PDLIM4*, and *EBI3* in lymph nodes (Fig. 2E). Extrathymic expression of the autoimmune regulator *AIRE* has been reported in humans (39, 40), however, its functional role in secondary lymphoid organs remains poorly understood and is a matter of intense research (41–43). We validated the presence of migDCs in lung-draining lymph nodes by immunofluorescence (fig. S16A) and *AIRE* expression by single-molecule fluorescence *in situ* hybridization (smFISH) (Fig. 2F). In addition, another subpopulation of migDCs found in lung and lung-draining lymph nodes up-regulated *CRLF2* [encoding thymic stromal lymphopoietin receptor (TSLPR)], chemokines (*CCL22* and *CCL17*), *CSF2RA*, and *GPR157* (Fig. 2E). TSLPR is involved in the induction of T helper 2 (T<sub>H</sub>2) cell responses in asthma (44). Expression of these genes in lung DCs was also observed in published scRNA-seq datasets (45, 46) (fig. S16, B and C). These observations suggest that dendritic cell activation coincides with the acquisition of tissue-specific markers that differ depending on the local microenvironment.

Overall, our analysis of the myeloid compartment has revealed shared and tissue-restricted features of mononuclear phagocytes, including alveolar macrophages in the lung, iron-recycling macrophages mostly localized in the spleen, and subtypes of migratory dendritic cells.

#### B cell subsets and immunoglobulin repertoires across tissues

B cells comprise progenitors in the bone marrow, developmental states in lymphoid tissues, and terminally differentiated memory and plasma cell states in lymphoid and nonlymphoid tissues. They play a central role in humoral immunity via the production of antibodies tailored to specific body sites. We first annotated the B cells using CellTypist and obtained six populations (fig. S17A). Manual curation revealed further heterogeneity in memory B cells and plasma cells, identifying 11 cell types in total (Fig. 3A and fig. S17B). As with the myeloid compartment, we cross-checked and verified these cell types in CellTypist training datasets (fig. S17, C and D) and in two independent immune datasets from gut (27) and lung (28) (fig. S17, E and F).

Naïve B cells expressed *TCL1A* and were primarily found in lymphoid tissues (Fig. 3, B to D). In addition, we identified two populations of germinal center B cells, express-

ing *AICDA* and *BCL6*, that differed in their proliferative states (marked by *MKI67*). Of note, we did not find differential expression of dark zone and light zone marker genes in these two populations, probably reflecting limited germinal center activity in our adult donors. Moreover, these germinal center populations were present in lymph nodes and diverse gut regions (Fig. 3, C and D), presumably representing Peyer's patches. Within memory B cells, which were characterized by expression of the B cell lineage markers (*MS4A1* and *CD19*) and *TNFRSF13B*, we found a transcriptomically distinct cluster positive for *ITGAX*, *TBX21*, and *FCRL2*, encoding CD11c, T-bet, and the Fc receptor-like protein 2, respectively (Fig. 3B). CD11c<sup>+</sup>T-bet<sup>+</sup> B cells, also known as age-associated B cells (ABCs), have been reported in autoimmunity and aging (47–49) and likely correspond to this *ITGAX<sup>+</sup>* memory B cell population. Unlike conventional memory B cells, ABCs showed relatively low expression of *CR2* (encoding CD21) and *CD27* (Fig. 3B). This subset was mainly present in the liver and bone marrow, while in secondary lymphoid organs, it was primarily found in the spleen, as confirmed by flow cytometry and immunofluorescence (Fig. 3, C and D, and fig. S18). This data deepens our understanding of the phenotype of this non-classical subtype of memory B cells and their tissue distribution.

We uncovered two populations of plasmablasts and plasma cells marked by expression of *CD38*, *XBP1*, and *SDC1*. Whereas the former expressed *MKI67* and were found in the spleen, liver, bone marrow, and blood, the latter expressed the integrin alpha-8-encoding gene *ITGA8* and the adhesion molecule *CERCAM* and were enriched in the jejunum and liver (Fig. 3, B to D). *ITGA8<sup>+</sup>* plasma cells have recently been reported in the context of an analysis of bone marrow plasma cells (50) and are likely a long-lived plasma cell population that is quiescent and tissue resident. Here we expand their tissue distribution to the liver and gut and describe their specific clonal distribution pattern.

B cells have an additional source of variability due to VDJ recombination, somatic hypermutation, and class switching, which can relate to cell phenotype. Therefore, we performed targeted enrichment and sequencing of B cell receptor (BCR) transcripts to assess isotypes, hypermutation levels, and clonal architecture of the B cell populations identified above. This analysis revealed an isotype and subclass usage pattern that related to cellular phenotype (fig. S19A). As expected, naïve B cells mainly showed a subclass preference for immunoglobulin M (IgM) and IgD. Notably, while evidence of class switching to IgA1 and IgG1 was seen within memory B cells (including ABCs), plasmablasts and

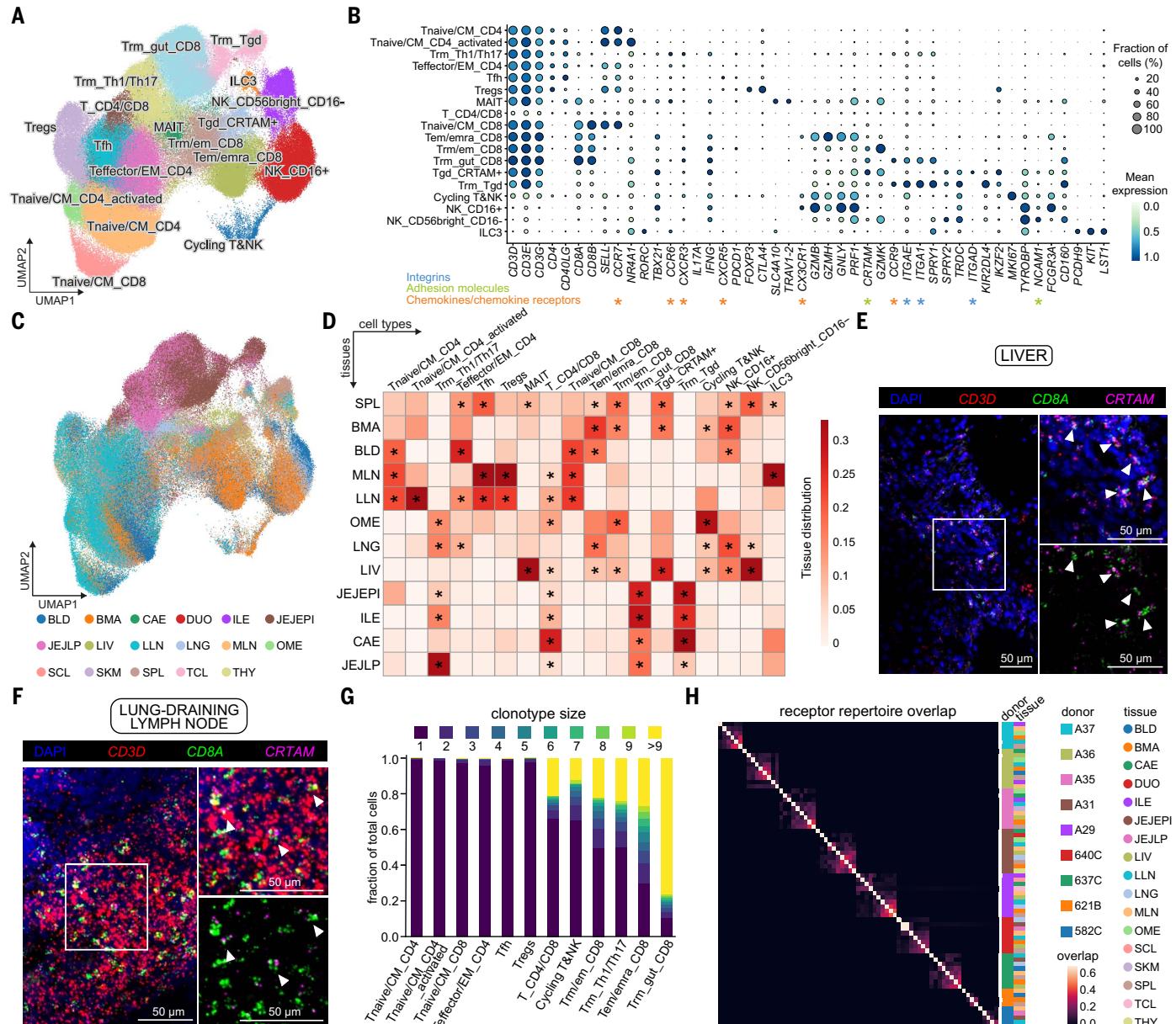
plasma cells also showed class switching to IgA2 and IgG2.

To determine to what extent this isotype subclass bias correlated with the tissue of origin, we assessed each cell state independently (requiring a minimum cell count of 50). Memory B cells showed a bias toward IgA1 in the mesenteric lymph nodes, and toward IgA2 in the ileum, where Peyer's patches are found (Fig. 3E). In the plasma cell compartment, we found an even more pronounced preference for IgA2 in the gut region (specifically in the jejunum lamina propria), consistent with the known dominance of this isotype at mucosal surfaces (Fig. 3E). Of note, plasma cells in the bone marrow, liver, and spleen were composed of >20% IgG2 subclass. With more limited numbers, we reported isotype distributions across tissues for other B cell populations (fig. S19, B and C). ABCs were dominated by IgM in both the spleen and lung-draining lymph nodes, consistent with previous reports (51).

Somatic hypermutation levels were, as expected, lowest in naïve B cells and highest in plasma cells (fig. S19D). Between isotypes and subclasses, somatic hypermutation did not differ significantly. Nonetheless, there was a tendency toward higher mutation rates in the distal classes IgG2, IgG4, and IgA2, which are downstream from the IgH locus and can thus accumulate more mutations during sequential switching (52) (Fig. 3F). We also explored the occurrence of sequential class switching events in our data by assessing the isotype frequency among expanded clonotypes (>10 cells). Mixed isotype clones were rare in our data (fig. S19E).

Next, we evaluated the distribution of expanded clones across tissues and cell types and found three major groups of clones—those present in only two tissues, three to four tissues, or five or more tissues, respectively (Fig. 3G), similar to previously reported patterns of B cell clonal tissue distribution (53). While the clones restricted to two tissues, typically between the spleen and the liver or bone marrow, were enriched in plasma cells, those distributed across more than five tissues, including lymph nodes, were over-represented in memory B cells. Together, these findings suggest that tissue-restricted clones may represent long-term immunological memory maintained by long-lived plasma cells resident in the bone marrow and spleen as well as liver in our data.

Overall, we characterized nine cell states in the B cell compartment and gained insights from in-depth characterization of both naïve and memory B cell as well as plasma cell subsets. We identified distinct clonal distribution patterns for the more tissue-restricted long-lived quiescent plasma cells versus the broad tissue distribution of classical memory B cell clones.



**Fig. 4. Tissue compartmentalization and site-specific adaptations of T cells and innate lymphoid cells.** (A) UMAP visualization of T cells and ILCs across human tissues colored by cell types. (B) Dot plot for expression of marker genes of the identified immune populations. Color represents maximum-normalized mean expression of cells expressing marker genes, and size represents the percentage of cells expressing these genes. (C) UMAP visualization of T cells and ILCs colored by tissues. (D) Heatmap showing the distribution of each T cell or ILC population across different tissues. Cell numbers are normalized within each tissue and later calculated as proportions across tissues. Only tissues

containing >50 ILC or T cells in at least two donors were included. Asterisks mark significant enrichment in a given tissue relative to the remaining tissues (Poisson regression stratified by donors,  $P < 0.05$  after BH correction). (E and F) smFISH visualization of *CD3D*, *CD8A*, and *CRTAM* transcripts, validating the tissue-resident memory  $CD8^+$  T cell population in the liver and lung-draining lymph nodes. (G) TCR repertoire analysis of T cells across tissues. Stacked bar plot shows the fraction of cells in a given cluster binned by clonotype size. (H) Heatmap showing the repertoire overlap between expanded clones (>1 cell) across donors as determined by Jaccard distance.

#### scRNA-seq analysis of T cells and innate lymphocytes reveals lineage and tissue-specific subsets

For annotation of the T cell and innate lymphocyte compartment, CellTypist identified 18 cell types (fig. S20A). After manual inspection, these clusters were further divided into additional subtypes (e.g., for cy-

totoxic T cells) (Fig. 4A and fig. S20B). As described above for the myeloid and B cell compartments, identities of the derived cell types were cross-validated in the immune compartment of the CellTypist training datasets (fig. S20, C and D) and the two independent studies of gut (27) and lung (28) (fig. S20, E and F).

Naïve/central memory  $CD4^+$  T cells were transcriptionally close to naïve  $CD8^+$  T cells as defined by high expression of *CCR7* and *SELL* and were mainly found in lymphoid sites (Fig. 4B). Other  $CD4^+$  T cells identified included T follicular helper ( $T_{FH}$ ) cells expressing *CXCR5*; regulatory T cells ( $T_{regs}$ ) expressing *FOXP3* and *CTLA4*; effector memory

CD4<sup>+</sup> T cells; and tissue-resident memory T<sub>H1</sub> and T<sub>H17</sub> cells expressing *CCR9*, *ITGAE*, and *ITGA1* found largely in intestinal sites (jejunum and ileum) and lungs (Fig. 4, B to D). Within the memory CD8<sup>+</sup> T compartment, we found three major subsets: T<sub>RM</sub>\_gut\_CD8 (resident memory T cell, T<sub>RM</sub>), T<sub>EM/EMRA</sub>\_CD8 (effector memory T cell, T<sub>EM</sub>; effector memory reexpressing CD45RA, T<sub>EMRA</sub>), and T<sub>RM/EM</sub>\_CD8. These subsets were characterized by differential expression of the chemokine receptors *CCR9* and *CX3CR1* and the activation marker *CRTAM* (Fig. 4B). The T<sub>RM</sub>\_gut\_CD8 population (*CCR9*<sup>+</sup>) expressed the tissue-residency markers *ITGAE* and *ITGA1*, encoding CD103 and CD49a, respectively, and localized to intestinal sites (Fig. 4, B to D). By contrast, the T<sub>EM/EMRA</sub>\_CD8 population expressing *CX3CR1* was found in blood-rich sites (blood, bone marrow, lung, and liver) and was excluded from lymph nodes and gut (Fig. 4, C and D), consistent with previous flow cytometry analysis of T<sub>EMRA</sub> cells (54) and results showing CX3CR1<sup>+</sup>CD8<sup>+</sup> T cells as blood-confined and absent from thoracic duct lymph (55). The T<sub>RM/EM</sub>\_CD8 population expressed high levels of *CRTAM*, a gene previously shown to be expressed by T<sub>RM</sub> cells (56) and was found in spleen, bone marrow, and to a lesser extent in lymph nodes and lungs. This may be a resident population more prevalent in lymphoid sites (16). We validated and mapped the T<sub>RM/EM</sub>\_CD8 population using smFISH in the liver (Fig. 4E) and lung-draining lymph nodes (Fig. 4F). Furthermore, we validated all three memory CD8<sup>+</sup> T cell populations at the protein level using flow cytometry of cells purified from human spleen (fig. S21). Although we could validate *CRTAM* at the RNA level by smFISH, the protein could not be detected without stimulation, suggesting that *CRTAM* is subject to posttranslational regulation upon T cell receptor (TCR) activation. These three distinct populations represent different states of tissue adaptation and maturation between effector memory and tissue-resident T cell memory states.

We also detected invariant T cell subsets such as MAIT cells, characterized by expression of *TRAV1-2* and *SLC4A10*, and two populations of  $\gamma\delta$  T cells: T<sub>RM</sub>\_T<sub>GD</sub> and T<sub>GD</sub>\_CRTAM<sup>+</sup>. The *CCR9*<sup>+</sup> T<sub>RM</sub>\_T<sub>GD</sub> population populated the gut and expressed the tissue-residency markers *ITGAE* and *ITGA1*, whereas the T<sub>GD</sub>\_CRTAM<sup>+</sup> population overexpressed *CRTAM*, *IKZF2* (encoding HELIOS), and the integrin molecule *ITGAD* (encoding CD11d) and was found primarily in the spleen, bone marrow, and liver (Fig. 4, B to D, and fig. S22, A and B). We validated the latter population by quantitative PCR (qPCR) of flow-sorted CD3<sup>+</sup>TCR $\gamma\delta$ <sup>+</sup> and CD3<sup>+</sup>TCR $\alpha\beta$ <sup>+</sup> cells from cryopreserved spleen samples from three donors (fig. S22, C

and D). As a small fraction of  $\alpha\beta$  T cells, marked by low expression of CD52 and CD127, were also noted to express *ITGAD*, the CD3<sup>+</sup>TCR $\alpha\beta$  population was split into CD52<sup>-</sup>CD127<sup>-</sup> and CD52<sup>+</sup>CD127<sup>+</sup> subpopulations. In keeping with our scRNA-seq data, *ITGAD* expression was high in CD3<sup>+</sup>TCR $\gamma\delta$  and CD52<sup>-</sup>CD127<sup>-</sup>CD3<sup>+</sup>TCR $\alpha\beta$ , providing additional evidence for the specific expression of this integrin alpha subunit in this subpopulation of  $\gamma\delta$  T cells.

Lastly, NK cells in our data were represented by two clusters with high expression of either *FCGR3A* (encoding CD16) or *NCAM1* (encoding CD56). We also defined an ILC3 population within a small cluster mixed with NK cells, via expression of markers including *PCDH9* (Fig. 4, A and B). Analyses of the tissue distribution of these populations revealed that, whereas most of the CD4<sup>+</sup> T and ILC3 cells were located in the lymph nodes and to some extent in the spleen, cytotoxic T and NK cells were more abundant in the bone marrow, spleen, and nonlymphoid tissues (Fig. 4, C and D).

#### TCR repertoire analysis shows clonal expansion and distribution patterns within and across tissues

To understand T cell-mediated protection in more depth, we analyzed T cell clonal distribution in a subset of the data within different tissues of a single individual and across different individuals. Chain pairing analysis showed that cells from the T cell clusters mostly contained a single pair of TCR $\alpha\beta$  chains (50 to 60%), with orphan (5 to 20%) and extra (5 to 10%) chains present in small fractions of cells (fig. S23A). Notably, the frequency of extra  $\alpha$  chains (extra VJ) was higher than that of  $\beta$  chains (extra VDJ), potentially because of more stringent and multilayered allelic exclusion mechanisms at the TCR $\beta$  locus compared with TCR $\alpha$  (57). As expected, the NK and ILC clusters held no productive TCR chains. Within the  $\gamma\delta$  T cell clusters, only a small proportion had a productive TCR chain, which may result from cytotoxic T cells co-clustering with  $\gamma\delta$  T cells. We also carried out  $\gamma\delta$  TCR sequencing in selected spleen, bone marrow, and liver samples. The  $\gamma\delta$  TCR sequencing data was subjected to a customized analysis pipeline that we developed and optimized using cellranger followed by contig reannotation with dandelion (see materials and methods), facilitating the full recovery of  $\gamma\delta$  chains in our data. This analysis confirmed that most of the productive  $\gamma\delta$  TCR chains originated from the *ITGAD*-expressing  $\gamma\delta$  T cells (fig. S23B), supporting the robust identification of this population. The T<sub>RM</sub>\_T<sub>GD</sub> population could not be confirmed by  $\gamma\delta$  TCR sequencing owing to the lack of sample availability.

We next examined V(D)J gene usage in relation to T cell identity. In the MAIT population,

we detected significant enrichment of *TRAV1-2*, as expected (fig. S23C). Selecting only the *TRAV1-2*<sup>+</sup> cells (MAIT cluster and other clusters) revealed a notable tissue-specific distribution of *TRAJ* segments, with *TRAJ33* in spleen and liver, *TRAJ12* in liver, and *TRAJ29/TRAJ36* in jejunum (fig. S23D). This suggests that there may be different antigens for MAIT cells in the spleen, liver, and gut corresponding to the different metabolomes in these tissues. In addition, full analysis of the TCR repertoire of the MAIT cells revealed previously unappreciated diversity of V segment usage in the  $\beta$  chain (fig. S23D).

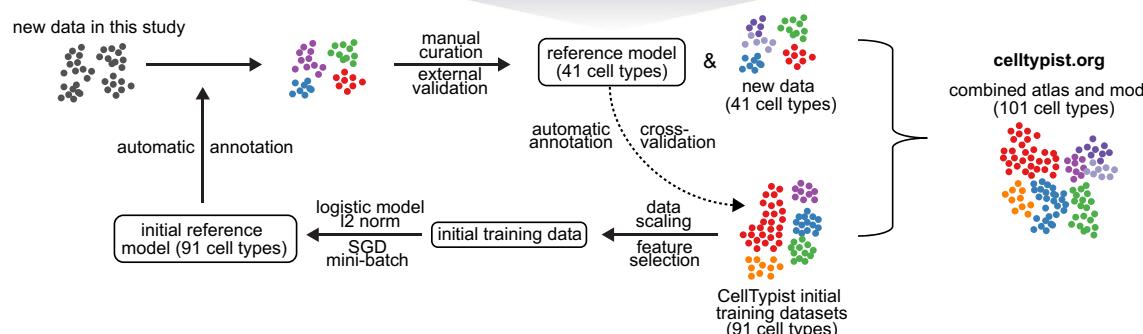
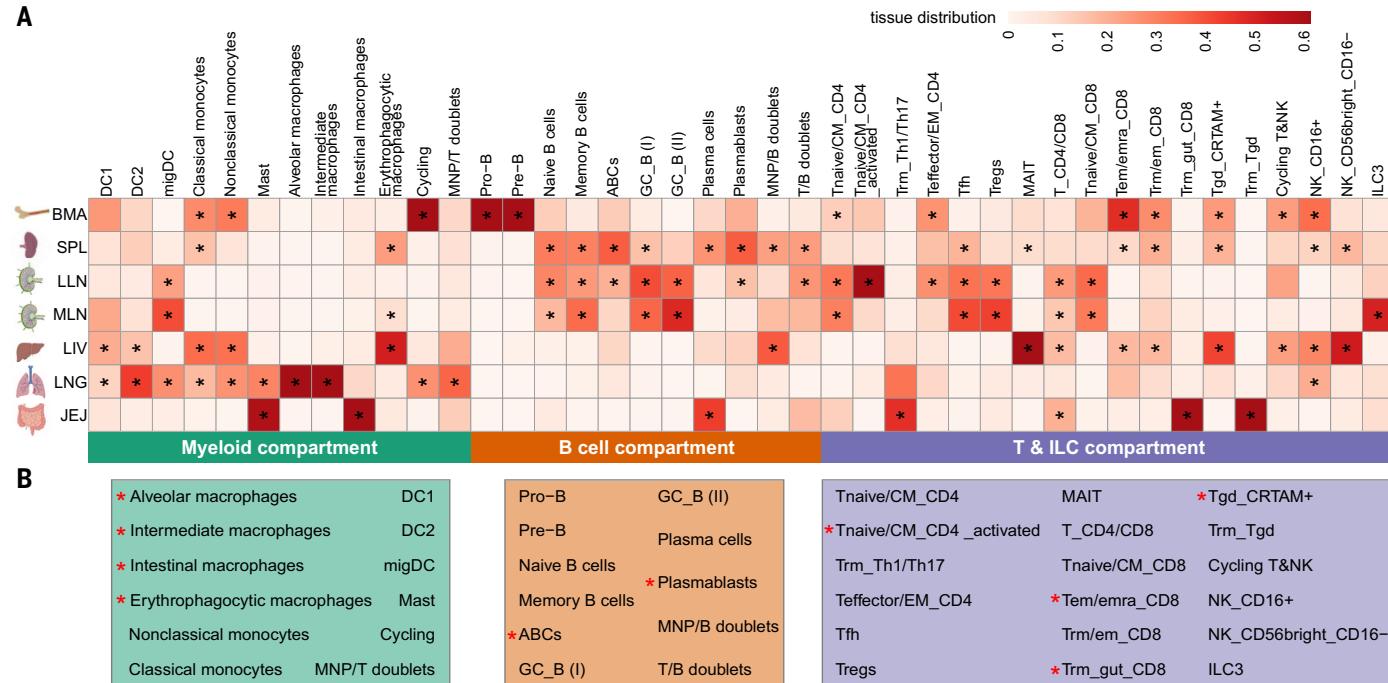
We then defined clonally related cells on the basis of identical CDR3 nucleotide sequences to investigate their TCR repertoires. Using this approach, we found that clonally expanded cells were primarily from the resident memory T cell compartment, including the T<sub>H1</sub> and T<sub>H17</sub> populations mentioned above (Fig. 4G and fig. S23E). As expected, these clonotypes were restricted to single individuals, and within an individual they were distributed across tissues and subsets (Fig. 4H and fig. S23, F to H). We found a small number of isolated CD4<sup>+</sup> T cell clones that shared T<sub>regs</sub> and effector T cell phenotypes, possibly owing to low levels of plasticity or to (trans)differentiation from the same naïve precursor cell in the periphery (fig. S23H). Focusing on the most expanded clonotypes (>20 cells), many were widespread across five or more tissues, supporting the systemic nature of tissue-resident immune memory (Fig. 4G). Moreover, we found that several clonotypes present across tissues consisted of a mixture of cells from the T<sub>EM/EMRA</sub>\_CD8 and T<sub>RM/EM</sub>\_CD8 populations (fig. S23H), suggesting that a single naïve CD8<sup>+</sup> T cell precursor can give rise to diverse cytotoxic T cell states, which harbor immune memory across multiple nonlymphoid tissues, emphasizing the plasticity of phenotype and location within a clone.

In summary, we have described 18 T or innate cell states in our data by integrating CellTypist logistic regression models, manual curation, and V(D)J sequencing. This analysis has yielded insights into the MAIT cell compartment and its antigen receptor repertoire distribution that differed between spleen, liver, and gut. For the cytotoxic T cell memory compartment, there was broad sharing of clones across gut regions for T<sub>RM</sub>\_gut\_CD8, and mixed T<sub>EM/EMRA</sub>\_CD8 and T<sub>RM/EM</sub>\_CD8 T cell clonotypes with broad tissue distributions.

#### A cross-tissue updatable reference of immune cell types and states

After focusing on individual immune compartments, we next took a combined approach to better understand the immune landscape of selected tissues. As shown in Fig. 5A, each tissue has its own immune neighborhood, for

A



**Fig. 5. A cross-tissue updatable reference of immune cell types and cell states.** **(A)** Heatmap showing the distribution of manually curated cell types across selected tissues. Cell numbers are normalized within each tissue and later calculated as proportions across tissues. Asterisks mark significant enrichment in a given tissue relative to the remaining tissues (Poisson regression stratified by donors,  $P < 0.05$  after BH correction). **(B)** Workflow for the iterative update of CellTypist through the periodic incorporation of curated cell type labels.

example, while spleen and lymph nodes are rich in B cells, composition of their myeloid compartment varies. In particular, a large population of erythropagocytic macrophages, known as red pulp macrophages, are evident in the spleen (in keeping with their role in red blood cell turnover), whereas lymph nodes are rich in dendritic cells. As expected, bone marrow uniquely contains progenitor populations. Furthermore, lung and liver contain notable numbers of monocytes, including *CX3CR1*<sup>+</sup> nonclassical monocytes, whereas these cells are absent from the jejunum, perhaps reflecting different degrees of vascularization. In contrast, the jejunum has an abundance of resident memory T cells (CD8<sup>+</sup> T cells, T<sub>H</sub>1 cells, and T<sub>H</sub>17 cells) as well as plasma cells.

Our long-term vision for CellTypist is to provide a reference atlas with deeply curated cell types publicly available to the community. Therefore, via a semiautomatic process, we fed the identities of the 41 immune cell types identified in our dataset (including both shared and novel cell type labels) back into CellTypist, demonstrating how CellTypist can be updated and improved over time. Combined with the initial 91 cell types and states included in the reference datasets, CellTypist now comprises 101 annotated cell types (Fig. 5B).

## Discussion

Here, we present a multitissue study of immune cells across the human body in diverse organ donors. By sampling multiple organs

from the same individuals, which allows for robust control of age, sex, medical history, drug exposure and sampling backgrounds, we reveal tissue-specific expression patterns across the myeloid and lymphoid compartments.

We also introduce CellTypist, a publicly available and updatable framework for automated immune cell type annotation that, in addition to identifying major cell types, can perform fine-grained cell subtype annotation—normally a time-consuming process that requires expert knowledge. We developed CellTypist by integrating and curating data obtained from 19 studies performed across a range of tissues, with in-depth immune cell analysis of 91 harmonized cell type labels. However, as demonstrated here, for example

in the  $\gamma\delta$  T cell compartment, manual curation after automated annotation still has a role to play in revealing specific cell subtypes that may be absent from the training set. To reduce the need for this manual curation, in the longer term, the CellTypist models will be periodically updated and extended to include further immune and nonimmune subpopulations as more data become available.

Within the myeloid compartment, macrophages showed the most prominent features of tissue specificity. Erythrophagocytic macrophages in the liver and spleen shared features related to iron recycling (58) with macrophages in other locations, such as the mesenteric lymph nodes, suggesting that macrophages participate in iron metabolism across a range of tissues. In addition, we characterized subsets of migratory dendritic cells ( $CCR7^+$ ) revealing specific expression of *CRLF2*, *CSF2RA*, and *GPR157* in the lung and lung-draining lymph nodes and expression of *AIRE* in the mesenteric and lung-draining lymph nodes. These migratory dendritic cell states are interesting targets for future in-depth functional characterization in the context of allergy, asthma, and other related pathologies (59, 60).

In the lymphoid compartment, we combined single-cell transcriptome and VDJ analysis, which allowed the phenotype of adaptive immune cells to be dissected using complementary layers of single-cell genomics data. Of note, we detected a subset of memory B cells expressing *ITGAX* (CD11c) and *TBX21* (T-bet) that resemble ABCs previously reported to be expanded in aging (48), following malaria vaccination (61) and in systemic lupus erythematosus patients (62). In our data, these B cells did not show clonal expansion, and at least 50% showed IgM subclass, suggesting that they may be present at low levels in healthy tissues and expand upon challenge as well as aging. BCR analysis revealed isotype usage biased toward IgA2 in gut plasma cells, which may be related to structural differences (63) or higher resistance to microbial degradation as compared with IgA1 (64).

In the T cell compartment, our results provided insights into the heterogeneity of T cell subtypes and their tissue adaptations. Notably, we identified subsets of  $CD4^+ T_{RM}$  on the basis of functional capacity for interferon- $\gamma$  or IL-17 production that were mostly localized to intestinal sites, analogous to mouse  $CD4^+ T_{RM}$  generated from IL-17-producing effector T cells in the gut (65). We also identified different subsets of  $CD8^+ T_{RM}$  including a gut-adapted subset expressing *CCR9*, which mediates homing to intestinal sites via binding to *CCL25* (66) and another  $T_{RM}$ -like subset more targeted to lymphoid sites. TCR clone sharing between memory subtypes of  $CD8^+$  T cells suggests their origin from a

common precursor, or their differentiation or conversion during migration or maintenance, such as conversion of  $T_{EM}$  cells to  $T_{RM}$  cells (56). We also identified distinct subsets of  $\gamma\delta$  T cells on the basis of tissue-specific gene expression patterns, showing distinct integrin gene expression and tissue distributions.

In summary, using this dataset of nearly 360,000 single-cell transcriptomes (of which ~330,000 were immune cells) from donor-matched tissues from 12 deceased individuals, we have shown how a combination of CellTypist-based automated annotation, expert-driven cluster analysis, and antigen receptor sequencing can synergize to dissect specific and functionally relevant aspects of immune cells across the human body. We have revealed previously unrecognized features of tissue-specific immunity in the myeloid and lymphoid compartments, and have provided a comprehensive framework for future cross-tissue cell type analysis. Further investigation of human tissue-resident immunity is needed to determine the effect of important covariates such as underlying critical illness, donor age, and gender, as well as considering the immune cell activation status, to gain a defining picture of how human biology influences immune functions. Our deeply characterized cross-tissue immune cell dataset has implications for the engineering of cells for therapeutic purposes and addressing cells to intended tissue locations, as well as for understanding tissue-specific features of infection as well as distinct modes of vaccine delivery to tissues.

## Materials and methods summary

### *Tissue acquisition, processing, and single-cell sequencing*

Tissue was obtained from deceased organ donors through the Cambridge Biorepository for Translational Medicine (CBTM, <https://www.cbtm.group.cam.ac.uk/>), REC 15/EE/0152. Detailed sample locations taken can be found in Fig. 1, and protocols are described in detail in supplementary materials. Additional tissue samples were from Columbia University and were obtained from deceased organ donors at the time of organ acquisition for clinical transplantation through an approved protocol and material transfer agreement with LiveOnNY.

Six donors were processed with a uniform protocol at Cambridge University, where solid tissues were cut into small pieces and then homogenized with two rounds of enzymatic digestion at 37°C for 15 min with mixing. The remaining six donors were subjected to a tissue-adapted protocol with the aim of improving immune cell recovery, and this protocol was harmonized as closely as possible between the two collection sites.

For scRNA-seq experiments, single cells were loaded onto the channels of a Chromium chip (10x Genomics). cDNA synthesis, amplification, and sequencing libraries were generated using either the Single Cell 5' Reagent (v1 and v2) (Cambridge University) or 3' Reagent (v3) (Columbia University) Kit. TCR $\alpha\beta$ , BCR, and TCR $\gamma\delta$  paired VDJ libraries were prepared from samples made with the 5' Reagent kit. All libraries were sequenced on either a HiSeq 4000 or NovaSeq 6000 instrument.

### *scRNA-seq and scVDJ-seq data analysis*

scRNA-seq data was aligned and quantified using the cellranger software (version 6.1.1, 10x Genomics). Cells from hashtagged samples were demultiplexed using HashSolo (67). Cells with fewer than 1000 unique molecular identifier counts and 600 detected genes were excluded. Doublets were detected using Scrublet (68). Downstream analysis from data normalization to graph-based clustering were performed using Scanpy (version 1.6.0) (69), with details described in supplementary materials. Data integration was done using BBKNN (70) and scVI (71), and the results were compared using kBET (72).

Single-cell TCR sequencing and single-cell BCR sequencing data were aligned and quantified using the cellranger-vdj software (versions 2.1.1 and 4.0, respectively). For TCR $\gamma\delta$ , we implemented a customized pipeline ([https://sc-dandelion.readthedocs.io/en/latest/notebooks/gamma\\_delta.html](https://sc-dandelion.readthedocs.io/en/latest/notebooks/gamma_delta.html)) owing to cellranger being tuned toward alpha/beta TCR chains. Single-cell TCR sequencing analysis including productive TCR chain pairing, and clonotype detection was performed using the scirpy package (73).

### *CellTypist*

Details of CellTypist, including cross-data cell type label harmonization and automated cell annotation, can be found in the supplementary text in the supplementary materials. Briefly, immune cells from 20 tissues of 19 studies were collected and harmonized into consistent labels. These cells were split into equal-sized mini-batches, and these batches were sequentially trained by the l2-regularized logistic regression using stochastic gradient descent learning. Feature selection was performed to choose the top 300 genes from each cell type, and the union of these genes was supplied as the input for a second round of training.

### *Single-molecule FISH, flow cytometry, qPCR, and immunofluorescence*

For smFISH, samples were run using the RNAscope 2.5 LS fluorescent multiplex assay (automated). Slides were imaged on the Perkin Elmer Opera Phenix High-Content Screening System, in confocal mode with 1- $\mu$ m z-step size, using 20 $\times$  (NA 0.16, 0.299  $\mu$ m per pixel) and

40× (NA 1.1, 0.149 μm per pixel) water-immersion objectives.

For flow cytometry, mononuclear cells were either stained ex vivo or after activation with phorbol 12-myristate 13-acetate/ionomycin for 2 hours. Cells were stained with the live/dead marker Zombie Aqua for 10 min at room temperature and then washed with phosphate-buffered saline (PBS) and 0.5% fetal calf serum (FCS), with the CD8 and B cell panels of antibodies.

qPCR was performed in three spleen samples. Cells were stained with the live/dead marker Zombie Aqua for 10 min at room temperature and then washed with PBS +0.5%FCS, followed by staining with the antibodies at 4°C for 45 min. Cell sorting was performed on a BD Fusion 4 laser sorter, and RNA was extracted using a Zymo Research RNA micro kit.

For immunofluorescence, samples were fixed in 1% paraformaldehyde for 24 hours followed by 8 hours in 30% sucrose in PBS and were stained for 2 hours at room temperature with the appropriate antibodies, washed three times in PBS, and mounted in Fluoromount-G (Southern Biotech). Images were acquired using a TCS SP8 (Leica, Milton Keynes, UK) confocal microscope.

## REFERENCES AND NOTES

- D.-M. Popescu et al., Decoding human fetal liver haematopoiesis. *Nature* **574**, 365–371 (2019). doi: [10.1038/s41586-019-1652-y](https://doi.org/10.1038/s41586-019-1652-y); pmid: [31597962](https://pubmed.ncbi.nlm.nih.gov/31597962/)
- L. C. Davies, S. J. Jenkins, J. E. Allen, P. R. Taylor, Tissue-resident macrophages. *Nat. Immunol.* **14**, 986–995 (2013). doi: [10.1038/ni.2705](https://doi.org/10.1038/ni.2705); pmid: [24048120](https://pubmed.ncbi.nlm.nih.gov/24048120/)
- E. Gomez Perdiguero et al., Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. *Nature* **518**, 547–551 (2015). doi: [10.1038/nature13989](https://doi.org/10.1038/nature13989); pmid: [25470051](https://pubmed.ncbi.nlm.nih.gov/25470051/)
- X. Fan, A. Y. Rudensky, Hallmarks of tissue-resident lymphocytes. *Cell* **164**, 1198–1211 (2016). doi: [10.1016/j.cell.2016.02.048](https://doi.org/10.1016/j.cell.2016.02.048); pmid: [26967286](https://pubmed.ncbi.nlm.nih.gov/26967286/)
- D. J. Topham, E. C. Reilly, Tissue-resident memory CD8<sup>+</sup> T cells: From phenotype to function. *Front. Immunol.* **9**, 515 (2018). doi: [10.3389/fimmu.2018.00515](https://doi.org/10.3389/fimmu.2018.00515); pmid: [29632527](https://pubmed.ncbi.nlm.nih.gov/29632527/)
- D. Masopust, A. G. Soerens, Tissue-resident T cells and other resident leukocytes. *Annu. Rev. Immunol.* **37**, 521–546 (2019). doi: [10.1146/annurev-immunol-042617-053214](https://doi.org/10.1146/annurev-immunol-042617-053214); pmid: [30726153](https://pubmed.ncbi.nlm.nih.gov/30726153/)
- P. A. Szabo, M. Miron, D. L. Farber, Location, location, location: Tissue resident memory T cells in mice and humans. *Sci. Immunol.* **4**, eaas9673 (2019). doi: [10.1126/sciimmunol.aaas9673](https://doi.org/10.1126/sciimmunol.aaas9673); pmid: [30952804](https://pubmed.ncbi.nlm.nih.gov/30952804/)
- N. Alzari et al., A human liver cell atlas reveals heterogeneity and epithelial progenitors. *Nature* **572**, 199–204 (2019). doi: [10.1038/s41586-019-1373-2](https://doi.org/10.1038/s41586-019-1373-2); pmid: [31292543](https://pubmed.ncbi.nlm.nih.gov/31292543/)
- J. Zhao et al., Single-cell RNA sequencing reveals the heterogeneity of liver-resident immune cells in human. *Cell Discov.* **6**, 22 (2020). doi: [10.1038/s41421-020-0157-z](https://doi.org/10.1038/s41421-020-0157-z); pmid: [32351704](https://pubmed.ncbi.nlm.nih.gov/32351704/)
- J.-E. Park et al., A cell atlas of human thymic development defines T cell repertoire formation. *Science* **367**, eaay3224 (2020). doi: [10.1126/science.aay3224](https://doi.org/10.1126/science.aay3224); pmid: [32079746](https://pubmed.ncbi.nlm.nih.gov/32079746/)
- K. R. James et al., Distinct microbial and immune niches of the human colon. *Nat. Immunol.* **21**, 343–353 (2020). doi: [10.1038/s41590-020-0602-z](https://doi.org/10.1038/s41590-020-0602-z); pmid: [32066951](https://pubmed.ncbi.nlm.nih.gov/32066951/)
- B. J. Stewart et al., Spatiotemporal immune zonation of the human kidney. *Science* **365**, 1461–1466 (2019). doi: [10.1126/science.aat5031](https://doi.org/10.1126/science.aat5031); pmid: [31604275](https://pubmed.ncbi.nlm.nih.gov/31604275/)
- M. T. Wong et al., A high-dimensional atlas of human T cell diversity reveals tissue-specific trafficking and cytokine signatures. *Immunity* **45**, 442–456 (2016). doi: [10.1016/j.immuni.2016.07.007](https://doi.org/10.1016/j.immuni.2016.07.007); pmid: [27521270](https://pubmed.ncbi.nlm.nih.gov/27521270/)
- P. Dogra et al., Tissue determinants of human NK cell development, function, and residence. *Cell* **180**, 749–763.e13 (2020). doi: [10.1016/j.cell.2020.01.022](https://doi.org/10.1016/j.cell.2020.01.022); pmid: [32059780](https://pubmed.ncbi.nlm.nih.gov/32059780/)
- N. A. Yudanin et al., Spatial and temporal mapping of human innate lymphoid cells reveals elements of tissue specificity. *Immunity* **50**, 505–519.e4 (2019). doi: [10.1016/j.immuni.2019.01.012](https://doi.org/10.1016/j.immuni.2019.01.012); pmid: [30770247](https://pubmed.ncbi.nlm.nih.gov/30770247/)
- B. V. Kumar, T. J. Connors, D. L. Farber, Human T cell development, localization, and function throughout life. *Immunity* **48**, 202–213 (2018). doi: [10.1016/j.immuni.2018.01.007](https://doi.org/10.1016/j.immuni.2018.01.007); pmid: [29466753](https://pubmed.ncbi.nlm.nih.gov/29466753/)
- D. L. Farber, Tissues, not blood, are where immune cells function. *Nature* **593**, 506–509 (2021). doi: [10.1038/d41586-021-01396-y](https://doi.org/10.1038/d41586-021-01396-y); pmid: [34035530](https://pubmed.ncbi.nlm.nih.gov/34035530/)
- S. P. Weisberg, B. B. Ural, D. L. Farber, Tissue-specific immunity for a changing world. *Cell* **184**, 1517–1529 (2021). doi: [10.1016/j.cell.2021.01.042](https://doi.org/10.1016/j.cell.2021.01.042); pmid: [33740452](https://pubmed.ncbi.nlm.nih.gov/33740452/)
- M. M. L. Poon, D. L. Farber, The whole body as the system in systems immunology. *iScience* **23**, 101509 (2020). doi: [10.1016/j.isci.2020.101509](https://doi.org/10.1016/j.isci.2020.101509); pmid: [32920485](https://pubmed.ncbi.nlm.nih.gov/32920485/)
- P. A. Szabo et al., Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat. Commun.* **10**, 4706 (2019). doi: [10.1038/s41467-019-12464-3](https://doi.org/10.1038/s41467-019-12464-3); pmid: [31624246](https://pubmed.ncbi.nlm.nih.gov/31624246/)
- Tabula Muris Consortium, Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature* **562**, 367–372 (2018). doi: [10.1038/s41586-018-0590-4](https://doi.org/10.1038/s41586-018-0590-4); pmid: [30283141](https://pubmed.ncbi.nlm.nih.gov/30283141/)
- S. He et al., Single-cell transcriptome profiling of an adult human cell atlas of 15 major organs. *Genome Biol.* **21**, 294 (2020). doi: [10.1186/s13059-020-02210-0](https://doi.org/10.1186/s13059-020-02210-0); pmid: [33287869](https://pubmed.ncbi.nlm.nih.gov/33287869/)
- X. Han et al., Construction of a human cell landscape at single-cell level. *Nature* **581**, 303–309 (2020). doi: [10.1038/s41586-020-2157-4](https://doi.org/10.1038/s41586-020-2157-4); pmid: [32214235](https://pubmed.ncbi.nlm.nih.gov/32214235/)
- C. Xu, Prete, Teichlab/celltypist; celltypist 0.1.6, Zenodo (2021); <https://zenodo.org/record/5736560>.
- C. B. Brown et al., Transcriptional basis of mouse and human dendritic cell heterogeneity. *Cell* **179**, 846–863.e24 (2019). doi: [10.1016/j.cell.2019.09.035](https://doi.org/10.1016/j.cell.2019.09.035); pmid: [31668803](https://pubmed.ncbi.nlm.nih.gov/31668803/)
- A.-C. Villani et al., Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, eaah4573 (2017). doi: [10.1126/science.aah4573](https://doi.org/10.1126/science.aah4573); pmid: [28428369](https://pubmed.ncbi.nlm.nih.gov/28428369/)
- R. Elmentaita et al., Cells of the human intestinal tract mapped across space and time. *Nature* **597**, 250–255 (2021). doi: [10.1038/s41586-021-03852-1](https://doi.org/10.1038/s41586-021-03852-1); pmid: [34497389](https://pubmed.ncbi.nlm.nih.gov/34497389/)
- E. Madisoona et al., A spatial multi-omics atlas of the human lung reveals a novel immune cell survival niche. *bioRxiv* 2021.11.26.470108 [Preprint] (2021). <https://doi.org/10.1101/2021.11.26.470108>
- E. Evren et al., Distinct developmental pathways from blood monocytes generate human lung macrophage diversity. *Immunity* **54**, 259–275.e7 (2021). doi: [10.1016/j.immuni.2020.12.003](https://doi.org/10.1016/j.immuni.2020.12.003); pmid: [33382972](https://pubmed.ncbi.nlm.nih.gov/33382972/)
- A. Deczkowska, A. Weiner, I. Amit, The physiology, pathology, and potential therapeutic applications of the TREM2 signaling pathway. *Cell* **181**, 1207–1217 (2020). doi: [10.1016/j.cell.2020.05.003](https://doi.org/10.1016/j.cell.2020.05.003); pmid: [32531244](https://pubmed.ncbi.nlm.nih.gov/32531244/)
- A. Martens, G. van Loo, A20 at the crossroads of cell death, inflammation, and autoimmunity. *Cold Spring Harb. Perspect. Biol.* **12**, a036418 (2020). doi: [10.1101/cshperspect.a036418](https://doi.org/10.1101/cshperspect.a036418); pmid: [31427375](https://pubmed.ncbi.nlm.nih.gov/31427375/)
- M. Haldar et al., Heme-mediated SPI-C induction promotes monocyte differentiation into iron-recycling macrophages. *Cell* **156**, 1223–1234 (2014). doi: [10.1016/j.cell.2014.01.069](https://doi.org/10.1016/j.cell.2014.01.069); pmid: [24630724](https://pubmed.ncbi.nlm.nih.gov/24630724/)
- A. Bjukj et al., Transcriptional and functional profiling defines human small intestinal macrophage subsets. *J. Exp. Med.* **215**, 441–458 (2018). doi: [10.1084/jem.20170057](https://doi.org/10.1084/jem.20170057); pmid: [29273642](https://pubmed.ncbi.nlm.nih.gov/29273642/)
- O. Spadaro et al., IGF1 shapes macrophage activation in response to immunometabolic challenge. *Cell Rep.* **19**, 225–234 (2017). doi: [10.1016/j.celrep.2017.03.046](https://doi.org/10.1016/j.celrep.2017.03.046); pmid: [28402847](https://pubmed.ncbi.nlm.nih.gov/28402847/)
- F. Sallusto et al., Rapid and coordinated switch in chemokine receptor expression during dendritic cell maturation. *Eur. J. Immunol.* **28**, 2760–2769 (1998). doi: [10.1002/\(SICI\)1521-4141\(199809\)28:9<2760::AID-IMMU2760>3.0.CO;2-N](https://doi.org/10.1002/(SICI)1521-4141(199809)28:9<2760::AID-IMMU2760>3.0.CO;2-N); pmid: [9754536](https://pubmed.ncbi.nlm.nih.gov/9754536/)
- M. R. Clatworthy et al., Immune complexes stimulate CCR7-dependent dendritic cell migration to lymph nodes. *Nat. Med.* **20**, 1458–1463 (2014). doi: [10.1038/nm.3709](https://doi.org/10.1038/nm.3709); pmid: [25384086](https://pubmed.ncbi.nlm.nih.gov/25384086/)
- G. J. Randolph, V. Angel, M. A. Swartz, Dendritic-cell trafficking to lymph nodes through lymphatic vessels. *Nat. Rev. Immunol.* **5**, 617–628 (2005). doi: [10.1038/nri1670](https://doi.org/10.1038/nri1670); pmid: [16056255](https://pubmed.ncbi.nlm.nih.gov/16056255/)
- J. G. Cyster, Chemokines and the homing of dendritic cells to the T cell areas of lymphoid organs. *J. Exp. Med.* **189**, 447–450 (1999). doi: [10.1084/jem.189.3.447](https://doi.org/10.1084/jem.189.3.447); pmid: [9927098](https://pubmed.ncbi.nlm.nih.gov/9927098/)
- J. R. Ferguson et al., Maturing human CD127<sup>+</sup> CCR7<sup>+</sup> PDL1<sup>+</sup> dendritic cells express AIRE in the absence of tissue restricted antigens. *Front. Immunol.* **9**, 2902 (2018). doi: [10.3389/fimmu.2018.02902](https://doi.org/10.3389/fimmu.2018.02902); pmid: [30692988](https://pubmed.ncbi.nlm.nih.gov/30692988/)
- P. L. Poliani et al., Human peripheral lymphoid tissues contain autoimmune regulator-expressing dendritic cells. *Am. J. Pathol.* **176**, 1104–1112 (2010). doi: [10.2353/ajpath.2010.090956](https://doi.org/10.2353/ajpath.2010.090956); pmid: [20093495](https://pubmed.ncbi.nlm.nih.gov/20093495/)
- J. Wang et al., Single-cell multiomics defines tolerogenic extrathymic Aire-expressing populations with unique homology to thymic epithelium. *Sci. Immunol.* **6**, eab10503 (2021). doi: [10.1126/sciimmunol.ab10503](https://doi.org/10.1126/sciimmunol.ab10503); pmid: [34767455](https://pubmed.ncbi.nlm.nih.gov/34767455/)
- J. M. Gardner et al., Deletional tolerance mediated by extrathymic Aire-expressing cells. *Science* **321**, 843–847 (2008). doi: [10.1126/science.1159407](https://doi.org/10.1126/science.1159407); pmid: [18687966](https://pubmed.ncbi.nlm.nih.gov/18687966/)
- J. M. Gardner et al., Extrathymic Aire-expressing cells are a distinct bone marrow-derived population that induce functional inactivation of CD4<sup>+</sup> T cells. *Immunity* **39**, 560–572 (2013). doi: [10.1016/j.immuni.2013.08.005](https://doi.org/10.1016/j.immuni.2013.08.005); pmid: [23993652](https://pubmed.ncbi.nlm.nih.gov/23993652/)
- B. N. Lambrecht, H. Hammad, Biology of lung dendritic cells at the origin of asthma. *Immunity* **31**, 412–424 (2009). doi: [10.1016/j.immuni.2009.08.008](https://doi.org/10.1016/j.immuni.2009.08.008); pmid: [19766084](https://pubmed.ncbi.nlm.nih.gov/19766084/)
- F. A. Vieira Braga et al., A cellular census of human lungs identifies novel cell states in health and in asthma. *Nat. Med.* **25**, 1153–1163 (2019). doi: [10.1038/s41591-019-0468-5](https://doi.org/10.1038/s41591-019-0468-5); pmid: [31209336](https://pubmed.ncbi.nlm.nih.gov/31209336/)
- W. Sungnak et al., SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nat. Med.* **26**, 681–687 (2020). doi: [10.1038/s41591-020-0208-6](https://doi.org/10.1038/s41591-020-0208-6); pmid: [32327558](https://pubmed.ncbi.nlm.nih.gov/32327558/)
- M.-L. Golinski et al., CD11c<sup>+</sup> B cells are mainly memory cells, precursors of antibody secreting cells in healthy donors. *Front. Immunol.* **11**, 32 (2020). doi: [10.3389/fimmu.2020.00032](https://doi.org/10.3389/fimmu.2020.00032); pmid: [32158442](https://pubmed.ncbi.nlm.nih.gov/32158442/)
- E. Blanco et al., Age-associated distribution of normal B-cell and plasma cell subsets in peripheral blood. *J. Allergy Clin. Immunol.* **141**, 2208–2219.e16 (2018). doi: [10.1016/j.jaci.2018.02.017](https://doi.org/10.1016/j.jaci.2018.02.017); pmid: [29505809](https://pubmed.ncbi.nlm.nih.gov/29505809/)
- C. Wei et al., A new population of cells lacking expression of CD27 represents a notable component of the B cell memory compartment in systemic lupus erythematosus. *J. Immunol.* **178**, 6624–6633 (2007). doi: [10.4049/jimmunol.178.10.6624](https://doi.org/10.4049/jimmunol.178.10.6624); pmid: [17475894](https://pubmed.ncbi.nlm.nih.gov/17475894/)
- A. Kassambara et al., GenomicScape: An easy-to-use web tool for gene expression data analysis. Application to investigate the molecular events in the differentiation of B cells into plasma cells. *PLOS Comput. Biol.* **11**, e1004077 (2015). doi: [10.1371/journal.pcbi.1004077](https://doi.org/10.1371/journal.pcbi.1004077); pmid: [25633866](https://pubmed.ncbi.nlm.nih.gov/25633866/)
- M. P. Cancro, Age-associated B cells. *Annu. Rev. Immunol.* **38**, 315–340 (2020). doi: [10.1146/annurev-immunol-091310](https://doi.org/10.1146/annurev-immunol-091310); pmid: [31986068](https://pubmed.ncbi.nlm.nih.gov/31986068/)
- K. Kitaura et al., Different somatic hypermutation levels among antibody subclasses disclosed by a new next-generation sequencing-based antibody repertoire analysis. *Front. Immunol.* **8**, 389 (2017). doi: [10.3389/fimmu.2017.00389](https://doi.org/10.3389/fimmu.2017.00389); pmid: [28515723](https://pubmed.ncbi.nlm.nih.gov/28515723/)
- W. Meng et al., An atlas of B-cell clonal distribution in the human body. *Nat. Biotechnol.* **35**, 879–884 (2017). doi: [10.1038/nbt.3942](https://doi.org/10.1038/nbt.3942); pmid: [28829438](https://pubmed.ncbi.nlm.nih.gov/28829438/)
- J. J. C. Thorne et al., Spatial map of human T cell compartmentalization and maintenance over decades of life. *Cell* **159**, 814–828 (2014). doi: [10.1016/j.cell.2014.10.026](https://doi.org/10.1016/j.cell.2014.10.026); pmid: [25417158](https://pubmed.ncbi.nlm.nih.gov/25417158/)
- M. Buggert et al., The identity of human tissue-emigrant CD8<sup>+</sup> T cells. *Cell* **183**, 1946–1961.e15 (2020). doi: [10.1016/j.cell.2020.11.019](https://doi.org/10.1016/j.cell.2020.11.019); pmid: [3306960](https://pubmed.ncbi.nlm.nih.gov/3306960)
- B. V. Kumar et al., Human Tissue-Resident Memory T Cells Are Defined by Core Transcriptional and Functional Signatures in Lymphoid and Mucosal Sites. *Cell Rep.* **20**, 2921–2934 (2017). doi: [10.1016/j.celrep.2017.08.078](https://doi.org/10.1016/j.celrep.2017.08.078); pmid: [28930685](https://pubmed.ncbi.nlm.nih.gov/28930685/)
- N. J. Schuldt, B. A. Binstadt, Dual TCR T cells: Identity crisis or multitaskers? *J. Immunol.* **202**, 637–644 (2019). doi: [10.4049/jimmunol.1800904](https://doi.org/10.4049/jimmunol.1800904); pmid: [30670579](https://pubmed.ncbi.nlm.nih.gov/30670579/)
- M. P. Soares, I. Hamza, Macrophages and iron metabolism. *Immunity* **44**, 492–504 (2016). doi: [10.1016/j.immuni.2016.02.016](https://doi.org/10.1016/j.immuni.2016.02.016); pmid: [26982356](https://pubmed.ncbi.nlm.nih.gov/26982356/)

59. T. Ito *et al.*, TSLP-activated dendritic cells induce an inflammatory T helper type 2 cell response through OX40 ligand. *J. Exp. Med.* **202**, 1213–1223 (2005). doi: [10.1084/jem.20051135](https://doi.org/10.1084/jem.20051135); pmid: [16275760](https://pubmed.ncbi.nlm.nih.gov/16275760/)
60. V. Soumelis *et al.*, Human epithelial cells trigger dendritic cell-mediated allergic inflammation by producing TSLP. *Nat. Immunol.* **3**, 673–680 (2002). doi: [10.1038/ni805](https://doi.org/10.1038/ni805); pmid: [12055625](https://pubmed.ncbi.nlm.nih.gov/12055625/)
61. H. J. Sutton *et al.*, Atypical B cells are part of an alternative lineage of B cells that participates in responses to vaccination and infection in humans. *Cell Rep.* **34**, 108684 (2021). doi: [10.1016/j.celrep.2020.108684](https://doi.org/10.1016/j.celrep.2020.108684); pmid: [33567273](https://pubmed.ncbi.nlm.nih.gov/33567273/)
62. D. Nehar-Belaid *et al.*, Mapping systemic lupus erythematosus heterogeneity at the single-cell level. *Nat. Immunol.* **21**, 1094–1106 (2020). doi: [10.1038/s41590-020-0743-0](https://doi.org/10.1038/s41590-020-0743-0); pmid: [32747814](https://pubmed.ncbi.nlm.nih.gov/32747814/)
63. A. Bonner, A. Almogren, P. B. Furtado, M. A. Kerr, S. J. Perkins, The nonplanar secretory IgA2 and near planar secretory IgA1 solution structures rationalize their different mucosal immune responses. *J. Biol. Chem.* **284**, 5077–5087 (2009). doi: [10.1074/jbc.M807529200](https://doi.org/10.1074/jbc.M807529200); pmid: [19109255](https://pubmed.ncbi.nlm.nih.gov/19109255/)
64. M. Kilian, J. Reinholdt, H. Lomholt, K. Poulsen, E. V. Frandsen, Biological significance of IgA1 proteases in bacterial colonization and pathogenesis: Critical evaluation of experimental evidence. *APMIS* **104**, 321–338 (1996). doi: [10.1111/j.1699-0463.1996.tb00724.x](https://doi.org/10.1111/j.1699-0463.1996.tb00724.x); pmid: [8703438](https://pubmed.ncbi.nlm.nih.gov/8703438/)
65. Y. Zhao *et al.*, Clonal expansion and activation of tissue-resident memory-like T<sub>H</sub>17 cells expressing GM-CSF in the lungs of patients with severe COVID-19. *Sci. Immunol.* **6**, eabf6692 (2021). doi: [10.1126/sciimmunol.abf6692](https://doi.org/10.1126/sciimmunol.abf6692); pmid: [33622974](https://pubmed.ncbi.nlm.nih.gov/33622974/)
66. E. J. Kunkel *et al.*, Lymphocyte CC chemokine receptor 9 and epithelial thymus-expressed chemokine (TECK) expression distinguish the small intestinal immune compartment: Epithelial expression of tissue-specific chemokines as an organizing principle in regional immunity. *J. Exp. Med.* **192**, 761–768 (2000). doi: [10.1084/jem.192.5.761](https://doi.org/10.1084/jem.192.5.761); pmid: [10974041](https://pubmed.ncbi.nlm.nih.gov/10974041/)
67. N. J. Bernstein *et al.*, Solo: Doublet identification in single-cell RNA-seq via semi-supervised deep learning. *Cell Syst.* **11**, 95–101.e5 (2020). doi: [10.1016/j.cels.2020.05.010](https://doi.org/10.1016/j.cels.2020.05.010); pmid: [32592658](https://pubmed.ncbi.nlm.nih.gov/32592658/)
68. S. L. Wolock, R. Lopez, A. M. Klein, Scrublet: Computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst.* **8**, 281–291.e9 (2019). doi: [10.1016/j.cels.2018.11.005](https://doi.org/10.1016/j.cels.2018.11.005); pmid: [30954476](https://pubmed.ncbi.nlm.nih.gov/30954476/)
69. F. A. Wolf, P. Angerer, F. J. Theis, SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018). doi: [10.1186/s13059-017-1382-0](https://doi.org/10.1186/s13059-017-1382-0); pmid: [29409532](https://pubmed.ncbi.nlm.nih.gov/29409532/)
70. K. Polański *et al.*, BBKNN: Fast batch alignment of single cell transcriptomes. *Bioinformatics* **36**, 964–965 (2020). doi: [10.1093/bioinformatics/btaa019](https://doi.org/10.1093/bioinformatics/btaa019); pmid: [31400197](https://pubmed.ncbi.nlm.nih.gov/31400197/)
71. R. Lopez, J. Regier, M. B. Cole, M. I. Jordan, N. Yosef, Deep generative modeling for single-cell transcriptomics. *Nat. Methods* **15**, 1053–1058 (2018). doi: [10.1038/s41592-018-0229-2](https://doi.org/10.1038/s41592-018-0229-2); pmid: [30504886](https://pubmed.ncbi.nlm.nih.gov/30504886/)
72. M. Büttner, Z. Miao, F. A. Wolf, S. A. Teichmann, F. J. Theis, A test metric for assessing single-cell RNA-seq batch correction. *Nat. Methods* **16**, 43–49 (2019). doi: [10.1038/s41592-018-0254-1](https://doi.org/10.1038/s41592-018-0254-1); pmid: [30573817](https://pubmed.ncbi.nlm.nih.gov/30573817/)
73. G. Sturm *et al.*, Scirpy: A Scanpy extension for analyzing single-cell T-cell receptor-sequencing data. *Bioinformatics* **36**, 4817–4818 (2020). doi: [10.1093/bioinformatics/btaa611](https://doi.org/10.1093/bioinformatics/btaa611); pmid: [32614448](https://pubmed.ncbi.nlm.nih.gov/32614448/)
74. C. Domínguez Conde, Teichlab/TissueImmuneCellAtlas, Zenodo (2022); <https://zenodo.org/record/6334988>.

#### ACKNOWLEDGMENTS

We thank the deceased organ donors, donor families, the extended Cambridge Biorepository for Translational Medicine (CBTM) team, and the transplant coordinators at LiveOnNY for access to the tissue samples. We thank the CellTypist Annotation Team [composed of S. Webb, L. Jardine (M. Haniffa laboratory), B. Stewart, Z.K.T. (M.R.C. laboratory), R. Hoo (R. Vento-Tormo laboratory), K. James, J.P., E. Madisoona, W. Sungnak, R.E., and P. He (S.A.T. laboratory) for contributions to the harmonization of CellTypist labels. We thank V. Svensson for the initial idea of CellTypist. CellTypist is publicly available, with user-friendly documentation for application to any single-cell transcriptomics datasets (<https://github.com/Teichlab/celltypist#interactive-tutorials>). We thank A. Maartens for providing feedback on the text. We acknowledge C. Usher, J. Eliasova, and BioRender for graphical images. We thank M. Anderson, J. Gardner, C. Benoit, and M. Haniffa for discussions on data interpretation, and we thank C. Talavera-López and N. Kumashika for discussions on data analysis. We also acknowledge the support received from the Cellular Generation and Phenotyping (CGaP) core facility, the Cellular Genetics Informatics team, and Core DNA Pipelines at the Wellcome Sanger Institute and the Cambridge NIHR BRC Cell Phenotyping Hub (Department of Medicine, University of Cambridge). This publication is part of the Human Cell Atlas - [www.humancellatlas.org/publications](http://www.humancellatlas.org/publications).

**Funding:** This work was funded, in part, by the Wellcome Trust (grants 206194 and 108413/A/15/D to S.A.T., grant WSSS 211276/D/18/Z to the CBTM, and grant 105924/Z/14/Z; RG79413 to J.L.J.), Chan Zuckerberg Initiative Seed Network (CZIF2019-002452) (to D.L.F., P.A.Si., J.L.J., and N.Y.), and NIH grants HL145547, AI128949, and AI06697 (to D.L.F. and P.A.Si.). This work was also supported by funding from the European Research Council (grant 646794 ThDEFINE to S.A.T.) and by the NIHR Cambridge Biomedical Research Centre (BRC-1215-20014). The CBTM is also funded by the Chan Zuckerberg Initiative. For

the purpose of open access, the authors have applied a CC BY public copyright license to any author accepted manuscript version arising from this submission. The views expressed here are those of the author(s) and not necessarily those of the NIHR, Department of Health and Social Care, or National Institutes of Health (NIH). T.G. was supported by the European Union's H2020 research and innovation program "ENLIGHT-TEN" under Marie Skłodowska-Curie grant 675395. **Author contributions:** N.Y., P.A.Si., D.L.F., K.S.-P., J.L.J., and S.A.T. designed the study. K.T.M. and K.S.-P. obtained tissues from deceased organ donors. D.K.M. and E.J.N. provided premortem blood samples. C.D.C., L.B.J., D.B.R., S.B.W., S.K.H., O.S., L.M., P.A.Sz., L.R., L.B., E.S.F., H.S.M., J.P., and D.C. performed tissue dissociation and prepared single-cell sequencing libraries. C.D.C., C.X., K.P., H.W.K., Z.K.T., R.E., and E.R. analyzed and interpreted the data. C.X. and T.G. developed CellTypist. C.X., T.G., and N.H. collected CellTypist training datasets. M.P. provided support with the CellTypist package and website. L.B.J., D.B.R., S.K.H., L.T., S.P., and N.R. performed validation experiments. L.C. provided help with histological analysis. T.L. provided help with smFISH image analysis. L.K.J., O.A.B., K.B.M., M.R.C., N.Y., P.A.Si., D.L.F., J.L.J., and S.A.T. supervised experiments and data analysis. C.D.C., C.X., L.B.J., S.A.T., and J.L.J. wrote the manuscript with contributions from D.L.F., M.R.C., D.B.R., and T.G. All authors read, provided input on, and approved the manuscript. **Competing interests:** In the past 3 years, S.A.T. has worked as a consultant for Genentech, Roche, and Transition Bio and is a remunerated member of the scientific advisory boards of Qiagen, GlaxoSmithKline, and Foresite Labs and an equity holder of Transition Bio. J.L.J. reports receiving consultancy fees and grant support from Sanofi Genzyme. All other authors have no competing interests. **Data and materials availability:** Raw single-cell sequencing data have been deposited in the ArrayExpress database at EMBL-EBI ([www.ebi.ac.uk/arrayexpress/](https://www.ebi.ac.uk/arrayexpress/)) under accession number E-MTAB-11536. Processed data can be downloaded and interactively explored at <https://www.tissueimmuneatlases.org>. Code used for data analysis can be found on Zenodo (74) and <https://github.com/Teichlab/TissueImmuneCellAtlas>.

#### SUPPLEMENTARY MATERIALS

[science.org/doi/10.1126/science.abl5197](https://science.org/doi/10.1126/science.abl5197)

Materials and Methods

Supplementary Text

Figs. S1 to S26

Tables S1 to S3

References (75–94)

MDAR Reproducibility Checklist

[View/request a protocol for this paper from Bio-protocol](https://science.org/protocol).

Submitted 19 July 2021; accepted 22 March 2022  
10.1126/science.abl5197