RESEARCH ARTICLE SUMMARY

IMMUNOLOGY

Mapping the developing human immune system across organs

Chenqu Suo†, Emma Dann†, Issac Goh, Laura Jardine, Vitalii Kleshchevnikov, Jong-Eun Park, Rachel A. Botting, Emily Stephenson, Justin Engelbert, Zewen Kelvin Tuong, Krzysztof Polanski, Nadav Yayon, Chuan Xu, Ondrej Suchanek, Rasa Elmentaite, Cecilia Domínguez Conde, Peng He, Sophie Pritchard, Mohi Miah, Corina Moldovan, Alexander S. Steemers, Pavel Mazin, Martin Prete, Dave Horsfall, John C. Marioni, Menna R. Clatworthy*, Muzlifah Haniffa*, Sarah A. Teichmann*

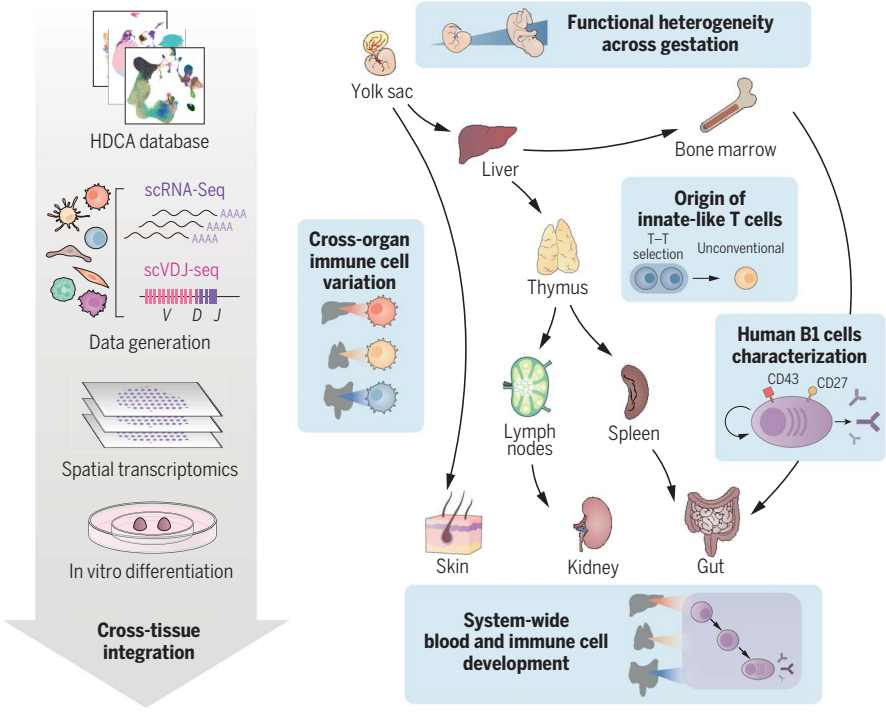
INTRODUCTION: The human immune system develops across several anatomical sites throughout gestation. Immune cells differentiate initially from extra-embryonic yolk sac progenitors and subsequently from aorto-gonad-mesonephros–derived hematopoietic stem cells before liver and bone marrow take over as the primary sites of hematopoiesis. Immune cells from these primary hematopoietic sites then seed developing lymphoid organs and peripheral nonlymphoid organs. Recent advances in single-cell genomics technologies have facilitated studies on the developing immune system at unprecedented scale and resolution. However, these studies have focused on one or a few organs rather than reconstructing the entire immune system as a distributed network across tissues.

RATIONALE: To provide a detailed characterization of the developing immune system across multiple organs, we performed single-cell RNA sequencing (scRNA-seq) using dissociated cells from yolk sac, prenatal spleen, and skin, and integrated publicly available cell atlases of six additional organs, spanning weeks 4 to 17 after conception. To further characterize developmental B and T cells and explore their antigen receptor repertoire, we also generated paired γδ T cell receptor (γδTCR)–, αβ T cell receptor (αβTCR)–, and B cell receptor (BCR)–sequencing data. Finally, to study the spatial localizations of cell populations in early hematopoietic tissue and lymphoid organs critical for B and T cell development, we performed spatial transcriptomics on prenatal spleen, liver, and thymus and used the scRNA-seq data as a reference to map the cells in situ.

RESULTS: We have integrated a cross-tissue single-cell atlas of developing human immune cells across prenatal hematopoietic, lymphoid, and nonlymphoid peripheral organs. This includes over 900,000 cells from which we identified over 100 cell states.

Using cross-gestation analysis, we revealed the acquisition of immune-effector functions of myeloid and lymphoid cell types from the second trimester, and their early transcriptomic signatures suggested a role in tissue morphogenesis. Through cross-organ analysis, we identified conserved processes of proliferation and maturation for monocytes and T cells before their migration from the bone marrow and thymus, respectively, into peripheral tissues. We discovered system-wide blood and immune cell development, in particular B lymphopoiesis, across all sampled peripheral organs. This expands on previous understanding of conventional hematopoietic organs (yolk sac, liver, and bone marrow) as the only sites for immune cell development. We validated the presence and location of lineage-committed progenitors spatially using 10X Genomics Visium Spatial Gene Expression and single-molecule fluorescence in situ hybridization. Finally, we identified and functionally validated the properties of human prenatal innate-like B and T cells, providing an extensive characterization of human B1 cells with single-cell transcriptomic and BCR information, as well as functional validation of spontaneous antibody secretion. Integrating the transcriptome profiles of human prenatal unconventional T cells, their αβTCR V(D)J usage, and data from an in vitro thymic organoid culture model, we supply additional evidence for thymocyte–thymocyte selection during unconventional T cell development.

CONCLUSION: Our comprehensive single-cell and spatial atlas of the developing human immune system provides valuable resources and biological insights to facilitate in vitro cell engineering and regenerative medicine and to enhance our understanding of congenital disorders affecting the immune system. ■



Cross-tissue mapping of the developing human immune system. We reconstructed the process of immune cell development, analyzing cells across prenatal hematopoietic, lymphoid, and peripheral organs, combining scRNA-seq, scVDJ-seq, and spatial transcriptomics. With this integrated dataset, we studied variation in cellular phenotypes across development and between tissues and the distribution of blood and immune cell progenitors across tissues and characterized fetal-specific innate-like B and T cells.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: mrc38@cam.ac.uk (M.C.), m.a.haniffa@newcastle.ac.uk (M.H.); st9@sanger.ac.uk (S.A.T.)
†These authors contributed equally to this work.
Cite this article as C. Suo *et al.*, *Science* **376**, eabo0510 (2022). DOI: 10.1126/science.abo0510

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abo0510

RESEARCH ARTICLE

IMMUNOLOGY

Mapping the developing human immune system across organs

Chenqu Suo[1,2]†, Emma Dann[1]†, Issac Goh[3], Laura Jardine[3,4], Vitalii Kleshchevnikov[1], Jong-Eun Park[1,5], Rachel A. Botting[3], Emily Stephenson[3], Justin Engelbert[3], Zewen Kelvin Tuong[1,6], Krzysztof Polanski[1], Nadav Yayon[1,7], Chuan Xu[1], Ondrej Suchanek[6], Rasa Elmentaite[1], Cecilia Domínguez Conde[1], Peng He[1,7], Sophie Pritchard[1], Mohi Miah[3], Corina Moldovan[8], Alexander S. Steemers[1], Pavel Mazin[1], Martin Prete[1], Dave Horsfall[3], John C. Marioni[1,7,9], Menna R. Clatworthy[1,6]*, Muzlifah Haniffa[1,3,10]*, Sarah A. Teichmann[1,11]*

Single-cell genomics studies have decoded the immune cell composition of several human prenatal organs but were limited in describing the developing immune system as a distributed network across tissues. We profiled nine prenatal tissues combining single-cell RNA sequencing, antigen-receptor sequencing, and spatial transcriptomics to reconstruct the developing human immune system. This revealed the late acquisition of immune-effector functions by myeloid and lymphoid cell subsets and the maturation of monocytes and T cells before peripheral tissue seeding. Moreover, we uncovered system-wide blood and immune cell development beyond primary hematopoietic organs, characterized human prenatal B1 cells, and shed light on the origin of unconventional T cells. Our atlas provides both valuable data resources and biological insights that will facilitate cell engineering, regenerative medicine, and disease understanding.

The human immune system develops across several anatomical sites throughout gestation. Immune cells differentiate initially from extra-embryonic yolk sac progenitors, and subsequently from aorto-gonad-mesonephros–derived hematopoietic stem cells (HSCs), before the liver and bone marrow take over as the primary sites of hematopoiesis (*1*, *2*). Immune cells from these primary hematopoietic sites seed developing lymphoid organs such as the thymus, spleen, and lymph nodes, as well as peripheral non-lymphoid organs.

Recent advances in single-cell genomics technologies have revolutionized our understanding of the developing human organs (*3*–*11*). However, these studies have focused on one or a few organs rather than reconstructing the entire immune system as a distributed network across all organs. Such a holistic understanding of the developing human immune system would have far-reaching implications for health and disease, including cellular engineering, regenerative medicine, and a deeper mechanistic understanding of congenital disorders affecting the immune system.

Here, we present a cross-tissue single-cell and spatial atlas of developing human immune cells across prenatal hematopoietic organs (yolk sac, liver, and bone marrow), lymphoid organs (thymus, spleen, and lymph nodes), and nonlymphoid peripheral organs (skin, kidney, and gut) to provide a detailed characterization of generic and tissue-specific properties of the developing immune system. We generated single-cell RNA-sequencing (scRNA-seq) data from yolk sac, prenatal spleen, and skin and integrated publicly available cell atlases of six additional organs spanning weeks 4 to 17 after conception (*3*, *4*, *7*, *8*, *10*, *11*). We also generated single-cell γδ T cell receptor (γδTCR)–sequencing data and additional, αβTCR–, and B cell receptor (BCR)–sequencing data. Finally, we integrated the single-cell transcriptome profiles with in situ tissue location using spatial transcriptomics.

This study reveals the acquisition of immune-effector functions of myeloid and lymphoid lineages from the second trimester, the maturation of developing monocytes and T cells before peripheral tissue seeding, and system-wide blood and immune cell development during human prenatal development. Moreover, we identified, characterized, and functionally validated the properties of human prenatal B1 cells and the origin of unconventional T cells.

[1]Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK. [2]Department of Paediatrics, Cambridge University Hospitals, Cambridge, UK. [3]Biosciences Institute, Newcastle University, Newcastle upon Tyne, UK. [4]Haematology Department, Freeman Hospital, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [5]Graduate School of Medical Science and Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea. [6]Molecular Immunity Unit, University of Cambridge Department of Medicine, Cambridge, UK. [7]European Molecular Biology Laboratory European Bioinformatics Institute, Hinxton, Cambridge, UK. [8]Department of Cellular Pathology, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [9]Cancer Research UK Cambridge Institute, Li Ka Shing Centre, University of Cambridge, Cambridge, UK. [10]Department of Dermatology and National Institute for Health Research (NIHR) Newcastle Biomedical Research Centre, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [11]Theory of Condensed Matter, Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge, UK.
*Corresponding author. Email: mrc38@cam.ac.uk (M.C.), m.a.haniffa@newcastle.ac.uk (M.H.); st9@sanger.ac.uk (S.A.T.)
†These authors contributed equally to this work.

Integrated cross-organ map of prenatal cell states in distinct tissue microenvironments

To systematically assess the heterogeneity of immune cell populations across human prenatal hematopoietic organs, lymphoid, and nonlymphoid tissues, we generated scRNA-seq data from prenatal spleen, yolk sac, and skin, which were integrated with a collection of publicly available single-cell datasets from the Human Developmental Cell Atlas initiative (*3*, *4*, *7*, *8*, *10*, *11*). In total, our dataset comprised samples from 25 embryos or fetuses between 4 and 17 postconception weeks (pcw) (Fig. 1A) profiled in 221 scRNA-seq libraries. For 65 of these libraries, paired antigen-receptor–sequencing data were available for αβTCR, γδTCR, or BCR (Fig. 1B). After mapping and preprocessing with a unified pipeline, a total of 908,178 cells were retained after quality control.

To facilitate joint analysis of the data, we integrated all libraries using single-cell variational inference (scVI) (*12*), minimizing protocol- and embryo-associated variation (fig. S1A) while retaining differences between organs. In keeping with previous single-cell atlases of immune cells of prenatal and adult tissues (*3*, *11*, *13*), our data captured the emergence of myeloid and lymphoid lineages, as well as closely linked megakaryocytes and erythroid and non-neutrophilic granulocyte lineages from hematopoietic progenitors (Fig. 1C and figs. S1B to S3). Linking transcriptional phenotypes to paired antigen receptor sequence expression, we paired αβTCR sequences for 28,739 cells, paired γδTCR sequences for 813 cells, and paired BCR sequences for 14,506 cells (fig. S1C).

We repeated dimensionality reduction and clustering on subsets of cells from different lineages and used marker gene analysis and comparison with existing cell labels to comprehensively annotate cell types across tissues. In total, we defined 127 high-quality cell populations (figs. S4 and S5). Cross-tissue integration enabled the identification of cell populations that were too rare to be resolved by the analysis of datasets from individual tissues, such as *AXL*- and *SIGLEC6*-expressing dendritic cells (DCs) (*14*) and plasma B cells (fig. S4). To facilitate the rapid reuse of our atlas for the analysis of newly collected samples, we made the weights from trained scVI models available to enable mapping of external scRNA-seq datasets using transfer learning with single-cell architectural surgery (scArches) (*15*).

To study the spatial localizations of the cell populations in an early hematopoietic tissue and lymphoid organs critical for B and T cell development, we profiled developing liver,

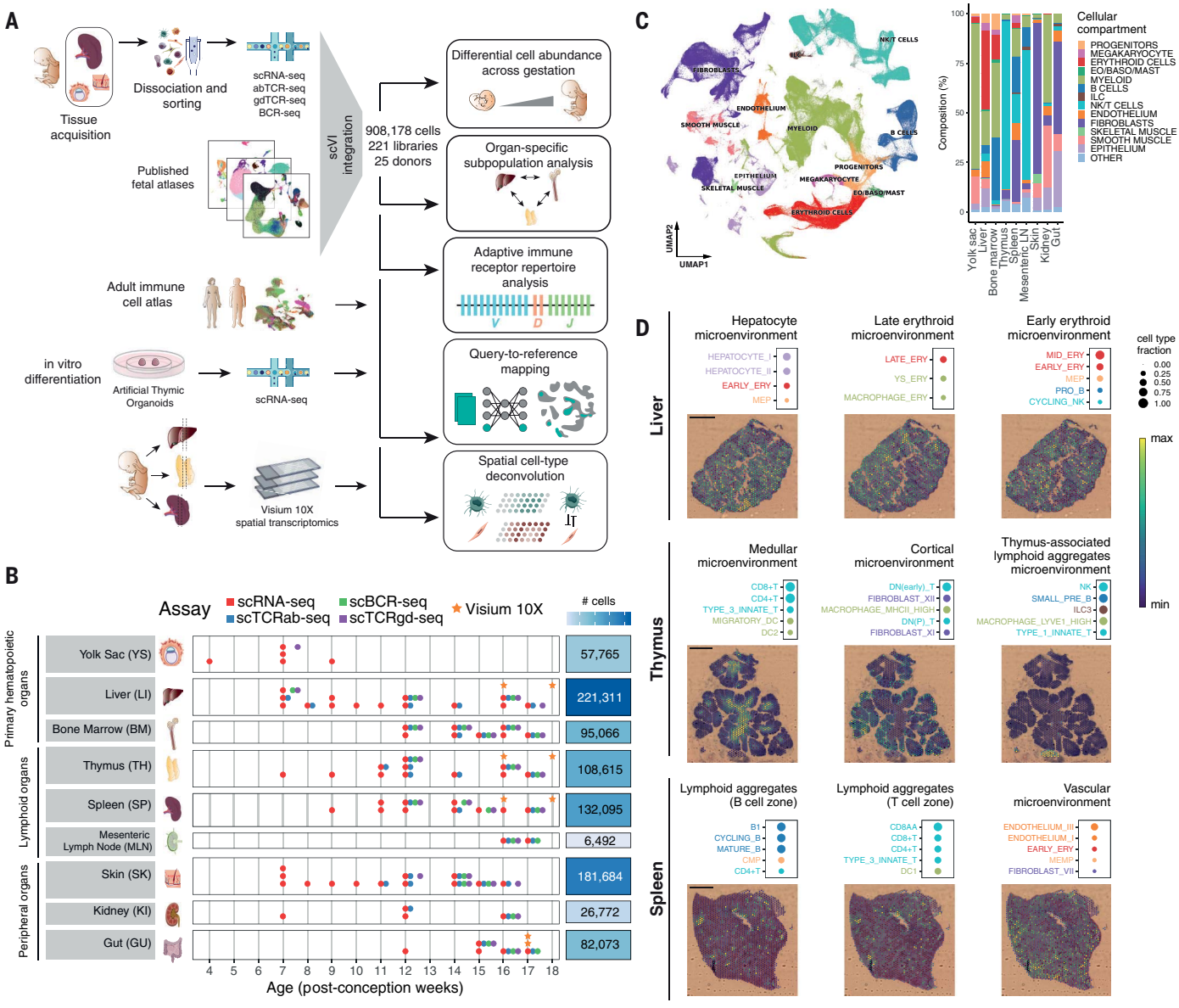


Fig. 1. Cross-tissue cellular atlas of the developing human immune system. (**A**) Overview of study design and analysis pipeline. scRNA-seq and scVDJ-seq data were generated from prenatal spleen, yolk sac, and skin, which were integrated using scVI with a collection of publicly available scRNA-seq datasets. This cell atlas was used for (i) differential abundance analysis across gestation and organs with Milo, (ii) antigen receptor repertoire analysis with scirpy and dandelion, (iii) comparison with adult immune cells and in vitro differentiated cells with scArches and CellTypist, and (iv) spatial cell-type deconvolution on 10X Genomics Visium data of hematopoietic and lymphoid organs using cell2location. (**B**) Summary of analyzed samples by gestational stage (*x*-axis) and organ (*y*-axis). Colors denote the types of molecular assays performed for each sample. The side bar indicates the total number of cells collected for each organ (after quality control). (**C**) Left: UMAP embedding of scRNA-seq profiles in prenatal tissues (908,178 cells) colored by broad cellular compartments. Right: bar plot of percentage of cells assigned to each broad compartment for each of the profiled organs. Raw cell proportions are adjusted to account for FACS-based CD45 enrichment. The category "other" denotes clusters annotated as low-quality cells. Eo/Baso/Mast, eosinophils/basophils/mast cells. (**D**) Representative colocalization patterns identified with NMF of spatial cell-type abundances estimated with cell2location. For each annotated microenvironment, a dot plot of relative contribution of cell types to microenvironment (top; dot size) and spatial locations of microenvironments on tissue slides (bottom) are shown, with the color representing the weighted contribution of each microenvironment to each spot. Scale bars, 1 mm.

thymus, and spleen from two donors at 16 and 18 pcw with spatial transcriptomics (10X Genomics Visium Spatial Gene Expression). Using our multiorgan scRNA-seq dataset as a reference, we performed spatial cell-type deconvolution with cell2location (*16*) to map cells in tissue (fig. S6). We used nonnegative matrix factorization (NMF) of the cell-type abundance estimates in tissue spots to identify microenvironments of colocalized cell types in the profiled tissues in an unbiased manner (Fig. 1D and figs. S7 to S10).

In the developing liver, we recovered expected signatures of tissue-specific parenchymal cells such as hepatocytes. In addition, we observed spatial segregation of early and late erythrocytes, suggesting distinctive hematopoietic zones (Fig. 1D and fig. S8). In the developing thymus, we recovered the localization of cell types in known histological structures. Developing T cells, for example, were largely

localized to the thymic cortex, whereas mature T cells were consistently mapped to the thymic medulla. Furthermore, in two of the thymic tissue sections, we observed aggregates of lymphoid tissue (hereafter referred to as thymus-associated lymphoid aggregates). Within these, we mapped B cell subsets, innate lymphoid cells (ILCs), and macrophage subtypes (Fig. 1D and fig. S9). In the developing spleen, most of the tissue was highly vascularized. In addition, within splenic lymphoid aggregates, we were able to distinguish partially overlapping B and T cell zones (Fig. 1D and fig. S10).

Heterogeneity of prenatal myeloid cells across organs and gestation

We first examined the main compartments of immune cells in our multiorgan dataset to identify gestation-specific and organ-specific variability within cell populations.

The myeloid compartment captured the development from committed myeloid progenitors to neutrophils, monocytes, macrophages, and DCs (fig. S4, G and H). Our cross-tissue analysis distinguished three distinct subsets of monocytes, which were characterized by a differential distribution between prenatal bone marrow and peripheral tissues and by the expression of *CXCR4*, *CCR2*, or *IL1B* (*17*). Among macrophages, we identified eight broad macrophage groupings on the basis of their transcriptome profile (fig. S4H): "LYVE1^{hi}", "iron-recycling", "MHC class II^{hi}", "Kupffer-like" (*18*), "microglia-like TREM2^{hi}" (*19*), "osteoclasts" (*11*, *20*), and "proliferating" macrophages. Assigning proliferating cells to the other identified subsets, we observed a high fraction of proliferating macrophages in the yolk sac and within the LYVE1^{hi} subset across organs, suggesting an increased self-renewal potential (fig. S11).

We compared prenatal and adult immune cell populations by mapping a cross-tissue adult dataset of immune cells (*21*) onto our prenatal myeloid reference (fig. S12, A and B). We found that the transcriptional profiles of DC subsets were conserved between adult and prenatal counterparts (fig. S12C). Adult monocytes were most similar to the IL1B^{hi} and CCR2^{hi} prenatal populations, and no CXCR4^{hi} monocytes in nonlymphoid adult tissues were observed (fig. S13). Most adult macrophages clustered separately from the prenatal macrophages, with the exception of erythrophagocytic macrophages (fig. S12, B and C). This population includes macrophages primarily from the spleen and liver that perform iron-recycling functions (*21*).

To quantify changes in cellular composition across gestation, we performed differential abundance analysis on cell neighborhoods using Milo (Fig. 2A and fig. S14A) (*22*). This analysis reaffirmed well-known compositional shifts that happen during gestation. For example, myeloid progenitor cells decreased in the liver but increased in the bone marrow, recapitulating the transition from liver to bone marrow hematopoiesis. DCs increased in proportional abundance across multiple tissues, as previously described in the liver and bone marrow (*3*). For several cell populations, we found that some neighborhoods were enriched and others depleted across gestation, suggesting evolving transcriptional heterogeneity during development. This was especially evident in the macrophage compartment in the skin and liver (Fig. 2A), with a large fraction of neighborhoods overlapping the LYVE1^{hi} and proliferating macrophages enriched in early gestation. Differential expression analysis revealed the up-regulation of a proinflammatory gene signature with chemokines and cytokines specific to early stages in all macrophage subtypes across tissues (Fig. 2B and fig. S14B). Tumor necrosis factor (TNF) and nuclear factor κB (NF-κB) have been implicated in lymphoid tissue organogenesis (*23*), and the chemokines noted here have been associated with angiogenesis (*24*–*27*). Conversely, a large fraction of neighborhoods within the iron-recycling and MHCII^{hi} macrophage populations were enriched in later stages of gestation. We found that these subpopulations up-regulated genes encoding for immune-effector functions (Fig. 2B, fig. S14C, and table S1). In parallel to macrophages, we observed similar transcriptional heterogeneity during gestation in mast cells (Fig. 2A). Specifically, early mast cells in yolk sac, liver, and skin displayed a similar proinflammatory phenotype characterized by expression of *TNF* and *NF-κB* subunits, as well as chemokines associated with endothelial cell recruitment (*CXCL3*, *CXCL2*, and *CXCL8*) (*26*) (fig. S15).

These findings suggest that early macrophages and mast cells may contribute to angiogenesis, tissue morphogenesis, and homeostasis, as previously reported (*28*–*30*), before adopting traditional immunological functions. The acquisition of macrophage antigen-presentation properties (e.g., MHCII up-regulation) between 10 and 12 pcw coincided with the expansion of adaptive lymphocytes (fig. S1E) and the development of lymphatic vessels and lymph nodes (*31*).

Differential abundance analysis on cell neighborhoods to test for organ-specific enrichment (fig. S16A) revealed that CXCR4^{hi} monocytes were enriched in bone marrow and IL1B^{hi} monocytes were enriched in peripheral organs. Among CCR2^{hi} monocytes, we distinguished bone marrow– and peripheral organ–specific subpopulations (Fig. 2C). Bone marrow CCR2^{hi} monocytes expressed proliferation genes, whereas peripheral organ CCR2^{hi} monocytes up-regulated *IL1B* and other TNF-α–signaling genes (Fig. 2D, fig. S16B, and table S2). This suggests that a CXCR4^{hi} to CCR2^{hi} transition accompanies monocyte egress from the bone marrow to seed peripheral tissues, and CCR2^{hi} monocytes further mature in tissues into IL1B^{hi} monocytes (Fig. 2, D and E). In mouse bone marrow, interactions between monocyte CXCR4 and stromal cell CXCL12 retain monocytes in situ until CCR2-CCL2 interactions predominate, potentially enabling monocyte egress (*17*). Here, we observed *CXCL12* expression in bone marrow fibroblasts and osteoblasts (fig. S16C). By contrast, the proportion of CXCR4^{hi} monocytes in the developing liver was much lower (fig. S16D), in keeping with reports that alternative mechanisms of monocyte retention and release operate in the murine developing liver (*32*).

Heterogeneity of prenatal lymphoid cells across organs and gestation

The lymphoid compartment captures the development of B and T cells, together with ILC and natural killer (NK) cell subsets (fig. S4, I to L).

Mapping adult cells onto our prenatal lymphoid reference, NK cells and type 3 ILCs (ILC3) displayed high similarity between adult and prenatal counterparts (fig. S17, A and B). Among adult T cells, naive populations and regulatory T cells (T_{regs}) closely matched prenatal conventional T cells ($CD4^+$ T, $CD8^+$ T, and T_{regs}), whereas resident and effector memory T cells did not have a developmental equivalent (fig. S17, C and D), although we cannot exclude the possibility that memory T cells appear after 17 pcw, as previously reported (*33*, *34*). We did not find a clear matching between adult T cell subsets and prenatal unconventional T cells (type 1 and type 3 innate T cells and CD8AA T cells in our annotation). All adult B cell progenitors, naive B cells, and memory B cells had prenatal counterparts, but no adult B cells were transcriptionally similar to prenatal putative B1 cells (fig. S17C).

Differential abundance analysis across gestation identified a broad shift from innate to adaptive immune populations (Fig. 3A and fig. S18A). ILCs and NK cells included cell neighborhoods that were both enriched and depleted across gestation. Genes involved in inflammatory responses, including TNF signaling, were overexpressed in <12 pcw liver and skin NK cells, although late splenic NK cells also expressed these genes. Conversely, late NK cells across organs overexpressed genes involved in cytokine signaling and granzyme genes (Fig. 3B; fig. S18, B and C; and table S3). As is the case for macrophages, these results suggest the progressive development of immune-effector function by NK cells.

We next tested for organ-specific cell neighborhoods in the lymphoid compartment (fig. S19A). Although certain populations of mature T cells were exclusively enriched in the thymus [ABT(entry), CD8AA], we found that neighborhoods of conventional and unconventional

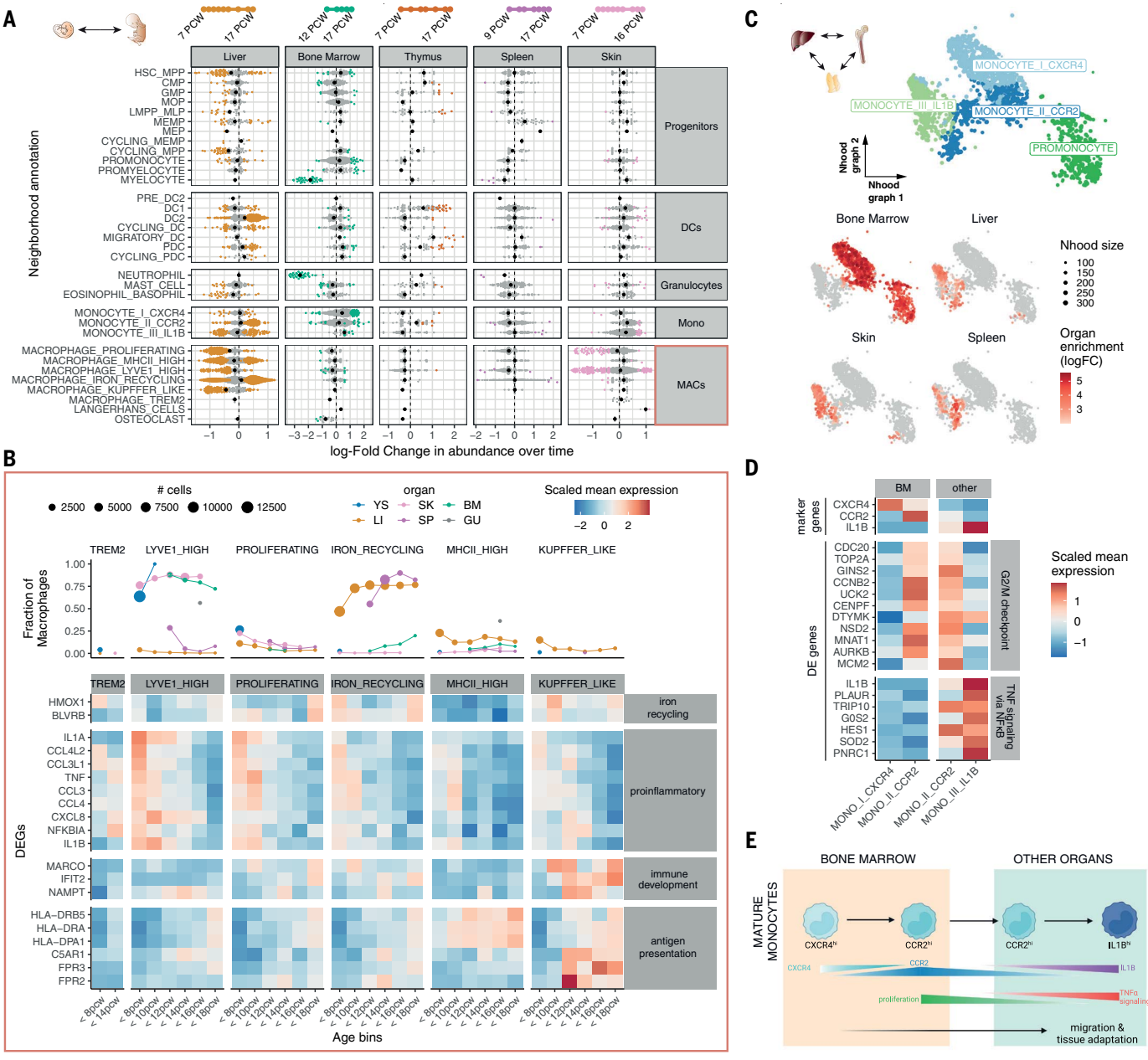


Fig. 2. Myeloid variation across time and tissues. (**A**) Bee-swarm plot of log-fold change (*x*-axis) in cell abundance across gestational stages in Milo neighborhoods of myeloid cells. Results from five organs are shown. Neighborhoods overlapping the same cell population are grouped together (*y*-axis) and colored if displaying significant differential abundance (DA) (spatial FDR < 10%). The black dot denotes the median log-fold change. The top bar denotes the range of gestational stages of the organ samples analyzed. (**B**) Heatmap of average expression across time of a selection of markers of stage-specific macrophage neighborhoods. Mean log-normalized expression for each gene is scaled (*z*-score). Gestational ages are grouped in six age bins. Age bins in which <30 cells of a given subset were present are not shown. The top panel shows the fraction of all macrophages belonging to the specified macrophage population in each time point and each organ (color). (**C**) Close-up view of monocytes on Milo neighborhood embedding of myeloid cells (subset from fig. S16). Top: neighborhoods are colored by overlapping cell population. Bottom: neighborhoods displaying significant DA (spatialFDR < 10%) are colored by log-fold change in abundance between the specified organ and all other organs. (**D**) Mean expression of a selection of differentially expressed genes between CCR2^{hi} monocytes from bone marrow (BM) and other organs. Log-normalized expression for each gene is scaled (*z*-score). Genes up-regulated in bone marrow associated with G_2/M checkpoint and genes down-regulated in bone marrow associated with TNF signaling are shown (from MSigDB Hallmark 2020 gene sets). (**E**) Schematic of the proposed process of monocyte egression from the bone marrow mediated by CXCR4 and CCR2 expression: CXCR4^{hi} monocytes are retained in the bone marrow until they switch to a proliferative state with increased expression of CCR2, mediating tissue egression. CCR2^{hi} monocytes seed peripheral tissues and then mature further to the periphery-specific IL1B-expressing subtype.

T cells could be subdivided into a subset enriched in the thymus and other subsets enriched in peripheral organs (Fig. 3C). Thymic mature T cells overexpressed genes involved in interferon-α (IFN-α) signaling, whereas peripheral mature T cells had higher expression of genes associated with TNF and NF-κB signaling (Fig. 3D, fig. S19B, and table S4). Both pathways have been implicated in the last stages of functional maturation of murine T cells right before emigration out of the thymus (*35*, *36*). In addition to the increase in type I IFN and

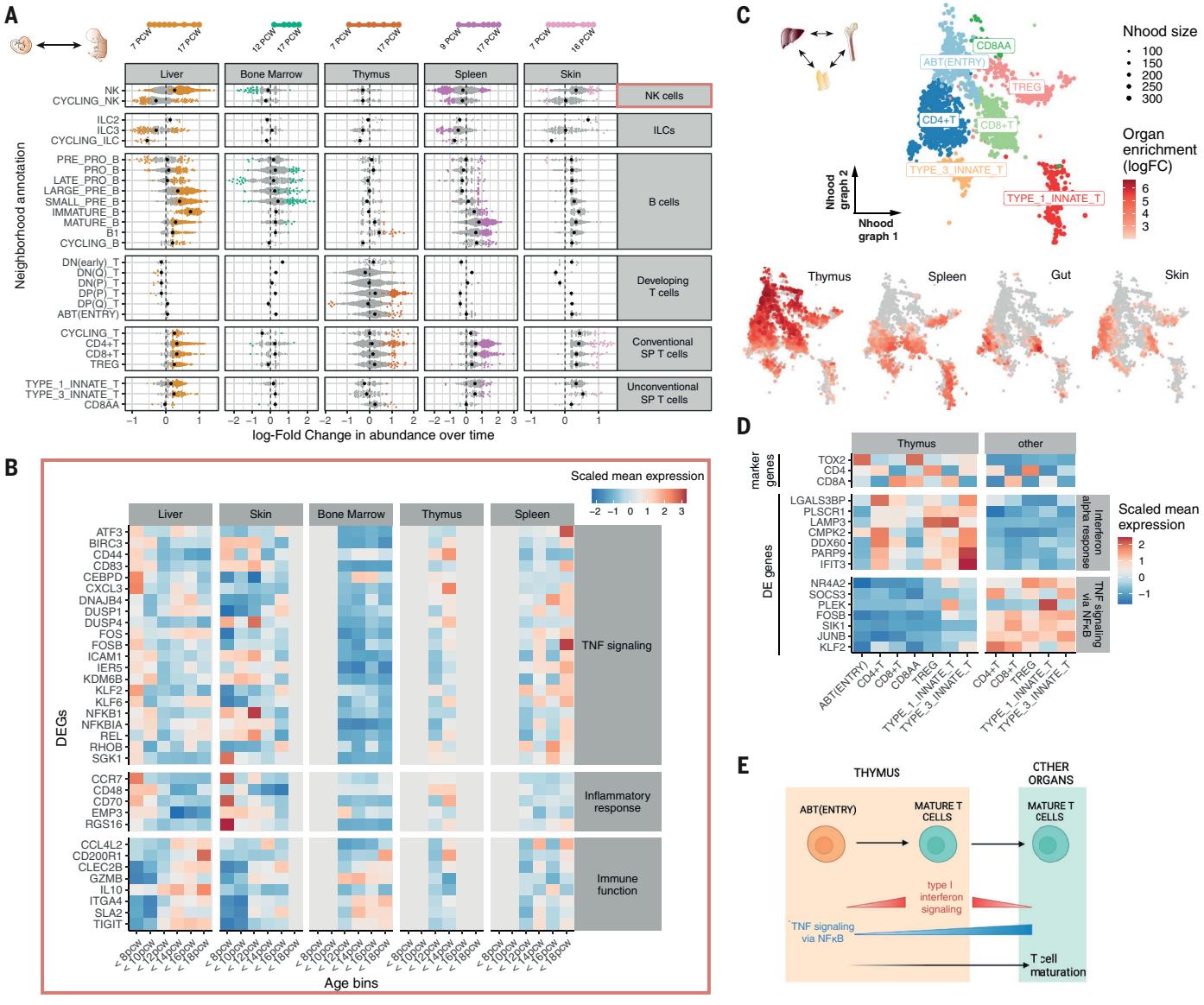


Fig. 3. Lymphoid variation across time and tissues. (**A**) Bee-swarm plot of log-fold change (*x*-axis) in cell abundance across gestational stages in Milo neighborhoods of lymphoid cells (as in Fig. 2A). (**B**) Heatmap showing average expression across time of a selection of genes identified as markers of early-specific and late-specific NK neighborhoods (as in Fig. 2B): NK cells identified in liver and skin before 12 pcw express TNF proinflammatory genes, whereas the expression of immune-effector genes such as cytokines, chemokines, and granzyme genes increases after 12 pcw. Age bins in which <30 NK cells were present in a given organ are grayed out. (**C**) Close-up view of single-positive T cells on Milo neighborhood embedding of lymphoid cells. Each point represents a neighborhood, and the layout of points is determined by the position of the neighborhood index cell in the UMAP in fig. S4I. Top: neighborhoods are colored by the cell population they overlap. Bottom: neighborhoods are colored by their log-fold change in abundance between the specified organ and all other organs. Only neighborhoods displaying significant differential abundance (spatialFDR < 10%) are colored. (**D**) Mean expression of a selection of differentially expressed genes between single-positive T cells from thymus (TH) and other organs. Genes down-regulated in the thymus associated with TNF signaling (using MSigDB Hallmark 2020 gene sets) and genes up-regulated in the thymus associated with an IFN-α response are shown. (**E**) Schematic of the proposed mechanism of thymocyte maturation and egression from thymus mediated by type I IFN and NF-κB signaling.

NF-κB signaling accompanying ABT(entry) to thymic mature T cells, expression of NF-κB-signaling genes continued to increase when mature T cells migrate out to peripheral tissues (Fig. 3E).

System-wide blood and immune cell development

While examining the distribution of various cell types across different organ systems, we were surprised to find that lineage-committed hematopoietic progenitors were present in nonhematopoietic organs. In particular, we detected B cell progenitors in almost all prenatal organs, megakaryocyte/erythroid progenitors in developing spleen and skin, and myeloid progenitors in the thymus, spleen, skin, and kidney (Fig. 4A). By contrast, T cell progenitors were restricted to the thymus, potentially reflecting more stringent niche requirements for T cell development and consistent with the observed absence of T cells in children with congenital athymia (*37*). This finding suggests that hematopoiesis is not restricted to developing liver and bone marrow between 7 and 17 pcw (*38*) and that other organs can also support blood and immune cell differentiation during prenatal development.

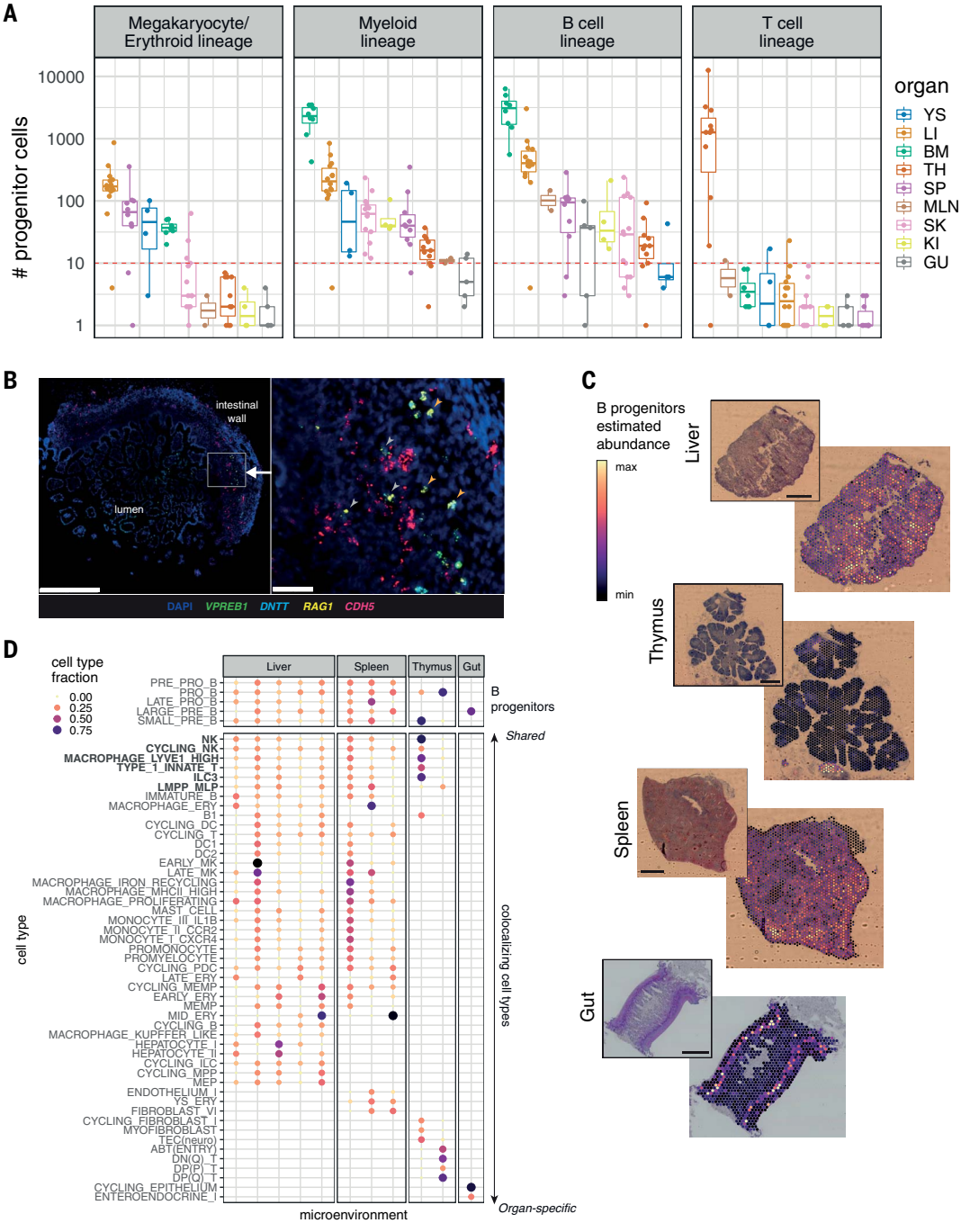


Fig. 4. System-wide blood and immune cell development. (**A**) Boxplots of the number of progenitor cells in all donors across organs. Each point represents a donor, color coded by organ. YS, yolk sac; LI, liver; BM, bone marrow; TH, thymus; SP, spleen; MLN, mesenteric lymph node; SK, skin; GU, gut; KI, kidney. The red dashed line marks the threshold of 10 cells for potential technical artifacts. Detailed cell types included in each lineage are shown in table S5. Boxes capture the first-to-third quartile of the cell number, and whisks span a further 1.5× interquartile range on each side of the box. (**B**) Multiplex smFISH staining with DAPI, *CDH5* for endothelial cells, and *VPREB1*, *DNTT*, and *RAG1* for B cell progenitors in the human prenatal intestine at 15 pcw. Left panel shows a zoomed-out view with the area of interest boxed in white. Scale bar, 500 μm. Right panel shows a detailed view of the area of interest. Scale bar, 50 μm. Gray arrows point to B cell progenitors associated with blood vessels, and orange arrows point to B cell progenitors away from blood vessels. (**C**) Scaled sum of abundances of B progenitor cell types estimated with cell2location, shown on representative slides for each organ, with the corresponding H&E staining. Scale bars, 1 mm. (**D**) Cell-type contributions to microenvironments containing B cell progenitors in different organs identified with nonnegative matrix factorization of spatial cell-type abundances estimated with cell2location. The color and the size of the dots represent the relative fraction of cells of a type assigned to the microenvironment.

In addition, across progenitor lineages, cells of different developmental stages were simultaneously present in peripheral organs (fig. S20, A to D). Cell-fate prediction analysis delineated a continuum of cells between HSCs and different lineages of immune cells in multiple organs (fig. S20, C and D), supporting the conclusion that lineage-committed differentiation takes place within peripheral organs.

Single-molecule fluorescence in situ hybridization (smFISH) staining confirmed the existence of lineage-committed progenitors in multiple organs. Cells simultaneously expressing *VPREB1* and *RAG1*, with or without *DNTT* were present in the prenatal gut, spleen, and thymus (Fig. 4B and fig. S21, A to C), consistent with B cell progenitors. Although some B cell progenitors in the prenatal gut were associated with *CDH5*-expressing blood vessels, many could be detected extravascularly (Fig. 4B), further supporting the conclusion that B cells develop in prenatal peripheral organs. We also validated the presence of megakaryocyte/erythroid lineage progenitors in the prenatal spleen and thymus (fig. S22, A to C) and of myeloid lineage progenitors in the prenatal gut and thymus (fig. S23, A to C).

Focusing on B lymphopoiesis given its widespread nature, we used cell2location (*16*) on 10X Genomics Visium spatial transcriptomic data and found that B cell progenitors were localized in the submucosa of the gut, in thymus-associated lymphoid aggregates, and proximal

to lymphoid aggregates in the spleen (fig. S24B), in addition to their expected presence in the developing liver (Fig. 4C and fig. S24A). The widespread nature of B lymphopoiesis suggests that the cellular environments supporting B cell development are much more widely available than previously thought. Spatial transcriptomic data identified cells colocalizing with B cell progenitors across multiple organs, including ILC3, LYVE1^{hi} macrophages, NK cells, type 1 innate T cells, and LMPP_MLP cells (see fig. S24C for predicted cell-cell interactions), whereas other colocalizing cell types were organ specific (Fig. 4D).

Identification of putative prenatal B1 cells

Among prenatal nonprogenitor B cells that had productive BCR light chains and low *IL7R* expression (fig. S25A), we identified immature B, mature B, cycling B, plasma B, and putative B1 cells (Fig. 5A and fig. S25B). These putative B1 cells had the highest expression of *CD5*, *CD27*, and *SPN* (CD43), consistent with previously reported markers (*39*–*41*), as well as *CCR10*, a highly specific marker that was expressed in a subset of B1 cells (Fig. 5A).

We next evaluated characteristics typical of murine B1 cells, including self-renewal (*42*, *43*), high immunoglobulin M (IgM) and low IgD expression (*44*), emergence in early development (*45*), low levels of nontemplated nucleotide BCR insertions (*46*, *47*), tonic BCR signaling (*39*), and spontaneous antibody secretion (*42*).

With regard to B1 cell self-renewal, we calculated the percentage of cycling cells (as indicated by nonzero *MKI67* expression) within immature B, mature B, B1, and plasma B cells, respectively (Fig. 5B and fig. S26A). The proportion of cycling B1 cells was significantly higher than cycling mature B cells, consistent with their capacity for self-renewal. B1 cells expressed lower levels of *IGHD* and higher levels of *IGHM* compared with mature B cells (Fig. 5B). Moreover, the highest frequency of B1 cells was found in the early embryonic stages. These were gradually replaced by other subsets of nonprogenitor B cells over time. The ratio of B1 to mature B cells showed a general decrease from the first to second trimester across most organs except the thymus (fig. S26B), where B1 cells persisted, consistent with a previous report of a shared phenotype between thymic B cells and B1 cells (*48*).

Analysis of nontemplated nucleotide insertions in the BCR showed that both N/P additions and CDR3 junctions in heavy and light chains were shorter in B1 cells compared with mature B cells (Fig. 5C). Moreover, a lower mutation frequency was observed in light chains of B1 cells compared with those in mature B cells, and the average mutation frequency was lower than that observed in adult B cells (*21*, *49*). We next examined the V(D)J usage within different B cell subtypes along the developmental path (fig. S26C). Prenatal B1 and mature B cells both exhibited a varied BCR repertoire with minimal clonal expansion (fig. S26D) and had differing preferential usage of V(D)J segments (Fig. 5D).

Our putative B1 cells showed features of tonic BCR signaling, with higher B cell activation scores (fig. S26E), as well as higher transcription factor (TF) activity in the TNF-α– and NF-κB–signaling pathway (fig. S26F), which is downstream of BCR signaling (*50*), compared with mature B cells.

We assessed spontaneous antibody secretion capacity in B1 cells by flow-sorting B cell subsets (fig. S26G) and assessing spontaneous IgM secretion using the enzyme-linked immune absorbent spot (ELISpot) assay. The normalized antibody-secreting spot counts were higher in the two B1 fractions than in the two mature B fractions, with the CCR10^{hi} B1 fraction showing the highest spot counts (Fig. 5E). scRNA-seq of the sorted B cell fractions on a different sample using the same gating strategy further confirmed that the two sorted B1 fractions were indeed B1 cell enriched compared with the mature B fractions (fig. S26H). We also explored the potential role of *CCR10* in prenatal B1 cells and observed the expression of one of its ligands, *CCL28*, in bone marrow stroma (chondrocytes and osteoblasts), in gut epithelium, and in keratinocytes and melanocytes in the skin (fig. S26I). Thus, *CCR10* may play a role in the tissue localization of prenatal B1 cells.

Overall, our scRNA-seq, paired V(D)J–sequencing data, and functional assay provide an extended characterization of human prenatal B1 cells (Fig. 5F).

Human unconventional T cells are trained by thymocyte-thymocyte selection

The mature T cell compartment consisted of conventional T cells (CD4$^+$ T cells, CD8$^+$ T cells, and T_{regs}) and unconventional T cells. The origin of the latter in humans is poorly understood. Unconventional T cells expressed the key innate marker *ZBTB16* (PLZF) (fig. S27A) (*51*) and could be further separated into three different subtypes: *RORC*- and *CCR6*-expressing type 3 innate T cells; *EOMES*- and *TBX21*-expressing type 1 innate T cells; and *PDCD1*-expressing and thymus-restricted CD8AA cells (figs. S2 and S4L), corresponding, respectively, to T-helper 17 (T_H17)–like cells, NK T cell (NKT)–like cells, and CD8αα$^+$ T cells (*7*).

The proportions of unconventional T cells among all mature T cells exhibited a decreasing trend from 7 to 9 pcw to 10 to 12 pcw across most of the organs surveyed here (fig. S27B). Type 1 and type 3 innate T cells were almost negligible in postnatal thymus, whereas CD8AA T cell abundance rebounded in pediatric age groups before a further decline in adulthood (fig. S27B). Thus, type 1 and type 3 innate T cells, but not CD8AA T cells, appear to be developmental-specific, unconventional T cells.

Spatially, we found that mature T cells segregated into two microenvironments in the thymic medulla (fig. S27C). Conventional CD4$^+$ T and CD8$^+$ T cells colocalized with medullary thymic epithelial cells (mTECs) close to the inner medulla, whereas CD8AA and type 1 innate T cells colocalized with type 1 DCs (DC1s) near the corticomedullary junction (fig. S27, D and E). T_{regs} and type 3 innate T cells were located within both microenvironments. Thus, in contrast to conventional T cells, CD8AA and type I innate T cells likely undergo distinct negative selection processes mediated by DCs rather than mTECs and may also be involved in DC activation, as previously suggested (*7*).

Single-cell sequencing of γδTCR and αβTCR was performed on a subset of samples to characterize antigen-receptor repertoires in unconventional T cells (Fig. 1B). By far, most unconventional T cells expressed paired αβTCR, but some of these cells expressed paired γδTCR (Fig. 6A). Most γδT cells expressed *TRGV9* and *TRDV2* (Fig. 6B), consistent with previous reports (*52*, *53*). However, there was also a large proportion of γδT cells, particularly those of the CD8AA and type 3 innate T cell subtypes, that expressed *TRGV8* or *TRGV10* instead (Fig. 6B). Thus, the γδTCR showed a relatively restricted repertoire and substantial clonal expansion (fig. S28B).

Prenatal unconventional T cells expressed a varied αβTCR repertoire (Fig. 6C and fig. S28C) with minimal clonal expansion (fig. S28D), unlike well-described unconventional T cell subsets [e.g., type I NKT and mucosal-associated invariant T (MAIT) cells] in adults (*54*). V-J gene usage in TCRα was previously observed to have a strong association with T cell developmental timing (*7*, *55*). Specifically, double-positive (DP) T cells tend to use proximal TRAV, TRAJ gene segments, whereas mature T cells tend to use more distal pairs, governed by the processive depletion of proximal segments in V-J gene recombination (*55*). The V-J gene usage of αβTCR-expressing unconventional T cells lies between that of DP cells and conventional T cells, as shown by the more proximal gene usage in unconventional T cells (Fig. 6C) and principal component analysis of the TCR repertoire (Fig. 6D). This suggests that unconventional T cells are developmentally closer to DP cells (Fig. 6E) and undergo fewer recombinations before positive selection.

Previous studies have suggested that these PLZF-expressing unconventional T cells may originate from positive selection on neighboring T cells (*51*, *56*–*58*), in contrast to conventional T cells arising from positive selection on cortical TECs (cTECs). After β-selection, DP T cells undergo proliferation before recombination of TCRα (*7*, *59*, *60*). Each DP cell is thus

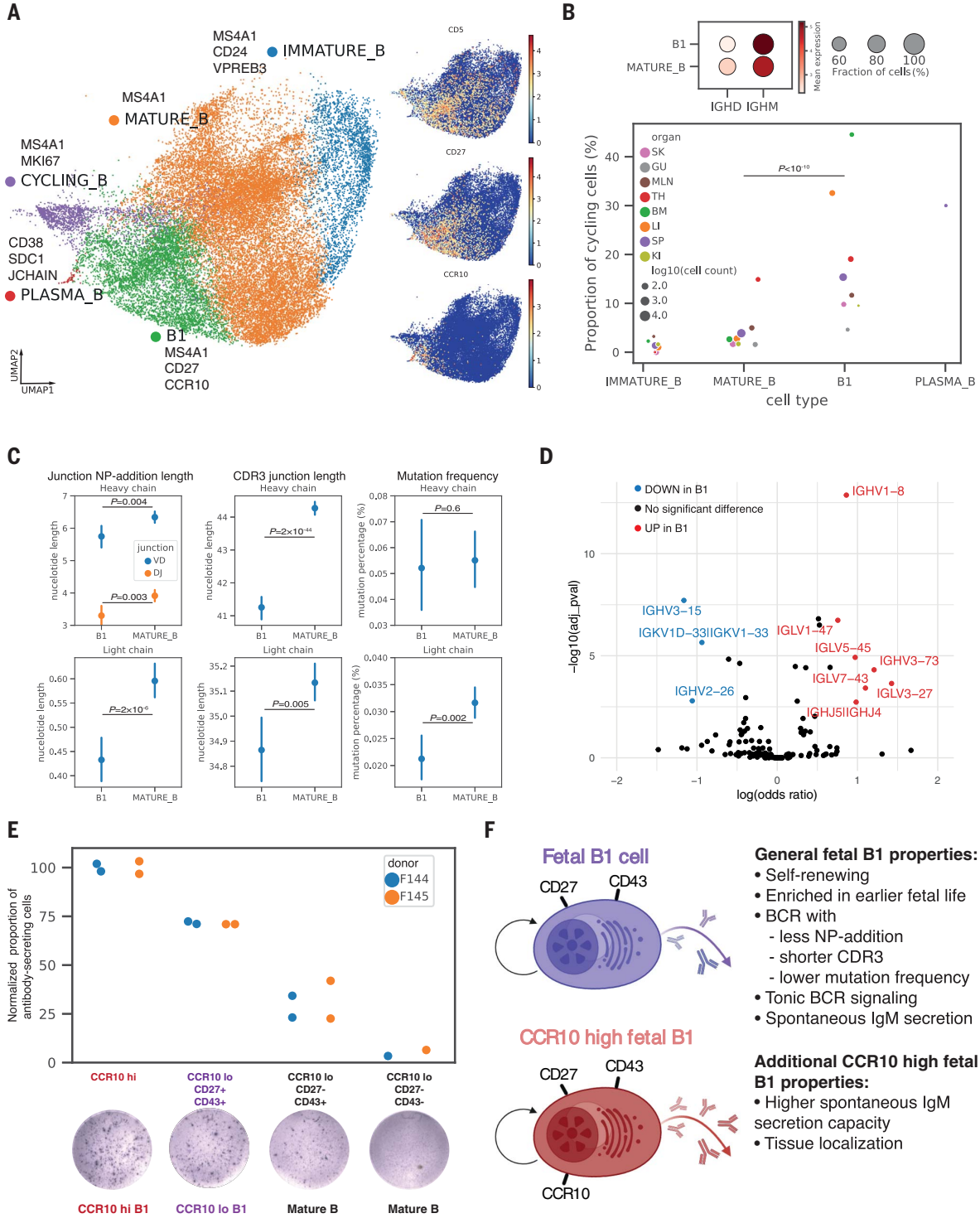


Fig. 5. Identification of putative prenatal B1 cells. (**A**) Left: Close-up view of nonprogenitor B cell populations on UMAP embedding of all lymphoid cells (fig. S4I), with marker genes listed next to each cell type. Right: expression of B1 marker genes on UMAP. (**B**) Top: dot plot of *IGHM* and *IGHD* expressions in B1 and mature B cells, with the color of dots representing the mean expression and size representing the fraction of cells expressing the gene. Bottom: cycling cell proportions within each B cell subtype colored by organs, with dot size representing $\log_{10}$(cell count) and only dots with at least 10 cells shown. B1 cells had significantly higher cycling proportions than mature B cells in a logistic regression controlling for donors and organs. (**C**) Point plots of NP-addition length, CDR3 junction length, and mutation frequency in BCR heavy chains or light chains in B1 cells (n = 2357) and mature B cells (n = 7387), with points representing the mean and lines representing 95% confidence intervals. Heavy-chain VD and DJ junction NP-addition lengths are only calculated for cells with high-quality D gene mapping (B1 cells: n = 615; mature B cells: n = 2430). Difference in characteristics were tested with linear regressions controlling for donors and organs. (**D**) Volcano plot summarizing results of BCR heavy- and light-chain V,J gene segment usage comparison between B1 and mature B cells. The y-axis is the $-\log_{10}$(Benjamini-Hochberg–adjusted P value), and the x-axis is log(odds ratio) computed using logistic

regression controlling for donors and organs. (**E**) Normalized proportions of antibody-secreting cells in different sorted fractions of the ELISpot experiments (raw counts in table S6), colored by donor. Each point represents a reaction well. The proportions of antibody-secreting cells were normalized against the average proportion in CCR10^{hi} wells for each donor to remove donor-specific effects. A representative well image for each sorted fraction is shown on the bottom. (**F**) Schematic illustration summarizing the features of all human prenatal B1 cells and additional features specific to CCR10^{hi} prenatal B1 cells.

surrounded by several neighboring DP cells from the same clone. It is therefore plausible that it requires less physical migration and thus is quicker for a DP T cell to receive positive signaling from a neighboring DP T cell rather than having to migrate to meet a nearby cTEC. Thus, the fact that unconventional T cells have a more similar TCR usage to DP cells agrees with the thymocyte-thymocyte (T-T) origin hypothesis.

To test our hypothesis for T-T–mediated selection of unconventional T cells, we differentiated induced pluripotent stem cells (iPSCs) into mature T cells using the artificial thymic organoid (ATO) (*61*). There were no human TECs present in the ATO system (Fig. 6F). scRNA-seq analysis of differentiated cells harvested at weeks 3, 5, and 7 from two iPSC lines confirmed that the in vitro culture system recapitulated T cell development from double-negative (DN) and DP, to ABT(ENTRY), and then to single-positive mature T cells (SP_T) (Fig. 6F and fig. S29, A to C). SP_T cells differentiated in vitro were dominated by *ZBTB16*-expressing unconventional T cells (Fig. 6G). Both label transfer (Fig. 6G) and similarity scoring on merged embeddings (fig. S29D) showed that the in vitro SP_T were most similar to in vivo type 1 innate T cells. Thus, our in vitro experiments support the T-T origin hypothesis of unconventional T cells.

Discussion

Our study provides a comprehensive single-cell dataset of the developing human immune system, spanning >900,000 single-cell profiles from nine tissues and encompassing >100 cell states. Compared with previous multiorgan developmental atlases (*9*), we increased coverage of developmental organs, gestation stages, and sequencing depth and generated paired BCR, αβTCR, and γδTCR datasets. Moreover, we demonstrate the utility of scRNA-seq reference to delineate tissue organization and cellular communication in spatial transcriptomics, providing a proof-of-concept study of the localizations of immune cells across prenatal tissues. Our preprocessed data and pretrained models (scVI and CellTypist models) will facilitate the alignment of new data to our dataset and streamline future expansion and analysis of human developmental atlases.

Our cross-organ analysis revealed several important biological phenomena. First, human macrophages, mast cells, and NK cells transcriptomically acquire immune-effector functions between 10 and 12 pcw. Their transcriptomic signatures before this time point suggest a role in tissue morphogenesis, consistent with previous findings for murine macrophages (*62*), and may explain why these cells appear in early development. The coincidental development of the lymphatic system around 12 pcw (*31*) raises the possibility of its potential role in initiating this transcriptional switch. Second, there are conserved processes of proliferation and maturation for monocytes and T cells before their migration from the bone marrow and thymus, respectively, into peripheral tissues. Third, in contrast to the previous dogma of hematopoiesis being restricted to the yolk sac, liver, and bone marrow during human development, system-wide blood and immune cell development takes place in peripheral organs, although at varying extents in different lineages. It is possible that hematopoiesis is supported to varying levels in prenatal organs, including the adrenal gland (*9*), before the onset of functional organ maturation, as exemplified by the fetal liver, which transitions from a hematopoietic to a metabolic organ. The potential for other peripheral organs to support hematopoiesis is evidenced by the reemergence of extramedullary hematopoiesis in adults, primarily in pathological settings (*63*–*66*), as well as the recent description of B lymphopoiesis in murine and nonhuman primate meninges (*67*–*69*).

Finally, this work identifies and functionally validates the properties of human prenatal innate-like B and T cells and provides an extensive characterization of human B1 cells. Our in vivo αβTCR V(D)J usage patterns and in vitro T cell differentiation data proposes T-T selection underpinning unconventional T cell development. Further studies are required to establish whether B1 cells arise from different progenitors (lineage model) (*70*–*72*) or from the same progenitors but with different signaling (selection model) (*73*, *74*), similar to the conventional and unconventional T cell model proposed here. Both innate-like B and T cells were abundant during early development, and their precise role at this developmental time point warrants further investigation. Their reported debris-removal (*41*, *42*), antigen-reactivity (*41*, *42*, *54*), and regulatory functions (*42*) may confer these prenatal innate-like B and T cells with tissue-homeostatic and important immunological roles.

In summary, this comprehensive atlas of the developing human immune system provides valuable resources and biological insights to facilitate in vitro cell engineering and regenerative medicine and to enhance our understanding of congenital disorders affecting the immune system.

Materials and Methods

A more detailed version of the materials and methods is provided in the supplementary materials.

Tissue acquisition and processing

Human developmental tissue samples (4 to 17 pcw; see metadata in table S7) used for this study were obtained from the MRC-Wellcome Trust–funded Human Developmental Biology Resource (HDBR; http://www.hdbr.org) with written consent and approval from the Newcastle and North Tyneside NHS Health Authority Joint Ethics Committee (08/H0906/21+5). All tissues were digested into single-cell suspensions with 1.6 mg/ml type IV collagenase (Worthington).

scRNA-seq experiment

Dissociated cells were stained with anti-CD45 antibody and 4′,6-diamidino-2-phenylindole (DAPI) before sorting. For scRNA-seq experiments, either the Chromium Single Cell 3′ Reagent Kit or the Chromium Single Cell V(D)J Reagent Kit from 10X Genomics was used. Unsorted, DAPI$^-$CD45$^+$, or DAPI$^-$CD45$^-$ fluorescence-activated cell sorting (FACS)–isolated cells were loaded onto each channel of the Chromium chip. Single-cell cDNA synthesis, amplification, gene expression, and targeted BCR and TCR libraries were generated. Targeted enrichment for γδTCR was performed following the TCR enrichment protocol from 10X Genomics with customized primers (table S8) (*75*). Sequencing was performed on the Illumina Novaseq 6000 system. The gene expression libraries were sequenced at a target depth of 50,000 reads per cell using the following parameters: read 1: 26 cycles, i7: eight cycles, i5: zero cycles; read 2: 91 cycles to generate 75-bp paired-end reads. BCR and TCR libraries were sequenced at a target depth of 5000 reads per cell.

ATO cell cultures

The PSC-ATO protocol was followed as previously described (*61*) (for more details, see the supplementary materials). Two iPSC lines, HPSI0114i-kolf_2 (Kolf) and HPSI0514i-fiaj_1 (Fiaj), obtained from the Human Induced Pluripotent Stem Cell Initiative (HipSci; www.hipsci.org) collection, were used.

Visium

Optimal cutting temperature (OCT) medium–embedded freshly frozen samples (table S9) were used for 10X Genomics Visium. All tissues were sectioned with a thickness of 15 μm.

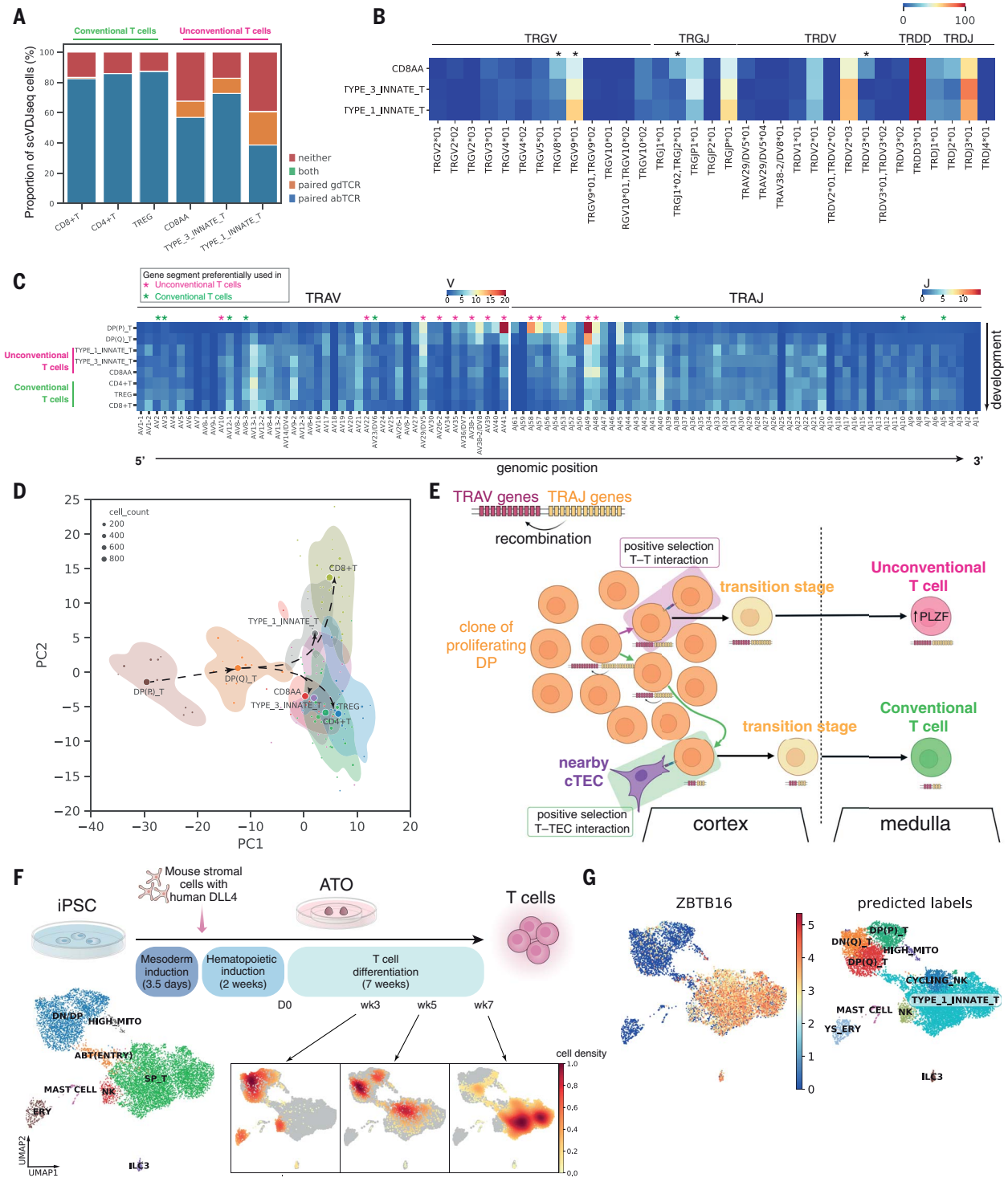


Fig. 6. Deep characterization of human unconventional T cells. (**A**) Proportions of cells expressing paired γδTCR, paired αβTCR, both, or neither. The proportions were calculated over cells that had both single-cell αβTCR and γδTCR sequencing. Expression of neither paired αβTCR nor paired γδTCR in some cells could be due to dropouts in single-cell TCR sequencing because >50% of these contained orphan VDJ or VJ chains of αβTCR or γδTCR (fig. S28A). (**B**) Heatmap showing the percentage of each γδTCR gene segment present in different T cell subtypes. Differential usage between cell subtypes was computed using the chi-squared test. Gene segments with Benjamini-Hochberg–adjusted *P* values < 0.05 are marked with asterisks. (**C**) Heatmap showing the proportion of each TCRα gene segment present in different T cell subtypes. The gene segment usage in unconventional T cells and conventional T cells was compared using logistic regression, controlling for donors and organs. Gene segments with Benjamini-Hochberg–adjusted *P* values < 0.05 are marked with magenta asterisks for preferential usage in unconventional T cells and green asterisks for preferential usage in conventional T cells. (**D**) Principal component analysis plot summarizing TRAV, TRAJ, TRBV, and TRBJ gene segment usage proportion in different T cell subtypes. Each dot represents a sample of at least 20 cells, with dot size representing the cell count. The centroid of each cell type is shown as a filled circle, and 80% confidence

contours are shown around the centroids. Arrows illustrate the proposed developmental trajectories. (**E**) Schematic illustration showing the T-T training origin of unconventional T cells in contrast to the T-TEC training origin of conventional T cells. (**F**) Top: schematic showing the experimental setup of T cell differentiation from iPSCs in ATOs. Bottom left: UMAP visualization of different cell types in the ATO. Bottom right: density plots of cells from each time point over UMAP embedding. (**G**) Left: predicted annotations from logistic regression overlaid on the same UMAP plot as in (F); right: *ZBTB16* expression pattern overlaid onto the same UMAP plot.

Optimal 18-min permeabilization was selected for fetal spleen and liver, and a 24-min permeabilization was used for fetal thymus. The spatial gene expression library was then generated following the manufacturer's protocol. All images for this process were acquired with an Axio Imager (Carl Zeiss Microscopy) and a 20× air objective [0.8 numerical aperture (NA)] using either fluorescence (Zeiss Axiocam 503 monochrome camera) for optimization or bright-field mode (Zeiss Axiocam 105 color camera) for hematoxylin & eosin (H&E) imaging. ZEN (blue edition) v.3.1 software was used for acquisition and stitching of the image tiles.

smFISH

smFISH was performed on thymus, spleen, and gut sections using the RNAScope 2.5 LS multiplex fluorescent assay (ACD, Bio-Techne) on the automated BOND RX system (Leica). Slides were stained for DAPI (nuclei), and three or four probes of interest were stained with fluorophores atto 425, opal 520, opal 570, and opal 650. Positive and negative control probes were used to optimize staining conditions for all tissues.

For fetal gut and spleen, OCT-embedded, freshly frozen, 10-μm-thick sections were pretreated offline for 15 min with chilled 4% paraformaldehyde and dehydrated through an ethanol series (50, 70, 100, and 100% ethanol) before processing on the Leica BOND RX with protease IV for 30 min at room temperature. The sections were imaged on a PerkinElmer Opera Phenix High Content Screening System (16-bit sCMOS camera, PerkinElmer) with a 20× water objective (High NA, PerkinElmer). Because of the high levels of endogenous autofluorescence, one of the spleen sections (fig. S21A) was imaged with a confocal microscope (Leica SP8) with a 40× 1.3 NA oil immersion objective and SP8 Leica HyD and PMT detectors.

Because of the high cellular density in thymic sections, 3-μm-thick formalin-fixed, paraffin-embedded sections that were treated on the Leica Bond RX with epitope retrieval 2 for 15 min at 95°C and protease III for 15 min at 40°C were used. Imaging was performed on an Operetta CLS High Content Screening System (16-bit sCMOS camera, PerkinElmer) with a 40× water objective (High NA, PerkinElmer) and 2-μm *z*-steps.

scRNA-seq analysis

Preprocessing

The gene expression data were mapped with cellranger 3.0.2 to an Ensembl 93–based GRCh38 reference (10X Genomics–distributed 3.0.0 version). Ambient RNA was removed with cellbender v0.2.0 (*76*). Low-quality cells were filtered out [minimum number of reads = 2000, minimum number of genes = 500, Scrublet (v0.2.3) (*77*) doublet detection score <0.4]. Possible maternal contamination was identified using the souporcell pipeline for genotyping (v.2.4.0) (*78*) (for more details, see the supplementary materials).

Data integration and annotation

Data normalization and preprocessing were performed using the Scanpy workflow (v1.8.1) (*79*). Raw gene read counts were normalized by sequencing depth in each cell (*scanpy.pp.normalize_per_cell*, with parameters *counts_per_cell_after=10e4*) and performed *ln(x)+1* transformation. Highly variable genes were then selected for joint embedding by dispersion (*scanpy.pp.highly_variable_genes* with parameters *min_mean = 0.001*, *max_mean = 10*). Dimensionality reduction and batch correction were performed using the scVI model (*12*) as implemented in scvi-tools (v0.14.5) (*80*), considering 10X Genomics chemistry (5′ and 3′) and the donor ID for each cell as the technical covariates to correct for (training parameters: *dropout_rate = 0.2*, *n_layers = 2*). The model was trained on raw counts of the 7500 most highly variable genes, excluding cell cycle genes and TCR/BCR genes (*7*) with 20 latent dimensions. To verify conservation of biological variation after integration, the available cell-type labels from the published datasets (66% of cells) were collected and harmonized, and the agreement between labels across different datasets was quantified in the cell clusters identified after integration using the normalized mutual information score, as implemented in *scikit-learn* (*81*). Unless otherwise specified, cell clustering was performed using the Leiden algorithm (*82*) (*resolution = 1.5*, *n_neighbors* = 30). To verify robustness to the choice of integration method, integration was performed in parallel using batched-balanced k-nearest neighbor (BBKNN) (*83*), as previously described (*7*) (fig. S30A). It was verified that clustering after integration with both scVI and BBKNN was consistent with previous annotations (fig. S30B).

To annotate fine cell populations across tissues, cells were clustered in the scVI latent space and preliminarily assigned to broad lineages using the expression of marker genes and previous annotations. For each broad lineage, scVI integration and clustering were repeated as described above and further subsets were defined (see hierarchy in fig. S5). Leiden clusters for the highest-resolution subsets (stroma, megakaryocyte/erythroid, progenitors, lymphoid, and myeloid) were annotated manually using the marker panels shown in fig. S4 (for a more detailed description of annotation strategy, see the supplementary materials). It was verified that refined annotations were highly consistent with unsupervised clustering after integration on the full dataset both with scVI and BBKNN (fig. S30C).

After full annotation, 23,156 cells (2.5% of total) were assigned to low-quality clusters (doublet clusters, maternal contaminants clusters, and clusters displaying a high percentage of reads from mitochondrial genes).

Differential abundance analysis

Differences in cell abundances associated with gestational age or organ were tested for using the Milo framework for differential abundance testing (*22*), with the Python implementation milopy (https://github.com/emdann/milopy). A more detailed description of this analysis can be found in the supplementary materials and methods.

Briefly, the dataset was subsetted to cells from libraries obtained with $CD45^+$ FACS, $CD45^-$ FACS, or no FACS, excluding FACS-isolated samples for which the true sorting fraction quantification could not be recovered. In total, 228,731 lymphoid cells and 214,874 myeloid cells were retained. To further minimize the differences in cell numbers driven by FACS efficiency, a FACS correction factor was calculated for each sample to use as a confounding covariate in differential abundance testing (fig. S32 and supplementary materials and methods). A KNN graph was constructed using similarity in the scVI embedding (k = 30 for test across gestation, k = 100 for test across tissues) and cells were assigned to neighborhoods (*milopy.core.make_nhoods*, parameters: *prop = 0.05*). The cells belonging to each sample in each neighborhood were then counted (*milopy.core.count_cells*). Each neighborhood was assigned a cell-type label on the basis of majority voting of the cells belonging to that neighborhood. A "mixed" label was assigned if the most abundant label was present in <50% of cells within that neighborhood.

To test for differential abundance across gestational age, the sample ages were divided into six equally sized bins (bin size = 2 pcw) and samples from organs in which fewer than three consecutive age bins were profiled were excluded (yolk sac, mesenteric lymph node,

kidney, and gut). The cell count in neighborhoods was modeled as a negative-binomial generalized linear model using a log-linear model to model the effects of age on cell counts while accounting for the FACS correction factor and the total number of cells over all neighborhoods. Multiple testing was controlled for using the weighted Benjamini-Hochberg correction, as described in (*22*). To detect markers of early-specific neighborhoods [spatial false discovery rate (spatialFDR) < 0.1, logFC < 0] and/or late-specific neighborhoods [spatialFDR < 0.1, logFC > 0] in cell type c and organ o, differential expression was tested for between cells from organ o assigned to the significant neighborhoods labeled as cell type c and cells belonging to all other neighborhoods labeled as cell type c. The t-test implementation in scanpy was used (*scanpy.tl.rank_genes_groups*, *method = "t-test_overestim_var"*). Genes expressed in >70% of tested cells were excluded. Genes were considered as significantly overexpressed (i.e., markers) if the differential expression logFC > 1 and FDR < 0.1%. Gene set enrichment analysis was performed using the implementation of the EnrichR workflow (*84*) in the Python package gseapy (https://gseapy.readthedocs.io/). The list of significantly overexpressed genes for all organs and cell types in which differential expression testing was performed can be found in tables S1 and S3.

To test for differential abundance between organs, the cell count was modeled in neighborhoods as above, using a log-linear model to model the effects of organ on cell counts while accounting for FACS correction factor, library prep protocol, and the total number of cells over all the neighborhoods. Neighborhoods in which $\beta_n^o > 0$ and spatialFDR < 0.01 were considered to be the cell subpopulations that showed organ-specific transcriptional signatures.

Having identified a subset of neighborhoods overlapping a cell type that was enriched in a certain organ, differential expression analysis was performed between these cells and cells from the same cell type (for more details, see the supplementary materials and methods). Briefly, single-cell expression profiles were first aggregated into pseudobulk expression profiles $\hat{x}$ for each cell type c and sample s [as recommended by (*85*, *86*)].

The mRNA counts of gene g in pseudobulk p were then modeled by a negative-binomial generalized linear model:

$$\bar{x}^{g,p} = NB\left(\mu_{g,p}\phi_{g,p}\right)$$

The expected count value $\mu_{n,p}$ is given by the following log-linear model:

$$\log \mu_{g,p} = \beta_0 + d_p\beta_g^{\text{donor}} + o_p\beta_g^{\text{organ}} + c_p\beta_g^{\text{celltype}} + c_p o_p\beta_g^{\text{organ}\times\text{celltype}} + \log L_p$$

The log-fold change $\beta_g^{\text{organ}\times\text{celltype}}$ in expression in a given cell type for organ $\hat{o}$ was estimated using the quasi-likelihood method (*87*) implemented in the R package glmGamPoi (*85*). The estimated logFC from the test on a set of control cell types (where organ-specific differences would not be expected) was used to filter out genes in which differential expression is driven by technical differences in tissue processing. The full results for the differential expression analysis between organs in mature T cells and monocytes are provided in tables S2 and S4.

TCR analysis

Single-cell αβTCR-sequencing data were mapped with cellranger-vdj (v.6.0.0). The output file *filtered_contig_annotations.csv* was used and analyzed with scirpy (v.0.6.0) (*88*). Single-cell γδTCR-sequencing data were mapped with cellranger-vdj (v.4.0.0). All contigs deemed high quality were selected and re-annotated with igblastn (v.1.17.1) against IMGT (international ImMunoGeneTics) reference sequences (last downloaded: 01/08/2021) through a workflow provided in dandelion (v0.2.0) (*89*) (https://github.com/zktuong/dandelion). The output file *all_contig_dandelion.tsv* was used and analyzed with scirpy (v0.6.0).

BCR analysis

Single-cell BCR data were initially processed with cellranger-vdj (v.6.0.0). BCR contigs contained in *all_contigs.fasta* and *all_contig_annotations.csv* were then processed further using dandelion (*89*) singularity container (v.0.2.0). BCR mutation frequencies were obtained using the *observedMutations* function in shazam (v.1.0.2) (*90*) with default settings.

B cell activation scoring

The Gene Ontology B Cell Activation gene list was downloaded from the Gene Set Enrichment Analysis website (http://www.gsea-msigdb.org/gsea/msigdb/genesets.jsp). Cells were scored according to expression values of all genes in this gene list, apart from three genes that were not present in the dataset using *scanpy.tl.score_genes()* function.

Transcription factor activity inference

The Python package DoRothEA (v.1.0.5) (*91*) was used to infer TF activities in B1 cells and mature B cells. TFs that had higher activities (positive "mean change") in B1 cells were then ranked according to their adjusted P values, and only the top 25 TFs are shown in fig. S26F.

Cell-cell interaction analysis

The Python package CellPhoneDB (v.3.0) (*92*, *93*) was used to infer cell-cell interactions. The scRNA-seq dataset was split by organ, and cell types with <20 cells in a given organ were filtered out. CellPhoneDB was run separately to infer cell-cell interactions in each organ using default parameters. To explore cell-cell interactions between B cell progenitors and colocalizing cell types (fig. S24D), the interactions predicted between each colocalizing cell type were aggregated by averaging the means and using the minimum of the P values to filter for significance. The ligand-receptor pairs that were significant ($P < 0.05$) across all three organs, liver, spleen, and thymus, were filtered and ranked by the maximum aggregated means. Only the top 60 ligand-receptor pairs are shown in fig. S24C.

Query-to-reference mapping

Query data were mapped to our prenatal data embeddings using online update of the scVI models following the scArches method (*15*), as implemented in the *scvi-tools* package (*80*). The model was trained for 200 epochs and by setting *weight_decay = 0* to ensure that the latent representation of the reference cells remained exactly the same. Reference genes missing in the query were set to zero, as recommended in (*15*). To generate a joint embedding of query and reference cells, the latent dimensions learned for query cells were concatenated to the latent dimensions used for the reference embedding and the KNN graph and uniform manifold approximation and projection (UMAP) were computed as described above. To assess that the mapping to the developmental reference conserved biological variation while minimizing technical variation in the query data, query cell-type labels and batch labels were compared with clusters obtained from Leiden clustering on the learned latent dimensions using the normalized mutual information score (see fig. S33 for mapping of adult query data).

Annotation prediction using CellTypist

The Python package CellTypist (v.0.1.9) (*21*) was used to perform annotation prediction with logistic regression models. For prediction on cycling B cells, the rest of the nonprogenitor B cells, including immature B, mature B, B1, and plasma B cells, were used as a training dataset. Default parameters were used for model building, and prediction was made without majority voting for accurate enumeration of predicted B cell subtypes within cycling B cells.

Comparison with human adult immune cells

scRNA-seq data from adult immune cells were generated and preprocessed as described previously (*21*). The dataset, including cell-type annotations, was downloaded from https://www.tissueimmunecellatlas.org/. A total of 264,929 adult lymphoid cells were mapped to the lymphoid embeddings of our developmental dataset and 54,047 adult myeloid cells to our myeloid embedding. To use cell annotations in our developmental dataset to predict adult cell types in the joint embedding,

the KNN-classifier approach described in (*15*) was adapted and the similarity to prenatal cells labeled was calculated taking the Euclidean distance in the joint embedding weighted by a Gaussian kernel.

Blood and immune cell progenitor scRNA-seq data analysis

For the cell fate prediction analysis shown in fig. S20, C and D, the Palantir method as implemented in CellRank was used (*94*, *95*). Briefly, from the scVI embedding on all immune cells (fig. S20A), cells belonging to progenitor populations were selected and a KNN graph on scVI latent dimensions on these cells was computed (k = 30). Then, transition probabilities were calculated using the *ConnectivityKernel* in the cellrank package. Coarse-grained macrostates were calculcated with the Generalized Perron Cluster Cluster Analysis, setting the number of macrostates to the number of annotated progenitor cell populations. The four target terminal states were set manually for each lineage (small pre B cells, DN(Q) T cells, early megakaryocytes, and promonocytes) and the probability of each cell to transition to one of the four terminal states was calculated. The fate simplex visualization in fig. S20, C and D, was generated using the function *cellrank.pl.circular_projection*.

ATO scRNA-seq data analysis

Raw scRNA-seq reads were mapped with cellranger 3.0.2 with combined human reference of GRCh38.93 and mouse reference of mm10-3.1.0. Low-quality cells were filtered out [minimum number of reads = 2000, minimum number of genes = 500, minimum Scrublet (*77*) doublet detection score <0.4]. Cells in which the percentage of counts from human genes was <90% were considered as mouse cells and were excluded from downstream analysis. Cells were assigned to different cell lines (Kolf and Fiaj) using genotype prediction with souporcell (v.2.4.0) (*78*). Batch correction was performed to minimize the differences between cells from different cell lines using scVI and clustered cells using the Leiden algorithm on the latent embedding as described above. The Python package CellTypist (v.0.1.9) (*21*) was used to perform annotation prediction with logistic regression using the whole in vivo scRNA-seq developmental dataset for training. For the in vivo to in vitro similarity analysis in fig. S29D, in vitro cells were mapped to the scVI model of lymphoid cells as described above. For each cell in the in vitro dataset, the Euclidean distance weighted by a Gaussian kernel to the closest in vivo cell from each in vivo cell population was calculated.

Spatial data analysis

Spatial transcriptomics data were mapped using spaceranger (v.1.2.1), and a custom image-processing script was used to identify regions overlapping tissues. To map cell types identified by scRNA-seq in the profiled spatial transcriptomics slides, the cell2location method was used (*16*) (see the supplementary materials and methods). Briefly, for the reference model training step, very lowly expressed genes were excluded using a recommended filtering strategy (*16*). Cell types in which <20 cells were profiled in the organ of interest and cell types labeled as low-quality cells were excluded from the reference. For the analysis of unconventional T cell localization in thymus (fig. S27C), a reference adding all the prenatal TECs from a thymus cell atlas was trained (*7*) [data were downloaded from Zenodo (*96*)]. For the spatial cell-type deconvolution step, all slides representing a given organ were analyzed jointly. To identify microenvironments of colocalizing cell types, NMF was used on the matrix of estimated cell-type abundances. Here, latent factors correspond to tissue microenvironments defined by a set of colocalized cell types. The NMF implementation in scikit-learn was used (*81*), setting the number of factors at 10. For downstream analysis, cell types in which the 99% quantile of cell abundance across locations in every slide from the same organ was always below the detection threshold of 0.15 were excluded. Unless otherwise specified, a cell type was considered to be part of a microenvironment if the cell-type fraction was >0.2.

For analysis of mature T cell localization in the thymic medulla (fig. S27, D and E), factors in which the sum of the cell-type fractions for mature T cells ($CD4^+$, $CD8^+$, T_{reg}, type 1 innate, type 3 innate, and CD8AA T cells) was >0.8 were retained. Spots were assigned to the inner medulla or corticomedullary microenvironment if the factor value in the spot was above the 90% quantile of all values in the slide. To annotate cortex and medulla from histology images, image features were extracted from the high-resolution images of H&E staining using the Python package squidpy (v1.1.2) (*97*), and Leiden clustering was performed on image features. The corticomedullary junction was then defined using spatial neighbor graph functionality in squidpy (see the supplementary materials and methods).

B1 functional validation experiment

Cryopreserved single-cell suspensions from F144 (17 pcw) and F145 (15 pcw) spleen samples were used for the ELISpot experiment. B cells were gated as singlet $DAPI^-CD3^-CD20^+$ cells. Plasma cells should generally be $CD20^{lo}$ and therefore are not included. To further exclude plasma cell contamination, the top 1% of B cells expressing the highest level of CD38 were gated out. The rest of the B cells were then sorted into four fractions: $CCR10^{hi}$, $CCR10^{lo}$ $CD27^+CD43^+$, $CCR10^{lo}CD27^-CD43^+$, and $CCR10^{lo}$ $CD27^-CD43^-$. CD27 and CD43 gates were chosen on the basis of fluorescence minus one controls.

The ELISpot experiment was performed with the Human IgM $ELISpot^{BASIC}$ kit (ALP) from Mabtech AB. After sorting, 7000 to 8000 cells were added into an ELISpot plate precoated with anti-IgM antibody and incubated at 37°C for 22 hours. The plate was then washed and incubated with biotinylated anti-IgM for 2 hours at room temperature, followed by a 1-hour incubation with streptavidin-ALP. The colored spots were developed with a 15-min incubation of 5-bromo-4-chloro-3-indolyl phosphate (BCIP)/nitro blue tetrazolium (NBT) substrate solution. Spots were counted with the AID ELISpot reader and iSpot software version 4.

In addition, scRNA-seq of the sorted B cell fractions was performed on a different donor (F149, 18 pcw fetal spleen) using the same gating strategy to further confirm the identity of sorted cells. The scRNA-seq data were preprocessed with scVI as above. Cell annotations were predicted using CellTypist v.0.1.9 (*21*).

REFERENCES AND NOTES

1. J.-E. Park, L. Jardine, B. Gottgens, S. A. Teichmann, M. Haniffa, Prenatal development of human immunity. *Science* **368**, 600–603 (2020). doi: 10.1126/science.aaz9330; pmid: 32381715
2. M. Jagannathan-Bogdan, L. I. Zon, Hematopoiesis. *Development* **140**, 2463–2467 (2013). doi: 10.1242/dev.083147; pmid: 23715539
3. D.-M. Popescu *et al*., Decoding human fetal liver haematopoiesis. *Nature* **574**, 365–371 (2019). doi: 10.1038/s41586-019-1652-y; pmid: 31597962
4. B. J. Stewart *et al*., Spatiotemporal immune zonation of the human kidney. *Science* **365**, 1461–1466 (2019). doi: 10.1126/science.aat5031; pmid: 31604275
5. Y. Zeng *et al*., Tracing the first hematopoietic stem cell generation in human embryo by single-cell RNA sequencing. *Cell Res.* **29**, 881–894 (2019). doi: 10.1038/s41422-019-0228-6; pmid: 31501518
6. Y. Zeng *et al*., Single-cell RNA sequencing resolves spatiotemporal development of pre-thymic lymphoid progenitors and thymus organogenesis in human embryos. *Immunity* **51**, 930–948.e6 (2019). doi: 10.1016/j.immuni.2019.09.008; pmid: 31604687
7. 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; pmid: 32079746
8. R. Elmentaite *et al*., Single-cell sequencing of developing human gut reveals transcriptional links to childhood Crohn's disease. *Dev. Cell* **55**, 771–783.e5 (2020). doi: 10.1016/j.devcel.2020.11.010; pmid: 33290721
9. J. Cao *et al*., A human cell atlas of fetal gene expression. *Science* **370**, eaba7721 (2020). doi: 10.1126/science.aba7721; pmid: 33184181
10. G. Reynolds *et al*., Developmental cell programs are co-opted in inflammatory skin disease. *Science* **371**, eaba6500 (2021). doi: 10.1126/science.aba6500; pmid: 33479125
11. L. Jardine *et al*., Blood and immune development in human fetal bone marrow and Down syndrome. *Nature* **598**, 327–331 (2021). doi: 10.1038/s41586-021-03929-x; pmid: 34588693
12. 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; pmid: 30504886
13. D. Pellin *et al*., A comprehensive single cell transcriptional landscape of human hematopoietic progenitors. *Nat. Commun.* **10**, 2395 (2019). doi: 10.1038/s41467-019-10291-0; pmid: 31160568
14. 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; pmid: 28428369

15. M. Lotfollahi *et al*., Mapping single-cell data to reference atlases by transfer learning. *Nat. Biotechnol.* **40**, 121–130 (2022). doi: 10.1038/s41587-021-01001-7; pmid: 34462589
16. V. Kleshchevnikov *et al*., Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat. Biotechnol.*10.1038/s41587-021-01139-4 (2022). doi: 10.1038/s41587-021-01139-4; pmid: 35027729
17. S. Z. Chong *et al*., CXCR4 identifies transitional bone marrow premonocytes that replenish the mature monocyte pool for peripheral responses. *J. Exp. Med.* **213**, 2293–2314 (2016). doi: 10.1084/jem.20160800; pmid: 27811056
18. S. A. MacParland *et al*., Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. *Nat. Commun.* **9**, 4383 (2018). doi: 10.1038/s41467-018-06318-7; pmid: 30348985
19. E. Gerrits, Y. Heng, E. W. G. M. Boddeke, B. J. L. Eggen, Transcriptional profiling of microglia; current state of the art and future perspectives. *Glia* **68**, 740–755 (2020). doi: 10.1002/glia.23767; pmid: 31846124
20. S. M. Toor, S. Wani, O. M. E. Albagha, Comprehensive transcriptomic profiling of murine osteoclast differentiation reveals novel differentially expressed genes and lncRNAs. *Front. Genet.* **12**, 781272 (2021). doi: 10.3389/fgene.2021.781272; pmid: 34868271
21. C. Domínguez Conde *et al*., Cross-tissue immune cell analysis reveals tissue-specific features in humans. *Science* **376**, eabl5197 (2022). doi: 10.1126/science.abl5197; pmid: 35549406
22. E. Dann, N. C. Henderson, S. A. Teichmann, M. D. Morgan, J. C. Marioni, Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nat. Biotechnol.* **40**, 245–253 (2022). pmid: 34594043
23. K. C. M. Jeucken, J. J. Koning, R. E. Mebius, S. W. Tas, The role of endothelial cells and TNF-receptor superfamily members in lymphoid organogenesis and function during health and inflammation. *Front. Immunol.* **10**, 2700 (2019). doi: 10.3389/fimmu.2019.02700; pmid: 31824495
24. X. Yang *et al*., Essential contribution of a chemokine, CCL3, and its receptor, CCR1, to hepatocellular carcinoma progression. *Int. J. Cancer* **118**, 1869–1876 (2006). doi: 10.1002/ijc.21596; pmid: 16284949
25. F. Hua, Y. Tian, CCL4 promotes the cell proliferation, invasion and migration of endometrial carcinoma by targeting the VEGF-A signal pathway. *Int. J. Clin. Exp. Pathol.* **10**, 11288–11299 (2017). pmid: 31966483
26. E. C. Keeley, B. Mehrad, R. M. Strieter, CXC chemokines in cancer angiogenesis and metastases. *Adv. Cancer Res.* **106**, 91–111 (2010). doi: 10.1016/S0065-230X(10)06003-3; pmid: 20399957
27. J. Heidemann *et al*., Angiogenic effects of interleukin 8 (CXCL8) in human intestinal microvascular endothelial cells are mediated by CXCR2. *J. Biol. Chem.* **278**, 8508–8515 (2003). doi: 10.1074/jbc.M208231200; pmid: 12496258
28. K. Norrby, Mast cells and angiogenesis. *APMIS* **110**, 355–371 (2002). doi: 10.1034/j.1600-0463.2002.100501.x; pmid: 12076253
29. D. Ribatti, E. Crivellato, The role of mast cell in tissue morphogenesis. Thymus, duodenum, and mammary gland as examples. *Exp. Cell Res.* **341**, 105–109 (2016). doi: 10.1016/j.yexcr.2015.11.022; pmid: 26615957
30. W. Wood, P. Martin, Macrophage functions in tissue patterning and disease: New insights from the fly. *Dev. Cell* **40**, 221–233 (2017). doi: 10.1016/j.devcel.2017.01.001; pmid: 28171746
31. K. Hoorweg, T. Cupedo, Development of human lymph nodes and Peyer's patches. *Semin. Immunol.* **20**, 164–170 (2008). doi: 10.1016/j.smim.2008.02.003; pmid: 18424165
32. P. Rantakari *et al*., Fetal liver endothelium regulates the seeding of tissue-resident macrophages. *Nature* **538**, 392–396 (2016). doi: 10.1038/nature19814; pmid: 27732581
33. N. Li *et al*., Memory $CD4^+$ T cells are generated in the human fetal intestine. *Nat. Immunol.* **20**, 301–312 (2019). doi: 10.1038/s41590-018-0294-9; pmid: 30664737
34. A. Mishra *et al*., Microbial exposure during early human development primes fetal immune cells. *Cell* **184**, 3394–3409.e20 (2021). doi: 10.1016/j.cell.2021.04.039; pmid: 34077752
35. Y. Xing, X. Wang, S. C. Jameson, K. A. Hogquist, Late stages of T cell maturation in the thymus involve NF-κB and tonic type I interferon signaling. *Nat. Immunol.* **17**, 565–573 (2016). doi: 10.1038/ni.3419; pmid: 27043411
36. L. V. Webb, S. C. Ley, B. Seddon, TNF activation of NF-κB is essential for development of single-positive thymocytes. *J. Exp. Med.* **213**, 1399–1407 (2016). doi: 10.1084/jem.20151604; pmid: 27432943
37. C. Collins, E. Sharpe, A. Silber, S. Kulke, E. W. Y. Hsieh, Congenital athymia: Genetic etiologies, clinical manifestations, diagnosis, and treatment. *J. Clin. Immunol.* **41**, 881–895 (2021). doi: 10.1007/s10875-021-01059-7; pmid: 33987750
38. P. G. Holt, C. A. Jones, The development of the immune system during pregnancy and early life. *Allergy* **55**, 688–697 (2000). doi: 10.1034/j.1398-9995.2000.00118.x; pmid: 10955693
39. D. O. Griffin, N. E. Holodick, T. L. Rothstein, Human B1 cells in umbilical cord and adult peripheral blood express the novel phenotype CD20+ CD27+ CD43+ CD70-. *J. Exp. Med.* **208**, 67–80 (2011). doi: 10.1084/jem.20101499; pmid: 21220451
40. D. O. Griffin, T. L. Rothstein, Human b1 cell frequency: Isolation and analysis of human b1 cells. *Front. Immunol.* **3**, 122 (2012). doi: 10.3389/fimmu.2012.00122; pmid: 22654880
41. T. L. Rothstein, D. O. Griffin, N. E. Holodick, T. D. Quach, H. Kaku, Human B-1 cells take the stage. *Ann. N. Y. Acad. Sci.* **1285**, 97–114 (2013). doi: 10.1111/nyas.12137; pmid: 23692567
42. N. Baumgarth, The double life of a B-1 cell: Self-reactivity selects for protective effector functions. *Nat. Rev. Immunol.* **11**, 34–46 (2011). doi: 10.1038/nri2901; pmid: 21151033
43. P. A. Lalor, L. A. Herzenberg, S. Adams, A. M. Stall, Feedback regulation of murine Ly-1 B cell development. *Eur. J. Immunol.* **19**, 507–513 (1989). doi: 10.1002/eji.1830190315; pmid: 2785046
44. K. Hayakawa, R. R. Hardy, D. R. Parks, L. A. Herzenberg, The "Ly-1 B" cell subpopulation in normal immunodefective, and autoimmune mice. *J. Exp. Med.* **157**, 202–218 (1983). doi: 10.1084/jem.157.1.202; pmid: 6600267
45. E. Montecino-Rodriguez, K. Dorshkind, B-1 B cell development in the fetus and adult. *Immunity* **36**, 13–21 (2012). doi: 10.1016/j.immuni.2011.11.017; pmid: 22284417
46. A. B. Kantor, C. E. Merrill, L. A. Herzenberg, J. L. Hillson, An unbiased analysis of V(H)-D-J(H) sequences from B-1a, B-1b, and conventional B cells. *J. Immunol.* **158**, 1175–1186 (1997). pmid: 9013957
47. U. C. Tornberg, D. Holmberg, B-1a, B-1b and B-2 B cells display unique VHDJH repertoires formed at different stages of ontogeny and under different selection pressures. *EMBO J.* **14**, 1680–1689 (1995). doi: 10.1002/j.1460-2075.1995.tb07157.x; pmid: 7737121
48. M. Miyama-Inaba *et al*., Unusual phenotype of B cells in the thymus of normal mice. *J. Exp. Med.* **168**, 811–816 (1988). doi: 10.1084/jem.168.2.811; pmid: 3261779
49. R. Elmentaite *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; pmid: 34497389
50. J. Schulze-Luehrmann, S. Ghosh, Antigen-receptor signaling to nuclear factor kappa B. *Immunity* **25**, 701–715 (2006). doi: 10.1016/j.immuni.2006.10.010; pmid: 17098202
51. E. S. Alonzo, D. B. Sant'Angelo, Development of PLZF-expressing innate T cells. *Curr. Opin. Immunol.* **23**, 220–227 (2011). doi: 10.1016/j.coi.2010.12.016; pmid: 21257299
52. T. Dimova *et al*., Effector $V\gamma9V\delta2$ T cells dominate the human fetal γδ T-cell repertoire. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E556–E565 (2015). doi: 10.1073/pnas.1412058112; pmid: 25617367
53. L. Tan *et al*., A fetal wave of human type 3 effector γδ cells with restricted TCR diversity persists into adulthood. *Sci. Immunol.* **6**, eabf0125 (2021). doi: 10.1126/sciimmunol.abf0125; pmid: 33893173
54. T. Mayassi, L. B. Barreiro, J. Rossjohn, B. Jabri, A multilayered immune system through the lens of unconventional T cells. *Nature* **595**, 501–510 (2021). doi: 10.1038/s41586-021-03578-0; pmid: 34290426
55. Z. M. Carico, K. Roy Choudhury, B. Zhang, Y. Zhuang, M. S. Krangel, Tcrd rearrangement redirects a processive Tcra recombination program to expand the Tcra repertoire. *Cell Rep.* **19**, 2157–2173 (2017). doi: 10.1016/j.celrep.2017.05.045; pmid: 28591585
56. Y. J. Lee *et al*., Generation of PLZF+ CD4+ T cells via MHC class II-dependent thymocyte-thymocyte interaction is a physiological process in humans. *J. Exp. Med.* **207**, 237–246 (2010). doi: 10.1084/jem.20091519; pmid: 20038602
57. H. Cho, Y. Bediako, H. Xu, H.-J. Choi, C.-R. Wang, Positive selecting cell type determines the phenotype of MHC class Ib-restricted CD8+ T cells. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13241–13246 (2011). doi: 10.1073/pnas.1105118108; pmid: 21788511
58. H. Georgiev, C. Peng, M. A. Huggins, S. C. Jameson, K. A. Hogquist, Classical MHC expression by DP thymocytes impairs the selection of non-classical MHC restricted innate-like T cells. *Nat. Commun.* **12**, 2308 (2021). doi: 10.1038/s41467-021-22589-z; pmid: 33863906
59. E. S. Hoffman *et al*., Productive T-cell receptor beta-chain gene rearrangement: Coincident regulation of cell cycle and clonality during development in vivo. *Genes Dev.* **10**, 948–962 (1996). doi: 10.1101/gad.10.8.948; pmid: 8608942
60. H. Spits, Development of alphabeta T cells in the human thymus. *Nat. Rev. Immunol.* **2**, 760–772 (2002). doi: 10.1038/nri913; pmid: 12360214
61. A. Montel-Hagen *et al*., Organoid-induced differentiation of conventional T cells from human pluripotent stem cells. *Cell Stem Cell* **24**, 376–389.e8 (2019). doi: 10.1016/j.stem.2018.12.011; pmid: 30661959
62. E. Mass *et al*., Specification of tissue-resident macrophages during organogenesis. *Science* **353**, aaf4238 (2016). doi: 10.1126/science.aaf4238; pmid: 27492475
63. N. Mende, E. Laurenti, Hematopoietic stem and progenitor cells outside the bone marrow: Where, when, and why. *Exp. Hematol.* **104**, 9–16 (2021). doi: 10.1016/j.exphem.2021.10.002; pmid: 34687807
64. N. Mende *et al*., Unique molecular and functional features of extramedullary hematopoietic stem and progenitor cell reservoirs in humans. *Blood*blood.2021013450 (2022). doi: 10.1182/blood.2021013450; pmid: 35073399
65. S. Krishnan *et al*., Hematopoietic stem and progenitor cells are present in healthy gingiva tissue. *J. Exp. Med.* **218**, e20200737 (2021). doi: 10.1084/jem.20200737; pmid: 33635312
66. C. H. Kim, Homeostatic and pathogenic extramedullary hematopoiesis. *J. Blood Med.* **1**, 13–19 (2010). doi: 10.2147/JBM.S7224; pmid: 22282679
67. S. Brioschi *et al*., Heterogeneity of meningeal B cells reveals a lymphopoietic niche at the CNS borders. *Science* **373**, eabf9277 (2021). doi: 10.1126/science.abf9277; pmid: 34083450
68. D. Schafflick *et al*., Single-cell profiling of CNS border compartment leukocytes reveals that B cells and their progenitors reside in non-diseased meninges. *Nat. Neurosci.* **24**, 1225–1234 (2021). doi: 10.1038/s41593-021-00880-y; pmid: 34253922
69. Y. Wang *et al*., Early developing B cells undergo negative selection by central nervous system-specific antigens in the meninges. *Immunity* **54**, 2784–2794.e6 (2021). doi: 10.1016/j.immuni.2021.09.016; pmid: 34626548
70. E. Montecino-Rodriguez, H. Leathers, K. Dorshkind, Identification of a B-1 B cell-specified progenitor. *Nat. Immunol.* **7**, 293–301 (2006). doi: 10.1038/ni1301; pmid: 16429139
71. B. L. Esplin, R. S. Welner, Q. Zhang, L. A. Borghesi, P. W. Kincade, A differentiation pathway for B1 cells in adult bone marrow. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5773–5778 (2009). doi: 10.1073/pnas.0811632106; pmid: 19307589
72. M. Yoshimoto *et al*., Embryonic day 9 yolk sac and intra-embryonic hemogenic endothelium independently generate a B-1 and marginal zone progenitor lacking B-2 potential. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1468–1473 (2011). doi: 10.1073/pnas.1015841108; pmid: 21209332
73. T. Kreslavsky, J. B. Wong, M. Fischer, J. A. Skok, M. Busslinger, Control of B-1a cell development by instructive BCR signaling. *Curr. Opin. Immunol.* **51**, 24–31 (2018). doi: 10.1016/j.coi.2018.01.001; pmid: 29414528
74. R. Graf *et al*., BCR-dependent lineage plasticity in mature B cells. *Science* **363**, 748–753 (2019). doi: 10.1126/science.aau8475; pmid: 30765568
75. E. P. Mimitou *et al*., Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat. Methods* **16**, 409–412 (2019). doi: 10.1038/s41592-019-0392-0; pmid: 31011186
76. S. J. Fleming, J. C. Marioni, M. Babadi, CellBender remove-background: a deep generative model for unsupervised removal of background noise from scRNA-seq datasets, bioRxiv 791699 [Preprint] (2019); .doi: 10.1101/791699
77. 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; pmid: 30954476
78. H. Heaton *et al*., Souporcell: Robust clustering of single-cell RNA-seq data by genotype without reference genotypes. *Nat. Methods* **17**, 615–620 (2020). doi: 10.1038/s41592-020-0820-1; pmid: 32366989
79. 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; pmid: 29409532
80. A. Gayoso *et al*., A Python library for probabilistic analysis of single-cell omics data. *Nat. Biotechnol.* **40**, 163–166 (2022). doi: 10.1038/s41587-021-01206-w; pmid: 35132262
81. F. Pedregosa *et al*., Scikit-learn: Machine learning in python. *j. mach. learn. res.* **12**, 2825–2830 (2011).
82. V. A. Traag, L. Waltman, N. J. van Eck, From Louvain to Leiden: Guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019). doi: 10.1038/s41598-019-41695-z; pmid: 30914743

83. K. Polański *et al*., BBKNN: Fast batch alignment of single cell transcriptomes. *Bioinformatics* **36**, 964–965 (2020). pmid: 31400197
84. E. Y. Chen *et al*., Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013). doi: 10.1186/1471-2105-14-128; pmid: 23586463
85. C. Ahlmann-Eltze, W. Huber, glmGamPoi: Fitting Gamma-Poisson generalized linear models on single cell count data. *Bioinformatics* **36**, 5701–5702 (2021). doi: 10.1093/bioinformatics/btaa1009; pmid: 33295604
86. J. W. Squair *et al*., Confronting false discoveries in single-cell differential expression. *Nat. Commun.* **12**, 5692 (2021). doi: 10.1038/s41467-021-25960-2; pmid: 34584091
87. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010). doi: 10.1093/bioinformatics/btp616; pmid: 19910308
88. 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; pmid: 32614448
89. E. Stephenson *et al*., Single-cell multi-omics analysis of the immune response in COVID-19. *Nat. Med.* **27**, 904–916 (2021). doi: 10.1038/s41591-021-01329-2; pmid: 33879890
90. N. T. Gupta *et al*., Change-O: A toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics* **31**, 3356–3358 (2015). doi: 10.1093/bioinformatics/btv359; pmid: 26069265
91. C. H. Holland *et al*., Robustness and applicability of transcription factor and pathway analysis tools on single-cell RNA-seq data. *Genome Biol.* **21**, 36 (2020). doi: 10.1186/s13059-020-1949-z; pmid: 32051003
92. M. Efremova, M. Vento-Tormo, S. A. Teichmann, R. Vento-Tormo, CellPhoneDB: Inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020). doi: 10.1038/s41596-020-0292-x; pmid: 32103204
93. L. Garcia-Alonso *et al*., Mapping the temporal and spatial dynamics of the human endometrium in vivo and in vitro. *Nat. Genet.* **53**, 1698–1711 (2021). doi: 10.1038/s41588-021-00972-2; pmid: 34857954
94. M. Setty *et al*., Characterization of cell fate probabilities in single-cell data with Palantir. *Nat. Biotechnol.* **37**, 451–460 (2019). doi: 10.1038/s41587-019-0068-4; pmid: 30899105
95. M. Lange *et al*., CellRank for directed single-cell fate mapping. *Nat. Methods* **19**, 159–170 (2022). pmid: 35027767
96. J.-E. Park, S. Teichmann, M. Haniffa, T. Taghon, Collection of codes and annotated matrix for the paper "A cell atlas of human thymic development defines T cell repertoire formation" (2021), doi: 10.5281/zenodo.5500511
97. G. Palla *et al*., Squidpy: A scalable framework for spatial omics analysis. *Nat. Methods* **19**, 171–178 (2022). doi: 10.1038/s41592-021-01358-2; pmid: 35102346
98. I. Virshup, S. Rybakov, F. J. Theis, P. Angerer, F. A. Wolf, anndata: Annotated data, bioRxiv 473007 [Preprint] (2021); doi: 10.1101/2021.12.16.473007
99. E. Dann, C. Suo, I. Goh, V. Kleshchevnikov, Teichlab/Pan_fetal_immune: Analysis code for publication: Mapping the developing human immune system across organs, Zenodo (2022); https://zenodo.org/record/6481461#.YnLK6drMKUk.

ACKNOWLEDGMENTS

We thank members of the Cooks laboratory (especially A. Montel-Hagen, S. Lopez, and G. Cooks) for their kind help in setting up the ATO experiments; R. Lindeboom, C. Talavera-Lopez, and K. Kanemaru for helpful discussions; and J. Eliasova, A. Garcia, and BioRender.com for graphical illustrations. We gratefully acknowledge the Sanger Flow Cytometry Facility, Newcastle University Flow Cytometry Core Facility, Sanger Cellular Generation and Phenotyping (CGaP) Core Facility, and the Sanger Core Sequencing pipeline for support with sample processing and sequencing library preparation. The human embryonic and fetal material was provided by the MRC-Wellcome Trust–funded Human Developmental Biology Resource (HDBR; http://www.hdbr.org). We are grateful to the donors and donor families for granting access to the tissue samples. This publication is part of the Human Cell Atlas (www.humancellatlas.org/publications). We acknowledge the Wellcome Trust Sanger Institute as the source of HPSI0114i-kolf_2 and HPSI0514i-fiaj_1 human induced pluripotent cell lines, which were generated under the Human Induced Pluripotent Stem Cell Initiative funded by a grant from the Wellcome Trust and the Medical Research Council (MRC), supported by the Wellcome Trust (WT098051) and the NIHR/Wellcome Trust Clinical Research Facility, and we also acknowledge Life Science Technologies Corporation as the provider of Cytotune. **Funding:** This work was supported by the Wellcome Human Cell Atlas Strategic Science Support (grant WT211276/Z/18/Z), CZI Seed Networks for the Human Cell Atlas (Thymus award CZF2019-002445), a MRC Human Cell Atlas award, and the Wellcome Human Developmental Biology Initiative. M.H. is supported by Wellcome (grant WT107931/Z/15/Z), the Lister Institute for Preventive Medicine, NIHR, and the Newcastle Biomedical Research Centre. S.A.T. is supported by Wellcome (grant WT206194 and 108413/A/15/D) and ERC Consolidator Grant ThDEFINE (646794). C.S. is supported by a Wellcome Trust Ph.D. Fellowship for Clinicians. Z.K.T and M.R.C are supported by a MRC Research Project Grant (MR/S035842/1). M.R.C. is supported by an NIHR Research Professorship (RP-2017-08-ST2-002) and a Wellcome Investigator Award (220268/Z/20/Z). **Author contributions:** Conceptualization: S.A.T., M.H., M.R.C., C.S., E.D. Data curation: C.S., E.D., I.G. Formal analysis: E.D., C.S., I.G., L.J., J.E.P., V.K., Z.K.T., K.P., C.X., N.Y., R.E., C.D.C., P.H., C.M., J.C.M. Funding acquisition: S.A.T., M.H. Methodology: C.S., I.G., R.A.B., E.S., J.E., M.M., A.S.S. Project administration: S.A.T., M.H., C.S., E.D., I.G. Software: E.D., K.P., Z.K.T., C.X., M.P., P.M., D.H. Supervision: S.A.T., M.H., M.R.C. Validation: C.S., S.P., N.Y., O.S. Visualization: C.S., E.D., N.Y. Writing – original draft: C.S., E.D., M.H., L.J., I.G., V.K., N.Y., K.P., Z.K.T., S.P. Writing – review and editing: all authors. **Competing interests:** In the past 3 years, S.A.T. has consulted for Genentech and Roche; sits on scientific advisory boards for Qiagen, Foresite Labs, Biogen, and GlaxoSmithKline; and is a cofounder and equity holder of Transition Bio. R.E. is a paid consultant of Foresite Capital. The remaining authors declare no competing interests. **Data and materials availability:** Raw sequencing data for newly generated sequencing libraries have been deposited in ArrayExpress (scRNA-seq libraries: accession no. E-MTAB-11343; scVDJ-seq libraries: accession no. E-MTAB-11388; 10X Genomics Visium libraries: accession no. E-MTAB-11341). Processed data objects are available for online visualization and download in AnnData format (*98*), as well as trained scVI models for query to reference mapping and trained Celltypist models for cell annotation (https://developmental.cellatlas.io/fetal-immune). All code scripts and notebooks for analysis presented in the manuscript are available at Zenodo (*99*) and https://github.com/Teichlab/Pan_fetal_immune.

SUPPLEMENTARY MATERIALS

View/request a protocol for this paper from *Bio-protocol*.

Submitted 12 January 2022; accepted 2 May 2022
Published online 12 May 2022
10.1126/science.abo0510

Mapping the developing human immune system across organs

Chenqu Suo, Emma Dann, Issac Goh, Laura Jardine, Vitalii Kleshchevnikov, Jong-Eun Park, Rachel A. Botting, Emily Stephenson, Justin Engelbert, Zewen Kelvin Tuong, Krzysztof Polanski, Nadav Yayon, Chuan Xu, Ondrej Suchanek, Rasa Elmentaite, Cecilia Domnguez Conde, Peng He, Sophie Pritchard, Mohi Miah, Corina Moldovan, Alexander S. Steemers, Pavel Mazin, Martin Prete, Dave Horsfall, John C. Marioni, Menna R. Clatworthy, Muzlifah Haniffa, and Sarah A. Teichmann

Science, **376** (6597), eabo0510.
DOI: 10.1126/science.abo0510

An expansive view of immunity's start

Although recent single-cell genomics studies have offered profound insights into the developing human immune system, they have not conceptualized the immune system as a distributed network across many tissues. Suo *et al*. integrated single-cell RNA sequencing, antigen-receptor sequencing, and spatial transcriptomics of nine prenatal tissues to reconstruct the immune system's development through time and space. They describe late acquisition of immune effector functions by macrophages and natural killer cells and the maturation of monocytes and T cells before peripheral tissue seeding. Moreover, they describe how blood and immune cell development occurs, not just in primary hematopoietic organs, but across peripheral tissues. Finally, the authors characterize the development of various prenatal innate-like B and T cell populations, including B1 cells. —STS