RESEARCH ARTICLE SUMMARY

HUMAN GENOMICS

A human cell atlas of fetal gene expression

Junyue Cao, Diana R. O'Day, Hannah A. Pliner, Paul D. Kingsley, Mei Deng, Riza M. Daza, Michael A. Zager, Kimberly A. Aldinger, Ronnie Blecher-Gonen, Fan Zhang, Malte Spielmann, James Palis, Dan Doherty, Frank J. Steemers, Ian A. Glass, Cole Trapnell*, Jay Shendure*

INTRODUCTION: A reference atlas of human cell types is a major goal for the field. Here, we set out to generate single-cell atlases of both gene expression (this study) and chromatin accessibility (Domcke *et al.*, this issue) using diverse human tissues obtained during midgestation.

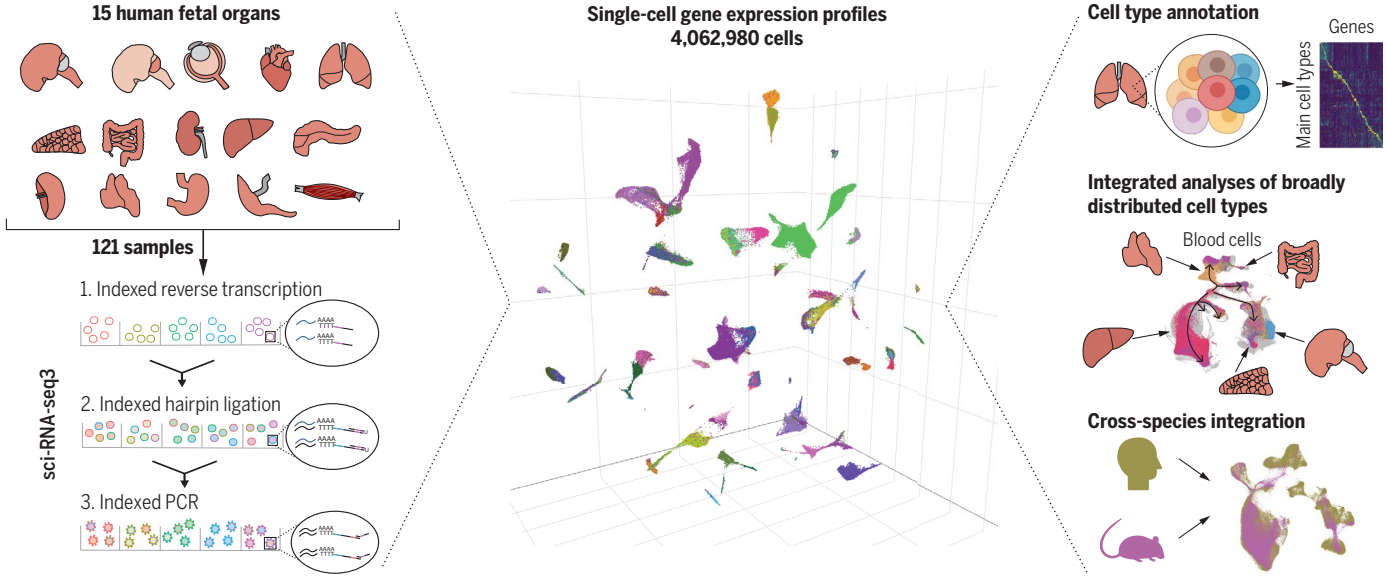
RATIONALE: Contemporary knowledge of the molecular basis of in vivo human development mostly derives from a combination of human genetics, in vivo investigations of model organisms, and in vitro studies of differentiating human cell lines, rather than through direct investigations of developing human tissues. Several challenges have historically limited the study of developing human tissues at the molecular level, including limited access, tissue degradation, and cell type heterogeneity. For this and the companion study (Domcke *et al.*, this issue), we were able to overcome these challenges.

RESULTS: We applied three-level single-cell combinatorial indexing for gene expression (sci-RNA-seq3) to 121 human fetal samples ranging from 72 to 129 days in estimated postconceptual age and representing 15 organs, altogether profiling 4 million single cells. We developed and applied a framework for quantifying cell type specificity, identifying 657 cell subtypes, which we preliminarily annotated based on cross-matching to mouse cell atlases. We identified and validated potentially circulating trophoblast-like and hepatoblast-like cells in unexpected tissues. Profiling gene expression in diverse tissues facilitated the cross-tissue analyses of broadly distributed cell types, including blood, endothelial, and epithelial cells. For blood cells, this yielded a multiorgan map of cell state trajectories from hematopoietic stem cells to all major sublineages. Multiple lines of evidence support the adrenal gland as a normal, albeit minor, site of erythropoiesis during fetal development. It was notably straightforward to integrate these human fetal data with a mouse embryonic cell atlas, despite differences in species and developmental stage. For some systems, this essentially permitted us to bridge gene expression dynamics from the embryonic to the fetal stages of mammalian development.

CONCLUSION: The single-cell data resource presented here is notable for its scale, its focus on human fetal development, the breadth of tissues analyzed, and the parallel generation of gene expression (this study) and chromatin accessibility data (Domcke *et al.*, this issue). We furthermore consolidate the technical framework for individual laboratories to generate and analyze gene expression and chromatin accessibility data from millions of single cells. Looking forward, we envision that the somewhat narrow window of midgestational human development studied here will be complemented by additional atlases of earlier and later time points, as well as similarly comprehensive profiling and integration of data from model organisms. The continued development and application of methods for ascertaining gene expression and chromatin accessibility—in concert with spatial, epigenetic, proteomic, lineage history, and other information—will be necessary to obtain a comprehensive view of the temporal unfolding of human cell type diversity that begins at the single-cell zygote. An interactive website facilitates the exploration of these freely available data by tissue, cell type, or gene (descartes.brotmanbaty.org). ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: coletrap@uw.edu (C.T.); shendure@uw.edu (J.S.)
Cite this article as J. Cao *et al.*, *Science* **370**, eaba7721 (2020). DOI: 10.1126/science.aba7721

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.aba7721



A human cell atlas of fetal gene expression enables the exploration of in vivo gene expression across diverse cell types. We used a three-level combinatorial indexing assay (sci-RNA-seq3) to profile gene expression in ~4,000,000 single cells from 15 fetal organs. This rich resource enables, for example, the identification and annotation of cell types, cross-tissue integration of broadly distributed cell types (e.g., blood, endothelial, and epithelial), and interspecies integration of mouse embryonic and human fetal cell atlases. PCR, polymerase chain reaction.

RESEARCH ARTICLE

HUMAN GENOMICS

A human cell atlas of fetal gene expression

Junyue Cao[1]*, Diana R. O'Day[2], Hannah A. Pliner[3], Paul D. Kingsley[4], Mei Deng[2], Riza M. Daza[1], Michael A. Zager[3,5], Kimberly A. Aldinger[2,6], Ronnie Blecher-Gonen[1], Fan Zhang[7], Malte Spielmann[8,9], James Palis[4], Dan Doherty[2,3,6], Frank J. Steemers[7], Ian A. Glass[2,3,6], Cole Trapnell[1,3,10]†, Jay Shendure[1,3,10,11]†

The gene expression program underlying the specification of human cell types is of fundamental interest. We generated human cell atlases of gene expression and chromatin accessibility in fetal tissues. For gene expression, we applied three-level combinatorial indexing to >110 samples representing 15 organs, ultimately profiling ~4 million single cells. We leveraged the literature and other atlases to identify and annotate hundreds of cell types and subtypes, both within and across tissues. Our analyses focused on organ-specific specializations of broadly distributed cell types (such as blood, endothelial, and epithelial), sites of fetal erythropoiesis (which notably included the adrenal gland), and integration with mouse developmental atlases (such as conserved specification of blood cells). These data represent a rich resource for the exploration of in vivo human gene expression in diverse tissues and cell types.

To date, most investigations of human development have been anatomical or histological in nature (*1*–*3*). However, it is clear that variation in the genetic and molecular programs unfolding within cells during development can cause disease. For example, most Mendelian disorders have a major developmental component (*4*). More-common and often devastating developmental conditions to which genetic factors substantially contribute include congenital heart defects, other birth defects, intellectual disabilities, and autism (*5*).

Several challenges have historically limited the study of developing human tissues at the molecular level. First, access to human embryonic and fetal tissues is limited. Second, even when available, the tissues are usually fixed and nucleic acids are degraded. Third, until recently, most molecular studies of complex tissues have been confounded by cell type heterogeneity. For these reasons, contemporary knowledge of the molecular basis of in vivo human development mostly derives from a combination of human genetics (in particular, of Mendelian disorders), in vivo investigations of model organisms (in particular, of the mouse), and in vitro studies of differentiating human cell lines (in particular, of embryonic or induced pluripotent stem cells), rather than from direct investigations of developing human tissues.

A reference human cell atlas based on developing tissues could serve as the foundation for a systematic effort to better understand the molecular and cellular events that give rise to all rare and common disorders of development, which collectively account for a major proportion of pediatric morbidity and mortality (*6*, *7*). Furthermore, although pioneering cell atlases have already been reported for many adult human organs (*8*, *9*), developing tissues may provide better opportunities to study the in vivo emergence and differentiation of human cell types. Relative to embryonic and fetal tissues, adult tissues are dominated by differentiated cells, and many cell states are not represented. By better resolving cell types and their trajectories, single-cell atlases generated from developing tissues could broadly inform our basic understanding of human biology as well as strategies for cell reprogramming and cell therapy.

As one step toward a comprehensive cell atlas of human development (*10*), we set out to generate single-cell atlases of both gene expression and chromatin accessibility using diverse human tissues obtained during mid-gestation (DESCARTES, Developmental Single Cell Atlas of gene Regulation and Expression; descartes.brotmanbaty.org). For gene expression, we applied three-level single-cell combinatorial indexing (sci-RNA-seq3) (*11*) to 121 fetal tissues representing 15 organs, altogether profiling gene expression in 5 million cells (Fig. 1A and table S1). We also measured chromatin accessibility in 1.6 million cells from the same organs using an overlapping set of samples (*12*). The profiled organs span diverse systems; however, some systems were not accessible—bone marrow, bone, gonads, and skin are notably absent.

Tissues were obtained from 28 fetuses ranging from 72 to 129 days in estimated post-conceptual age. We applied a method for extracting nuclei directly from cryopreserved tissues that works across a variety of tissue types and produces homogenates suitable for both sci-RNA-seq3 and sci-ATAC-seq3 (single-cell combinatorial indexing assay for transposase-accessible chromatin with high-throughput sequencing) (*12*). For most organs, extracted nuclei were fixed with paraformaldehyde. For renal and digestive organs where ribonucleases (RNases) and proteases are abundant, we used fixed cells rather than nuclei, which increased cell and mRNA recovery (*13*). For each experiment, nuclei or cells from a given tissue were deposited to different wells, such that the first index of sci-RNA-seq3 protocol also identified the source. As a batch control for experiments on nuclei, we spiked a mixture of human HEK293T and mouse NIH/3T3 nuclei, or nuclei from a common sentinel tissue (trisomy 21 cerebrum), into one or several wells. As a batch control for experiments on cells, we spiked cells derived from a tissue (pancreas) into one or several wells.

We sequenced sci-RNA-seq3 libraries from seven experiments across seven Illumina NovaSeq 6000 sequencer runs, altogether generating 68.6 billion raw reads. Processing data as previously described (*11*), we recovered 4,979,593 single-cell gene expression profiles [unique molecular identifier (UMI) > 250] [see files S1 to S3 at the Gene Expression Omnibus (GEO) (accession no. GSE156793)]. Single-cell transcriptomes from human-mouse control wells were overwhelmingly species coherent (~5% collisions) (fig. S1A). Uniform manifold approximation and projection (UMAP) (*14*) of nuclei or cells from the sentinel tissues indicated that cell type differences dominated over interexperimental batch effects (fig. S1, B and C). Integrated analysis (*15*) of nuclei and cells from the common pancreatic tissue also resulted in highly overlapping distributions (fig. S1D).

We profiled a median of 72,241 cells or nuclei per organ [Fig. 1A; maximum, 2,005,512 (cerebrum); minimum, 12,611 (thymus)]. Despite shallow sequencing (~14,000 raw reads per cell) relative to other large-scale single-cell RNA sequencing (scRNA-seq) atlases (*16*–*19*), we recovered a comparable number of UMIs per cell or nucleus (median 863 UMIs and 524 genes, not including cultured cells; fig. S1E). As expected, nuclei exhibited a higher proportion of UMIs mapping to introns than cells (56% for nuclei; 45% for cells; $P < 2.2 \times 10^{-16}$, two-sided Wilcoxon rank sum test). We

[1]Department of Genome Sciences, University of Washington School of Medicine, Seattle, WA, USA. [2]Department of Pediatrics, University of Washington School of Medicine, Seattle, WA, USA. [3]Brotman Baty Institute for Precision Medicine, Seattle, WA, USA. [4]Department of Pediatrics, University of Rochester Medical Center, Rochester, NY, USA. [5]Center for Data Visualization, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. [6]Center for Integrative Brain Research, Seattle Children's Research Institute, Seattle, WA, USA. [7]Illumina Inc., San Diego, CA, USA. [8]Human Molecular Genomics Group, Max Planck Institute for Molecular Genetics, Berlin, Germany. [9]Institute of Human Genetics, University of Lübeck, Lübeck, Germany. [10]Allen Discovery Center for Cell Lineage Tracing, Seattle, WA, USA. [11]Howard Hughes Medical Institute, Seattle, WA, USA.
*Present address: Laboratory of Single-Cell Genomics and Population Dynamics, The Rockefeller University, New York, NY, USA.
†Corresponding author. Email: coletrap@uw.edu (C.T.); shendure@uw.edu (J.S.)

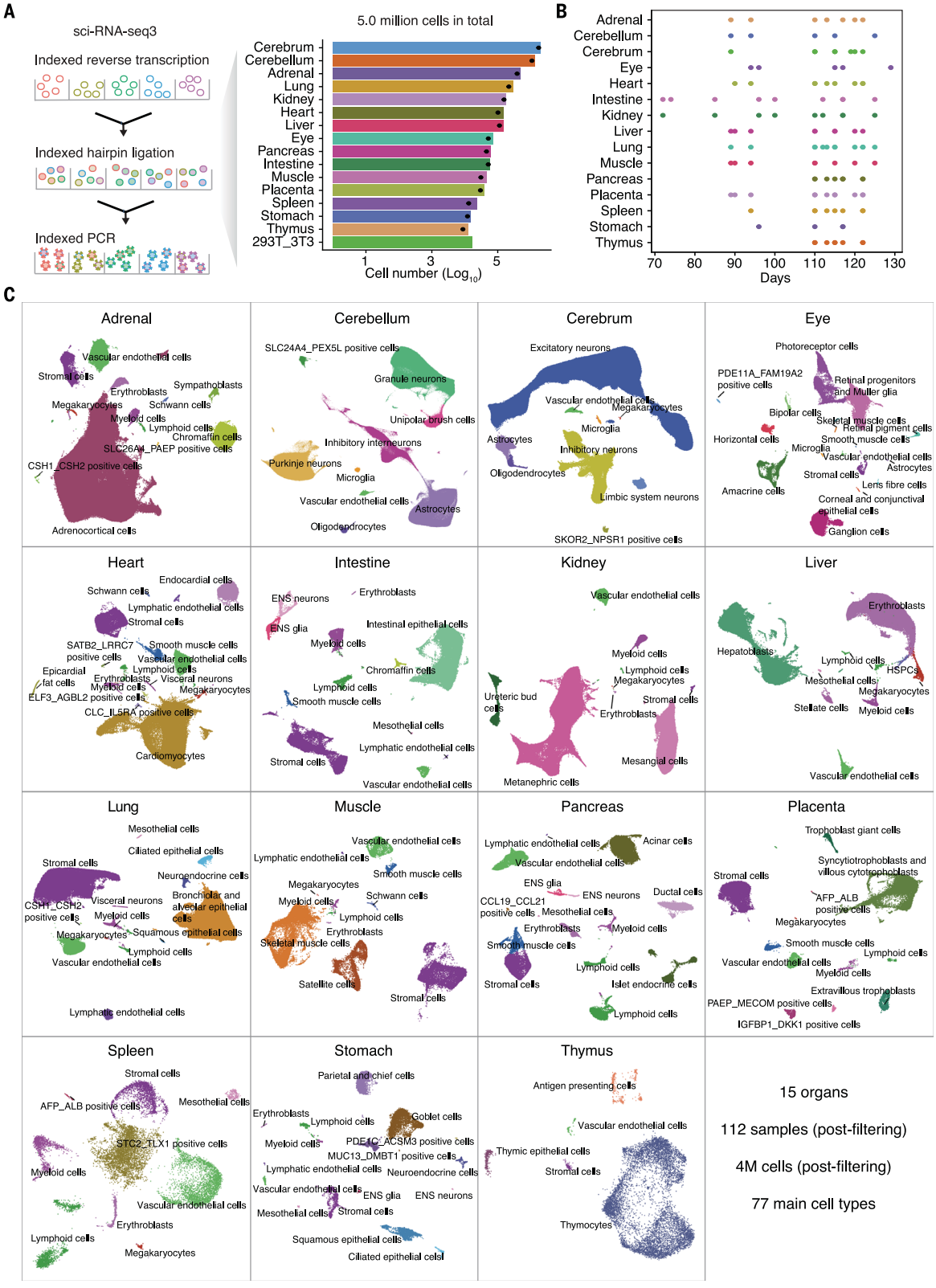


Fig. 1. Data generation and identifying cell types across 15 human organs. (**A**) Project workflow (left) and bar plot (right) showing the number of cells profiled per organ on a $\log_{10}$ scale. Dots indicate the number of cells remaining for downstream analysis after quality control (QC) filtering procedures. PCR, polymerase chain reaction. (**B**) Bar plot showing the distribution of estimated postconceptual ages for tissue samples corresponding to each organ. (**C**) After filtering against low-quality cells and doublet-enriched clusters, 4 million single-cell gene expression profiles were subjected to UMAP visualization and Louvain clustering with Monocle 3 on a per-organ basis. Clusters were initially annotated on a per-organ basis as well, utilizing recent organ-specific cell atlas efforts, which yielded 172 main cell types (colors and labels). Because many cell type annotations appear in multiple organs (e.g., vascular endothelial cells), we consolidated these to 77 main cell types.

henceforth use the word cells to refer to both cells and nuclei, unless otherwise stated.

Tissues were readily identified as deriving from a male (n = 14) or a female (n = 14) by sex-specific gene expression (fig. S1F). Each of the 15 organs was represented by multiple samples (median 8) that included at least two of each sex (fig. S1G) and a range of estimated postconceptual ages (Fig. 1B). Pseudobulk transcriptomes clustered by organ rather than individual or experiment [fig. S1H; files S4 and S5 at GEO (GSE156793)]. About half of the expressed, protein-coding transcripts were differentially expressed across pseudobulk transcriptomes [11,766 of 20,033; false discovery rate (FDR) of 5%; table S2].

We applied Scrublet (*20*) to detect 6.4% likely doublet cells, which corresponded to a doublet estimate of 12.6% including both within-cluster and between-cluster doublets (fig. S1I). We then applied a scalable strategy that we had previously developed (*11*) to remove low-quality cells, doublet-enriched clusters, and the spiked-in HEK293T and NIH/3T3 cells. All analyses below focus on the 4,062,980 human single-cell gene expression profiles derived from 112 fetal tissue samples that remained after this filtering step.

Identification and annotation of 77 main cell types

Using Monocle 3 (*11*), we subjected single-cell gene expression profiles to UMAP visualization and Louvain clustering on a per-organ basis. Altogether, we initially identified and annotated 172 cell types on the basis of cell type–specific marker gene expression (*16*, *21*–*84*) [Fig. 1C and table S3; files S6 and S7 at GEO (GSE156793)]. After collapsing common annotations across tissues, these reduced to 77 main cell types, 54 of which were observed in only a single organ (e.g., Purkinje neurons in the cerebellum) and 23 of which were observed in multiple organs (e.g., vascular endothelial cells in every organ). There were 15 cell types that we were unable to annotate during our manual, organ-by-organ review (the subset named by a pair of markers in Fig. 1C); these are discussed further below and in (*85*). Each of these 77 main cell types was represented by a median of 4829 cells, ranging from 1,258,818 cells (excitatory neurons in the cerebrum) to only 68 cells (*SLC26A4*- and *PAEP*-positive cells in the adrenal gland) (fig. S2A). Each main cell type was observed in multiple individuals (median 9; fig. S2B). We recovered nearly all major cell types identified by previous atlasing efforts directed at the same organs, despite differences with respect to species, stage of development, and technology (*16*, *23*, *28*, *33*, *35*, *51*, *69*, *72*, *86*–*88*). We identified a median of 12 main cell types per organ, ranging from 5 (thymus) to 16 (eye, heart, and stomach). We did not observe a correlation between the number of profiled cells and the number of identified cell types (Spearman $\rho = -0.10$, $P = 0.74$).

On average, we identified 11 marker genes per main cell type (minimum, 0; maximum, 294; defined as differentially expressed genes with at least a fivefold difference between first and second ranked cell type with respect to expression; FDR of 5%; fig. S2C and table S4). There were several cell types that lacked marker genes at this threshold because of highly related cell types in other organs (e.g., enteric glia versus Schwann cells). For this reason, we also report sets of within-tissue marker genes, determined by the same procedure but on an organ-by-organ basis (average 147 markers per cell type; minimum, 12; maximum, 778; fig. S2D and table S5). An interactive website facilitates the exploration of these data by tissue, cell type, or gene (descartes.brotmanbaty.org).

Although canonical markers were generally observed and were critical for our annotation process, to our knowledge, most of the observed markers have not been identified in prior studies. For example, *OLR1*, *SIGLEC10*, and noncoding RNA *RP11-480C22.1* are among the strongest markers of microglia, along with more-established microglial markers such as *CLEC7A* (*89*), *TLR7* (*90*), and *CCL3* (*91*). As anticipated, given that these tissues are undergoing development, many of the 77 main cell types include states progressing from precursors to one or several terminally differentiated cell types. For example, cerebral excitatory neurons exhibited a continuous trajectory from *PAX6*+ neuronal progenitors to *NEUROD6*+ differentiating neurons (*92*) to *SLC17A7*+ mature neurons (*93*) (fig. S2, E and F). In the liver, hepatic progenitors (*DLK1*+, *KRT8*+, and *KRT18*+) (*94*, *95*) exhibited a continuous trajectory to functional hepatoblasts (*SLC22A25*+, *ACSS2*+, and *ASS1*+) (fig. S2, G and H) (*96*–*98*). In contrast with mouse organogenesis—wherein the maturation of the transcriptional program is tightly coupled to developmental time (*11*)—cell state trajectories were inconsistently correlated with estimated postconceptual ages in these data (fig. S2, I and J). A potential explanation for this is that gene expression is markedly more dynamic during embryonic than during fetal development. However, it is also possible that inaccuracies in the estimated postconceptual ages confound our resolution.

In addition to these manual annotations of cell types, we also generated semiautomated classifiers for each organ using Garnett (*99*). The Garnett classifiers were generated agnostic of previous clustering, with marker genes separately compiled from the literature (*99*). Classifications by Garnett were concordant with manual classifications (fig. S3A). Using the Garnett models trained on these data, we were able to accurately classify cell types from other single-cell datasets, including data generated with different methods as well as those from adult organs. When we applied the classifier for pancreas to inDrop scRNA-seq data (*100*), Garnett correctly annotated 82% of the cells (cluster-extended; 11% incorrect, 8% unclassified) (fig. S3B). These models can broadly be used for the automated cell type classification of single-cell data from diverse organs (fig. S3C; descartes.brotmanbaty.org).

We next evaluated the specificity of our main cell types by intradataset cross-validation with a support vector machine (SVM) classifier (*101*). In this framework, high cross-validation precision and recall values indicate that cells derived from a given cluster can robustly be reassigned to that cluster; we thus use high F1 scores as a proxy for identifying cell clusters as valid types, at least in the setting of the tissue in which they were identified. We first evaluated this approach on the kidney. As expected, annotated kidney cell types have much higher specificity scores (median 0.99) than control cell types, in which cell labels are permuted before cross-validation (median 0.17) [Fig. 2A (leftmost panel only), Fig. 2B (left panel only), fig. S4A, and table S3].

We then applied this approach to cells from each organ. Once again, annotated main cell types exhibited much higher specificity scores than permuted cell types (Fig. 2C and fig. S4B; median 0.99 versus 0.10; $P < 2.2 \times 10^{-16}$, two-sided Wilcoxon rank sum test). Despite smaller numbers of cells, most of the 15 initially unannotated cell types also exhibited high specificity scores (median 0.98). The exceptions are probably better described as subtypes of other cell types [discussed further below and in (*85*)]. We also applied this method to the consolidated set of 77 main cell types (i.e., rather than organ-by-organ) with similar results (fig. S4C).

Automated preliminary annotation of cell subtypes

To identify cell subtypes, we performed unsupervised clustering on main cell types with >1000 cells in any given tissue. For each main cell type in each tissue, we first applied batch correction (*102*) followed by dimensionality reduction and Louvain clustering (Fig. 2A). After merging clusters that were not readily distinguishable by the intradataset cross-validation procedure described above, a total of 657 cell subtypes were identified across the 15 tissues, with a median of 824 cells in each. All subtypes were composed of cells contributed by at least two individuals (median 7). Unsurprisingly, given the procedure used for merging clusters, these subtypes have higher specificity scores than permuted controls (median 0.77 versus 0.13; $P < 2.2 \times 10^{-16}$, two-sided Wilcoxon rank sum test; Fig. 2C).

We next sought to leverage existing mouse cell atlases to annotate these human subtypes

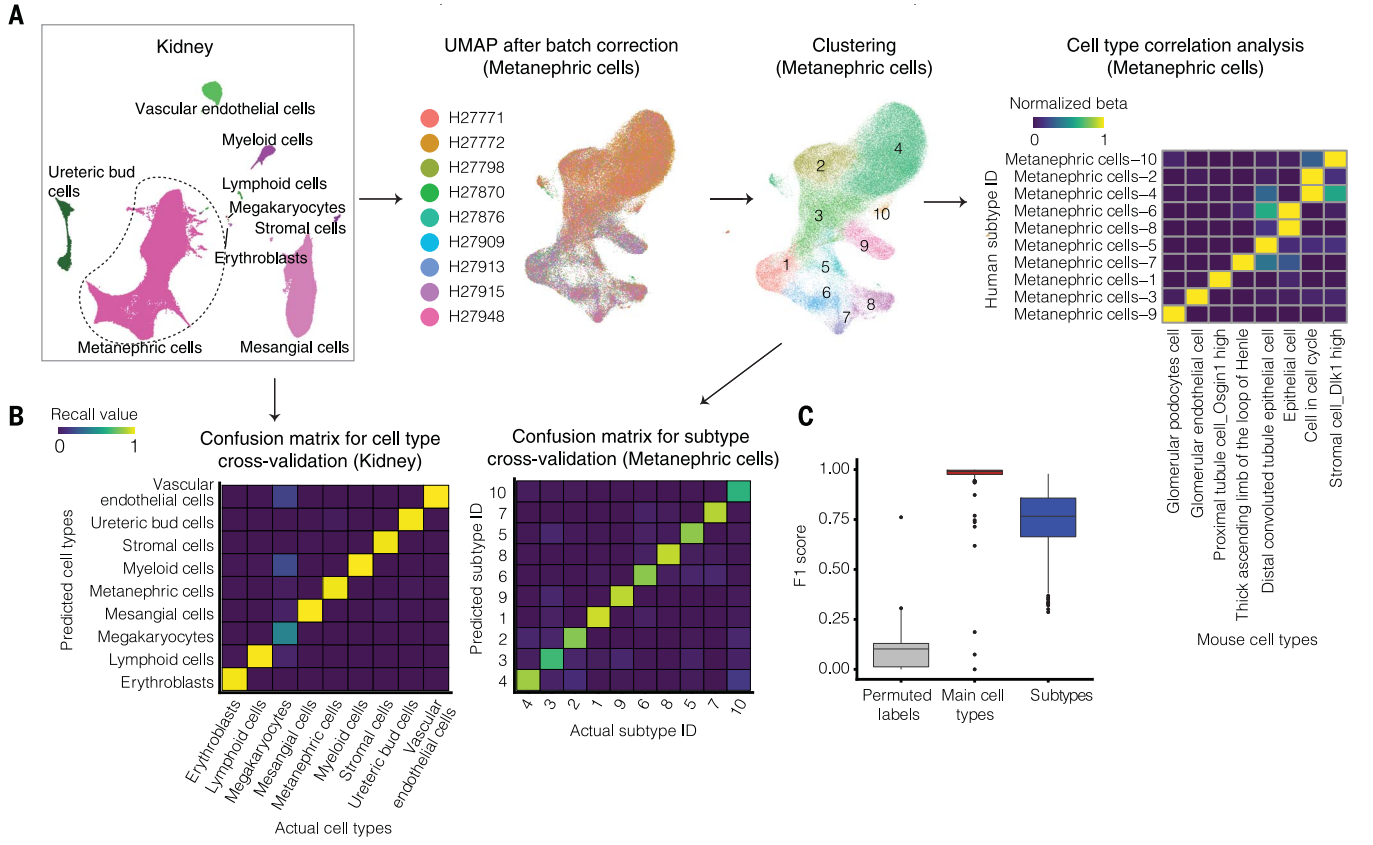


Fig. 2. Identification of cell subtypes. (**A**) Pipeline for cell subtype identification. Briefly, on a tissue-by-tissue basis, we subjected each main cell type with >1000 cells to batch correction (*102*), UMAP visualization, and Louvain clustering. Clusters with similar transcriptomes were merged by an automated procedure. Briefly, we applied an intradataset cross-validation approach (*101*) to evaluate their specificity and iteratively merged similar clusters. We then compared putative human cell subtypes identified in our data (rows) against annotated mouse cell types from the corresponding tissues (*16*) (columns) by cell type correlation analysis. Colors correspond to beta values, normalized by the maximum beta value per row. All MCA cell types with a beta of a matched human cell type >0.01—i.e., also the maximum beta for that human cell type—are shown for the kidney metanephric cells. (**B**) Confusion matrix for intradataset cell type cross-validation with an SVM classifier for main cell types (left) and metanephric subtypes (right) in the kidney. In total, 2000 cells (or all cells for cell types with <2000 cells profiled) are randomly sampled for each cell type or subtype before cross-validation analysis. (**C**) Box plot showing the cell specificity score (F1 score) distribution for permuted controls, main cell types, and subtypes from intradataset cross-validation.

in an automated fashion. With a cell type cross-matching method that we had previously developed (*11*), we could match 605 of 606 (99%) human cell subtypes to at least one cell type in corresponding fetal and/or adult tissues from the mouse cell atlas (MCA) (*16*) (specificity score beta > 0.01, the same threshold that we used to align against MCA previously; 51 adrenal subtypes were excluded because corresponding MCA tissue was not available) (table S6 and figs. S5 to S8). Additionally, 77 of 148 (52%) cerebral or cerebellar subtypes matched to at least one adult cell type from the mouse brain cell atlas (MBCA) (fig. S9) (*50*).

Despite the species difference, many human cell subtypes matched 1:1 with mouse cell types. For example, diverse epithelial subtypes in the human kidney matched 1:1 with annotated MCA cell types (Fig. 2A), and diverse neuronal subtypes in the human cerebrum matched 1:1 with annotated MBCA cell types (fig. S9). Notably, although there were many sets of human subtypes that matched a single MCA or MBCA cell type (e.g., hepatoblasts in fig. S5 and oligodendrocytes in fig. S9), these likely reflect bonafide heterogeneity as evidenced by their specificity scores (Fig. 2C). Additional work is necessary to annotate such subtypes with greater granularity.

Integration across tissues and investigation of initially unannotated cell types

We next sought to integrate data and compare cell types across all 15 organs. To mitigate the effects of gross differences in sampling, we randomly sampled 5000 cells per cell type per organ (or in cases where <5000 cells of a given cell type were represented in a given organ, all cells were taken), and we performed UMAP visualization (Fig. 3A and fig. S10A). As expected, cell types represented in multiple organs, as well as developmentally related cell types, tended to colocalize. Many surface proteins (4565 of 5480), secreted proteins (2491 of 2933), transcription factors (1715 of 1984), and noncoding RNAs (3130 of 10,695) were differentially expressed across the 77 main cell types (FDR of 0.05; Fig. 3B and table S4; descartes.brotmanbaty.org). The expression patterns of noncoding RNAs were notably sufficient to separate cell types into developmentally coherent groups (fig. S10, B and C).

As noted above, there were 15 cell types that we were unable to annotate during our manual, organ-by-organ review (the subset named by pairs of markers in Fig. 1C). To shed light on these, we examined their distribution in the global UMAP (Fig. 3A), whether they matched to annotated cell types in MCA or MBCA (figs. S5 to S9), their distribution across tissues derived from different individuals (fig. S11A), and their potential for maternal origin (fig. S11B).

These further analyses enabled us to annotate 8 of the 15 cell types (*85*). For example, rare *CSH1*- and *CSH2*-positive cells in the lung and adrenal gland (two of the most deeply profiled organs)

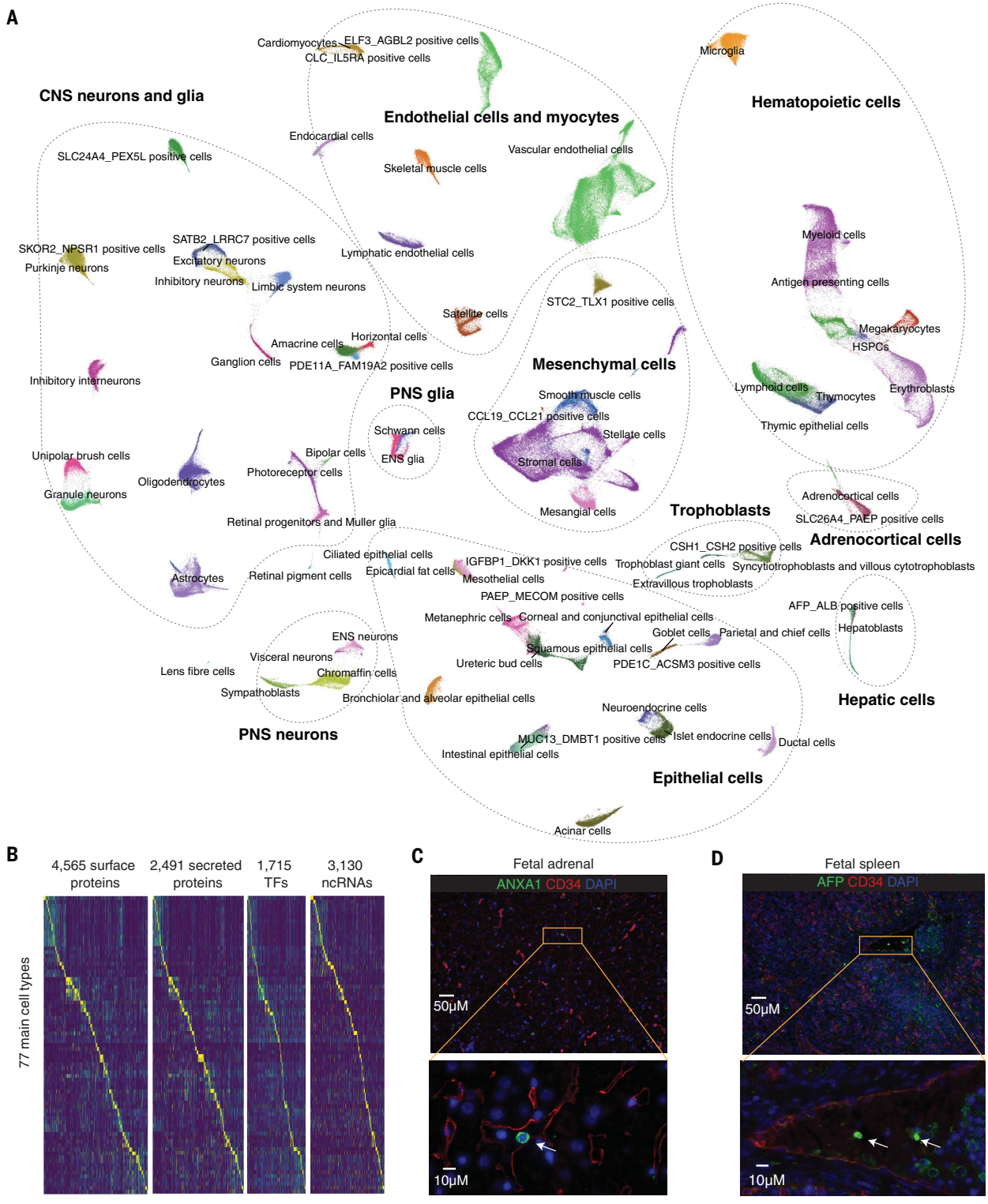


Fig. 3. Integrated visualization of cell types across all profiled tissues. (**A**) From each organ, we sampled 5000 cells from each cell type (or all cells for cell types with <5000 cells in a given organ). These were subjected to UMAP visualization on the basis of the top differentially expressed genes across cell types within each organ. Here, they are colored by cell type labels, with colors as in Fig. 1C. In fig. S10A, the same UMAP visualization is colored by tissue of origin. (**B**) Heatmap showing the relative expression of surface and secreted protein-coding genes, noncoding RNAs (ncRNAs), and TFs (columns) in 77 main cell types (rows). UMI counts for genes are scaled for library size, log-transformed, and then mapped to *Z* scores and capped to [0, 3]. (**C** and **D**) Representative fluorescence microscopy images of (C) human fetal adrenal or (D) spleen tissue, staining for (C) endothelium (CD34+), *CSH1*-, and *CSH2*-positive cells (ANXA1+; labeled by arrowhead) or (D) *AFP*- and *ALB*-positive cells (AFP+ is indicated with arrows). Nuclei are stained with blue 4′,6-diamidino-2-phenylindole (DAPI). Bottom panels correspond to inset zooms. Scale bars, 50 μm (top) and 10 μm (bottom).

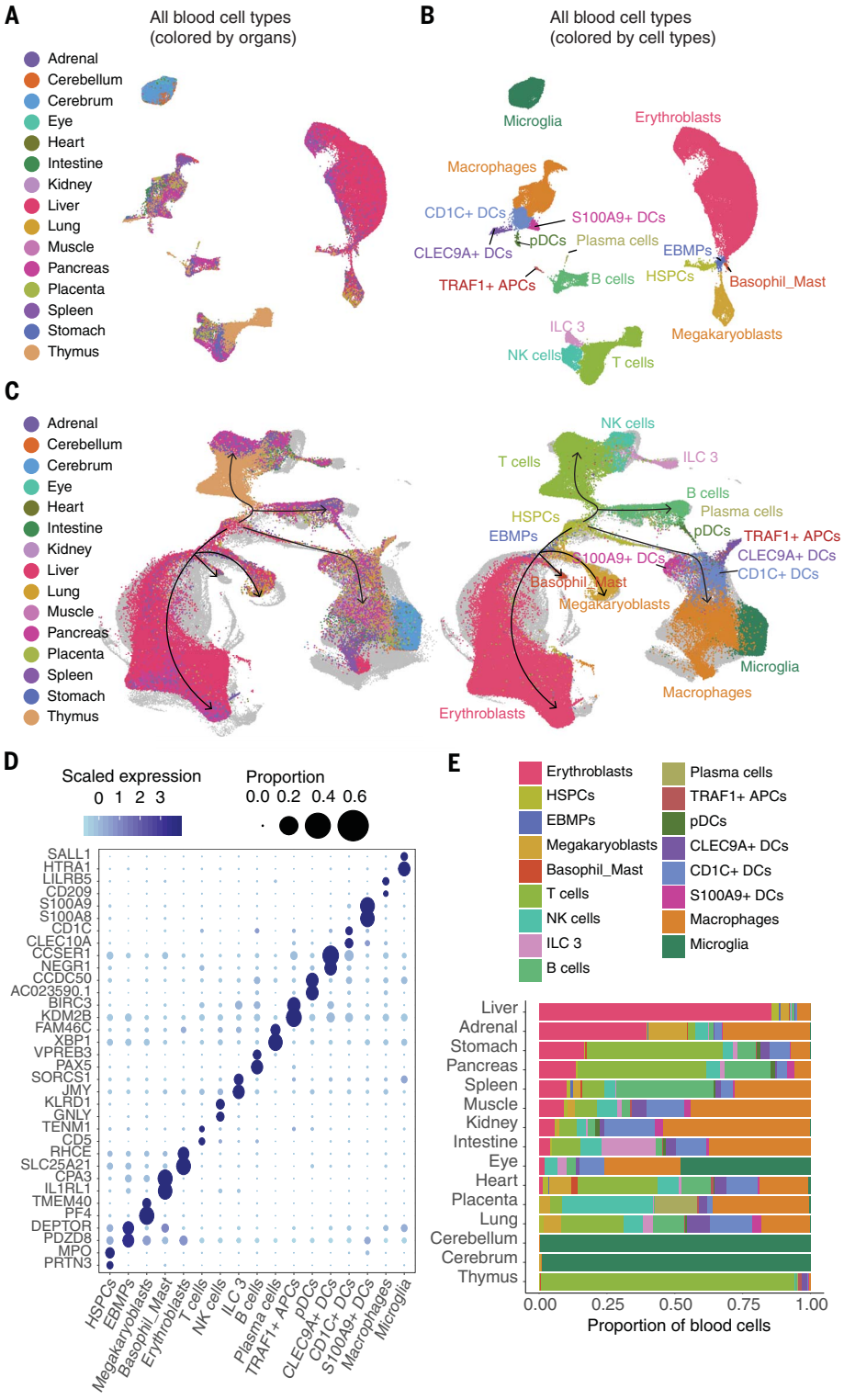


Fig. 4. Identification and characterization of blood cell subtypes and developmental trajectories. (**A** and **B**) UMAP visualization and marker-based annotation of blood cell types colored by organ type (A) and cell type (B). (**C**) UMAP visualization of blood cells, integrating across all profiled organs of this study and an scRNA-seq atlas of blood cells from human fetal liver (*108*). Cells from (*108*) are colored in light gray, and cells from our study are colored by tissue of origin (left) or blood cell types (right). Black arrows indicate inferred cell state transition directions from HSPCs to all main blood lineages. (**D**) Dot plot showing expression of two selected marker genes per cell type. The size of the dot encodes the percentage of cells within a cell type in which that marker was detected, and its color encodes the average expression level. (**E**) Bar plot showing the estimated fraction of cells per organ derived from each of the 17 annotated blood cell types.

are highly similar to placental trophoblasts—e.g., expressing high levels of placental lactogen, chorionic gonadotropin, and aromatase (Fig. 3A) (*85*). *AFP*- and *ALB*-positive cells in the placenta and spleen resemble hepatoblasts—e.g., expressing high levels of serum albumin, alpha fetoprotein, and apolipoproteins (Fig. 3A) (*85*) [at least in the placenta, similar hepatoblast-like *AFP*- and *ALB*-positive cells were observed in the mouse (fig. S5)]. Follow-up immunostaining studies supported the presence of these trophoblast-like and hepatoblast-like cells in the adrenal gland and spleen, respectively (Fig. 3, C and D, and fig. S12). Given that these cell types are rarely but recurrently observed in several organs, they potentially correspond to circulating trophoblasts and circulating hepatoblasts.

In males, both *IGFBP1*- and *DKK1*-positive as well as *PAEP*- and *MECOM*-positive cells in the placenta expressed appreciable levels of *XIST* or *TSIX* (fig. S12B); on further review of markers, these correspond to maternal decidualized stromal cells and maternal endometrial epithelial cells, respectively. This conclusion is supported by maternal genotypes in the corresponding cell types in chromatin accessibility data (*12*).

Several additional cell types were annotated through strong matches to MCA or MBCA (fig. S13) or through their position in the global UMAP coupled with additional literature review (Fig. 3A) (*85*); these include *STC2*- and *TLX1*-positive cells, which are abundant in the spleen and express genes associated with mesenchymal precursor or stem cells (*103*–*105*). Of the remaining seven initially unannotated cell types, four would likely better be classified as subtypes (and correspondingly, these tended to have lower specificity scores), and three have high specificity scores but remain ambiguous (*85*).

Characterization of blood lineage development across organs

The nature of this dataset creates an opportunity to systematically investigate organ-specific differences in gene expression within broadly distributed cell types—for example, blood cells. We reclustered 103,766 cells, derived from all 15 organs, that corresponded to hematopoietic cell types (Fig. 4A). We then performed Louvain clustering and further annotated fine-grained blood cell types, in some cases identifying very rare cell types (Fig. 4B). For example, myeloid cells separate into microglia, macrophages, and diverse dendritic cell subtypes [*CD1C*+, *S100A9*+, *CLEC9A*+, and plasmacytoid dendritic cells (pDCs)] (*106*). The microglial cluster primarily derives from brain tissues, and it is well separated from macrophages, which is consistent with their distinct developmental trajectories (*107*). Lymphoid cells clustered into several groups, including B cells, natural killer (NK) cells, ILC 3 cells, and T cells, the latter of which

includes the thymopoiesis trajectory. We also recovered very rare cell types such as plasma cells (139 cells, mostly in placenta and making up 0.1% of all blood cells or 0.003% of the full dataset) and *TRAF1*+ antigen-presenting cells (APCs) (189 cells, mostly in thymus and heart and making up 0.2% of all blood cells or 0.005% of the full dataset).

To validate these annotations, we integrated fetal blood cells from all organs with an scRNA-seq atlas of blood cells from the fetal liver (*108*) (Fig. 4C, left, and fig. S14A). Despite different methods, corresponding cell types from two datasets were highly overlapping; this was also the case upon integration analysis with another scRNA-seq dataset of 1231 human embryonic blood cells (*109*) (fig. S14B). Notably, some extremely rare cell types identified through CD45+ fluorescence-activated cell sorting (FACS) enrichment (e.g., *VCAM1*+ EI macrophages, monocyte precursors, and neutrophil-myeloid progenitors) were not annotated in our data. On the other hand, we captured fetal blood cells derived from tissues other than the liver—e.g., microglia in the brain and T and B cells in the thymus and spleen, respectively. Furthermore, as they span multiple organs, we are better able to capture cell state transition paths from hematopoietic stem and progenitor cells (HSPCs) to lymphoid cells than a single-organ study (Fig. 4C, right).

Although gene expression markers for different immune cell types have been extensively studied, these may be limited by their definition via a restricted set of organs or cell types. Here, we find that many conventional immune cell markers were expressed in multiple cell types. For example, conventional markers for T cells (*110*–*112*) were also expressed in macrophages and dendritic cells (*CD4*) or NK cells (*CD8A*), consistent with other studies (*113*) (fig. S14C). We computed pan-organ cell type–specific markers across 14 blood cell types (Fig. 4D and table S7). From this we observed that T cells specifically expressed *CD8B* and *CD5* (*114*) as expected, but also *TENM1* (Fig. 4D and fig. S14C). ILC 3 cells, whose annotation was determined on the basis of their expression of *RORC* (*115*) and *KIT* (*116*), were more specifically marked by *SORCS1* and *JMY* (Fig. 4D and fig. S14C). These and other markers identified by pan-organ analysis may be useful for labeling and purifying specific blood cell types.

As expected, different organs showed varying proportions of blood cells (Fig. 4E). For example, the liver contained the highest proportion of erythroblasts, consistent with its role as the primary site of fetal erythropoiesis (*117*), whereas T cells were enriched in the thymus and B cells in the spleen. Nearly all blood cells recovered from the cerebellum and cerebrum were microglia. The tissue distribution of ILC 3 cells as well as subtypes of dendritic cells was captured as well (Fig. 4E and fig. S14D). Pan-organ analysis also enabled the identification of rare cell populations in specific organs. We identified rare HSPCs in the liver but also rare cells that are transcriptionally similar to HSPCs in the lung, spleen, thymus, heart, intestine, adrenal gland, and other organs (fig. S15). Subclustering analyses showed that HSPCs outside of the liver, as well as a subset of liver HSPCs, expressed differentiation markers such as *LYZ* (*118*), *ACTG1* (*119*), and *ANK1* (*120*), whereas most liver HSPCs expressed *MECOM* and *NRIP1*, both of which are required for the maintenance and function of normal quiescent HSPCs (*121*, *122*) (fig. S15).

Focusing on erythropoiesis, we observed a continuous trajectory from HSPCs to an intermediate cell type, erythroid-basophil-megakaryocyte biased progenitors (EBMPs), which then split into erythroid, basophilic, and megakaryocytic trajectories (Fig. 5A and table S8), consistent with a recent study of the mouse fetal liver (*123*, *124*). This consistency was despite differences in species (human versus mouse), techniques (sci-RNA-seq3 versus 10x Genomics), and tissues (pan-organ versus liver only). With unsupervised clustering and adopting terminology from that study (*123*), we further partitioned the continuum of erythroid states into three stages: early erythroid progenitors (EEPs) (marked by *SLC16A9* and *FAM178B*), committed erythroid progenitors (CEPs) (marked by *KIF18B* and *KIF15*), and cells in the erythroid terminal differentiation state (ETDs) (marked by *TMCC2* and *HBB*) (Fig. 5B). Early and late stages of megakaryocytic cells were also readily identified (Fig. 5, A and B).

As expected, given their established role in fetal erythropoiesis, a portion of blood cells in the liver and spleen corresponded to EEPs, CEPs, and megakaryocyte progenitors (*125*). Notably, we also observed EEPs, CEPs, and megakaryocyte progenitors in the adrenal gland in every sample studied (Fig. 5C and fig. S16A). Because we do not observe cell types that are more common in the liver and spleen, trivial contamination during recovery of the adrenal glands is an unlikely explanation. Although occasional islands of extramedullary hematopoiesis have been observed in the adrenal glands of human embryos (*126*, *127*), the consistency across individuals led us to further investigate whether the adrenal glands may serve as a normal site of erythropoiesis in mammals. Immunohistochemical analysis of human fetal adrenal tissues showed nucleated GYPA+ cells outside CD34+ blood vessels (Fig. 5D and fig. S16B). We further used imaging flow cytometry to visualize and enumerate maturing erythroid precursors and enucleated erythrocytes (*128*) in the perinatal period of the mouse. Approximately 8% of viable dissociated cells from the adrenal gland consisted of maturing erythroblasts, compared with 0.2% of viable dissociated cells in the kidney (Fig. 5E). Also consistent with the adrenal gland being a site of ongoing erythropoiesis, its distribution of immature to mature erythroblasts matched closely with that of the bone marrow of adult mice (Fig. 5, E and F).

Macrophages were even more widely distributed. We collated all macrophages, together with microglia from the brain, and subjected them to UMAP visualization and Louvain clustering, independent of other cell types (Fig. 5, G and H; fig. S16C; and table S9). Notably, microglia were divided into three subclusters, one of which, marked by *IL1B* and *TNFRSF10D*, likely represents activated microglia expressing proinflammatory cytokines involved in the normal development of the nervous system (*129*, *130*). The other microglial clusters were marked by expression of *TMEM119* and *CX3CR1* (*131*) (more common in the cerebrum) or *PTPRG* and *CDC14B* (*132*) (more common in the cerebellum).

The macrophages outside the brain clustered into three major groups (Fig. 5, G and H; fig. S16C; and table S9): (i) antigen-presenting macrophages, found mostly in gastrointestinal (GI) tract organs (intestine and stomach) and marked by high expression of antigen-presenting (e.g., *HLA-DPB1* and *HLA-DQA1*) and inflammatory activation genes [e.g., *AHR* (*133*)]; (ii) perivascular macrophages, found in most organs, with specific expression of markers such as *F13A1* (*134*) and *COLEC12* (*135*), as well as markers such as *RNASE1* and *LYVE1*; and (iii) phagocytic macrophages, enriched in the liver, spleen, and adrenal gland (Fig. 5I), with specific expression of markers such as *CD5L* (*136*), *TIMD4* (*137*), and *VCAM1* (*138*). Phagocytic macrophages are critical for removing the pyrenocytes (the so-called extruded nucleus) after enucleation of late-stage erythroblasts to form reticulocytes; their observation in the adrenal gland is consistent with its aforementioned potential role as an additional site of normal fetal erythropoiesis. Below, we leverage integration with a mouse atlas of organogenesis (*11*) to investigate the conserved program of blood cell specification and developmental origins of microglia and macrophages.

Characterization of endothelial and epithelial cells across organs

As a second analysis of a single class of cells across many organs, we reclustered 89,291 endothelial cells (ECs) that correspond to vascular endothelium (VECs), lymphatic endothelium (LECs), or endocardium. These three groups readily separated from one another, and VECs further clustered, at least to some degree, by organ (fig. S17, A to C). That organ-specific differences are more readily detected than differences between arteries, capillaries, and veins is consistent with previous cell atlases of the adult mouse (*16*, *28*). We performed

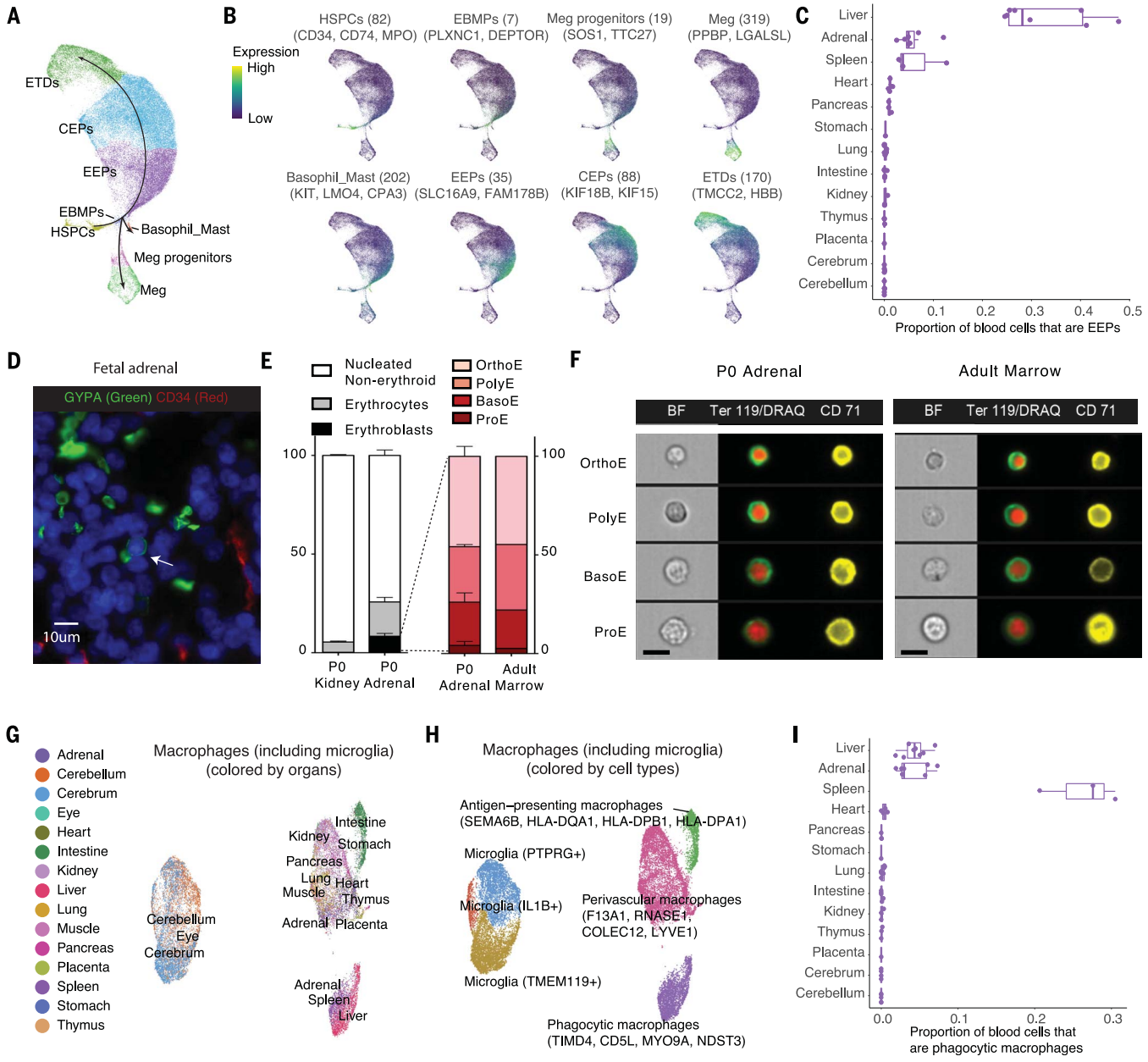


Fig. 5. Identification and characterization of erythropoiesis and macrophage differentiation in adrenal gland. (**A**) Zoomed view of the erythropoiesis trajectory portion of Fig. 4B, colored by erythroid or megakaryocyte subtype. Black arrows show trajectory directionalities defined by (*123*). (**B**) Plots similar to (A), colored by the normalized expression of cell type–specific genes (FDR of 0.05 and more than twofold expression difference between first and second ranked cell type), with the number of cell type–specific genes used and names of the top few genes shown. UMI counts for these genes are scaled for library size, log-transformed, aggregated, and then mapped to *Z* scores. (**C**) Point and box plot showing the proportion of blood cells that are EEPs for individual samples of different organs. Samples with low recovery of blood cells (≤200) are excluded. (**D**) Representative fluorescence microscopy of human fetal adrenal tissue, staining for endothelium (CD34+) and erythroblasts (nucleated and GYPA+); nuclei stained with blue DAPI. The arrow indicates a GYPA+ erythroblast outside a CD34+ blood vessel. Scale bars, 10 μm. (**E**) (Left) Percentage of dissociated kidney and adrenal glands from newborn (P0) mice composed of enucleated erythrocytes and maturing erythroblasts. (Right) Distribution of maturing erythroblasts (proerythroblasts, ProE; basophilic erythroblasts, BasoE; polychromatophilic erythroblasts, PolyE; and orthochromatic erythroblasts, OrthoE) in the adrenal gland at P0 and in adult bone marrow. Error bars represent means + SEM, $n = 3$. (**F**) Representative images of maturing erythroblasts in the P0 adrenal gland and the adult bone marrow. Scale bars, 10 μm. (**G** and **H**) UMAP visualization and marker-based annotation of macrophage subtypes colored by organ type (G) and subtype name (H). (**I**) Point and box plot showing the proportion of blood cells that are phagocytic macrophages for individual samples of different organs. Samples with low recovery of blood cells (≤200) are excluded.

an integrative analysis of ECs from human fetal tissues (this study) and mouse adult tissues (*139*) (fig. S17, D and E). Both human and mouse ECs were separated first by vascular versus lymphatic versus endocardial, and then by organ. VECs from the same tissue were generally clustered together, despite differences with respect to species, developmental stage, and technique. Conserved markers of organ-specific ECs were readily identified (fig. S17F) (*139*).

Differential gene expression analysis identified 700 markers that are specifically expressed

in a subset of ECs (FDR of 0.05 and more than twofold expression difference between first and second ranked cluster) (fig. S17G and table S10). About one-third of these encoded membrane proteins, many of which appeared to correspond to potential specialized functions (*12*, *140*–*142*). In agreement with observations in mice (*139*), brain ECs specifically expressed gene sets involved in amino acid transport ($q = 5.6 \times 10^{-10}$) and carboxylic acid transport ($q = 4.2 \times 10^{-8}$); lung ECs specifically expressed gene sets involved in adenosine 3′,5′-monophosphate (cAMP) ($q = 8.2 \times 10^{-3}$) and cyclic nucleotide ($q = 1.4 \times 10^{-2}$) catabolism, and vascular ECs from the GI tract, heart, and muscle specifically expressed gene sets involved in stem cell differentiation ($q = 3.7 \times 10^{-2}$). Potentially underlying these differences, human fetal ECs expressed distinct sets of transcription factors (TFs) (fig. S17H). For example, LECs specifically expressed *TBX1*, brain VECs specifically expressed *FOXQ1* and *FOXF2*, and liver VECs specifically expressed *DAB2*, all of which are consistent with observations in mice (*139*, *143*, *144*).

As a third analysis of a broadly distributed type of cell, we reclustered 282,262 epithelial cells, derived from all organs, and subjected these to UMAP visualization (fig. S18, A and B). Although some epithelial cell types were highly organ specific—e.g., acinar (pancreas) and alveolar cells (lung)—epithelial cells with similar functions generally clustered together (fig. S18C).

Within epithelial cells, two neuroendocrine cell clusters were identified (fig. S18C). The simpler of these corresponded to adrenal chromaffin cells and was marked by the specific expression of *HMX1* (*NKX-5-3*), a TF involved in sympathetic neuron diversification (*145*). The other cluster comprised neuroendocrine cells from multiple organs (stomach, intestine, pancreas, and lung) and was marked by specific expression of *NKX2-2*, a TF with a key role in pancreatic islet and enteroendocrine differentiation (*146*). We performed further analysis on the latter group, identifying five subsets (fig. S18, D to F): (i) pancreatic islet beta cells, marked by insulin expression; (ii) pancreatic islet alpha and gamma cells, marked by pancreatic polypeptide and glucagon expression; (iii) pancreatic islet delta cells, marked by somatostatin expression; (iv) pulmonary neuroendocrine cells (PNECs), marked by expression of *ASCL1* and *NKX2-1*, both TFs with key roles in specifying this lineage in the lung (*147*, *148*); and (v) enteroendocrine cells. Enteroendocrine cells further comprised several subsets, including *NEUROG*-expressing pancreatic islet epsilon progenitors (*149*, *150*), *TPH1*-expressing enterochromaffin cells in both the stomach and intestine (*151*), and gastrin- or cholecystokinin-expressing G, L, K, and I cells (*151*). Finally, we observed ghrelin-expressing enteroendocrine progenitors in the stomach and intestine (*150*, *152*), but also ghrelin-expressing endocrine cells in the developing lung (*153*) (fig. S18F). The diverse functions of neuroendocrine cells are closely linked with their secreted proteins; we identified 1086 secreted protein-coding genes differentially expressed across neuroendocrine cells (FDR of 0.05) (fig. S18G and table S11). For example, PNECs showed specific expression of trefoil factor 3, which is involved in mucosal protection and lung ciliated cell differentiation (*154*); gastrin-releasing peptide, which stimulates gastrin release from G cells in the stomach (*155*); and *SCGB3A2*, a surfactant associated with lung development (*156*).

As an illustrative example of how these data can be used to explore cell trajectories, we further investigated the path of epithelial cell diversification leading to renal tubule cells. Combining and reclustering ureteric bud and metanephric cells, we identified both progenitor and terminal renal epithelial cell types, with differentiation paths that are highly consistent with a recent study of the human fetal kidney (*157*) (fig. S19A). By differential gene expression analysis, we further identified TFs potentially regulating their specification (fig. S19B and table S12). For example, nephron progenitors in the metanephric trajectory specifically expressed high levels of mesenchyme and meis homeobox genes (*MEOX1*, *MEIS1*, and *MEIS2*) (*158*), whereas podocytes specifically expressed *MAFB* and *TCF21/POD1* (*159*, *160*). As another example, *HNF4A* was specifically expressed in proximal tubule cells—a mutation of this gene causes Fanconi renotubular syndrome, a disease that specifically affects the proximal tubule—and *HNF4A* was recently shown to be required for formation of the proximal tubule in mice (*161*).

Integration of human and mouse developmental atlases

The transition from embryonic to fetal development is of considerable interest, but access to human embryonic tissues is even more limited than access to fetal tissues. To again leverage the mouse, we sought to integrate these human fetal data with a mouse organogenesis cell atlas (MOCA), for which we had previously profiled 2 million cells from undissected embryos spanning E9.5 to E13.5 (*11*). For context, this window corresponds to days 22 to 44 of human development (*162*, *163*), whereas the tissues studied here are estimated to derive from days 72 to 129.

First, we compared the 77 main cell types defined here against the developmental trajectories of organogenesis defined by MOCA by means of a cell type cross-matching method (*11*). Most human cell types strongly matched to a single major mouse trajectory and subtrajectory (fig. S20 and tables S13 and S14). These generally corresponded to expectation, although a few discrepancies facilitated corrections to MOCA (see legends of figs. S20 and S21). Many human cell types and mouse trajectories that lacked strong 1:1 matches [summed nonnegative least squares (NNLS) regression coefficients < 0.6] corresponded to tissues excluded in the other dataset (e.g., mouse placenta and human skin and gonads). Other ambiguities probably follow from the gap between the developmental windows studied (e.g., adrenal cell types), rarity (e.g., bipolar cells), and/or complex developmental relationships (e.g., fetal cell types that derive from multiple embryonic trajectories).

Second, we sought to directly coembed human and mouse cells together. In brief, we sampled 100,000 mouse embryonic cells from MOCA (randomly) and ~65,000 human fetal cells (maximum 1000 cells from each of 77 cell types) and subjected these to integrated analysis (*15*). The distribution of mouse cells in the resulting UMAP visualization was similar to our global analysis of MOCA (Fig. 6, A to C, and figs. S21 to S23) (*11*). Furthermore, despite the species difference, human fetal cells were overwhelmingly distributed in a manner that respected developmental relationships between cell types. For example, human fetal endothelial, hematopoietic, hepatic, epithelial, and mesenchymal cells all mapped to the corresponding mouse embryonic trajectories (Fig. 6B and fig. S21). Within each major trajectory, mouse cells order by successive time point (*11*), whereas human fetal cells appear to project from the last (E13.5) mouse embryonic time point (Fig. 6C). At the subtrajectory level, seniscal mappings include human fetal intestinal epithelial cells emanating from the mouse midgut-hindgut subtrajectory; human fetal parietal and chief cells (stomach) and acinar and ductal cells (pancreas) emanating from the mouse foregut epithelial subtrajectory; human fetal bronchiolar and alveolar epithelial cells emanating from the mouse lung epithelial trajectory; human fetal ureteric bud and metanephric cells emerging separately from the mouse embryonic renal epithelial trajectory; and many others (figs. S21 to S23).

However, there were also a few surprises. For example, although central nervous system (CNS) neurons mapped to the neural tube trajectory and enteric nervous system (ENS) glia and Schwann cells mapped to peripheral nervous system (PNS) glial trajectories, some neural crest derivatives—including ENS neurons, visceral neurons, sympathoblasts, and chromaffin cells—clustered separately from the corresponding mouse embryonic trajectories (figs. S21 to S23), potentially because of excessive differences between the developmental stages or between the species. Human fetal astrocytes clustered with the mouse embryonic neural epithelial trajectory [mouse astrocytes

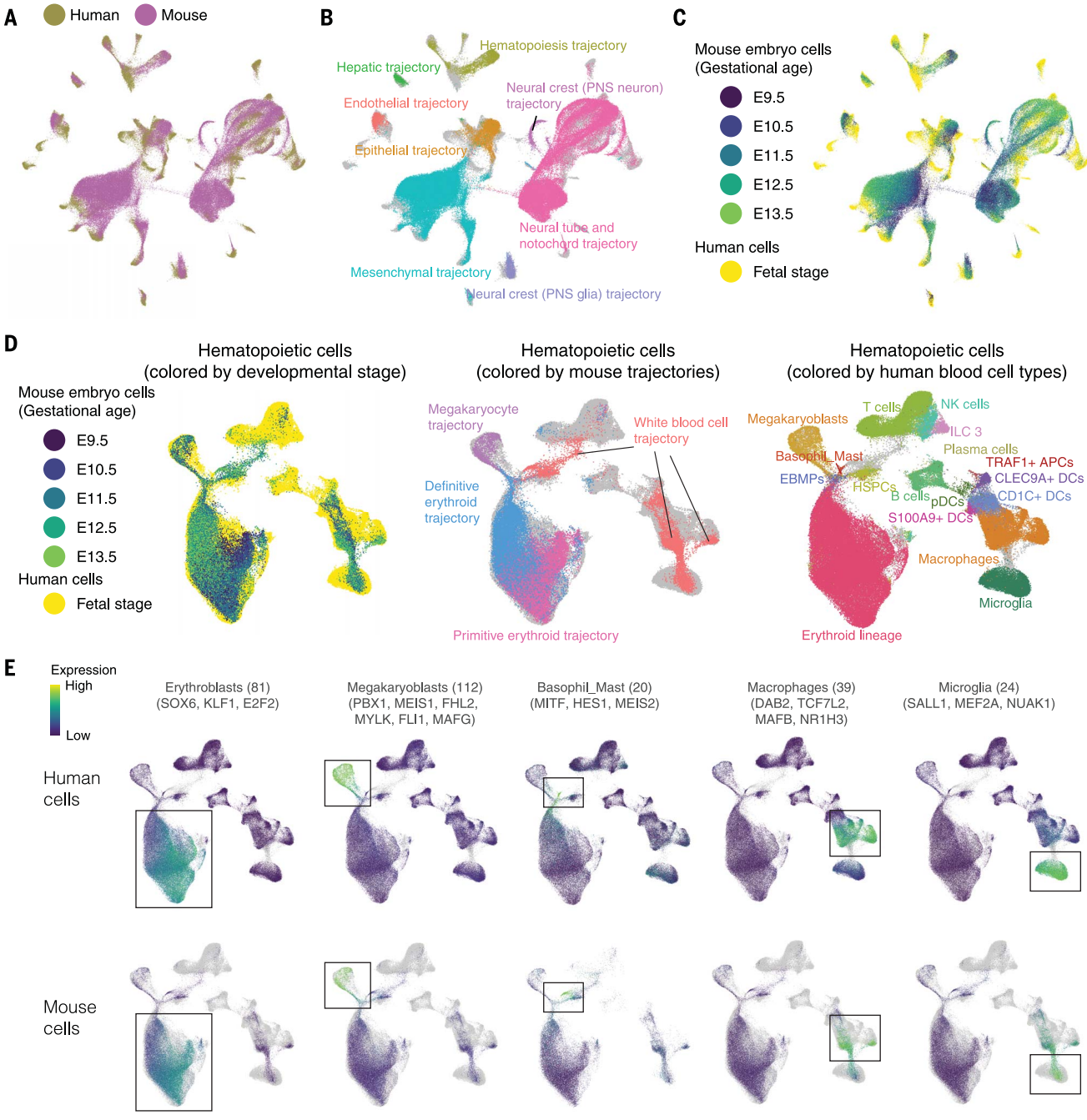


Fig. 6. Integration of human fetal and mouse embryonic cell atlases. (**A** to **C**) After downsampling as described in the text, we applied Seurat (*15*) to jointly analyze human fetal and mouse embryonic cells (*11*). (A) Cells are colored by source species. (B) Mouse cells are colored by the identity of the main mouse embryonic trajectory (*11*). Human cells are colored in gray. (C) Cells are colored by source and development stage. Within each major trajectory and as has been shown previously (*11*), mouse cells order by successive time points, and human fetal cells appear to project from the last (E13.5) mouse embryonic time point. (**D**) We applied Seurat (*15*) to jointly analyze 103,766 human and 40,606 mouse hematopoietic cells. The same UMAP visualization is shown in all panels. (Left) Cells are colored by source and development stage. (Middle) Mouse cells are colored by the identity of mouse subtrajectory (*11*). Human cells are colored in gray. (Right) Human cells are colored according to annotations from Fig. 4B. Mouse cells are colored in gray. (**E**) Plot similar to (D), colored by the normalized expression of human-mouse conserved cell type–specific genes, with their number listed and top TFs named. UMI counts for these genes are scaled for library size, log-transformed, aggregated, and then mapped to *Z* scores.

do not develop until E18.5 (*164*)]. Human fetal oligodendrocytes overlap a rare mouse embryonic subtrajectory (*Pdgfra*+ glia) that, in retrospect, is more likely to correspond to oligodendrocyte precursors (*Olig1*+, *Olig2*+, and *Brinp3*+) (*165*, *166*), which calls into question our previous annotation of a different *Olig1*+ subtrajectory as oligodendrocyte precursors (*11*). These and other unexpected relationships merit further investigation.

To assess relationships between mouse embryonic and human fetal cells in greater detail, we applied the same strategy to extracted cells from the hematopoietic (Fig. 6D and fig. S24), endothelial (fig. S25), and epithelial (fig. S26) trajectories. In these visualizations, we observe examples of the organ-resolved human data deconvoluting the whole-embryo mouse data

into more fine-grained subsets. For example, subsets of the mouse white blood cell embryonic subtrajectory (*11*) map to specific human blood cell types such as HSPCs, microglia, macrophages (liver and spleen), macrophages (other organs), and dendritic cells (DCs) (Fig. 6D). These subsets were further validated by the expression of related blood cell markers (fig. S24C) and annotated on the basis of their human *k*-nearest neighbors ($k = 3$) in the co-embedding (fig. S24D).

Out of 1087 human fetal blood cell type–specific gene markers that are also differentially expressed across mouse blood cell types, 337 genes were differentially expressed (FDR of 0.05) in the same cell type (Fig. 6E and table S15; for comparison, only 12 genes intersected after permutations of labels). In total, 28 of these 337 conserved markers were TFs, 24 of which have been previously reported to be involved in early blood cell differentiation or maintenance for target cell types—e.g., *HLF* as a critical regulator of HSPCs quiescence (*167*), *MITF* as driving mast cell differentiation (*168*), *PAX5* as a master regulator of B cell development (*169*), and *SOX6* as enhancing the differentiation of erythroid progenitors (*170*). However, 4 of the 28 conserved marker TFs have not been previously characterized in the relevant context: *NR1D2* in IL 3 cells, *TCF7L2* in macrophages, *FHL2* in megakaryoblasts, and *NUAK1* in microglia.

In this same analysis, human fetal macrophage and microglia form distinct clusters, but they are connected by a subset of mouse cells from the white blood cell trajectory (Fig. 6D), consistent with previous studies showing that both cell types differentiate from yolk sac progenitors (*171*). To explore this further, we extracted and reanalyzed 4327 mouse embryonic microglia and macrophages by means of unsupervised trajectory analysis (*172*). We observed three smooth cell differentiation trajectories from a common progenitor to microglia in the brain, phagocytic macrophages (*TIMD4*+ and *CD5L*+; mostly in liver, spleen, and adrenal), and perivascular macrophages (*F13A1*+ and *LYVE1*+; widely distributed) (fig. S27A and Fig. 5). The directionality of progression through pseudotime along each macrophage trajectory was consistent with real developmental time (fig. S27B). In total, 1412 genes, including 111 TFs, were differentially expressed in the three macrophage branches (table S16). For example, the microglial trajectory showed elevated expression of *BACH2* and *RUNX3* as well as known microglial regulators *SALL1* (*173*) and *MEF2A* (*173*, *174*), perivascular macrophages of *DAB2*, and *TCF7L2*, and phagocytic macrophages of *MAFB* and *NR1H3* (fig. S27C). Overall, these analyses illustrate how fetal annotations can be used to identify and characterize progenitors of specific lineages at developmental time points where they may be difficult to resolve on their own, even across species.

Discussion

Two centuries after the formulation of the cell theory—the assertion that all living things consist of cells and that the cell is the most basic unit of life (*175*)—we are on the cusp of cataloging and characterizing all cell types that constitute a human body, both in health and disease. To this end, the field of single-cell biology is progressing at an astonishing rate, propelled by a synergy between new technologies and new computational methods to make sense of the data produced by those technologies. In the past few years alone, this synergy has enabled compelling and informative single-cell atlases of many human organs as well as of entire model organisms (*11*, *51*, *69*, *108*, *152*, *176*–*182*).

Human development is a remarkable process that begins with a fertilized zygote and proceeds through a germinal stage followed by embryogenesis. By the end of the 10th week, the embryo has acquired its basic form and is termed a fetus. For the following 30 weeks, all organs continue to grow and mature, with diverse terminally differentiated cell types arising from their progenitors. Although the germinal and embryogenesis stages have been intensively profiled with single-cell methods in humans and mice (*11*, *180*, *181*), it has been more challenging to profile the fetal stage. Although several single-cell studies of human fetal development have recently appeared (*152*, *182*–*184*), these are restricted to individual organs or cell lineages and do not obtain a comprehensive view.

In this study, together with (*12*), we set out to generate single-cell atlases of gene expression and chromatin accessibility using diverse tissues obtained during human fetal development. From 15 distinct organs, we successfully profiled gene expression in ~4 million single cells and chromatin accessibility in ~800,000 single cells. Limitations of these datasets include nonuniform sampling (i.e., more cells profiled in some organs than others), missing tissues (most notably, bone marrow, skin, bone, and gonads), relatively low sequencing depth, and the sparsity of single-cell molecular profiles. Nonetheless, we identified hundreds of cell types and subtypes that are supported by a framework for quantifying specificity as well as by matching nearly all of them to cell types or subtypes from published mouse atlases.

In contrast with organ-specific studies, the diversity of tissues profiled here enabled cross-tissue comparisons of broadly distributed cell types. We emphasize that our process for annotating cell types benefited tremendously from the myriad single-cell atlases of specific human organs or other mammals that have been generated to date (*8*, *9*, *11*, *16*, *28*, *50*, *108*, *139*). Of course, decisions in the annotation process can be subjective (e.g., over- versus under-clustering), and both cell type and subtype annotations made here should be considered preliminary and subject to revision.

The apparent hematopoiesis that we observe in the fetal adrenal gland is consistent with the fact that the adrenal gland, along with many other organs (e.g., spleen, liver, and lymph nodes), can serve as a site of extramedullary hematopoiesis in adults with pathologic conditions that lead to an increased demand for blood cell production, particularly hemoglobinopathies (*185*, *186*). Although occasional islands of extramedullary hematopoiesis have been seen in the adrenal glands of human embryos (*126*, *127*), our findings in both the human and mouse provide quantitative evidence that the adrenal gland serves as a normal, albeit minor, site of erythropoiesis during a developmental window that overlaps with the transition of hematopoiesis from the liver to the marrow.

The ease with which we were able to integrate single-cell profiles from mouse organogenesis and human fetal development is notable, particularly given that these represent different stages of mammalian development, not to mention our separation from mice by >100 million years of evolution. The relatively straightforward alignment of the datasets highlights the extent of evolutionary constraint on the molecular programs of individual cell types, and it furthermore lends support to long-standing use of the mouse as a powerful model system for studying human development.

Looking forward, we envision that the somewhat narrow window of midgestational human development studied here will be complemented by additional atlases of earlier and later time points (e.g., embryonic and adult) as well as by similarly comprehensive profiling and integration of data from model organisms. The continued development and application of methods for ascertaining gene expression and chromatin accessibility—in concert with spatial, epigenetic, proteomic, lineage history, and other information—will be necessary to obtain a comprehensive view of temporal unfolding of human cell type diversity that begins at the single-cell zygote.

To date, investigations of human development have largely been indirect, with key molecular factors nominated by human genetics and then investigated in model organisms and/or in vitro systems. Knowledge of the in vivo landscape of gene expression and regulation has been limited. In filling part of this gap, we hope that this atlas will enable a better understanding of the molecular and cellular basis of both rare and common disorders of human development, while also informing the path to successful therapies.

Materials and methods

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

sci-RNA-seq3

A more detailed version of the full sci-RNA-seq3 workflow is available on protocols.io (*187*) and in the supplementary materials.

Preparation of nuclei

Human fetal tissues (89 to 125 days estimated postconceptual age) were obtained by the University of Washington Birth Defects Research Laboratory (BDRL) under a protocol approved by the University of Washington Institutional Review Board. Tissues of interest were isolated and rinsed in 1X HBSS. Dried tissue was snap frozen in liquid nitrogen, manually pulverized on dry ice with a chilled hammer, aliquoted, and stored at −80°C until further processing. A subset of these aliquots were used for sci-RNA-seq3, and others for sci-ATAC-seq3, as described in the companion paper. For RNA-seq, nuclei from tissues and control cell lines were lysed in the cell lysis buffer and fixed with ice-cold 4% paraformaldehyde (EMS, 15-4-100) on the basis of the published sci-RNA-seq3 protocol (*11*). For human cell extraction in renal and digestive organs (kidney, pancreas, intestine, and stomach) and paraformaldehyde fixation, we followed the procedure described in (*13*).

Immunohistochemistry

Fetal tissues were fixed in formalin and embedded in paraffin. Sections of 4- to 5-μm thickness were cut and placed on Superfrost Plus slides (12-550-17, FisherBrand). For immunohistochemistry, sections were subjected to heat-mediated antigen retrieval (pH 6.0) followed by blocking with normal serum. Primary antibodies were incubated overnight at 4°C. The primary antibody we used: GYPA (R&D, MAB1228, 1:250), CD34 (R&D, AF7227, 1:250), CD34 (Novus, NBP2-32933, 1:250), ANXA1 (R&D, AF3770, 1:500), TNFRS10C (R&D, MAB6301, 1:500), AFP (Novus, NBP1-76275, 1:400), ALB (R&D, MAB1455, 1:10K), AHSG (R&D, AF1184, 1:400), and APOA1 (R&D, MAB36641, 1:250). Species and subtype-appropriate fluorescent dye-labeled secondary antibodies were used (Alexa Fluor 488 and 594, 1:400, Jackson ImmunoResearch Lab) or biotinylated secondary antibody were used followed by ABC Elite Systems (PK-6100, Vector Lab) for 3,3′-diaminobenzidine (DAB) chromogen staining.

sci-RNA-seq3 library construction and sequencing

The paraformaldehyde fixed nuclei were processed similarly to the published sci-RNA-seq3 protocol (*11*). For paraformaldehyde fixed cells, frozen fixed cells were thawed on 37°C water bath, spun down at 500 × *g* for 5 min, and incubated with 500 μl PBSI [1 x phosphate-buffered saline (PBS), pH 7.4, 1% bovine serum albumin (BSA), 1% SuperRnaseIn] including 0.2% Triton X-100 for 3 min on ice. Cells were pelleted and resuspended in 500 μl nuclease-free water including 1% SuperRnaseIn. 3 ml 0.1N HCl were added into the cells for 5min incubation on ice (*17*). 3.5 ml Tris-HCl (pH 8.0) and 35 μl 10% Triton X-100 were added into cells to neutralize HCl. Cells were pelleted and washed with 1 ml PBSR. Cells were pelleted and resuspended in 100 μl PBSI. The following steps were similar with the sci-RNA-seq3 protocol (with paraformaldehyde fixed nuclei) with slight modifications: (i) We distributed 20,000 fixed cells (instead of 80,000 nuclei) per well for reverse transcription (RT). (ii) We replaced all nuclei wash buffer in following steps with PBSI. (iii) All nuclei dilution buffer were replaced with PBS + 1% BSA.

Processing of sequencing reads

Read alignment and gene count matrix generation for the scRNA-seq was performed using the pipeline that we developed for sci-RNA-seq3 (*11*) with minor modifications: Duplicates were removed using the UMI sequence (ED < 2, including insertions and deletions), RT index, hairpin ligation adaptor index, and read 2 end-coordinate.

After the single-cell gene count matrix was generated, cells with <250 UMIs were filtered out. Each cell was assigned to its original human fetal sample on the basis of the RT barcode. Reads mapping to each fetus individual were aggregated to generate pseudobulk RNA-seq datasets. For sex assignments, we counted reads mapping to female-specific noncoding RNA (*TSIX* and *XIST*) or chrY genes (except genes *TBL1Y*, *RP11-424G14.1*, *NLGN4Y*, *AC010084.1*, *CD24P4*, *PCDH11Y*, and *TTTY14*, which are detected in both males and females). Fetuses were readily separated into females (more reads mapping to *TSIX* and *XIST* than chrY genes) and males (more reads mapping to chrY genes than *TSIX* and *XIST*).

Clustering analysis of pseudobulk transcriptomes was done with Monocle 3/alpha (*11*). Briefly, an aggregated gene expression matrix was constructed as described above for human fetal organs from each individual. Samples with >5000 total UMIs were selected. The dimensionality of the data was reduced by principal components analysis (PCA) (10 components), first on the top 500 most highly dispersed genes and then with UMAP (max_components = 2, n_neighbors = 10, min_dist = 0.5, metric = 'cosine').

Cell filtering, clustering and marker gene identification

For the detection of potential doublet cells and doublet-derived subclusters from each organ, we used an iterative clustering strategy as shown before (*11*). For data visualization, cells labeled as doublets [by scrublet/v0.1 pipeline (*188*)] or from doublet-derived subclusters were filtered out. For each cell, we only retain protein-coding genes, lincRNA genes and pseudogenes. Genes expressed in <10 cells and cells expressing <100 genes were further filtered out. The downstream dimension reduction and clustering analysis were done by Monocle 3/alpha with similar settings (*11*). Clusters were assigned to known cell types on the basis of cell type–specific markers (table S3). We found the above Scrublet and iterative clustering-based approach is limited in marking cell doublets between abundant cell clusters and rare cell clusters (e.g., <1% of total cell population). To further remove such doublet cells, we took the cell clusters identified by Monocle 3 and first computed differentially expressed genes across cell clusters (within-organ) with the differentialGeneTest() function of Monocle 3. We then selected a gene set combining the top ten gene markers for each cell cluster (ordered by *q* value and fold expression difference between first and second ranked cell cluster). Cells from each main cell cluster were selected for dimension reduction by PCA (10 components) first on the selected gene set of top cluster specific gene markers, and then by UMAP (max_components = 2, n_neighbors = 50, min_dist = 0.1, metric = 'cosine'), followed by clustering identification using the density peak clustering algorithm implemented in Monocle 3 (rho_thresh = 5, delta_thresh = 0.2 for most clustering analysis). Subclusters showing low expression of target cell cluster specific markers and enriched expression of nontarget cell cluster specific markers were annotated as doublets derived subclusters and filtered out in visualization and downstream analysis. Differentially expressed genes across cell types (within-organ) were recomputed with the differentialGeneTest() function of Monocle 3 after removing all doublets or cells from doublet-derived subclusters.

Adjudication of the 15 initially unannotated cell types

As noted in the main text, our first round of annotation was performed on a tissue-by-tissue basis by comparing observed cell types with those expected from prior knowledge of the same tissue. In general, we recovered all or nearly all main cell types identified by previous atlasing efforts directed at the same organs, despite differences with respect to species, stage of development and/or technology. Additionally, we identified 15 cell types that we did not at least initially expect to observe in a given tissue. We labeled these on the basis of the top enriched differentially expressed gene markers within that tissue, e.g., CSH1_CSH2 positive cells. After the initial round of annotation, we reexamined these 15 cell types on the basis of their distribution

in the global UMAP, whether they matched annotated cell types in mouse atlases, their distribution across tissues derived from different individuals, and their potential for maternal origin. Our updated interpretations are summarized in the supplementary materials.

Clustering analysis of cells across organs

For clustering analysis of 77 main cell types across 15 organs, we sampled 5000 cells from each cell type (or all cells for cell types with <5000 cells in a given organ). The dimensionality of the data was reduced first by PCA (50 components) on the gene set combining top cell type–specific gene markers identified above (table S5, q value = 0) and then with UMAP (max_components = 2, n_neighbors = 50, min_dist = 0.1, metric = 'cosine'). Differentially expressed genes across cell types were identified with the differentialGeneTest() function of Monocle 3. For annotating cell type–specific gene features, we intersected the cell type–specific genes identified above with the predicted secreted and membrane protein coding gene sets from the Human Protein Atlas (*189*), as well as the TF set annotated in the "motifAnnotations_hgnc" data from package RcisTarget/v1.2.1 (*190*).

For clustering analysis of blood cell across 15 organs, we extracted all blood cells corresponding to annotated clusters of myeloid cells, lymphoid cells, thymocytes, megakaryocytes, microglia, antigen presenting cells, erythroblasts, and HSPCs. The dimensionality of the data was reduced first by PCA (40 components) on the expression of a gene set combining the top 3000 blood cell type–specific gene markers (table S5, only genes specifically expressed in at least one blood cell type were selected (q < 0.05, fold expression difference between first and second ranked cell cluster > 2) and ordered by median q value across organs) and then with UMAP (max_components = 2, n_neighbors = 50, min_dist = 0.1, metric = 'cosine'). Cell clusters were identified using the Louvain algorithm implemented in Monocle 3 (louvain_res = 1×10^{-4}). Clusters were assigned to known cell types on the basis of cell type–specific markers. We then coembedded the human fetal blood cells and a scRNA-seq atlas of blood cells from the fetal liver (*108*), using the Seurat v3 integration method (FindAnchors and IntegrateData) (*15*) with a chosen dimensionality of 30 on the top 3000 highly variable genes with shared gene names in both datasets.

We then applied a similar analysis strategy as above for clustering analysis of endothelial or epithelial cells across organs. For endothelial cells, we first extracted cells corresponding to annotated clusters of vascular endothelial cells, lymphatic endothelial cells and endocardial cells across organs. The dimensionality of the data was reduced first by PCA (30 components) on the gene set combining top 1000 endothelial cell type–specific gene markers identified above (table S5, only genes specifically expressed in at least one endothelial cell type were selected (q < 0.05, fold expression difference between first and second ranked cell cluster > 2) and ordered by median q value across organs) and then with UMAP with the same parameters used for blood cells. Cell clusters were identified using the Louvain algorithm implemented in Monocle 3 (louvain_res = 1×10^{-4}), and then annotated on the basis of the tissue origin of endothelial cells. For epithelial cells, we first extracted cells from the epithelial cell cluster in fig. S4B, followed by dimension reduction first by PCA (50 components) first on the top 5000 most highly dispersed genes and then with UMAP (max_components = 2, n_neighbors = 50, min_dist = 0.1, metric = 'cosine'). For validating the tissue specific endothelial cells, we then coembedded the human fetal endothelial cells and a scRNA-seq atlas of endothelial cells from mouse adult tissues (*139*), using the Seurat v3 integration method (FindAnchors and IntegrateData) (*15*) with a chosen dimensionality of 30 on the top 3000 highly variable genes with shared gene names in both datasets.

Intradataset cross-validation analysis

For cells from each organ, we randomly sampled up to 2000 cells from each main cell type. We then followed the same process (*101*). Briefly, we combined all sampled cells from each organ and evaluated cell type specificity by applying a fivefold cross-validation to the dataset, with an SVM classifier (with linear kernel). Whole transcriptome was used in cell type prediction. We then computed the cross-validation F1 value as cell type specificity score. As control, we randomly permuted the cell type labels, followed by the same analysis pipeline. For cell type specificity analysis across all organs, we applied the same analysis strategy to the full dataset after sampling up to 2000 cells of each main cell type.

Subclustering analysis

For each main cell type (with >1000 cells) in each organ, we applied Harmony/v1.0 for batch correction and dimension reduction (*102*). Briefly, the dimensionality of the data was reduced by PCA (30 components, or 10 components for cell types with <5000 cells) first on the top 3000 (or 1000 for cell types with <5000 cells) most highly variable genes, followed by batch correction on sample ID. Cell clusters were identified using the Louvain algorithm implemented in Seurat/v3.1.4 (*15*) (resolution = 0.5). We then applied the intradataset cross-validation approach to evaluate the specificity of subclusters within each main cell type. For every subcluster pair, A and B, we computed the number of A cells mislabeled as B cells in cross-validation analysis with the true dataset (mislabeled cell number: n) or the permuted dataset (mislabeled cell number: m). A large n value suggests the two subclusters are not well separated by the full transcriptome. We thus iteratively merged similar subcluster pairs ($n > m$), and identified a total of 657 subtypes across 15 organs. The intradataset cross-validation approach was applied to evaluating subtype specificity within each main cell type in each organ. To annotate the identity of subtypes, we applied the same cell type correlation analysis strategy described in (*11*) to compare cell subtypes from this study with cell types of the same organ from the Microwell-seq based Mouse Cell Atlas (MCA) (*16*). A similar comparison was performed for all subtypes from the brain against cell types annotated in a recent mouse brain atlas (MBCA) (*50*).

Validating erythropoiesis in the adrenal tissues from newborn mice

Adrenals and kidneys were harvested from CD1 Swiss albino mice (Charles River) on the day of birth (P0), and bone marrow cells were flushed from the femurs of the dams. Solid tissues were dissociated using collagenase and stained for imaging flow cytometry using the markers Ter119 (AF488), CD117 (PE-CF594), CD71 (PE), CD45 (EF450), and DRAQ5. Gating of maturing erythroblast populations was performed using published methods (*128*) and analyzed with IDEAS (Luminex) software.

Comparison of human and mouse developmental atlases

We first applied a slightly modified version of the strategy described in (*11*) to identify correlated cell types between this human fetal cell atlas and the mouse organogenesis cell atlas (MOCA) (*11*). As a different approach, we coembedded the human fetal cell atlas and the mouse organogenesis cell atlas (MOCA) (*11*) using the Seurat v3 integration method (FindAnchors and IntegrateData) (*15*) with a chosen dimensionality of 30 on the top 3000 highly variable genes with shared gene names in both human and mouse. We first integrated 65,000 human fetal cells (up to 1000 cells randomly sampled from each of 77 cell types) and 100,000 mouse embryonic cells (randomly sampled from MOCA) with default parameters. We then applied the same integrative analysis strategy to extracted human and mouse cells from the hematopoietic, endothelial, and epithelial trajectories.

For the coembedded human and mouse hematopoietic cells, we annotated each mouse cell on the basis of its k-nearest neighbors of human cells. We chose a small k value (k = 3) such that rare cell types were also annotated. Differentially expressed genes across mouse hematopoietic cells were computed with the

differentialGeneTest() function of Monocle 3/alpha.

Pseudotemporal ordering of mouse macrophage/microglia cells was done with Monocle 3/alpha with the reduction method of "DDRTree." Briefly, the top 3 principal components on the top 500 highly variable genes were used to construct the DDRTree pseudotime trajectory with UMI number per cell as a covariate [param.gamma = 120, norm_method = "log," residualModelFormulaStr = "~ sm.ns(Total_mRNAs, df = 3)"]. The cells are separated into three branch trajectories in the DDRTree space. Differentially expressed genes across the three branches were computed with the differentialGeneTest() function of Monocle 3/alpha. We then clustered cells with k means clustering (k = 10) and computed the average development time for each cluster. The progenitor cell group was annotated on the basis of the lowest average development time and appeared at the center of the three branches. Each cell was assigned a pseudotime value on the basis of its distance from the progenitor cells.

Using the Garnett models trained on this human cell atlas for cell type classification

The R package Garnett for Monocle 3 (version 0.2.9) was used to generate cell type classifiers for each of the 15 tissues. Marker genes for each cell type were assembled from literature searches, and models were trained using train_cell_classifier using default parameters and the gene database org.Hs.eg.db (version 3.10.0). Models were trained on the entirety of each tissue dataset with the exception of cerebrum, where 100,000 cells were randomly sampled for training for computational efficiency. The automatic and manual annotation processes were conducted independently. To compare cell type assignments with those obtained through manual annotation (i.e., the 77 main cell types), we applied the function classifiy_cells using the trained models with the following nondefault parameters: cluster_extend = TRUE, cluster_extend_max_frac_incorrect = 0.25, cluster_extend_max_frac_unknown = 0.95. Garnett cell type assignments that matched the cell type assignment from manual annotation were considered correct with the following exceptions: Garnett classification of "Chromaffin cells" was considered correct when manual annotation was "Sympathoblasts;" Garnett classification of "B cells" or "T cells" was considered correct when manual annotation was "Lymphoid cells;" Garnett classification of "Cap mesenchyme cells," "Collecting duct cells," "Distal tubule cells," "Loop of Henle cells," "Proximal tubule cells," and "Podocytes" were considered correct when manual annotation was "Metanephric cells;" Garnett classification of "Ureter cells" and "Collecting duct cells" were considered correct when manual annotation was "Ureteric bud cells;" Garnett classification of "Pancreatic Alpha cells," "Pancreatic Beta cells," and "Pancreatic Delta cells" was considered correct when manual annotation was "Islet endocrine cells;" Garnett classification of "D cells" was considered correct with manual annotation of "Neuroendocrine cells."

To test the applicability of Garnett trained models to future data, we applied the pancreas model to human adult pancreas scRNA-seq data from reference (*100*). The model was applied using the function classify_cells with the same parameters as above. When comparing cell type assignments with those provided by the authors, we considered the following cell types to be equivalent: acinar, Acinar cells; ductal, Ductal cells; endothelial, Endothelial cells; mast, Myeloid cells; macrophage, Myeloid cells; schwann, Glia; alpha, Pancreatic Alpha cells; beta, Pancreatic Beta cells; delta, Pancreatic Delta cells; activated_stellate, Pancreatic stellate cells; quiescent_stellate, Pancreatic stellate cells; and t_cell, T cells.

REFERENCES AND NOTES

1. B. S. de Bakker *et al*., An interactive three-dimensional digital atlas and quantitative database of human development. *Science* **354**, aag0053 (2016). doi: 10.1126/science.aag0053; pmid: 27884980
2. S. Yamada *et al*., Developmental atlas of the early first trimester human embryo. *Dev. Dyn.* **239**, 1585–1595 (2010). doi: 10.1002/dvdy.22316; pmid: 20503356
3. J. E. Jirasek, *An Atlas of the Human Embryo and Fetus: A Photographic Review of Human Prenatal Development* (CRC Press, 2000).
4. V. A. McKusick, Mendelian Inheritance in Man and its online version, OMIM. *Am. J. Hum. Genet.* **80**, 588–604 (2007). doi: 10.1086/514346; pmid: 17357067
5. H. C. Mefford, M. L. Batshaw, E. P. Hoffman, Genomics, intellectual disability, and autism. *N. Engl. J. Med.* **366**, 733–743 (2012). doi: 10.1056/NEJMra1114194; pmid: 22356326
6. S. E. McCandless, J. W. Brunger, S. B. Cassidy, The burden of genetic disease on inpatient care in a children's hospital. *Am. J. Hum. Genet.* **74**, 121–127 (2004). doi: 10.1086/381053; pmid: 14681831
7. M. H. Wojcik *et al*., Genetic disorders and mortality in infancy and early childhood: Delayed diagnoses and missed opportunities. *Genet. Med.* **20**, 1396–1404 (2018). doi: 10.1038/gim.2018.17; pmid: 29790870
8. S. Darmanis *et al*., A survey of human brain transcriptome diversity at the single cell level. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7285–7290 (2015). doi: 10.1073/pnas.1507125112; pmid: 26060301
9. E. Madissoon *et al*., scRNA-seq assessment of the human lung, spleen, and esophagus tissue stability after cold preservation. *Genome Biol.* **21**, 1 (2019). doi: 10.1186/s13059-019-1906-x; pmid: 31892341
10. S. Behjati, S. Lindsay, S. A. Teichmann, M. Haniffa, Mapping human development at single-cell resolution. *Development* **145**, dev152561 (2018). doi: 10.1242/dev.152561; pmid: 29439135
11. J. Cao *et al*., The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019). doi: 10.1038/s41586-019-0969-x; pmid: 30787437
12. S. Domcke *et al*., A human cell atlas of fetal chromatin accessibility. *Science* **370**, eaba7612 (2020). doi: 10.1126/science.aba7612
13. E. R. Thomsen *et al*., Fixed single-cell transcriptomic characterization of human radial glial diversity. *Nat. Methods* **13**, 87–93 (2016). doi: 10.1038/nmeth.3629; pmid: 26524239
14. L. McInnes, J. Healy, N. Saul, L. Großberger, UMAP: Uniform Manifold Approximation and Projection. *J. Open Source Softw.* **3**, 861 (2018). doi: 10.21105/joss.00861
15. T. Stuart *et al*., Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888–1902.e21 (2019). doi: 10.1016/j.cell.2019.05.031; pmid: 31178118
16. X. Han *et al*., Mapping the Mouse Cell Atlas by Microwell-Seq. *Cell* **172**, 1091–1107.e17 (2018). doi: 10.1016/j.cell.2018.05.012; pmid: 29775597
17. A. B. Rosenberg *et al*., Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* **360**, 176–182 (2018). doi: 10.1126/science.aam8999; pmid: 29545511
18. J. Cao *et al*., Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **357**, 661–667 (2017). doi: 10.1126/science.aam8940; pmid: 28818938
19. M. Plass *et al*., Cell type atlas and lineage tree of a whole complex animal by single-cell transcriptomics. *Science* **360**, eaaq1723 (2018). doi: 10.1126/science.aaq1723; pmid: 29674432
20. 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
21. G. Stelzer *et al*., The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr. Protoc. Bioinformatics* **54**, 1.30.1–1.30.33 (2016). doi: 10.1002/cpbi.5
22. G. Teitelman, H. Baker, T. H. Joh, D. J. Reis, in *Catecholamines: Basic and Clinical Frontiers*, E. Usdin, I. J. Kopin, J. Barchas, Eds. (Pergamon, 1979), pp. 794–796.
23. Y. Lu, W. Yi, Q. Wu, S. Zhong, Z. Zuo, F. Zhao, M. Zhang, N. Tsai, Y. Zhuo, S. He, J. Zhang, X. Duan, X. Wang, T. Xue, Single-cell RNA-seq analysis maps the development of human fetal retina. bioRxiv 423830 [Preprint]. 23 September 2018. doi: 10.1101/423830
24. The UniProt Consortium, UniProt: The universal protein knowledgebase. *Nucleic Acids Res.* **45**, D158–D169 (2017). doi: 10.1093/nar/gkw1099; pmid: 27899622
25. Gene_Cryab (2020); www.ncbi.nlm.nih.gov/gene/12955.
26. CRBB3_HUMAN, UniProt accession no. P26998; www.uniprot.org/uniprot/P26998.
27. K1C12_HUMAN, UniProt accession no. Q99456; www.uniprot.org/uniprot/Q99456.
28. Tabula Muris Consortium, Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature* **562**, 367–372 (2018). doi: 10.1038/s41586-018-0590-4; pmid: 30283141
29. B. R. Kim *et al*., SOX2 and PI3K Cooperate to Induce and Stabilize a Squamous-Committed Stem Cell Injury State during Lung Squamous Cell Carcinoma Pathogenesis. *PLOS Biol.* **14**, e1002581 (2016). doi: 10.1371/journal.pbio.1002581; pmid: 27880766
30. S. Dauger *et al*., Phox2b controls the development of peripheral chemoreceptors and afferent visceral pathways. *Development* **130**, 6635–6642 (2003). doi: 10.1242/dev.00866; pmid: 14627719
31. D. G. Jackson, R. Prevo, S. Clasper, S. Banerji, LYVE-1, the lymphatic system and tumor lymphangiogenesis. *Trends Immunol.* **22**, 317–321 (2001). doi: 10.1016/S1471-4906(01)01936-6; pmid: 11377291
32. Y. Rinkevich *et al*., Identification and prospective isolation of a mesothelial precursor lineage giving rise to smooth muscle cells and fibroblasts for mammalian internal organs, and their vasculature. *Nat. Cell Biol.* **14**, 1251–1260 (2012). doi: 10.1038/ncb2610; pmid: 23143399
33. Y. Cui *et al*., Single-Cell Transcriptome Analysis Maps the Developmental Track of the Human Heart. *Cell Rep.* **26**, 1934–1950.e5 (2019). doi: 10.1016/j.celrep.2019.01.079; pmid: 30759401
34. M. E. Shy *et al*., Schwann cell expression of PLP1 but not DM20 is necessary to prevent neuropathy. *Ann. Neurol.* **53**, 354–365 (2003). doi: 10.1002/ana.10466; pmid: 12601703
35. H. Suryawanshi *et al*., A single-cell survey of the human first-trimester placenta and decidua. *Sci. Adv.* **4**, eaau4788 (2018). doi: 10.1126/sciadv.aau4788; pmid: 30402542
36. F. E. Thorngate, P. A. Strockbine, S. K. Erickson, D. L. Williams, Altered adrenal gland cholesterol metabolism in the apoE-deficient mouse. *J. Lipid Res.* **43**, 1920–1926 (2002). doi: 10.1194/jlr.M200205-JLR200; pmid: 12401891
37. A. Furlan *et al*., Multipotent peripheral glial cells generate neuroendocrine cells of the adrenal medulla. *Science* **357**, eaal3753 (2017). doi: 10.1126/science.aal3753; pmid: 28684471
38. M. E. Fomin, A. I. Beyer, M. O. Muench, Human fetal liver cultures support multiple cell lineages that can engraft immunodeficient mice. *Open Biol.* **7**, 170108 (2017). doi: 10.1098/rsob.170108; pmid: 29237808
39. S. Matsuoka *et al*., CD34 expression on long-term repopulating hematopoietic stem cells changes during developmental stages. *Blood* **97**, 419–425 (2001). doi: 10.1182/blood.v97.2.419; pmid: 11154218

40. J. D. Miller, S. M. Lankford, K. B. Adler, A. R. Brody, Mesenchymal stem cells require MARCKS protein for directed chemotaxis in vitro. *Am. J. Respir. Cell Mol. Biol.* **43**, 253–258 (2010). doi: 10.1165/rcmb.2010-0015RC; pmid: 20224071
41. F. Dijk, A. A. B. Bergen, W. Kamphuis, GAP-43 expression is upregulated in retinal ganglion cells after ischemia/reperfusion-induced damage. *Exp. Eye Res.* **84**, 858–867 (2007). doi: 10.1016/j.exer.2007.01.006; pmid: 17343850
42. R. P. de Melo Guimarães *et al*., Evidence of Müller Glia Conversion Into Retina Ganglion Cells Using Neurogenin2. *Front. Cell. Neurosci.* **12**, 410 (2018). doi: 10.3389/fncel.2018.00410; pmid: 30483060
43. R. A. Poché *et al*., Lim1 is essential for the correct laminar positioning of retinal horizontal cells. *J. Neurosci.* **27**, 14099–14107 (2007). doi: 10.1523/JNEUROSCI.4046-07.2007; pmid: 18094249
44. P. A. Ruzycki, X. Zhang, S. Chen, CRX directs photoreceptor differentiation by accelerating chromatin remodeling at specific target sites. *Epigenetics Chromatin* **11**, 42 (2018). doi: 10.1186/s13072-018-0212-2; pmid: 30068366
45. Z. Shi *et al*., Vsx1 regulates terminal differentiation of type 7 ON bipolar cells. *J. Neurosci.* **31**, 13118–13127 (2011). doi: 10.1523/JNEUROSCI.2331-11.2011; pmid: 21917795
46. M. V. Sofroniew, H. V. Vinters, Astrocytes: Biology and pathology. *Acta Neuropathol.* **119**, 7–35 (2010). doi: 10.1007/s00401-009-0619-8; pmid: 20012068
47. F. Beleggia *et al*., CRIM1 haploinsufficiency causes defects in eye development in human and mouse. *Hum. Mol. Genet.* **24**, 2267–2273 (2015). doi: 10.1093/hmg/ddu744; pmid: 25561690
48. C. J. Jaworski *et al*., Expression analysis of human pterygium shows a predominance of conjunctival and limbal markers and genes associated with cell migration. *Mol. Vis.* **15**, 2421–2434 (2009). pmid: 19956562
49. S. Rétaux, M. Rogard, I. Bach, V. Failli, M. J. Besson, *Lhx9*: A novel LIM-homeodomain gene expressed in the developing forebrain. *J. Neurosci.* **19**, 783–793 (1999). doi: 10.1523/JNEUROSCI.19-02-00783.1999; pmid: 9880598
50. A. Zeisel *et al*., Molecular Architecture of the Mouse Nervous System. *Cell* **174**, 999–1014.e22 (2018). doi: 10.1016/j.cell.2018.06.021; pmid: 30096314
51. R. Menon *et al*., Single-cell analysis of progenitor cell dynamics and lineage specification in the human fetal kidney. *Development* **145**, dev164038 (2018). doi: 10.1242/dev.164038; pmid: 30166318
52. M. Ben Khelifa *et al*., Mutations in *DNAH1*, which encodes an inner arm heavy chain dynein, lead to male infertility from multiple morphological abnormalities of the sperm flagella. *Am. J. Hum. Genet.* **94**, 95–104 (2014). doi: 10.1016/j.ajhg.2013.11.017; pmid: 24360805
53. H. He *et al*., Krüppel-like Factor 4 Promotes Esophageal Squamous Cell Carcinoma Differentiation by Up-regulating Keratin 13 Expression. *J. Biol. Chem.* **290**, 13567–13577 (2015). doi: 10.1074/jbc.M114.629717; pmid: 25851906
54. Y. Pan, W.-D. Wang, T. Yago, Transcriptional regulation of podoplanin expression by Prox1 in lymphatic endothelial cells. *Microvasc. Res.* **94**, 96–102 (2014). doi: 10.1016/j.mvr.2014.05.006; pmid: 24944097
55. U. Hopfer *et al*., Targeted disruption of *Col8a1* and *Col8a2* genes in mice leads to anterior segment abnormalities in the eye. *FASEB J.* **19**, 1232–1244 (2005). doi: 10.1096/fj.04-3019com; pmid: 16051690
56. R. Vento-Tormo *et al*., Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* **563**, 347–353 (2018). doi: 10.1038/s41586-018-0698-6; pmid: 30429548
57. J. A. Gomez-Sanchez *et al*., Schwann cell autophagy, myelinophagy, initiates myelin clearance from injured nerves. *J. Cell Biol.* **210**, 153–168 (2015). doi: 10.1083/jcb.201503019; pmid: 26150392
58. M. Hoekstra, M. Van Eck, HDL is redundant for adrenal steroidogenesis in LDLR knockout mice with a human-like lipoprotein profile. *J. Lipid Res.* **57**, 631–637 (2016). doi: 10.1194/jlr.M066019; pmid: 26891738
59. K. Huber, N. Karch, U. Ernsberger, C. Goridis, K. Unsicker, The role of Phox2B in chromaffin cell development. *Dev. Biol.* **279**, 501–508 (2005). doi: 10.1016/j.ydbio.2005.01.007; pmid: 15733675
60. X. Pei, Who is hematopoietic stem cell: CD34+ or CD34-? *Int. J. Hematol.* **70**, 213–215 (1999). pmid: 10643145
61. J. R. Lavoie *et al*., Brief Report: Elastin Microfibril Interface 1 and Integrin-Linked Protein Kinase Are Novel Markers of Islet Regenerative Function in Human Multipotent Mesenchymal Stromal Cells. *Stem Cells* **34**, 2249–2255 (2016). doi: 10.1002/stem.2385; pmid: 27090767
62. D. S. Wall *et al*., Progenitor cell proliferation in the retina is dependent on Notch-independent Sonic hedgehog/Hes1 activity. *J. Cell Biol.* **184**, 101–112 (2009). doi: 10.1083/jcb.200805155; pmid: 19124651
63. F. Wu *et al*., Two transcription factors, Pou4f2 and Isl1, are sufficient to specify the retinal ganglion cell fate. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E1559–E1568 (2015). doi: 10.1073/pnas.1421535112; pmid: 25775587
64. R. Balasubramanian, A. Bui, X. Dong, L. Gan, Lhx9 Is Required for the Development of Retinal Nitric Oxide-Synthesizing Amacrine Cell Subtype. *Mol. Neurobiol.* **55**, 2922–2933 (2018). doi: 10.1007/s12035-017-0554-y; pmid: 28456934
65. F. Wu *et al*., Onecut1 is essential for horizontal cell genesis and retinal integrity. *J. Neurosci.* **33**, 13053–13065 (2013). doi: 10.1523/JNEUROSCI.0116-13.2013; pmid: 23926259
66. M. Vitorino *et al*., Vsx2 in the zebrafish retina: Restricted lineages through derepression. *Neural Dev.* **4**, 14 (2009). doi: 10.1186/1749-8104-4-14; pmid: 19344499
67. S. Kolandaivelu, V. Ramamurthy, AIPL1 protein and its indispensable role in cone photoreceptor function and survival. *Adv. Exp. Med. Biol.* **801**, 43–48 (2014). doi: 10.1007/978-1-4614-3209-8_6; pmid: 24664679
68. I. R. Holtman *et al*., Induction of a common microglia gene expression signature by aging and neurodegenerative conditions: A co-expression meta-analysis. *Acta Neuropathol. Commun.* **3**, 31 (2015). doi: 10.1186/s40478-015-0203-5; pmid: 26001565
69. S. Zhong *et al*., A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. *Nature* **555**, 524–528 (2018). doi: 10.1038/nature25980; pmid: 29539641
70. M. Lee *et al*., Tcf7l2 plays crucial roles in forebrain development through regulation of thalamic and habenular neuron identity and connectivity. *Dev. Biol.* **424**, 62–76 (2017). doi: 10.1016/j.ydbio.2017.02.010; pmid: 28219675
71. J. Holzschuh *et al*., Noradrenergic neurons in the zebrafish hindbrain are induced by retinoic acid and require tfap2a for expression of the neurotransmitter phenotype. *Development* **130**, 5741–5754 (2003). doi: 10.1242/dev.00816; pmid: 14534139
72. R. A. Carter *et al*., A Single-Cell Transcriptional Atlas of the Developing Murine Cerebellum. *Curr. Biol.* **28**, 2910–2920.e2 (2018). doi: 10.1016/j.cub.2018.07.062; pmid: 30220501
73. J. Park *et al*., Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science* **360**, 758–763 (2018). doi: 10.1126/science.aar2131; pmid: 29622724
74. J. Madsen, O. Nielsen, I. Tornøe, L. Thim, U. Holmskov, Tissue localization of human trefoil factors 1, 2, and 3. *J. Histochem. Cytochem.* **55**, 505–513 (2007). doi: 10.1369/jhc.6A7100.2007; pmid: 17242463
75. P. Zhang *et al*., Dissecting the Single-Cell Transcriptome Network Underlying Gastric Premalignant Lesions and Early Gastric Cancer. *Cell Rep.* **27**, 1934–1947.e5 (2019). doi: 10.1016/j.celrep.2019.04.052; pmid: 31067475
76. T. F. Griggs *et al*., Rhinovirus C targets ciliated airway epithelial cells. *Respir. Res.* **18**, 84 (2017). doi: 10.1186/s12931-017-0567-0; pmid: 28472984
77. P. A. Reyfman *et al*., Single-Cell Transcriptomic Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis. *Am. J. Respir. Crit. Care Med.* **199**, 1517–1536 (2019). doi: 10.1164/rccm.201712-24100C; pmid: 30554520
78. S. M. Brierley, T. J. Hibberd, N. J. Spencer, Spinal Afferent Innervation of the Colon and Rectum. *Front. Cell. Neurosci.* **12**, 467 (2018). doi: 10.3389/fncel.2018.00467; pmid: 30564102
79. Y. Liu *et al*., Single-cell RNA-seq reveals the diversity of trophoblast subtypes and patterns of differentiation in the human placenta. *Cell Res.* **28**, 819–832 (2018). doi: 10.1038/s41422-018-0066-y; pmid: 30042384
80. P. S. Zammit *et al*., Pax7 and myogenic progression in skeletal muscle satellite cells. *J. Cell Sci.* **119**, 1824–1832 (2006). doi: 10.1242/jcs.02908; pmid: 16608873
81. T. Schöneberg, J. Meister, A. B. Knierim, A. Schulz, The G protein-coupled receptor GPR34 - The past 20 years of a grownup. *Pharmacol. Ther.* **189**, 71–88 (2018). doi: 10.1016/j.pharmthera.2018.04.008; pmid: 29684466
82. C. Mallika, Q. Guo, J. Y. H. Li, Gbx2 is essential for maintaining thalamic neuron identity and repressing habenular characters in the developing thalamus. *Dev. Biol.* **407**, 26–39 (2015). doi: 10.1016/j.ydbio.2015.08.010; pmid: 26297811
83. J. Cao *et al*., Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science* **361**, 1380–1385 (2018). doi: 10.1126/science.aau0730; pmid: 30166440
84. B. J. Pepe-Mooney *et al*., Single-Cell Analysis of the Liver Epithelium Reveals Dynamic Heterogeneity and an Essential Role for YAP in Homeostasis and Regeneration. *Cell Stem Cell* **25**, 23–38.e8 (2019). doi: 10.1016/j.stem.2019.04.004; pmid: 31080134
85. See supplementary materials.
86. P. A. Reyfman *et al*., Single-Cell Transcriptomic Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis. *Am. J. Respir. Crit. Care Med.* **199**, 1517–1536 (2019). doi: 10.1164/rccm.201712-24100C; pmid: 30554520
87. A. L. Haber *et al*., A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333–339 (2017). doi: 10.1038/nature24489; pmid: 29144463
88. 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
89. S. Krasemann *et al*., The TREM2-APOE Pathway Drives the Transcriptional Phenotype of Dysfunctional Microglia in Neurodegenerative Diseases. *Immunity* **47**, 566–581.e9 (2017). doi: 10.1016/j.immuni.2017.08.008; pmid: 28930663
90. T. Kielian, Toll-like receptors in central nervous system glial inflammation and homeostasis. *J. Neurosci. Res.* **83**, 711–730 (2006). doi: 10.1002/jnr.20767; pmid: 16541438
91. H. Kohno, T. Maeda, L. Perusek, E. Pearlman, A. Maeda, CCL3 production by microglial cells modulates disease severity in murine models of retinal degeneration. *J. Immunol.* **192**, 3816–3827 (2014). doi: 10.4049/jimmunol.1301738; pmid: 24639355
92. M. Uittenbogaard, K. K. Baxter, A. Chiaramello, NeuroD6 genomic signature bridging neuronal differentiation to survival via the molecular chaperone network. *J. Neurosci. Res.* **88**, 33–54 (2010). doi: 10.1002/jnr.22182; pmid: 19610105
93. T. Kodama *et al*., Neuronal classification and marker gene identification via single-cell expression profiling of brainstem vestibular neurons subserving cerebellar learning. *J. Neurosci.* **32**, 7819–7831 (2012). doi: 10.1523/JNEUROSCI.0543-12.2012; pmid: 22674258
94. H. Yanai *et al*., Dlk-1, a cell surface antigen on foetal hepatic stem/progenitor cells, is expressed in hepatocellular, colon, pancreas and breast carcinomas at a high frequency. *J. Biochem.* **148**, 85–92 (2010). doi: 10.1093/jb/mvq034; pmid: 20356822
95. D. Zhao *et al*., Derivation and characterization of hepatic progenitor cells from human embryonic stem cells. *PLOS ONE* **4**, e6468 (2009). doi: 10.1371/journal.pone.0006468; pmid: 19649295
96. H. Mutlu-Albayrak *et al*., Identification of SLC22A5 Gene Mutation in a Family with Carnitine Uptake Defect. *Case Rep. Genet.* **2015**, 259627 (2015). doi: 10.1155/2015/259627; pmid: 26075114
97. Z. Huang *et al*., ACSS2 promotes systemic fat storage and utilization through selective regulation of genes involved in lipid metabolism. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E9499–E9506 (2018). doi: 10.1073/pnas.1806635115; pmid: 30228117
98. S. C. S. Nagamani, A. Erez, A metabolic link between the urea cycle and cancer cell proliferation. *Mol. Cell. Oncol.* **3**, e1127314 (2016). doi: 10.1080/23723556.2015.1127314; pmid: 27308634
99. H. A. Pliner, J. Shendure, C. Trapnell, Supervised classification enables rapid annotation of cell atlases. *Nat. Methods* **16**, 983–986 (2019). doi: 10.1038/s41592-019-0535-3; pmid: 31501545
100. M. Baron *et al*., A Single-Cell Transcriptomic Map of the Human and Mouse Pancreas Reveals Inter- and Intra-cell Population Structure. *Cell Syst.* **3**, 346–360.e4 (2016). doi: 10.1016/j.cels.2016.08.011; pmid: 27667365
101. T. Abdelaal *et al*., A comparison of automatic cell identification methods for single-cell RNA sequencing data. *Genome Biol.* **20**, 194 (2019). doi: 10.1186/s13059-019-1795-z; pmid: 31500660
102. I. Korsunsky *et al*., Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019). doi: 10.1038/s41592-019-0619-0; pmid: 31740819

103. L. Castagnaro *et al*., Nkx2-5^{+}islet1^{+} mesenchymal precursors generate distinct spleen stromal cell subsets and participate in restoring stromal network integrity. *Immunity* **38**, 782–791 (2013). doi: 10.1016/j.immuni.2012.12.005; pmid: 23601687
104. A. Oda *et al*., Niche-induced extramedullary hematopoiesis in the spleen is regulated by the transcription factor Tlx1. *Sci. Rep.* **8**, 8308 (2018). doi: 10.1038/s41598-018-26693-x; pmid: 29844356
105. P.-H. Kim, S.-S. Na, B. Lee, J.-H. Kim, J.-Y. Cho, Stanniocalcin 2 enhances mesenchymal stem cell survival by suppressing oxidative stress. *BMB Rep.* **48**, 702–707 (2015). doi: 10.5483/BMBRep.2015.48.12.158; pmid: 26424558
106. 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
107. S. Epelman, K. J. Lavine, G. J. Randolph, Origin and functions of tissue macrophages. *Immunity* **41**, 21–35 (2014). doi: 10.1016/j.immuni.2014.06.013; pmid: 25035951
108. D.-M. Popescu *et al*., Decoding human fetal liver haematopoiesis. *Nature* **574**, 365–371 (2019). doi: 10.1038/s41586-019-1652-y; pmid: 31597962
109. Z. Bian *et al*., Deciphering human macrophage development at single-cell resolution. *Nature* **582**, 571–576 (2020). doi: 10.1038/s41586-020-2316-7; pmid: 32499656
110. Immune cell markers (BD Biosciences); www.bdbiosciences.com/us/s/cdmarkers.
111. Immune cell markers (proteintech); www.ptglab.com/products/featured-products/markers-for-immune-cells/.
112. Immune cell marker (Invitrogen eBioscience); www.thermofisher.com/us/en/home/life-science/cell-analysis/cell-analysis-learning-center/cell-analysis-resource-library/ebioscience-resources/human-cd-other-cellular-antigens.html.
113. S. Kitchen *et al*., Ligation of CD4 on human blood monocytes triggers macrophage differentiation (158.8). *J. Immunol.* **186**, 158.8 (2011); www.jimmunol.org/content/186/1_Supplement/158.8.
114. R. G. Domingues *et al*., CD5 expression is regulated during human T-cell activation by alternative polyadenylation, PTBP1, and miR-204. *Eur. J. Immunol.* **46**, 1490–1503 (2016). doi: 10.1002/eji.201545663; pmid: 27005442
115. F. Melo-Gonzalez, M. R. Hepworth, Functional and phenotypic heterogeneity of group 3 innate lymphoid cells. *Immunology* **150**, 265–275 (2017). doi: 10.1111/imm.12697; pmid: 27935637
116. M. Nagasawa *et al*., KLRG1 and NKp46 discriminate subpopulations of human CD117^{+}CRTH2^{-} ILCs biased toward ILC2 or ILC3. *J. Exp. Med.* **216**, 1762–1776 (2019). doi: 10.1084/jem.20190490; pmid: 31201208
117. E. Dzierzak, S. Philipsen, Erythropoiesis: Development and differentiation. *Cold Spring Harb. Perspect. Med.* **3**, a011601 (2013). doi: 10.1101/cshperspect.a011601; pmid: 23545573
118. I. Kobayashi *et al*., Enrichment of hematopoietic stem/progenitor cells in the zebrafish kidney. *Sci. Rep.* **9**, 14205 (2019). doi: 10.1038/s41598-019-50672-5; pmid: 31578390
119. K. J. Knudsen *et al*., ERG promotes the maintenance of hematopoietic stem cells by restricting their differentiation. *Genes Dev.* **29**, 1915–1929 (2015). doi: 10.1101/gad.268409.115; pmid: 26385962
120. B. Psaila *et al*., Single-cell profiling of human megakaryocyte-erythroid progenitors identifies distinct megakaryocyte and erythroid differentiation pathways. *Genome Biol.* **17**, 83 (2016). doi: 10.1186/s13059-016-0939-7; pmid: 27142433
121. K. Kataoka *et al*., Evi1 is essential for hematopoietic stem cell self-renewal, and its expression marks hematopoietic cells with long-term multilineage repopulating activity. *J. Exp. Med.* **208**, 2403–2416 (2011). doi: 10.1084/jem.20110447; pmid: 22084405
122. M. Lapierre *et al*., Expression and role of RIP140/NRIP1 in chronic lymphocytic leukemia. *J. Hematol. Oncol.* **8**, 20 (2015). doi: 10.1186/s13045-015-0116-6; pmid: 25879677
123. B. K. Tusi *et al*., Population snapshots predict early haematopoietic and erythroid hierarchies. *Nature* **555**, 54–60 (2018). doi: 10.1038/nature25741; pmid: 29466336
124. 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
125. A. Yumine, S. T. Fraser, D. Sugiyama, Regulation of the embryonic erythropoietic niche: A future perspective. *Blood Res.* **52**, 10–17 (2017). doi: 10.5045/br.2017.52.1.10; pmid: 28401096
126. L. M. Ernst, E. D. Ruchelli, D. S. Huff, *Color Atlas of Fetal and Neonatal Histology* (Springer, 2011).
127. L. I. Gardner, Adrenal function in the premature and newborn. *Pediatrics* **17**, 414–417 (1956). pmid: 13297523
128. T. Kalfa, K. E. McGrath, Analysis of Erythropoiesis Using Imaging Flow Cytometry. *Methods Mol. Biol.* **1698**, 175–192 (2018). doi: 10.1007/978-1-4939-7428-3_10; pmid: 29076090
129. J. E. Merrill, Tumor necrosis factor alpha, interleukin 1 and related cytokines in brain development: Normal and pathological. *Dev. Neurosci.* **14**, 1–10 (1992). doi: 10.1159/000111642; pmid: 1350976
130. S.-Y. Park, M.-J. Kang, J.-S. Han, Interleukin-1 beta promotes neuronal differentiation through the Wnt5a/RhoA/JNK pathway in cortical neural precursor cells. *Mol. Brain* **11**, 39 (2018). doi: 10.1186/s13041-018-0383-6; pmid: 29973222
131. P. R. Ormel *et al*., Microglia innately develop within cerebral organoids. *Nat. Commun.* **9**, 4167 (2018). doi: 10.1038/s41467-018-06684-2; pmid: 30301888
132. T. F. Pais *et al*., The NAD-dependent deacetylase sirtuin 2 is a suppressor of microglial activation and brain inflammation. *EMBO J.* **32**, 2603–2616 (2013). doi: 10.1038/emboj.2013.200; pmid: 24013120
133. J. Zhu *et al*., Aryl Hydrocarbon Receptor Promotes IL-10 Expression in Inflammatory Macrophages Through Src-STAT3 Signaling Pathway. *Front. Immunol.* **9**, 2033 (2018). doi: 10.3389/fimmu.2018.02033; pmid: 30283437
134. A. Lapenna, M. De Palma, C. E. Lewis, Perivascular macrophages in health and disease. *Nat. Rev. Immunol.* **18**, 689–702 (2018). doi: 10.1038/s41577-018-0056-9; pmid: 30127389
135. H. Hu *et al*., Prenatal diagnosis of de novo partial trisomy 18p and partial monosomy 18q recurrent in a family with fatal aortic coarctation. *Gene* **517**, 132–136 (2013). doi: 10.1016/j.gene.2012.12.001; pmid: 23232358
136. L. Sanjurjo *et al*., The human CD5L/AIM-CD36 axis: A novel autophagy inducer in macrophages that modulates inflammatory responses. *Autophagy* **11**, 487–502 (2015). doi: 10.1080/15548627.2015.1017183; pmid: 25713983
137. E. Zigmond *et al*., Infiltrating monocyte-derived macrophages and resident kupffer cells display different ontogeny and functions in acute liver injury. *J. Immunol.* **193**, 344–353 (2014). doi: 10.4049/jimmunol.1400574; pmid: 24890723
138. P. Dutta *et al*., Macrophages retain hematopoietic stem cells in the spleen via VCAM-1. *J. Exp. Med.* **212**, 497–512 (2015). doi: 10.1084/jem.20141642; pmid: 25800955
139. J. Kalucka *et al*., Single-Cell Transcriptome Atlas of Murine Endothelial Cells. *Cell* **180**, 764–779.e20 (2020). doi: 10.1016/j.cell.2020.01.015; pmid: 32059779
140. K. P. Gannon, S. E. McKey, D. E. Stec, H. A. Drummond, Altered myogenic vasoconstriction and regulation of whole kidney blood flow in the ASIC2 knockout mouse. *Am. J. Physiol. Renal Physiol.* **308**, F339–F348 (2015). doi: 10.1152/ajprenal.00572.2014; pmid: 25520010
141. X. Lian *et al*., RXFP1 Receptor Activation by Relaxin-2 Induces Vascular Relaxation in Mice *via* a $G\alpha_{i2}$-Protein/PI3Kß/γ/Nitric Oxide-Coupled Pathway. *Front. Physiol.* **9**, 1234 (2018). doi: 10.3389/fphys.2018.01234; pmid: 30233409
142. B. J. Andreone *et al*., Blood-Brain Barrier Permeability Is Regulated by Lipid Transport-Dependent Suppression of Caveolae-Mediated Transcytosis. *Neuron* **94**, 581–594.e5 (2017). doi: 10.1016/j.neuron.2017.03.043; pmid: 28416077
143. L. Chen *et al*., Tbx1 regulates Vegfr3 and is required for lymphatic vessel development. *J. Cell Biol.* **189**, 417–424 (2010). doi: 10.1083/jcb.200912037; pmid: 20439995
144. W. Tao, R. Moore, Y. Meng, E. R. Smith, X.-X. Xu, Endocytic adaptors Arh and Dab2 control homeostasis of circulatory cholesterol. *J. Lipid Res.* **57**, 809–817 (2016). doi: 10.1194/jlr.M063065; pmid: 27005486
145. A. Furlan, M. Lübke, I. Adameyko, F. Lallemend, P. Ernfors, The transcription factor Hmx1 and growth factor receptor activities control sympathetic neurons diversification. *EMBO J.* **32**, 1613–1625 (2013). doi: 10.1038/emboj.2013.85; pmid: 23591430
146. Y.-C. Wang *et al*., Homeodomain transcription factor NKX2.2 functions in immature cells to control enteroendocrine differentiation and is expressed in gastrointestinal neuroendocrine tumors. *Endocr. Relat. Cancer* **16**, 267–279 (2009). doi: 10.1677/ERC-08-0127; pmid: 18987169
147. Y. Kobayashi, P. R. Tata, Pulmonary Neuroendocrine Cells: Sensors and Sentinels of the Lung. *Dev. Cell* **45**, 425–426 (2018). doi: 10.1016/j.devcel.2018.05.009; pmid: 29787707
148. C. Li *et al*., Apc deficiency alters pulmonary epithelial cell fate and inhibits Nkx2.1 via triggering TGF-beta signaling. *Dev. Biol.* **378**, 13–24 (2013). doi: 10.1016/j.ydbio.2013.03.018; pmid: 23562608
149. A. J. Churchill *et al*., Genetic evidence that Nkx2.2 acts primarily downstream of Neurog3 in pancreatic endocrine lineage development. *eLife* **6**, e20010 (2017). doi: 10.7554/eLife.20010; pmid: 28071588
150. L. Arnes, J. T. Hill, S. Gross, M. A. Magnuson, L. Sussel, Ghrelin expression in the mouse pancreas defines a unique multipotent progenitor population. *PLOS ONE* **7**, e52026 (2012). doi: 10.1371/journal.pone.0052026; pmid: 23251675
151. F. M. Gribble, F. Reimann, Enteroendocrine Cells: Chemosensors in the Intestinal Epithelium. *Annu. Rev. Physiol.* **78**, 277–299 (2016). doi: 10.1146/annurev-physiol-021115-105439; pmid: 26442437
152. C. Ramond *et al*., Understanding human fetal pancreas development using subpopulation sorting, RNA sequencing and single-cell profiling. *Development* **145**, dev165480 (2018). doi: 10.1242/dev.165480; pmid: 30042179
153. M. Volante *et al*., Ghrelin expression in fetal, infant, and adult human lung. *J. Histochem. Cytochem.* **50**, 1013–1021 (2002). doi: 10.1177/002215540205000803; pmid: 12133904
154. P. LeSimple *et al*., Trefoil factor family 3 peptide promotes human airway epithelial ciliated cell differentiation. *Am. J. Respir. Cell Mol. Biol.* **36**, 296–303 (2007). doi: 10.1165/rcmb.2006-0270OC; pmid: 17008636
155. Y. Taché, W. Marki, J. Rivier, W. Vale, M. Brown, Central nervous system inhibition of gastric secretion in the rat by gastrin-releasing peptide, a mammalian bombesin. *Gastroenterology* **81**, 298–302 (1981). doi: 10.1016/S0016-5085(81)80061-3; pmid: 7239137
156. R. Kurotani *et al*., Role of secretoglobin 3A2 in lung development. *Am. J. Respir. Crit. Care Med.* **178**, 389–398 (2008). doi: 10.1164/rccm.200707-1104OC; pmid: 18535256
157. M. Hochane *et al*., Single-cell transcriptomics reveals gene expression dynamics of human fetal kidney development. *PLOS Biol.* **17**, e3000152 (2019). doi: 10.1371/journal.pbio.3000152; pmid: 30789893
158. B. Dekel *et al*., Multiple imprinted and stemness genes provide a link between normal and tumor progenitor cells of the developing human kidney. *Cancer Res.* **66**, 6040–6049 (2006). doi: 10.1158/0008-5472.CAN-05-4528; pmid: 16778176
159. T. Moriguchi *et al*., MafB is essential for renal development and F4/80 expression in macrophages. *Mol. Cell. Biol.* **26**, 5715–5727 (2006). doi: 10.1128/MCB.00001-06; pmid: 16847325
160. Y. Maezawa *et al*., Loss of the podocyte-expressed transcription factor Tcf21/Pod1 results in podocyte differentiation defects and FSGS. *J. Am. Soc. Nephrol.* **25**, 2459–2470 (2014). doi: 10.1681/ASN.2013121307; pmid: 24904088
161. S. S. Marable, E. Chung, M. Adam, S. S. Potter, J.-S. Park, Hnf4a deletion in the mouse kidney phenocopies Fanconi renotubular syndrome. *JCI Insight* **3**, e97497 (2018). doi: 10.1172/jci.insight.97497; pmid: 30046000
162. R. O'Rahilly, Early human development and the chief sources of information on staged human embryos. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **9**, 273–280 (1979). doi: 10.1016/0028-2243(79)90068-6; pmid: 400868
163. K. Theiler, *The House Mouse: Atlas of Embryonic Development* (Springer, 2013).
164. K. Reemst, S. C. Noctor, P. J. Lucassen, E. M. Hol, The Indispensable Roles of Microglia and Astrocytes during Brain Development. *Front. Hum. Neurosci.* **10**, 566 (2016). doi: 10.3389/fnhum.2016.00566; pmid: 27877121
165. J. Dai, K. K. Bercury, J. T. Ahrendsen, W. B. Macklin, Olig1 function is required for oligodendrocyte differentiation in the mouse brain. *J. Neurosci.* **35**, 4386–4402 (2015). doi: 10.1523/JNEUROSCI.4962-14.2015; pmid: 25762682
166. J. See *et al*., Oligodendrocyte maturation is inhibited by bone morphogenetic protein. *Mol. Cell. Neurosci.* **26**, 481–492 (2004). doi: 10.1016/j.mcn.2004.04.004; pmid: 15276151
167. K. Komorowska *et al*., Hepatic Leukemia Factor Maintains Quiescence of Hematopoietic Stem Cells and Protects the Stem Cell Pool during Regeneration. *Cell Rep.* **21**, 3514–3523 (2017). doi: 10.1016/j.celrep.2017.11.084; pmid: 29262330
168. H. Huang, Y. Li, B. Liu, Transcriptional regulation of mast cell and basophil lineage commitment. *Semin. Immunopathol.* **38**, 539–548 (2016). doi: 10.1007/s00281-016-0562-4; pmid: 27126100
169. J. Medvedovic, A. Ebert, H. Tagoh, M. Busslinger, Pax5: A master regulator of B cell development and leukemogenesis. *Adv. Immunol.* **111**, 179–206 (2011). doi: 10.1016/B978-0-12-385991-4.00005-2; pmid: 21970955

170. C. Cantù *et al*., Sox6 enhances erythroid differentiation in human erythroid progenitors. *Blood* **117**, 3669–3679 (2011). doi: 10.1182/blood-2010-04-282350; pmid: 21263153
171. Y. Lavin, A. Mortha, A. Rahman, M. Merad, Regulation of macrophage development and function in peripheral tissues. *Nat. Rev. Immunol.* **15**, 731–744 (2015). doi: 10.1038/nri3920; pmid: 26603899
172. X. Qiu *et al*., Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017). doi: 10.1038/nmeth.4402; pmid: 28825705
173. A. Buttgereit *et al*., Sall1 is a transcriptional regulator defining microglia identity and function. *Nat. Immunol.* **17**, 1397–1406 (2016). doi: 10.1038/ni.3585; pmid: 27776109
174. I. R. Holtman, D. Skola, C. K. Glass, Transcriptional control of microglia phenotypes in health and disease. *J. Clin. Invest.* **127**, 3220–3229 (2017). doi: 10.1172/JCI90604; pmid: 28758903
175. P. Mazzarello, A unifying concept: The history of cell theory. *Nat. Cell Biol.* **1**, E13–E15 (1999). doi: 10.1038/8964; pmid: 10559875
176. M. Asp *et al*., A Spatiotemporal Organ-Wide Gene Expression and Cell Atlas of the Developing Human Heart. *Cell* **179**, 1647–1660.e19 (2019). doi: 10.1016/j.cell.2019.11.025; pmid: 31835037
177. D. E. Wagner *et al*., Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* **360**, 981–987 (2018). doi: 10.1126/science.aar4362; pmid: 29700229
178. J. A. Briggs *et al*., The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science* **360**, eaar5780 (2018). doi: 10.1126/science.aar5780; pmid: 29700227
179. N. Karaiskos *et al*., The *Drosophila* embryo at single-cell transcriptome resolution. *Science* **358**, 194–199 (2017). doi: 10.1126/science.aan3235; pmid: 28860209
180. P. Blakeley *et al*., Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Development* **142**, 3613 (2015). doi: 10.1242/dev.131235; pmid: 26487783
181. B. Pijuan-Sala *et al*., A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature* **566**, 490–495 (2019). doi: 10.1038/s41586-019-0933-9; pmid: 30787436
182. P. Wang *et al*., Dissecting the Global Dynamic Molecular Profiles of Human Fetal Kidney Development by Single-Cell RNA Sequencing. *Cell Rep.* **24**, 3554–3567.e3 (2018). doi: 10.1016/j.celrep.2018.08.056; pmid: 30257215
183. J. M. Segal *et al*., Single cell analysis of human foetal liver captures the transcriptional profile of hepatobiliary hybrid progenitors. *Nat. Commun.* **10**, 3350 (2019). doi: 10.1038/s41467-019-11266-x; pmid: 31350390
184. S. Gao *et al*., Tracing the temporal-spatial transcriptome landscapes of the human fetal digestive tract using single-cell RNA-sequencing. *Nat. Cell Biol.* **20**, 721–734 (2018). doi: 10.1038/s41556-018-0105-4; pmid: 29802404
185. J. S. Banerji, K. M. Ramani, A. Devasia, Extramedullary hematopoiesis in the adrenal: Case report and review of literature. *Can. Urol. Assoc. J.* **7**, E436–E438 (2013). doi: 10.5489/cuaj.1389; pmid: 23826059
186. H. Al-Thani, M. Al-Sulaiti, G. El-Mabrok, A. Tabeb, A. El-Menyar, Adrenal extramedullary hematopoiesis associated with beta-thalassemia trait in an adult woman: A case report and review of literature. *Int. J. Surg. Case Rep.* **24**, 83–87 (2016). doi: 10.1016/j.ijscr.2016.05.024; pmid: 27232291
187. J. Cao, J. Shendure, sci-RNA-seq3, protocols.io (2019); https://dx.doi.org/10.17504/protocols.io.9yih7ue.
188. 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
189. P. J. Thul, C. Lindskog, The human protein atlas: A spatial map of the human proteome. *Protein Sci.* **27**, 233–244 (2018). doi: 10.1002/pro.3307; pmid: 28940711
190. S. Aibar *et al*., SCENIC: Single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017). doi: 10.1038/nmeth.4463; pmid: 28991892
191. S. Amini *et al*., Haplotype-resolved whole-genome sequencing by contiguity-preserving transposition and combinatorial indexing. *Nat. Genet.* **46**, 1343–1349 (2014). doi: 10.1038/ng.3119; pmid: 25326703
192. J. Cao, sci-RNA-seq3 processing pipeline, Zenodo (2020); http://doi.org/10.5281/zenodo.4013713.

ACKNOWLEDGMENTS

We thank past and present members of the Shendure and Trapnell laboratories, A. Adey, the Brotman Baty Institute Advanced Technology Lab, and the University of Washington Birth Defects Research Laboratory. We thank C. Spurrell and D. Ahrendsen for reviewing the sci-RNA-seq3 protocol. **Funding:** M.S. was supported by grants from the Deutsche Forschungsgemeinschaft (SP1532/3-1, 4-1, and 5-1). Aspects of this work were supported by funding from the Brotman Baty Institute for Precision Medicine, the Paul G. Allen Frontiers Foundation (Allen Discovery Center grant to J.S. and C.T.), the Chan Zuckerberg Initiative (to C.T.), and the NIH (DK119285 to J.P.; HD000836 to I.A.G.). J.S. is an investigator of the Howard Hughes Medical Institute. **Author contributions:** J.C. developed techniques and performed sci-RNA-seq3 experiments with assistance from R.M.D., R.B., F.Z., and F.J.S.; D.R.O. and M.D. performed tissue collection, nuclei extraction, and immunohistochemistry staining under the supervision of D.D. and I.A.G.; K.A.A. performed genotyping; P.D.K. and J.P. performed erythropoiesis validation in mice; J.C. performed computational analyses with assistance from H.A.P. and M.S.; M.A.Z. developed the website with assistance from J.C.; J.C. and J.S. wrote the manuscript with input from all coauthors; and J.S. and C.T. supervised the project. **Competing interests:** F.Z. and F.J.S. declare competing financial interests in the form of stock ownership and paid employment by Illumina, Inc. J.S. has competing financial interests (paid consulting and/or equity) with Guardant Health, Maze Therapeutics, Camp4 Therapeutics, Nanostring, Phase Genomics, Adaptive Biotechnologies, and Stratos Genomics. One or more embodiments of one or more patents and patent applications filed by Illumina and the University of Washington may encompass the methods, reagents, and data disclosed in this manuscript. The authors declare no other competing interests. **Data and materials availability**: A detailed version of the sci-RNA-seq3 protocol is available on protocols.io (*187*). All methods for making the transposase complexes are described in (*191*), and Tn5 transposase is available from Lucigen; however, Illumina will provide transposase complexes in response to reasonable requests from the scientific community, subject to a material transfer agreement. A demultiplexing script and tutorial are provided on Zenodo (*192*). Raw data are provided at dbGaP (accession no. phs002003.v1.p1). Supplementary files S1 to S7 and table S5 (which is also provided in the supplementary materials) are provided at GEO (accession no. GSE156793). Garnett models generated from these data are also posted to our website, where they can broadly be used for the automated cell type classification of single-cell data from diverse organs. An interactive version of the full dataset facilitates the exploration of these data by tissue, cell type, or gene (descartes.brotmanbaty.org).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6518/eaba7721/suppl/DC1
Supplementary Materials and Methods
Figs. S1 to S27
Tables S1 to S16
References (*193*–*207*)
MDAR Reproducibility Checklist

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

3 January 2020; accepted 10 September 2020
10.1126/science.aba7721

Science

A human cell atlas of fetal gene expression

Junyue Cao, Diana R. O'Day, Hannah A. Pliner, Paul D. Kingsley, Mei Deng, Riza M. Daza, Michael A. Zager, Kimberly A. Aldinger, Ronnie Blecher-Gonen, Fan Zhang, Malte Spielmann, James Palis, Dan Doherty, Frank J. Steemers, Ian A. Glass, Cole Trapnell and Jay Shendure

Science **370** (6518), eaba7721.
DOI: 10.1126/science.aba7721

The genomics of human development

Understanding the trajectory of a developing human requires an understanding of how genes are regulated and expressed. Two papers now present a pooled approach using three levels of combinatorial indexing to examine the single-cell gene expression and chromatin landscapes from 15 organs in fetal samples. Cao *et al.* focus on measurements of RNA in broadly distributed cell types and provide insights into organ specificity. Domcke *et al.* examined the chromatin accessibility of cells from these organs and identify the regulatory elements that regulate gene expression. Together, these analyses generate comprehensive atlases of early human development.

Science, this issue p. eaba7721, p. eaba7612

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.