Single-cell transcriptome profiling an adult human cell atlas of 15 major organs

Shuai He[1,2,3,4,#], Lin-he Wang[2,3,4,#], Yang Liu[1,#], Yi-qi Li[1], Haitian Chen[2,3,4], Jinghong Xu[2,3,4], Wan Peng[1], Guo-Wang Lin[1], Pan-Pan Wei[1], Bo Li[5,6], Xiaojun Xia[1], Dan Wang[1], Jin-Xin Bei[1,7,*], Xiaoshun He[2,3,4,*], Zhiyong Guo[2,3,4,*]

[1]Sun Yat-sen University Cancer Centre, State Key Laboratory of Oncology in South China, Collaborative Innovation Centre for Cancer Medicine, Guangzhou 510060, P. R. China.

[2]Organ Transplant Center, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou 510080, China

[3]Guangdong Provincial Key Laboratory of Organ Donation and Transplant Immunology, Guangzhou 510080, China

[4]Guangdong Provincial International Cooperation Base of Science and Technology (Organ Transplantation), Guangzhou 510080, China

[5]Department of Biochemistry and Molecular Biology, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou 510080, P. R. China

[6]RNA Biomedical Institute, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou 510120, P. R. China

[7]Centre for Precision Medicine, Sun Yat-sen University, Guangzhou 510080, P. R. China.

Email addresses of authors:

Shuai He, heshuai@sysucc.org.cn; Lin-he Wang, wanglh6@mail2.sysu.edu.cn;

Yang Liu, liuyang3@sysucc.org.cn; Yi-qi Li, liyq2@sysucc.org.cn;

Haitian Chen, chenht3@mail2.sysu.edu.cn; Jinghong Xu, xujh38@mail2.sysu.edu.cn;

Wan Peng, pengwan@sysucc.org.cn; Guo-Wang Lin, lingw@sysucc.org.cn;

Pan-Pan Wei, weipp@sysucc.org.cn; Bo Li, libo47@mail.sysu.edu.cn;

Xiaojun Xia, xiaxj@sysucc.org.cn; Dan Wang, wangdan@sysucc.org.cn;

Jin-Xin Bei, beijx@sysucc.org.cn; Xiaoshun He, gdtrc@163.com;

Zhiyong Guo, rockyucsf1981@126.com

[#]These authors contributed equally.

*Correspondence should be addressed to:

beijx@sysucc.org.cn (JXB), or gdtrc@163.com (XH), or rockyucsf1981@126.com (ZG).

ABSTRACT

Background: As core units of organ tissues, cells of various types play their harmonious rhythms to maintain the homeostasis of human body. It is essential to identify characteristics of the cells in human organs and their regulatory networks for understanding biological mechanisms related to health and disease. However, a systematic and comprehensive single-cell transcriptional profile across multiple organs of normal human adults has been pending.

Results: we performed single-cell transcriptomes of 88,622 cells derived from 15 tissue organs of one adult donor and generated an adult human cell atlas (AHCA). The AHCA depicted 234 subtypes of cells, including major cell types such as T, B, myeloid, epithelial, and stromal cells, as well as novel cell types in skin, each of which was distinguished by multiple marker genes and transcriptional profiles and collectively contributed to the heterogeneity of major human organs. Moreover, TCR and BCR repertoire comparison and trajectory analyses revealed direct clonal sharing of T and B cells with various developmental states among different tissues. Furthermore, novel cell markers, transcription factors and ligand-receptor pairs were identified with potential functional regulations on maintaining the homeostasis of human cells among tissues.

Conclusions: The AHCA reveals the inter- and intra-organ heterogeneity of cell characteristics and provides a useful resource to uncover key events during the development of human diseases such as the recent outbreak of coronavirus disease 2019 (COVID-19) in the context of heterogeneity of cells and organs.

Keywords

Human cell atlas, single-cell RNA sequencing, TCR, BCR, transcriptome

INTRODUCTION

Human body consists of multiple organs, where cells of different types are the core units of structure and function. Like instruments from different families in a symphony orchestra, cells and organs play their harmonious rhythms to maintain the homeostasis of human body. Perturbation of the homeostasis leads to various pathological conditions. Therefore, it is essential to identify characteristics of the cells in human organs and their regulatory networks for understanding biological mechanisms related to health and disease.

Recent technological innovations in the single-cell transcriptional profiling using single-cell RNA sequencing (scRNA-seq) provide a promising strategy to quantify gene expression at the genome-wide level in thousands of individual cells simultaneously[1-3], which have expanded our knowledge regarding cellular heterogeneity and networks as well as development in human tissues and organs at single-cell resolution[4-11]. Previous studies have demonstrated cell composition for many tissues of human and mouse, such as brain[12], kidney[13], lung[14], skin[15]. The strategy empowers the identification of novel cell types. A cell type marked by cystic fibrosis transmembrane conductance regulator (*CFTR*) was identified in the lung of human and mouse, which has ability to regulate luminal pH and was implicated in the pathogenesis of cystic fibrosis[16]. Non-genetic cellular heterogeneity has been revealed in hematopoietic progenitor cells and keratinocytes, which play important roles in maintaining hematopoiesis[17] and compartmentalizing crucial molecular activities in human epidermis[15], respectively. For the development of human embryo, transcriptome analysis of about 70,000 single cells from the first-trimester placentas with matched maternal blood and decidual cells uncover the cellular organization of decidua and placenta, as well as distinctive immunomodulatory and chemokine profiles of decidual nature killer (NK) cells[18]. In addition, the single-cell transcriptional profiles of embryonic and adult organs in mice have been reported, which reveal the landscape of organogenesis and cellular heterogeneity in organs[8, 9, 19]. However, a systematic and comprehensive single-cell transcriptional profile across multiple organs of normal human adults has been pending. Previous studies with scRNA-seq on human samples were mostly restricted to a few specific organs with disease conditions and did not attempt to characterize the heterogeneity and connexity among multiple organs in the same individuals.

Here we aimed to investigate the transcriptional heterogeneity and interactions of cells from human adult organs at single-cell resolution. Using single-cell RNA sequencing, we profiled the transcriptomes of more than 88,000 cells of 15 organs from one individual donor. Comprehensive comparisons within and across tissues for distinct cell types were performed to reveal the complexity of inter-cells in gene profiles, active transcription factors and potential biological functions. The resulting high-resolution adult human cell atlas (AHCA) provides a global view of various cell populations and connexity in human body, and also a useful resource to investigate the biology of normal human cells and the development of diseases affecting different organs such as the distribution of potential viral receptor for the most recent outbreak of coronavirus disease 2019 (COVID-19).

RESULTS

Global view of single cell RNA sequencing of 15 organ samples

Viable single cells were prepared from tissue samples of 15 different organs of a research-consented adult donor (**Fig. 1A and Additional file 19: Table S1**). mRNA transcripts from each sample were ligated with barcoded indexes at 5'-end and reverse transcribed into cDNA, using GemCode technology (10X Genomics, USA). cDNA libraries including the enriched fragments spanning the full-length V(D)J segments of T cell receptors (TCR) or B cell receptor (BCR), and 5'-end fragments for gene expression were separately constructed, which were subsequently subjected for high-throughput sequencing. In average, we obtained more than 400 million sequencing reads for each organ sample, which resulted in the median of sequencing saturation (covering the fraction of library complexity) at 88% (61.6%-97%) for each sample (**Additional file 20: Table S2**). After primary quality control (QC) filters, 91,393 cells were identified (**Additional file 20: Table S2**). Higher unique molecular indices (UMIs) and genes were detected in the skin and trachea samples (with median UMIs of 4,022.5 and 4,100.5 and genes of 1,528 and 1,653, respectively) compared to the other organs, indicating their higher transcriptional activation (**Fig. 1B, Additional file 1: Figure S1, and Additional file 20: Table S2**). The cells in each organ were classified using unsupervised clustering and cell types were assigned based on canonical marker genes[20] (**Additional file 21: Table S3**). Next, visualization of the cells by t-distributed

stochastic neighbor embedding (t-SNE) revealed multiple subpopulations of cells in each organ, with numbers of clusters ranging from 9 in the blood to 23 in the skin (**Additional file 2: Figure S2, Additional file 3: Figure S2 continued, and Additional file 21: Table S3**). Clusters due to cell doublets or unassigned cell types were identified and excluded manually in each organ, which resulted in a total of 88,622 cells for the downstream analysis (**Additional file 21: Table S3**).

With transcriptional profiles of such large number of cells, we identified some rare and novel cell populations. For instance, a group of Langerhans cells was observed in the skin sample (1% of all cells in the sample), which specifically expressed *CD207* and *CD1A*[21]. A fibroblast subtype with high expression of *COCH* gene (fold change, FC > 116), and a very small group of 26 epithelial cells (0.31%) that specifically highly expressed *SCGB2A2, SCGB1B2P* and *PIP* genes (FC > 546, 392 and 179, respectively, compared to all other cells; **Additional file 4: Figure S3A**) were identified in the skin sample. Both *SCGB2A2* and *PIP* genes have been found specifically expressed in breast tissue and over-expressed in breast cancer[22]. However, the functions of these genes in skin epithelial cells await further investigations.

We combined all the 88,622 cells in clustering analysis and identified 43 clusters in 15 organs (**Additional file 4: Figure S3B**). We observed close clustering of cells from different organs (more than seven organs) for major cell types, including T cells, B cells, NK cells, fibroblasts, endothelial cells, smooth muscle cells, macrophages and monocytes (**Fig. 1C, Additional file 4: Figure S3B, and Additional file 5: Figure S4**). This is consistent with the understanding that cells derived from a same linage distribute across human body, especially for the circulating immune cells. Moreover, multiple clusters were further identified for several major cell types (T cells, B cells, fibroblasts, myeloid cells and endothelial cells), reflecting their heterogenous transcriptional profiles (**Additional file 6: Figure S5A, B**).

The heterogeneity of T cells in development state and clonalities over the body

In this study, we identified a total of 20,494 T cells that prevailing in the immune cells of most tissues including bladder, common bile duct, liver, marrow, muscle, rectum, small intestine, spleen and stomach (**Additional file 22: Table S4**), which is consistent with a previous

finding[23]. The cells were divided into $CD4^+$ (7,122) and $CD8^+$ (13,372) T cells according their gene profiles and further grouped into 11 and 22 major unsupervised clusters, respectively (**Fig. 2A, B**), including naive T (T_N), central memory T (T_{CM}), effector memory T (T_{EM}), regulatory T (T_{reg}), tissue-resident memory T (T_{RM}), effector T (T_{EFF}), intraepithelial lymphocytes (IEL) T, mucosal-associated invariant T (MAIT) and γδ T cells, based on known markers[24]. Both $CD4^+$ and $CD8^+$ T cell clusters showed a distribution pattern in organ-specific manner (**Additional file 7: Figure S6A, B**) and each of them had differentially expressed genes (**Additional file 7: S6C, D and Additional file 22: Table S4**). Overlapping of clusters was also observed between organs (such as blood and marrow), suggesting shared T cell subtypes (**Additional file 7: Figure S6A, B**).

To better understand the development state of T cells, we performed trajectory analyses of $CD4^+$ and $CD8^+$ T cells. We observed that the trajectory tree rooted from T_N cells, sprouting into T_{EFF}, T_{CM} and T_{RM} branches for $CD4^+$ T cells and T_{EFF}, T_{CM}, IEL, γδ and T_{RM} branches for $CD8^+$ T cell (**Fig. 2C and Additional file 7: Figure S6E, F**). T_{RM} cells and IEL cells with higher pseudo-time scores were found in $CD4^+$ and $CD8^+$ T cells, respectively, suggesting their terminal development state (**Fig. 2C and Additional file 7: Figure S6E, F**). Moreover, several T_{RM} clusters at the end of other branches with mediate scores for both $CD4^+$ and $CD8^+$ T cells, indicating the middle state of these clusters. In addition, the clusters with different states showed an organ specific pattern (**Additional file 7: Figure S6E, F**). These observations reveal the existence of heterogeneity in the developmental states of both $CD4^+$ and $CD8^+$ cells in human organs.

Transcription factors (TFs) have been demonstrated as important regulators of gene expression and to shape different phenotype of T cells[25]. We therefore performed Single-Cell Regulatory Network Inference and Clustering (SCENIC) analysis to assess TFs underlying differential gene expression in T cells. We identified well-defined and cell subtype-specific TFs for $CD4^+$ (**Fig. 2D**) and $CD8^+$ (**Fig. 2E**) clusters (**Additional file 22: Table S4**), such as higher activity of *FOXP3*[25] and *NFATC1[26]* in T_{reg} cells*,* upregulation of *LEF1*, *MYC*, *TCF7*[27] and *KLF2*[28] in T_N cells, *TBX21*, *STAT1 and IRF1* in T_{EFF} cells[25, 29, 30] (**Fig. 2D, E**). In addition, many other poorly investigated TFs were also observed in $CD4^+$ and $CD8^+$ clusters, such as upregulation of *FOSB*, *JUN*, *ID2*, *ATF4*, *FOSL2*, *JUND* in $CD4^+$

T_{RM} cells, and higher activity of *FOS*, *JUND*, *ATF3*, *JUN*, *JUNB*, *GTF2B* and *CEBPD in* $CD8^+$ T_{RM} *cells*. Collectively, these results indicate that the multiple TFs combinations regulate T cells development to maintain heterogeneous states of $CD4^+$ and $CD8^+$ T cells.

To better investigate the clonalities and dynamic relationships among T cell subtypes across tissues, we preformed TCR clonal typing accompanied with transcriptome analysis (**Fig. 1E**). After stringent QC filters, we identified 4,150 TCR clonotypes with unique heterodimer α and β chains among 6,203 T cells (30.3% of the whole 20,494 T cells), including 2,301 $CD4^+$ T cells and 3,902 $CD8^+$ T cells. Among them, 3,728 cells (1,641 $CD4^+$ and 2,087 $CD8^+$ T cells) had a unique TCR clonotype for each, while the remaining 2,475 cells (214 $CD4^+$ and 2,261 $CD8^+$ T cells) shared two or more of the 422 TCR clonotypes (**Additional file 22: Table S4**). We observed similar numbers of V and J segments for TCR α chain in both $CD4^+$ and $CD8^+$ T cells, which shared 50% and 40% of the top 10 frequent V (*TRAV12-1*, *TRAV21*, *TRAV29/DV5*, *TARV13-1* and *TRAV12-2*) and J (*TRAVJ49*, *TRAVJ48*, *TRAVJ9*, and *TRAVJ39*) segments, respectively. By contrast, the diversity of V segment was much higher than that of J segment for β chain in both $CD4^+$ and $CD8^+$ T cells, which shared 40 % and 80% of the top 10 frequent V (*TRBV6-5*, *TRBV20-1*, *TRBV7-2*, and *TRBV2*) and J (*TRBJ2-1*, *TRBJ2-7*, *TRBJ1-5*, *TRBJ1-1*, *TRBJ2-3*, *TRBJ1-2*, *TRBJ2-5*, and *TRBJ2-2*) segments, respectively (**Additional file 8: Figure S7A, B**). Although more $CD8^+$ T cells were detected than $CD4^+$ T cells among all tissues, no significant difference in clone sizes (numbers of unique clonotypes) was observed between the two cell populations (**Additional file 8: Figure S7C**). Singular cells of unique TCR clonotypes were prevalent for $CD4^+$ T cells among tissues, except that higher proportion of multiple cells with identical TCR clonotypes or clonal expansions were observed in the stomach and marrow (**Additional file 8: Figure S7D top panel**). Clonal expansions were much commonly found for $CD8^+$ T cells in all tissues (**Additional file 8: Figure S7D bottom panel**).

To investigate clonotype distribution of T cells across tissues, we evaluated the ability of sharing TCR for each tissue with others. We observed more intensive and broader sharing of TCR clonotype for $CD8^+$ than $CD4^+$ T cells across tissues (**Fig. 2F**). Higher migration capacity as reflected by migration-index score of tissues was found in $CD8^+$ T cells than $CD4^+$ T cells (**Fig. 2G top panel**). Moreover, higher expansion ability but lower diversity was observed in

$CD8^+$ T cells in each tissue compared to $CD4^+$ T cells (**Fig. 2G bottom panel and Additional file 8: Figure S7E**). Cells with clonal expansion had considerable proportions in T_{EM}, T_{EFF}, T_{RM} and IEL clusters of $CD8^+$ T cells (**Additional file 8: Figure S7F bottom panel**), and T_{EFF} and *TRBV12-3*_T_{RM} of $CD4^+$ T cells (**Additional file 8: Figure S7F top pane**l). We further evaluated the clonal expansion and transition (clonotype sharing ability of each subpopulation) of each T cell cluster, which revealed significantly stronger expansion and transition ability of $CD8^+$ compared with $CD4^+$ T cells (**Fig. 2H and Additional file 8: Figure S7G**). This is consistent with the finding of stronger links between subtypes of $CD8^+$ than $CD4^+$ T cells (**Fig. 2F and Additional file 8: Figure S7H, I**). Moreover, we observed that the clonal expansion of certain $CD8^+$ T cells (T_{EFF}, T_{RM}, T_{EM}, γδ and IEL T cells) distributed across multiple tissues with different development states (**Fig. 2I and Additional file 8: Figure S7I**), suggesting that most of these cell types recognizing the same antigens propagated and migrated more intensively than other cell types. In addition, we observed consistent results of TCR diversity using β chains (**Additional file 8: Figure S7J-I**).

The heterogeneity of B cells and plasma cells

Clustering analysis revealed 15 distinct cell clusters among 10,655 B and plasma cells from 11 organ samples, including nine B cell (*CD20*) and six plasma cell (*SDC1*) clusters, respectively (**Fig. 3A**). We observed the predominance of B cells over plasma cells in all tested organs except for esophagus and rectum (**Fig. 3B**). Differential gene expression (DEG) analysis revealed that B cells exhibited distinct gene profiles from plasma cells (**Additional file 9: Figure S8A and Additional file 23: Table S5**). Although *CD27* is as a canonical marker of memory B cells[31], we observed low transcription level of this gene in memory B cells, but higher transcription level in plasma cells (**Fig. 3C**). *TCL1A* was significantly expressed in two naive B cell clusters with distinct gene expression profiles, one from lymph nodes (TCL1A_ly_naive_B) and the other (TCL1A_naive_B) from multiple tissues, compared to the other B and plasma cell clusters (**Fig. 3A and Additional file 9: Figure S8A**). Moreover, *TCL1A* was exclusively expressed in non-CD27-expressing B cell clusters and was highly reversely correlated with *CD27* transcription (Person's R = -0.83, P = 0; **Fig. 3C and Additional file 9: Figure S8B**). *TCL1A* has been reported as an important gene in B

cells lymphoma[32]. These suggest that *TCL1A* might be a novel marker of naive B cells.

B cells are professional antigen-presenting cells (APCs) with high expression of *CIITA* and *MHC* class II genes, which are silenced during the differentiation to plasma cells[33]. Consistently, plasma cells had much lower presenting-antigen score (PAS) compared to B cell clusters (**Additional file 9: Figure S8C**), suggesting different abilities to present antigen between the two cell types. We also investigated the potential biological function of different cell clusters using GO analysis. The potential biological functions of B cells were found be enriched in immune response (for example, *responding to activation of immune*) and that of plasma cells were associated with protein synthesis (including *signal peptide processing*, *protein localization to endoplasmic reticulum* and *protein folding*; **Fig. 3D**). In addition, gene set variation analysis (GSEA) showed that TCL1A_naive_B was enriched in *oxidative phosphorylation* and *fatty acid metabolism* (**Additional file 9: Figure S8D**).

TFs play a critical role in the differentiation of B cells to plasma cells, engaging B cells with effector or memory functions[34]. SCENIC analysis revealed that TFs exhibited similar activities within B and plasma cells but with distinct patterns between the two populations (**Additional file 23: Table S5**). TFs with higher activity were found in B cells, including *MYC[35]* and *REL*[36], which have been known to modulate B cell development, as well as other TFs, such as EGF receptors (*EGFR1/2/3*), which has not been characterized in B cells. Likewise, TFs enriched in plasma cells included *PRDM1*, *XBP1*, *FOS* and *IRF4*, which play important roles in the development of plasma cells[34]. In addition, many TFs with unknown roles were found, including *CREB3L2*, *ATF4*, *ATF3* and *FOXO3*. Taken together, these results suggest that various TFs regulate the development of B cells to plasma cells.

To explore the clonalities of B and plasma cell clusters across organ tissues, we performed a single-cell BCR sequencing analysis. After stringent QC filters, 4,001 of 10,655 cells were assigned with 3,844 clonotypes, among which 3,654 clonotypes were presented on singular cells and 190 were presented by multiple cells (**Additional file 23: Table S5**). We observed various usage of V and J gene segments for both heavy and light chains of immunoglobulin genes, with preferred usage of some particular V and J segments (**Additional file 9: Figure S8E-G**). Unlike T cells, clonal diversity was common for B cells among all organs, while clonal expansion of B cells was limited and restricted in the spleen

and stomach (**Additional file 9: Figure S8H**). Moreover, less sharing of BCR clonotypes between B and plasma cells was observed across organs compared to T cells (**Fig. 2F and Fig. 3F**). In addition, BCR analysis showed lower expansion and transition abilities of B cells compared to plasma cells and T cells (**Fig. 2H, Fig. 3G, and Additional file 9: Figure S8I**).

The heterogeneity of myeloid cells and the origin of macrophages

We obtained 5,605 myeloid cells from 14 organ tissues, which were grouped into 17 distinct clusters (**Fig. 4A, B**). Based on the differential expressions of marker genes, we further identified seven monocyte clusters, nine macrophage clusters, and one small dendritic cell (DC) cluster (**Fig. 4A, B, and Additional file 10: Figure S9A**). The hierarchical clustering analysis revealed that all classical monocytes were closely related as a branch node, while the intermediate and non-classical monocytes were closely related to macrophages (**Fig. 4C**). DEG analysis showed that each myeloid cell cluster had a specific gene signature, suggesting the inter-cell heterogeneity among monocytes, macrophages and DC cells (**Fig. 4D and Additional file 24: Table S6**). All clusters contained cells from multiple tissues, except for FN1_Intermediate_Mon (C5) and CCL20_Intermediate_Mon (C10) in the rectum and Langerhans clusters (C16) in the skin, suggesting similar transcriptional profiles and origins of these myeloid cells **(Fig. 4A, B**). Monocytes showed predominance in 10 of the 14 tested tissues, except for the esophagus, heart, small intestine and skin, where macrophages accounted for more than 50% of all myeloid cells (**Fig. 4B**).

TFs such as *SPI1* are involved in the development of monocytes to macrophages[37, 38]. However, how TFs regulate the development in normal human body is still unclear. TF activity analysis revealed similar activation of certain TFs among the four classical monocytes, which is consistent with their close similarity in gene signatures (**Fig. 4D, E and Additional file 24: Table S6**). Higher activation of *SPI1* were determined in cell clusters of classical monocytes, non-classical monocytes and LIPA_macrophage, suggesting its involvement in the development of these cells. We also identified several TFs with higher activity in the four classical monocyte clusters, including *NFIL3*, *RXRA*, *ELF2*, *ETV6*, *ELK3*, *NFE2* and *EP300*. Their roles in maintaining the classical state of monocytes have yet to be investigated. Non-classical monocytes and macrophages shared several active TFs, such as *TCF7L2*,

STAT1, *KLF11*, *ELF1*, *NR1H3*, *SPIC* and *HDAC2*, which is consistent with their close pattern in the hierarchical clustering. We observed multiple not-well-characterized TFs activated in the intermediate, non-classical monocytes, macrophages and DCs other than classical monocytes, including *JUN*, *KLF2*, *PRDM1*, *EGFR1 and ATF3*. We also observed cluster-specific TFs for myeloid cell clusters **(Fig. 4E and Additional file 24: Table S6**). For instance, activation of *ZNF217* and *NR1H2* was higher in C14 macrophages. These results indicated that TFs combinations help shape the different states of myeloid cells in a steady state.

Considering the previous understanding that tissue-resident macrophages are originated from circulating monocytes[39], we examined the connection between macrophages and monocytes. First, we observed the coexistence of macrophages and monocytes in the same organs (**Fig. 4A, B**). Second, trajectory analysis revealed that the classical monocytes had initial state at the root, sprouting into branches with more developed states of monocytes and then macrophages either directly or via the two intermediate monocytes (**Fig. 4F and Additional file 10: Figure S9B, C**). Notably, high expression of proliferation markers including *MKI67*, *PCNA* and *TYMS* were exclusively detected in the C12 macrophages (mainly from the rectum) and Langerhans cells (C16, skin; **Additional file 10: Figure S9D**). GSEA analysis also revealed that most of the macrophage clusters had higher enrichment score of *G2M checkpoint and E2F targets* pathways (**Additional file 10: Figure S9E**). Additional trajectory analysis of macrophages in the rectum showed that C12 positioned in the middle position in the development of other clusters, suggesting that it was the origin of these clusters (**Additional file 10: Figure S9F**). Taken together, these observations suggest that circulating monocytes give rise to the macrophages in the organs and local microenvironment also contributes to the expansion and proliferation of different macrophage populations.

It has been demonstrated that myeloid cells could act as professional APCs, with the strongest antigen-presenting ability for DCs[40]. We observed various antigen-presenting abilities as reflected by PAS for different myeloid clusters (**Additional file 10: Figure S9G**). Langerhans cells had stronger antigen-presenting ability than other macrophages and monocytes ($P < 2.2 \times 10^{-16}$; **Additional file 10: Figure S9G**). In average, the classical

monocytes had the lowest PAS among myeloid cell clusters ($P < 2.2 \times 10^{-16}$). Surprisingly, DCs had similar PAS with macrophages ($P = 0.16$), which might be due to the limited number of DCs included in the analysis.

The similarity and heterogeneity of epithelial cells intra- and inter-tissues

We obtained a total of 18,090 epithelial cells from nine tested organ tissues (**Fig. 5A**). Differential gene expression analysis revealed a clear distinct pattern of gene expression among the tissue samples (**Additional file 11: Figure S10A and Additional file 25: Table S7**). There were 189 genes with tissue-specific expression (FC ≥ 5, pct.1 ≥ 0.2; **Additional file 25: Table S7 and Additional file 11: Figure S10A**), indicating the heterogeneity of epithelial cells among organ tissues. GO analysis also revealed various biological function of epithelial cells from different tissues, among which *epithelial cell differentiation*, *humoral immune response* and *regulation of myeloid leukocyte activation* were common pathways enriched in the majority of epithelial cells, suggesting that the cells share common functions in epithelial cell development and regulating immune response (**Additional file 11: Figure S10B**).

The epithelial cells were further grouped into 33 clusters (**Fig. 5A**). Grouping of close clusters was seen in all organ tissues, expect for C31 (liver and common bile duct) and C29 (the stomach, trachea and small intestine), suggesting the presence of heterogeneity in epithelial cells at both intra- and inter-organ contexts. The 33 clusters showed different expression profiles with 375 signature genes (FC ≥ 5, pct.1 ≥ 0.2 and pct.2 ≤ 0.2), most of which were expressed exclusively in one cluster (**Fig. 5B, Additional file 11: Figure S10C, D, and Additional file 25: Table S7**. Hierarchical clustering analysis revealed closer grouping of cell clusters within tissues than cross tissues (**Fig. 5C**). Cells from digestive organs including the small intestine, stomach, rectum and common bile duct were clustered closely, as were cells from non-digestive tissues including the skin, trachea and bladder (**Fig. 5C**). Although the esophagus is a digestive organ, the epithelial cells were clustered much closer to the skin cells than cells from digestive organs.

To explore the potential functions of cell from each cluster, we preformed GO analysis for two groups of cells from the 13 digestion-related clusters and the 20 non-digestion-related

clusters, separately. For the cells of digestive system, biological functions related to metabolic process, energy synthesis and digestion pathways were commonly observed (**Fig. 5D**). For the non-digestion-related cells, *cell cycle regulation* and *epidermis development* were the strongly enriched pathways (**Fig. 5E**). As we observed the enrichment of immune-related pathways in most of the cell clusters (**Fig. 5D, E**), we examined their antigen-presenting abilities, which revealed weak antigen-presenting ability for epithelial cells from the skin and esophagus (**Additional file 11: Figure S10E**).

Next, we investigated the contribution of TFs in regulating the heterogenous transcriptional profiles of epithelial cells in and between organs. SCENIC analysis revealed that the regulation of TFs was similar among cell clusters of a same tissue but was very different for among clusters between tissues (**Fig. 5F**). Interestingly, the digestion-associated epithelial clusters exhibited similar activation of TFs, as did the cell clusters belong to non-digestive tissues including the trachea, skin and esophagus (**Fig. 5F**). In addition, we also observed some cluster-specific TFs in skin (*CEBPA*, *KLF4*, *EGR1*, *GATA3* and *NFIC)*, small intestine epithelial cells (*HNF4G*, *NR1H3*, *NR1I2*, *VDR* and *GATA4)*, tuff cells (*MTA3*, *SOX13*, *POU2F1*, *ETV7* and *POU2F3*), absorptive cells (*HES4*), and stem cells (*ASCL2*).

The similarity and heterogeneity of stromal cells

Endothelial cells (ECs) line up in a monolayer and form the interior surface of blood and lymphatic vessels as well as heart chambers. We identified a total of 7,137 ECs, including 6,863 blood endothelial cells (BECs, marked with *VWF*) and 274 lymphatic endothelial cells (LECs, marked with *LYVE1*). BECs and LECs could be further grouped into 13 and two clusters, respectively, which had unique gene signatures (**Additional file 12: Figure S11A-C and Additional file 26: Table S8**). Hierarchical cluster analysis revealed that the cell clusters from a same tissue were grouped closer than those between two different tissues (**Additional file 12: Figure S11D**), such as SELE_BECs and APOD_BECs from the skin, as well as MADCAM1_BEC and CD320_BEC from the common bile duct. Interestingly, a group of LECs (FCN3_LECs) were strictly identified in the liver and expressed only two of the four genes for LECs (*PECAM1* and *LYVE1*) and other liver-specific markers (*CD4*, *CD14*, *FCN2/3*, *OIT3* and *CLEC4G*), but not the other two genes for LECs (*PDPN* and *PROX1*[41]; **Additional file**

12: Figure S11B, E). The other group of LECs (CCL21_LEC) were from tissues other than the liver, with exclusively high expression of *CCL21*, the protein of which binds the chemokine receptor 7 (*CCR7*) and thereby promotes adhesion and migration of various immune cells[42]. This suggests that these LECs have higher potential to attract immune cells than the LECs in the liver. GO analysis results revealed that BECs and LECs from most of the clusters had common endothelial functions including *vasculature development* and *response to wounding*, as well as functions regulating immune response (**Additional file 12: Figure S11F**). Moreover, the cells from each cluster were shown to have specific biological functions, including *endocrine and hormone regulation* for liver BECs (TIMP1_BEC), *adaptive immune response* for heart BECs (ACKR1_BEC) and *regulation of inflammatory response* for liver LECs (FCN3_LEC). We further examined the antigen-presenting ability of cells from each cluster, which revealed higher ability for BECs from the skin (SELE_BEC and APOD_BEC, $P < 2.2\text{x}10^{-16}$ and $P = 0.003354$, respectively) than that from other organs, and higher ability for the LECs in the liver (FCN3_LEC; $P = 2.61\text{x}10^{-13}$) than the LECs from other tissues (**Additional file 12: Figure S11G**).

We also identified a total of 17,835 fibroblasts and smooth muscle cells from nine tissues. These cells were further grouped into 14 fibroblast clusters (11,767 cells, *MMP2*), four smooth muscle cell clusters (3,201 cells, *ACTA2*), and another five clusters assigned as FibSmo (2,867 cells; marked with *MMP2* and *ACTA2*; **Additional file 13: Figure S12A, B**). We observed organ-specific distribution for fibroblasts of different clusters, but mixture of multiple organs for smooth muscle cell clusters (**Additional file 13: Figure S12A, B**), which is consistent with the previous findings in a mouse model[9]. DEG analysis revealed that cells from each cluster had unique signature (**Additional file 13: Figure S12C and Additional file 26: Table S8**). Hierarchical cluster analysis showed that cells from clusters were grouped as organ-specific manner, namely close distance for those within a same tissue (**Additional file 13: Figure S12D**). GO analysis revealed that fibroblasts, smooth muscle cells and FibSmo cells shared classical functions, including *blood vessel development*, *negative regulation of cell proliferation*, *response to wounding*, *regulation of cellular component movement* and *cell-substrate adhesion*, together with additional functions for cells from each cluster (**Additional file 13: Figure S12E**).

ACE2 expression across different tissues

ACE2 is considered as a potential receptor for the most recent outbreaking coronavirus (HCoV-19)[43, 44]. We therefore examined the expression of *ACE2* in our dataset containing 43 cell clusters and observed extremely high expression of *ACE2* in only two clusters, namely small intestine epithelial cells (C15) and common bile duct epithelial cells (C26; **Fig. 1C, Additional file 4: Figure S3B and Additional file 14: Figure S13A, B**). The majority of ACE2-expressing cells were epithelial cells, which were detected in 11 of 15 organs, including common bile duct, rectum, skin and muscle, and other seven tissues recently reported (esophagus, trachea, small intestine, heart, liver, stomach and bladder[45]; **Additional file 14: Figure S13B bottom panel**). High expression of *ACE2* was observed in the heart (~3.01%), small intestine (~24.7%) and common bile duct (~19.09%), while low expression was observed in most of other tissues (< 1% *ACE2* positive cells; **Additional file 14: Figure S13B bottom panel**). Specifically, we observed high expression of *ACE2* in the fibroblasts of the heart and the smooth muscle cells of the heart and muscle (**Additional file 14: Figure S13A, B**). These observations are consistent with the enquiry results from the Human Protein Atlas database (https://www.proteinatlas.org) and recent reports[46-49]. Moreover, in common bile duct tissue, GO enrichment analysis revealed enrichment of cell adhesion and immune-related pathways in ACE2-positive cells, including *neutrophil activation involved in immune response* (*ANXA2*, *ASAH1*, *B2M*, *FTL*, *LYZ*, *MGST1*, *CEACAM6*, *DNAJC3*, *TCN1*, and *ATP6AP2*), *cell-cell adhesion via plasma-membrane adhesion molecules* (*DSG2*, *ITGB1*, *EPCAM*, *CEACAM6*, *PTPRF*, and *CD164*; **Additional file 26: Table S8**).

Complex and broad intercellular communication networks within tissues

Since we observed heterogeneity for cells in each organ tissue, we explored their potential intercellular communication network. First, we examined molecular interactions based on ligand-receptor pairs in each tissue using CellPhoneDB. We observed a total of 19,742 significant interactions based on 468 ligand-receptor pairs among cell types and tissues, which varied from 252 in the lymph node to 2,725 in the trachea (**Fig. 6A and Additional file**

27: Table S9). Among them, *MIF_CD74*, *HBEGF_CD44*, *CD55_ADGRE5*, *MIF_TNFRSF14* and *APP_CD74* were the top five frequent interacting pairs detected across different cell types (**Fig. 6B**), suggesting their important roles in mediating crosstalk between different cell types. Next, we focused on the interactions between any pairs of the major cell types including $CD4^+$ T, $CD8^+$ T, NK, B, plasma, myeloid, epithelial, endothelial, smooth muscle cells, and fibroblasts (**Fig. 6C**). We observed that myeloid cell was the most active cell type interacting with the other types of cells (6,695 inter-cell interactions), among which the most active intensive interactions were between myeloid cells and epithelial cells (8.1% of the total inter-cell interactions) in the skin and trachea (52.6%). Interestingly, the interactions were common between myeloid cells and all kinds of epithelial cells, with the most frequent ligand-receptor pair *CD74_MIF* (**Additional file 15: Figure S14**), dysregulations of which have been reported in several tumours[50, 51], suggesting their potential roles in tumorigenesis. $CD8^+$ T cells were among the most active cell types interacting with other types of cells (total inter-cell 4511 interactions), among which the most active interactions were between $CD8^+$ T and myeloid cells (2.2% of the total inter-cell interactions). The interactions were found mainly in the liver, rectum, and spleen (52.1%). The frequent interaction pairs between $CD8^+$ T cell and myeloid cells were *RPS19_C5AR1*, *MIF_CD74*, *CD55_ADGRE5*, *ITGAL_ICAM1*, *aLb2 complex_ICAM1*, *CD99_PILRA* and *ANXA1_FPR1*, most of which play important roles in immune regulation[52-55] (**Additional file 16: Figure S15**). $CD8^+$ T cells also had broad interactions with non-immune cells (**Additional file 17: Figure S16**). For instance, the most frequent interacting cytokine and receptor pair *CXCL12_CXCR4*[56] was observed between fibroblasts and $CD8^+$ T cells, suggesting the attractive potential of fibroblasts for T cells migration into tissues (**Additional file 17: Figure S16**).

DISCUSSION

Here we generated for the first time to our knowledge an adult human cell atlas (AHCA), by profiling the single cell transcriptome for 88,622 cells from 15 organs. The AHCA included 234 subtypes of cells, each of which was distinguished by multiple marker genes and transcriptional profiles and collectively contributed to the heterogeneity of major human

organs. The AHCA empowered us to explore the developmental trajectories of major cell types and identify regulators and interacting networks that might play important roles in maintaining the homeostasis of human body. We have provided the AHCA publicly available as a resource to uncover key events during the development of human disease in the context of heterogeneity of cells and organs.

It has been demonstrated that T_N cells are generated at the thymus and populate lymphoid tissues sites where they differentiate to T_{EFF} cells upon antigen stimulus, and subsequently develop into long-lived memory T cells[57]. However, how T cells develop into different state throughout human organs and the links across T cells of different states as well as the underlying regulatory networks, are largely unknown, especially in the context of one individual body. In our study, trajectory analysis revealed a clear development route from T_N cells into T_{EFF} cells and then $CD4^+$ and $CD8^+$ T_{RM} cells in non-lymphoid organs with terminal development states, which is consistent with the previous findings[57]. We also noted that T_{RM} cells exhibit development states at organ-specific patterns and consistently these cells were regulated by different types of transcription factors, suggesting that tissue microenvironment might regulate gene expression by affecting transcription factor's activity, which in turn shape the specific T cell phenotypes. TCR analysis tracking cell of same linage revealed widespread links among subpopulations of T_{RM} cells, which however, are considered as non-recirculating[58]. Moreover, intensive sharing of TCRs were observed among T_{RM} cells, T_{EM} cell and T_{EFF} cells (**Additional file 8: Figure S7H, I**). Together with their development states, these results suggest that T_{EM} cells and T_{EFF} cells might enter the tissues and developed into T_{RM} cells and IEL T cells.

Infiltrating macrophages come from classical monocytes in pathological settings, such as cancer[59], while various origins of adult macrophages among tissues in the steady state have been reported[60]. As such, how macrophages are renewed in maintenance of their hemostasis in tissues has been debated over whether the tissue-resident macrophages undergo local proliferation or recruitment of monocytes from peripheral blood[61, 62]. In our study, multiple observations suggest that macrophages in organs are derived from either circulating monocytes or in-site expansion and proliferation of macrophage populations coping with local microenvironment. Consistently, entry of monocytes to steady-state

nonlymphoid organs and self-maintenance of tissue macrophages have been reported in mouse models [63, 64]. Moreover, we observed a more isolated clustering of epithelial cells and stromal cells compared to that of the immune cells, which have circulating capability. Because epithelial cells and stromal cells are a fundamental component forming protective barriers and supporting matrix for many organs, disruptions of their homeostasis have been implicated in various diseases[65-68]. Accumulating studies have demonstrated remarkably functional heterogeneity of epithelial cells among tissues[69-71]. Consistently, our gene ontology enrichment analyses revealed very diverse functions of epithelial cells among different tissues, as well as stromal cells. Taken together, higher degree of heterogeneity might reflect higher degree of terminal differentiation states and distinct specific functions of these cells among different organs.

The AHCA not only brings more detailed understanding of cell development and heterogeneity, but also reveals novel cell types and genes as well as regulatory factors that might be important for cell development. We identified a small proportion of novel cells in the skin tissue, in addition to the majority of immune cells, myeloid cells, epithelial cells and stromal cells. Transcription factors are known as the 'master regulators' for gene expression[72, 73] and we identified numerous additional novel transcription factors in regulating the development of different cell state of major cell types, such as *CREM* in $CD4^+$ T_{RM}, *MAFF* in $CD8^+$ T_{RM}, and *ELF1* in both $CD4^+$ and $CD8^+$ T_{RM}, *CTCF* and *TGIF1* in B cells, *KLF13* in plasma cells, as well as *BHLHE40*, *MYC*, *YY1*, *FOS*, *SOX2*, *BCL11A* and *KLF5* in non-digestive tissues containing the trachea, skin and esophagus. Such TFs networks not only extend our understanding for how they regulate gene expression and shape the different phenotypes, but also provide potential gene combinations in reprogramming application. The AHCA also provides useful data to explore the cellular networks at a single-cell resolution. We note that a large number of interactions between immune cells and other cells in all tissues, reflecting essential and broad communications between immune cells and other cell types in human body. In addition, we also observed that epithelial cells had the most frequent intra-cell interactions among all cell types, suggesting that epithelial cells could interact with each other intensively to regulate their biology functions in each organ (**Additional file 18: Figure S17**).

Concerning the recent outbreak of COVID-19 caused by the new human coronavirus HCoV-19[74], we attempted to explore the usage of the AHCA in understanding the infection of this virus and the disease. The clinical course of HCoV-19 infected patients remains to be fully characterized, the disease pathogenesis is still largely unknown, and no pharmacological therapies of proven efficacy yet exist. The common symptoms include fever, fatigue, myalgia, dry cough, anorexia and shortness of breath[75]. The complications include acute respiratory distress syndrome, arrhythmia, acute cardiac injury, shock and acute kidney injury. Some severe patients died of multiple organs failure. Our dataset revealed *ACE2* expression in multiple cell types and organs, which overlapped with the suffered organs (lung, heart, trachea, stomach, small intestine, artery and muscle) of HCoV-19 infection[75]. Moreover, we observed GO enrichment pathway of *neutrophil activation involved in immune response* in ACE2-positive cells, which might contribute to neutrophilia in considerable proportion of patients (38%)[76] and much severe neutrophilia in non-survivors[75]. Given that *ACE2* is likely a receptor for the coronavirus[43, 44], these data suggest that ACE2-positive cells are the targets of the virus, by blocking the *ACE2* functions or triggering cell death. The skin and common bile duct also contain ACE2-positie cells, and thereby we note that these organs might also be the targets of the virus. Collectively, these findings provide hints for understanding the pathogenesis of HCoV-19 infection.

We acknowledged that current AHCA has several limitations. First, although we obtained a sufficient number of sequencing reads for each sample, the number of genes detected in each cell was limited. This might underestimate the roles of some low expressed genes, such as long non-coding RNAs. Second, we obtained around 5,000 cells in average for each organ, which might limit our ability to identify rare cell types and thus underestimate the heterogeneity of intercell interactions in organs. Third, we included only 15 organ tissues from a single donor in our study. Further studies on gene expression profiling at both transcriptional and protein levels, as well as functional characterization with more organs from a larger number of donors, would provide much broader and detailed global view of human cell atlas and cell biology.

Conclusions

We generated an adult human cell atlas, by profiling the single cell transcriptome for more than 88,000 cells of 15 organs from one research-consent donor. The AHCA uncovers the heterogeneity of cells in major human organs, containing more than 234 subtypes of cells. Comprehensive analyses of the AHCA enabled us to delineate the developmental trajectories of major cell types and to identify novel cell types, regulators, and key molecular events that might play important roles in maintaining the homeostasis of human body or otherwise developing into human diseases.

MATERIALS AND METHODS

Organ tissue collection

An adult male donor who died of traumatic brain injury was recruited with written consent from his family. The management of organs was complied with the organ procurement guidelines and the study was approved by the Ethics Committee of the First Affiliated Hospital of Sun Yat-sen University. Besides the organs for transplantation purpose, we collected tissues from 15 organs in sequence, including blood, bone marrow, liver, common bile duct, lymph node (hilar and mesenteric), spleen, heart (apical), urinary bladder, trachea, esophagus, stomach, small intestine, rectum, skin and muscle (thigh). All hollow viscera tissues were dissected according to the whole layer structure, and all parenchymal viscera tissues were obtained from the organ lower pole. All the tissue collection procedures were accomplished within 20 minutes to maximize cell viability. To avoid cross-contamination, we used different sets of sterilized surgical instruments. The blood and bone marrow samples were loaded into 10 ml anticoagulation tubes containing EDTA (BD Biosciences, Cat. no. BD-366643) and other tissues were placed in physiological saline (4℃) to wash away the blood and secretions, and then immediately in a D10 resuspension buffer, containing culture medium (DMEM medium; Gibco™, Cat. no. 11965092) with 10% fetal bovine serum (FBS; Gibco™, Cat. no. 10099141). All tissue samples were kept on ice and delivered to the laboratory within 40 mins for further processing.

Tissue dissociation and cell purification

All tissues were dissociated within 1.5 hours and viable cells were collected at the end using fluorescence-activated cell sorting (FACS; BD FACS Aria™ III). For solid tissues exclusive of liver, each fresh tissue was cut into 1 mm pieces and incubated with a proper digestive solution of enzyme cocktail (**Additional file 19: Table S1**), followed by neutralization with D10 and then passed through a 40 μm cell strainer (BD, Cat. no. 352340). The cell suspension was centrifuged at 300x g for 5 min at 4 °C, and the pellet was resuspended with 0.8% NH_4Cl (Sigma-Aldrich, Cat. no. 254134-5G) red blood cells lysis buffer (RBCL) on ice for 10 min, followed by an additional wash with D10. The liver tissue was cut into pieces of 3-4 mm and incubated with 1 mM EGTA (Sigma-Aldrich, Cat. no. E0396-10G) in 1 x DBPS (Gibco™, Cat. no. 14190250) for 10 min at 37 °C with rotation at 50 rpm. After a washing with 1 x DPBS to remove EGTA, tissue was then incubated in pre-warmed digestion buffer (Supplementary information, Table S1) with rotation at 100 rpm at 37 °C for 30 min. The liver cell suspension was carefully passed through a 70 mm nylon cell strainer (BD, Cat. no. 352350), which was further centrifuged at 50× g for 3 min at 4°C to pellet hepatocytes. The supernatant was centrifuged at 300× g for 5 min at 4°C to pellet non-parenchymal cells. The pellet was resuspended and treated with RBCL. The blood and bone marrow samples were pelleted by centrifugation at 300x g for 5 min at 4 °C and was resuspended with RBCL on ice for 10 min, followed by an additional wash with D10. All cells from each tissue were resuspended with D10 to a concentration of 50-500 million cells per milliliter and stained with Calcein AM (Component A: AM) and Ethidium homodimer-1 (Component B: EH) in LIVE/DEAD Viability/Cytotoxicity Kit (Invitrogen, Cat. no. L3224) for 20 min on ice. Only the AM^+EH^- cells were collected by FACS for each tissue.

cDNA library preparation

The concentration of single cell suspension was determined using Cellometer Auto 2000 (Cellometer) and adjusted to 1,000 cells/μl. Appropriate number of cells were loaded into CHROMIUM instrument (10X Genomics, CA, USA) according to the standard protocol of the Chromium single cell V(D)J kit in order to capture 5,000 cells to 10,000 cells/chip position. In brief, mRNA transcripts from each sample were ligated with barcoded indexes at 5'-end and

reverse transcribed into cDNA, using GemCode technology (10X Genomics, USA). cDNA libraries including the enriched fragments spanning the full-length V(D)J segments of T cell receptors (TCR) or B cell receptor (BCR), and 5'-end fragments for gene expression were separately constructed, which were subsequently subjected for high-throughput sequencing.

Single cell RNA-seq data processing

5'-end cDNA, TCR and BCR libraries were mixed and subjected for sequencing on Illumina HiSeq XTen instruments with pared-end 150 bp. Raw data (BCL files) from HiSeq platform was converted to fastq files using Illumina-implemented software bcl2fastq (version v2.19.0.316). cDNA reads were aligned to the human reference genome (hg38) and digital gene expression matrix was built using STAR algorithm in CellRanger ("count" option; version 3.0.1; 10X Genomics)[77]. TCR and BCR reads were aligned to human reference VDJ dataset (http://cf.10Xgenomics.com/supp/cell-vdj/refdata-cellranger-vdj-GRCh38-alts-ensembl-2.0.0.tar.gz) using CellRanger ("vdj" option; version 3.1.0; 10X Genomics). Parameters were set as default except for "force-cells" as 13,000. Raw digital gene expression matrix in the "filtered_feature_bc_matrix" file folder generated by CellRanger was used for further analysis.

Cell clustering and differential gene expression analysis

Quality control filtering, variable gene selection, dimensionality reduction and clustering for cells were performed using R software (version 3.6.1; https://www.r-project.org) and Seurat package[6] (version 3.0.1; https://satijalab.org/seurat) with default settings unless otherwise stated. For each tissue, we removed cells with low quality cells (UMI < 1,000, gene number < 500, and mitochondrial genome fragments ≥ 0.25) and genes with rare frequencies (0.1% of all cells). From the remaining cells, gene expression data for each sample was subtracted and scaled using "ScaleData" function (negative binomial model). Principal Component Analysis (PCA) and t-SNE implemented in the "RunPCA" and "RunTSNE" functions, respectively, were used to identify the deviations among cells. Genes with high variations were identified using "FindVariableGenes" and included for PCA (mean.cutoff ≥ 0.1, dispersion.cutoff ≥ 0.5). We used different number of Principal Components (PCs) and value

of perplexity determined by elbow plot for each tissue and cell type (**Additional file 28: Table S10**). Cell clusters were identified using “FindClusters” function and shown using t-SNE with the same number of PCs. Differential expression markers or genes were determined using Wilcoxon test implemented in “FindAllMarkers” function, which was considered significant with an average natural logarithm (fold-change) of at least 0.25 and a Bonferroni-adjusted p-value lower than 0.05. Subsequently, the candidate markers were reviewed and were used to annotate cell clusters. For analyses of the merged data from all tissues, we used 30 PCs and resolution parameter set to 1 for cell clustering. For further analyses, we removed cell clusters that had multiple well-defined marker genes of different cell types, and with an extremely high number of UMIs (**Additional file 21: Table S3**).

Identification and removal of highly transcribed genes with contamination potentials

Because we observed that some cell-specific genes were broadly expressed among all types of cell in a tissue, for example, *APOC3* in enterocytes cells in small intestine with an average more than 200 UMIs in a cell, we suspected that if a fraction of a certain type of cells were broken during the sample processing, cell-specific genes with high transcriptions would be released and thus contaminated all cell droplets. Especially, such genes would screw the differential expression analysis of cell type among all tissues. Therefore, we identified these genes and removed them for comparation analyses within major cell types. We assumed that non-epithelial cells from a same linage have similar gene profiles at a certain degree and such that a few genes would have modest expression in only one organ. Here given an example for T cells, we grouped together the cells previously labeled with NK/T, T, and immune cells (**Additional file 21: Table S3**) in each tissue. Next, we determined the top 1% of genes with high transcription in each tissue sample based on the total number of UMIs and marked these as potentially contaminating genes. We further randomly sampled 300 cells for each cell type in a tissue and as such generated artificial data for all tissues, with which differential expression gene analysis was performed using “FindAllMarker” function. Any gene with normal logarithm of FC above 0.25 and with an expression in less than 5% cells of all other tissue samples was considered contaminating gene. After removing the above contaminating genes, we performed cluster analysis with the first 20 PCs and resolution of

1.5. After further removal of cell clusters that had multiple well-defined marker genes of different cell types, we repeated cluster analysis using a lower resolution setting (**Additional file 28: Table S10**) to identify $CD4^+$, $CD8^+$ and NK cell (**Additional file 22: Table S4**) clusters.

For B cells, plasma cells, endothelial cells, macrophages, monocytes and fibroblasts, we applied a similar strategy to remove contaminating genes and cell clusters with multiple cell-specific markers.

Trajectories analysis

We performed trajectories analysis using Monocle3 alpha[4] for all tissue-derived T cells, macrophages/monocytes, according to the general pipeline (http://cole-trapnell-lab.github.io/Monocle-release/Monocle3/). For T cells, after identification of T cells clusters for $CD4^+$ and $CD8^+$ T, raw gene expression counts of cells were imported to the software. Only genes matching the thresholds (both mean expression and dispersion ratio are greater than 0.1 and 0.2, respectively.) were used for cell ordering and training the pseudo-time trajectory. For trajectory analysis of macrophages/monocytes, we used different thresholds for genes selection (mean expression ≥ 0.3, dispersion ration ≥ 0.1).

TCR and BCR analysis

We assessed the enrichment of TCR and BCR in various organs using R package Startrac (version 0.1.0)[24], which included only the cells with the certain clonotypes assigned by Cellranger and with paired chains (α and β for T cells, heavy and light chains for B cells). In brief, cells sharing identical TCR or BCR clones between tissues were measured using migration-index score, and the degree of cell linking between different clusters of T cells or B cells was determined by transition-index score. TCR or BCR diversity (Shannon-index score) was calculated using a “1 - expansion-index score”. For detailed pipeline, please referred to the website (https://github.com/Japrin/STARTRAC/blob/master/vignettes/startrac.html).

Presenting-antigen and proliferated scores

Presenting-antigen ability score (PAS) and proliferated score of each cell were calculated

using “AddModuleScore” function implemented in the Seurat package, with gene sets in REACTOME database “MHC CLASS II ANTIGEN PRESENTATION” pathway, and proliferation-association genes (including *ZWINT*, *E2F1*, *FEN1*, *FOXM1*, *H2AFZ*, *HMGB2*, *MCM2*, *MCM3*, *MCM4*, *MCM5*, *MCM6*, *MKI67*, *MYBL2*, *PCNA*, *PLK1*, *CCND1*, *AURKA*, *BUB1*, *TOP2A*, *TYMS*, *DEK*, *CCNB1* and *CCNE1*)[78], respectively.

Gene Ontology (GO) pathway enrichment analysis

All GO enrichment analysis was performed using the online tools metascape[79] with “multiple gene list mode” (http://metascape.org/gp/index.html). For epithelial cells, we selected genes with FC ≥ 2 and p.adjust < 0.05 for each tissue, while FC ≥ 1.5, p.adjust < 0.05, and ptc.1 > 0.2 for each cluster. For non-epithelial cells, we selected genes with lnFC ≥ 0.25, p.adjust < 0.05 and ptc.1 > 0.2 in each cluster. Only the genes ranked top 150 according the FC were used for comparisons in each tissue and subpopulation. The background was given as all the genes expressed in corresponding cell types. In addition, only the gene sets in the “GO Biological Processes” were considered.

Cellular interaction analysis

To investigate the cellular interaction, we identified the inferred paired molecules using CellphoneDB software (version 2.0)[18] with default parameters. To evaluate the interaction between major cell types, we first merged clusters within major cell types including T, B, NK, myeloid, endothelial cells, fibroblasts, smooth muscle cells and FibSmo cells within each tissue. Next, the merged cell types were as follows, $CD4^+$ T cells (T_N, T_{CM}, T_{reg}, T_{EM}, T_{EFF}, T_{RM}), $CD8^+$ T cells (T_N, T_{CM}, T_{EM}, T_{EFF}, T_{RM}, IEL, γδ, MAIT). B cells (naïve and memory B cells), NK cells, plasma cells, myeloid (monocytes: Mon, macrophages: Mac and DCs), Endothelial cells (BECs and IECs), fibroblasts (Fib), smooth muscle cells (Smo), and FibSmo cells (FibSmo). We did not merge subpopulations of epithelial cells because of their high degree of heterogeneity in each tissue. Considering the test efficiency and computational burden, we focused on cell types with more than 30 cells and only randomly selected 250 cells of each cell type for analysis in each tissue. The significant ligand-receptor pairs were filtered with a *P* value less than 0.05. All the analyses above were performed as tissue independent.

Visualization of interaction network was done using Cytoscape (version 3.7.0).

Single-cell regulatory network inference and clustering (SCENIC) analysis

We conducted SCENIC analysis on cells passed the filtering for each major cell types, using R package SCENIC (version 1.1.2.2) as previously described[80]. Regions for TFs searching were restricted to 10k distance centered the transcriptional start site (TSS) or 500bp upstream of the TSSs. Transcription factor binding motifs (TFBS) over-represented on a gene list and networks inferring were done using R package RcisTarget (version 1.4.0) and GENIE3 (version 1.6.0), respectively, with 20-thousand motifs database. We randomly selected no more than 250 cells for each cell cluster. The input matrix was the normalized expression matrix from Seurat. The cluster-specific TFs of one cluster were defined as the top 10 or 15 highly enriched TFs according to decreasing fold change compared to all other cell clusters using a Wilcoxon rank-sum test.

Analysis of differential pathway

Gene set variation analysis (GSVA) or gene set enrichment analysis (GSEA) was performed to identify significantly enriched or depleted groups of genes in each transcriptional dataset, using R package GSVA (version 1.32.0) or GSEA software (version 4.0.3) (https://www.gsea-msigdb.org/gsea/index.jsp) on the 50 hallmark pathways with default parameters, respectively.

Data sharing and code availability

The key raw data have been deposited in the Research Data Deposit (RDD; http://www.researchdata.org.cn). The R scripts are available at https://github.com/bei-lab or upon request.

Conflict of Interests

The authors declare no conflict of interest.

AUTHORSHIP CONTRIBUTIONS

J.-X.B., X.S.H. and Z.Y.G. conceived and supervised this study. S.H., L.-H.W., Y.L., Y.-Q.L., J.H.X., H.T.C., W.P., G.-W.L., P.P.W., B.L., X.X.J. and D.W. performed sample preparation, library construction, and data generation. S.H. performed bioinformatics analyses. S.H., J.-X.B. and Z.Y.G. wrote the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

We acknowledge Chunlin Luo, Xixi Chen, Xiangyu Xiong, Shuqiang Liu, Xiaochen Xu, Yaqing Zhou, Yanchun Feng, Jingjing He and Liping Xu for performing the tissues dissociation. This work was supported by grants as follows: the National Natural Science Foundation of China (81970564, 81471583 and 81570587), Guangdong Innovative and Entrepreneurial Research Team Program (2016ZT06S638), the Guangdong Provincial Key Laboratory Construction Projection on Organ Donation and Transplant Immunology (2013A061401007 and 2017B030314018), Guangdong Provincial Natural Science Funds for Major Basic Science Culture Project (2015A030308010), Guangdong Provincial Natural Science Funds for Distinguished Young Scholars (2015A030306025), Special Support Program for Training High Level Talents in Guangdong Province (2015TQ01R168), Pearl River Nova Program of Guangzhou (201506010014), and Science and Technology Program of Guangzhou (201704020150), National Program for Support of Top-Notch Young Professionals (J.-X.B.), Chang Jiang Scholars Program (J.-X.B.), Special Support Program of Guangdong (J.-X.B.) and Sun Yat-sen University Young Teacher Key Cultivate Project (17ykzd29).

Reference

1. Dalerba P, Kalisky T, Sahoo D, Rajendran PS, Rothenberg ME, Leyrat AA, Sim S, Okamoto J, Johnston DM, Qian D, et al: **Single-cell dissection of transcriptional heterogeneity in human colon tumors.** *Nat Biotechnol* 2011, **29:**1120-1127.
2. Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, Peshkin L, Weitz DA, Kirschner MW: **Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells.** *Cell* 2015, **161:**1187-1201.
3. Wu AR, Neff NF, Kalisky T, Dalerba P, Treutlein B, Rothenberg ME, Mburu FM, Mantalas GL, Sim S, Clarke MF, Quake SR: **Quantitative assessment of single-cell RNA-sequencing methods.** *Nat Methods* 2014, **11:**41-46.
4. Qiu X, Hill A, Packer J, Lin D, Ma YA, Trapnell C: **Single-cell mRNA quantification and differential analysis with Census.** *Nat Methods* 2017, **14:**309-315.
5. Fabre PJ, Leleu M, Mascrez B, Lo Giudice Q, Cobb J, Duboule D: **Heterogeneous combinatorial expression of Hoxd genes in single cells during limb development.** *BMC Biol* 2018, **16:**101.
6. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R: **Integrating single-cell transcriptomic data across different conditions, technologies, and species.** *Nat Biotechnol* 2018, **36:**411-420.
7. Haghverdi L, Lun ATL, Morgan MD, Marioni JC: **Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors.** *Nat Biotechnol* 2018, **36:**421-427.
8. Tabula Muris C, Overall c, Logistical c, Organ c, processing, Library p, sequencing, Computational data a, Cell type a, Writing g, et al: **Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris.** *Nature* 2018, **562:**367-372.

9. Han X, Wang R, Zhou Y, Fei L, Sun H, Lai S, Saadatpour A, Zhou Z, Chen H, Ye F, et al: **Mapping the Mouse Cell Atlas by Microwell-Seq.** *Cell* 2018, **173:**1307.

10. Wagner DE, Weinreb C, Collins ZM, Briggs JA, Megason SG, Klein AM: **Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo.** *Science* 2018, **360:**981-987.

11. Briggs JA, Weinreb C, Wagner DE, Megason S, Peshkin L, Kirschner MW, Klein AM: **The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution.** *Science* 2018, **360**.

12. Rosenberg AB, Roco CM, Muscat RA, Kuchina A, Sample P, Yao Z, Graybuck LT, Peeler DJ, Mukherjee S, Chen W, et al: **Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding.** *Science* 2018, **360:**176-182.

13. Young MD, Mitchell TJ, Vieira Braga FA, Tran MGB, Stewart BJ, Ferdinand JR, Collord G, Botting RA, Popescu DM, Loudon KW, et al: **Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors.** *Science* 2018, **361:**594-599.

14. Schiller HB, Montoro DT, Simon LM, Rawlins EL, Meyer KB, Strunz M, Vieira Braga FA, Timens W, Koppelman GH, Budinger GRS, et al: **The Human Lung Cell Atlas: A High-Resolution Reference Map of the Human Lung in Health and Disease.** *Am J Respir Cell Mol Biol* 2019, **61:**31-41.

15. Cheng JB, Sedgewick AJ, Finnegan AI, Harirchian P, Lee J, Kwon S, Fassett MS, Golovato J, Gray M, Ghadially R, et al: **Transcriptional Programming of Normal and Inflamed Human Epidermis at Single-Cell Resolution.** *Cell Rep* 2018, **25:**871-883.

16. Plasschaert LW, Zilionis R, Choo-Wing R, Savova V, Knehr J, Roma G, Klein AM, Jaffe AB: **A**

single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature* 2018, **560:**377-381.

17. Giladi A, Paul F, Herzog Y, Lubling Y, Weiner A, Yofe I, Jaitin D, Cabezas-Wallscheid N, Dress R, Ginhoux F, et al: **Single-cell characterization of haematopoietic progenitors and their trajectories in homeostasis and perturbed haematopoiesis.** *Nat Cell Biol* 2018, **20:**836-846.

18. Vento-Tormo R, Efremova M, Botting RA, Turco MY, Vento-Tormo M, Meyer KB, Park JE, Stephenson E, Polanski K, Goncalves A, et al: **Single-cell reconstruction of the early maternal-fetal interface in humans.** *Nature* 2018, **563:**347-353.

19. Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, et al: **The single-cell transcriptional landscape of mammalian organogenesis.** *Nature* 2019, **566:**496-502.

20. Zhang X, Lan Y, Xu J, Quan F, Zhao E, Deng C, Luo T, Xu L, Liao G, Yan M, et al: **CellMarker: a manually curated resource of cell markers in human and mouse.** *Nucleic Acids Res* 2019, **47:**D721-d728.

21. Mizumoto N, Takashima A: **CD1a and langerin: acting as more than Langerhans cell markers.** *J Clin Invest* 2004, **113:**658-660.

22. Lacroix M: **Significance, detection and markers of disseminated breast cancer cells.** *Endocr Relat Cancer* 2006, **13:**1033-1067.

23. Wong MT, Ong DE, Lim FS, Teng KW, McGovern N, Narayanan S, Ho WQ, Cerny D, Tan HK, Anicete R, et al: **A High-Dimensional Atlas of Human T Cell Diversity Reveals Tissue-Specific Trafficking and Cytokine Signatures.** *Immunity* 2016, **45:**442-456.

24. Zhang L, Yu X, Zheng L, Zhang Y, Li Y, Fang Q, Gao R, Kang B, Zhang Q, Huang JY, et al:

Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. *Nature* 2018, **564**:268-272.

25. Rothenberg EV: **The chromatin landscape and transcription factors in T cell programming.** *Trends Immunol* 2014, **35**:195-204.

26. Vaeth M, Gogishvili T, Bopp T, Klein M, Berberich-Siebelt F, Gattenloehner S, Avots A, Sparwasser T, Grebe N, Schmitt E, et al: **Regulatory T cells facilitate the nuclear accumulation of inducible cAMP early repressor (ICER) and suppress nuclear factor of activated T cell c1 (NFATc1).** *Proc Natl Acad Sci U S A* 2011, **108**:2480-2485.

27. Willinger T, Freeman T, Herbert M, Hasegawa H, McMichael AJ, Callan MF: **Human naive CD8 T cells down-regulate expression of the WNT pathway transcription factors lymphoid enhancer binding factor 1 and transcription factor 7 (T cell factor-1) following antigen encounter in vitro and in vivo.** *J Immunol* 2006, **176**:1439-1446.

28. Sebzda E, Zou Z, Lee JS, Wang T, Kahn ML: **Transcription factor KLF2 regulates the migration of naive T cells by restricting chemokine receptor expression patterns.** *Nat Immunol* 2008, **9**:292-300.

29. Karwacz K, Miraldi ER, Pokrovskii M, Madi A, Yosef N, Wortman I, Chen X, Watters A, Carriero N, Awasthi A, et al: **Critical role of IRF1 and BATF in forming chromatin landscape during type 1 regulatory cell differentiation.** *Nat Immunol* 2017, **18**:412-421.

30. Wang D, Diao H, Getzler AJ, Rogal W, Frederick MA, Milner J, Yu B, Crotty S, Goldrath AW, Pipkin ME: **The Transcription Factor Runx3 Establishes Chromatin Accessibility of cis-Regulatory Landscapes that Drive Memory Cytotoxic T Lymphocyte Formation.** *Immunity* 2018, **48**:659-674.e656.

31. Agematsu K, Hokibara S, Nagumo H, Komiyama A: **CD27: a memory B-cell marker.** *Immunol Today* 2000, **21:**204-206.

32. Herling M, Patel KA, Weit N, Lilienthal N, Hallek M, Keating MJ, Jones D: **High TCL1 levels are a marker of B-cell receptor pathway responsiveness and adverse outcome in chronic lymphocytic leukemia.** *Blood* 2009, **114:**4675-4686.

33. Yoon HS, Scharer CD, Majumder P, Davis CW, Butler R, Zinzow-Kramer W, Skountzou I, Koutsonanos DG, Ahmed R, Boss JM: **ZBTB32 is an early repressor of the CIITA and MHC class II gene expression during B cell differentiation to plasma cells.** *Journal of immunology (Baltimore, Md : 1950)* 2012, **189:**2393-2403.

34. Shinnakasu R, Kurosaki T: **Regulation of memory B and plasma cell differentiation.** *Curr Opin Immunol* 2017, **45:**126-131.

35. Fernandez D, Ortiz M, Rodriguez L, Garcia A, Martinez D, Moreno de Alboran I: **The proto-oncogene c-myc regulates antibody secretion and Ig class switch recombination.** *J Immunol* 2013, **190:**6135-6144.

36. Roy K, Mitchell S, Liu Y, Ohta S, Lin YS, Metzig MO, Nutt SL, Hoffmann A: **A Regulatory Circuit Controlling the Dynamics of NFkappaB cRel Transitions B Cells from Proliferation to Plasma Cell Differentiation.** *Immunity* 2019, **50:**616-628 e616.

37. Valledor AF, Borràs FE, Cullell-Young M, Celada A: **Transcription factors that regulate monocyte/macrophage differentiation.** *J Leukoc Biol* 1998, **63:**405-417.

38. Kueh HY, Champhekar A, Champhekhar A, Nutt SL, Elowitz MB, Rothenberg EV: **Positive feedback between PU.1 and the cell cycle controls myeloid differentiation.** *Science* 2013, **341:**670-673.

39. Davies LC, Jenkins SJ, Allen JE, Taylor PR: **Tissue-resident macrophages.** *Nature Immunology* 2013, **14:**986.

40. Roche PA, Furuta K: **The ins and outs of MHC class II-mediated antigen processing and presentation.** *Nat Rev Immunol* 2015, **15:**203-216.

41. Amatschek S, Kriehuber E, Bauer W, Reininger B, Meraner P, Wolpl A, Schweifer N, Haslinger C, Stingl G, Maurer D: **Blood and lymphatic endothelial cell-specific differentiation programs are stringently controlled by the tissue environment.** *Blood* 2007, **109:**4777-4785.

42. Mueller SN, Germain RN: **Stromal cell contributions to the homeostasis and functionality of the immune system.** *Nat Rev Immunol* 2009, **9:**618-629.

43. Xu X, Chen P, Wang J, Feng J, Zhou H, Li X, Zhong W, Hao P: **Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission.** *Sci China Life Sci* 2020.

44. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, Wang W, Song H, Huang B, Zhu N, et al: **Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding.** *Lancet* 2020.

45. Xin Zou KC, Jiawei Zou, Peiyi Han, Jie Hao, Zeguang Han: **The single-cell RNA-seq data analysis on the receptor ACE2 expression reveals the potential risk of different human organs vulnerable to Wuhan 2019-nCoV infection.** *Front Med*:0-.

46. Hamming I, Timens W, Bulthuis ML, Lely AT, Navis G, van Goor H: **Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis.** *J Pathol* 2004, **203:**631-637.

47. Crackower MA, Sarao R, Oudit GY, Yagil C, Kozieradzki I, Scanga SE, Oliveira-dos-Santos AJ,

da Costa J, Zhang L, Pei Y, et al: **Angiotensin-converting enzyme 2 is an essential regulator of heart function.** *Nature* 2002, **417**:822-828.

48. Kowalczuk S, Broer A, Tietze N, Vanslambrouck JM, Rasko JE, Broer S: **A protein complex in the brush-border membrane explains a Hartnup disorder allele.** *FASEB J* 2008, **22**:2880-2887.

49. Harmer D, Gilbert M, Borman R, Clark KL: **Quantitative mRNA expression profiling of ACE 2, a novel homologue of angiotensin converting enzyme.** *FEBS Lett* 2002, **532**:107-110.

50. Bach JP, Deuster O, Balzer-Geldsetzer M, Meyer B, Dodel R, Bacher M: **The role of macrophage inhibitory factor in tumorigenesis and central nervous system tumors.** *Cancer* 2009, **115**:2031-2040.

51. Liu YH, Lin CY, Lin WC, Tang SW, Lai MK, Lin JY: **Up-regulation of vascular endothelial growth factor-D expression in clear cell renal cell carcinoma by CD74: a critical role in cancer cell tumorigenesis.** *J Immunol* 2008, **181**:6584-6594.

52. Calame DG, Mueller-Ortiz SL, Morales JE, Wetsel RA: **The C5a anaphylatoxin receptor (C5aR1) protects against Listeria monocytogenes infection by inhibiting type 1 IFN expression.** *J Immunol* 2014, **193**:5099-5107.

53. Stein R, Mattes MJ, Cardillo TM, Hansen HJ, Chang CH, Burton J, Govindan S, Goldenberg DM: **CD74: a new candidate target for the immunotherapy of B-cell neoplasms.** *Clin Cancer Res* 2007, **13**:5556s-5563s.

54. Long EO: **ICAM-1: getting a grip on leukocyte adhesion.** *J Immunol* 2011, **186**:5021-5023.

55. Osei-Owusu P, Charlton TM, Kim HK, Missiakas D, Schneewind O: **FPR1 is the plague receptor on host immune cells.** *Nature* 2019, **574**:57-62.

56. Contento RL, Molon B, Boularan C, Pozzan T, Manes S, Marullo S, Viola A: **CXCR4-CCR5: a**

couple modulating T cell functions. *Proc Natl Acad Sci U S A* 2008, **105:**10101-10106.

57. Thome JJ, Yudanin N, Ohmura Y, Kubota M, Grinshpun B, Sathaliyawala T, Kato T, Lerner H, Shen Y, Farber DL: **Spatial map of human T cell compartmentalization and maintenance over decades of life.** *Cell* 2014, **159:**814-828.

58. Schenkel JM, Masopust D: **Tissue-resident memory T cells.** *Immunity* 2014, **41:**886-897.

59. Qian BZ, Li J, Zhang H, Kitamura T, Zhang J, Campion LR, Kaiser EA, Snyder LA, Pollard JW: **CCL2 recruits inflammatory monocytes to facilitate breast-tumour metastasis.** *Nature* 2011, **475:**222-225.

60. Yona S, Kim KW, Wolf Y, Mildner A, Varol D, Breker M, Strauss-Ayali D, Viukov S, Guilliams M, Misharin A, et al: **Fate mapping reveals origins and dynamics of monocytes and tissue macrophages under homeostasis.** *Immunity* 2013, **38:**79-91.

61. Hume DA, Ross IL, Himes SR, Sasmono RT, Wells CA, Ravasi T: **The mononuclear phagocyte system revisited.** *J Leukoc Biol* 2002, **72:**621-627.

62. Hume DA: **The mononuclear phagocyte system.** *Curr Opin Immunol* 2006, **18:**49-53.

63. Jakubzick C, Gautier EL, Gibbings SL, Sojka DK, Schlitzer A, Johnson TE, Ivanov S, Duan Q, Bala S, Condon T, et al: **Minimal differentiation of classical monocytes as they survey steady-state tissues and transport antigen to lymph nodes.** *Immunity* 2013, **39:**599-610.

64. Hashimoto D, Chow A, Noizat C, Teo P, Beasley MB, Leboeuf M, Becker CD, See P, Price J, Lucas D, et al: **Tissue-resident macrophages self-maintain locally throughout adult life with minimal contribution from circulating monocytes.** *Immunity* 2013, **38:**792-804.

65. Bryant DM, Mostov KE: **From cells to organs: building polarized tissue.** *Nat Rev Mol Cell Biol* 2008, **9:**887-901.

66. Macara IG, Guyer R, Richardson G, Huo Y, Ahmed SM: **Epithelial homeostasis.** *Curr Biol* 2014, **24:**R815-825.

67. Lynch MD, Watt FM: **Fibroblast heterogeneity: implications for human disease.** *J Clin Invest* 2018, **128:**26-35.

68. Eckers A, Haendeler J: **Endothelial cells in health and disease.** *Antioxid Redox Signal* 2015, **22:**1209-1211.

69. Haber AL, Biton M, Rogel N, Herbst RH, Shekhar K, Smillie C, Burgin G, Delorey TM, Howitt MR, Katz Y, et al: **A single-cell survey of the small intestinal epithelium.** *Nature* 2017, **551:**333-339.

70. Nguyen QH, Pervolarakis N, Blake K, Ma D, Davis RT, James N, Phung AT, Willey E, Kumar R, Jabart E, et al: **Profiling human breast epithelial cells using single cell RNA sequencing identifies cell diversity.** *Nature Communications* 2018, **9:**2028.

71. Dong J, Hu Y, Fan X, Wu X, Mao Y, Hu B, Guo H, Wen L, Tang F: **Single-cell RNA-seq analysis unveils a prevalent epithelial/mesenchymal hybrid state during mouse organogenesis.** *Genome Biol* 2018, **19:**31.

72. Lee TI, Young RA: **Transcriptional regulation and its misregulation in disease.** *Cell* 2013, **152:**1237-1251.

73. Singh H, Khan AA, Dinner AR: **Gene regulatory networks in the immune system.** *Trends Immunol* 2014, **35:**211-218.

74. Jiang S, Shi Z, Shu Y, Song J, Gao GF, Tan W, Guo D: **A distinct name is needed for the new coronavirus.** *Lancet* 2020.

75. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, Wang B, Xiang H, Cheng Z, Xiong Y, et al:

Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA* 2020.

76. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, Qiu Y, Wang J, Liu Y, Wei Y, et al: **Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study.** *Lancet* 2020, **395:**507-513.

77. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR: **STAR: ultrafast universal RNA-seq aligner.** *Bioinformatics* 2013, **29:**15-21.

78. Whitfield ML, George LK, Grant GD, Perou CM: **Common markers of proliferation.** *Nat Rev Cancer* 2006, **6:**99-106.

79. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK: **Metascape provides a biologist-oriented resource for the analysis of systems-level datasets.** *Nat Commun* 2019, **10:**1523.

80. Aibar S, Gonzalez-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, Rambow F, Marine JC, Geurts P, Aerts J, et al: **SCENIC: single-cell regulatory network inference and clustering.** *Nat Methods* 2017, **14:**1083-1086.

Figure legends

Figure 1. Overview of single-cell RNA sequencing of 15 organ tissues from a male adult donor.

A. An experiment schematic diagram highlights the sites of organs for tissue collection and sample processing. Live cells were collected using flow cytometry sorting (FACS) and subjected for cell barcoding. cDNA libraries for TCR, BCR and 5'-mRNA expression were constructed independently, followed by high throughput sequencing and downstream analyses.

B. Bar plot showing the numbers of high-quality cells (UMI ≥ 1000, Gene number ≥ 500 and the proportion of mitochondrial gene UMI < 0.25) in each organ (colored).

C. t-SNE visualization of all cells (88,622) in organs (colored as in **B**). The labeled cell types were the predominant cell types in each cluster.

Figure 2. The heterogeneity, development and clonality of T cells in human organs.

A, **B.** t-SNE plots of 7,122 $CD4^+$ T (11 clusters) and 13,372 $CD8^+$ T cells (22 clusters) from 15 organ tissues, respectively. Each dot represents a cell. Each color-coded region represents one cluster, which is indicated at the right panel of the plot.

C. Pseudo-time trajectory analysis of all $CD4^+$ (top panel) and $CD8^+$ T cells (bottom panel) with high variable genes. Each dot represents a single cell and is colored according to their clustering at a for $CD4^+$ and b for $CD8^+$. The inlet t-SNE plot at each plot shows each cell with a pseudo-time score from dark blue to yellow, indicating early and terminal states, respectively.

D, E. Heatmaps of the active scores of each T cell cluster as numbered on top, of which expression was regulated by transcription factors (TFs), as estimated using SCENIC analysis. Shown only the top 15 TFs for $CD4^+$ T cells (D) and the top 10 for $CD8^+$ T cells (E), which had the highest difference in expression regulation estimates between each cluster and all other cells, under a Wilcoxon rank-sum test.

F. Sharing intensity of TCR clones in $CD4^+$ (top panel) and $CD8^+$ (bottom panel) T cells between different organ samples. Each line represents a sharing of TCR between two organs at the ends and the thickness of the line represents a migration-index score between paired

organs calculated by STARTRAC. The sizes of the dots were shown according to the size of T cell clones in organs with different colors.

G. Migration- (top panel) and expansion-index (bottom panel) scores of $CD4^+$ and $CD8^+$ T cells of each tissue calculated by STARTRAC with a paired Student's t-test.

H. Expansion- (top panel) and transition-index (bottom panel) scores of each $CD4^+$ and $CD8^+$ T cell cluster calculated by STARTRAC.

I. Alluvial diagram showing the $CD8^+$ T cells with same TCRs detected in at least six tissues and different cell types. Each line represents a sharing of a cell, colored by tissue.

Figure 3. The heterogeneity and clonality of B cells in human organs.

A. t-SNE plots showed 15 clusters (10,655 cells) of B and plasma cells colored by tissues (top panel) and cell subtypes (bottom panel). Each dot represents a cell.

B. Distribution of B and plasma cells in each organ. Pie charts on top illustrate the composition of B and plasma cells in each organ. The stacked bars represent the percentage of each cluster in the indicated organ.

C. Violin plots showed the normalized expression of maker genes for B (*MS4A1*), plasma cells (*SDC1*), naïve B cell (*TCL1A*) and memory B cells (*CD27*).

D. Gene Ontology enrichment analysis results of B and plasma cell clusters. Cell clustered as numbered below are colored according to their $-\log_{10}P$ values. Only the top 20 significant terms (p-value < 0.05) were shown.

E. Heatmap of the active scores of each B and plasma cell subtype as numbered on top, of which expression was regulated by transcription factors (TFs), as estimated using SCENIC analysis. Shown are the top 10 TFs having the highest difference in expression regulation estimated between each cluster and all other cells, tested with a Wilcoxon rank-sum test.

F. Sharing intensity of BCR clones between different organs. Each line represents a sharing of BCR between two organs at the ends and the thickness of the line represents a migration-index score between paired organs calculated by STARTRAC. The size of the dot is shown as the size of B and plasma cell clones, colored by organ.

G. Expansion- (top panel) and transition-index (bottom panel) scores of each B and plasma cell cluster calculated by STARTRAC.

Figure 4. Heterogeneity and developmental stages of myeloid cells.

A. t-SNE plots showed 5,605 myeloid cells, colored according to their origins of tissues (top panel) or cell clusters (bottom panel). Each dot represents a cell.

B. Distribution of different myeloid cell subpopulations in each tissue. Pie charts on top illustrate the composition of macrophages, monocytes and DC cells in each tissue. The stacked bars represent the percentage of each cluster in the indicated organ.

C. Dendrogram of 15 clusters based on their normalized mean expression values (correlation distance metric, average linkage). Only genes with ln(fold-change) above 0.25, p.adjust < 0.05 and pct.1 ≥ 0.2 in each cluster were included in the calculations.

D. Heatmap of top 20 differentially expressed genes (239 genes ranked descending according to fold change) in each cell cluster as numbered on top.

E. Heatmap of the active scores of each myeloid cell subtype as numbered on top, of which expression was regulated by transcription factors (TFs), as estimated using SCENIC analysis. Shown are the top 10 TFs having the highest difference in expression regulation estimates between each cluster and all other cells, tested with a Wilcoxon rank-sum test.

F. Pseudo-time trajectory analysis of all myeloid cells with high variable genes. Each dot represents a single cell and is colored according to their clustering in **A**. The inlet t-SNE plot at each plot shows each cell with a pseudo-time score from dark blue to yellow, indicating early and terminal states, respectively.

Figure 5. The heterogeneity of epithelial cells inter- and intra-organ tissues.

A. t-SNE plots showed 18,090 epithelial cells, colored according to their origins of tissues (top panel) or cell clusters (bottom panel).

B. Dot plotting visualization of the normalized expression of marker genes of each epithelial cluster. The size of the dot represented the percentage of cells with a cell type, and the color represented the average expression level.

C. Dendrogram of 33 clusters based on their normalized mean expression values (correlation distance metric, average linkage). Only genes with ln(fold-change) above 0.25, p.adjust < 0.05 and pct.1 ≥ 0.2 in each cluster were included in the analysis.

D, E. Gene Ontology enrichment analysis results of each epithelial cell cluster in digestive organs (D) and non-digestive organs (E). Cell clustered as numbered below were colored according to their $-\log_{10}P$ values. Only the top 20 significant terms (p-value < 0.05) were shown.

F. Heatmap of the active scores of epithelial cell subtypes as numbered on top, of which expression was regulated by transcription factors (TFs), as estimated using SCENIC analysis. Shown are the top 10 TFs having the highest difference in expression regulation estimates between each cluster and all other cells, tested with a Wilcoxon rank-sum test.

Figure 6. Intercellular communication networks within tissues.

A. Counts of significant ligand-receptor interactions in each organ as indicated in the x axis.

B. Frequency of each ligand-receptor interaction in among all cell types. Shown are interactions with a frequency larger than 50.

C. Counts of interactions between a major cell type as subtitled to other major cell types as indicated in the colored circles, as well as that within the major cell type. CD4, $CD4^+$ T cells; CD8, $CD8^+$ T cells; B, B cells; Plasma, plasma cells; Myeloid, myeloid cells; NK, NK cells; Epi, epithelial cells; Fib, fibroblasts; Smo, smooth muscle cells; FibSmo, FibSmo cells; and Endo, endothelial cells). Numbers in red indicate the counts of ligand-receptor pairs for each intercellular link.

Figure S1. Quality information of single-cell RNA sequencing, related to figure 1.

Box plots showed the numbers of UMI and genes detected in each cell of each organ.

Figure S2. **Distribution of cells in each organ tissue, related to figure 1.**

Bar plots on the left panels showed the numbers of cells of different cell types, and the t-SNE plots on the right panels showed their relevant clustering. Cell types were determined by differential gene expression of known markers between clusters. The colors for each cell type are consistent between the bar plot and t-SNE plot for a same organ. Each dot in the t-SNE plot represented a single cell.

Figure S2. Distribution of cells in each organ tissue (continued).

Figure S3. **t-SNE plot of the cell clusters in different organ samples, related to figure 1.**

A. t-SNE plots showed cells in skin colored by normalized expression of *COCH*, *SCGB2A2*, *SCGB1B2P*, *PIP*, *MUCL1,* and *DCD* genes.

B. t-SNE plot showed cell clusters in 15 organs (88,622 cells). Each dot in the t-SNE plot represented a single cell, and each color-coded region indicated one cluster.

Figure S4. **Correlation analysis of cell types determined by whole dataset (all cell of 15 organs) and manual annotation in each organ, related to figure 1.**

x axis indicated clusters from main **Figure 1C**, and the y axis indicated clusters annotated manually, which were determined using cells from all tested 15 organs or from individual organ by genes with high expression, respectively. The result indicates similar cell types from distinct organs were clustered into a major cell type. In particular T cells, NK cell, B cells, plasma cells, endothelial cells, fibroblasts, smooth muscle cells, macrophages, monocytes and Schwan cells from different organs were grouped into limited clusters (regions outlined in dashed boxes).

Figure S5. Expression of maker genes in major cell types from all 15 organ samples, related to figure 1.

A. Dot plots showed the most highly expressed marker genes (x axis) of major cell types (y axis) in main **Figure 1C**. The depth of the color from white to blue and the size of the dot represented the average expression from low to high and the percent of cells expressing the gene, respectively.

B. t-SNE plots showed the normalized expression of representative markers for T/NK cells (*CD3E*), B cells (*MS4A1*), macrophages/monocytes (*CD14*), endothelial cells (*PECAM1*), fibroblasts (*MMP2*) and smooth muscle cells (*ACTA2*). Each dot represented a single cell. Color keys from grey to blue indicated relative expression levels from low to high.

Figure S6. The heterogeneity and development trajectory analysis of T cells, **related to**

figure 2.

A. t-SNE plot of 7,122 $CD4^+$ T (top panel) and 13,372 $CD8^+$ T cells (bottom panel) from 15 organ tissues. Each point represented a cell, colored according to their origins of tissues as indicated on the right legend.

B. Bar charts showed the proportions of cell types in each tissue for $CD4^+$ T cells (left panel) and $CD8^+$ T cells (right panel).

C, D. Heatmap showed the expression of the signature genes for each of the $CD4^+$ (C) and $CD8^+$ (D) T cell clusters as indicated by numbers at the bottom.

E, F. Pseudo-time trajectory analysis of each of the $CD4^+$ (E) and $CD8^+$ T cells (F) with high variable genes inferred by Monocle3 alpha. Each dot represented a single cell and colored according to their origin of tissue (top panels) or cluster (bottom panels).

Figure S7. The heterogeneity and clonalities of T cell clones in human body, related to figure 2.

A, B. Frequencies of the V and J segments of the TCR α (top panels) and β chains (bottom panels) for $CD4^+$ T cells (A) and $CD8^+$ T cells (B).

C. Comparison of TCR clone sizes of $CD4^+$ and $CD8^+$ T cells in each tissue as indicated by different colors. Two-sided paired Student's t-test was done.

D. Bar plots showed the numbers of clonal cells in each organ for $CD4^+$ (top panel) and $CD8^+$ (bottom panel) T cells. The clonotypes are categorized as unique ($n = 1$), duplicated clonal ($n = 2$) and clonal ($n \geq 3$) based on their numbers of cells sharing a TCR.

E. Comparison of diversity between $CD4^+$ and $CD8^+$ T cells in each organ using Shannon entropy. A two-sided paired Student's t-test was done.

F. Bar plots showed the numbers of clonal cells in each cell cluster for $CD4^+$ (top panel) and $CD8^+$ (bottom panel) T cells. The definitions for clonotypes are as same as in **D**.

G. Box plot showed the expansion- and transition-index scores of each $CD4^+$ or $CD8^+$ T cell cluster. The comparison was tested using a Wilcoxon rank-sum test.

H, I. Heatmaps showed the pairwise transition-index scores for either two of the $CD4^+$ T cells (H) and $CD8^+$ T cells (I) clusters.

J, K. Diversities of $CD4^+$ (J) and $CD8^+$ (K) T cells in each organ with color illustration in **E**.

Shannon-index score was measured using TCR β chain or paired TCR α-β chains. Comparison was tested with a two-sided paired Student's t-test.

L. Diversities of subtypes of $CD4^+$ (top panel) and $CD8^+$ (bottom panel) T cell. Shannon-index score was measured using TCR β chain or paired TCR α-β chains.

Figure S8. The heterogeneity of B cells and plasma cells, related to figure 3.

A. Expression profiles of each B and plasma cell clusters as indicated by the numbers on top and illustrated at the right inlet box. 1,013 genes with FC ≥ 1.5 and p.adjust < 0.05 were shown and IGK/L/H genes were removed.

B. Scatter plot showed the reverse correlation between the expressions of *CD27* and *TCL1A* in all B and plasma cells. Cells with no expression of both the genes were excluded. Person's R = -0.83, *P* = 0.

C. Violin plot showed the presenting-antigen ability scores of each B and plasma cell cluster as indicated at the x axis.

D. GSEA analysis results showed the significant enriched pathways of *oxidative phosphorylation* and *fatty acid metabolism* in naïve B in non-lymph node anatomical position.

E-G. Frequencies of V and J gene segments of IGH (E), IGK (F) and IGL (G) chains.

H, I. Bar plots showed the numbers of clonal cells in each organ (H) and cell cluster (I). The clonotypes are categorized as unique (n = 1), duplicated clonal (n = 2) and clonal (n ≥ 3) based on their numbers of cells sharing a BCR.

Figure S9. Heterogeneity and trajectory analysis of macrophages, related to figure 4.

A. t-SNE plots showed the normalized expression of marker genes of monocytes (*S100A8/9/12 and VCAN*), macrophages (pan-marker: *C1QC, C1QA, C1QB, VSIG4*; Langerhans: *CD207*), DC (*CD1C*) and subpopulation-specific genes (*CD14* and *CD16*). Each dot represented a single cell and the depth of color from grey to blue represented the expression level from low to high.

B, C. Pseudo-time trajectory analysis of myeloid cells in each tissue (B) or cluster (C). Each dot represented a single cell.

D. Violin plots showed the normalized expression of proliferated marker genes (*MKI67*,

TYMS and *PCAN*) in each myeloid cell cluster as indicated at the x axis.

E. GSVA analysis showed the pathways with significant difference of activation between monocytes and macrophages clusters. Activation was determined using GSVA. Comparison was tested with an R package limma.

F. Violin plot showed the presenting-antigen ability scores of each myeloid cell cluster. comparisons were done using a Wilcoxon rank-sum test.

G. Pseudo-time trajectory analysis of myeloid cells in rectum. Each point represented a single cell and was colored by cluster corresponding to **C**. High variable genes included in the analyses were inferred using Monocle3 alpha.

Figure S10. Expression and antigen presenting ability of epithelial cells, related to figure 5.

A. Scaled heatmap showed the expression of 189 tissue-specific genes in epithelial cells from different tissues as indicated at the bottom. The tissue-specific genes were defined as FC ≥ 5, pct.1 ≥ 0.2 and p.adj < 0.05. The expression levels for high and low were indicated in red and blue, respectively.

B. Top 20 pathways enriched significantly according to the GO analysis ($p < 0.05$) in epithelial cell from each organ as indicated at the bottom. The depth of color represented -log10*P* calculated by online tool “metascape”.

C. Scaled heatmap showed the expression of 375 signature genes in epithelial cell clusters as indicated by numbers on top. The signature genes were determined by FC ≥ 5, pct.1 ≥ 0.2, pct.2 ≤ 0.2 and p.adj < 0.05.

D. Heatmap showed the numbers of cluster-specific genes shared between clusters.

E. Violin plot showed the presenting-antigen ability score of each epithelial cell cluster as indicated at the x axis.

Figure S11. The heterogenous expression pattern of endothelial cells.

A. t-SNE showed 7,137 endothelial cells, colored by tissue (top panel) or cluster (15 clusters, bottom panel).

B. Violin plots showed the log normalized expression of marker genes for pan-endothelial

cells (*PECAM1*), blood endothelial cells (*VWF*), and lymphatic endothelial cells (*LYEV1* and *PDPN*) in each endothelial cell cluster as indicated at the bottom.

C. Scaled heatmap showed the expression of the top 10 cluster-specific genes in endothelial cell clusters as indicated at the bottom. The cluster-specific genes were defined as pct.1 ≥ 0.2 and p.adj < 0.05. The expression levels for high and low were indicated in red and blue, respectively.

D. Dendrogram of cell clusters based on their normalized mean expression values of 15 clusters (correlation distance metric and average linkage). Only genes with lnFC above 0.25, p.adjust < 0.05 and pct.1 ≥ 0.2 in each cluster are used to calculate the distance. Bar plots at the right illustrate the proportions of cells in corresponding organs, colored by tissue as in **A** (top panel).

E. Scatter plot showed the differential gene expression of LECs in the liver (blue) and other organs (red) compared to all other endothelial cells. Selected signature genes for all LECs were indicated in grey box.

F. GO enrichment analysis in each endothelial cell cluster. The point for each cluster in a pathway was colored by -log10*P* calculated using online tool “metascape”.

G. Violin plot showed the presenting-antigen ability scores of each endothelial cell cluster, comparison was tested with a Wilcoxon rank-sum test. P-value was calculated as in the comparison of one cluster to all other cells.

Figure S12. The similarity and heterogeneity of fibroblasts and smooth muscle cells.

A. t-SNE showed a total of 17,835 stromal cells, colored according to their origin of tissue (left panel) or cell cluster (15 clusters, right panel).

B. Violin plots showed the log normalized expression of marker genes for fibroblasts (*MMP2*), and smooth muscle cells (*ACTA2*) in each cell cluster.

C. Scaled heatmap showed the expression of the top 10 cluster-specific genes in each stromal cell (fibroblasts, smooth muscle cell and FibSmo cell) cluster as indicated on top. The cluster-specific genes were defined as pct.1 ≥ 0.2 and p.adj < 0.05. The expression levels for high and low were indicated in red and blue, respectively.

D. Dendrogram of clusters based on the normalized mean expression values of 23 clusters

(correlation distance metric, average linkage). Only genes with lnFC above 0.25, p.adjust < 0.05 and pct.1 ≥ 0.2 in each cluster are used to calculate the distance. Bar plots at the right illustrate the proportions of cells in corresponding organs, colored by tissue as in **A** (left panel).

E. Selected pathways in the GO enrichment analysis in each stromal cell cluster as indicated at the bottom. The point for each cluster in a pathway was colored by -log10*P* calculated using online tool "metascape".

Figure S13. Distribution of *ACE2* positive cells in human body.

A. t-SNE plot showed the normalized expression of *ACE2* in 15 organs. The expression levels for high and low were indicated in red and blue, respectively. Arrows indicated higher proportion of *ACE2* in common bile duct, small intestine, heart and muscle tissues.

B. Violin plot (top panel) showing the normalized expression of *ACE2* in 43 cell clusters identified in 15 organs. Table (bottom panel) listed the proportions and cell types of *ACE2* positive cells in each organ.

Figure S14. Overview of ligand-receptor interactions between epithelial cells and myeloid cells, related to figure 6.

The significance of *P*-values was indicated as circle size (permutation test). The naming system was as follow, taking an example of "IL10_receptor IL10" in "Mac.Mucosal_Squamous_epithelial_cells_esophagus_10.Esophagus": The interaction pair A_B, C.D.E: the means of the average expression levels of IL10 receptor in Mac cluster and *IL10* in cluster Mucosal_Squamous_epithelial_cells_esophagus_10 in tissue Esophagus.

Figure S15. Overview of ligand-receptor interactions between $CD8^+$ T and myeloid cells, related to figure 6.

Notes as in **Figure S14**.

Figure S16. Overview of ligand-receptor interactions between $CD8^+$ T and fibroblasts, related to figure 6.

Notes as in **Figure S14**.

Figure S17. Overview of ligand-receptor interactions within epithelia cells.

Notes as in **Figure S14**.

Figure. 1

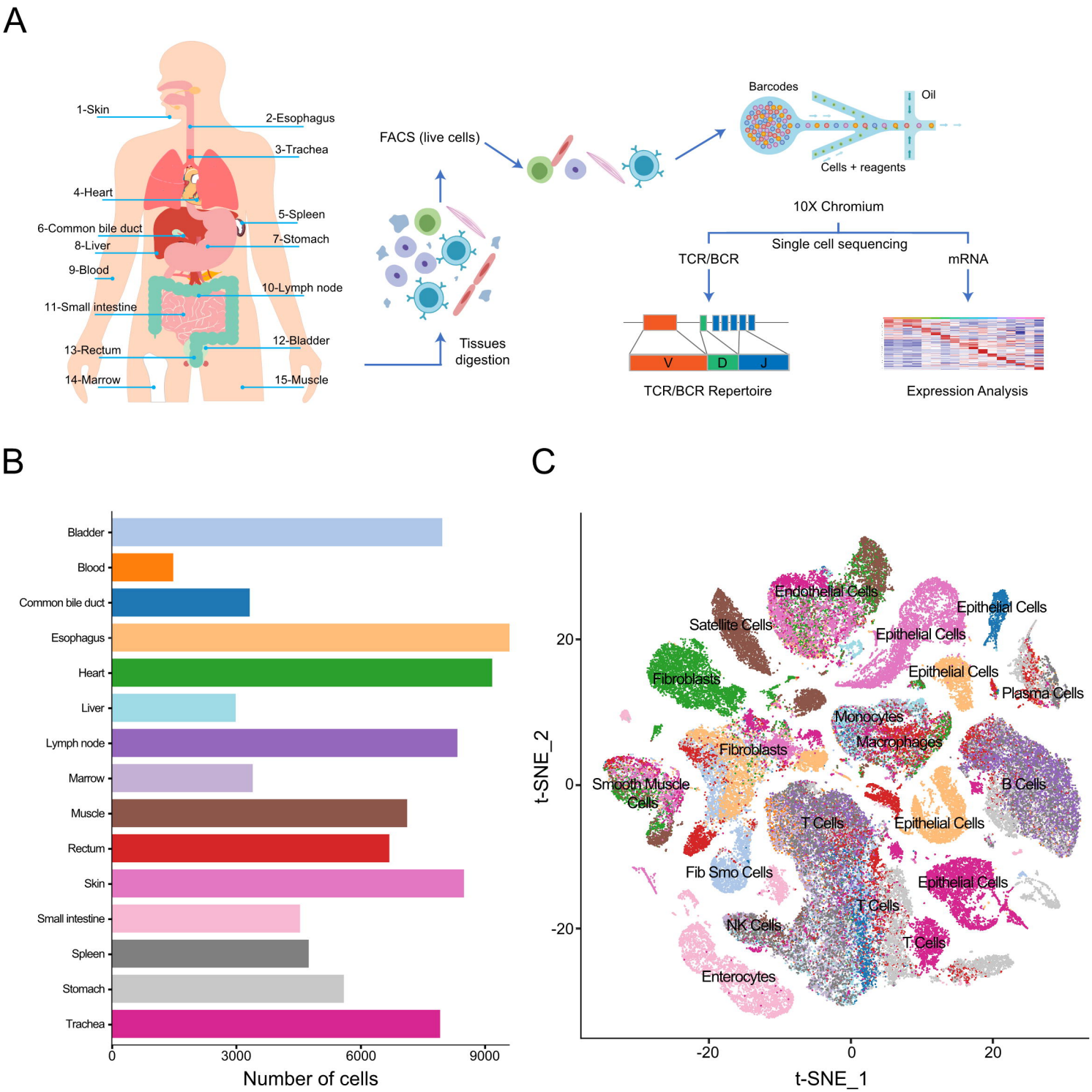
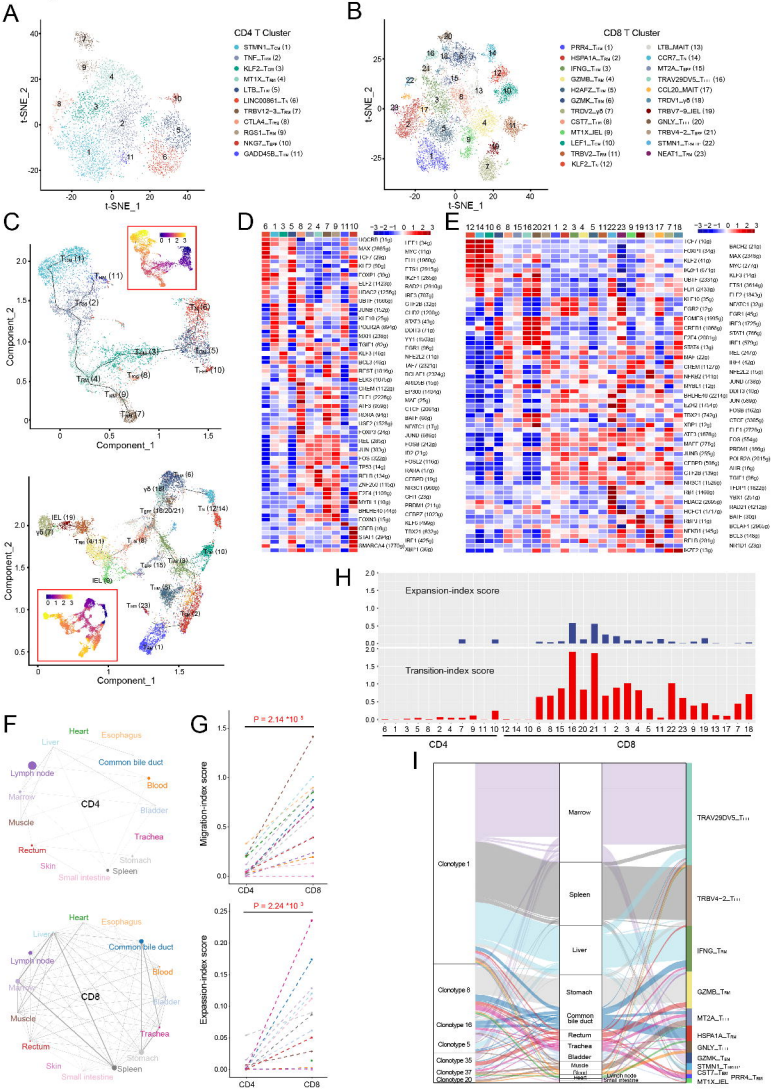

A
1-Skin
2-Esophagus
3-Trachea
4-Heart
5-Spleen
6-Common bile duct
7-Stomach
8-Liver
9-Blood
10-Lymph node
11-Small intestine
12-Bladder
13-Rectum
14-Marrow
15-Muscle
FACS (live cells)
Tissues digestion
Barcodes
Oil
Cells + reagents
10X Chromium
Single cell sequencing
TCR/BCR
mRNA
V
D
J
TCR/BCR Repertoire
Expression Analysis
B
Bladder
Blood
Common bile duct
Esophagus
Heart
Liver
Lymph node
Marrow
Muscle
Rectum
Skin
Small intestine
Spleen
Stomach
Trachea
0
3000
6000
9000
Number of cells
C
Endothelial Cells
Epithelial Cells
Satellite Cells
Epithelial Cells
Epithelial Cells
Fibroblasts
Plasma Cells
Monocytes
Macrophages
Fibroblasts
Smooth Muscle Cells
B Cells
T Cells
Epithelial Cells
Fib Smo Cells
Epithelial Cells
T Cells
NK Cells
T Cells
Enterocytes
20
0
-20
t-SNE_2
-20
0
20
t-SNE_1

Figure. 2



A
CD4 T Cluster
STMN1_Tcu (1)
TNF_Tcu (2)
KLF2_Tcu (3)
MT1X_Tcu (4)
LTB_Tcu (5)
LINC00861_Tc (6)
TRBV12-3_Tcu (7)
CTLA4_Tcu (8)
RGS1_Tcu (9)
NKG7_Tcu (10)
GADD45B_Tcu (11)
t-SNE_1
t-SNE_2
B
CD8 T Cluster
PRR4_Tcu (1)
HSPA1A_Tcu (2)
IFNG_Tcu (3)
GZMB_Tcu (4)
H2AF2_Tcu (5)
GZMK_Tcu (6)
TRDV2_γδ (7)
CST7_Tcu (8)
MT1X_IEL (9)
LEF1_Tcu (10)
TRBV2_Tcu (11)
KLF2_Tcu (12)
LTB_MAIT (13)
CCR7_Tc (14)
MT2A_Tcu (15)
TRAV29DV5_Tc (16)
CCL20_MAIT (17)
TRDV1_γδ (18)
TRBV7-9_IEL (19)
GNLY_Tc (20)
TRBV4-2_Tcu (21)
STMN1_Tcu (22)
NEAT1_Tcu (23)
t-SNE_1
t-SNE_2
C
Component_1
Component_2
D
E
F
Heart
Esophagus
Liver
Common bile duct
Lymph node
Blood
Marrow
CD4
Bladder
Muscle
Trachea
Rectum
Stomach
Skin
Spleen
Small intestine
CD8
G
P = 2.14 *10 5
Migration-index score
P = 2.24 *10 3
Expassion-index score
CD4
CD8
H
Expansion-index score
Transition-index score
CD4
CD8
I
Clonotype 1
Clonotype 8
Clonotype 16
Clonotype 5
Clonotype 35
Clonotype 37
Clonotype 20
Marrow
Spleen
Liver
Stomach
Common bile duct
Rectum
Trachea
Bladder
Muscle
Skin
Small intestine
TRAV29DV5_Tc
TRBV4-2_Tc
IFNG_Tcu
GZMB_Tcu
MT2A_Tc
HSPA1A_Tcu
GNLY_Tc
GZMK_Tcu
STMN1_Tcu
CST7_Tcu
MT1X_IEL
PRR4_Tcu

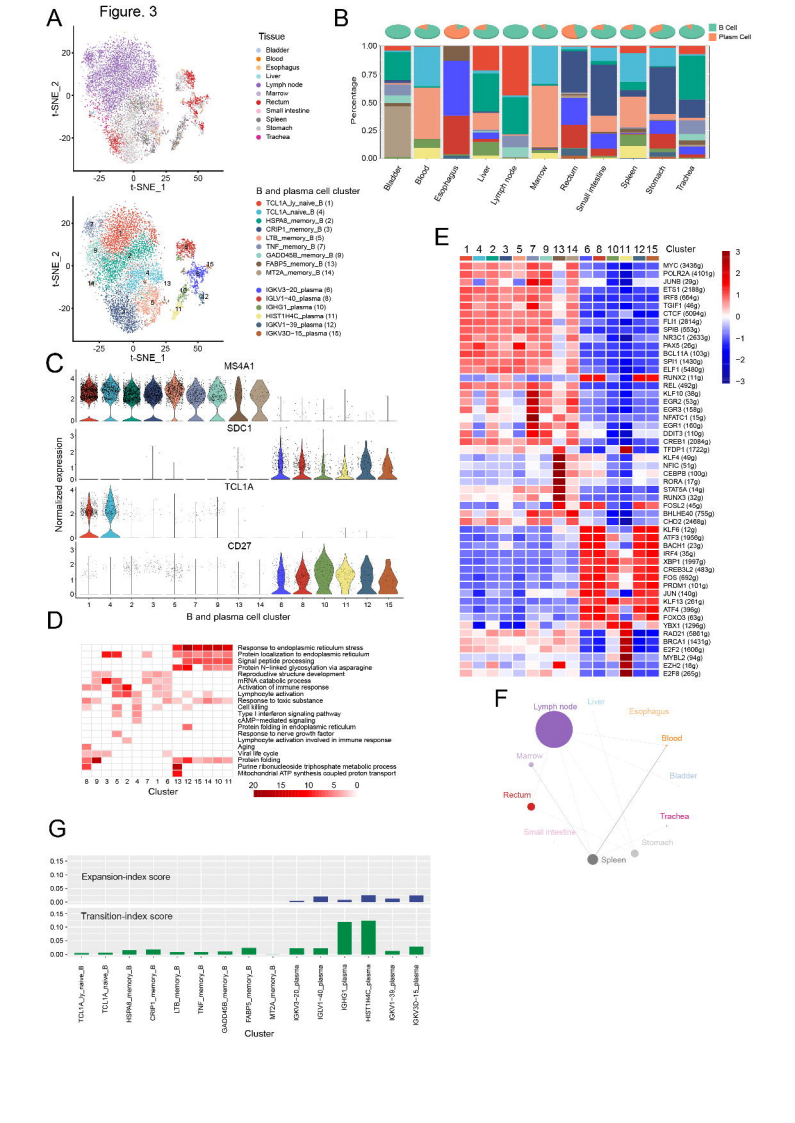


Figure. 3
A
t-SNE_2
t-SNE_1
Tissue
Bladder
Blood
Esophagus
Liver
Lymph node
Marrow
Rectum
Small intestine
Spleen
Stomach
Trachea
B and plasma cell cluster
TCL1A_ly_naive_B (1)
TCL1A_naive_B (4)
HSPA8_memory_B (2)
CRIP1_memory_B (3)
LTB_memory_B (5)
TNF_memory_B (7)
GADD45B_memory_B (9)
FABP5_memory_B (13)
MT2A_memory_B (14)
IGKV3−20_plasma (6)
IGLV1−40_plasma (8)
IGHG1_plasma (10)
HIST1H4C_plasma (11)
IGKV1−39_plasma (12)
IGKV3D−15_plasma (15)
B
B Cell
Plasm Cell
Percentage
Bladder
Blood
Esophagus
Liver
Lymph node
Marrow
Rectum
Small intestine
Spleen
Stomach
Trachea
C
MS4A1
SDC1
TCL1A
CD27
Normalized expression
B and plasma cell cluster
D
Response to endoplasmic reticulum stress
Protein localization to endoplasmic reticulum
Signal peptide processing
Protein N−linked glycosylation via asparagine
Reproductive structure development
mRNA catabolic process
Activation of immune response
Lymphocyte activation
Response to toxic substance
Type I IFN
Type I interferon signaling pathway
cAMP−mediated signaling
Protein folding in endoplasmic reticulum
Response to nerve growth factor
Lymphocyte activation involved in immune response
Aging
Viral life cycle
Protein folding
Purine ribonucleoside triphosphate metabolic process
Mitochondrial ATP synthesis coupled proton transport
Cluster
E
Cluster
MYC (3436g)
PCLR2A (4101g)
JUNB (29g)
ETS1 (2188g)
IRF8 (604g)
TGIF1 (46g)
CTCF (5094g)
FLI1 (2814g)
SPIB (803g)
NR3C1 (2633g)
PAX5 (26g)
BCL11A (103g)
SPI1 (1430g)
ELF1 (5480g)
RUNX2 (11g)
REL (462g)
KLF10 (38g)
EGR2 (53g)
EGR3 (158g)
NFATC1 (15g)
EGR1 (160g)
DDIT3 (110g)
CREB1 (2084g)
TFDP1 (1722g)
KLF4 (46g)
NFIC (51g)
CEBPB (100g)
RORA (17g)
STAT5A (14g)
RUNX3 (32g)
FOSL2 (45g)
BHLHE40 (755g)
CHD2 (2468g)
KLF6 (12g)
ATF3 (1956g)
BACH1 (23g)
IRF4 (35g)
XBP1 (1997g)
CREB3L2 (483g)
FOS (892g)
PRDM1 (101g)
JUN (140g)
KLF13 (261g)
ATF4 (396g)
FOXO3 (63g)
YBX1 (1296g)
RAD21 (5861g)
BRCA1 (1431g)
E2F2 (1606g)
MYBL2 (94g)
E2F3 (16g)
E2F8 (265g)
F
Lymph node
Liver
Esophagus
Blood
Marrow
Bladder
Rectum
Trachea
Small intestine
Stomach
Spleen
G
Expansion−index score
Transition−index score
TCL1A_ly_naive_B
TCL1A_naive_B
HSPA8_memory_B
CRIP1_memory_B
LTB_memory_B
TNF_memory_B
GADD45B_memory_B
FABP5_memory_B
MT2A_memory_B
IGKV3−20_plasma
IGLV1−40_plasma
IGHG1_plasma
HIST1H4C_plasma
IGKV1−39_plasma
IGKV3D−15_plasma
Cluster

Figure. 4

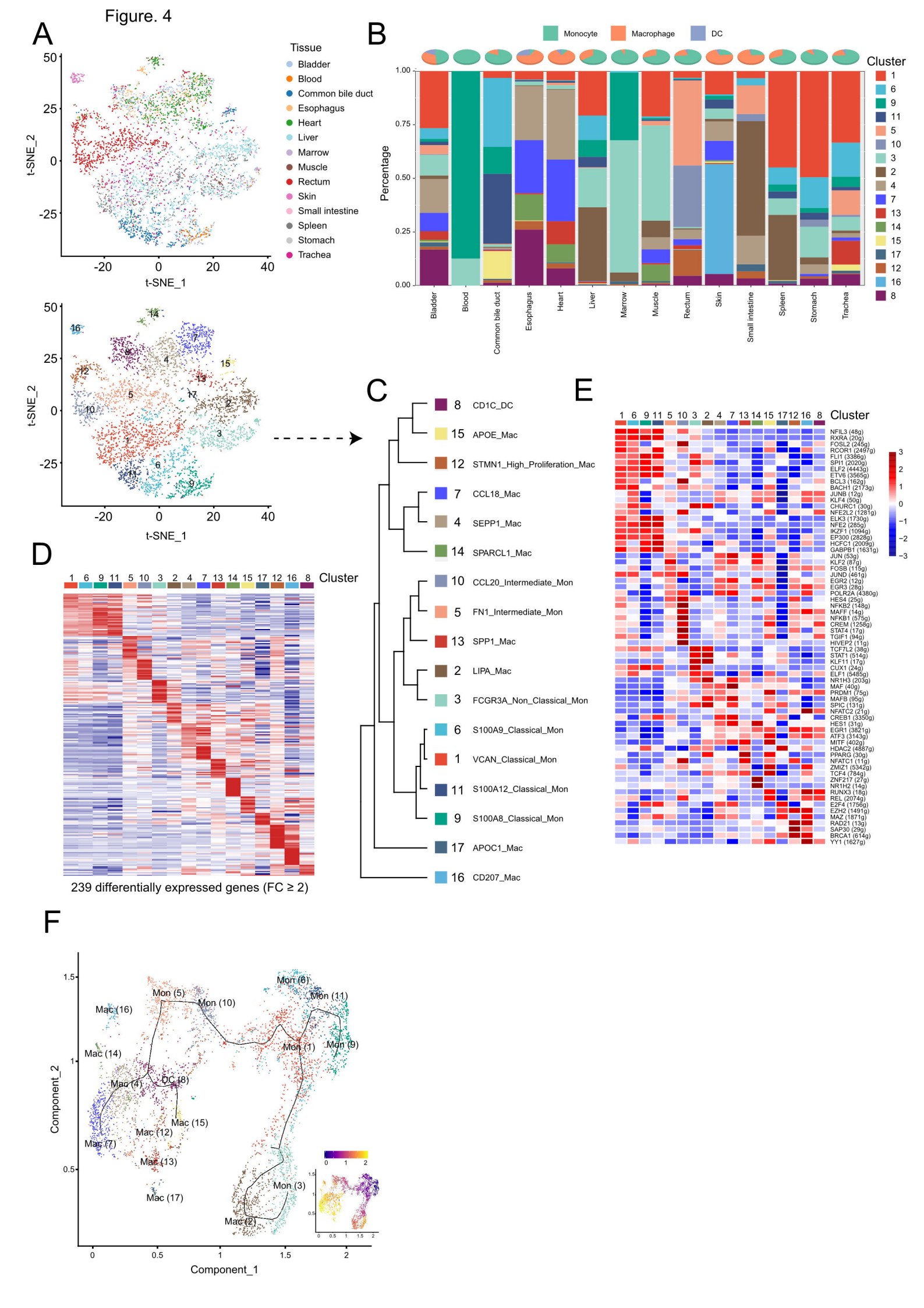

A
Tissue
Bladder
Blood
Common bile duct
Esophagus
Heart
Liver
Marrow
Muscle
Rectum
Skin
Small intestine
Spleen
Stomach
Trachea
t-SNE_1
t-SNE_2
B
Monocyte
Macrophage
DC
Cluster
Percentage
Bladder
Blood
Common bile duct
Esophagus
Heart
Liver
Marrow
Muscle
Rectum
Skin
Small intestine
Spleen
Stomach
Trachea
C
8 CD1C_DC
15 APOE_Mac
12 STMN1_High_Proliferation_Mac
7 CCL18_Mac
4 SEPP1_Mac
14 SPARCL1_Mac
10 CCL20_Intermediate_Mon
5 FN1_Intermediate_Mon
13 SPP1_Mac
2 LIPA_Mac
3 FCGR3A_Non_Classical_Mon
6 S100A9_Classical_Mon
1 VCAN_Classical_Mon
11 S100A12_Classical_Mon
9 S100A8_Classical_Mon
17 APOC1_Mac
16 CD207_Mac
D
Cluster
239 differentially expressed genes (FC ≥ 2)
E
Cluster
F
Mon (6)
Mon (5)
Mon (10)
Mon (11)
Mac (16)
Mon (1)
Mon (9)
Mac (14)
Mac (4)
DC (8)
Mac (15)
Mac (12)
Mac (7)
Mac (13)
Mon (3)
Mac (17)
Mac (2)
Component_1
Component_2

Figure. 5

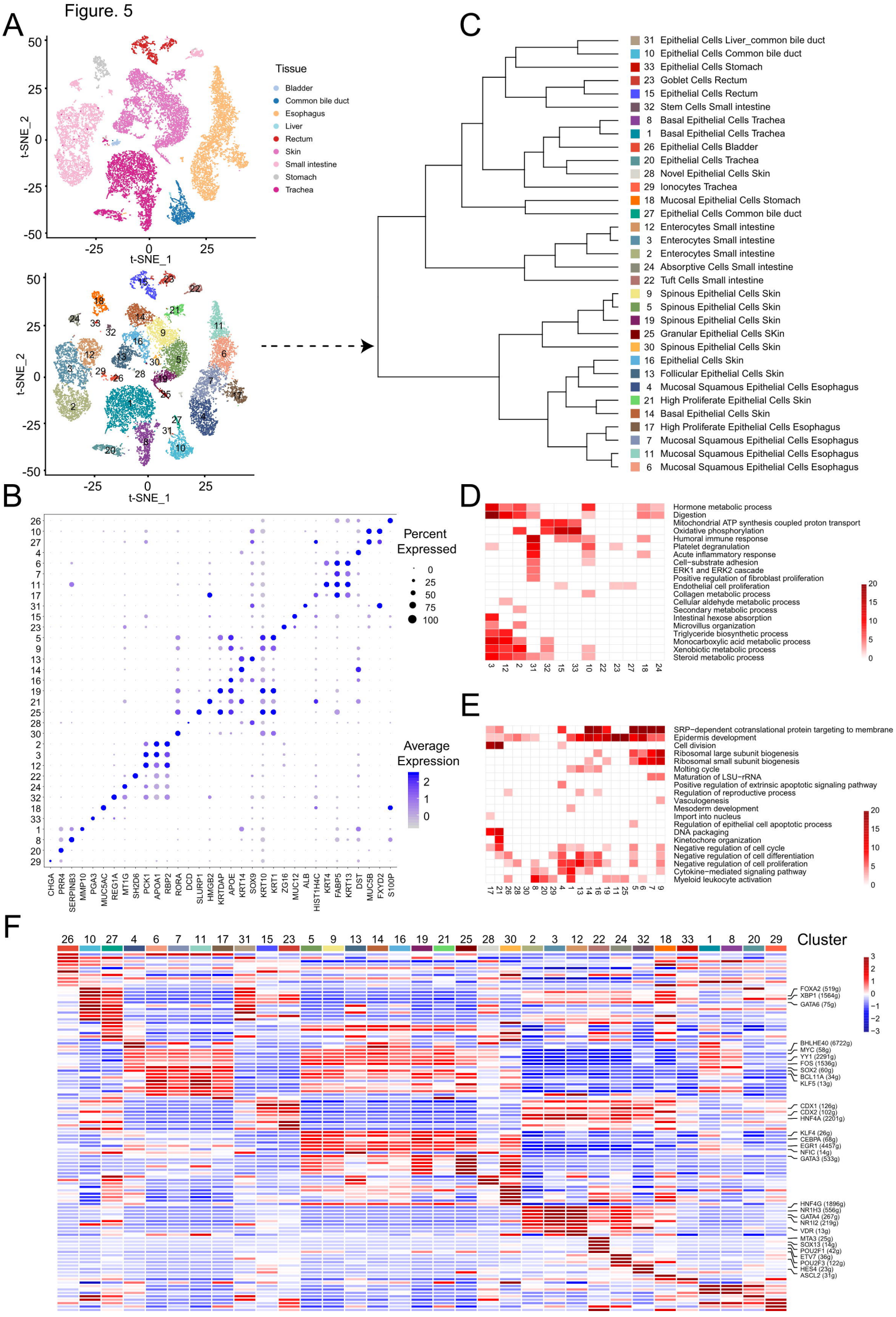
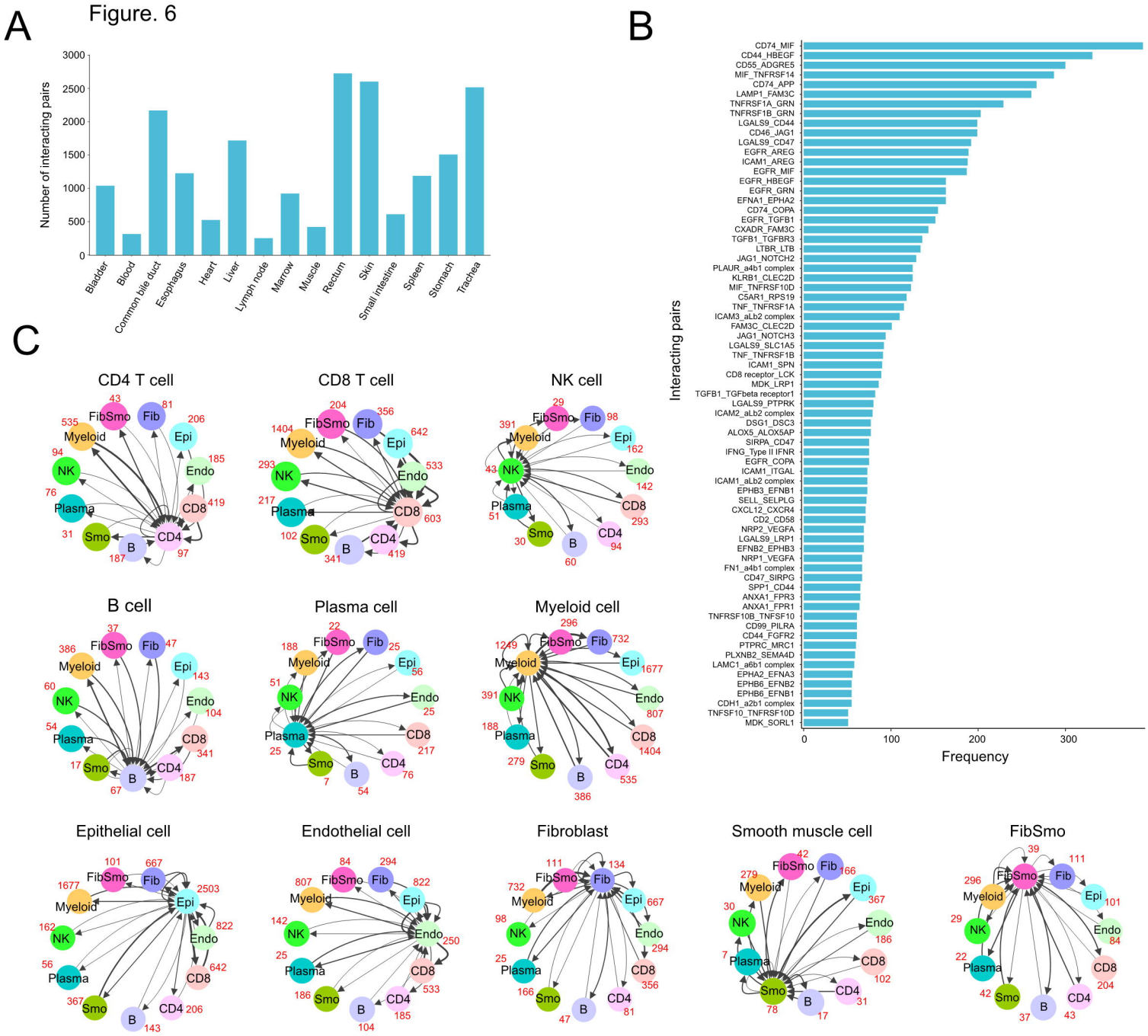

A
t-SNE_1
t-SNE_2
Tissue
Bladder
Common bile duct
Esophagus
Liver
Rectum
Skin
Small intestine
Stomach
Trachea
C
31 Epithelial Cells Liver_common bile duct
10 Epithelial Cells Common bile duct
33 Epithelial Cells Stomach
23 Goblet Cells Rectum
15 Epithelial Cells Rectum
32 Stem Cells Small intestine
8 Basal Epithelial Cells Trachea
1 Basal Epithelial Cells Trachea
26 Epithelial Cells Bladder
20 Epithelial Cells Trachea
28 Novel Epithelial Cells Skin
29 Ionocytes Trachea
18 Mucosal Epithelial Cells Stomach
27 Epithelial Cells Common bile duct
12 Enterocytes Small intestine
3 Enterocytes Small intestine
2 Enterocytes Small intestine
24 Absorptive Cells Small intestine
22 Tuft Cells Small intestine
9 Spinous Epithelial Cells Skin
5 Spinous Epithelial Cells Skin
19 Spinous Epithelial Cells Skin
25 Granular Epithelial Cells SKin
30 Spinous Epithelial Cells Skin
16 Epithelial Cells Skin
13 Follicular Epithelial Cells Skin
4 Mucosal Squamous Epithelial Cells Esophagus
21 High Proliferate Epithelial Cells Skin
14 Basal Epithelial Cells Skin
17 High Proliferate Epithelial Cells Esophagus
7 Mucosal Squamous Epithelial Cells Esophagus
11 Mucosal Squamous Epithelial Cells Esophagus
6 Mucosal Squamous Epithelial Cells Esophagus
B
Percent Expressed
Average Expression
CHGA
PRR4
SERPINB3
MMP10
PGA3
MUC5AC
REG1A
MT1G
SH2D6
PCK1
APOA1
RBP2
RORA
DCD
SLURP1
HMGB2
KRTDAP
APOE
KRT14
SOX9
KRT10
KRT1
ZG16
MUC12
ALB
HIST1H4C
KRT4
FABP5
KRT13
DST
MUC5B
FXYD2
S100P
D
Hormone metabolic process
Digestion
Mitochondrial ATP synthesis coupled proton transport
Oxidative phosphorylation
Humoral immune response
Platelet degranulation
Acute inflammatory response
Cell−substrate adhesion
ERK1 and ERK2 cascade
Positive regulation of fibroblast proliferation
Endothelial cell proliferation
Collagen metabolic process
Cellular aldehyde metabolic process
Secondary metabolic process
Intestinal hexose absorption
Microvillus organization
Triglyceride biosynthetic process
Monocarboxylic acid metabolic process
Xenobiotic metabolic process
Steroid metabolic process
E
SRP−dependent cotranslational protein targeting to membrane
Epidermis development
Cell division
Ribosomal large subunit biogenesis
Ribosomal small subunit biogenesis
Molting cycle
Maturation of LSU−rRNA
Positive regulation of extrinsic apoptotic signaling pathway
Regulation of reproductive process
Vasculogenesis
Mesoderm development
Import into nucleus
Regulation of epithelial cell apoptotic process
DNA packaging
Kinetochore organization
Negative regulation of cell cycle
Negative regulation of cell differentiation
Negative regulation of cell proliferation
Cytokine−mediated signaling pathway
Myeloid leukocyte activation
F
Cluster
FOXA2 (519g)
XBP1 (1564g)
GATA6 (75g)
BHLHE40 (6722g)
MYC (58g)
YY1 (2291g)
FOS (1536g)
SOX2 (60g)
BCL11A (34g)
KLF5 (13g)
CDX1 (126g)
CDX2 (102g)
HNF4A (2201g)
KLF4 (26g)
CEBPA (68g)
EGR1 (4457g)
NFIC (14g)
GATA3 (533g)
HNF4G (1896g)
NR1H3 (556g)
GATA4 (267g)
NR1I2 (219g)
VDR (13g)
MTA3 (25g)
SOX13 (14g)
POU2F1 (42g)
ETV7 (36g)
POU2F3 (122g)
HES4 (23g)
ASCL2 (31g)

Figure. 6


A
Number of interacting pairs
3000
2500
2000
1500
1000
500
0
Bladder
Blood
Common bile duct
Esophagus
Heart
Liver
Lymph node
Marrow
Muscle
Rectum
Skin
Small intestine
Spleen
Stomach
Trachea
B
Interacting pairs
CD74_MIF
CD44_HBEGF
CD55_ADGRE5
MIF_TNFRSF14
CD74_APP
LAMP1_FAM3C
TNFRSF1A_GRN
TNFRSF1B_GRN
LGALS9_CD44
CD46_JAG1
LGALS9_CD47
EGFR_AREG
ICAM1_AREG
EGFR_MIF
EGFR_HBEGF
EGFR_GRN
EFNA1_EPHA2
CD74_COPA
EGFR_TGFB1
CXADR_FAM3C
TGFB1_TGFBR3
LTBR_LTB
JAG1_NOTCH2
PLAUR_a4b1 complex
KLRB1_CLEC2D
MIF_TNFRSF10D
C5AR1_RPS19
TNF_TNFRSF1A
ICAM3_aLb2 complex
FAM3C_CLEC2D
JAG1_NOTCH3
LGALS9_SLC1A5
TNF_TNFRSF1B
ICAM1_SPN
CD8 receptor_LCK
MDK_LRP1
TGFB1_TGFbeta receptor1
LGALS9_PTPRK
ICAM2_aLb2 complex
DSG1_DSC3
ALOX5_ALOX5AP
SIRPA_CD47
IFNG_Type II IFNR
EGFR_COPA
ICAM1_ITGAL
ICAM1_aLb2 complex
EPHB3_EFNB1
SELL_SELPLG
CXCL12_CXCR4
CD2_CD58
NRP2_VEGFA
LGALS9_LRP1
EFNB2_EPHB3
NRP1_VEGFA
FN1_a4b1 complex
CD47_SIRPG
SPP1_CD44
ANXA1_FPR3
ANXA1_FPR1
TNFRSF10B_TNFSF10
CD99_PILRA
CD44_FGFR2
PTPRC_MRC1
PLXNB2_SEMA4D
LAMC1_a6b1 complex
EPHA2_EFNA3
EPHB6_EFNB2
EPHB6_EFNB1
CDH1_a2b1 complex
TNFSF10_TNFRSF10D
MDK_SORL1
0
100
200
300
Frequency
C
CD4 T cell
CD8 T cell
NK cell
B cell
Plasma cell
Myeloid cell
Epithelial cell
Endothelial cell
Fibroblast
Smooth muscle cell
FibSmo
Fib
Myeloid
Epi
Endo
NK
Plasma
CD8
CD4
Smo
B