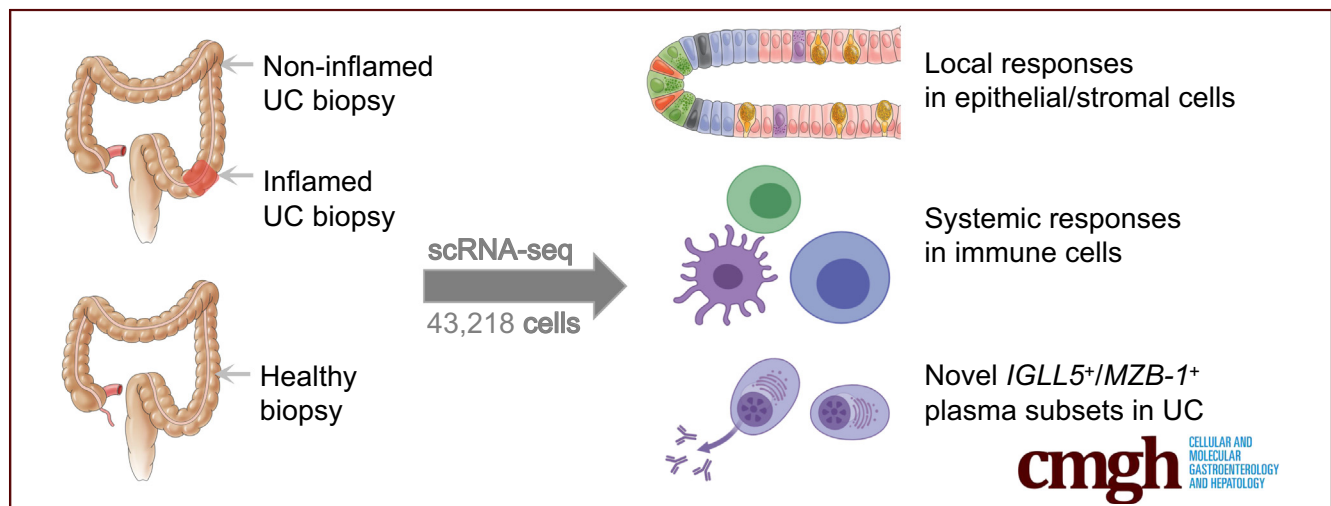


## ORIGINAL RESEARCH

## Identification of Novel Population-Specific Cell Subsets in Chinese Ulcerative Colitis Patients Using Single-Cell RNA Sequencing

Guang Li,<sup>1,a</sup> Bowen Zhang,<sup>2,a</sup> Jianyu Hao,<sup>1,a</sup> Xiaojing Chu,<sup>2,3</sup> Miriam Wiestler,<sup>4</sup> Markus Cornberg,<sup>2,4</sup> Cheng-Jian Xu,<sup>2,4,5</sup> Xinjuan Liu,<sup>1</sup> and Yang Li<sup>2,5</sup>

<sup>1</sup>Department of Gastroenterology, Beijing Chaoyang Hospital, Capital Medical University, Chaoyang District, Beijing, China; <sup>2</sup>Centre for Individualised Infection Medicine and TWINCORE, joint ventures between the Hannover Medical School and the Helmholtz Centre for Infection Research, Hannover, Germany; <sup>3</sup>Department of Genetics, University Medical Centre Groningen, University of Groningen, Groningen, the Netherlands; <sup>4</sup>Department of Gastroenterology, Hepatology and Endocrinology, Hannover Medical School, Hannover, Germany; and <sup>5</sup>Department of Internal Medicine and Radboud Center for Infectious Diseases, Radboud University Medical Center, Nijmegen, the Netherlands



## SUMMARY

We performed a single-cell RNA sequencing analysis of colon biopsies from Chinese ulcerative colitis (UC) patients. We provided novel insights of UC and its molecular signatures. It serves as an important reference for improving our understanding of genetic risks underlying UC.

**BACKGROUND & AIMS:** Genome-wide association studies (GWAS) and transcriptome analyses have been performed to better understand the pathogenesis of ulcerative colitis (UC). However, current studies mainly focus on European ancestry, highlighting a great need to identify the key genes, pathways and cell types in colonic mucosal cells of adult UC patients from other ancestries. Here we aimed to identify key genes and cell types in colonic mucosal of UC.

**METHODS:** We performed Single-cell RNA sequencing (scRNA-seq) analysis of 12 colon biopsies of UC patients and healthy controls from Chinese Han ancestry.

**RESULTS:** Two novel plasma subsets were identified. Five epithelial/stromal and three immune cell subsets show significant difference in abundance between inflamed and non-inflamed samples. In general, UC risk genes show consistent expression alteration in both Immune cells of inflamed and non-inflamed tissues. As one of the exceptions, IgA deficiency, marking the signal of immune dysfunction, is specific to the inflamed area. Moreover, Th17 derived activation was observed in both epithelial cell lineage and immune cell lineage of UC patients as compared to controls, suggesting a systemic change of immune activities driven by Th17. The UC risk genes show enrichment in progenitors, glial cells and immune cells, and drug-target genes are differentially expressed in antigen presenting cells.

**CONCLUSIONS:** Our work identifies novel population-specific plasma cell molecular signatures of UC. The transcriptional signature of UC is shared in immune cells from both inflamed and non-inflamed tissues, whereas the transcriptional response to disease is a local effect only in inflamed epithelial/stromal cells. (*Cell Mol Gastroenterol Hepatol* 2021;12:99–117; <https://doi.org/10.1016/j.jcmgh.2021.01.020>)

**Keywords:** Ulcerative Colitis; Single-Cell RNA Sequencing; Genome-Wide Association Studies.

Ulcerative colitis (UC) is a chronic inflammatory disease, which is characterized by relapsing and remitting mucosal inflammation in the rectum and extending to proximal segments of the colon. The pathogenesis of UC is complex and multifactorial,<sup>1</sup> including genetic predisposition,<sup>2</sup> epithelial barrier defects,<sup>3,4</sup> and deregulated immune responses.<sup>5</sup> In general, UC is thought to arise from an inappropriate activation of the intestinal mucosal immune system in response to commensal bacteria in a genetically susceptible host.<sup>6</sup> The breakdown of the epithelial barrier and mucosal immune barrier homeostasis at the cellular level has been revealed to play an important role in the onset of UC.<sup>7,8</sup>

UC is common in industrialized locations including North America and Western Europe.<sup>9</sup> Recently, there is also an increasing incidence of UC in Asia due to urbanization.<sup>10</sup> The recent epidemiological report showed that India and China had the highest inflammatory bowel disease (IBD) incidence in Asia. In China, the incidence of UC was reported to be positively associated with gross domestic product.<sup>11</sup> Genome-wide association studies (GWASs) and transcriptome analyses have been widely used to dissect the molecular mechanisms of UC, by identifying risk alleles as well as transcriptional alterations aiming at defining the functional consequences of associated alleles for both coding and noncoding genetic variation.<sup>12,13</sup> To better understand the pathogenesis of UC, recent studies have been focused on the transcriptome at a cellular resolution and states within the tissue of IBD lesions.<sup>14–17</sup> However, most of the current studies were performed in individuals of European ancestry, with only 1 study focusing on pediatric-onset colitis in Chinese,<sup>17</sup> highlighting a great demand for studying on the key genes, pathways and cell types in colonic mucosal cells of adult UC patients from other ancestries.

Here, we performed a single-cell RNA sequencing (scRNA-seq) analysis of 12 colon biopsies from 5 UC patients including 4 inflamed (UC) biopsies, 4 noninflamed (self-control [SC]) biopsies and 4 healthy biopsies (HCs) from healthy individuals from Chinese Han ancestry (Figure 1A, Supplementary Table S1A), to uncover the pathogenesis of UC on a cellular level.

## Results

### Two Novel Plasma Cells Were Identified From All Colon Biopsies

After quality control, 43,218 cells with an average of 1053 genes per cell remained and clustered into 21 subpopulations (Figure 1B), in which cells from each of the UC, SC, and HCs were mapped to a comparable space in UMAP (Uniform Manifold Approximation and Projection) after batch correction (Figure 2). We calculated marker genes for each cluster (Supplementary Table S1B), then annotated the clusters with both data-derived methods (Figures 1C and 3–5) and literature-derived methods (Figures 1D, 6, and 7), and ended up with 10 epithelial and stromal cell types (eg,

enterocyte, enterocyte progenitors, goblet, goblet progenitors, LGR5+ stem cell, CLP/Paneth-like cells, fibroblasts, TRMP5+ tuft cell, glial cells and endothelial cells) and 11 immune cells (eg, naive T cells, memory T cells, CD8+ T cell/natural killer cell, CD8+ T cell, naive B cell, memory B cell, monocytes/dendritic cells, mast cells, and 3 clusters of plasma cells). Among them, 16 of 21 subsets were replicated in the scRNA-seq study from American/European descent.<sup>14</sup> Interestingly, we observed 2 novel immune subsets (plasma-2 and plasma-3 cells) in our data.

The proportions of 5 epithelial and stromal and 3 immune cell subsets significantly differed between UC and SC (or HCs). Glial cells, fibroblasts, goblet progenitors, and enterocytes with their progenitors were significantly decreased in inflamed UC tissues. CD8+ T cell as well as 2 types of plasma cells were increased in UC (Figure 1E and F).

On the one hand, tuft cells are the chemosensory cells in the gut and are enriched for taste-sensing molecules.<sup>18</sup> We observed a cluster of mature tuft cells (named as TRMP5+ tuft), which highly expressed immune-related genes (*AZGP1*, *PTPN18*, and *BMX*) and neuronal signaling genes (*AVIL*, *HTR3E*, and *ITPR2*). Of note, TRMP5+ tuft cell also showed high expression level on *HPGDS*, *ALOX5*, and *PTGS1*, which function in the metabolism of arachidonic acid and prostaglandin.<sup>19–21</sup> Moreover, TRMP5+ tuft cells also expressed *IL-17RB*, which may mediate the cross-talk with *ILC2*,<sup>22</sup> and acted as a marker of tuft cell-like human colorectal cancer stem cells,<sup>23</sup> as previously identified in the Chinese pediatric UC.<sup>17</sup> Notably, we showed a significantly lower abundance of TRMP5+ tuft cells in UC compared with that in SC (paired Student's *t* test,  $P = .048$ ) (Figure 1E and Supplementary Table S2A).

On the other hand, 2 of 3 plasma subsets were observed with a higher proportion in UC, compared with that in either HCs or SC (Figure 1F and Table S2A). Among the mentioned 2 plasma subsets, plasma-2 cells specifically expressed the *MZB-1* gene (Figure 8A, Supplementary Table S1B and C), which was significantly upregulated in UC compared with HCs ( $P_{\text{adjusted}} = 1.18 \times 10^{-65}$ ) or SC ( $P_{\text{adjusted}} = 3.45 \times 10^{-37}$ ). Plasma-3 cells specifically expressed the *IGLL5* gene, which was also significantly upregulated in UC compared with HCs ( $P_{\text{adjusted}} = 1.16 \times 10^{-18}$ ) or SC ( $P_{\text{adjusted}} = 1.83 \times 10^{-15}$ ). Consistent with previous findings, the upregulation of *MZB-1* and *IGLL5* in

<sup>a</sup>Authors share co-first authorship.

**Abbreviations used in this paper:** CD, Crohn's disease; DEG, differentially expressed gene; DGE, digital gene expression; GWAS, genome-wide association study; HC, healthy biopsy; IBD, inflammatory bowel disease; Ig, immunoglobulin; IL, interleukin; ISC, intestinal stem cell; MHC, major histocompatibility complex; PBMC, peripheral blood mononuclear cell; PBS, phosphate-buffered saline; PC, principal component; PCA, principal component analysis; SC, self-control; scRNA-seq, single-cell RNA sequencing; TNF- $\alpha$ , tumor necrosis factor alpha; UC, ulcerative colitis.

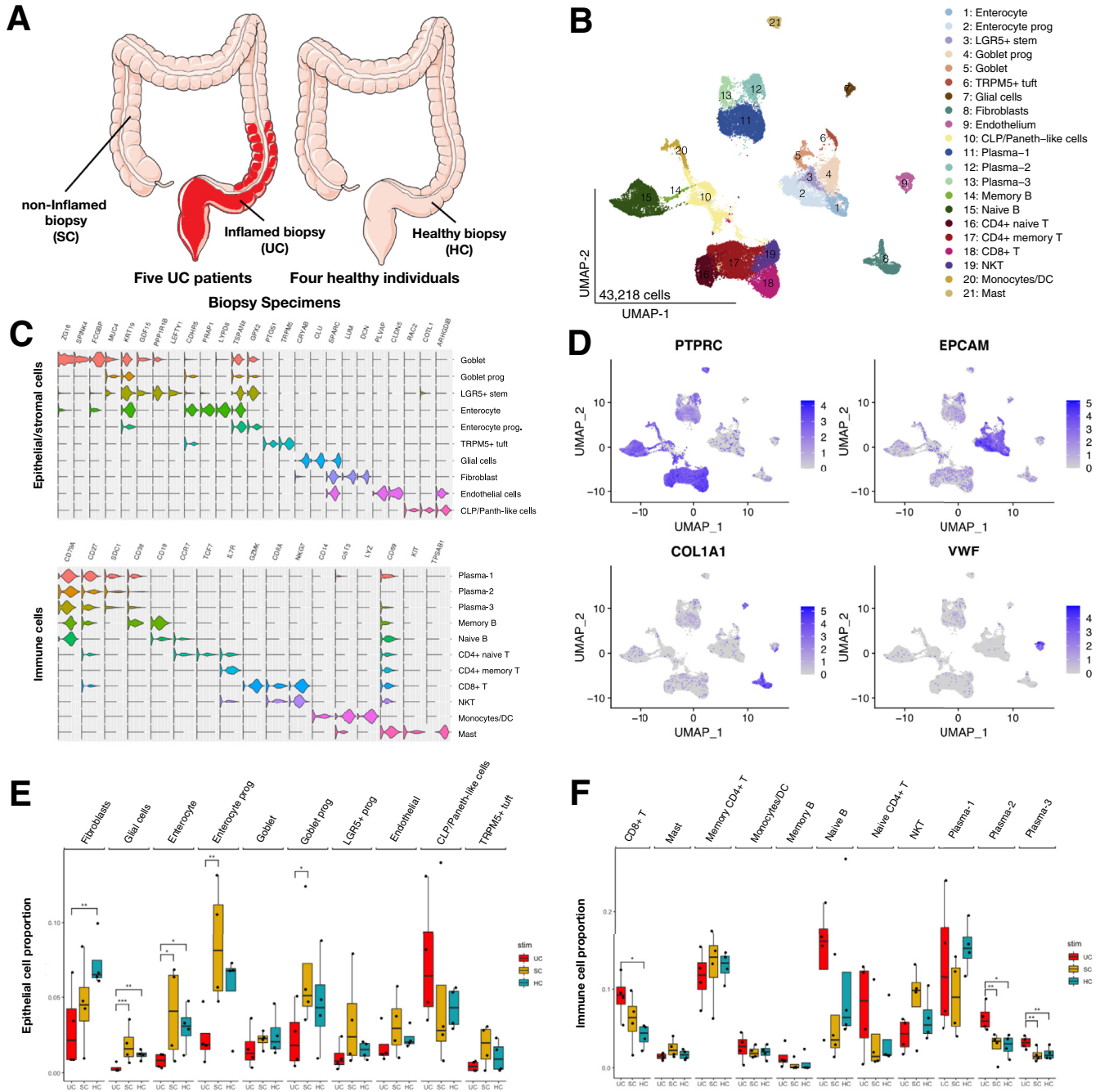


Most current article

© 2021 The Authors. Published by Elsevier Inc. on behalf of the AGA Institute. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

2352-345X

<https://doi.org/10.1016/j.jcmgh.2021.01.020>



**Figure 1. Single-cell expression atlas and cell typing in biopsies of Chinese UC patients and control samples.** (A) Experimental design. Fresh biopsy specimens were disassociated and single-cell suspensions were obtained from 4 samples from inflamed sigmoid colon (UC) with the noninflamed ascending colon biopsy specimens (SC), and 4 healthy samples (HCs). (B) The UMAP (Uniform Manifold Approximation and Projection) plot identifies 10 epithelial cell clusters and 11 immune cells from 12 colon biopsies. (C) Violin plots showing the expression distribution of selected marker genes across cell clusters. (D) UMAP shows the lineage markers, *PTPRC* for immune cells, *EPCAM* for epithelial cells, *COL1A1* for fibroblasts, and *VWF* for endothelial cells. (E, F) Boxplots showing percentage of epithelial cell (E) and immune cell (F) clusters of total specimens (per biopsy) in inflamed samples relative to noninflamed samples and healthy samples (significant changes of Dirichlet-multinomial regression adjusted with the Benjamini-Hochberg method were marked out). \* $P < .05$ , \*\* $P < .01$ .

Crohn’s disease (CD) has been reported in mesenteric adipose tissue.<sup>8</sup> Furthermore, genes coexpressed with *MZB-1* and *IGLL5* in those cell types were enriched in endoplasmic reticulum, which play important roles in cytosol transport,

suggesting its involvement in the inflammatory cascades of plasma cells (Figure 11B).

In order to directly compare our plasma signatures with those of American or European ancestry, we intersected



identified DEGs between UC and SC and HCs in plasmas of our study with those found in Smillie et al (Supplementary Table S3A and B).<sup>14</sup> On the one hand, most upregulated genes ( $n = 270$  of 502) found in our study, such as *XBP-1* and major histocompatibility complex (MHC) class II genes, were found consistently upregulated in Smillie et al (Supplementary Table S3C).<sup>14</sup> On the other hand, 2 plasma maker genes, *IgLL5* and *MZB-1*, showed upregulation in UC patients in our study, but were found to be downregulated (*IgLL5*) and not changed (*MZB-1*) in Smillie et al.<sup>14</sup> This observation indicates that expression alteration in these 2 genes could be specific to the plasma of Chinese UC patients (Supplementary Table S3).

### The Antigen Presentation Pathway and IL-17 Signaling Pathway Were Activated in the Inflamed Tissue of UC

In epithelial and stromal lineages, upregulated genes between inflamed tissues and HCs were enriched in the antigen processing and presentation pathway and MHC class II complex activity. These pathways were also showed up in UC compared with SC (Figure 9A and B).

Differentially expressed genes (DEGs) between UC and HCs in LGR5+ stem cells, enterocyte progenitors, and goblet progenitors overrepresented the functions of rejective immunity, but mature cells (enterocyte and goblet) lose the functional difference of rejective reaction. Of note, the majority of genes involved in MHC class II molecules, such as *HLA-DRA*, *HLA-DRB1*, and *HLA-DRC*, were upregulated in UC, but the MHC class I molecular *HLA-B* was downregulated.

Moreover, the interleukin (IL)-17 signaling pathway was significantly enriched in the upregulated genes of goblet progenitor, endothelial cell, and fibroblasts in the comparison of UC vs HCs, suggesting that the intestinal epithelial barrier was involved in IL-17 cytokine responses in UC. One of the major genes of the IL-17 signaling pathway in these cell types was *LCN2* (Figure 9C). *LCN2*, a bacteriostatic molecule, has tissue destructive effects and is proinflammatory with chemo-attractant molecule binding properties.<sup>24</sup> Earlier studies, using both DNA microarrays and RNA-seq data, showed that the *LCN2* is among the top 10 upregulated genes in UC and is correlated with disease severity.<sup>25-27</sup> In addition, IL-17A showed a synergistic effect with IL-22 and tumor necrosis factor alpha (TNF- $\alpha$ ) in inducing colonic epithelial expression of *LCN2*.<sup>28</sup>

Next, we investigated whether the DEGs showing consistent pattern between Chinese UC patients and patients of American or European ancestry. We intersected DEGs found between UC and SC and HCs in epithelial cells of our study with those found in Smillie et al.<sup>14</sup> Most upregulated genes ( $n = 299$  of 517) found in our study also showed upregulation in the UC patients of Smillie et al (Supplementary Table S3C), including *LCN2*, *JUN*, and MHC class genes, etc. The upregulated genes found in our study were enriched in pathways like translation initiation and structural molecule activity (Figure 10).

### The Th17 Cell Was the Major Immune Characteristic of UC, But Immunoglobulin A Was the Key Local Immune Component of Inflamed Tissue

Similar to the activation of IL-17 signaling pathway in the epithelial lineages, the Th17 cell differentiation, as well as the T cell receptor signaling and TNF signaling pathway, were activated in CD4+ cell lineages in UC compared with HCs. In general, in immune-mediated inflammation, the presented antigen leads to the differentiation of CD4+ helper T cells and stimulates Th17 cells to produce proinflammatory cytokines, such as TNF and IL-17.<sup>29</sup> On the other hand, we found that the Th17 cell differentiation and its cytokines related pathways were activated in natural killer T cells, which is in line with the previous finding that CD8+IL-17+ cells were increased in American UC patients.<sup>2</sup> Additionally, we found that *NFKBIA* and *JUN* were predominantly presented in plasma cells, mast cells, monocytes, and dendritic cells in HCs, but were only expressed in B cells and T cells in both UC and SC, whereas additional MHC class II genes (eg, *HLA-DRA*) were found in CD8+ T and natural killer T cells (Figure 11C).

The intestinal immunoglobulin A (IgA) production was found to be enriched in the upregulated genes of goblet and enterocytes cells, their progenitor cells, and glial cells in comparison between UC and SC (Figure 11B). Most of the IgA molecules presented at mucosal sites, where they are produced by locally residing plasma cells.<sup>30</sup> As a key local immune component, IgA protects the integrity of intestinal mucosal barrier by coating the bacteria.<sup>31</sup> In line with these findings, abnormal coating capacity of IgA was reported in UC in Chinese young patients.<sup>32</sup>

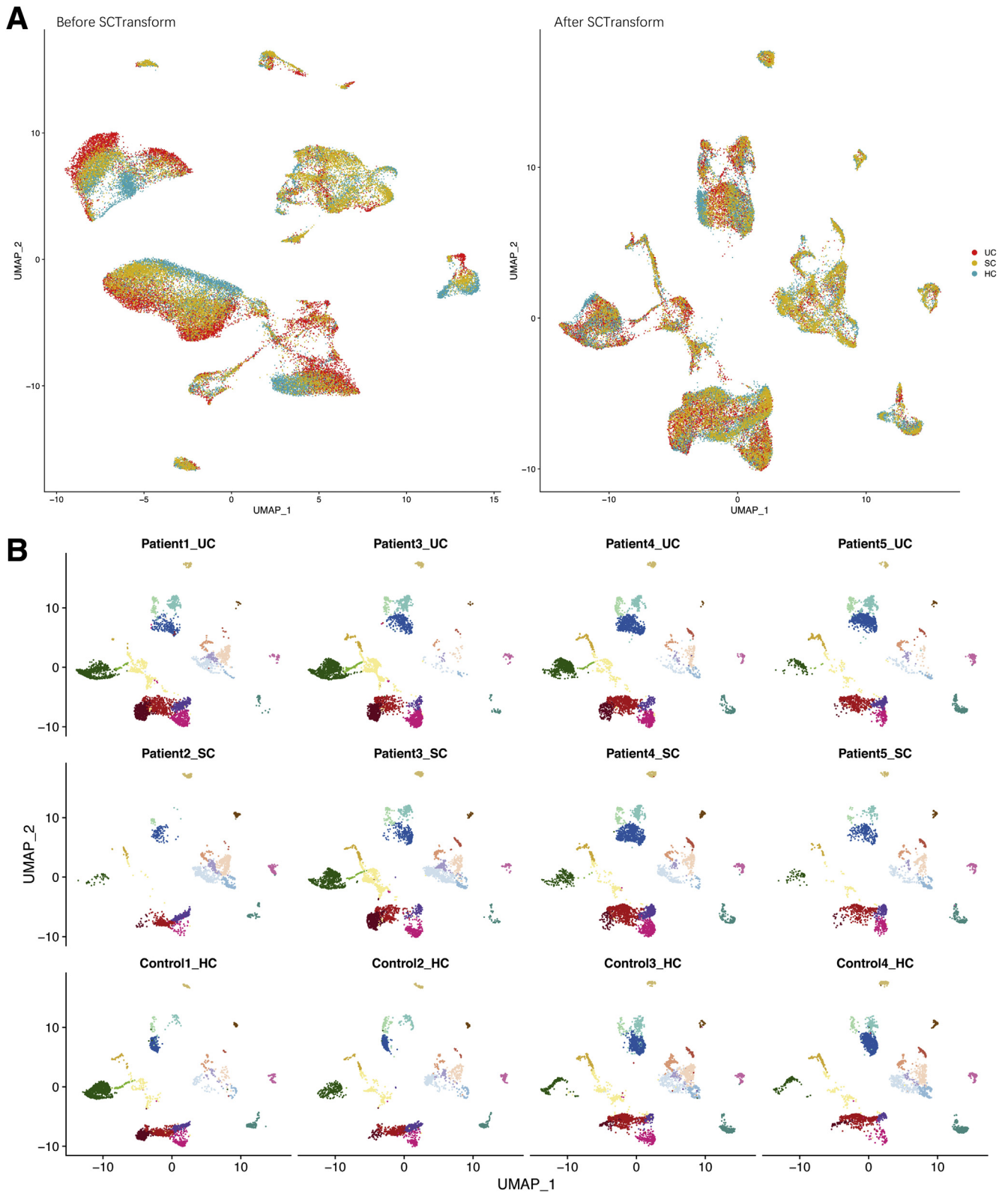
### A Total of 195 UC Risk Genes Were Enriched in the Epithelia Progenitors and Immune Cells With Druggable Targets in Antigen-Presenting Cells

Among 511 UC risk genes published in the GWAS catalog (<https://www.ebi.ac.uk/gwas/>), 141 were found in 195 DEGs between UC and control subjects, of which 90% showed upregulation in UC. Compared with the number of height-associated loci overlapped with our DEGs, UC risk loci showed enrichment in epithelial cells (goblet, goblet progenitors, LGR5+ stem cell, and enterocyte), endothelial cell, glial cells, fibroblasts, and various immune cells (plasma-1, naive and memory B cell, CD4+ T cell, and natural killer T cell) (Supplementary Table S2B).

Among these UC risk DEGs, 8 genes were previously reported in the UC GWASs in the Asian population: *CFB*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DRA*, *HLA-DRB1*, *IRF8*, *PTPRC*, and *SLC26A3*. Particularly, *HLA-C*, *HLA-DRA*, and *HLA-DRB1* were upregulated in the epithelial cells of UC compared with both SC and HCs (Figure 12A and B and Figure 13), suggesting an increased activity of antigen presentation during inflammation in the inflamed tissue, which is consistent with the enrichment pathways of DEGs reported previously.

In T cell lineages, *FYN*, *PTPRC*, and *CDC42SE2* were increased in UC, and they are known to be involved in the T cell receptor signaling via regulating the receptor-like

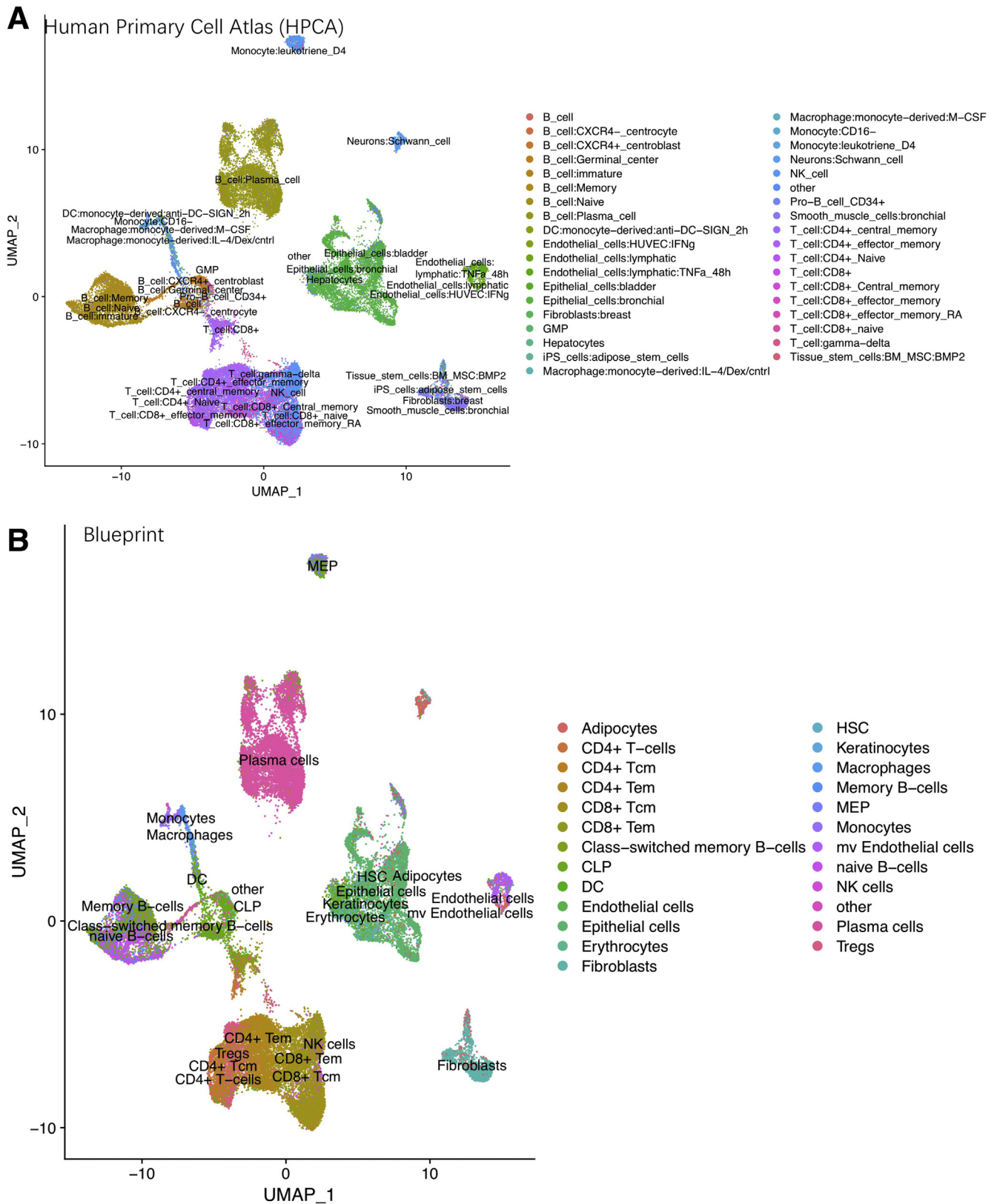


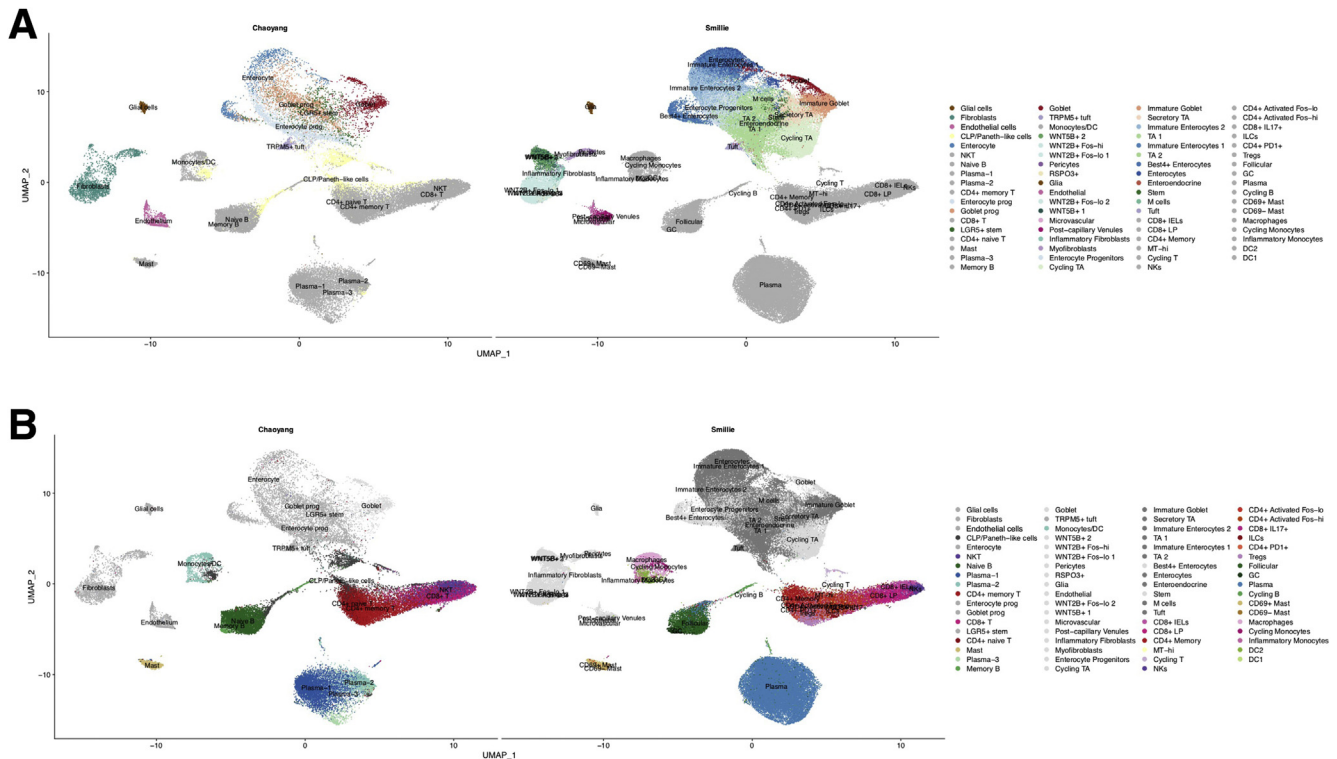


**Figure 2.** The UMAP plots for batch correction and donor effects. (A) UMAP before and after SCTransform batch correction. (B) UMAP of each donor shows minimal donor effects after batch correction.

tyrosine phosphatase.<sup>33</sup> Similarly, immunosuppressive costimulatory molecules, *CTLA4* with its receptor *ICOS*, were increased in CD4<sup>+</sup> naive T cell and CD4<sup>+</sup> memory T cell,

respectively. Of note, *STAT3* is a recognized proinflammatory transcripts.<sup>34</sup> It is known that the differentiation of Th17 cells typically requires the cytokines signaling





**Figure 4.** UMAP shows the overlap of this study and Smillie et al<sup>14</sup> separated by (A) epithelial and stromal cells and (B) immune cells.

via the *STAT3* transcription activator.<sup>35</sup> *STAT3* was upregulated in CD4+ naive T cell of UC and SC, which was also consistent with the change of Th17 cell and IL-17 signal in systemic inflammation.

Our analysis showed that many MHC molecules were upregulated in the epithelial and stromal cells of UC compared with either HCs or SC. In immune cells, however, many dysregulated risk genes were identified between UC and HCs but not between UC and SC. For example, *LTB*, *DUSP1*, and *IRF8* of memory B cell and T cells were increased in UC compared with HCs, but not with SC. Therefore, we performed a systematic comparison of significance of differentially expressed UC risk genes in immune and epithelial and stromal cells. It reveals contrasting patterns: the identified differentially expressed risk genes between UC and HCs are more likely to be differentially expressed between UC and SC in epithelial and stromal cells than in immune cells ( $P_{\chi^2 \text{ test}} = 1.35 \times 10^{-3}$ ) (Figure 12A and B and 13).

Interestingly, we also noticed that 25% of the DEGs are druggable targets.<sup>36</sup> Of note, the enrichment of DEGs in the known UC drug targets was found for basiliximab ( $P = .049$ ) and adalimumab ( $P = .042$ ) in glial cells, and abatacept in monocytes and dendritic cells ( $P = .008$ ).

## Discussion

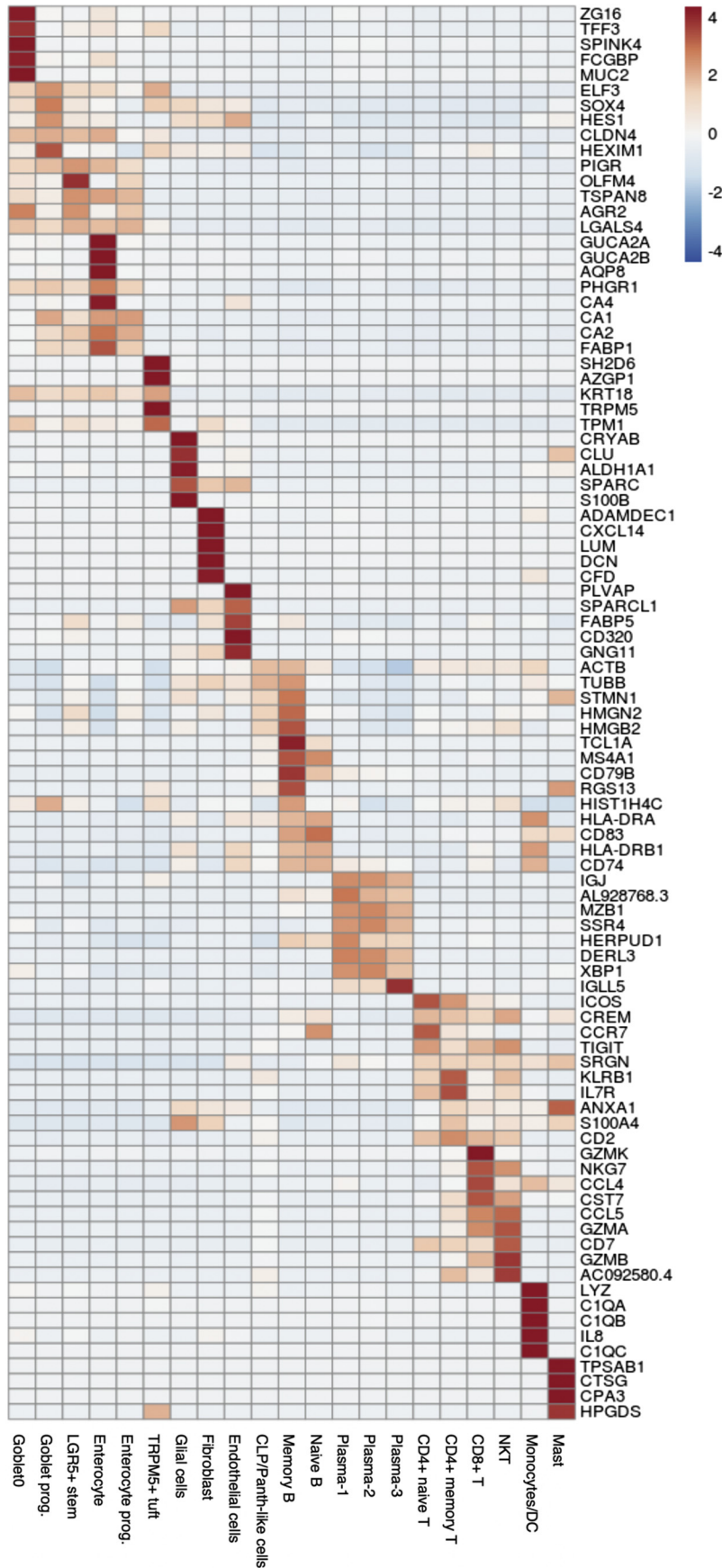
Here, we demonstrate novel molecular signatures of adult Chinese UC patients at a single-cell level. Most of the

cell types we identified can be replicated in the previous studies in American UC patients. The enteric nervous system plays a pivotal role in rectifying and orchestrating the inflammatory responses in gut tract.<sup>37</sup> Enteric glial cells have been recognized as antigen presenting cells, express substance P, and produce  $\text{TNF-}\alpha$ , IL-1 $\beta$ , and IL-6, which could induce the activation of mast cells, macrophages, and T cells, and promote lymphocyte proliferation.<sup>38,39</sup> We identified the decreased glial cells in inflamed tissues compared with both SC and HCs, similar to American UC patients<sup>14</sup> but not detected in Chinese pediatric UC patients.<sup>17</sup>

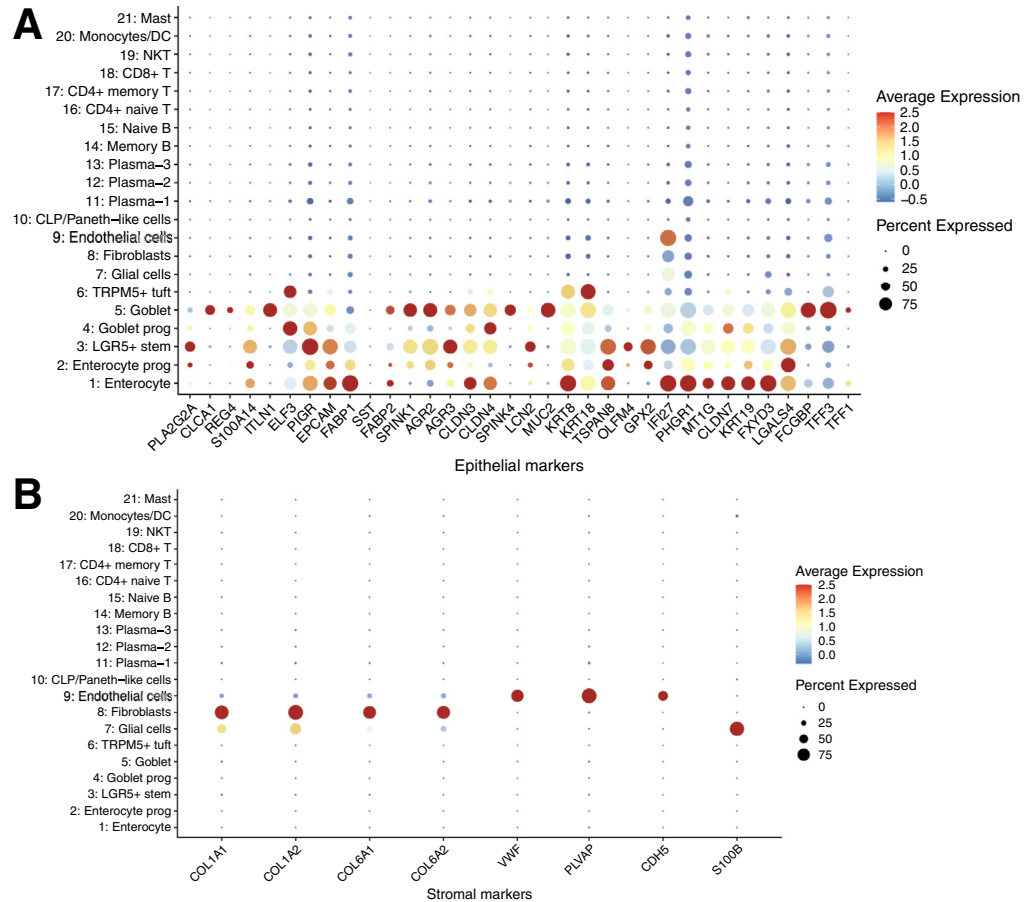
Tuft cells have also been identified from UC patients. Theoretically, tuft cells are absent from the stratified squamous epithelium of the anal canal and esophagus but may increase following replacement with a metaplastic, intestine-like columnar epithelium.<sup>21</sup> A tuft cell signature based on bulk profiles of TRPM5+ tuft cells contained both neuronal and inflammatory gene programs; this could reflect either coexpression in the same cells or distinct subsets.<sup>40</sup> TRPM5 plays a crucial role for chemosensation in promoting tuft cell expansion in response to infection.<sup>22</sup> TRPM5+ tuft cells also interact closely with immune cells, playing a crucial role in the cellular regulatory network coordinating responses to luminal parasites.<sup>41</sup>

In the epithelial lineages, most progenitor cells were decreased in UC. We also identified LGR5+ stem cells, the major intestinal stem cell (ISC), which play a key role in regeneration of intestinal injury,<sup>42</sup> in our data. In general, the intestinal epithelium is maintained by long-lived ISCs,





**Figure 5. Heatmap showing the top 5 DEGs in each cluster.** DEGs were obtained by comparing expression level in cells of one cluster against to that in the rest of cells.



**Figure 6.** Dot heatmap showing the expression of epithelial and stromal lineage markers.

residing the crypt base, the specifically expressed marker is LGR5+. <sup>43</sup> Above the ISC zone, there are short-lived progenitors that normally give rise to lineage-specific differentiated cell types but can de-differentiate into ISCs in the case of injury. <sup>44</sup> The short-lived enterocyte precursors just serve as a large reservoir of potential stem cells during crypt regeneration. <sup>45</sup> Thus, the reducing enterocyte progenitors and matured cells suggested the deficiency of mucosal regeneration in inflamed UC.

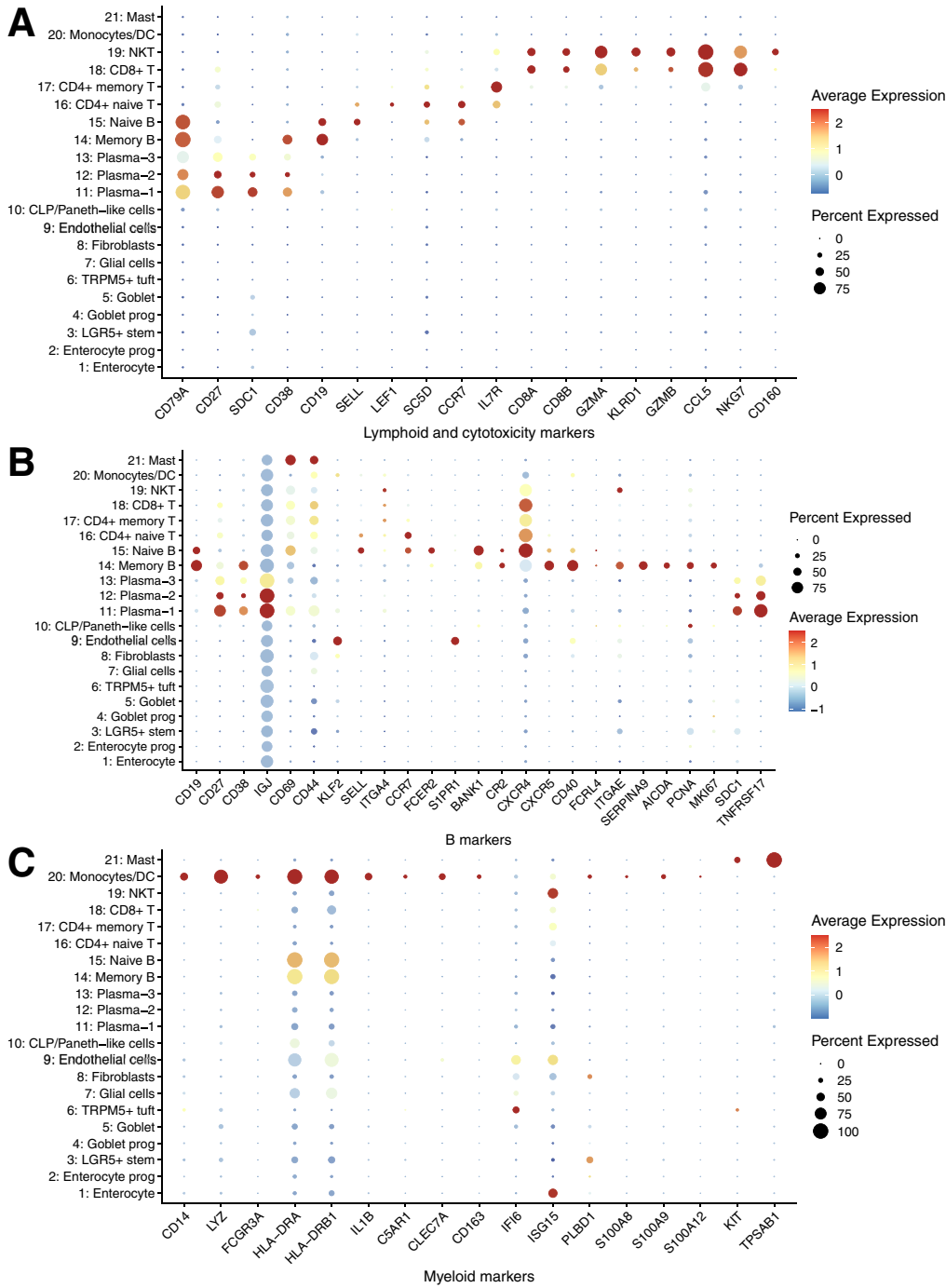
Pericytes and enteroendocrine cells have been identified in the previous study by Smillie et al <sup>14</sup> but were largely missed in our data. A small number of pericytes and enteroendocrine cells were found but mixed with fibroblasts and LGR5+ stem cells, respectively. The lack of resolution in detecting pericytes and enteroendocrine cells is likely due to the smaller sample size of our study compared with that of Smillie et al. <sup>14</sup>

Epithelial barrier and immune barrier defects are strongly implicated in the pathogenesis of UC with significant dysbiosis, <sup>1</sup> although there were no UC-related specific bacteria identified. <sup>46</sup> The DEGs were enriched in the function of antigen presentation and MHC class II complex activity in epithelial lineages of UC, and DEGs of the most progenitors were involved in the function of rejective immunity. Most differential genes code the antimicrobial proteins, such as LCN, a bacteriostatic molecule, involved in the

antimicrobial immune response. *CXCL1* and *CXCL2* are involved as the antimicrobial humoral immune response mediated by antimicrobial peptide, acting as the recruiters of immune cells. Both of them are the risk genes and were verified upregulated in UC. <sup>47</sup>

Type 17 immunity involved in the inflammation in the pathogenesis of ulcerative colitis has been implicated, <sup>48,49</sup> and type 17 immunity were derived by anticommensal response or antigen presentation in UC. <sup>50,51</sup> We found in the DEGs in epithelial and stromal lineages that Th17 signals were increased not only in the inflamed tissue, but also in the uninfamed control tissue, with the bacteriostatic molecule *LCN2* expression. In addition, Th17 cell differentiation and T cell receptor signaling were enriched in the CD4+ T cell lineage, which were consistent with the functional change in epithelial lineages. Moreover, the UC risk gene *STAT3* was also increasingly expressed in CD4+ naive T cell of UC and SC.

Many studies have reported that the increased levels of Th17 cells and IL-17 not only in the intestinal mucosa but also in the peripheral blood mononuclear cells (PBMCs) and serum of active UC in Chinese population. <sup>52,53</sup> It is also known that the level of Th17 cell was increased in PBMCs, and the expression of IL-17A messenger RNA were increased in the PBMCs, mesenteric lymph node, and lamina propria of colon of dextran sulfate sodium colitis mice. <sup>54</sup>



**Figure 7.** Dot heatmap showing the expression of lymphoid, B cells, and myeloid lineage markers.

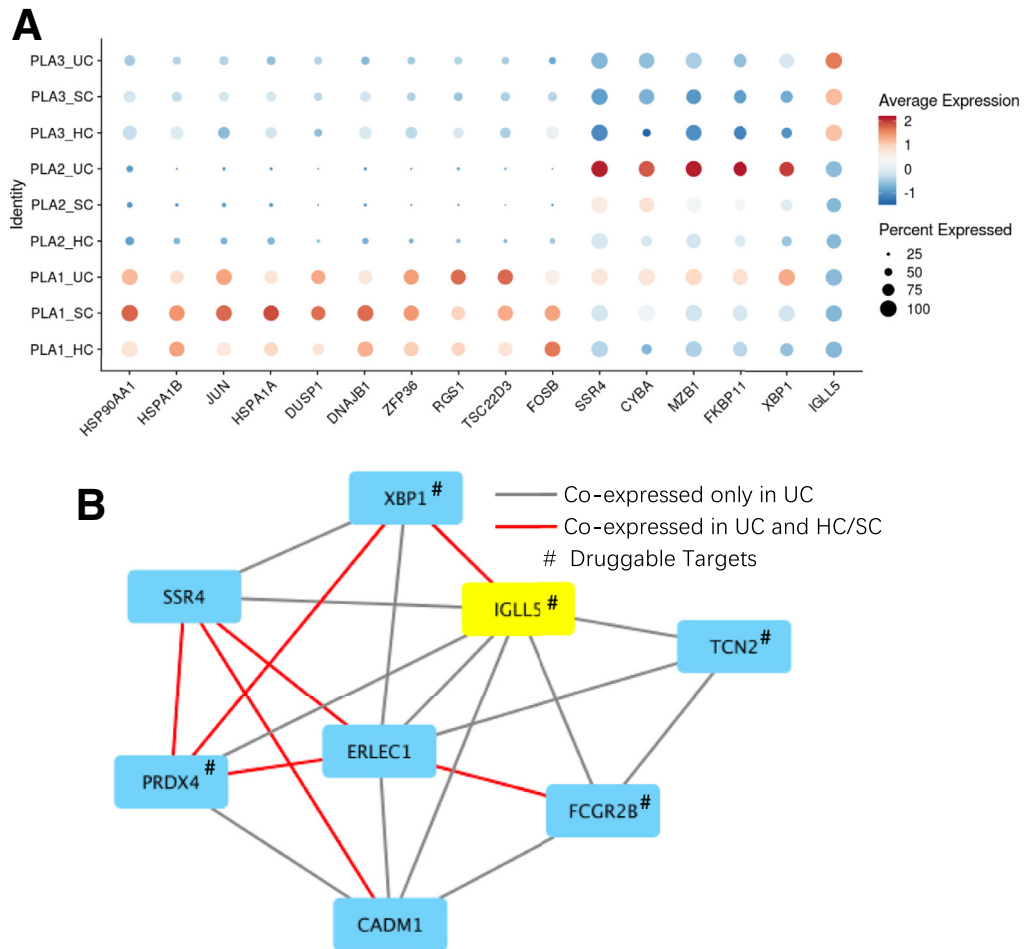
Above all, type 17 immunity is one of the important systemic changes in UC.

In epithelial lineages, upregulated genes between inflamed tissues and HCs were replicated in the comparison between UC and SC, but in immune cell lineages, most DEGs identified between inflamed intestines and HCs cannot be found as DEGs in the comparison between inflamed intestines and SC. These observations indicated the local intestinal damage as well as a systemic inflammation were involved in the pathogenesis of UC.<sup>55</sup> Particularly in line with that, type IL-17 immunity was activated overall the

colon, and the defective IgA secretion has been confirmed as a local immunological character of UC.<sup>9</sup>

In summary, most cell types could be found in the colon of both Chinese UC patients and healthy control subjects, while we identified 2 novel plasma subsets in our data. The transcriptional signature of UC is shared in immune cells from both inflamed and noninflamed tissues, except the defection of IgA, which is a located immunity dysfunction of inflamed area. However, in epithelial and stromal cells, the transcriptional response to disease is a local effect seen only in the inflamed biopsy. The activation of Th17 driven by





**Figure 8. Gene signatures found in 3 plasma subsets.** (A) Dot heatmap showing the expression levels of marker genes detected in the 3 plasma cell clusters, and their expression differences in the inflamed samples (UC) compared with non-inflamed samples (SC) and healthy samples (HCs). (B) Coexpression network showing the difference of coexpressed genes to plasma-3 specific marker gene (*IGLL5*) between inflamed UC and HCs or SC.

anticommensal inflammatory response is the core of the pathogenesis of UC, nearly including all the changes of epithelial cell lineage and immune cell lineage. In addition, the drug target genes were differentially expressed in antigen presenting cells. Our study serves as an important reference of the molecular mechanism behind the genetic risks of UC from transcriptional aspects, and identifies cell type-specific drug targets.

## Materials and Methods

### Patient Selection and Sample Collection

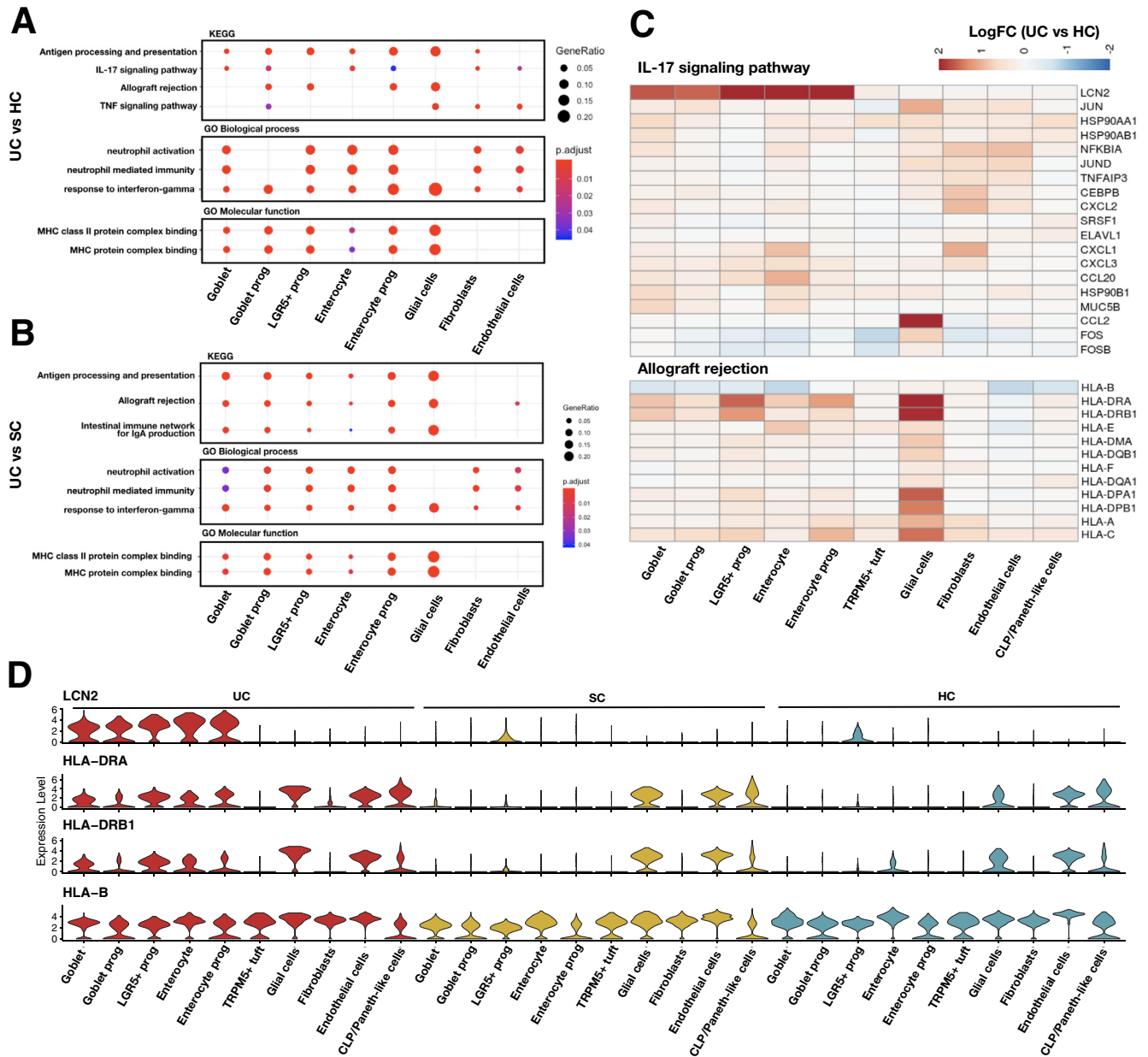
The biopsy samples were collected from 4 male and 1 female Han Chinese patients at the Department of Gastroenterology, Beijing Chaoyang Hospital of Capital Medical University (Beijing, China). All patients have been diagnosed as left-sided moderate UC for at least 3 years, with relapsing course, without receiving any treatment in recent 3 months. None of them had undergone surgical resections. UC was diagnosed by the conventional clinical, radiological, and endoscopic features, and eventually confirmed by histological examination of colonic biopsies.<sup>56</sup> Four Han Chinese healthy control subjects were enrolled at the Health Examination Center of Beijing Chaoyang Hospital of Capital Medical University. Among them, 3 healthy control subjects were age- and sex-matched with 3 UC patients. Written informed

consent was obtained and ethical approval was granted by the ethics committee of the Beijing Chaoyang Hospital, Capital Medical University. Characteristics of all patients are represented in [Supplementary Table S1A](#). For each of the 5 patients, 1 pinch biopsy specimen was collected from the inflamed sigmoid colon (the most common site of inflammation in UC) as the UC group, and 1 pinch biopsy specimen from the normal ascending colon of patients served as SC (SC group), as well as 4 pinch biopsy specimens from sigmoids of 4 healthy volunteers served as healthy control (HC group). Biopsy specimens were collected into RPMI 1640 medium on ice and processed immediately.

### Tissue Processing

Tissues were washed twice with phosphate-buffered saline (PBS). The biopsy specimens were cut into 1-mm<sup>3</sup> pieces using sterile scalpel blades and put into a Petri dish. A total of 2-mg/mL collagenase II and 10-U/mL DNAase were added and rotated at 37°C for a period of time.

After standing for 2–3 minutes, decant the supernatant and remove the large lumps with the filter membrane. After centrifuging the cells, the supernatant portion was poured out and discarded. The cells were suspended again with erythrocyte lysis buffer, cultured at room temperature for 2~3 minutes, and then centrifuged at 120 *g* for 3 minutes at 4°C. Samples were lastly resuspended with PBS.



**Figure 9. Gene signatures found in the inflamed samples (UC) compared with noninflamed samples (SC) and healthy samples (HCs) in epithelial/stromal cells.** Enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) pathways of DEGs between (A) UC vs HCs and (B) UC vs SC. (C) Heatmaps showing the expression changes (UC vs HCs) of detected genes in IL-17 signaling and allograft rejection pathway in each Epithelial/stromal cell cluster. (D) Violin plots showing the expression distribution of *LCN2* and MHC class II genes in each epithelial and stromal cell cluster of UC, SC, and HCs.

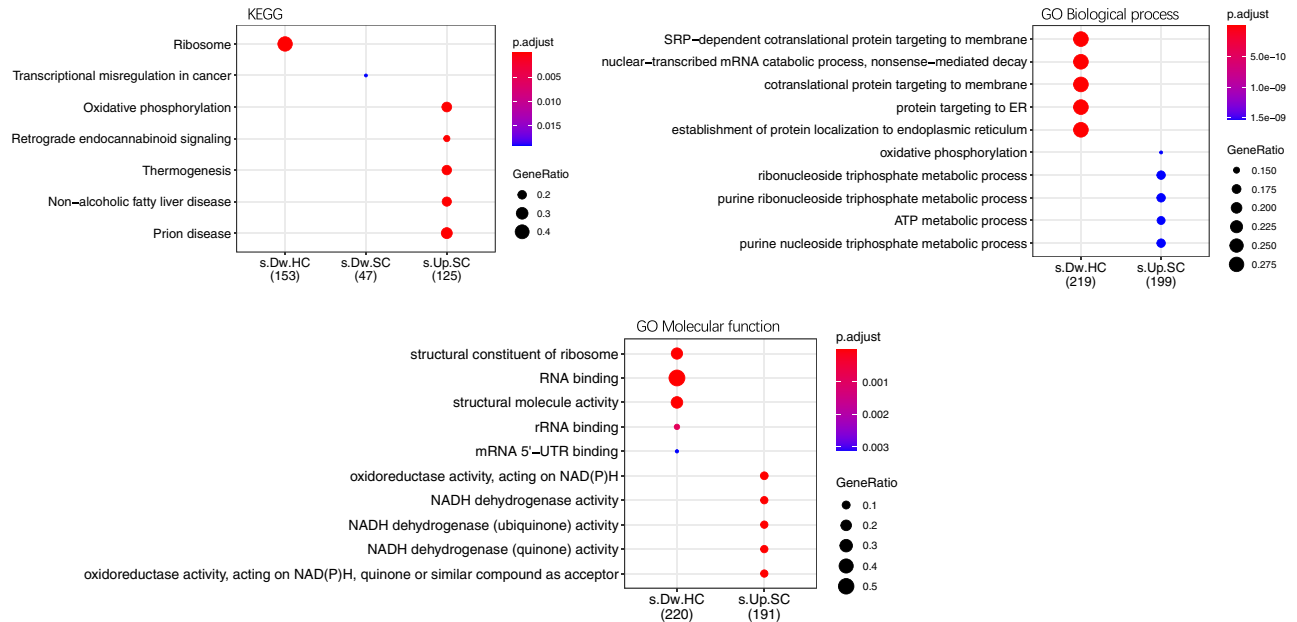
**Single-Cell RNA-Seq**

Cell capture and complementary DNA synthesis was performed using single-cell 3' Library and Gel Bead Kit V2 (10x Genomics, Pleasanton, CA; 120237) and Chromium Single Cell A Chip Kit (10x Genomics; 120236). The cell suspension (300–600 living cells/ $\mu$ L determined by Count Star) was loaded onto the Chromium Single Cell Controller (10x Genomics) to generate single-cell gel beads in the emulsion according to the manufacturer's protocol. In short, single cells were suspended in PBS containing 0.04% bovine

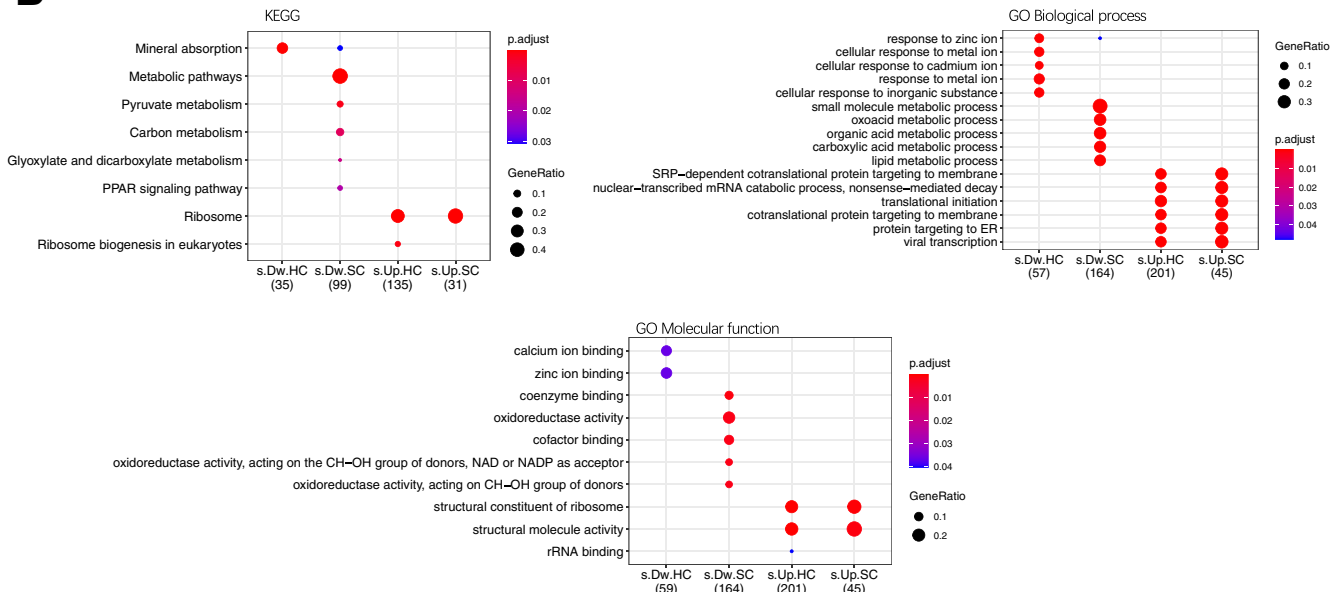
serum albumin. About 7000 cells were added to each channel, and the target cell recovery rate was estimated to be 3000 cells. Captured cells were lysed and the released RNAs were barcoded through reverse transcription in individual GEMs (Gel Bead-In Emulsion).

Using a S1000 Touch Thermal Cycler (Bio-Rad, Hercules, CA) to reverse-transcribe, the GEMs were programmed at 53°C for 45 minutes, followed by 85°C for 5 minutes, and were held at 4°C. The complementary DNA was obtained and amplified, and the quality was assessed using the

## A Enrichment of Chaoyang cohort specific DEGs in plasma cells



## B Enrichment of Chaoyang cohort specific DEGs in Epithelial cells



**Figure 10.** GO and KEGG enrichment terms of the DEGs in plasma or epithelial cells found specifically in this study but not in Smillie et al.<sup>14</sup> ATP, adenosine triphosphate; mRNA, messenger RNA; rRNA, ribosomal RNA; s.Dw.HC, specific downregulated when comparing UC with HCs; s.Dw.SC, specific downregulated when comparing UC with SC; s.Up.HC, specific upregulated when comparing UC with HCs; s.Up.SC: specific upregulated when comparing UC with SC.

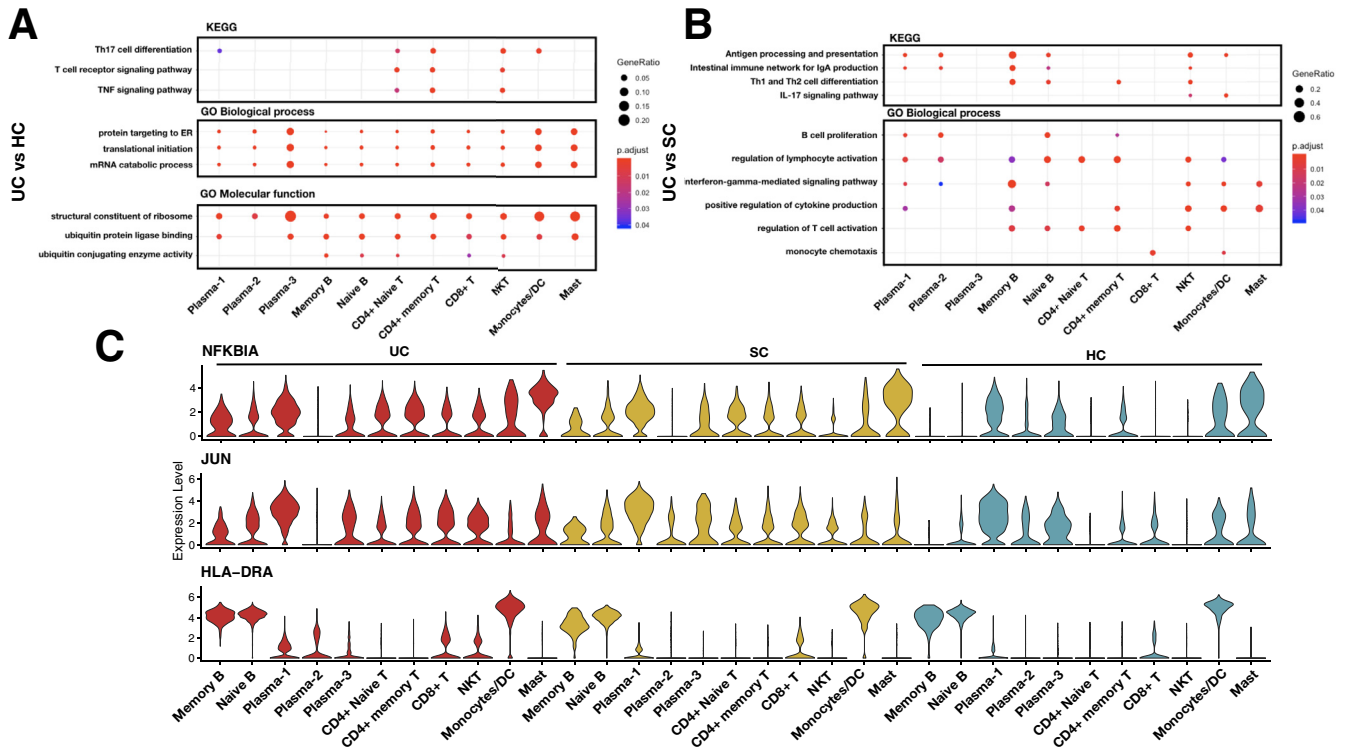
Agilent 4200 (Agilent, Santa Clara, CA; performed by CapitalBio, Beijing, China).

scRNA-seq libraries were prepared according to the manufacturer's introduction and the Single Cell 3' Library Gel Bead Kit V2 was used. Sequencing was performed on the Illumina NovaSeq 6000 sequencer (Illumina, San Diego, CA) with a sequencing depth of at least 69,000 reads per cell and

150 bp (PE150) paired-end reads (performed by CapitalBio).

In addition, 1 UC (patient 2) and 1 SC (patient 1) sample were excluded for further analyses because a high percentage of mitochondrial genes (>0.25) and the lower number of recovered cells (n = 688) were estimated in them.



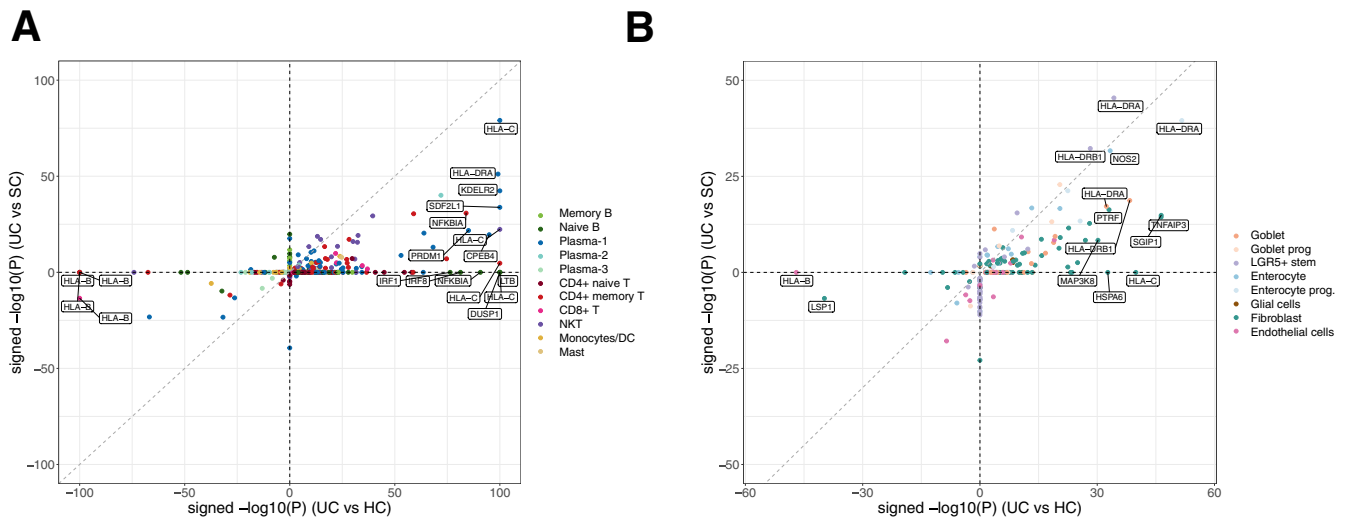


**Figure 11. Gene signatures found in the inflamed samples (UC) compared with noninflamed samples (SC) and healthy samples (HCs) in immune cells.** Enriched KEGG and GO pathways of DEGs between (A) UC vs HCs and (B) UC vs SC in immune cells. (C) Violin plots showing the expression distribution of NFKB1A, JUN, and HLA-DRA in each immune cell cluster of UC, SC, and HCs.

**Reads Processing and Quality Control**

Cell Ranger v3.0.2 (10x Genomics) was used to process scRNA-seq reads. To generate a digital gene expression (DGE) matrix for each sample, we mapped their reads to hg19 human transcriptome and recorded the number of UMIs (Unique Molecular Identifiers) for each gene in each cell. For each DGE matrices, the estimated number of cells and number of genes per cell were examined based on cell barcode and UMI barcodes.

Seurat v3.1 was applied to analyze DGE matrix from each sample.<sup>57,58</sup> To filter out low-quality cells and doublets, empirically filtering criteria were applied to each cell: number of estimated genes should be higher than 100 and lower than 6000 and the ratio of reads mapping to the mitochondria should be lower than 25%. Only genes detected in at least 5 cells were maintained for subsequent analyses.



**Figure 12. Comparison of signed log P values of DEGs from UC risk loci in (A) epithelial and stromal cell clusters and (B) immune cells.**



phenotypes, rather than by cell types or cell subsets. Principal component analysis (PCA) has then been performed on the integrated datasets, followed by shared nearest neighbor graph construction using PC1 to PC20 and  $k = 20$  nearest neighbors to identify unsupervised cell clusters. Finally, UMAP was used to visualize the cell clusters.

In order to keep the biological differences for downstream analyses, the above-mentioned batch correction was only used in the cell clustering and PCA-related steps. For the other analyses, we used standard LogNormalization methods. The original gene counts for each cell were normalized by total UMI counts and multiplied by 10,000 (TP10K), and then log transformed by  $\log(TP10K+1)$ .

In order to annotate cell identity to each cluster, we used a double-checking strategy for the inference<sup>59</sup> by comparing data-derived marker genes with public databases, and by directly visualizing the expression pattern of literature-derived marker genes. First, we used the automatic cell annotation tool, SingleR,<sup>60</sup> with 2 reference dataset, Human Primary Cell Atlas Data and Blueprint Encode data, to generate primary annotation. Next, data-derived marker genes were detected by applying differential expression tests between cells in one cluster and all other cells in the dataset. Upregulated genes from the cluster of interest were ranked by the Wilcoxon rank sum test and compared with their reported cell types in human large intestines in CellMarker databases.<sup>61</sup> Then, marker genes and cell surface markers reported in other human intestine or immune cell analyses were regarded as literature-derived markers, and we visualized their expression levels in each of the identified cell clusters to manually check the cell identities. Finally, we used the harmony algorithm<sup>62</sup> to integrate our data with well-annotated cells from Smillie et al<sup>14</sup> to further validate the identities.

To test significant changes of cell proportion, the Dirichlet-multinomial regression, which takes compositional dependencies into account, was used. For comparison and validation, Student's  $t$  test, Wilcoxon signed rank test, and paired  $t$  test (when applicable) were also applied.

### DEGs and Enrichment Analysis

For each annotated cell type, DEGs were estimated between UC and SC, UC and HCs, and SC and HCs, using FindMarkers function in Seurat v3.1 with the default Wilcoxon rank sum test, and with MAST<sup>63</sup> for validation. For each comparison, differential expression tests were performed only on genes that were detected in more than 10% of cells in any groups.  $P$  value adjustment was performed using Bonferroni correction based on the total number of genes in the tested dataset.

R package clusterProfiler v3.10.1<sup>64</sup> was used for Kyoto Encyclopedia of Genes and Genomes and Gene Ontology enrichment analysis for over-represented pathways and Gene Ontology terms on the DEGs found in UC vs HCs and UC vs SC in each cell identity.

### DEGs in GWAS Loci

UC GWAS data (2020-02-22-EFO\_0000729) were downloaded from NHGRI-EBI GWAS catalog,<sup>65</sup> which contains 281

associated loci with 605 reported risk genes. Risk genes that were found to be significantly differentially expressed in any of our comparison were extracted and classified as Asian and non-Asian risk genes based on whether they are initially estimated or have been replicated in Asian samples. In total, we identified 195 DEGs that are reported to be associated with ulcerative colitis, and only 8 of them have been reported in Asian GWASs, which are *CFB*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DRA*, *HLA-DRB1*, *IRF8*, *PTPRC*, and *SLC26A3*. Fisher's exact test was applied to test the overrepresentation of risk gene in differentially expressed genes in any cell types, using all genes expressed in >10% cells in estimated cell types as reference.

### Drug Target Analysis

Druggable gene list was obtained from Finan C's work.<sup>36</sup> The UC drugs information was collected from literatures.<sup>66,67</sup> Drug targets were obtained from DrugBank.<sup>68</sup> Fisher's exact test was applied for enrichment analysis.

### Reference

1. Ungaro R, Mehandru S, Allen PB, Peyrin-Biroulet L, Colombel JF. Ulcerative colitis. *Lancet* 2017; 389:1756–1770.
2. Liu TC, Stappenbeck TS. Genetics and pathogenesis of inflammatory bowel disease. *Annu Rev Pathol* 2016; 11:127–148.
3. Meir M, Burkard N, Ungewiß H, Diefenbacher M, Flemming S, Kannapin F, Germer CT, Schweinlin M, Metzger M, Waschke J, Schlegel N. Neurotrophic factor GDNF regulates intestinal barrier function in inflammatory bowel disease. *J Clin Invest* 2019;129:2824–2840.
4. Chang J, Leong RW, Wasinger VC, Ip M, Yang M, Phan TG. Impaired intestinal permeability contributes to ongoing bowel symptoms in patients with inflammatory bowel disease and mucosal healing. *Gastroenterology* 2017;153:723–731 e1.
5. Neurath MF. Targeting immune cell circuits and trafficking in inflammatory bowel disease. *Nat Immunol* 2019;20:970–979.
6. Brown EM, Kenny DJ, Xavier RJ. Gut microbiota regulation of T cells during inflammation and autoimmunity. *Annu Rev Immunol* 2019;37:599–624.
7. Lassen KG, Xavier RJ. Genetic control of autophagy underlies pathogenesis of inflammatory bowel disease. *Mucosal Immunol* 2017;10:589–597.
8. Hibi T, Kobayashi T, Hisamatsu T. Pathogenesis of inflammatory bowel disease. *Nihon Rinsho* 2017; 75:364–369.
9. Feuerstein JD, Moss AC, Farraye FA. Ulcerative colitis. *Mayo Clin Proc* 2019;94:1357–1373.
10. Ng SC, Shi HY, Hamidi N, Underwood FE, Tang W, Benchimol EI, Panaccione R, Ghosh S, Wu JCY, Chan FKL, Sung JY, Kaplan GG. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *Lancet* 2018;390:2769–2778.
11. Ng SC, Kaplan GG, Tang W, Banerjee R, Adigopula B, Underwood FE, Tanyingoh D, Wei SC, Lin WC, Lin HH,



- Li J, Bell S, Niewiadomski O, Kamm MA, Zeng Z, Chen M, Hu P, Ong D, Ooi CJ, Ling KL, Miao Y, Miao J, Janaka de Silva H, Niriella M, Aniwani S, Limsrivilai J, Pisespongsa P, Wu K, Yang H, Ng KK, Yu HH, Wang Y, Ouyang Q, Abdullah M, Simadibrata M, Gunawan J, Hilmi I, Lee Goh K, Cao Q, Sheng H, Ong-Go A, Chong VH, Ching JYL, Wu JCY, Chan FKL, Sung JY. Population density and risk of inflammatory bowel disease: a prospective population-based study in 13 countries or regions in Asia-Pacific. *Am J Gastroenterol* 2019;114:107–115.
12. Planell N, Lozano JJ, Mora-Buch R, Masamunt MC, Jimeno M, Ordás I, Esteller M, Ricart E, Piqué JM, Panés J, Salas A. Transcriptional analysis of the intestinal mucosa of patients with ulcerative colitis in remission reveals lasting epithelial cell alterations. *Gut* 2013;62:967–976.
13. McGovern DP, Kugathasan S, Cho JH. Genetics of inflammatory bowel diseases. *Gastroenterology* 2015;149:1163–1176.e2.
14. Smillie CS, Biton M, Ordoñas-Montanes J, Sullivan KM, Burgin G, Graham DB, Herbst RH, Rogel N, Slyper M, Waldman J, Sud M, Andrews E, Velonias G, Haber AL, Jagadeesh K, Vickovic S, Yao J, Stevens C, Dionne D, Nguyen LT, Villani AC, Hofree M, Creasey EA, Huang H, Rozenblatt-Rosen O, Garber JJ, Khalili H, Desch AN, Daly MJ, Ananthakrishnan AN, Shalek AK, Xavier RJ, Regev A. Intra- and inter-cellular rewiring of the human colon during ulcerative colitis. *Cell* 2019;178:714–730.e22.
15. Uniken Venema WT, Voskuil MD, Vila AV, van der Vries G, Jansen BH, Jabri B, Faber KN, Dijkstra G, Xavier RJ, Wijmenga C, Graham DB, Weersma RK, Festen EA. Single-cell RNA sequencing of blood and ileal T cells from patients with Crohn's disease reveals tissue-specific characteristics and drug targets. *Gastroenterology* 2019;156:812–815.e22.
16. Parikh K, Antanaviciute A, Fawcner-Corbett D, Jagielowicz M, Alicino A, Lagerholm C, Davis S, Kinchen J, Chen HH, Alham NK, Ashley N, Johnson E, Hublitz P, Bao L, Lukomska J, Andev RS, Björklund E, Kessler BM, Fischer R, Goldin R, Koohy H, Simmons A. Colonic epithelial cell diversity in health and inflammatory bowel disease. *Nature* 2019;567:49–55.
17. Huang B, Chen Z, Geng L, Wang J, Liang H, Cao Y, Chen H, Huang W, Su M, Wang H, Xu Y, Liu Y, Lu B, Xian H, Li H, Li H, Ren L, Xie J, Ye L, Wang H, Zhao J, Chen P, Zhang L, Zhao S, Zhang T, Xu B, Che D, Si W, Gu X, Zeng L, Wang Y, Li D, Zhan Y, Delfouneso D, Lew AM, Cui J, Tang WH, Zhang Y, Gong S, Bai F, Yang M, Zhang Y. Mucosal profiling of pediatric-onset colitis and IBD reveals common pathogenics and therapeutic pathways. *Cell* 2019;179:1160–1176.e24.
18. Gerbe F, Jay P. Intestinal tuft cells: epithelial sentinels linking luminal cues to the immune system. *Mucosal Immunol* 2016;9:1353–1359.
19. Morii E, Oboki K. MITF is necessary for generation of prostaglandin D2 in mouse mast cells. *J Biol Chem* 2004;279:48923–48929.
20. Haeggstrom JZ. Leukotriene biosynthetic enzymes as therapeutic targets. *J Clin Invest* 2018;128:2680–2690.
21. O'Leary CE, Schneider C, Locksley RM. Tuft cells—systemically dispersed sensory epithelia integrating immune and neural circuitry. *Annu Rev Immunol* 2019;37:47–72.
22. Howitt MR, Lavoie S, Michaud M, Blum AM, Tran SV, Weinstock JV, Gallini CA, Redding K, Margolskee RF, Osborne LC, Artis D, Garrett WS. Tuft cells, taste-chemosensory cells, orchestrate parasite type 2 immunity in the gut. *Science* 2016;351:1329–1333.
23. Goto N, Fukuda A, Yamaga Y, Yoshikawa T, Maruno T, Maekawa H, Inamoto S, Kawada K, Sakai Y, Miyoshi H, Taketo MM, Chiba T, Seno H. Lineage tracing and targeting of IL17RB(+) tuft cell-like human colorectal cancer stem cells. *Proc Natl Acad Sci U S A* 2019;116:12996–13005.
24. Yeşil A, Gönen C, Senateş E, Paker N, Gökden Y, Koçhan K, Erdem ED, Gündüz F. Relationship between neutrophil gelatinase-associated lipocalin (NGAL) levels and inflammatory bowel disease type and activity. *Dig Dis Sci* 2013;58:2587–2593.
25. Østvik AE, Granlund AV, Torp SH, Flatberg A, Beisvåg V, Waldum HL, Flo TH, Espevik T, Damås JK, Sandvik AK. Expression of Toll-like receptor-3 is enhanced in active inflammatory bowel disease and mediates the excessive release of lipocalin 2. *Clin Exp Immunol* 2013;173:502–511.
26. Dooley TP, Curto EV, Reddy SP, Davis RL, Lambert GW, Wilborn TW, Elson CO. Regulation of gene expression in inflammatory bowel disease and correlation with IBD drugs: screening by DNA microarrays. *Inflamm Bowel Dis* 2004;10:1–14.
27. Budzynska A, Gawron-Kiszka M, Nowakowska-Dulawa E, Spiewak J, Lesinska M, Kukla M, Waluga M, Hartleb M. Serum neutrophil gelatinase-associated lipocalin (NGAL) correlates with clinical and endoscopic activity in ulcerative colitis but fails to predict activity in Crohn's disease. *J Physiol Pharmacol* 2017;68:859–865.
28. Stallhofer J, Friedrich M, Konrad-Zerna A, Wetzke M, Lohse P, Glas J, Tillack-Schreiber C, Schnitzler F, Beigel F, Brand S. Lipocalin-2 is a disease activity marker in inflammatory bowel disease regulated by IL-17A, IL-22, and TNF-alpha and modulated by IL23R genotype status. *Inflamm Bowel Dis* 2015;21:2327–2340.
29. Bunte K, Beikler T. Th17 cells and the IL-23/IL-17 axis in the pathogenesis of periodontitis and immune-mediated inflammatory diseases. *Int J Mol Sci* 2019;20:3394.
30. Macpherson AJ, Yilmaz B, Limenitakis JP, Ganai-Vonarburg SC. IgA function in relation to the intestinal microbiota. *Annu Rev Immunol* 2018;36:359–381.
31. Donaldson GP, Ladinsky MS, Yu KB, Sanders JG, Yoo BB, Chou WC, Conner ME, Earl AM, Knight R, Bjorkman PJ, Mazmanian SK. Gut microbiota utilize immunoglobulin A for mucosal colonization. *Science* 2018;360:795–800.
32. Lin R, Chen H, Shu W, Sun M, Fang L, Shi Y, Pang Z, Wu W, Liu Z. Clinical significance of soluble immunoglobulins A and G and their coated bacteria in feces of

- patients with inflammatory bowel disease. *J Transl Med* 2018;16:359.
33. Zikherman J, Jenne C, Watson S, Doan K, Raschke W, Goodnow CC, Weiss A. CD45-Csk phosphatase-kinase titration uncouples basal and inducible T cell receptor signaling during thymic development. *Immunity* 2010; 32:342–354.
  34. Arkteg CB, Goll R, Gundersen MD, Anderssen E, Fenton C, Florholmen J. Mucosal gene transcription of ulcerative colitis in endoscopic remission. *Scand J Gastroenterol* 2020;55:139–147.
  35. Lu D, Liu L, Ji X, Gao Y, Chen X, Liu Y, Liu Y, Zhao X, Li Y, Li Y, Jin Y, Zhang Y, McNutt MA, Yin Y. The phosphatase DUSP2 controls the activity of the transcription activator STAT3 and regulates TH17 differentiation. *Nat Immunol* 2015;16:1263–1273.
  36. Finan C, Gaulton A, Kruger FA, Lumbers RT, Shah T, Engmann J, Galver L, Kelley R, Karlsson A, Santos R, Overington JP, Hingorani AD, Casas JP. The druggable genome and support for target identification and validation in drug development. *Sci Transl Med* 2017;9: eaag1166.
  37. Li H, Fan C, Lu H, Feng C, He P, Yang X, Xiang C, Zuo J, Tang W. Protective role of berberine on ulcerative colitis through modulating enteric glial cells-intestinal epithelial cells-immune cells interactions. *Acta Pharm Sin B* 2020; 10:447–461.
  38. Yu YB, Li YQ. Enteric glial cells and their role in the intestinal epithelial barrier. *World J Gastroenterol* 2014; 20:11273–11280.
  39. O'Connor TM, O'Connell J, O'Brien DI, Goode T, Bredin CP, Shanahan F. The role of substance P in inflammatory disease. *J Cell Physiol* 2004;201:167–180.
  40. Bezencon C, Furholz A, Raymond F, Mansourian R, Metairon S, Le Coutre J, Damak S. Murine intestinal cells expressing Trpm5 are mostly brush cells and express markers of neuronal and inflammatory cells. *J Comp Neurol* 2008;509:514–525.
  41. Middelhoff M, Westphalen CB, Hayakawa Y, Yan KS, Gershon MD, Wang TC, Quante M. Dclk1-expressing tuft cells: critical modulators of the intestinal niche? *Am J Physiol Gastrointest Liver Physiol* 2017;313:G285–G299.
  42. Cheng CW, Biton M, Haber AL, Gunduz N, Eng G, Gaynor LT, Tripathi S, Calibasi-Kocal G, Rickelt S, Butty VL, Moreno-Serrano M, Iqbal AM, Bauer-Rowe KE, Imada S, Ulutas MS, Mylonas C, Whary MT, Levine SS, Basbinar Y, Hynes RO, Mino-Kenudson M, Deshpande V, Boyer LA, Fox JG, Terranova C, Rai K, Piwnicka-Worms H, Mihaylova MM, Regev A, Yilmaz ÖH. Ketone body signaling mediates intestinal stem cell homeostasis and adaptation to diet. *Cell* 2019; 178:1115–1131.e15.
  43. Barker N. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nat Rev Mol Cell Biol* 2014;15:19–33.
  44. Hayakawa Y, Tsuboi M, Asfaha S, Kinoshita H, Niikura R, Konishi M, Hata M, Oya Y, Kim W, Middelhoff M, Hikiba Y, Higashijima N, Ihara S, Ushiku T, Fukayama M, Tailor Y, Hirata Y, Guha C, Yan KS, Koike K, Wang TC. BHLHA15-positive secretory precursor cells can give rise to tumors in intestine and colon in mice. *Gastroenterology* 2019;156:1066–1081.e16.
  45. Tetteh PW, Basak O, Farin HF, Wiebrands K, Kretschmar K, Begthel H, van den Born M, Korving J, de Sauvage F, van Es JH, van Oudenaarden A, Clevers H. Replacement of lost Lgr5-positive stem cells through plasticity of their enterocyte-lineage daughters. *Cell Stem Cell* 2016;18:203–213.
  46. Guo XY, Liu XJ, Hao JY. Gut microbiota in ulcerative colitis: insights on pathogenesis and treatment. *J Dig Dis* 2020;21:147–159.
  47. Boshagh MA, Foroutan P, Moloudi MR, Fakhari S, Malakouti P, Nikkhoo B, Jalili A. ELR positive CXCL chemokines are highly expressed in an animal model of ulcerative colitis. *J Inflamm Res* 2019;12:167–174.
  48. Rosen MJ, Karns R, Vallance JE, Bezold R, Waddell A, Collins MH, Haberman Y, Minar P, Baldassano RN, Hyams JS, Baker SS, Kellermayer R, Noe JD, Griffiths AM, Rosh JR, Crandall WV, Heyman MB, Mack DR, Kappelman MD, Markowitz J, Moulton DE, Leleiko NS, Walters TD, Kugathasan S, Wilson KT, Hogan SP, Denson LA. Mucosal expression of type 2 and type 17 immune response genes distinguishes ulcerative colitis from colon-only Crohn's disease in treatment-naïve pediatric patients. *Gastroenterology* 2017;152:1345–1357.e7.
  49. Rampal R, Wari N, Singh AK, Das U, Bopanna S, Gupta V, Nayak B, Velapandian T, Kedia S, Kumar D, Awasthi A, Ahuja V. Retinoic acid is elevated in the mucosa of patients with active ulcerative colitis and displays a proinflammatory role by augmenting IL-17 and IFN $\gamma$  production. *Inflamm Bowel Dis* 2020; 27:74–83.
  50. Castro-Dopico T, Dennison TW, Ferdinand JR, Mathews RJ, Fleming A, Clift D, Stewart BJ, Jing C, Strongili K, Labzin LI, Monk EJM, Saeb-Parsy K, Bryant CE, Clare S, Parkes M, Clatworthy MR. Anti-commensal IgG drives intestinal inflammation and type 17 immunity in ulcerative colitis. *Immunity* 2019; 50:1099–1114.e10.
  51. Zheng K, Jia J, Yan S, Shen H, Zhu P, Yu J. Paeoniflorin ameliorates ulcerative colitis by modulating the dendritic cell-mediated TH17/Treg balance. *Inflammopharmacology* 2020;28:1705–1716.
  52. Fujino S, Andoh A, Bamba S, Ogawa A, Hata K, Araki Y, Bamba T, Fujiyama Y. Increased expression of interleukin 17 in inflammatory bowel disease. *Gut* 2003; 52:65–70.
  53. Long Y, Zhao X, Xia C, Li X, Fan C, Liu C, Wang C. Upregulated IL-17A secretion and CCR6 co-expression in Treg subsets are related to the imbalance of Treg/Th17 cells in active UC patients. *Scand J Immunol* 2020; 91:e12842.
  54. Ma YH, Zhang J, Chen X, Xie YF, Pang YH, Liu XJ. Increased CD4<sup>+</sup>CD45RA<sup>+</sup>FoxP3<sup>low</sup> cells alter the balance between Treg and Th17 cells in colitis mice. *World J Gastroenterol* 2016;22:9356–9367.
  55. Ueno A, Jeffery L, Kobayashi T, Hibi T, Ghosh S, Jijon H. Th17 plasticity and its relevance to inflammatory bowel disease. *J Autoimmun* 2018;87:38–49.

56. Magro F, Gionchetti P, Eliakim R, Ardizzone S, Armuzzi A, Barreiro-de Acosta M, Burisch J, Gecse KB, Hart AL, Hindryckx P, Langner C, Limdi JK, Pellino G, Zagórowicz E, Raine T, Harbord M, Rieder F; European Crohn's and Colitis Organisation [ECCO]. Third European Evidence-based Consensus on Diagnosis and Management of Ulcerative Colitis. Part 1: Definitions, Diagnosis, Extra-intestinal Manifestations, Pregnancy, Cancer Surveillance, Surgery, and Ileo-anal Pouch Disorders. *J Crohns Colitis* 2017;11:649–670.
57. 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.
58. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* 2019;20:296.
59. Luecken MD, Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol Syst Biol* 2019;15: e8746.
60. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, Chak S, Naikawadi RP, Wolters PJ, Abate AR. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol* 2019;20:163–172.
61. Zhang X, Lan Y, Xu J, Quan F, Zhao E, Deng C, Luo T, Xu L, Liao G, Yan M, Ping Y, Li F, Shi A, Bai J, Zhao T, Li X, Xiao Y. CellMarker: a manually curated resource of cell markers in human and mouse. *Nucleic Acids Res* 2019;47:D721–D728.
62. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, Brenner M, Loh PR, Raychaudhuri S. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* 2019;16:1289–1296.
63. Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, Slichter CK, Miller HW, McElrath MJ, Prlic M, Linsley PS, Gottardo R. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol* 2015;16:278.
64. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012;16:284–287.
65. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, McMahon A, Morales J, Mountjoy E, Sollis E, Suveges D, Vrousseau O, Whetzel PL, Amode R, Guillen JA, Riat HS, Trevanion SJ, Hall P, Junkins H, Flicek P, Burdett T, Hindorf LA, Cunningham F, Parkinson H. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* 2019;47:D1005–D1012.
66. Vaughn BP, Moss AC. Novel treatment options for ulcerative colitis. *Clin Investig (Lond)* 2013;3:1057–1069.
67. Pastorelli L, Pizarro TT, Cominelli F, Vecchi M. Emerging drugs for the treatment of ulcerative colitis. *Expert Opin Emerg Drugs* 2009;14:505–521.
68. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N, Iynkkaran I, Liu Y, Maciejewski A, Gale N, Wilson A, Chin L, Cummings R, Le D, Pon A, Knox C, Wilson M. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res* 2018;46:D1074–D1082.

---

Received September 16, 2020. Accepted January 27, 2021.

#### Correspondence

Address correspondence to: Xinjuan Liu, PhD, Department of Gastroenterology, Beijing Chaoyang Hospital, Capital Medical University, Chaoyang District, Beijing, 100020, China. e-mail: liuxinjuan@mail.ccmu.edu.cn.

#### Acknowledgments

The authors thank the patients-participants at the Beijing Chaoyang hospital for contributing intestinal biopsy samples.

#### CRedit Authorship Contributions

Guang Li (Conceptualization: Equal; Formal analysis: Equal; Resources: Equal; Writing – original draft: Lead; Writing – review & editing: Equal)

Bowen Zhang (Formal analysis: Equal; Writing – original draft: Lead; Writing – review & editing: Equal)

Jianyu Hao (Conceptualization: Equal; Formal analysis: Equal; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Xiaojing Chu (Formal analysis: Equal; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Miriam Wiestler (Formal analysis: Supporting; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Markus Cornberg (Formal analysis: Supporting; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Chengjian Xu (Formal analysis: Supporting; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Xinjuan Liu (Conceptualization: Equal; Formal analysis: Equal; Methodology: Equal; Resources: Equal; Writing – original draft: Supporting; Writing – review & editing: Lead)

Yang Li (Conceptualization: Equal; Formal analysis: Equal; Methodology: Equal; Writing – original draft: Supporting; Writing – review & editing: Supporting)

#### Conflicts of interest

The authors disclose no conflicts.

#### Funding

Xinjuan Liu was supported by the National Natural Science Foundation of China (82070559), National Natural Science Foundation of Beijing (7192072), and the Project of Digestive Medical Coordinated Development Center of Beijing Municipal Administration of Hospitals (XXT11). Yang Li was supported by an ERC starting Grant (948207) and a Radboud University Medical Centre Hypatia Grant [2018].