



OPEN Identification of hub genes and immune infiltration in ulcerative colitis using bioinformatics

Weitao Hu¹, Taiyong Fang², Mingxuan Zhou³ & Xiaoqing Chen¹✉

Ulcerative colitis (UC) is a chronic inflammatory disease of the intestine, whose pathogenesis is not fully understood. Given that immune infiltration plays a key role in UC progression, our study aimed to assess the level of immune cells in UC intestinal mucosal tissues and identify potential immune-related genes. The GSE65114 UC dataset was downloaded from the Gene Expression Omnibus database. Differentially expressed genes (DEGs) between healthy and UC tissues were identified using the “limma” package in R, while their Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were determined with the clusterProfiler package. Protein–protein interaction network analysis and visualization were performed with STRING and Cytoscape. Immune cell infiltration was calculated with CIBERSORT. The relationship between hub genes and immune-infiltrated cells in UC was determined by Pearson correlation. A total of 206 DEGs were identified, of which 174 were upregulated and 32 downregulated. GO and KEGG functional classification indicated DEG enrichment in immune response pathways, including Toll-like receptor signaling, IL-17 signaling, and immune system process and chemokine signaling. 13 hub genes were identified. Infiltration matrix analysis of immune cells showed abundant plasma cells, memory B cells, resting CD4 memory T cells, $\gamma\delta$ T cells, M0 and M1 macrophages, and neutrophils in UC intestinal tissues. Correlation analysis revealed 13 hub genes associated with immune-infiltrated cells in UC. 13 hub genes associated with immune-infiltrated cells in UC were identified; they included *CXCL13*, *CXCL10*, *CXCL9*, *CXCL8*, *CCL19*, *CTLA4*, *CCR1*, *CD69*, *CD163*, *IL7R*, *PECAM1*, *TLR8* and *TLR2*. These genes could potentially serve as markers for the diagnosis and treatment of UC.

Abbreviations

| | |
|--------|--|
| UC | Ulcerative colitis |
| DEGs | Differentially expressed genes |
| GO | Gene Ontology |
| KEGG | Kyoto Encyclopedia of Genes and Genomes |
| GEO | Gene Expression Omnibus |
| STRING | Search Tool for the Retrieval of Interacting Genes |
| PPI | Protein-protein interaction |
| MCODE | Molecular Complex Detection |
| CXCL13 | Chemokine (C-X-C motif) ligand 13 |
| CXCL10 | Chemokine (C-X-C motif) ligand 10 |
| CXCL9 | Chemokine (C-X-C motif) ligand 9 |
| CXCL8 | Chemokine (C-X-C motif) ligand 8 |
| CCL19 | C-C motif chemokine ligand 19 |
| CTLA4 | Cytotoxic T-lymphocyte associated protein 4 |
| CCR1 | C-C motif chemokine receptor 1 |

¹Department of Rheumatology, The Second Affiliated Hospital of Fujian Medical University, 34 North Zhongshan Road, Licheng District, Quanzhou 362000, Fujian, People’s Republic of China. ²Department of Gastroenterology, The Second Affiliated Hospital of Fujian Medical University, Quanzhou 362000, Fujian, People’s Republic of China. ³Department of General Practice, The Second Affiliated Hospital of Fujian Medical University, Quanzhou 362000, Fujian, People’s Republic of China. ✉email: chenxiaoqing202203@163.com

| | |
|--------|---|
| IL7R | Interleukin-7 receptor subunit alpha |
| TLR2 | Toll-like receptor 2 |
| TLR8 | Toll-like receptor 8 |
| PECAM1 | Platelet and endothelial cell adhesion molecule 1 |
| IBD | Inflammatory bowel disease |
| NK | Natural killer |

Ulcerative colitis (UC) is a chronic inflammatory disease of the intestine, whose etiology remains poorly understood¹. The diagnosis of UC is based on a combination of nonspecific symptoms, endoscopic findings, and histological features, which makes it sometimes difficult to discriminate UC from other diseases^{2,3}. Although several drugs are available for the treatment of UC⁴, up to 15% of patients do not respond to drug therapy or have chronic colitis secondary to dysplasia, which requires surgery⁵. Therefore, there is an urgent need to better understand the pathogenesis of UC and identify more effective treatments.

Autoimmune mechanisms have long been hypothesized to be involved in the pathogenesis of UC⁴. Previous studies have suggested the presence of immune cells in the intestinal mucosa of UC patients⁶. Multiple environmental factors can interfere with the microbial ecosystem of the colon and determine how gut microbes interact with immune cells, thereby provoking an uncontrolled inflammatory response and exacerbating UC symptoms⁷. Infection or dysbiosis may disrupt the natural immune tolerance of genetically susceptible individuals, causing immune imbalance in the intestinal mucosa and further affecting the onset and progression of UC⁸. It has been shown that the development of new therapeutic approaches to modulate the gut microbiota through the administration of bacterial strains that produce probiotic immune metabolites and the addition of specific prebiotics and fecal microbiota transplants can treat UC patients requiring ileal pouch ano-anastomosis (IPAA)⁹. All these findings suggest a critical role of immune cells in the pathogenesis of UC. Therefore, understanding the pathogenesis of UC from the perspective of immune infiltration may hold the key for early treatment and prevention of UC-induced deterioration. In particular, it may provide a new approach for targeted immunotherapy of UC.

Bioinformatics can reveal the molecular mechanisms of disease through large-scale gene or protein expression profiling in diseased vs. healthy tissues. Expression data are especially useful to study dynamic regulation among multiple immune cells. CIBERSORT is a commonly used analytical tool in studies of tumor immunity, whereby information about cell subsets is derived from bulk gene expression data. The method, however, has not been applied widely to investigate non-tumor immunity or to analyze immune infiltration in UC. The present study aimed to identify potential UC biomarkers and assess the level of immune cells in intestinal mucosal tissues of UC patients using CIBERSORT. Furthermore, to identify immune-related genes suitable for the diagnosis and treatment of UC, the correlation between immune cells and hub genes was calculated.

Materials and methods

Microarray data. The Gene Expression Omnibus (GEO; www.ncbi.nlm.nih.gov/geo/)¹⁰ database was used to search gene microarray data of UC intestinal tissues. The GSE65114 dataset (Contributed by Balfe A, Lennon G, O'Connell R et al.: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65114>), containing 12 healthy (control group) and 16 UC (experimental group) colon tissue samples was selected. Probes were converted to gene symbols according to the GPL16686 [HuGene-2_0-st] Affymetrix Human Gene 2.0 ST Array [transcript (gene) version] platform.

Identification of differentially expressed genes (DEGs). The gene expression matrix of the GSE65114 dataset was analyzed with the “limma” package in R to obtain DEGs between UC and healthy samples. Briefly, $|\log_2 \text{fold change (FC)}| > 1$ and $P < 0.05$ were set as the selection criteria for DEGs, with $|\log_2 \text{FC}| < 0$ indicating downregulated genes and $|\log_2 \text{FC}| > 0$ indicating upregulated genes when UC vs healthy individuals.

Protein–protein interaction (PPI) network construction and module analysis. The PPI network was constructed using the Search Tool for the Retrieval of Interacting Genes (STRING; <http://string-db.org>) (version 11.5)¹¹ online database, and the parameters were set as follows: meaning of network edges: confidence level; minimum required interaction score: medium confidence (0.400). Identification of functional interactions between proteins may provide insights into the mechanisms underlying disease development. The PPI network was drawn using Cytoscape (version 3.9.1), an open-source bioinformatics platform for visualizing molecular interactions¹². The most significant module was identified with the Cytoscape plug-in Molecular Complex Detection (MCODE) (version 2.0), which is used to identify densely connected regions by clustering a given network based on its topology¹³. The criteria for selection were as follows: MCODE score > 5 , degree cut-off = 2, node score cut-off = 0.2, maximum depth = 100, and k-score = 2. Using the cytoHubba plugin, the overlap of the top 20 genes based on algorithms such as MCC, maximum neighborhood component (MNC), Closeness, Degree, and edge percolated component (EPC) algorithms were identified as hub genes of UC.

Pathway enrichment analyses of DEGs. The clusterProfiler package in R was used to identify the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG)^{14–16} pathways characterizing DEGs, as well as to explore their potential biological processes, cellular components, molecular functions, and important signaling pathways. $P < 0.05$ and false detection rate < 0.1 were considered statistically significant.

Evaluation of subtype distribution among immune-infiltrated cells. CIBERSORT was shown to transform a normalized gene expression matrix into the composition of 22 immune cell types based on a decon-

volution algorithm¹⁷. Here, CIBERSORT was employed to calculate the composition of immune cells in UC and healthy samples. The algorithm employed the LM22 signature and 1000 permutations. Given $P < 0.05$, 16 UC and 12 healthy samples were selected for further analysis.

Correlation and differential analysis of immune-infiltrated cells. To assess the correlation among various immune cells, the Pearson correlation coefficient was calculated from sample data screened by CIBERSORT with $P < 0.05$. The rank-sum test was used to compare the UC and control groups.

Correlation between hub genes and immune-infiltrated cells in UC. The gene expression profile and the immune infiltrating cell profile of the CIBERSORT-analyzed GSE65114 dataset were subjected to Pearson correlation matrix analysis. Using Pearson correlation coefficient (r) > 0.6 and $P < 0.05$, the correlation between hub genes and immune-infiltrated cells was determined.

Results

Identification of DEGs in UC. The GSE65114 dataset containing gene expression profiles of 12 healthy and 16 UC active colonic mucosa tissue samples was retrieved from the GEO database. After standardization of microarray results, 206 DEGs were identified; they included 174 upregulated and 32 downregulated genes (Fig. 1).

PPI network construction, module analysis, and hub gene identification. PPI analysis of DEGs was based on the STRING database and results were visualized using Cytoscape (Fig. 2A). The MCODE plugin identified the most densely connected regions (15 nodes and 87 edges) in the PPI network (Fig. 2B). The complete Cytoscape's SIF file about the PPI network was in the Supplementary file. With the 5 algorithms of the Cytoscape plugin cytoHubba, we calculated the top 20 genes (Table 1). Subsequently, the top 13 intersecting genes analyzed based on these 5 algorithms were selected as the hub genes of UC, which included *CCL19*, *CCR1*, *CD163*, *CD69*, *CTLA4*, *CXCL10*, *CXCL13*, *CXCL8*, *CXCL9*, *IL7R*, *PECAM1*, *TLR2* and *TLR8* (Fig. 2C). Notably, in our study, all these hub genes were upregulated in patients with UC.

KEGG and GO enrichment in DEGs. To predict the biological function of DEGs, we performed functional enrichment analysis. GO analysis revealed that the DEGs were enriched mainly in extracellular region, regulation of response to stimulus, inflammatory response, immune system process and defense response (Fig. 3). KEGG pathway analysis indicated that upregulated genes were significantly enriched in viral protein interaction with cytokine and cytokine receptor, cytokine-cytokine receptor interaction, complement and coagulation cascades, chemokine signaling pathway, Toll-like receptor signaling pathway, and IL-17 signaling pathway (Fig. 4A); whereas downregulated genes were enriched mainly in proximal tubule bicarbonate reclamation and PPAR signaling pathway (Fig. 4B).

Distribution of immune-infiltrated cells. The microarray was screened using the CIBERSORT inverse convolution method with $P < 0.05$, resulting in 12 healthy intestinal tissues at the top of the heat map and 16 UC intestinal tissue groups at the bottom. Plasma cells, memory B cells, resting CD4 memory T cells, $\gamma\delta$ T cells, M0 macrophages, M1 macrophages, and neutrophils were all more abundant in colonic mucosal tissues of patients with active UC than in healthy controls (Fig. 5A). Figure 5B details the distribution of 22 immune cells in each sample.

Correlation and differential analysis of immune-infiltrated cells. A positive correlation was detected between naïve B cells and CD4 T cells ($r = 0.75$, $p < 0.05$), memory B cells and M0 macrophages ($r = 0.76$, $p < 0.05$), activated CD4 memory T cells and resting dendritic cells ($r = 0.98$, $p < 0.01$), follicular helper T cells and regulatory T cells (Tregs) ($r = 0.83$, $p < 0.05$), as well as M2 macrophages and activated dendritic cells ($r = 0.83$, $p < 0.05$). Instead, a negative correlation was detected between naïve B cells and plasma cells ($r = -0.78$, $p < 0.05$), memory B cells and monocytes ($r = -0.77$, $p < 0.05$), plasma cells and CD8 T cells ($r = -0.87$, $p < 0.05$), M0 and M2 macrophages ($r = -0.79$, $p < 0.05$), as well as activated dendritic cells and resting mast cells ($r = -0.86$, $p < 0.05$) (Fig. 6A).

Differences in immune infiltrating cells between the intestinal tissues of healthy and UC patients were visualized by a bar plot, with statistically significant differences at $P < 0.05$. Plasma cells, memory B cells, resting CD4 memory T cells, $\gamma\delta$ T cells, M0 macrophages, M1 macrophages, and neutrophils were all differentially elevated in the intestinal tissues of UC patients (Fig. 6B).

Correlation between hub genes and immune-infiltrated cells in UC. The relationship between hub genes and immune-infiltrated cells in UC, which differed between UC and control samples, was evaluated by Pearson correlation (Fig. 7). M1 macrophages displayed a positive correlation with *CXCL13* ($r = 0.69$), *CXCL10* ($r = 0.85$), *CXCL9* ($r = 0.95$), *CXCL8* ($r = 0.74$), *CCL19* ($r = 0.76$), *CCR1* ($r = 0.67$), *CD163* ($r = 0.72$), *CD69* ($r = 0.67$), *CTLA4* ($r = 0.80$), *IL7R* ($r = 0.80$), *PECAM1* ($r = 0.75$), *TLR8* ($r = 0.78$) and *TLR2* ($r = 0.80$). Memory B cells exhibited a positive correlation with *CTLA4* ($r = 0.78$). Plasma cells showed a positive correlation with *PECAM1* ($r = 0.86$). M0 macrophages exhibited a positive correlation with *CXCL8* ($r = 0.74$) and *CCR1* ($r = 0.67$). Therefore, hub genes were strongly correlated with immune-infiltrated cells in UC.

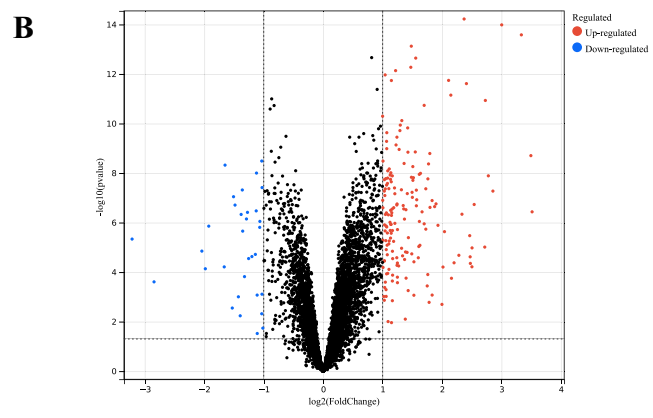
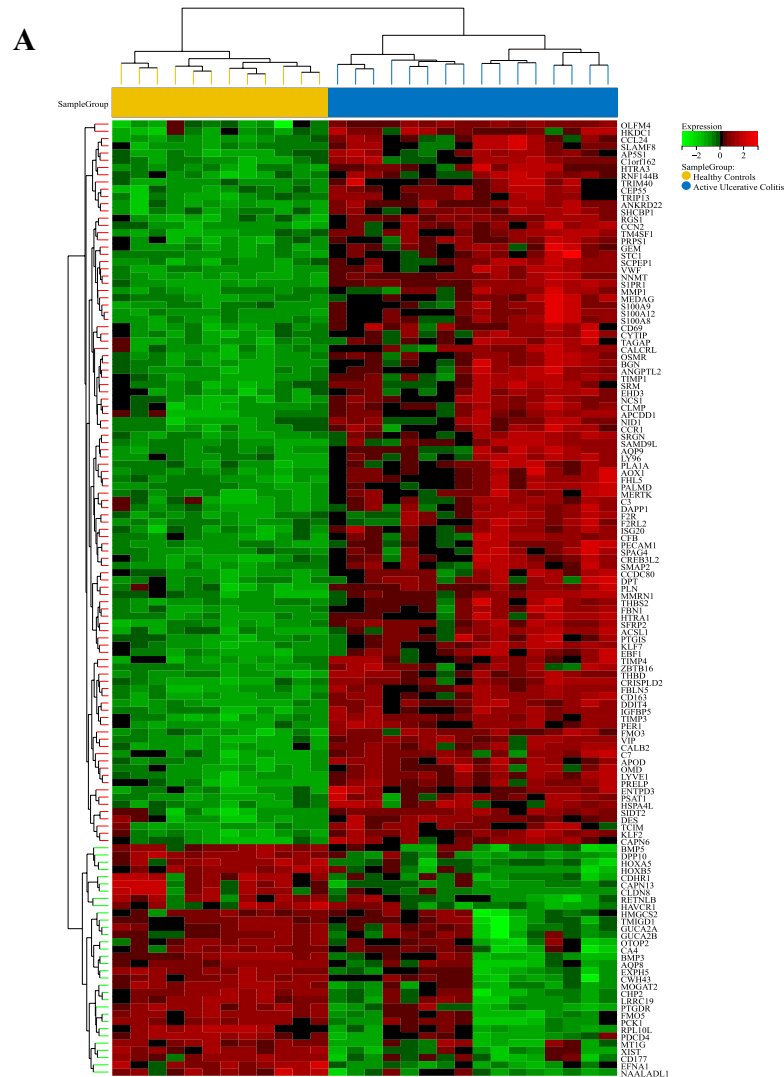


Figure 1. Differential gene expression in colonic mucosa tissue of UC patients. (A) Heat map of DEGs. Green corresponds to lower gene expression and red to higher gene expression. (B) Volcano plot of differentially expressed genes. $|\text{Log}_2\text{FC}| \geq 1$ and $P < 0.05$ were chosen as filtering conditions. Blue dots represent significantly downregulated genes, red dots represent significantly upregulated genes, and black dots denote genes not showing any significant difference.

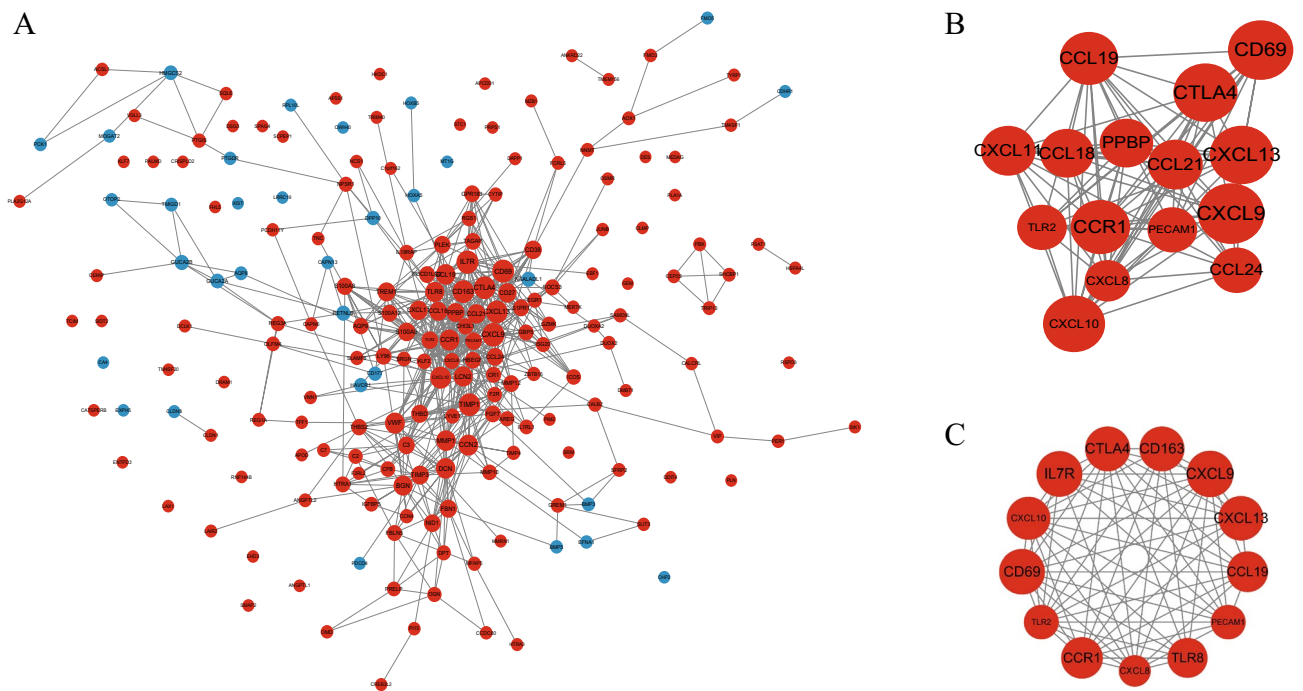


Figure 2. PPI network construction and module analysis. **(A)** PPI network of DEGs. The network consists of 206 nodes (39 isolates) and 603 edges, and network Diameter is 12, network Sparseness is 0.049, and Avg clustering coefficient is 0.424. Upregulated genes are marked in red and downregulated genes are marked in blue. The size of nodes represents the difference in expression; the larger is the size, the more significant is the *P* value. **(B)** Cytoscape-based identification of the densest connected regions (15 nodes and 87 edges) in the PPI network. **(C)** Hub genes identified in the densest connected regions. A darker red hue indicates a higher score.

| MCC | MNC | Degree | EPC | Closeness | Overlap |
|--------|--------|--------|--------|-----------|---------------|
| CXCL8 | DCN | DCN | DCN | CXCL8 | |
| CTLA4 | CXCL8 | CXCL8 | CXCL8 | CTLA4 | |
| IL7R | CTLA4 | CTLA4 | CTLA4 | MMP1 | |
| CCL21 | MMP1 | MMP1 | MMP1 | IL7R | |
| CXCL13 | IL7R | IL7R | IL7R | CXCL13 | CCL19 |
| CCR1 | CXCL13 | CXCL13 | CXCL13 | CCL21 | CCR1 |
| TLR8 | TLR8 | TLR8 | TLR8 | TLR8 | CD163 |
| CXCL9 | CCR1 | CCR1 | CCR1 | CCR1 | CD69 |
| CCL19 | VWF | VWF | VWF | VWF | CTLA4 |
| CCL24 | CXCL9 | CXCL9 | CXCL9 | CXCL9 | CXCL10 |
| TLR2 | CCN2 | CCN2 | CCN2 | CCL19 | CXCL13 |
| CCL18 | CCL19 | CCL19 | CCL19 | TLR2 | CXCL8 |
| CD163 | BGN | BGN | BGN | CD163 | CXCL9 |
| CD27 | TLR2 | TLR2 | TLR2 | TIMP1 | IL7R |
| TREM1 | CD163 | CD163 | CD163 | TREM1 | PECAM1 |
| CXCL10 | TIMP1 | TIMP1 | TIMP1 | CXCL10 | TLR2 |
| PECAM1 | CXCL10 | CXCL10 | CXCL10 | PECAM1 | TLR8 |
| PPBP | PECAM1 | PECAM1 | PECAM1 | PPBP | |
| CD69 | LCN2 | LCN2 | CD69 | CD69 | |
| CXCL11 | CD69 | CD69 | CXCL11 | CXCL11 | |

Table 1. The top 20 hub genes rank in cytoHubba. Significant values are in bold.

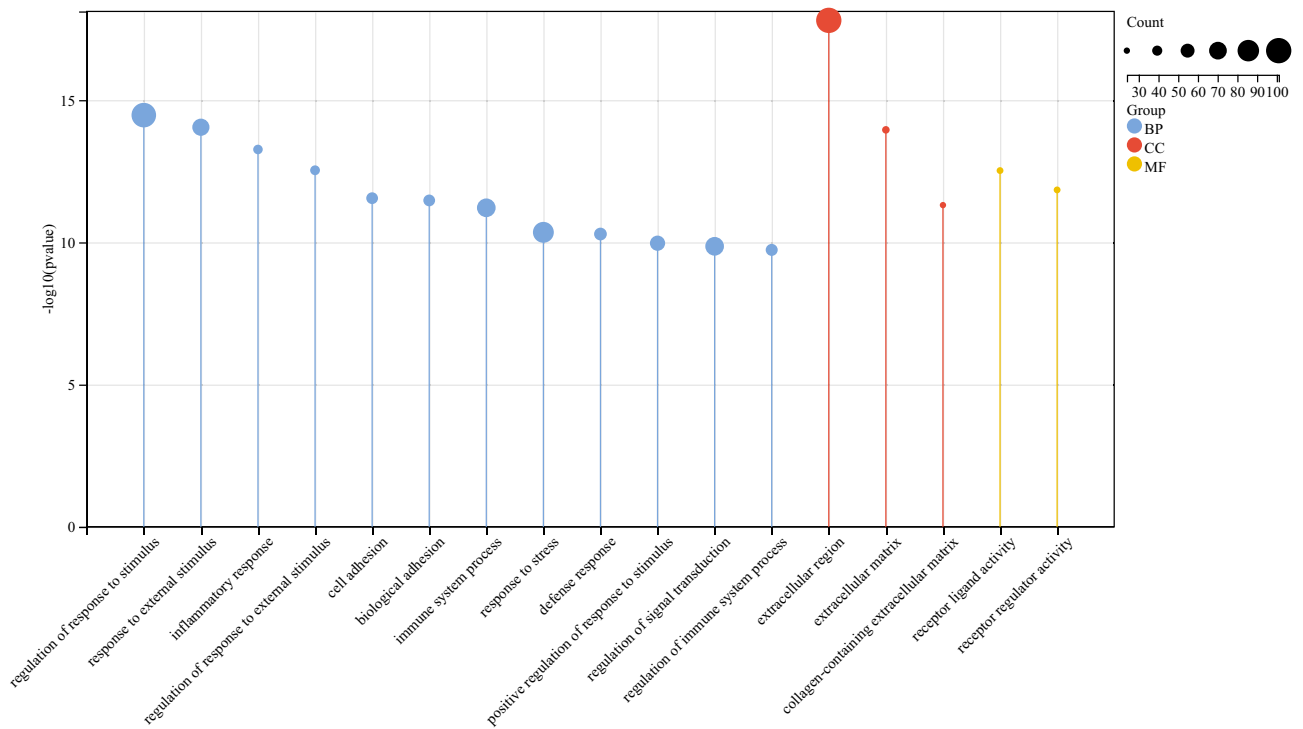


Figure 3. GO enrichment analysis of DEGs related to UC. A darker color and a larger bubble denote a more significant difference.

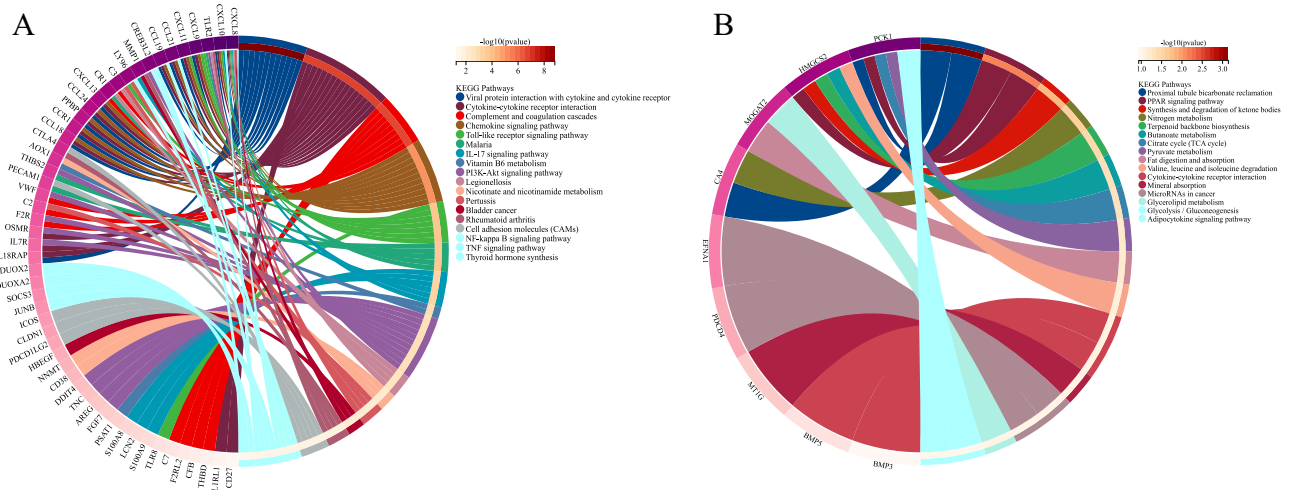


Figure 4. KEGG enrichment analysis of DEGs related to UC. (A) Upregulated DEGs. (B) Downregulated DEGs. The genes are linked to their assigned pathway terms via colored ribbons and are ordered according to the observed log₁₀ P-value, which is displayed in descending intensity of red squares next to the selected genes.

Discussion

UC originates from a disruption in the balance between the host’s mucosal immunity and intestinal bacterial flora, resulting in an abnormal immune response to commensal non-pathogenic bacteria⁶. In the present study, gene expression data were obtained from the GEO database, which identified 174 upregulated and 32 down-regulated DEGs in UC tissues. GO and KEGG functional classification indicated that DEGs were enriched mainly in immune response pathways, such as Toll-like receptor signaling, IL-17 signaling, and immune system process and chemokine signaling. Toll-like receptor signaling plays an important role in the pathogenesis of inflammatory bowel disease (IBD)¹⁸; whereas IL-17 signaling is involved in the development of colonic tissue damage and inflammation during UC¹⁹. CD4 T cells and natural killer (NK) T cells can promote the release of Th2-associated cytokines and Th17-associated pro-inflammatory cytokines, aggravating the UC intestinal

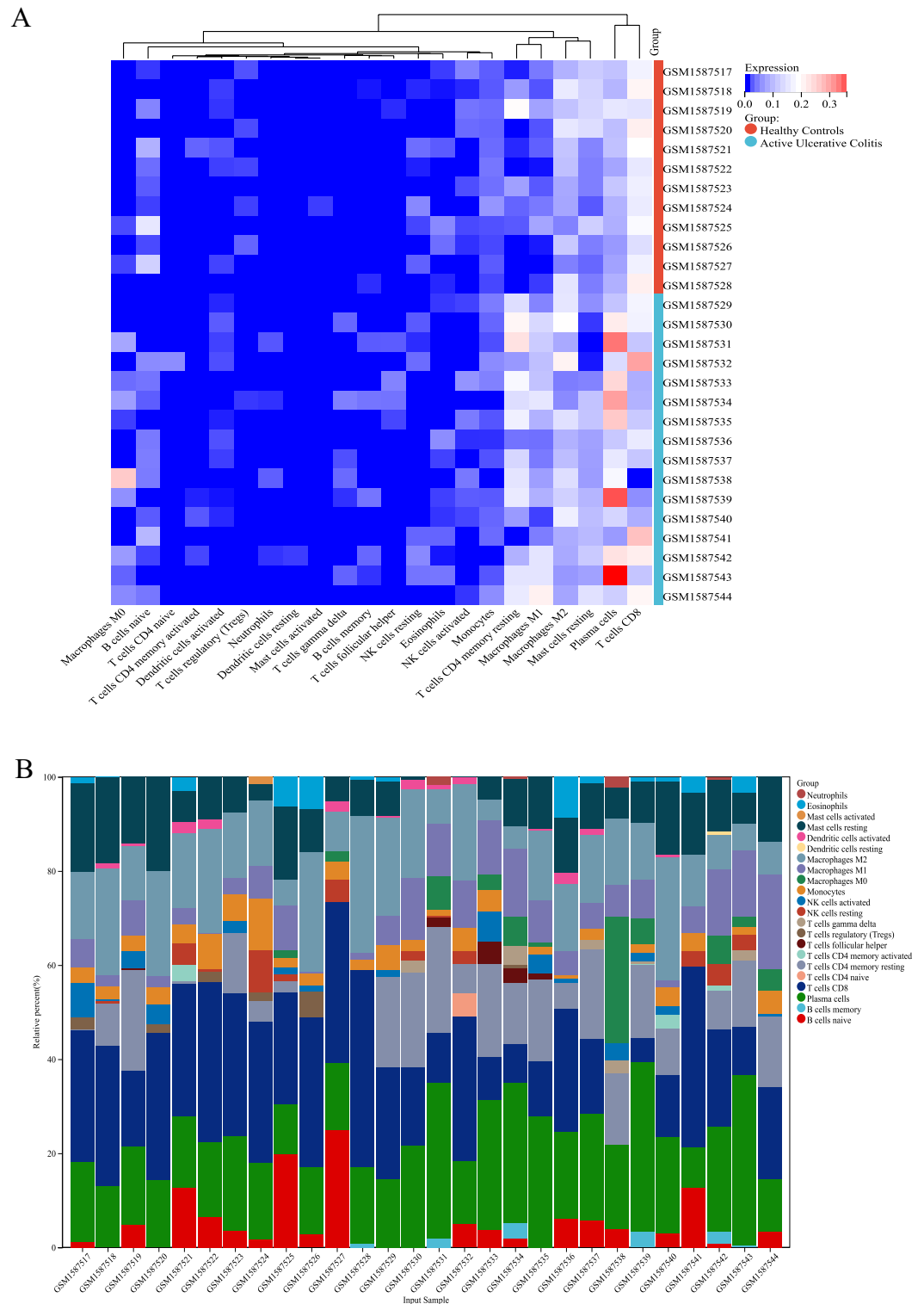


Figure 5. Infiltration of immune-associated cells in healthy and UC samples. **(A)** Immune cell content in each sample. **(B)** Relative percentage of 22 subpopulations of immune cells in 28 samples from the GSE65114 dataset.

inflammatory response^{20,21}. Growing evidence points to the role of chemokine signaling in UC^{22–24}. Overall, results of functional enrichment analysis suggested that immune factors played a critical role in UC pathogenesis.

UC occurs following dysregulation of the innate and adaptive immune responses to the intestinal microbiota in genetically susceptible hosts, with many different immune cell types involved^{4,25,26}. Plasma cells, memory B cells, resting CD4 memory T cells, $\gamma\delta$ T cells, M0 macrophages, M1 macrophages, and neutrophils were found to be more abundant in UC than healthy intestinal tissues. Plasma cells, B cells, and T helper cells are hallmarks

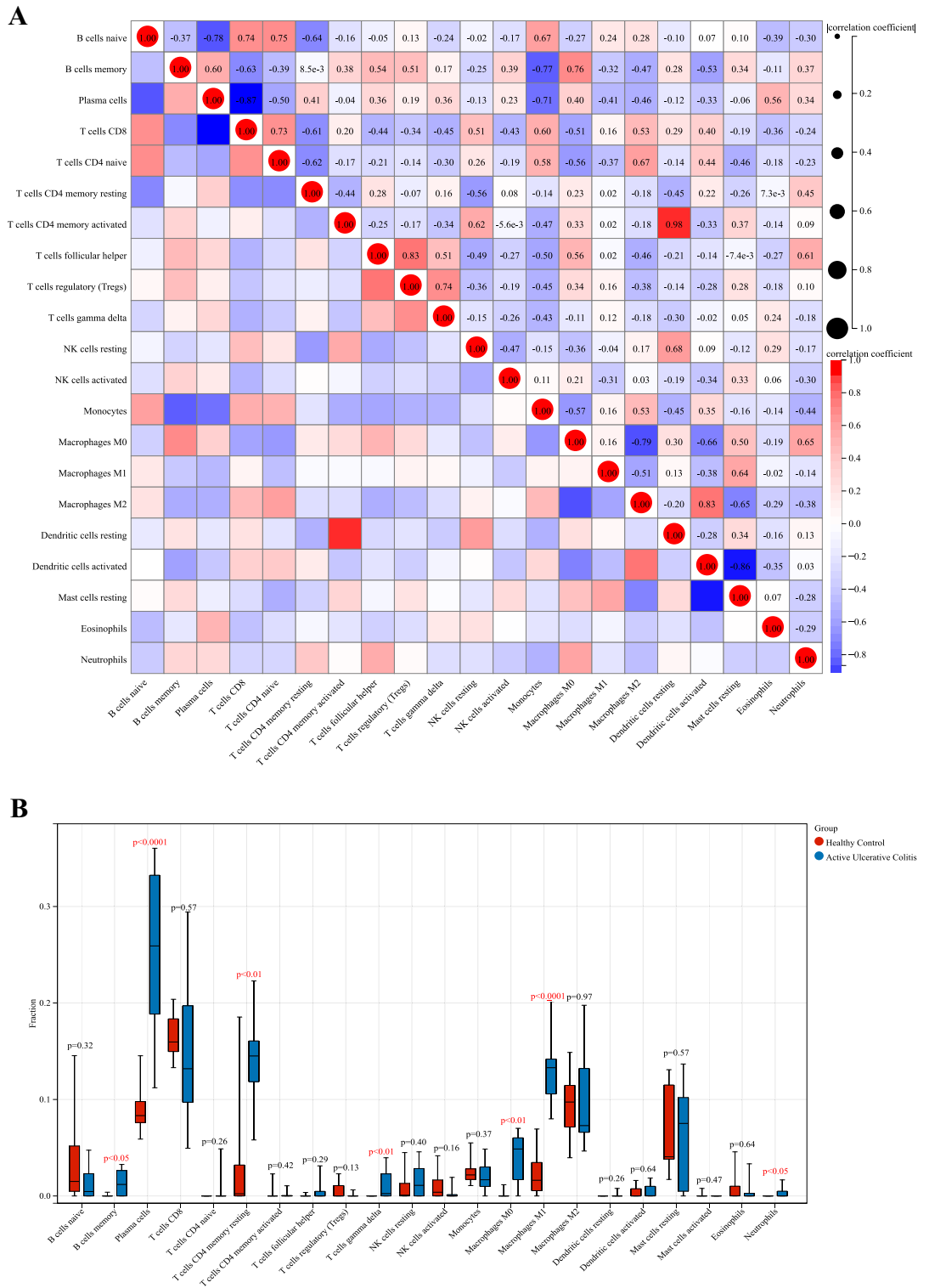


Figure 6. Correlation analysis and bar plot of differences among immune cells in the UC group. **(A)** Correlation analysis. Red indicates a positive correlation and blue indicates a negative correlation; the higher is the absolute value, the stronger is the correlation between immune cells. **(B)** Bar plot showing the proportion of each immune cell type between healthy and UC samples; red corresponds to healthy samples and blue to UC samples, $P < 0.05$.

of adaptive immune disorders associated with the onset and progression of UC intestinal inflammation^{27–29}. Wang et al. suggested that antigens activated memory B cells, prompting them to differentiate toward plasma

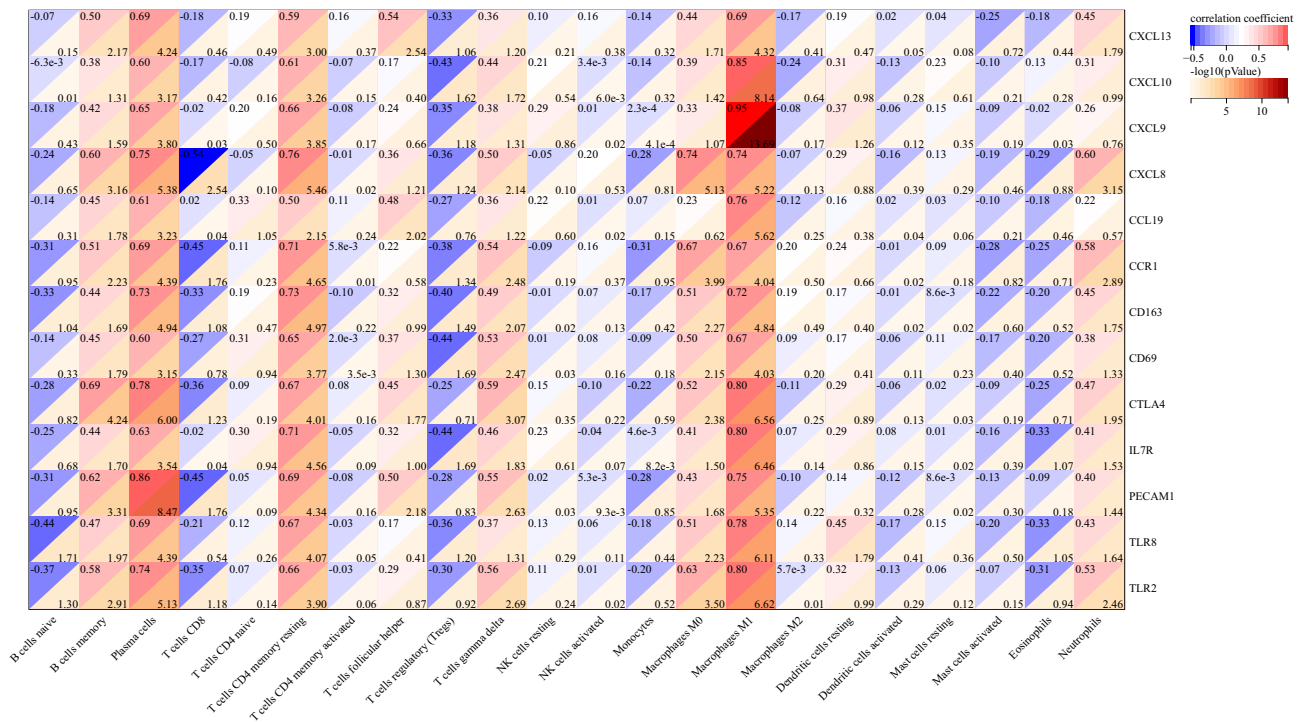


Figure 7. Correlation between hub genes and immune-infiltrated cells in UC. The darker is the red hue, the smaller is the *P* value.

cells and produce antigen-specific IgG that initiated UC pathogenesis³⁰. Hanai et al. suggested that leukocytapheresis acted therapeutically by reducing the ratio of CD4 + CD45RO + CD62L– effector memory T cells to CD4 + CD45RO + CD62L + central memory T cells in UC patients³¹. Earlier studies suggested that, in the presence of T cells, numerous types of immune cells might be triggered to produce more chemokines, followed by neutrophils infiltration into the colonic mucosa, thus directly and/or indirectly exacerbating the severity of UC-like chronic colitis³². $\gamma\delta$ T cells expressing the V δ 2 chain produce IL-17 in the intestine of patients with long-standing IBD and are involved in the chronic inflammatory process³³. Recent studies have shown that the global immune cell landscape of UC tissues is characterized by an increase in M0 macrophages and neutrophils³⁴. Macrophages are key effector cells of the innate immune system and are crucial for intestinal mucosal stability^{35,36}. In addition, they may serve as antigen-presenting cells and play a critical role in the initiation of the immune response³⁷. Zhuang et al. highlighted the abnormalities of M1/M2 macrophage polarization at the onset and during development of UC³⁸. Activation, migration, and degranulation of neutrophils are important mechanisms of intestinal injury in IBD³⁹. Here, the identified UC-associated immune-infiltrated cells were involved in the progression of UC, as indicated by correlation between such cells and the colonic mucosal tissue of patients with UC. Nevertheless, the intricate network of interactions and regulation among immune cells means that more research is required to determine their exact role in UC.

According to the PPI network of DEGs, 13 out of 446 genes displayed an elevated degree of interaction and were upregulated in UC patients: *CXCL13*, *CXCL10*, *CXCL9*, *CXCL8*, *CCL19*, *CTLA4*, *CD69*, *CD163*, *CCR1*, *PECAM1*, *IL7R*, *TLR8* and *TLR2*. Chemokines can significantly increase chronic inflammation and intestinal tissue destruction in IBD through their ability to induce chemotaxis and leukocyte activation⁴⁰. *CXCL8* (*IL-8*) is an ELR + chemokine secreted by neutrophils, macrophages, and intestinal epithelial cells^{41–43}. *CXCL8* has a strong chemotactic effect on neutrophils and activates their metabolism and degranulation⁴⁴. In addition to being a feature of UC pathophysiology, neutrophils infiltration in the intestinal mucosa is also a functional indicator of adaptive immunity⁴⁵. Both *CXCL9* and *CXCL10* belong to the family of ELR- chemokines, and target *CXCR3* as their receptor⁴⁶. *CXCL9* is highly expressed in the intestinal mucosa of mice with experimental colitis and in UC patients (especially in lymphocytes, macrophages, and epithelial cells)²². *CXCL10* is a potent chemokine that is primarily secreted by monocytes and macrophages (including M1 macrophages) in IBD^{47,48}. *CXCL9* and *CXCL10* recruit mostly Th1 cells, monocytes, and NK cells⁴⁹. *CXCL13* belongs to the ELR-chemokine family and targets *CXCR5* as its receptor, and its mRNA expression levels are elevated in intestinal tissues of UC models⁵⁰. In contrast to ELR + CXC chemokines, ELR- CXC chemokines lack chemotactic activity on neutrophils; instead, they are highly responsive to memory T cells and NK cells⁵⁰. A humanized anti-*CCL21* monoclonal antibody has been identified as a potent marker for the diagnosis of active IBD and as a possible therapeutic agent for the prevention of IBD recurrence⁵¹. *CCL19* is involved in the progression of UC^{52,53}. Given that *CCL19* induces the activation of MEK1-ERK1/2 and PI3K-AKT cascades in M1 macrophages⁵⁴, *CCL19* is believed to exacerbate the progression of UC by inducing chemotaxis in M1 macrophages. *CCR1* is expressed on neutrophils, contributing significantly to tissue damage and mucosal dysfunction in UC^{53,55}. In summary, the family of chemokines/

chemokine-receptors was found here to correlate positively with M0/M1 macrophages and neutrophils. This indicates that immune-infiltrated cells and UC hub genes collectively influence UC through immune factors.

CTLA4 serves as a negative regulator of the immune system, and is highly expressed on Tregs and activated T cells^{56,57}. Currently, the level of *CTLA4* expression in UC remains unclear. Wang et al.⁵⁸ and Magnusson et al.⁵⁹ suggested that *CTLA4* expression in T cells was lower in UC patients than in healthy individuals. However, other studies have shown that *CTLA4* mRNA expression was significantly higher in colonic mucosal tissue from patients with active UC compared with controls devoid of inflammation⁶⁰. Interestingly, in our study, *CTLA4* was upregulated in patients with UC and correlated positively with Tregs. Although Abatacept™ (*CTLA4*-Ig) has been effective against psoriasis⁶¹ and rheumatoid arthritis⁶² through its inhibitory effect on T cell activation, there is no evidence from a phase III clinical trial⁶³ of its beneficial effect on UC. *CTLA4* may be upregulated in patients with UC via a feedback mechanism, but may fail to exert negative immune regulatory effects due to defects in its downstream pathways or reduced activation. Furthermore, blocking *CTLA4* or a Treg-specific reduction of *CTLA4* expression lead to increased numbers of plasma cells and memory B cells after vaccination⁶⁴. This evidence points to the involvement of *CTLA4* in regulating the above immune cells, but whether it interacts and participates in the immune regulatory mechanisms of UC remains to be determined.

CD69 is one of the first surface antigens expressed by T lymphocytes after activation, and its expression can act as a co-stimulatory signal to promote further activation and proliferation of T cells⁶⁵. *CD69* is expressed on different leukocytes, including newly-activated lymphocytes, certain subtypes of memory T cells, infiltrating lymphocytes isolated from patients with chronic inflammatory disorders, Tregs, and NK cells^{65–67}. The immunoregulatory function of *CD69* involves controlling the differentiation balance of Th/Treg cells and enhancing the suppressive activity of Tregs⁶⁸. However, the exact mechanism of the interaction between *CD69* and UC requires further study. Soluble *CD163* is a specific macrophage activation marker that is reduced by anti-TNF- α antibody treatment in active inflammatory bowel disease⁶⁹. Defining the landscape of mononuclear phagocytes in mesenteric lymph nodes provides evidence for the expansion of CD163 + Mono/M Φ -like cells in UC, highlighting the distinction between UC and CD⁷⁰.

PECAM1 (*CD31*) is expressed on the surface of endothelial cells, platelets, monocytes, neutrophils, T cell subsets, B cells, and dendritic cells, and has been reported also in plasma cells^{71–73}. The interaction between endothelial cells and leukocytes is a key step in the inflammatory response, and *PECAM1* enables leukocytes to enter the site of inflammation and cause tissue damage⁷⁴. Thus, *PECAM1* has an important role in inflammatory microcirculation injury. In the present study, *PECAM1* correlated positively with plasma cells, which suggested that plasma cells probably contributed to inflammation and tissue damage in UC through *PECAM1* expression. A further possibility is that *PECAM1* may promote plasma cells infiltration into inflammatory sites of the intestine in UC.

Interleukin-7 (IL-7) is a cytokine produced mainly by epithelial and stromal cells that regulates T lymphocyte homeostasis⁷⁵. Almost all conventional mature T lymphocytes express high levels of IL-7 receptor (*IL-7R*), with naturally occurring Tregs being a special exception⁷⁶. In healthy colon biopsies, intestinal epithelial cells produce IL-7 and mucosal T lymphocytes express *IL7R*⁷⁷. A study has shown that genetic locus variants in the *IL7R* gene are associated with UC susceptibility⁷⁸. In addition, elevated expression of IL-7 signaling pathway genes in blood CD8 + T cells at diagnosis was significantly associated with the course of IBD disease⁷⁸.

TLR2 upregulation in the gut of UC patients has been associated with increased antigenic stimulation of inflammatory and immune pathways activated by ligand binding⁷⁹. Macrophages in the intestinal mucosa can rapidly engage in Toll-like receptor-mediated inflammatory responses to prevent pathogen invasion, but these innate immune responses can also trigger UC⁸⁰. Documented increased expression of *TLR2* in macrophages from IBD patients⁸¹ corroborates our finding showing a positive correlation between *TLR2* and M0/M1 macrophages. Hence, we hypothesize that M0/M1 macrophages may mediate the inflammatory response to UC through *TLR2*. *TLR8* is a key component of innate and adaptive immunity, and it has been shown that its expression is increased in UC patients and that mRNA levels are positively correlated with the severity of intestinal inflammation as well as the severity of inflammation⁷⁵.

We have mapped a proposed mechanism for the main results of this study (Fig. 8). Certainly, the above speculation requires further studies to verify the role of immune response in UC via the mutual regulation between hub genes and immune-infiltrated cells. The present study presented also some limitations. First, it was based on the GEO database, which is a secondary mining and analysis database of previously published datasets. Hence, the experimental results may differ from the conclusions of previous experiments, most likely due to biased data analysis caused by the small sample size. Second, the CIBERSORT deconvolution algorithm is based on limited genetic data, which may lead to inaccurate results due to different disease predisposing factors and the plasticity of disease phenotypes. Here, CIBERSORT was used to identify potential immune-related genes or immune infiltrating cells in UC. In addition, group centrality metrics⁸² are particularly useful because they describe the importance common to all hub genes and can be used to identification of hub genes by tools such as keyplayer^{83,84} (<http://www.analytictech.com/keyplayer/keyplayer.htm>) or Pyntacle⁸⁵. Although we did not use other network tools or methods in this study, perhaps we will enrich our study with these tools for other studies in the future. Nevertheless, our study may still provide compelling evidence for further research on the potential of the identified immune-infiltrated cells or immune-related genes for the treatment and diagnosis of UC (Supplementary file).

Conclusion

13 hub genes associated with immune-infiltrated cells in UC were identified in this study; they include *CXCL13*, *CXCL10*, *CXCL9*, *CXCL8*, *CCL19*, *CTLA4*, *CCR1*, *PECAM1*, *CD163*, *CD69*, *IL7R*, *TLR8* and *TLR2*. Additionally, bioinformatics analysis revealed M0/M1 macrophages, $\gamma\delta$ T cells, plasma cells, memory B cells, and neutrophils

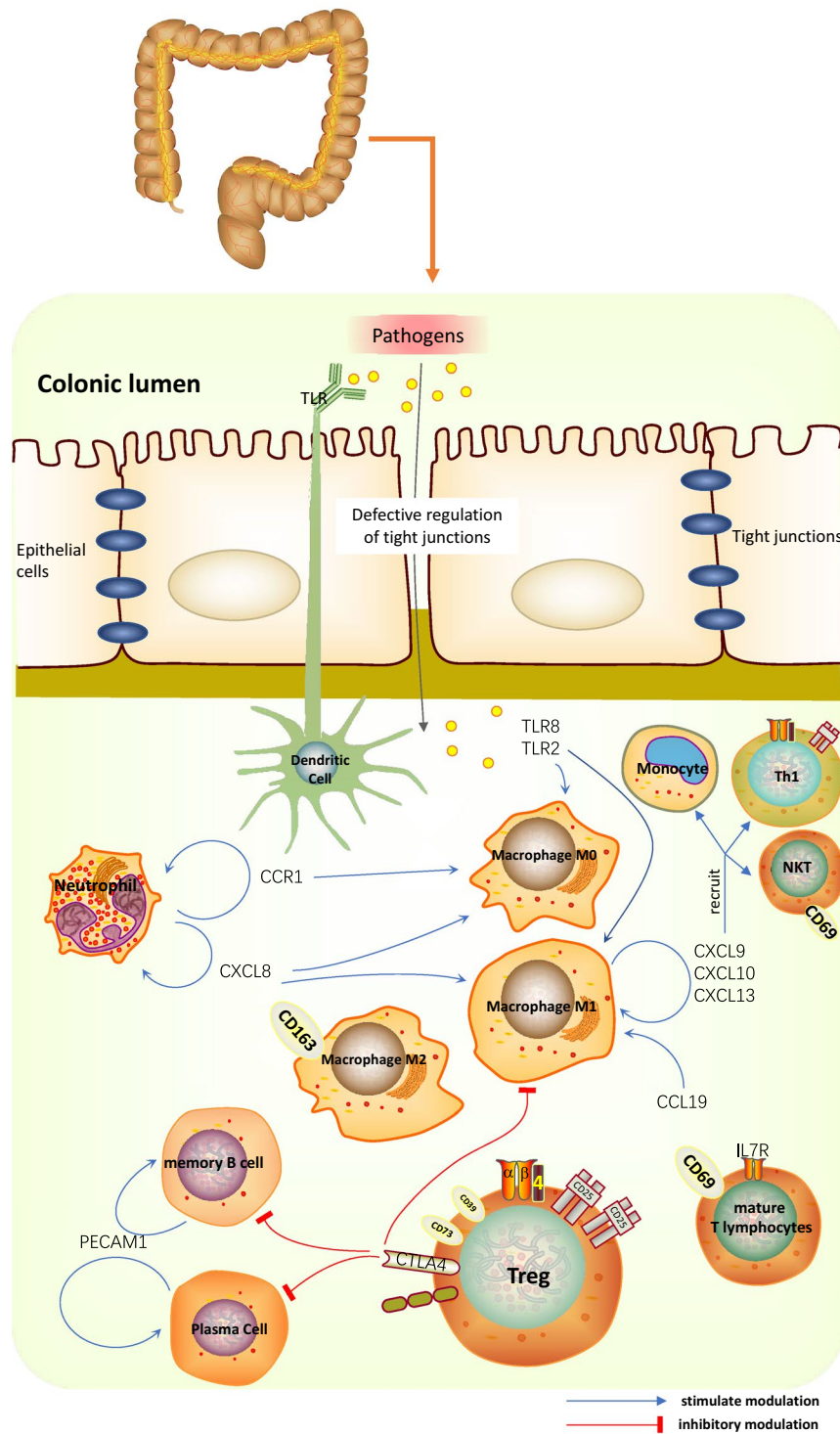


Figure 8. Hub genes and immune-infiltrated cells in UC.

to constitute the majority of immune-infiltrated cells in UC. These findings point to the immune response playing an important role in UC via mutual regulation between hub genes and immune-infiltrated cells.

Data availability

The datasets generated and/or analyzed during the current study are available in the [GEO] repository, [<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65114>].

References

- Danese, S. & Fiocchi, C. Ulcerative colitis. *N. Engl. J. Med.* **365**(18), 1713–1725 (2011).
- Kornbluth, A., Sachar, D. B., G. Practice Parameters Committee of the American College of. Ulcerative colitis practice guidelines in adults: American College of Gastroenterology, Practice Parameters Committee. *Am. J. Gastroenterol.* **105**(3), 501–523 (2010) (quiz 524).
- Maaser, C. *et al.* ECCO-ESGAR Guideline for Diagnostic Assessment in IBD Part 1: Initial diagnosis, monitoring of known IBD, detection of complications. *J. Crohns Colitis* **13**(2), 144–164 (2019).
- Ungaro, R. *et al.* Ulcerative colitis. *Lancet* **389**(10080), 1756–1770 (2017).
- Dignass, A. *et al.* Second European evidence-based consensus on the diagnosis and management of ulcerative colitis part 2: current management. *J. Crohns Colitis* **6**(10), 991–1030 (2012).
- Ordas, I. *et al.* Ulcerative colitis. *Lancet* **380**(9853), 1606–1619 (2012).
- Eisenstein, M. Ulcerative colitis: Towards remission. *Nature* **563**(7730), S33 (2018).
- Ramos, G. P. & Papadakis, K. A. Mechanisms of disease: Inflammatory bowel diseases. *Mayo Clin. Proc.* **94**(1), 155–165 (2019).
- Palmieri, O. *et al.* Microbiome analysis of mucosal ileoanal pouch in ulcerative colitis patients revealed impairment of the pouches immunometabolites. *Cells* **10**(11), 3243 (2021).
- Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**(1), 207–210 (2002).
- Franceschini, A. *et al.* STRING v9.1: Protein–protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* **41**(1), D808–D815 (2013).
- Smoot, M. E. *et al.* Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics* **27**(3), 431–432 (2011).
- Bandettini, W. P. *et al.* MultiContrast Delayed Enhancement (MCOE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: A clinical validation study. *J. Cardiovasc. Magn. Reson.* **14**, 83 (2012).
- Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**(1), 27–30 (2000).
- Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**(11), 1947–1951 (2019).
- Kanehisa, M. *et al.* KEGG: Integrating viruses and cellular organisms. *Nucleic Acids Res.* **49**(D1), D545–D551 (2021).
- Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **12**(5), 453–457 (2015).
- Kordjazy, N. *et al.* Role of toll-like receptors in inflammatory bowel disease. *Pharmacol. Res.* **129**, 204–215 (2018).
- Nanki, K. *et al.* Somatic inflammatory gene mutations in human ulcerative colitis epithelium. *Nature* **577**(7789), 254–259 (2020).
- Latella, G. & Papi, C. Crucial steps in the natural history of inflammatory bowel disease. *World J. Gastroenterol.* **18**(29), 3790–3799 (2012).
- Baumgart, D. C. & Sandborn, W. J. Crohn's disease. *Lancet* **380**(9853), 1590–1605 (2012).
- Elia, G. & Guglielmi, G. CXCL9 chemokine in ulcerative colitis. *Clin. Ter.* **169**(5), e235–e241 (2018).
- Lee, D. S. *et al.* Expression of chemokine CCL28 in ulcerative colitis patients. *Gut Liver* **15**(1), 70–76 (2021).
- Zhang, J. *et al.* Investigation of potential genetic biomarkers and molecular mechanism of ulcerative colitis utilizing bioinformatics analysis. *Biomed Res. Int.* **2020**, 4921387 (2020).
- Kobayashi, T. *et al.* Ulcerative colitis. *Nat. Rev. Dis. Primers* **6**(1), 74 (2020).
- Nakase, H. *et al.* The influence of cytokines on the complex pathology of ulcerative colitis. *Autoimmun. Rev.* **21**(3), 103017 (2022).
- Boland, B. S. *et al.* Heterogeneity and clonal relationships of adaptive immune cells in ulcerative colitis revealed by single-cell analyses. *Sci. Immunol.* **5**(50), eabb4432 (2020).
- Timmermans, W. M. *et al.* B-cell dysregulation in Crohn's disease is partially restored with infliximab therapy. *PLoS ONE* **11**(7), e0160103 (2016).
- Choy, M. C., Visvanathan, K. & De Cruz, P. An overview of the innate and adaptive immune system in inflammatory bowel disease. *Inflamm. Bowel Dis.* **23**(1), 2–13 (2017).
- Wang, X. *et al.* Circulating memory B cells and plasmablasts are associated with the levels of serum immunoglobulin in patients with ulcerative colitis. *J. Cell. Mol. Med.* **20**(5), 804–814 (2016).
- Hanai, H. *et al.* A new paradigm in ulcerative colitis: Regulatory T cells are key factor which induces/exacerbates UC through an immune imbalance. *Mol. Immunol.* **54**(2), 173–180 (2013).
- Nanno, M. *et al.* Exacerbating role of gammadelta T cells in chronic colitis of T-cell receptor alpha mutant mice. *Gastroenterology* **134**(2), 481–490 (2008).
- Lo Presti, E. *et al.* Characterization of gammadelta T cells in intestinal mucosa from patients with early-onset or long-standing inflammatory bowel disease and their correlation with clinical status. *J. Crohns Colitis* **13**(7), 873–883 (2019).
- Penrose, H. M. *et al.* Ulcerative colitis immune cell landscapes and differentially expressed gene signatures determine novel regulators and predict clinical response to biologic therapy. *Sci. Rep.* **11**(1), 9010 (2021).
- Bain, C. C. & Mowat, A. M. Macrophages in intestinal homeostasis and inflammation. *Immunol. Rev.* **260**(1), 102–117 (2014).
- Muller, P. A., Matheis, F. & Mucida, D. Gut macrophages: Key players in intestinal immunity and tissue physiology. *Curr. Opin. Immunol.* **62**, 54–61 (2020).
- Dharmasiri, S. *et al.* Human intestinal macrophages are involved in the pathology of both ulcerative colitis and Crohn disease. *Inflamm. Bowel Dis.* **27**(10), 1641–1652 (2021).
- Zhuang, H. *et al.* Tiliroside ameliorates ulcerative colitis by restoring the M1/M2 macrophage balance via the HIF-1 α /glycolysis pathway. *Front. Immunol.* **12**, 649463 (2021).
- Verspaget, H. W. *et al.* Diminished neutrophil function in Crohn's disease and ulcerative colitis identified by decreased oxidative metabolism and low superoxide dismutase content. *Gut* **29**(2), 223–228 (1988).
- MacDermott, R. P., Sanderson, I. R. & Reinecker, H. C. The central role of chemokines (chemotactic cytokines) in the immunopathogenesis of ulcerative colitis and Crohn's disease. *Inflamm. Bowel Dis.* **4**(1), 54–67 (1998).
- Bie, Y. *et al.* The crucial role of CXCL8 and its receptors in colorectal liver metastasis. *Dis. Markers* **2019**, 8023460 (2019).
- Wanninger, J. *et al.* Adiponectin-stimulated CXCL8 release in primary human hepatocytes is regulated by ERK1/ERK2, p38 MAPK, NF- κ B, and STAT3 signaling pathways. *Am. J. Physiol. Gastrointest. Liver Physiol.* **297**(3), G611–G618 (2009).
- Yoshida, N. *et al.* Interleukin-8 production via protease-activated receptor 2 in human esophageal epithelial cells. *Int. J. Mol. Med.* **19**(2), 335–340 (2007).
- Metzemaekers, M. *et al.* Truncation of CXCL8 to CXCL8(9–77) enhances actin polymerization and in vivo migration of neutrophils. *J. Leukoc. Biol.* **107**(6), 1167–1173 (2020).
- Espallat, M. P., Kew, R. R. & Obeid, L. M. Sphingolipids in neutrophil function and inflammatory responses: Mechanisms and implications for intestinal immunity and inflammation in ulcerative colitis. *Adv. Biol. Regul.* **63**, 140–155 (2017).
- Lasagni, L. *et al.* An alternatively spliced variant of CXCR3 mediates the inhibition of endothelial cell growth induced by IP-10, Mig, and I-TAC, and acts as functional receptor for platelet factor 4. *J. Exp. Med.* **197**(11), 1537–1549 (2003).

47. Uguccioni, M. *et al.* Increased expression of IP-10, IL-8, MCP-1, and MCP-3 in ulcerative colitis. *Am. J. Pathol.* **155**(2), 331–336 (1999).
48. Tarique, A. A. *et al.* Phenotypic, functional, and plasticity features of classical and alternatively activated human macrophages. *Am. J. Respir. Cell Mol. Biol.* **53**(5), 676–688 (2015).
49. Cole, K. E. *et al.* Interferon-inducible T cell alpha chemoattractant (I-TAC): A novel non-ELR CXC chemokine with potent activity on activated T cells through selective high affinity binding to CXCR3. *J. Exp. Med.* **187**(12), 2009–2021 (1998).
50. Boshagh, M. A. *et al.* ELR positive CXCL chemokines are highly expressed in an animal model of ulcerative colitis. *J. Inflamm. Res.* **12**, 167–174 (2019).
51. Capitano, M. L. *et al.* A humanized monoclonal antibody against the endothelial chemokine CCL21 for the diagnosis and treatment of inflammatory bowel disease. *PLoS ONE* **16**(7), e0252805 (2021).
52. Nguyen, V. Q. *et al.* Noncanonical NF-kappaB signaling upregulation in inflammatory bowel disease patients is associated with loss of response to anti-TNF agents. *Front. Pharmacol.* **12**, 655887 (2021).
53. Zhang, H. *et al.* Effect of secondary lymphoid tissue chemokine suppression on experimental ulcerative colitis in mice. *Genet. Mol. Res.* **13**(2), 3337–3345 (2014).
54. Xuan, W. *et al.* The chemotaxis of M1 and M2 macrophages is regulated by different chemokines. *J. Leukoc. Biol.* **97**(1), 61–69 (2015).
55. Gupta, R. A. *et al.* Effect of piperine on inhibition of FFA induced TLR4 mediated inflammation and amelioration of acetic acid induced ulcerative colitis in mice. *J. Ethnopharmacol.* **164**, 239–246 (2015).
56. Walker, L. S. & Sansom, D. M. The emerging role of CTLA4 as a cell-extrinsic regulator of T cell responses. *Nat. Rev. Immunol.* **11**(12), 852–863 (2011).
57. Hosseini, A. *et al.* CTLA-4: From mechanism to autoimmune therapy. *Int. Immunopharmacol.* **80**, 106221 (2020).
58. Wang, S. *et al.* Circulating follicular regulatory T cells could inhibit Ig production in a CTLA-4-dependent manner but are dysregulated in ulcerative colitis. *Mol. Immunol.* **114**, 323–329 (2019).
59. Magnusson, M. K. *et al.* Impaired butyrate induced regulation of T cell surface expression of CTLA-4 in patients with ulcerative colitis. *Int. J. Mol. Sci.* **22**(6), 3084 (2021).
60. Camarillo, G. F. *et al.* Gene expression profiling of mediators associated with the inflammatory pathways in the intestinal tissue from patients with ulcerative colitis. *Mediat. Inflamm.* **2020**, 9238970 (2020).
61. Mease, P. *et al.* Abatacept in the treatment of patients with psoriatic arthritis: Results of a six-month, multicenter, randomized, double-blind, placebo-controlled, phase II trial. *Arthritis Rheum.* **63**(4), 939–948 (2011).
62. Kremer, J. M. *et al.* Treatment of rheumatoid arthritis by selective inhibition of T-cell activation with fusion protein CTLA4Ig. *N. Engl. J. Med.* **349**(20), 1907–1915 (2003).
63. Sandborn, W. J. *et al.* Abatacept for Crohn's disease and ulcerative colitis. *Gastroenterology* **143**(1), 62–69.e4 (2012).
64. Wing, J. B. *et al.* Regulatory T cells control antigen-specific expansion of T_H cell number and humoral immune responses via the coreceptor CTLA-4. *Immunity* **41**(6), 1013–1025 (2014).
65. Cibrán, D. & Sanchez-Madrid, F. CD69: From activation marker to metabolic gatekeeper. *Eur. J. Immunol.* **47**(6), 946–953 (2017).
66. Gonzalez-Amaro, R. *et al.* Is CD69 an effective brake to control inflammatory diseases?. *Trends Mol. Med.* **19**(10), 625–632 (2013).
67. Okhrimenko, A. *et al.* Human memory T cells from the bone marrow are resting and maintain long-lasting systemic memory. *Proc. Natl. Acad. Sci. USA* **111**(25), 9229–9234 (2014).
68. Gorabi, A. M. *et al.* The pivotal role of CD69 in autoimmunity. *J. Autoimmun.* **111**, 102453 (2020).
69. Dige, A. *et al.* Soluble CD163, a specific macrophage activation marker, is decreased by anti-TNF-alpha antibody treatment in active inflammatory bowel disease. *Scand. J. Immunol.* **80**(6), 417–423 (2014).
70. Chapuy, L. *et al.* Transcriptomic analysis and high-dimensional phenotypic mapping of mononuclear phagocytes in mesenteric lymph nodes reveal differences between ulcerative colitis and Crohn's disease. *J. Crohns Colitis* **14**(3), 393–405 (2020).
71. Govender, D. *et al.* CD31 (JC70) expression in plasma cells: An immunohistochemical analysis of reactive and neoplastic plasma cells. *J. Clin. Pathol.* **50**(6), 490–493 (1997).
72. Jackson, D. E. The unfolding tale of PECAM-1. *FEBS Lett.* **540**(1–3), 7–14 (2003).
73. Baumann, C. I. *et al.* PECAM-1 is expressed on hematopoietic stem cells throughout ontogeny and identifies a population of erythroid progenitors. *Blood* **104**(4), 1010–1016 (2004).
74. Woodfin, A., Voisin, M. B. & Nourshargh, S. PECAM-1: A multi-functional molecule in inflammation and vascular biology. *Arterioscler. Thromb. Vasc. Biol.* **27**(12), 2514–2523 (2007).
75. Sanchez-Munoz, F. *et al.* Transcript levels of toll-like receptors 5, 8 and 9 correlate with inflammatory activity in ulcerative colitis. *BMC Gastroenterol.* **11**, 138 (2011).
76. Mazzucchelli, R. & Durum, S. K. Interleukin-7 receptor expression: Intelligent design. *Nat. Rev. Immunol.* **7**(2), 144–154 (2007).
77. Watanabe, M. *et al.* Interleukin 7 is produced by human intestinal epithelial cells and regulates the proliferation of intestinal mucosal lymphocytes. *J. Clin. Investig.* **95**(6), 2945–2953 (1995).
78. Anderson, C. A. *et al.* Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat. Genet.* **43**(3), 246–252 (2011).
79. Fan, Y. & Liu, B. Expression of Toll-like receptors in the mucosa of patients with ulcerative colitis. *Exp. Ther. Med.* **9**(4), 1455–1459 (2015).
80. Wang, L. *et al.* MTDH promotes intestinal inflammation by positively regulating TLR signalling. *J. Crohns Colitis* **15**(12), 2103–2117 (2021).
81. Campos, N. *et al.* Macrophages from IBD patients exhibit defective tumour necrosis factor-alpha secretion but otherwise normal or augmented pro-inflammatory responses to infection. *Immunobiology* **216**(8), 961–970 (2011).
82. Everett, M. G. & Borgatti, S. P. The centrality of groups and classes. *J. Math. Sociol.* **23**(3), 181–201 (1999).
83. Borgatti, S. P. Identifying sets of key players in a social network. *Comput. Math. Organ. Theory* **12**(1), 21–34 (2006).
84. An, W. & Liu, Y.-H. keyplayer: An R package for locating key players in social networks. *R Journal* **8**(1), 257–268 (2016).
85. Parca, L. *et al.* Pyntacle: A parallel computing-enabled framework for large-scale network biology analysis. *Gigascience* **9**(10), giaa115 (2020).

Acknowledgements

We would like to thank Editage (www.editage.cn) for English language editing.

Author contributions

Conception and design: W.T.H.; Administrative support: X.Q.C.; Provision of study materials: M.X.Z.; Collection and assembly of data: T.Y.F.; Data analysis and interpretation: W.T.H.; Manuscript writing: All authors; Final approval of manuscript: All authors.

Funding

This work was supported by Quanzhou City Science & Technology Program of China (No. 2019C071R) and Natural Science Foundation of Fujian Province (No. 2020J01231).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-33292-y>.

Correspondence and requests for materials should be addressed to X.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023