



OPEN ACCESS

EDITED BY

Ming-Chin Lin,
Taipei Medical University, Taiwan

REVIEWED BY

Rakeeb Ahmad Mir,
Central University of Kashmir, India
Sylvie Amu,
University College Cork, Ireland

*CORRESPONDENCE

Xudong Tang
txdly@sina.com

SPECIALTY SECTION

This article was submitted to
Family Medicine and Primary Care,
a section of the journal
Frontiers in Public Health

RECEIVED 25 July 2022

ACCEPTED 17 October 2022

PUBLISHED 31 October 2022

CITATION

Huang J, Zhang J, Wang F, Zhang B
and Tang X (2022) Revealing immune
infiltrate characteristics and potential
diagnostic value of immune-related
genes in ulcerative colitis: An
integrative genomic analysis.
Front. Public Health 10:1003002.
doi: 10.3389/fpubh.2022.1003002

COPYRIGHT

© 2022 Huang, Zhang, Wang, Zhang
and Tang. This is an open-access
article distributed under the terms of
the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution
or reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Revealing immune infiltrate characteristics and potential diagnostic value of immune-related genes in ulcerative colitis: An integrative genomic analysis

Jinke Huang¹, Jiaqi Zhang^{1,2}, Fengyun Wang^{1,2},
Beihua Zhang^{1,2} and Xudong Tang^{1,2*}

¹Department of Gastroenterology, Xiyuan Hospital of China Academy of Chinese Medical Sciences, Beijing, China, ²Institute of Digestive Diseases, Xiyuan Hospital of China Academy of Chinese Medical Sciences, Beijing, China

Objectives: Ulcerative colitis (UC) is an autoimmune disease of the colon. The aim of this study was to explore the characteristics of immune infiltrates in UC patients and identify immune-related diagnostic biomarkers for UC.

Methods: Three gene expression profiles were acquired from the GEO database, followed by identification of differentially expressed genes (DEGs) by Linear Modeling of Microarray Data. Enrichment analysis of Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) and Disease Ontology (DO) were performed to analyze the biological functions of DEGs. Subsequently, the single sample gene set enrichment analysis (ssGSEA) was performed to identify immune infiltration characteristics of UC. Correlations between diagnostic genes and immune infiltration were explored to identify markers with the greatest diagnostic potential, and a UC diagnostic model was subsequently constructed. Finally, the prediction performance of the model was quantified by nomogram, non-correlated nomogram, and ROC curve.

Results: A total of 3111 DEGs (1,608 up-regulated and 1,503 down-regulated genes) were identified. DEGs were significantly involved in the immune system and UC-related pathways. Immune infiltration profiles of colonic tissue were significantly different between healthy individuals and UC patients. High proportions of resting of aDCs, B cells, CD8⁺ T cells, DCs, iDCs, Macrophages, Neutrophils, pDCs, T helper cells, Tfh, Th1 cells, Th2 cells, TIL and Treg were found in UC samples. A 5-gene based diagnostic prediction model was constructed and the results of nomogram, non-correlated nomogram and ROC curve suggested the powerful diagnostic value of the model.

Conclusions: This study identified the immune infiltrate characteristics and 5 immune-related genes for UC. The model based on the immune-related genes facilitates the early diagnosis of UC and provides a basis for the evaluation of the prognosis of UC.

KEYWORDS

ulcerative colitis, immune infiltration, genes, diagnostic value, genomic analysis

Introduction

Ulcerative colitis (UC) is a complex disease characterized by chronic inflammation of the colon (1). Worldwide, UC is estimated to affect 9–100,000 people annually, and the incidence is increasing year by year (2). The growing number of UC patients places a heavy economic burden on society, with direct and indirect costs ranging from \$8.1–14.9 billion per year in the United States and \$12.5–29.1 billion in Europe (3). The treatment goal in UC is the induction and maintenance of remission. Although the therapeutic armamentarium is expanding, the treatment of UC is highly challenging because of its incompletely understood pathogenesis (4). Therefore, an in-depth understanding of disease pathogenesis and identification of biomarkers of disease progression at the molecular level may provide new ideas for the early diagnosis of UC.

The etiology and pathogenesis of UC are not fully understood, and it is mainly thought to be caused by an enhanced immune response to the gut microbiota in genetically susceptible individuals (5). Many studies have investigated the function of various immune cells, but it has been challenging to predict the role of all immune subsets in UC in an integrated manner. Initial activation of innate immunity causes

a non-specific response, and then, sustained stimulation of inflammation will activate adaptive immunity, which may lead to persistent chronic inflammation (6). Accumulating evidence suggests that both innate and adaptive immune abnormalities are responsible for the abnormal inflammatory response in the gut (7). Inflammation associated with inflammatory bowel disease (IBD) has been reported to be closely associated with aberrant immune response elicited by CD4 T cells and dendritic cells (8–10). IRF5 contributes to the regulation of T cell signaling and modulates cytokine secretion to promote inflammation in UC (11). Neutrophil HGF-MET signaling can also contribute to the progression of UC (12). Furthermore, infiltrating immune cells are present in the intestinal mucosa of individuals with UC (13), and increased immune cell infiltration may correlate with the severity and recurrence of UC (14, 15). All these findings suggest a key role of immune cells in the pathogenesis of UC, and molecules associated with these cells may serve as new biomarkers for UC.

Gene chip is a genetic detection technology that can detect all expression information of all genes from a sample and reveal numerous genes activated in different tissues and their physiological and pathological states (16, 17). Currently, microarray technology integrated with bioinformatics analysis

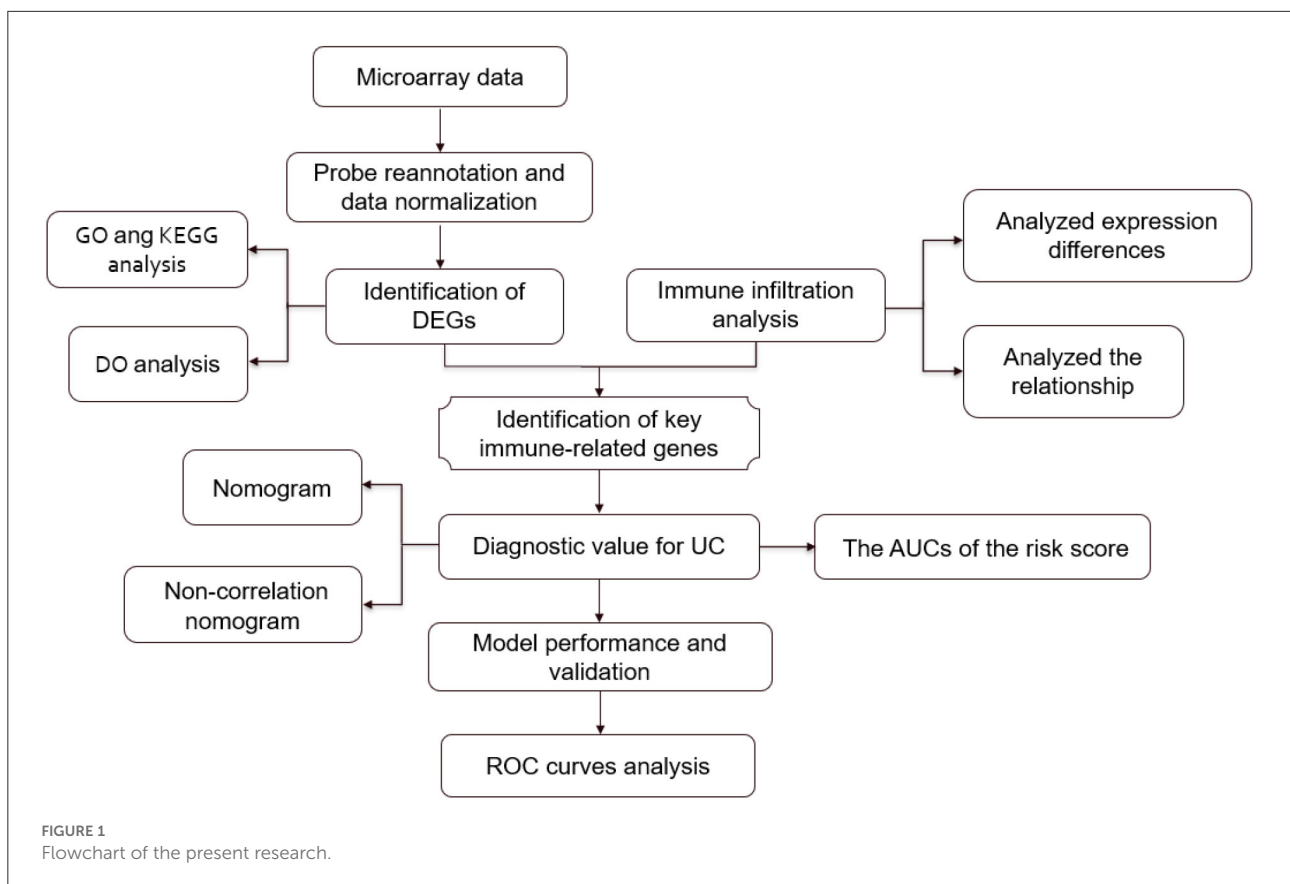


TABLE 1 Details of gene expression profiles.

Dataset	Platform	Tissue	Normal	UC	Reference (PMID)	Hyperlinks
GSE87473	GPL13158	Colon	106	21	29401083	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE87473
GSE75214	GPL6244	Colon	97	11	28885228	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE75214
GSE92415	GPL13158	Colon	87	21	23735746	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE92415

TABLE 2 R packages for bioinformatics analysis.

R package	Version	Description
Limma	3.50.3	Probe reannotation and data normalization, identification of DEGs
Pheatmap	1.0.12	Plotting heat map
Ggplot2	3.3.6	Plotting volcano plot, visualization of enrichment analysis results
Clusterprofiler	4.2.2	GO and KEGG enrichment analyses
DOSE	3.20.1	Disease ontology enrichment analysis
Corrplot	0.92	Correlation matrix visualization
Ggpubr	0.4.0	Plotting boxplot
Psych	2.2.5	Correlation analysis of DEGs and immune infiltration
RMS	6.3-0	Construction of diagnostic model
ROCR	1.0-11	ROC analysis

has been widely used to explore pathological features and identify potential novel biomarkers for various diseases (18, 19). Based on large-scale microarray gene expression data, this study applied integrated bioinformatics analysis to explore the molecular mechanisms of UC. Moreover, we focused on identifying core genes associated with immune infiltrating cells and used these core biomarkers to construct a risk prediction model for UC with the aim of providing new ideas for early diagnosis of UC. The flow chart of the present research is shown in Figure 1.

Materials and methods

Microarray data acquisition

Gene expression profiles were acquired from GEO database (www.ncbi.nlm.nih.gov/geo/) (20) with the following criteria: (a) patients were diagnosed as UC; (b) data on colonic tissue from healthy controls and UC patients from the same GEO platform; (c) datasets inclusion with at least 10 UC and healthy tissue samples; (d) GEO platforms containing >5,000 genes. Finally, three gene expression profiles (GSE87473, GSE92415, and GSE 75214) were included. Table 1 shows the details of the gene expression profiles.

Identification of differentially expressed genes

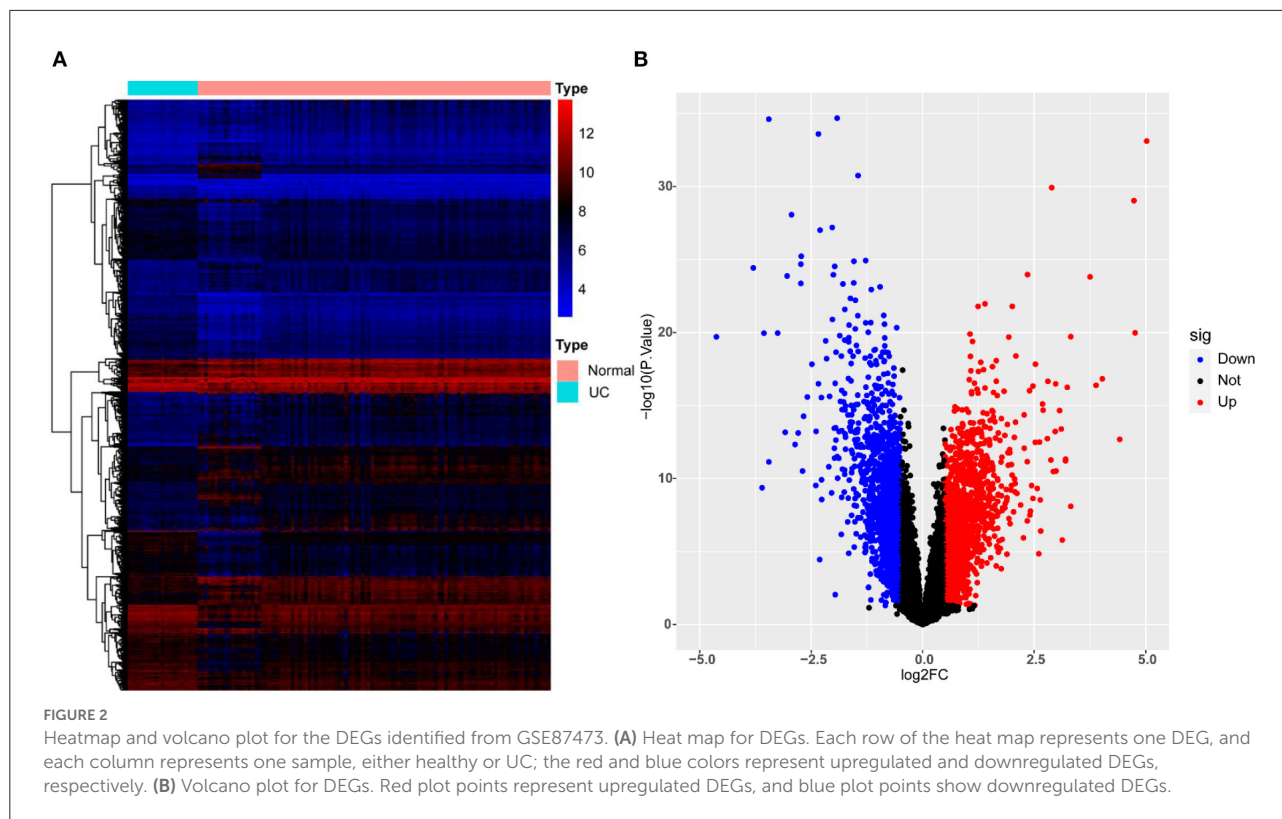
After downloading the microarray expression matrices, probe names were converted to gene symbols using R software (version 4.1.2). DEGs between UC colon tissue and healthy colon tissue were screened by the “limma” package in R software, and the threshold for DEG was set to $|\log_2 \text{Fold change (FC)}| > 0.5$ and $p\text{-value} < 0.05$.

Functional analysis of DEGs

Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Disease Ontology (DO) enrichment analyses were performed on the screened DEGs from GSE87473 by the “clusterProfiler,” and “DOSE” packages in R software. The threshold for enrichment analysis was set to $p\text{-value} < 0.05$. The “ggplot2” package was applied to visualize the enrichment results.

Immune infiltration analysis

Single sample gene set enrichment analysis (ssGSEA) method was applied to quantify the enrichment scores of immune cells (activated DCs (aDCs), B cells, CD8+ T cells, DCs, interdigitating DCs (iDCs), macrophages, mast cells, neutrophils, natural killer (NK) cells, plasmacytoid DCs (pDCs), T helper cells, T follicular helper (Tfh) cells, T helper1 (Th1) cells, T helper 2 (Th2) cells, tumor infiltrating lymphocytes (TIL), regulatory T (Treg) cells), and immune functions (antigen presenting cell (APC) co-inhibition, APC co-stimulation, chemokine receptors (CCR), check-point, cytolytic activity, human leukocyte antigen (HLA), inflammation-promoting, major histocompatibility complex (MHC) class I, Parainflammation, T cell co-inhibition, T cell co-stimulation, type-I interferon (IFN) response, type II IFN response) between control and UC group. A $p\text{-value} < 0.05$ was used to filter the samples. Heat map of the 29 types of immune cells and immune function in samples was produced by the “pheatmap” package. Levels of immune cells and immune function between UC and control samples were visualized by the “ggpubr” package. Correlative heat map was performed using



the “corrplot” package to reveal the correlation of immune cells and immune function.

The construction of diagnostic model

The top genes with the most significant differences in expression between healthy and UC samples were considered as diagnostic genes for UC, and they were evaluated by the “psych” package for their relevance to immune cell and immune function subtypes. After the feature selection, the diagnostic genes most strongly associated with immune infiltration were used to construct diagnostic models with “rms” package. The prediction performance of the model was quantified by nomogram, non-correlated nomogram, and receiver operating characteristic (ROC) curve which was performed with “ROCR” package.

Model performance and validation

The expression data of GSE75214 and GSE92415 were used to verify the robustness of diagnostic model. The area under the curve (AUC) from a ROC curve analysis was calculated to test the diagnostic performance of the model: $ROC-AUC \geq 0.9$ indicates outstanding discrimination; $0.8 \leq ROC-AUC < 0.9$ indicates excellent discrimination; $0.7 \leq ROC-AUC < 0.8$

indicates acceptable discrimination; and $ROC = 0.5$ indicates no discrimination (21).

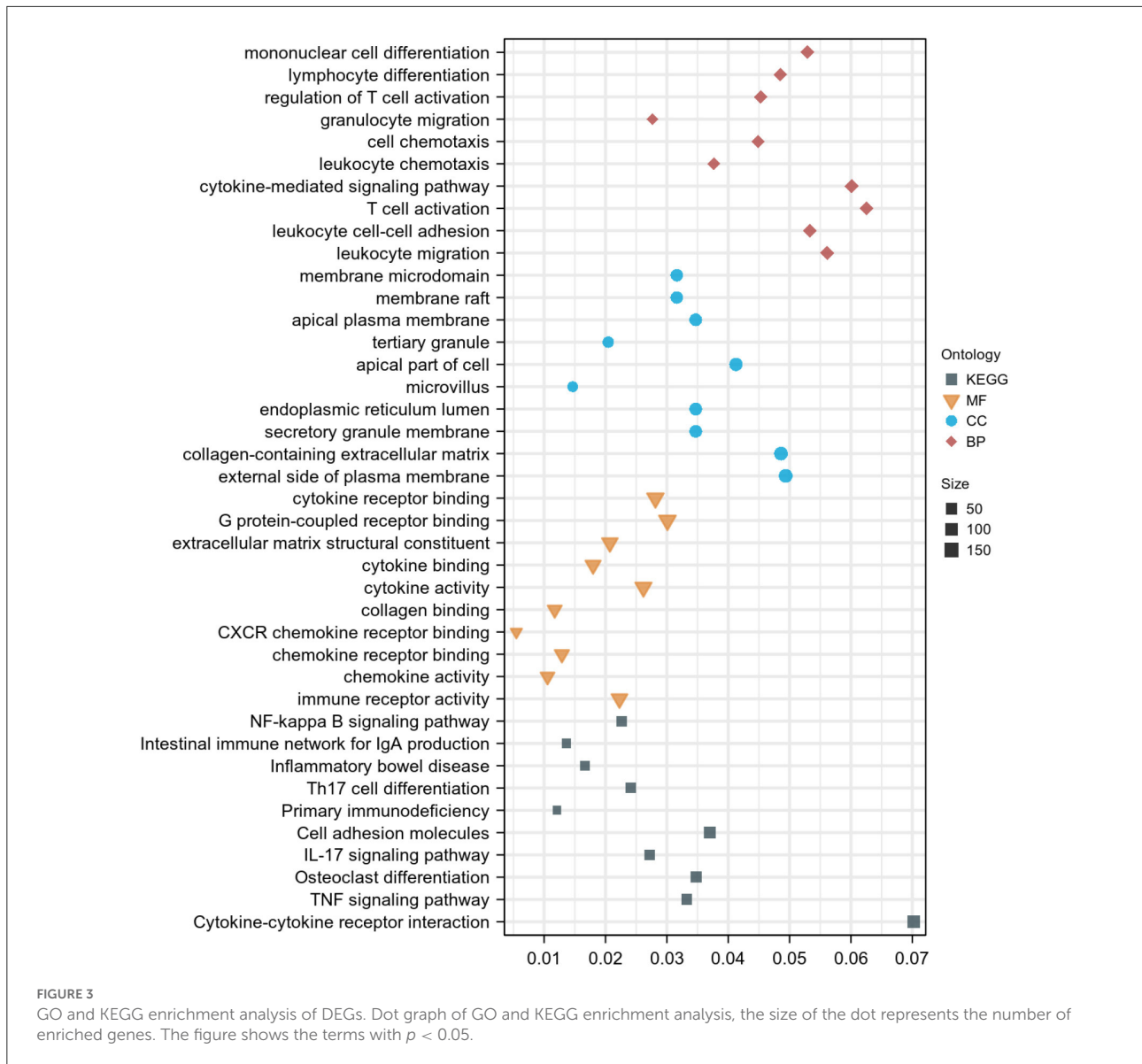
Statistical analysis

Categorical variables were presented as percentages, while continuous variables were presented as the mean \pm standard deviation. All data analyses in this study were performed using R software (version 4.1.2), and the main packages that were used for t bioinformatics analysis are provided in Table 2. A p -value < 0.05 was considered significant for screening DEGs, enrichment analysis, correlation analysis, and immune infiltration analysis.

Results

Differential gene screening

3111 DEGs (1,608 up-regulated and 1,503 down-regulated genes) were identified from GSE87473. The top 10 up-regulated DEGs involved were: DUOX2, MMP3, SLC6A14, DEFB4A, TNIP3, S100A8, CXCL1, DUOXA2, REG1A, and MMP10, CALU while the top 10 down-regulated DEGs were: AQP8, SLC51A, CLDN8, HMGCS2, DPP10-AS1, PCK1, ABCG2, SLC26A2, GBA3, and MEP1B. Figure 2 presents the details of the heatmap and volcano plot of DEGs.



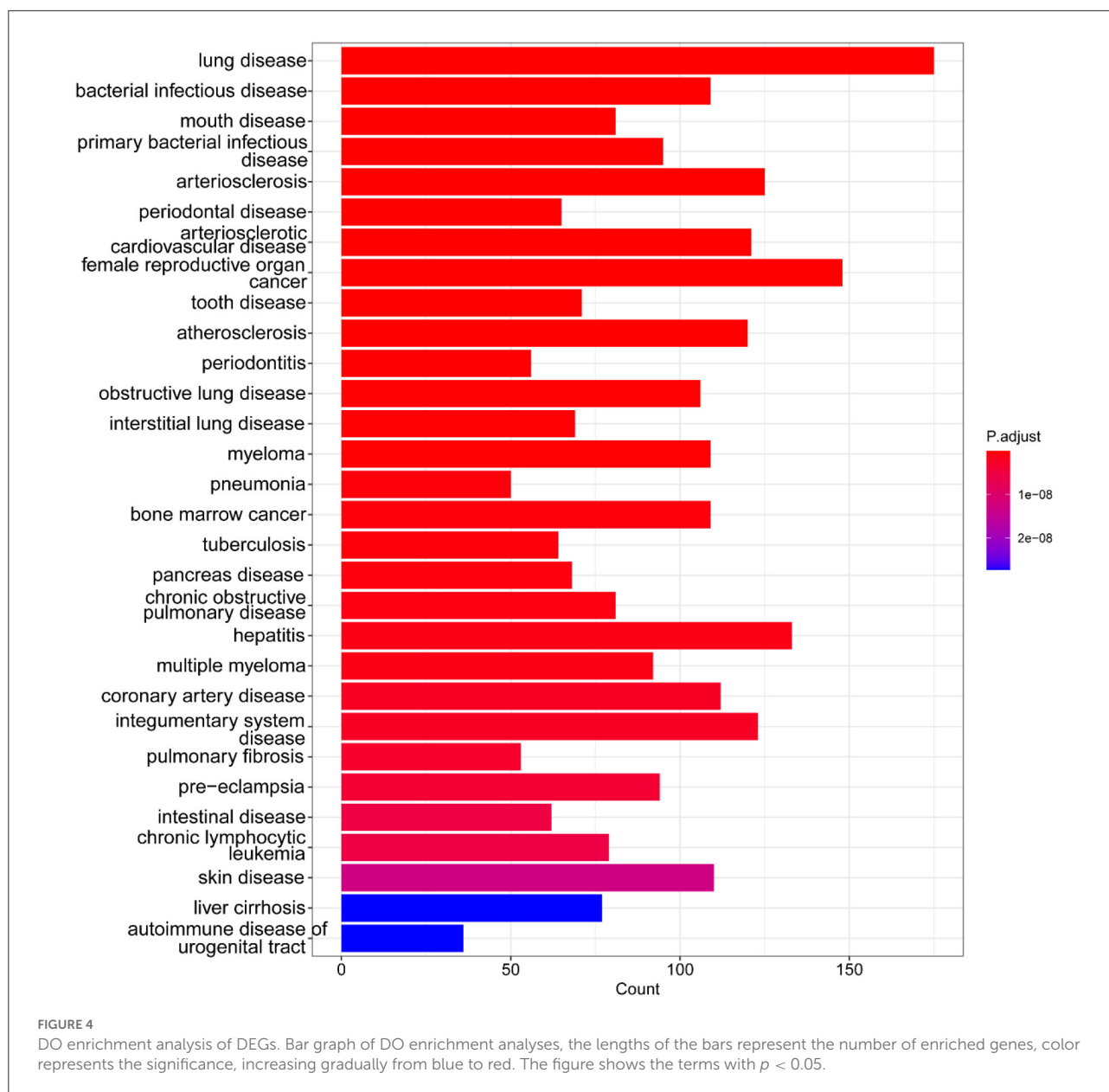
Functional enrichment analysis

For GO analysis, DEGs were significantly enriched in the following process: leukocyte migration, leukocyte cell-cell adhesion, T cell activation, cytokine-mediated signaling pathway, leukocyte chemotaxis, cell chemotaxis, granulocyte migration, regulation of T cell activation, lymphocyte differentiation, mononuclear cell differentiation, immune receptor activity, chemokine activity, chemokine receptor binding, CXCR chemokine receptor binding, collagen binding, cytokine activity, cytokine binding, extracellular matrix structural constituent, and G protein-coupled receptor binding (Figure 3). For KEGG analysis, genes were significantly enriched in immune-related pathways such as TNF signaling pathway,

Osteoclast differentiation, IL-17 signaling pathway, Th17 cell differentiation, and NF-kappa B signaling pathway (Figure 3). For DO analysis, DEGs were significantly enriched in infectious diseases, inflammatory diseases, and cancer (Figure 4).

Immune infiltration analysis

The normalized enrichment score of immune infiltrates is presented in the heat map (Figure 5). The results of differential analysis of immune cell revealed that UC patients had a higher level of aDCs, B cells, CD8⁺ T cells, DCs, iDCs, Macrophages, Neutrophils, pDCs, T helper cells, Tfh, Th1 cells, Th2 cells, TIL and Treg than healthy subjects (Table 3; Figure 6A).



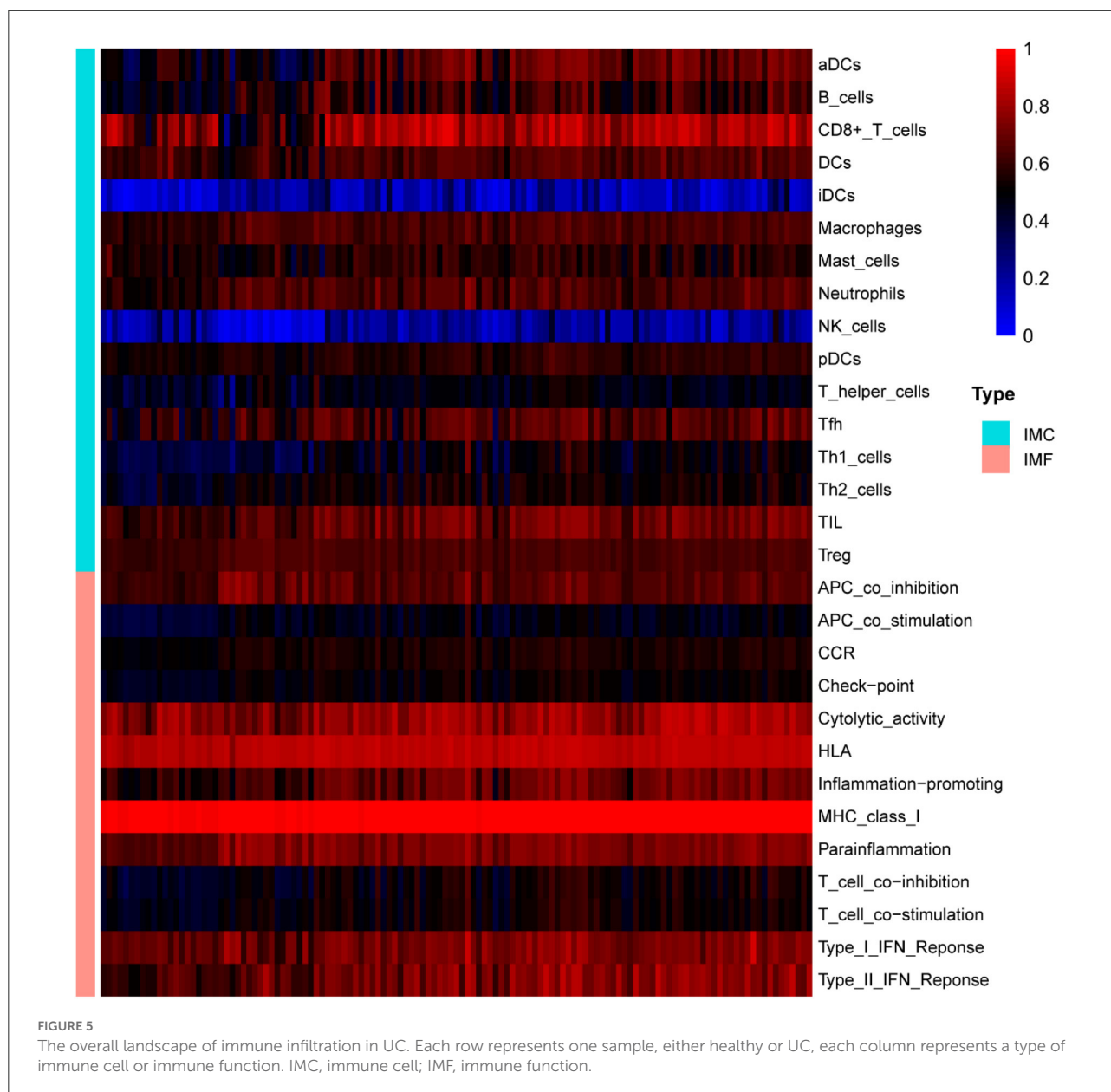
The results of differential immune function analysis revealed that significant APC co-inhibition, APC co-stimulation, CCR, Check-point, HLA, Inflammation-promoting, MHC class I, Parainflammation, T cell co-inhibition, T cell co-stimulation, Type I IFN Reponse, and Type II IFN Reponse were observed in UC patients (Table 3; Figure 6B). Details of these biomarkers in patients with UC are presented in Table 3.

The correlation analysis revealed that activated B cells were not related to Mast cells; $CD8^+$ T cells were not related to iDCs; iDCs were not related to NK cells; Macrophages were not related to Mast cells or NK cells; Mast cells were not related to Neutrophils, pDCs, Th2 cells or Treg; and NK cells were not related to Treg (Figure 7A). However, strong correlations

were observed for all other types of immune cells and immune function (Figure 7).

The construction of diagnostic model

Results of Pearson correlation analysis revealed that all upregulated diagnostic genes were significantly positively correlated with almost all immune cell subtypes and immune function subtypes (except $CD8^+$ T cells and NK cells). Similarly, almost all down-regulated diagnostic genes were negatively correlated with almost all immune cell subtypes and immune function subtypes (except $CD8^+$ T cells and NK cells) (Figure 8).



Inspired by the above results, five diagnostic genes most associated with immune infiltration (HMGCS2, CLDN8, AQP8, DEFB4A, MMP3) were used to construct a diagnostic model for UC. Details of these biomarkers in patients with UC are presented in [Table 4](#).

The nomogram showed the diagnostic efficacy of the model constructed with these predicted diagnostic genes for UC ([Figure 9A](#)). Based on the calibration curve predicted by the uncorrelated nomogram, the performance of the column line plot was close to the ideal model, suggesting that the predictive value of the model is credible ([Figure 9B](#)). Similarly, ROC-AUC of the risk score was 0.897, which indicates excellent discrimination of the model ([Figure 9C](#)).

Model performance and validation

To go step further validation, ROC curves were applied to assesses the prediction accuracy of the model. The ROC-AUC of the risk score was 0.871 in GSE75214 and 0.908 in GSE92415, respectively, indicating excellent discrimination of the model ([Figure 10](#)).

Discussion

The development of UC involves genetic susceptibility, environmental factors and disturbances in the gut microbiota

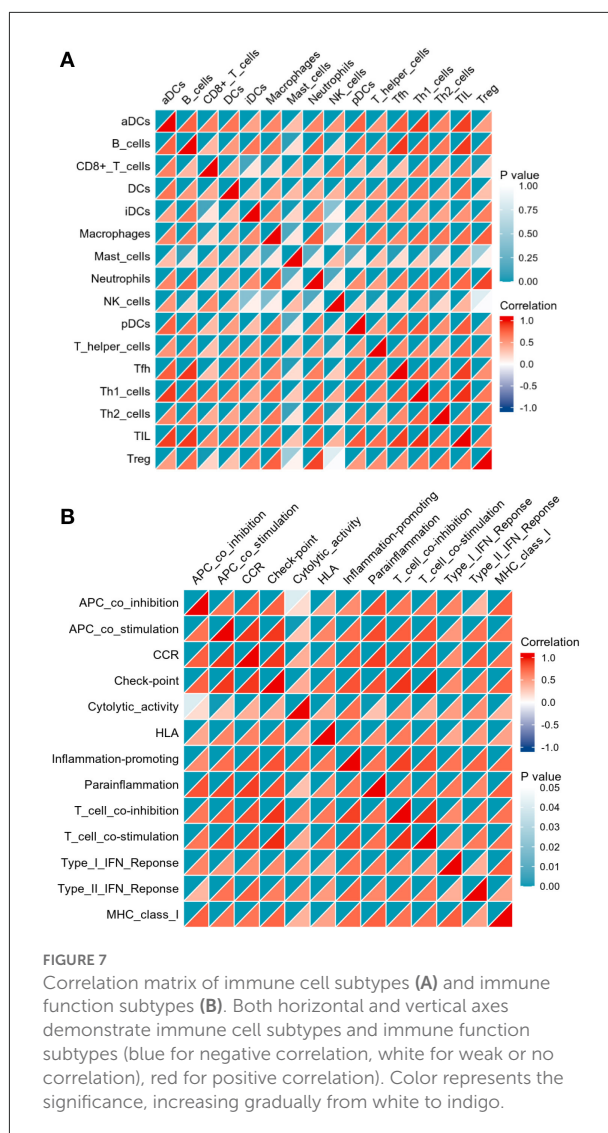
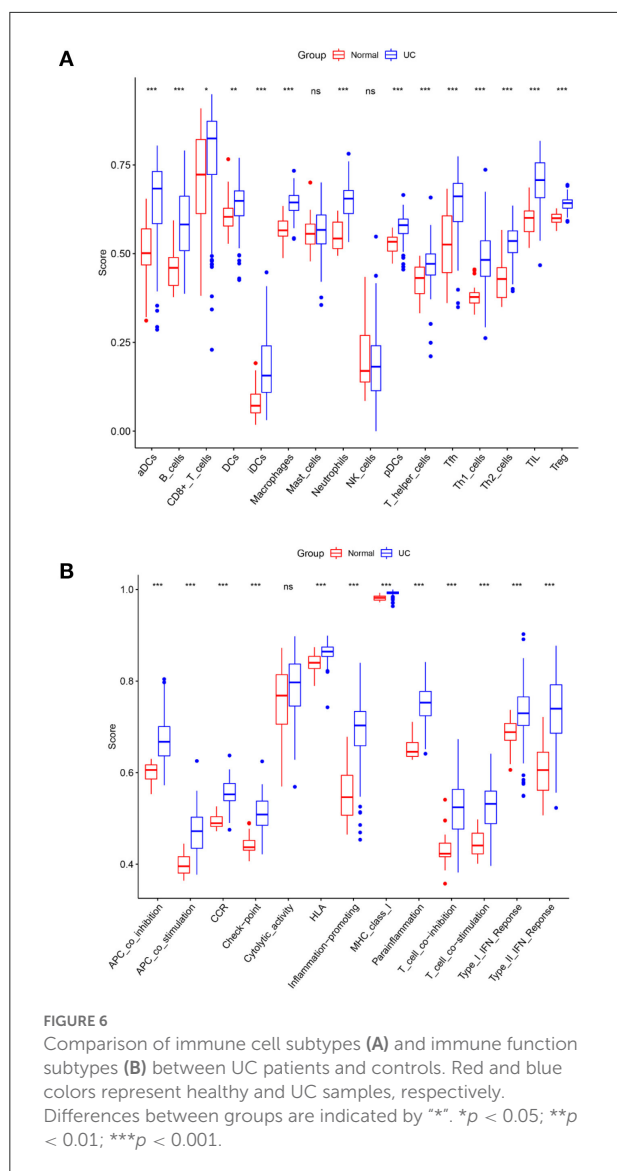
TABLE 3 Results of immune infiltration analysis.

Immune infiltrates	Control		UC		P-value
	Median, interquartile range	Mean \pm standard deviation	Median, interquartile range	Mean \pm standard deviation	
aDCs	0.501 [0.468, 0.570]	0.495 \pm 0.103	0.683 [0.584, 0.731]	0.641 \pm 0.120	<0.05
B cells	0.46 [0.410, 0.489]	0.459 \pm 0.058	0.582 [0.508, 0.662]	0.584 \pm 0.097	<0.05
CD8 ⁺ T cells	0.723 [0.613, 0.822]	0.716 \pm 0.136	0.825 [0.723, 0.873]	0.775 \pm 0.147	<0.05
DCs	0.603 [0.578, 0.628]	0.608 \pm 0.053	0.649 [0.607, 0.677]	0.636 \pm 0.063	<0.05
iDCs	0.072 [0.052, 0.104]	0.085 \pm 0.050	0.157 [0.109, 0.240]	0.180 \pm 0.091	<0.05
Macrophages	0.566 [0.549, 0.592]	0.571 \pm 0.034	0.644 [0.622, 0.664]	0.643 \pm 0.037	<0.05
Mast cells	0.556 [0.527, 0.583]	0.558 \pm 0.049	0.567 [0.527, 0.609]	0.562 \pm 0.067	>0.05
Neutrophils	0.543 [0.514, 0.589]	0.553 \pm 0.043	0.655 [0.613, 0.678]	0.646 \pm 0.050	<0.05
NK cells	0.17 [0.138, 0.270]	0.206 \pm 0.103	0.182 [0.114, 0.241]	0.185 \pm 0.101	>0.05
pDCs	0.534 [0.507, 0.546]	0.531 \pm 0.027	0.580 [0.557, 0.597]	0.575 \pm 0.039	<0.05
T helper cells	0.431 [0.387, 0.462]	0.424 \pm 0.048	0.471 [0.440, 0.499]	0.467 \pm 0.058	<0.05
Tfh	0.526 [0.447, 0.606]	0.523 \pm 0.104	0.661 [0.590, 0.698]	0.641 \pm 0.084	<0.05
Th1 cells	0.378 [0.361, 0.390]	0.381 \pm 0.036	0.482 [0.436, 0.536]	0.479 \pm 0.083	<0.05
Th2 cells	0.429 [0.376, 0.460]	0.426 \pm 0.057	0.535 [0.503, 0.564]	0.528 \pm 0.053	<0.05
TIL	0.601 [0.562, 0.620]	0.600 \pm 0.045	0.707 [0.658, 0.756]	0.701 \pm 0.068	<0.05
Treg	0.600 [0.589, 0.611]	0.600 \pm 0.015	0.642 [0.628, 0.652]	0.641 \pm 0.021	<0.05
APC co-inhibition	0.606 [0.586, 0.617]	0.601 \pm 0.023	0.667 [0.637, 0.701]	0.671 \pm 0.051	<0.05
APC co-stimulation	0.396 [0.381, 0.416]	0.396 \pm 0.021	0.472 [0.435, 0.503]	0.470 \pm 0.045	<0.05
CCR	0.490 [0.482, 0.504]	0.492 \pm 0.014	0.553 [0.539, 0.576]	0.554 \pm 0.029	<0.05
Check-point	0.437 [0.430, 0.452]	0.444 \pm 0.022	0.509 [0.485, 0.538]	0.509 \pm 0.038	<0.05
Cytolytic activity	0.768 [0.706, 0.814]	0.755 \pm 0.076	0.797 [0.746, 0.837]	0.787 \pm 0.069	>0.05
HLA	0.840 [0.827, 0.854]	0.839 \pm 0.020	0.864 [0.854, 0.874]	0.863 \pm 0.020	<0.05
Inflammation-promoting	0.546 [0.507, 0.594]	0.552 \pm 0.062	0.703 [0.659, 0.733]	0.687 \pm 0.075	<0.05
MHC class I	0.982 [0.977, 0.986]	0.981 \pm 0.006	0.993 [0.991, 0.995]	0.992 \pm 0.006	<0.05
Parainflammation	0.645 [0.635, 0.666]	0.651 \pm 0.022	0.753 [0.724, 0.777]	0.748 \pm 0.041	<0.05
T cell co-inhibition	0.423 [0.416, 0.446]	0.434 \pm 0.039	0.524 [0.477, 0.563]	0.518 \pm 0.066	<0.05
T cell co-stimulation	0.441 [0.423, 0.468]	0.445 \pm 0.031	0.532 [0.489, 0.560]	0.526 \pm 0.052	<0.05
Type I IFN response	0.689 [0.671, 0.707]	0.682 \pm 0.038	0.730 [0.703, 0.766]	0.729 \pm 0.064	<0.05
Type II IFN response	0.606 [0.561, 0.644]	0.608 \pm 0.063	0.740 [0.686, 0.792]	0.734 \pm 0.080	<0.05

and is characterized by an abnormal mucosal immune response and a persistent inflammatory infiltrate. Pathological mechanisms that have not been fully elucidated have limited the development of early diagnosis and treatment of UC. Currently, integrated microarray-based bioinformatics analysis has been widely used to identify potential novel biomarkers for various diseases, which is important for specifying molecular markers and diagnosing UC early in the disease process. In this study, DEGs were identified to be mainly enriched in immune response-related pathways. Immune infiltration analysis suggested significant differences in immune cell and immune function types between UC patients and controls. Correlation analysis of DEGs and immune microenvironment was performed, and five immune-related genes (HMGS2, CLDN8, AQP8, DEFB4A, MMP3) were identified and were

subsequently used to construct a diagnostic model of UC. Encouragingly, this model showed good diagnostic performance for UC, and data from the other two datasets further validated the accuracy of this model for UC diagnosis.

With data from GEO, 3111 DEGs (1,608 up-regulated genes and 1,503 down-regulated genes) were obtained in UC by comprehensive analysis of the microarray dataset (GSE87473). Further GO analysis indicated that these DEGs were significantly enriched in the leukocyte migration, leukocyte cell-cell adhesion, T cell activation, cytokine-mediated signaling pathway, leukocyte chemotaxis, cell chemotaxis, granulocyte migration, regulation of T cell activation, lymphocyte differentiation, mononuclear cell differentiation, immune receptor activity, chemokine activity, chemokine receptor binding, CXCR chemokine receptor binding, collagen binding,



cytokine activity, cytokine binding, extracellular matrix structural constituent, and G protein-coupled receptor binding. For KEGG analysis, genes were significantly enriched in immune-related pathways such as TNF signaling pathway, Osteoclast differentiation, IL-17 signaling pathway, Th17 cell differentiation, and NF-kappa B signaling pathway. For DO analysis, DEGs were significantly enriched in lung disease, intestinal disease, mouth disease, primary bacterial infectious disease, arteriosclerosis, IBD, arteriosclerotic cardiovascular disease, female reproductive organ cancer, and tooth disease. In summary, results of the bioinformatics analyses suggested that these DEGs were closely related to immune cell infiltration in UC. These findings further validate the key role of immune abnormalities in the pathological progression of UC (22).

Inspired by the results of functional analysis of DEGs, immune infiltration analysis was further performed. The results revealed that there were significant differences in aDCs, B cells, CD8⁺ T cells, DCs, iDCs, Macrophages, Neutrophils, pDCs, T helper cells, Tfh, Th1 cells, Th2 cells, TIL and Treg between colon tissue in UC patients and the healthy group. In fact, most of the above immune cells have been reported to be in an abnormal state in UC by previous studies (23, 24). Therefore, these identified immune cells may be involved in the development and progression of UC. Top genes with the most significant expression differences between healthy and UC samples were considered as potential diagnostic genes for UC. After feature selection, HMGCS2, CLDN8, AQP8, DEFB4A, and MMP3 were identified as having strong association with immune infiltration, and thus they may be key genes that

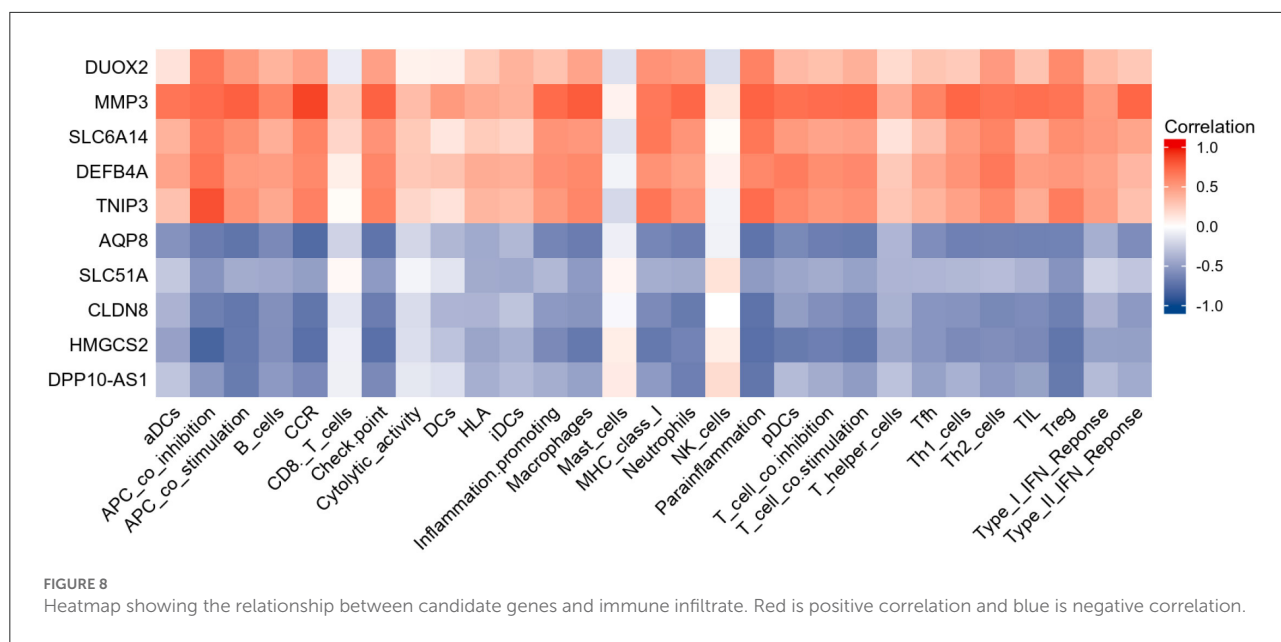


TABLE 4 Details of genes in diagnostic model.

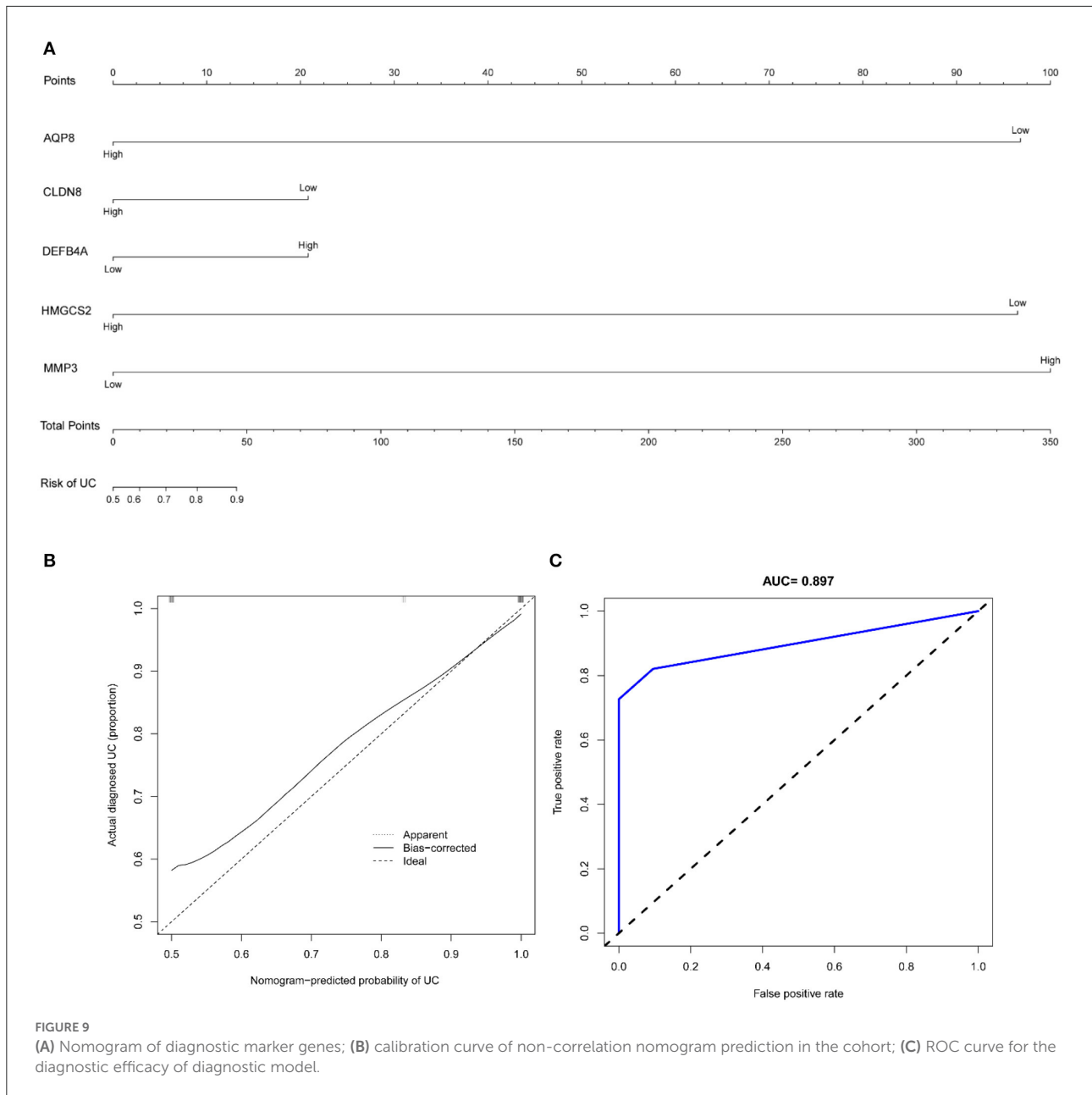
Genes	Sig	Log2FC	Average expression
AQP8	Down	-4.626352947	8.071990976
CLDN8	Down	-3.601376685	6.380647165
HMGCS2	Down	-3.559635503	7.584732929
DEFB4A	Up	4.418274083	8.68987552
MMP3	Up	4.763273249	9.167130008

identify the immune features of UC and are involved in UC immune regulation.

Inspired by the results of functional analysis of DEGs and immune infiltration analysis, HMGCS2, CLDN8, AQP8, DEFB4A, and MMP3 were applied to construct a diagnostic model for UC. The nomogram showed the well diagnostic efficacy of the model constructed with these predicted diagnostic genes for UC. The calibration curve for the uncorrelated nomogram prediction showed that the performance of the column line plot was close to the ideal model, suggesting that the predictive value of the model is credible. Similarly, ROC-AUC of the risk score was 0.897, suggesting a high diagnostic efficiency of the diagnostic marker gene model. To go step further validation, data from GSE75214 and GSE92415 were applied to assesses the prediction accuracy of the model. The results revealed that ROC-AUC of the risk score were 0.871 in GSE75214 and 0.908 in GSE92415, respectively, indicating excellent discrimination of the diagnostic model.

A total of five immune-related genes were included in the diagnostic model. The protein encoded by HMGCS2 belongs to the HMG-CoA synthase family of mitochondrial enzymes

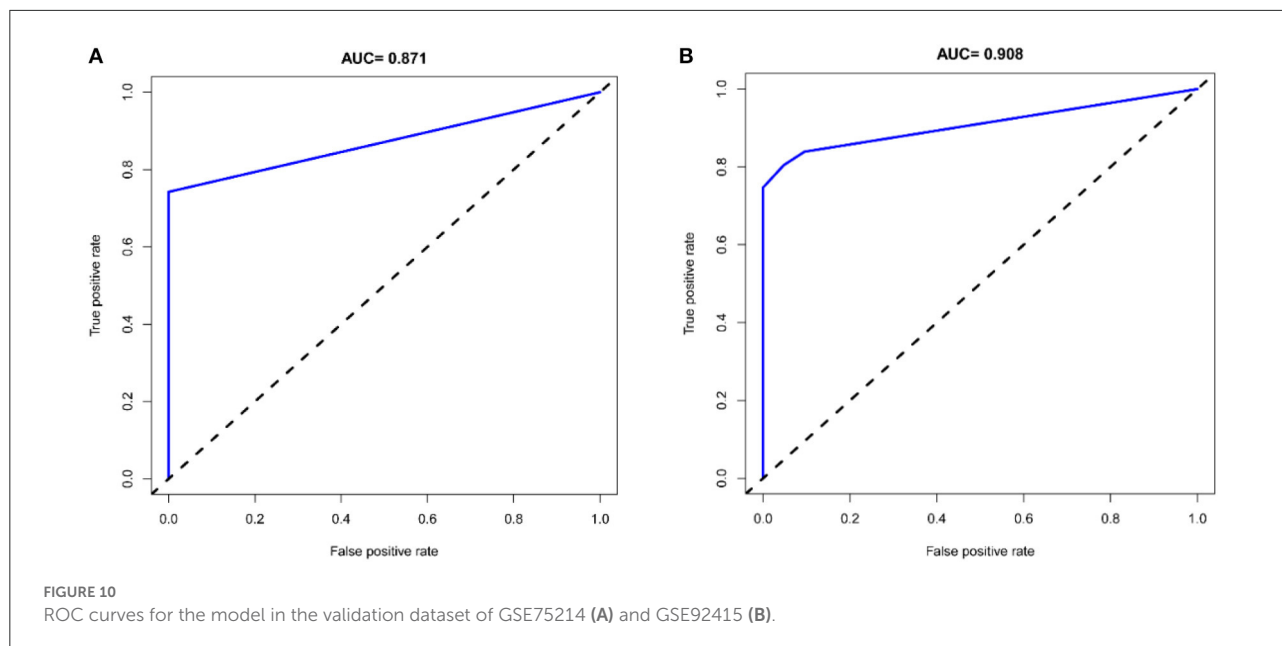
that catalyze the first reaction of ketogenesis, which is a crucial alternative metabolic pathway and is involved in the regulation of the body's immune function (25). Restoration of ketogenesis enhances immune cell effects (26) and attenuates the activation of pro-inflammatory macrophages (27). The protein levels of HMGCS2 in the intestinal epithelium of UC patients were reported to be sharply decreased compared to healthy samples (28). Increased ketogenesis may help to counteract intestinal inflammation, and conversely, its suppression may exacerbate intestinal pathology (28). The CLDN8 gene encodes a member protein of the claudin family. Claudins are integral membrane proteins and components of tight junction chains that play a key role in maintaining the integrity of the intestinal mucosal barrier. CLDN8 was reported to be significantly downregulated in the biological colon of IBD patients, and similar results were observed in colitis mice (29). AQP8 encodes an epithelial water transport protein specifically expressed in colonic absorptive cells, and it was found to be significantly downregulated in UC patients compared to healthy controls (30, 31). In addition, AQP8 was observed to promote H₂O₂ diffusion in experimental mouse models, which suggested its balance and regulatory effects on antioxidant pathways (32). DEFB4A encodes defensin, beta 4, an antibiotic peptide locally regulated by inflammation, which has been shown to be involved in the pathological process of IBD. Results of high-throughput sequencing suggested that the composition of the microbiota differs significantly between UC and non-IBD. Alterations in the microbiota can affect antimicrobial peptide expression, which in turn is involved in the progression of IBD (33). Various studies have highlighted the involvement of specific matrix metalloprotease in IBD: MMP3 transcript or protein levels are upregulated in the mucosa of inflammatory IBD or in the



serum of IBD patients, and MMP protein hydrolytic activity is increased in epithelial-derived cells of inflammatory IBD (34–36). It was found that MMP3 expression was significantly upregulated in inflammatory colonic segments of IBD patients compared to non-inflammatory regions (37). In addition, MMP3 serum assay possesses a suggestive role for early response to infliximab treatment of UC (38). In summary, these key genes are all involved in the development of UC, and their inclusion in our diagnostic model of UC is reasonable.

Although this study applied a relatively large sample size to characterize the immune microenvironment and

construct a diagnostic model for UC by integrating the GEO dataset, limitations should be acknowledged. First, this study explored the infiltration of immune cells by ssGSEA and found that immune cells play an important role in the pathological progression of UC. Therefore, it is crucial to validate our findings by flow cytometry. Second, although the present model may serve as a valid predictive tool for UC diagnosis, the true predictive value of the model should be prospectively validated in future independent and multicentered-studies with larger sample sizes.



Conclusion

In conclusion, this study identified the immune infiltrate characteristics and five immune-related genes for UC. The model based on the immune-related genes facilitates the early diagnosis of UC and provides a basis for the evaluation of the prognosis of UC.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author/s.

Author contributions

JH drafted the manuscript. JZ, FW, and BZ helped with implementation of this work. XT contributed to the methodology, review, and editing of the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (No. 81830118), China Academy of Chinese Medical Sciences Innovation Fund (No. CI 2021A01012), China Academy of Chinese Medical Sciences

Excellent Young Talent Cultivation Fund (No. ZZ 15-YQ-002), and Administration of Traditional Chinese Medicine Digestive Refractory Disease Inheritance and Innovation Team Project (No. ZYYCXTD-C-C202010).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2022.1003002/full#supplementary-material>

References

- Huang J, Zhang J, Ma J, Ma J, Liu J, Wang FJ, et al. Inhibiting ferroptosis: a novel approach for ulcerative colitis therapeutics. *Oxid Med Cell Longev*. (2022) 2022:9678625. doi: 10.1155/2022/9678625
- Talley NJ, Abreu MT, Achkar JP, Bernstein CN, Dubinsky MC, Hanauer SB, et al. An evidence-based systematic review on medical therapies for inflammatory bowel disease. *Am J Gastroenterol*. (2011) 106:S2–25. doi: 10.1038/ajg.2011.58
- Cohen RD, Yu AP, Wu EQ, Xie J, Mulani PM, Chao J. Systematic review: the costs of ulcerative colitis in Western countries. *Aliment Pharmacol Ther*. (2010) 31:693–707. doi: 10.1111/j.1365-2036.2010.04234.x
- Agrawal M, Spencer EA, Colombel JF, Ungaro RC. Approach to the management of recently diagnosed inflammatory bowel disease patients: a user's guide for adult and pediatric gastroenterologists. *Gastroenterology*. (2021) 161:47–65. doi: 10.1053/j.gastro.2021.04.063
- Du L, Ha C. Epidemiology and pathogenesis of ulcerative colitis. *Gastroenterol Clin North Am*. (2020) 49:643–54. doi: 10.1016/j.gtc.2020.07.005
- Bai X, Liu W, Chen H, Zuo T, Wu X. Immune cell landscaping reveals distinct immune signatures of inflammatory bowel disease. *Front Immunol*. (2022) 13:861790. doi: 10.3389/fimmu.2022.861790
- Kaluzna A, Olczyk P, Komosińska-Vashev K. The role of innate and adaptive immune cells in the pathogenesis and development of the inflammatory response in ulcerative colitis. *J Clin Med*. (2022) 11:400. doi: 10.3390/jcm11020400
- Wang Y, Zhang H, He H, Ai K, Yu W, Xiao X, et al. LRCH1 suppresses migration of CD4+ T cells and refers to disease activity in ulcerative colitis. *Int J Med Sci*. (2020) 17:599–608. doi: 10.7150/ijms.39106
- Yang W, Liu H, Xu L, Yu T, Zhao X, Yao S, et al. GPR120 inhibits colitis through regulation of CD4+ T cell interleukin 10 production. *Gastroenterology*. (2022) 162:150–65. doi: 10.1053/j.gastro.2021.09.018
- Yang ZJ, Wang BY, Wang TT, Wang FF, Guo YX, Hua RX, et al. Functions of Dendritic Cells and Its Association with Intestinal Diseases. *Cells*. (2021) 10:583. doi: 10.3390/cells10030583
- Yan J, Pandey SP, Barnes BJ, Turner JR, Abraham C, T. Cell-intrinsic IRF5 regulates T cell signaling, migration, and differentiation and promotes intestinal inflammation. *Cell Rep*. (2020) 31:107820. doi: 10.1016/j.celrep.2020.107820
- Stakenborg M, Verstockt B, Meroni E, Govers G, De Simone V, Verstockt S, et al. Neutrophilic HGF-MET signalling exacerbates intestinal inflammation. *J Crohns Colitis*. (2020) 14:1748–58. doi: 10.1093/ecco-jcc/jjaa121
- Hone Lopez S, Kats-Ugurlu G, Renken RJ, Buikema HJ, de Groot MR, Visschedijk MC, et al. Immune checkpoint inhibitor treatment induces colitis with heavy infiltration of CD8 + T cells and an infiltration pattern that resembles ulcerative colitis. *Virchows Arch*. (2021) 479:1119–29. doi: 10.1007/s00428-021-03170-x
- Hegazy AN, West NR, Stubbington MJT, Wendt E, Suijker KIM, Datsi A, et al. Circulating and tissue-resident CD4+ T cells with reactivity to intestinal microbiota are abundant in healthy individuals and function is altered during inflammation. *Gastroenterology*. (2017) 153:1320–37. doi: 10.1053/j.gastro.2017.07.047
- Hart AL, Al-Hassi HO, Rigby RJ, Bell SJ, Emmanuel AV, Knight SC, et al. Characteristics of intestinal dendritic cells in inflammatory bowel diseases. *Gastroenterology*. (2005) 129:50–65. doi: 10.1053/j.gastro.2005.05.013
- Lu Z, Su H. Employing gene chip technology for monitoring and assessing soil heavy metal pollution. *Environ Monit Assess*. (2021) 194:2. doi: 10.1007/s10661-021-09650-6
- Yu Z, Ma X, Zhang W, Chang X, An L, Niu M, et al. Microarray data mining and preliminary bioinformatics analysis of hepatitis D virus-associated hepatocellular carcinoma. *Biomed Res Int*. (2021) 2021:1093702. doi: 10.1155/2021/1093702
- Zhao E, Zhou C, Chen S, A. signature of 14 immune-related gene pairs predicts overall survival in gastric cancer. *Clin Transl Oncol*. (2021) 23:265–74. doi: 10.1007/s12094-020-02414-7
- Zhao E, Xie H, Zhang Y. Identification of differentially expressed genes associated with idiopathic pulmonary arterial hypertension by integrated bioinformatics approaches. *J Comput Biol*. (2021) 28:79–88. doi: 10.1089/cmb.2019.0433
- Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, et al. NCBI GEO: mining tens of millions of expression profiles—database and tools update. *Nucleic Acids Res*. (2007) 35:D760–5. doi: 10.1093/nar/gkl887
- Omlor W, Wahl AS, Sipilä P, Lütcke H, Laurency B, Chen JW, et al. Context-dependent limb movement encoding in neuronal populations of motor cortex. *Nat Commun*. (2019) 10:4812. doi: 10.1038/s41467-019-12670-z
- van Unen V, Ouboter LF, Li N, Schreurs M, Abdelaal T, Kooy-Winkelaar Y, et al. Identification of a disease-associated network of intestinal immune cells in treatment-naïve inflammatory bowel disease. *Front Immunol*. (2022) 13:893803. doi: 10.3389/fimmu.2022.893803
- Luo Y, Liu S, Li H, Hou J, Lin W, Xu Z, et al. Mass cytometry and single-cell transcriptome analyses reveal the immune cell characteristics of ulcerative colitis. *Front Mol Biosci*. (2022) 9:859645. doi: 10.3389/fmolb.2022.859645
- Mitsialis V, Wall S, Liu P, Ordovas-Montanes J, Parmet T, Vukovic M, et al. Single-cell analyses of colon and blood reveal distinct immune cell signatures of ulcerative colitis and crohn's disease. *Gastroenterology*. (2020) 159:591–608. doi: 10.1053/j.gastro.2020.04.074
- Goldberg EL, Shchukina I, Asher JL, Sidorov S, Artyomov MN, Dixit VD. Ketogenesis activates metabolically protective $\gamma\delta$ T cells in visceral adipose tissue. *Nat Metab*. (2020) 2:50–61. doi: 10.1038/s42255-019-0160-6
- Wei R, Zhou Y, Li C, Rychahou P, Zhang S, Titlow WB, et al. Ketogenesis attenuates KLF5-dependent production of CXCL12 to overcome the immunosuppressive tumor microenvironment in colorectal cancer. *Cancer Res*. (2022) 82:1575–88. doi: 10.1158/0008-5472.CAN-21-2778
- Zhang L, Shi J, Du D, Niu N, Liu S, Yang X, et al. Ketogenesis acts as an endogenous protective programme to restrain inflammatory macrophage activation during acute pancreatitis. *EBioMedicine*. (2022) 78:103959. doi: 10.1016/j.ebiom.2022.103959
- Kim JT, Napier DL, Kim J, Li C, Lee EY, Weiss HL, et al. Ketogenesis alleviates TNF α -induced apoptosis and inflammatory responses in intestinal cells. *Free Radic Biol Med*. (2021) 172:90–100. doi: 10.1016/j.freeradbiomed.2021.05.032
- Wang H, Chao K, Ng SC, Bai AH, Yu Q, Yu J, et al. Pro-inflammatory miR-223 mediates the cross-talk between the IL23 pathway and the intestinal barrier in inflammatory bowel disease. *Genome Biol*. (2016) 17:58. doi: 10.1186/s13059-016-0901-8
- Dotti I, Mora-Buch R, Ferrer-Picón E, Planell N, Jung P, Masamunt MC, et al. Alterations in the epithelial stem cell compartment could contribute to permanent changes in the mucosa of patients with ulcerative colitis. *Gut*. (2017) 66:2069–79. doi: 10.1136/gutjnl-2016-312609
- Min M, Peng LH, Sun G, Guo MZ, Qiu ZW, Yang YS. Aquaporin 8 expression is reduced and regulated by microRNAs in patients with ulcerative colitis. *Chin Med J*. (2013) 126:1532–7. doi: 10.3760/cma.j.issn.0366-6999.20122989
- Te Velde AA, Pronk I, de Kort F, Stokkers PC. Glutathione peroxidase 2 and aquaporin 8 as new markers for colonic inflammation in experimental colitis and inflammatory bowel diseases: an important role for H2O2? *Eur J Gastroenterol Hepatol*. (2008) 20:555–60. doi: 10.1097/MEG.0b013e3282f45751
- Jalanka J, Cheng J, Hiippala K, Ritari J, Salojärvi J, Ruuska T, et al. Colonic mucosal microbiota and association of bacterial taxa with the expression of host antimicrobial peptides in pediatric ulcerative colitis. *Int J Mol Sci*. (2020) 21:6044. doi: 10.3390/ijms21176044
- Meijer MJ, Mieremet-Ooms MA, van der Zon AM, van Duijn W, van Hogezaand RA, Sier CF, et al. Increased mucosal matrix metalloproteinase-1, -2, -3 and -9 activity in patients with inflammatory bowel disease and the relation with Crohn's disease phenotype. *Dig Liver Dis*. (2007) 39:733–9. doi: 10.1016/j.dld.2007.05.010
- Hu J, Van den Steen PE, Sang QX, Opendakker G. Matrix metalloproteinase inhibitors as therapy for inflammatory and vascular diseases. *Nat Rev Drug Discov*. (2007) 6:480–98. doi: 10.1038/nrd2308
- Lakatos G, Hritz I, Varga MZ, Juhász M, Miheller P, Cierny G, et al. The impact of matrix metalloproteinases and their tissue inhibitors in inflammatory bowel diseases. *Dig Dis*. (2012) 30:289–95. doi: 10.1159/000336995
- Medina C, Santana A, Paz-Cabrera MC, Parra-Blanco A, Nicolás D, Gimeno-García AZ, et al. Increased activity and expression of gelatinases in ischemic colitis. *Dig Dis Sci*. (2006) 51:2393–9. doi: 10.1007/s10620-006-9255-5
- Barberio B, D'Inca R, Facchin S, Dalla Gasperina M, Fohom Tagne CA, Cardin R, et al. Matrix metalloproteinase 3 predicts therapeutic response in inflammatory bowel disease patients treated with infliximab. *Inflamm Bowel Dis*. (2020) 26:756–63. doi: 10.1093/ibd/izz195