Original Article

Bioinformatics analysis-based mining of potential markers for inflammatory bowel disease and their immune relevance

Yuwen Zhu[1], Yanbin Pan[1], Lichao Fan[1], Meng Zou[1], Yingjie Liu[1], Jiayi Hu[1], Shijun Xia[1], Yue Li[1], Ruijie Dai[2], Wenjiang Wu[1]

[1]Department of Anorectal Surgery, Shenzhen Hospital of Guangzhou University of Chinese Medicine, Shenzhen, China; [2]Department of Anorectal Surgery, Shenzhen Traditional Chinese Medicine Anorectal Hospital, Shenzhen, China

Contributions: (I) Conception and design: Y Zhu, Y Pan; (II) Administrative support: W Wu; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: Y Zhu, Y Pan, L Fan; (V) Data analysis and interpretation: Y Zhu, L Fan, M Zou, Y Liu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Wenjiang Wu, MMed. Department of Anorectal Surgery, Shenzhen Hospital of Guangzhou University of Chinese Medicine, No. 6001, Beihuan Avenue, Futian District, Shenzhen 518000, China. Email: 1053660645@qq.com.

Background: The incidence of inflammatory bowel disease (IBD) is increasing every year and is characterized by a prolonged course, frequent relapses, difficulty in curing, and a lack of more efficacious therapeutic biomarkers. The aim of this study was to find key core genes as therapeutic biomarkers for IBD.

Methods: GSE75214 in Gene Expression Omnibus (GEO) was used as the experimental set. The genes in the top 25% of standard deviation of all samples in the experimental set were subjected to systematic weighted gene co-expression network analysis (WGCNA) to find candidate genes. Then, least absolute shrinkage and selection operator (LASSO) logistic regression was used to further screen the central genes. Finally, the validity of hub genes was verified on GEO dataset GSE179285 using "BiocManager" R package.

Results: Twelve well-preserved modules were identified in the experimental set using the WGCNA method. Among them, five modules significantly associated with IBD were screened as clinically significant modules, and four candidate genes were screened from these five modules. Then *TIMP1*, *GUCA2B*, and *HIF1A* were screened as hub genes. These hub genes successfully distinguished tumor samples from healthy tissues by artificial neural network algorithm in an independent test set with an area under the working characteristic curve of 0.946 for the subjects.

Conclusions: IBD differentially expressed gene (DEGs) are involved in immunoregulatory processes. *TIMP1*, *GUCA2B*, and *HIF1A*, as core genes of IBD, have the potential to be therapeutic targets for patients with IBD, and our findings may provide a new outlook on the future treatment of IBD.

Keywords: Inflammatory bowel disease (IBD); bioinformatics analysis; hub gene; potential marker; immune

Submitted Feb 21, 2024. Accepted for publication Jul 07, 2024. Published online Aug 27, 2024.
doi: 10.21037/tcr-24-274
View this article at: https://dx.doi.org/10.21037/tcr-24-274

Introduction

Inflammatory bowel disease (IBD) is a chronic autoimmune disease affecting the gut in which the body develops an immune response to the intestinal microbiota or antigens in the intestinal tract, leading to inflammation and disease in genetically susceptible individuals (1). IBD affects more than 3.5 million people and its incidence is on the rise worldwide (2). The most prevalent forms are Crohn's disease (CD) and ulcerative colitis (UC), which are characterized by recurrent, chronic inflammation of the gastrointestinal tract or colon that persists over time (3). The etiology and pathogenesis of the disease are still not fully defined, but may be related to genetic susceptibility, imbalance of intestinal microbial homeostasis, impaired intestinal mucosal barrier function, intestinal immunomodulatory

disorders, and stimulation of external environmental factors (4). Identified risk genes in IBD have been shown to be involved in the maintenance of a correct immune response (5), and when these risk genes are mutated, the mechanisms of the body's immune response may be compromised and leads to weakened intestinal mucosal barrier function, diminished intestinal autoantimicrobial activity, and impaired autophagy and antigen recognition (6). It has been shown that the colon of patients with UC is often associated with significant colonic epithelial cell death and destruction, as well as increased intestinal permeability, leading to damage to the intestinal mechanical barrier (7). CD, on the other hand, exhibits predominantly damage to the immune barrier, the severity of which correlates closely with the downregulation of immunoglobulin A (IgA) expression (8). With the development and application of immune-targeted drugs and the updating of therapeutic targets, the clinical treatment of IBD has improved significantly (9), but the clinical remission rate is still very low, so the development of new targets for the treatment of IBD has become an urgent need nowadays. And clinical data show that the use of traditional drug treatment has not reached the standard of clinical cure. About 47% of patients still have long-term chronic inflammatory reactions after treatment, and 40% of patients still need surgery eventually. This seriously affects the quality of life of patients, causing heavy social impact and medical burden (10). Therefore, the weighted gene co-expression network analysis (WGCNA) method and least absolute shrinkage and selection operator (LASSO) regression were used to construct a co-expression network to identify central genes in IBD, which can effectively distinguish IBD samples from normal tissues. Unlike polynomial regression, LASSO regression controls the complexity of the model through the L1 regularization term, which enables variable selection. Unlike ridge regression, LASSO regression allows for the selection of the most important variables and compresses the others to 0. Thus, LASSO regression combines the advantages of polynomial and ridge regression to some extent. These findings may provide potential diagnostic and therapeutic targets for future IBD research and clinical intervention. We present this article in accordance with the TRIPOD reporting checklist (available at https://tcr.amegroups.com/article/view/10.21037/tcr-24-274/rc).

Highlight box

Key findings

- Inflammatory bowel disease (IBD) differentially expressed genes (DEGs) are involved in immunoregulatory processes. *TIMP1*, *GUCA2B*, and *HIF1A*, as core genes of IBD, have the potential to be therapeutic targets for patients with IBD, and our findings may provide a new outlook on the future treatment of IBD.

What is known and what is new?

- IBD is a chronic autoimmune disease affecting the gut in which the body develops an immune response to the intestinal microbiota or antigens in the intestinal tract.
- *TIMP1*, *GUCA2B*, and *HIF1A*, as core genes of IBD, have the potential to be therapeutic targets for patients with IBD.

What is the implication, and what should change now?

- Our results may provide a new prospect for the future treatment of IBD, and the development of new drugs and therapeutic technologies in the future. We can do clinical trials to further validate the findings.

Methods

Data collection

Gene Expression Omnibus (GEO), a gene expression database created and maintained by the National Center for Biotechnology Information (NCBI). The GEO database was used to store gene expression datasets and platform records. IBD-related dataset GSE75214 (https://ftp.ncbi.nlm.nih.gov/geo/series/GSE75nnn/GSE75214/matrix/) was downloaded from the GEO database for analysis. GSE75214 (11) based on the GPL6244 platform [(HuGene-1_0-st) Affymetrix Human Gene 1.0 ST Array] included gene expression profiles of 172 cases of diseased colon tissues and 22 cases of normal colon tissues. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Screening for differential genes

After filtering and normalization of the two groups of samples, the R software Limma package was used to screen the two groups of samples for differentially expressed genes (DEGs). Adjusted P value <0.05 and $|\log_2$fold change ($\log_2$FC)| >1 were used as the screening conditions, where $\log_2$FC >1 represented up-regulation of gene expression, and $\log_2$FC <–1 represented down-regulation of gene expression, and volcano and heat maps were plotted by R.

Functional enrichment and pathway analysis of differential genes

In order to find out the main functions of the differential genes and their pathways, the "clusterprofile" package was used to perform the Gene Ontology (GO) function enrichment analysis and the Kyoto Encyclopedia of Genes and Genomes (KEGG) function enrichment analysis, and the "clusterprofile" package was used to perform the GO function enrichment analysis. The differences were statistically significant at P<0.05. The "Limma" package was used to analyze the functional enrichment of GO and KEGG pathways. The results of typing and immunocyte infiltration were integrated using the "Limma" R language package. Pathway difference analysis of DEGs was performed using "clusterProfiler", "enrichplot", "org.Hs.eg.db", and "DOSE", and the top 5 pathways were selected to draw line graphs (12).

Typing-WGCNA

WGCNA can be used to study correlations in biochip samples and to find appropriate biomarkers or therapeutic targets in different biological contexts (13). The core modules and core genes of IBD genes were screened using the "WGCNA" R language package. The top 25% of the most fluctuating genes were selected for WGCNA, and a neighbor-joining matrix was constructed to describe the strength of correlation between nodes. A topological overlap matrix is created from the neighbor matrix in order to quantify the similarity between nodes. The process of hierarchical clustering was used to find modules that contained at least 50 genes. From the WGCNA plates, the plates with the lowest adjust P value were chosen for analysis. Each plate's core genes were filtered using the criteria "geneSig Filter =0.5, moduleSig Filter =0.8". Gene significance (GS) and module significance (MS) were computed by combining modules with phenotypic data. This allowed for the analysis of the association between models and modules as well as the measurement of the importance of genes and clinics in the development process. For each gene, module membership (MM) was also computed in order to assess GS inside the module.

Wayne diagram and LASSO regression

As potential center genes, the genes with the strongest inter-module communication were chosen. Genes that are important to biology typically have high absolute GS values. The criteria for screening candidate genes were (absolute GS value >0.50; absolute MM value >0.80). The candidate center genes were then intersected with DEGs using the R package "Venn".

Next, we utilize LASSO for analysis. LASSO, which obtains a more refined model by constructing a penalty function that makes it compress some coefficients while setting some coefficients to zero. Thus, it retains the advantages of subset shrinkage and is a biased estimation for dealing with data with complex covariance. The LASSO algorithm was performed based on the candidate genes using the "glmnet" R software package, and the central genes were obtained from the intersections of the genes obtained by this gene learning technique using a Wayne diagram.

Immune cell infiltration

We investigated the differences in infiltration between high and low expression of core genes in 28 immune cells using single sample gene set enrichment analysis (ssGSEA), which applies the gene set variation analysis ("GSVA") and "GSEABase" R software packages to predict the proportions of multiple cell types in the gene expression profiles were visualized using the vioplot package; heatmaps and box plots were drawn using the "pheatmap" and "ggpubr" R packages to illustrate the results. Finally, R packages "tidyverse", "ggplot2", and "ggExtra" were used to derive the relationship between the expression levels of core genes and immune cells. The relationship between the expression level of the core genes and the immune cells was determined. The percentage of immune cells in the IBD immune microenvironment was determined by filtering at "P<0.05" and visualized using the vioplot package.

Database cross-validation

IBD-related dataset GSE179285 (14) was downloaded from the GEO database for validation, GSE179285 (https://ftp.ncbi.nlm.nih.gov/geo/series/GSE179nnn/GSE179285/matrix/) was based on the GPL6480 platform Agilent-014850 human whole genome microarray 4x44K G4112F expression array, which included 361 cases of diseased colon tissues and 58 cases of normal colon tissues.

Key gene validation

For the validation set GSE179285, the expression of key

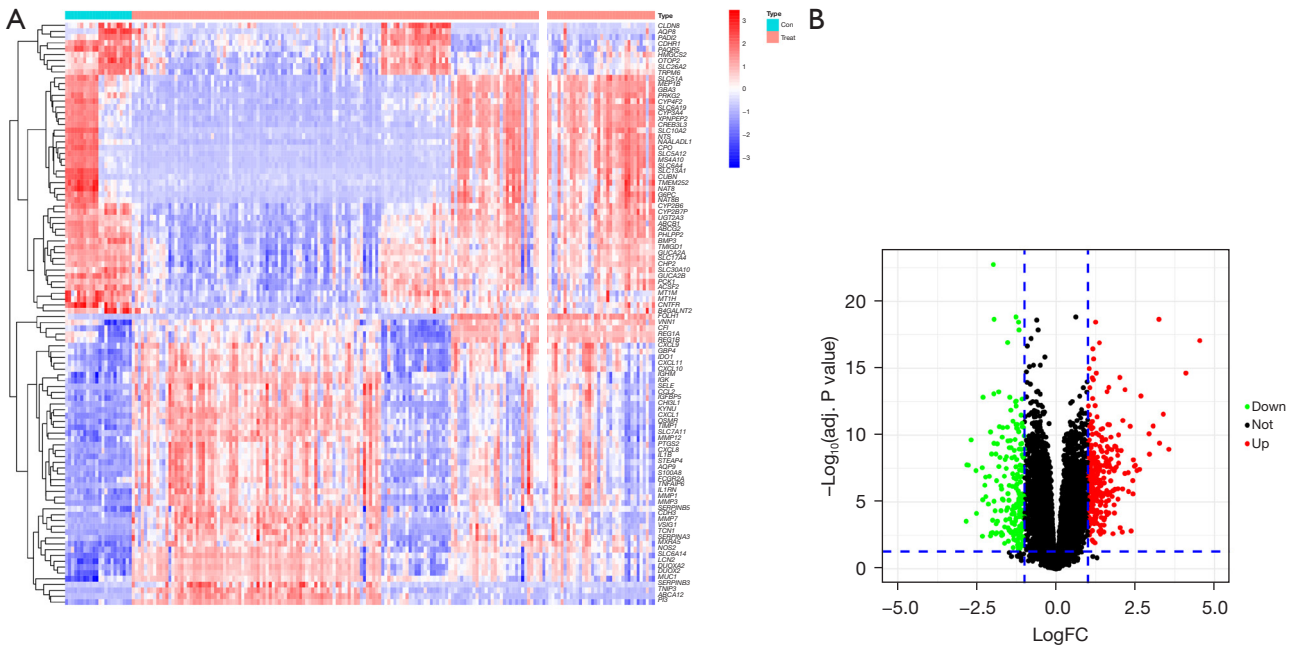

Figure 1 Differential gene expression analysis of IBD patients. (A) Heat map showing up-regulated genes in red and down-regulated genes in blue. (B) Volcano map showing 292 highly expressed genes and 198 lowly expressed genes. Red bars indicate up-regulated genes and green bars indicate down-regulated genes. FC, fold change; adj., adjusted; IBD, inflammatory bowel disease.

genes in positive and normal samples were obtained by probe technology and database; the results were visualized by "BiocManager" R package, and the results were visualized in the form of box plots. Finally, the key genes with statistically significant differences in the two validation sets were used as screening results, and it was concluded that these key genes could be potential markers of IBD.

Statistical analysis

All the bioinformatics analyses and statistical tests were performed with R software 4.2.1. Wilcoxon or Student's *t*-test was utilized for analysis of the difference between groups. The correlation between variables was determined using Pearson correlation test. All statistical P values were two-sided, and P<0.05 was regarded as statistically significant.

Results

Microarray data

The experimental set of IBD microarray GSE75214 contained 172 IBD tissues and 22 normal colon tissues. The validation set IBD microarray GSE179285 contained 223 IBD intestinal epithelial tissues and 31 normal intestinal epithelial tissues.

Screening of DEGs

There were 490 IBD DEGs obtained from the GSE75214 chip by comparative analysis with normal samples. Subsequently, volcano and heat maps were used to show the 490 differential genes that reached the threshold, including 292 significantly highly expressed genes and 198 significantly low expressed genes (*Figure 1*).

GO and KEGG enrichment analysis

To predict the biological functions of DEGs, we performed functional enrichment analysis. GO and KEGG enrichment analyses identified a number of statistically significant and relevant biological activities (*Figure 2*). In total, 926 significant GO terms and 30 significant pathways were enriched (P<0.05). Changes in terms of biological process (BP) were clearly associated with 762 functional categories including leukocyte migration, response to molecules of bacterial origin, response to lipopolysaccharides (LPSs),

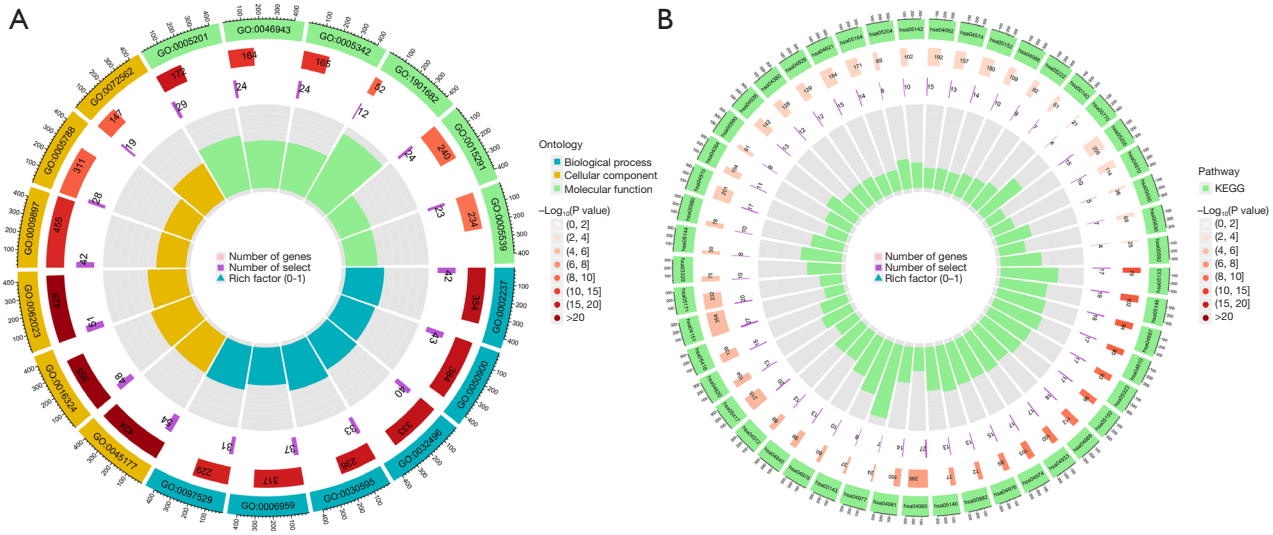

Figure 2 Functional enrichment analysis of IBD differential genes. (A) Nine hundred and twenty-six significant GO terms and 30 significant pathways were enriched. (B) KEGG pathway enrichment showed that DEGs were mainly enriched in 49 signaling pathways. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; IBD, inflammatory bowel disease; DEG, differentially expressed gene.

humoral immune response, and inflammatory response (cytotropism). In addition, there were 46 terms in cellular component (CC) analysis, mainly apical portion of the cell, collagen-containing extracellular matrix, apical plasma membrane, and outer membrane of the plasma membrane. There were also 118 terms in molecular function (MF), such as structural components of extracellular matrix, anion transmembrane transporter activity, and transmembrane transporter activities of carboxylic acids and organic acids. KEGG pathway enrichment showed that the identified targets were enriched mainly in 49 pathways, the including the PI3K-Akt signaling pathway, cytokine-cytokine receptor interactions, coronavirus disease 2019 (COVID-19), and lipids and atherosclerosis.

Results of GSEA enrichment analysis

GSEA enrichment analysis was used in the normal and disease groups, respectively. As shown in *Figure 3*, both groups were enriched for multiple immune function gene sets. The genes in the normal group were mainly enriched for normal immune response-related activities, such as macrophages, induced T-cell differentiation, CD4, CD8, and peripheral blood mononuclear cells (*Figure 3A*). As for the disease group, these genes were enriched in immune responses that may be mediated with inflammation, including CD7, natural killer (NK) cells, and triggering receptor expressed on myeloid cell-1 (TREM-1) (*Figure 3B*). These data suggest that immune response-related pathways play a key role in IBD.

WGCNA identification of immune-related candidate biomarker genes in IBD

We took the disease group dataset using hierarchical clustering and clustered it according to its relative Euclidean distance for gene analysis (*Figure 4*). Performing WGCNA and choosing a scale-free fit index of 0.85, we obtained the optimal soft threshold, i.e., 20, resulting in an average connectivity of around 10 (*Figure 4A*). Four modules were obtained using the dynamic shear tree algorithm (*Figure 4B*). Module-trait relationship analysis showed that the turquoise module was strongly associated with disease genes, and 753 genes in the turquoise module were considered to be IBD-associated genes (*Figure 4B*). Genes in the turquoise module showed high correlation between MM and GS (*Figure 4C*).

Identification of potential biomarkers by Venn diagram with LASSO regression

The results of differential genes and WGCNA were cross-analyzed to construct a Venn diagram, and four cross-over genes were obtained (*Figure 5A*). *TIMP1*, *GUCA2B*, and

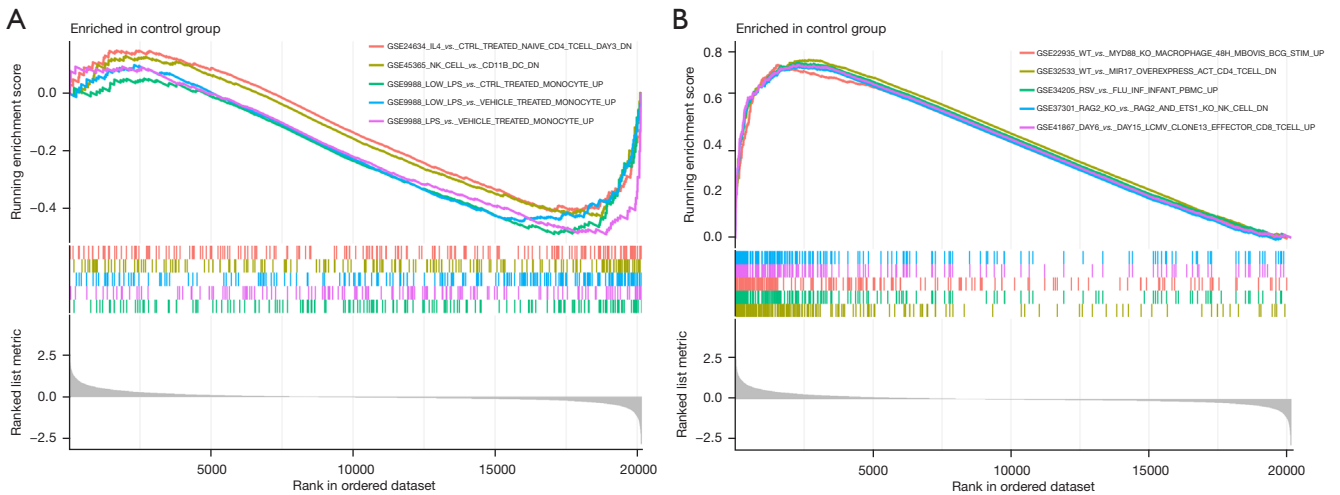

Figure 3 GSEA of IBD normal and disease group genes. (A) Enriched gene set of normal group genes in the GSEA collection. Each curve represents a pathway with a unique color, where genes up-regulated on the pathway are located close to the left of the origin of the coordinates, while down-regulated genes are located to the right of the X-axis. Only the first five most prominent pathways are shown. (B) Disease group gene enrichment pathways. GSEA, gene set enrichment analysis; IBD, inflammatory bowel disease.

HIF1A were shown to be three possible biomarkers after these four crossover genes were further evaluated using the LASSO regression model (*Figure 5B,5C*).

Expression of IBD core genes TIMP1, GUCA2B, and HIF1A

Boxplots were used to confirm the three main genes' expression levels. As illustrated in *Figure 6A-6C*, the expression levels of *GUCA2B* were significantly lower in IBD tissues compared to healthy controls, whereas the expression levels of *TIMP1* and *HIF1A* were significantly greater in IBD tissues than in healthy controls. The sensitivity and specificity of the three core genes for the diagnosis of IBD were evaluated by comparing their area under the curve (AUC) values using receiver operating characteristic (ROC) curve analysis. All three of the core genes' AUC values were higher than 0.7, demonstrating the strong diagnostic utility of these genes for IBD (*Figure 6D-6F*). The GSE179285 dataset was used to further validate the diagnostic value of the three core genes mentioned above in order to confirm their clinical relevance. While the AUC values of the *TIMP1* and *HIF1A* genes were no less than 0.7, the AUC value of *GUCA2B* was greater than 0.8 (*Figure 6G-6I*).

Immune cell infiltration and its correlation with hub genes

The ssGSEA method was utilized to evaluate the association between immune cell infiltration differences between healthy controls and IBD patients in order to look into these disparities further. The location of 28 immune cells in the GSE75214 sample is depicted in *Figure 7A*. The immune cell infiltration analysis results indicated that the following cells were present: neutrophils, plasmacytoid dendritic cells, regulatory T cells, T follicular helper cells type 1, T helper cells, type 2 T helper cells, eosinophils, γδ T cells, immature dendritic cells, myeloid-derived suppressor cells, macrophages, mast cells, NK T cells, NK cells, and compared to healthy tissues, there was a large increase in CD4 T cells, indicating a crucial role for these cells in the development of IBD (*Figure 7B*). Correlation analyses of 28 immune cells with pivotal genes showed that activated CD4. T cells, activated dendritic cells, eosinophils, γδ T cells, immature dendritic cells, bone marrow-derived suppressor cells, macrophages, mast cells, NK T cells, NK cells, neutrophils, plasmacytoid dendritic cells, regulatory T cells, T follicular helper cells type 1, T helper cells, T helper cells type 2, and central memory CD4 T cells were positively correlated with *TIMP1* and *HIF1A* (both P<0.05) and negatively correlated with *GUCA2B* (both P<0.05)

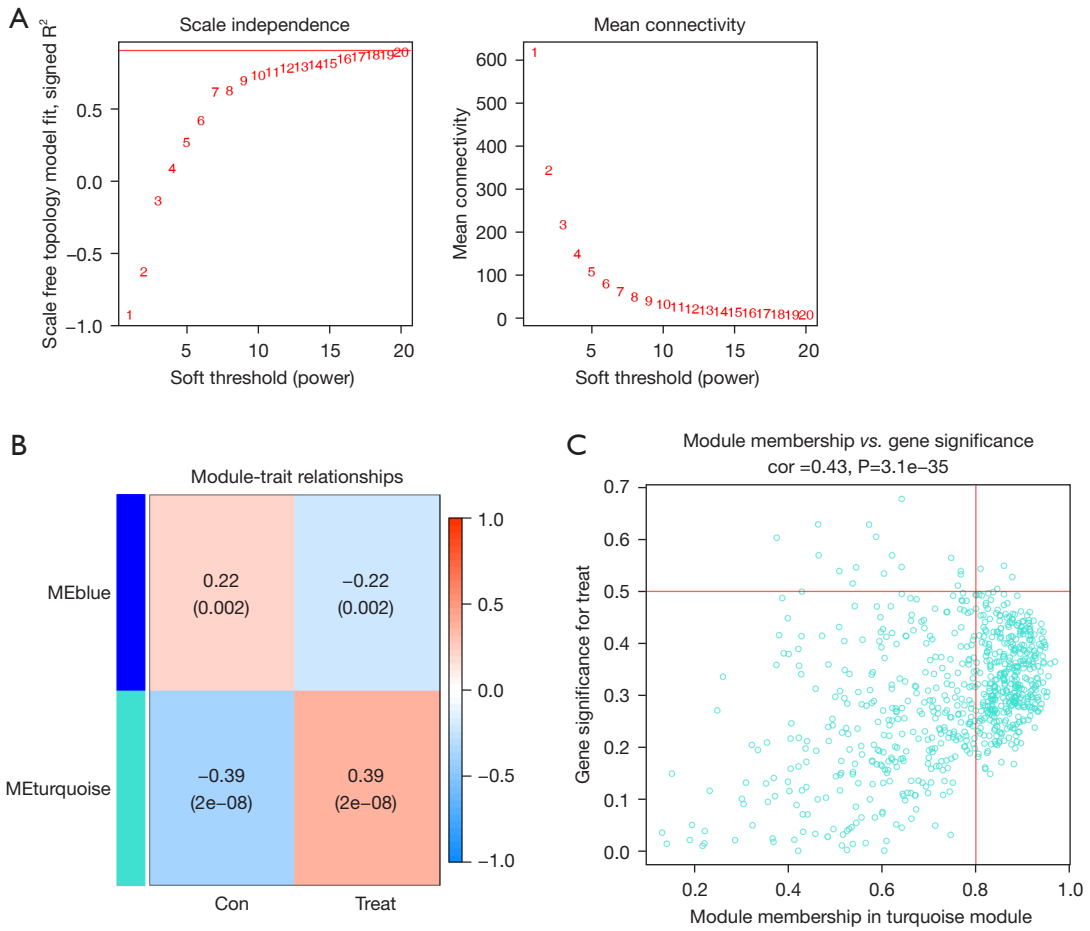

Figure 4 Weighted correlation network analysis identification of immune-related candidate biomarker genes in IBD. (A) Calculation of the soft threshold power (left) and the mean connectivity (right). (B) The clustered modules. (C) Correlation of genes in the turquoise module between MM and GS. Con, control; IBD, inflammatory bowel disease; MM, module membership; GS, gene significance.

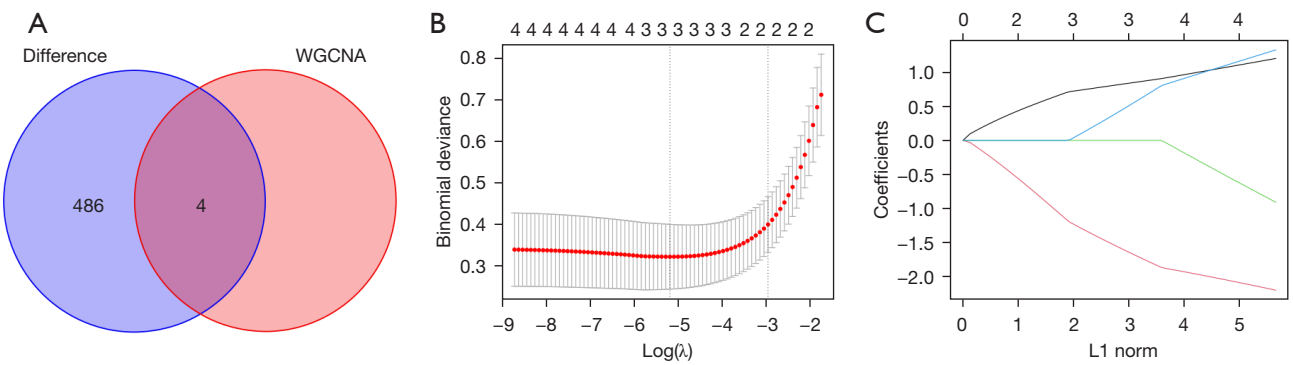

Figure 5 Venn and LASSO regression screening for potential biomarkers. (A) By extracting the intersection of turquoise module genes and DEGs of weighted correlation network analysis, four candidate hub genes were obtained. (B) Two dotted vertical lines denote the optimal parameter λ (1 – se) values based respectively. Lastly, three variables with nonzero coefficients were selected by verifying the optimal λ (1 – se) in the LASSO model. (C) The relationship between the selected features and the absolute value of the coefficient. WGCNA, weighted correlation network analysis; DEG, differentially expressed gene; LASSO, least absolute shrinkage and selection operator; se, standard error.

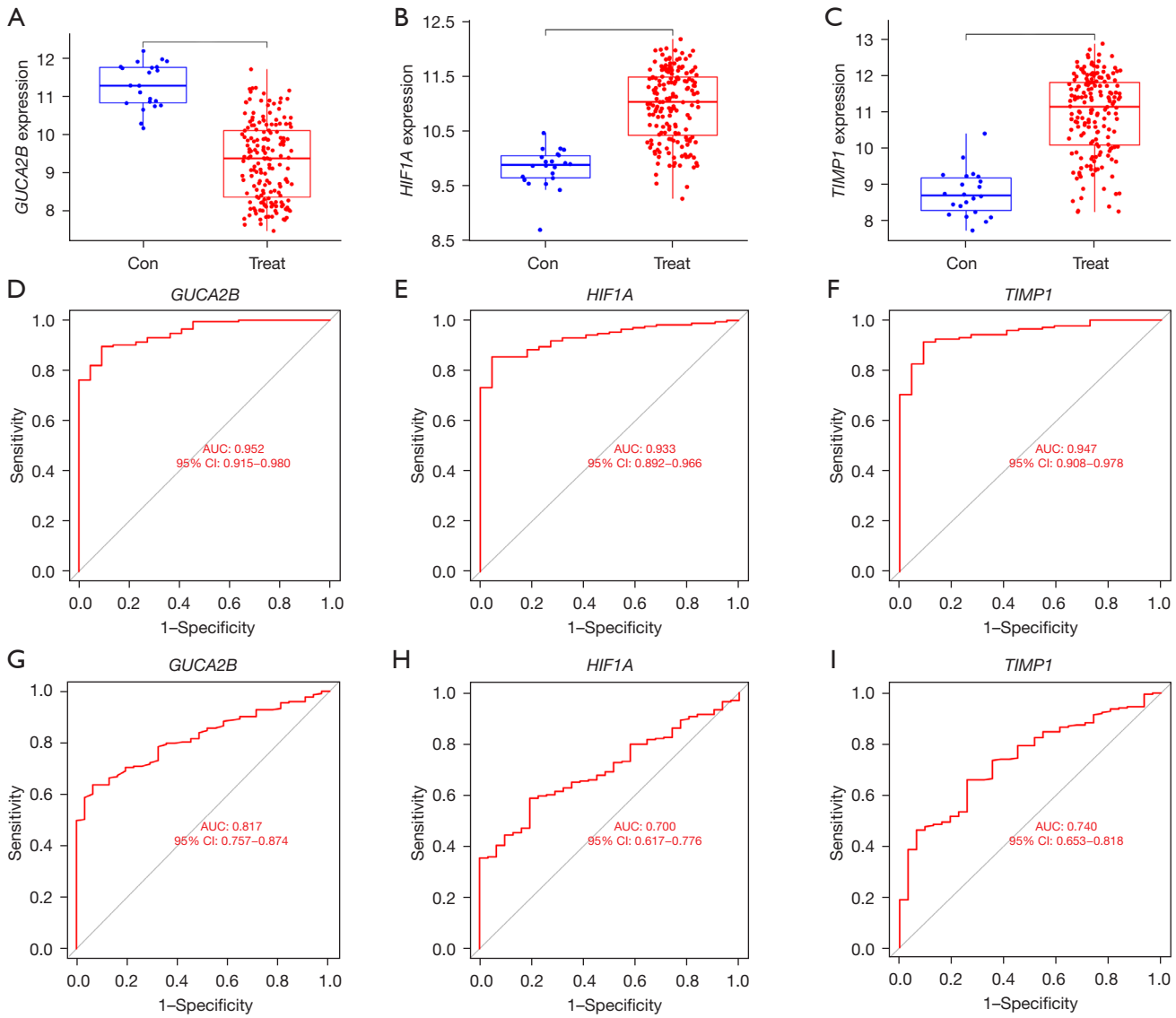

Figure 6 Three main genes' expression levels. (A-C) *GUCA2B*, *HIF1A*, and *TIMP1* expression levels in normal and IBD tissues, where blue represents normal tissues and red represents diseased tissues. (D-F) The sensitivity and specificity of the three core genes for the diagnosis of IBD were assessed using ROC curves in the dataset GSE75214. (G-I) The diagnostic value of the three center genes was further validated in the GSE179285 dataset using ROC curves. Con, control; AUC, area under the curve; CI, confidence interval; IBD, inflammatory bowel disease; ROC, receiver operating characteristic.

(*Figure 7C*). These results further demonstrated the critical role played by these immune cells in the progression of IBD.

Discussion

Two modules were found to be clinically significant and preserved in this investigation using WGCNA. The genes in these two modules were considerably enriched in BPs including macrophages, CD4, and CD8 T cells, according to the results of GO and KEGG analyses. There was a direct connection between IBD and all of these BPs. Following the identification of three important genes—*TIMP1*, *GUCA2B*, and *HIF1A*—by an artificial neural network method employing LASSO regression, the outcomes demonstrated that these three gene models could successfully distinguish between IBD tissues and normal tissues.

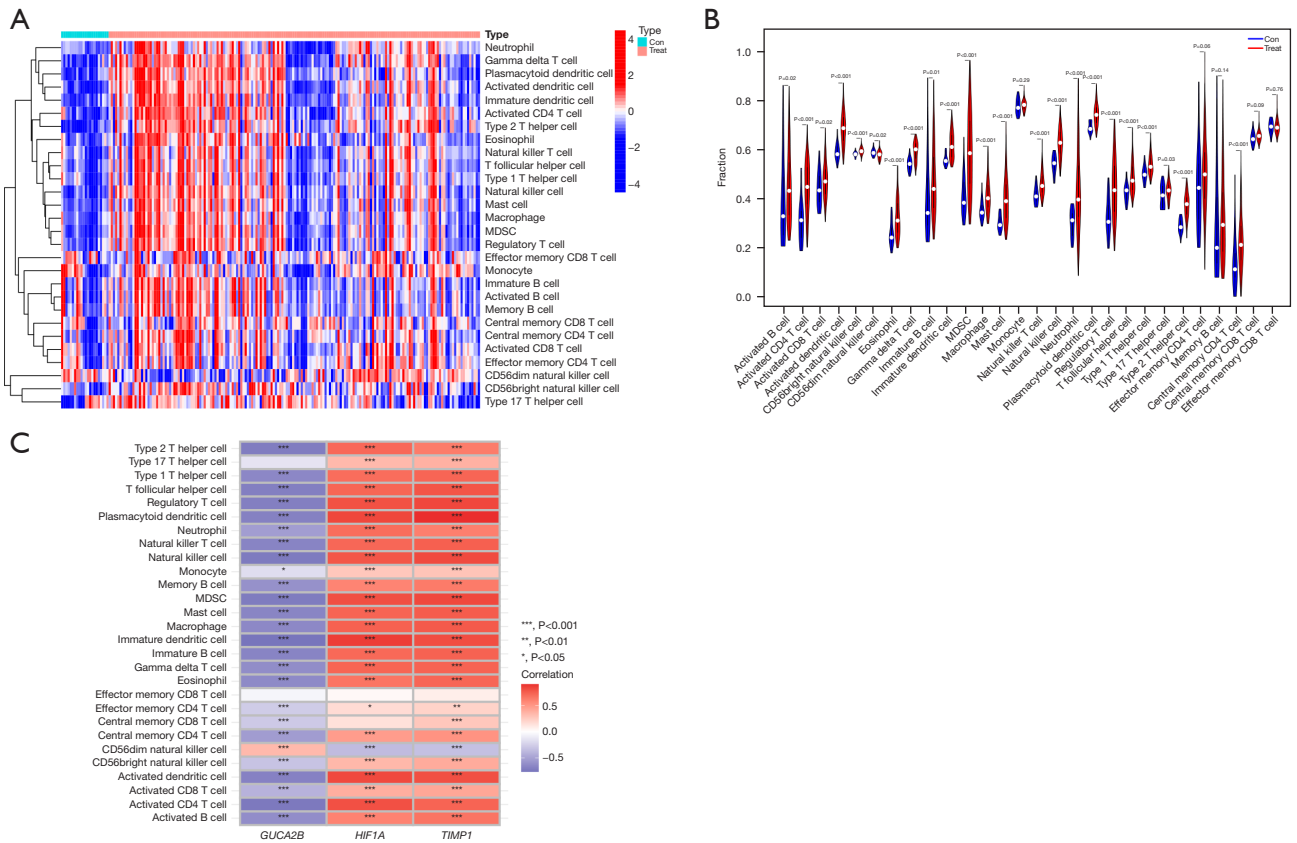

Figure 7 Immune infiltration analysis. (A) Differences in enrichment abundance between normal and IBD tissues in 28 immune cells in the GSE75214 dataset. (B) Difference in enrichment abundance of 28 immune cells in the GSE75214 dataset. Blue represents normal tissue, red represents IBD tissue; X-axis represents 28 immune cells; and Y-axis represents immune cell infiltration abundance. (C) Expression levels of the three core genes in 28 immune cells. Red represents high expression and purple represents low expression. Con, control; MDSC, myeloid-derived suppressor cell; IBD, inflammatory bowel disease.

GUCA2B gene

GUCA2B, also known as urinary guanosine (UGN), belongs to the guanosine lipopeptide family. Gastrointestinal epithelial cells are the main source of UGN, which has been demonstrated to be highly expressed in the normal gastrointestinal tract (15) and has been associated with the maintenance of gastrointestinal intestinal fluid homeostasis and intestinal function (16), which regulates salt and water homeostasis in the intestines and the kidneys (17). Brenna *et al.* observed that the dysregulatory mechanisms involved in the development of IBD include a *GUCA2B* transcription reduction, which lead to a significant downregulation of *GUCA2B* in the inflamed colonic mucosa of IBD (18). It has also been noted that it exhibits down-regulated gene expression levels in different subtypes of IBD (19). In addition, uroguanosine also plays an important role in the regulation of intestinal secretions and is an endogenous ligand for the guanylate cyclase-C [GC-C (*GUCY2C*)] receptor, which is considered to be an "intestinal natriuretic factor", and its progenitor hormone UGN (proUGN), which is secreted from the duodenal epithelium into the lumen, enters the circulation and activates the luminal membrane receptor of enterocytes, GC-C (*GUCY2C*). The enterocyte luminal membrane receptor GC-C has been shown to be involved in the maintenance of intestinal barrier function, regulation of satiety, irritable bowel syndrome (IBS), and tumor growth (20,21).

Uridoguanosine also plays an obvious role in inhibiting the development and progression of intestinal polyps by participating in the cyclic guanosine monophosphate (cGMP)-mediated signaling mechanism in intestinal cells and inducing normal cell death, which leads to the

regulation of renewal of intestinal mucosal cells and maintenance of the stability of the mucosal barrier. Various evidences have shown that uroguanosine can stimulate Cl^- secretion and K^+ efflux from epithelial cells by activating the receptor for guanylate cyclase (22), and K^+ efflux plays a significant role in inducing apoptosis (23,24). In other words, uroguanosine stimulates intestinal fluid secretion by activating the intracellular cGMP signaling pathway, which promotes K^+ cycling on the basolateral plasma membrane of enterocytes, and also activates anion channels on the surface of uroguanosine-targeted cells, and uroguanosine, which is produced and released locally in the intestinal mucosa, induces apoptosis of gastrointestinal tract mucosal lining epithelial cells through activation of GC-C. The treatment of UC with GC-C agonists has been described in earlier years, which serves to restore the intestinal homeostatic signaling pathway and promotes the integrity of the colonic mucosa (25). The mouse experiments of Han *et al.* (26) have also demonstrated that the UGN, and its signaling through the GC-C, is crucial for the maintenance of the small intestinal barrier function. It can be argued that lack of uroguanosine disrupts intestinal homeostasis, which may in turn lead to the development of IBD.

HIF1A gene

The hypoxia-inducible factor (HIF) family includes *HIF1A*. Almost all cells have HIF-1a, and in hypoxic cells—particularly those exposed to less than 6% O_2 concentration—the expression of HIF-1a rises exponentially (27). Numerous HIF-1-regulated target genes, such as glycolysis, erythropoiesis, angiogenesis, and vascular remodeling, have been found to be crucial for the cellular and systemic physiological responses to hypoxia (28).

Inflammation and hypoxia are hallmarks of IBD. Massive inflammatory cell infiltration and excessive immune response are hallmarks of IBD intestinal mucosal immunopathology (29,30). An imbalance of regulatory T (Treg) cells and effector T (Teff) cells, including type 1 T helper (Th1) cells, Th17, and Th2 cells, respectively, has been reported to exacerbate inflammatory injury in the intestinal mucosa of IBD. In addition, changes in the intestinal microenvironment can lead to intestinal inflammation. These changes induce hypoxia in the intestinal mucosa, i.e., oxygen delivery is lower than tissue demand. A study has shown that hypoxia is present in the environment of mucosal inflammatory diseases such as IBD (31). Under these circumstances, hypoxia controls a group of genes that initiate metabolic or oxygen supply pathways, allowing organs to adjust to the hypoxic environment. The HIF, which is made up of the b-subunit and the a-subunit (which includes the three isoforms HIF-1a, HIF-2a, and HIF-3a), mediates this modulation of hypoxia (32,33). hypoxia response elements (HREs) are bound by this inducible factor, which triggers a cascade of hypoxia reactions that regulate the survival and function of immune cells (34). According to earlier research, HIF-1 largely functions as a pro-inflammatory agent in macrophages during IBD and an anti-inflammatory agent in T cells, DCs, and epithelial cells (35-38). Increased expression of HIF-1a in the inflamed mucosa of IBD is regulated by a variety of inflammatory stimuli, such as pro-inflammatory cytokines and LPS (39,40). Tumor necrosis factor (TNF), one of the key pro-inflammatory cytokines in IBD, induces HIF-1a expression under normoxic conditions, and thus there may also be a correlation between the upregulation of HIF-1a and the development of colon cancer (41).

Moreover, angiogenesis is a defining characteristic of persistent inflammation (42), and unchecked vascular endothelial growth factor (VEGF) production causes excessive intestinal blood vessel growth and an imbalance in the capillary network's distribution, which lowers tissue perfusion. Elevated thrombopoietin levels, activated platelet predominance, and high platelet counts have all been shown to enhance the risk of intestinal ischemia in IBD patients (43,44). This can result in focal hypoxia and *HIF1A* overexpression. Additionally, cytokines stimulate HIF-1a protein, which is then produced by inflammatory cell components (45). The expression of angiogenic factor thymidine phosphorylase (TP) in stromal fibroblasts may also be upregulated due to cytokine stimulation (46). Additionally, the stress oxidative response triggered by 2-deoxyribose may either directly increase the expression of HIF-1a molecules or possibly facilitate the release of cytokines from fibroblasts, which in turn triggers the expression of HIF-1a molecules. Surgical specimens from IBD patients have also shown that activation of HIF-1 and HIF-2 correlates with increased vascular density in the lesion area (47). However, we do not know exactly what is the causal relationship between angiogenesis or local hypoxia and elevated *HIF1A* levels.

TIMP1 gene

TIMP1, which belongs to the *TIMP* gene family, plays an

important role in regulating cell proliferation and anti-apoptotic functions (48). *In vitro* studies have shown that overexpression of *TIMP1* can lead to a substantial increase in genes involved in proliferation, apoptosis, and signaling (49). Wang *et al.* demonstrated that *TIMP1* was positively expressed in 80-89% of UC patients with inflammatory ulcerative lesions and intact colonic mucosa cases (50). Rath *et al.* reported that inflammatory mucosa in adult IBD patients. Rath *et al.* reported that *TIMP1* levels were significantly higher in inflamed mucosa of adult IBD patients (51). In addition to this, *TIMP1* is also a therapeutic target of mesalazine, and the expression of *TIMP1* in UC was significantly reduced after mesalazine drug treatment (52).

The *TIMP* gene family is also a natural inhibitor of matrix metalloproteinases (MMPs), a group of peptidases involved in extracellular matrix degradation. It has been shown that MMP-2 and MMP-9 can be detected in the epithelium of the inferior crypt and in ulcerative-altered mucosa (53-55), and that MMP-9 activity can be detected in the luminal fluid of patients with IBD, and that interleukin 6 upregulates MMP-9 activity and protein expression during inflammation (56). It has been noted that MMP-9 is significantly upregulated in dextran sulfate sodium (DSS)-induced UC, and targeting MMP-9 alleviates experimental symptoms of UC (57). Whereas *TIMP1* inhibits the angiogenic activity of macrophages by forming a certain specific form with MMP-9, a study by de Bruyn *et al.* claimed that the upregulated MMP and *TIMP* genes were significantly decreased more than two-fold in active IBD tissues after infliximab treatment compared to pre-treatment (58). We conjectured, in light of these data, that *TIMP1* and MMP-9 interaction may be crucial for the development and course of IBD.

Furthermore, *TIMP1* has been shown to have anti-apoptotic properties and to activate downstream pathways, which decreases the sensitivity of tumor cells to a range of anticancer medications. Study has shown that *TIMP1* can bind to the CD63/integrin β1 complex and inhibit apoptosis (59). Through the development of transcription factors (e.g., *SLUG*, etc.) of the epithelial-mesenchymal transition (EMT) process, it has been discovered that *TIMP1* may promote colon cancer invasion and metastasis. This process results in the down-regulation of the epithelial marker E-cadherin and the up-regulation of the mesenchymal marker fiber connectin (60). It is reasonable to speculate that the overexpression of *TIMP1* may inhibit the normal apoptosis of intestinal mucosal cells, thereby affecting the self-regulatory renewal of the mucosa, disrupting the intestinal mucosal barrier, contributing to the development of IBD, and increasing the prevalence of colon cancer. Notably, *TIMP1* is also a secreted protein that can be detected in blood and body fluids by enzyme-linked immunosorbent assay (ELISA), which may make it a potential serum marker for IBD, but the exact expression level of *TIMP1* in the blood of IBD patients remains to be explored.

Conclusions

In summary, the results of this study show that the three core genes, *TIMP1*, *GUCA2B*, and *HIF1A*, can be used as biomarkers of IBD, as well as potential therapeutic targets. The limitation is the lack of experiments to validate the current theoretical results, but this lays a scientific theoretical basis for our future research, and our results may provide a new prospect for the future treatment of IBD, and the development of new drugs and therapeutic technologies in the future. This study also has some limitations: the findings are derived from data analysis only and have not been experimentally validated. However, it can provide a strong scientific theory and basis for further experimental studies in the future.

Acknowledgments

Funding: None.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://tcr.amegroups.com/article/view/10.21037/tcr-24-274/rc

Peer Review File: Available at https://tcr.amegroups.com/article/view/10.21037/tcr-24-274/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tcr.amegroups.com/article/view/10.21037/tcr-24-274/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as

revised in 2013).

References

1. Demir A, Kahraman R, Candan G, et al. The role of FAS gene variants in inflammatory bowel disease. Turk J Gastroenterol 2020;31:356-61.
2. Kaplan GG. The global burden of IBD: from 2015 to 2025. Nat Rev Gastroenterol Hepatol 2015;12:720-7.
3. Lloyd-Price J, Arze C, Ananthakrishnan AN, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature 2019;569:655-62.
4. Chang JT. Pathophysiology of Inflammatory Bowel Diseases. N Engl J Med 2020;383:2652-64.
5. Geremia A, Biancheri P, Allan P, et al. Innate and adaptive immunity in inflammatory bowel disease. Autoimmun Rev 2014;13:3-10.
6. Halling ML, Kjeldsen J, Knudsen T, et al. Patients with inflammatory bowel disease have increased risk of autoimmune and inflammatory diseases. World J Gastroenterol 2017;23:6137-46.
7. Yi J, Bergstrom K, Fu J, et al. Dclk1 in tuft cells promotes inflammation-driven epithelial restitution and mitigates chronic colitis. Cell Death Differ 2019;26:1656-69.
8. Harrington LE, Mangan PR, Weaver CT. Expanding the effector CD4 T-cell repertoire: the Th17 lineage. Curr Opin Immunol 2006;18:349-56.
9. Baumgart DC, Le Berre C. Newer Biologic and Small-Molecule Therapies for Inflammatory Bowel Disease. N Engl J Med 2021;385:1302-15.
10. Raine T, Danese S. Breaking Through the Therapeutic Ceiling: What Will It Take? Gastroenterology 2022;162:1507-11.
11. Vancamelbeke M, Vanuytsel T, Farré R, et al. Genetic and Transcriptomic Bases of Intestinal Epithelial Barrier Dysfunction in Inflammatory Bowel Disease. Inflamm Bowel Dis 2017;23:1718-29.
12. Godec J, Tan Y, Liberzon A, et al. Compendium of Immune Signatures Identifies Conserved and Species-Specific Biology in Response to Inflammation. Immunity 2016;44:194-206.
13. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008;9:559.
14. Keir ME, Fuh F, Ichikawa R, et al. Regulation and Role of αE Integrin and Gut Homing Integrins in Migration and Retention of Intestinal Lymphocytes during Inflammatory Bowel Disease. J Immunol 2021;207:2245-54.
15. Bełtowski J. Guanylin and related peptides. J Physiol Pharmacol 2001;52:351-75.
16. Cappelli K, Gialletti R, Tesei B, et al. Guanylin, Uroguanylin and Guanylate Cyclase-C Are Expressed in the Gastrointestinal Tract of Horses. Front Physiol 2019;10:1237.
17. Maake C, Auf der Maur F, Jovanovic K, et al. Occurrence and localization of uroguanylin in the aging human prostate. Histochem Cell Biol 2003;119:69-76.
18. Brenna Ø, Bruland T, Furnes MW, et al. The guanylate cyclase-C signaling pathway is down-regulated in inflammatory bowel disease. Scand J Gastroenterol 2015;50:1241-52.
19. Perez K, Ngollo M, Rabinowitz K, et al. Meta-Analysis of IBD Gut Samples Gene Expression Identifies Specific Markers of Ileal and Colonic Diseases. Inflamm Bowel Dis 2022;28:775-82.
20. Brenna Ø, Furnes MW, Munkvold B, et al. Cellular localization of guanylin and uroguanylin mRNAs in human and rat duodenal and colonic mucosa. Cell Tissue Res 2016;365:331-41.
21. Lorenz JN, Nieman M, Sabo J, et al. Uroguanylin knockout mice have increased blood pressure and impaired natriuretic response to enteral NaCl load. J Clin Invest 2003;112:1244-54.
22. Joo NS, London RM, Kim HD, et al. Regulation of intestinal Cl- and HCO3-secretion by uroguanylin. Am J Physiol 1998;274:G633-44.
23. Bortner CD, Hughes FM Jr, Cidlowski JA. A primary role for K+ and Na+ efflux in the activation of apoptosis. J Biol Chem 1997;272:32436-42.
24. Hughes FM Jr, Bortner CD, Purdy GD, et al. Intracellular K+ suppresses the activation of apoptosis in lymphocytes. J Biol Chem 1997;272:30567-76.
25. Pitari GM. Pharmacology and clinical potential of guanylyl cyclase C agonists in the treatment of ulcerative colitis. Drug Des Devel Ther 2013;7:351-60.
26. Han X, Mann E, Gilbert S, et al. Loss of guanylyl cyclase

C (GCC) signaling leads to dysfunctional intestinal barrier. PLoS One 2011;6:e16139.
27. Jiang BH, Semenza GL, Bauer C, et al. Hypoxia-inducible factor 1 levels vary exponentially over a physiologically relevant range of O2 tension. Am J Physiol 1996;271:C1172-80.
28. Semenza GL. HIF-1 and human disease: one highly involved factor. Genes Dev 2000;14:1983-91.
29. Knights D, Lassen KG, Xavier RJ. Advances in inflammatory bowel disease pathogenesis: linking host genetics and the microbiome. Gut 2013;62:1505-10.
30. Peterson CT, Sharma V, Elmén L, et al. Immune homeostasis, dysbiosis and therapeutic modulation of the gut microbiota. Clin Exp Immunol 2015;179:363-77.
31. Hatoum OA, Binion DG, Gutterman DD. Paradox of simultaneous intestinal ischaemia and hyperaemia in inflammatory bowel disease. Eur J Clin Invest 2005;35:599-609.
32. Makino Y, Cao R, Svensson K, et al. Inhibitory PAS domain protein is a negative regulator of hypoxia-inducible gene expression. Nature 2001;414:550-4.
33. Tian H, McKnight SL, Russell DW. Endothelial PAS domain protein 1 (EPAS1), a transcription factor selectively expressed in endothelial cells. Genes Dev 1997;11:72-82.
34. Taylor CT, Colgan SP. Hypoxia and gastrointestinal disease. J Mol Med (Berl) 2007;85:1295-300.
35. Bäcker V, Cheung FY, Siveke JT, et al. Knockdown of myeloid cell hypoxia-inducible factor-1α ameliorates the acute pathology in DSS-induced colitis. PLoS One 2017;12:e0190074.
36. Flück K, Breves G, Fandrey J, et al. Hypoxia-inducible factor 1 in dendritic cells is crucial for the activation of protective regulatory T cells in murine colitis. Mucosal Immunol 2016;9:379-90.
37. Higashiyama M, Hokari R, Hozumi H, et al. HIF-1 in T cells ameliorated dextran sodium sulfate-induced murine colitis. J Leukoc Biol 2012;91:901-9.
38. Karhausen J, Furuta GT, Tomaszewski JE, et al. Epithelial hypoxia-inducible factor-1 is protective in murine experimental colitis. J Clin Invest 2004;114:1098-106.
39. Hellwig-Bürgel T, Rutkowski K, Metzen E, et al. Interleukin-1beta and tumor necrosis factor-alpha stimulate DNA binding of hypoxia-inducible factor-1. Blood 1999;94:1561-7.
40. Frede S, Stockmann C, Freitag P, et al. Bacterial lipopolysaccharide induces HIF-1 activation in human monocytes via p44/42 MAPK and NF-kappaB. Biochem J 2006;396:517-27.
41. Sun M, He C, Wu W, et al. Hypoxia inducible factor-1α-induced interleukin-33 expression in intestinal epithelia contributes to mucosal homeostasis in inflammatory bowel disease. Clin Exp Immunol 2017;187:428-40.
42. Danese S, Sans M, de la Motte C, et al. Angiogenesis as a novel component of inflammatory bowel disease pathogenesis. Gastroenterology 2006;130:2060-73.
43. Kapsoritakis AN, Potamianos SP, Sfiridaki AI, et al. Elevated thrombopoietin serum levels in patients with inflammatory bowel disease. Am J Gastroenterol 2000;95:3478-81.
44. Kanazawa S, Tsunoda T, Onuma E, et al. VEGF, basic-FGF, and TGF-beta in Crohn's disease and ulcerative colitis: a novel mechanism of chronic intestinal inflammation. Am J Gastroenterol 2001;96:822-8.
45. Thornton RD, Lane P, Borghaei RC, et al. Interleukin 1 induces hypoxia-inducible factor 1 in human gingival and synovial fibroblasts. Biochem J 2000;350 Pt 1:307-12.
46. Braybrooke JP, Propper DJ, O'Byrne KJ, et al. Induction of thymidine phosphorylase as a pharmacodynamic end-point in patients with advanced carcinoma treated with 5-fluorouracil, folinic acid and interferon alpha. Br J Cancer 2000;83:219-24.
47. Giatromanolaki A, Sivridis E, Maltezos E, et al. Hypoxia inducible factor 1alpha and 2alpha overexpression in inflammatory bowel disease. J Clin Pathol 2003;56:209-13.
48. Nalluri S, Ghoshal-Gupta S, Kutiyanawalla A, et al. TIMP-1 Inhibits Apoptosis in Lung Adenocarcinoma Cells via Interaction with Bcl-2. PLoS One 2015;10:e0137673.
49. Batra J, Robinson J, Soares AS, et al. Matrix metalloproteinase-10 (MMP-10) interaction with tissue inhibitors of metalloproteinases TIMP-1 and TIMP-2: binding studies and crystal structure. J Biol Chem 2012;287:15935-46.
50. Wang YD, Yan PY. Expression of matrix metalloproteinase-1 and tissue inhibitor of metalloproteinase-1 in ulcerative colitis. World J Gastroenterol 2006;12:6050-3.
51. Rath T, Roderfeld M, Graf J, et al. Enhanced expression of MMP-7 and MMP-13 in inflammatory bowel disease: a precancerous potential? Inflamm Bowel Dis 2006;12:1025-35.
52. Huang R, Wang K, Gao L, et al. TIMP1 Is A Potential Key Gene Associated With The Pathogenesis And Prognosis Of Ulcerative Colitis-Associated Colorectal Cancer. Onco Targets Ther 2019;12:8895-904.
53. Seifert WF, Wobbes T, Hendriks T. Divergent patterns of

matrix metalloproteinase activity during wound healing in ileum and colon of rats. Gut 1996;39:114-9.
54. Schuppan D, Hahn EG. MMPs in the gut: inflammation hits the matrix. Gut 2000;47:12-4.
55. Tarlton JF, Whiting CV, Tunmore D, et al. The role of up-regulated serine proteases and matrix metalloproteinases in the pathogenesis of a murine model of colitis. Am J Pathol 2000;157:1927-35.
56. Castaneda FE, Walia B, Vijay-Kumar M, et al. Targeted deletion of metalloproteinase 9 attenuates experimental colitis in mice: central role of epithelial-derived MMP. Gastroenterology 2005;129:1991-2008.
57. Ishida K, Takai S, Murano M, et al. Role of chymase-dependent matrix metalloproteinase-9 activation in mice with dextran sodium sulfate-induced colitis. J Pharmacol Exp Ther 2008;324:422-6.
58. de Bruyn M, Machiels K, Vandooren J, et al. Infliximab restores the dysfunctional matrix remodeling protein and growth factor gene expression in patients with inflammatory bowel disease. Inflamm Bowel Dis 2014;20:339-52.
59. Toricelli M, Melo FH, Peres GB, et al. Timp1 interacts with beta-1 integrin and CD63 along melanoma genesis and confers anoikis resistance by activating PI3-K signaling pathway independently of Akt phosphorylation. Mol Cancer 2013;12:22.
60. Song G, Xu S, Zhang H, et al. TIMP1 is a prognostic marker for the progression and metastasis of colon cancer through FAK-PI3K/AKT and MAPK pathway. J Exp Clin Cancer Res 2016;35:148.

Cite this article as: Zhu Y, Pan Y, Fan L, Zou M, Liu Y, Hu J, Xia S, Li Y, Dai R, Wu W. Bioinformatics analysis-based mining of potential markers for inflammatory bowel disease and their immune relevance. Transl Cancer Res 2024;13(8):3960-3973. doi: 10.21037/tcr-24-274

