Special Issue: Gastrointestinal Genomics and Outcome Research

Journal of INTERNATIONAL MEDICAL RESEARCH

Journal of International Medical Research
48(5) 1–16

DOI: 10.1177/0300060519887268
journals.sagepub.com/home/imr
SAGE

Identification of differentially expressed genes between the colon and ileum of patients with inflammatory bowel disease by gene co-expression analysis

Yuting Zhang[1,2], Bo Shen[3], Liya Zhuge[1] and Yong Xie[1,4]

Abstract

Objective: We aimed to identify differentially expressed genes (DEG) in patients with inflammatory bowel disease (IBD).

Methods: RNA-seq data were obtained from the Array Express database. DEG were identified using the edgeR package. A co-expression network was constructed and key modules with the highest correlation with IBD inflammatory sites were identified for analysis. The Cytoscape MCODE plugin was used to identify key sub-modules of the protein–protein interaction (PPI) network. The genes in the sub-modules were considered hub genes, and functional enrichment analysis was performed. Furthermore, we constructed a drug–gene interaction network. Finally, we visualized the hub gene expression pattern between the colon and ileum of IBD using the ggpubr package and analyzed it using the Wilcoxon test.

Results: DEG were identified between the colon and ileum of IBD patients. Based on the co-expression network, the green module had the highest correlation with IBD inflammatory sites. In total, 379 DEG in the green module were identified for the PPI network. Nineteen hub genes were differentially expressed between the colon and ileum. The drug–gene network identified these hub genes as potential drug targets.

[1]Institute of Gastroenterology, the First Affiliated Hospital of Nanchang University, Nanchang, Jiangxi Province, P. R. China
[2]Department of Liver Diseases, People's Hospital of Yichun City, Yichun, Jiangxi Province, P. R. China
[3]Department of Hepatobiliary Surgery, People's Hospital of Yichun City, Yichun, Jiangxi Province, P R China
[4]Department of Gastroenterology, the First Affiliated Hospital of Nanchang University, Nanchang, Jiangxi Province, P R China

Corresponding author:
Yong Xie, Department of Gastroenterology, Institute of Gastroenterology, the First Affiliated Hospital of Nanchang University, No.17 Yongwaizheng Street, Nanchang 330006, Jiangxi Province, P R China.
Email: xieyong_tfahoncu@163.com

Conclusion: Nineteen DEG were identified between the colon and ileum of IBD patients.

Keywords
Inflammatory bowel disease, differentially expressed gene, lesion sites, weighted gene co-expression network analysis, module, protein–protein interaction network

Date received: 6 August 2019; accepted: 17 October 2019

Introduction

Inflammatory bowel disease (IBD) is a chronic idiopathic disease that causes inflammation of the gastrointestinal tract, with a heterogeneous presentation.[1,2] IBD has two subtypes: Crohn's disease (CD) and ulcerative colitis (UC), which differ in presentation and treatment processes.[3] CD is a chronic granulomatous inflammation and the lesions may involve different parts of the gastrointestinal tract.[4] Lesions mainly involve the terminal ileum and adjacent colon. In contrast, UC is a chronic nonspecific colonic inflammation,[5] and UC lesions mainly involve the colonic mucosa and submucosa. Lesions may be present in the distal colon and can be retrograde to the proximal segment, even involving the whole colon and the terminal ileum. The latest epidemiological studies show that the incidence rate of IBD is increasing throughout the world.[6] The mechanism that determines the inflammatory location of IBD remains unknown.

Genetic and environmental factors are thought to be the main risk factors for IBD, contributing to abnormal activity in the intestinal immune system; however, the pathogenesis of IBD remains unclear.[7] Genome-wide association studies (GWAS) have identified common genes affecting innate defense or cellular homeostasis that are closely related to IBD.[8,9] IBD patients often suffer from weight loss, bloody diarrhea, severe abdominal pain, and increased risk of cancer.[10] IBD gradually worsens over time, involves multiple recurrences, and persists for decades.[11] As a chronic, recurrent gastrointestinal disorder, IBD can have an adverse effect on mental health and quality of life, which may lead to suicidal behavior.[12] Therefore, it is important to clarify the molecular mechanisms of IBD that contribute to the development of clinical treatment decisions and reduce the burden of IBD in terms of clinical infrastructure and healthcare.

Gene expression changes affect many biological processes, including inflammation. It has been confirmed that T-helper (Th)17-related genes are differentially expressed between the inflammatory colon and ileum of patients with IBD, which suggests that gene expression patterns are regulated by different inflammatory sites.[13] Therefore, it is necessary to better understand gene expression differences between the colon and ileum of patients with IBD and determine whether this expression is regulated by different inflammatory sites.

High-throughput sequencing methods can simultaneously detect thousands of parameters for a single sample; however, systematically describing and analyzing these data and mining useful information remain challenging. Weighted gene co-expression network analysis (WGCNA) is a systematic biological tool for describing patterns related to gene expression in samples and it is widely applied in biomedical

research because of its powerful analytical performance.[14] However, few studies using WGCNA have been conducted on IBD.

Our research aimed to take advantage of the Array Express public database to screen differentially expressed genes (DEG) between the inflammatory colon and ileum of patients with IBD by using gene co-expression analysis, and to determine whether intestinal region-specific gene expression is regulated by different inflammatory sites.

Materials and methods

Ethics approval and consent to participate

This study used primarily bioinformatics methods and did not require ethical review.

Data acquisition and preprocessing

RNA sequencing (RNA-seq) data (accession no: E-MTAB-7604) were obtained from the Array Express database (https://www.ebi.ac.uk/arrayexpress/), a public database of European Bioinformatics Institute (EBI) microarray gene expression data.[15] Total RNA-seq data contained 43 IBD samples (26 colon samples and 17 ileum samples from patients with CD and UC). We downloaded the corresponding clinical sample information. The 43 raw gene expression files were stored in corresponding text files, which were merged into an expression matrix for subsequent analysis by the R language (www.r-project.org).

Gene information for all genes with protein products was downloaded from the HUGO Gene Nomenclature Committee (HGNC; https://www.genenames.org/).[16] Based on the obtained matrix file, the expression values of all genes with protein products were extracted according to the obtained gene information, and this file was used for further analyses.

Screening of differentially expressed genes

We identified DEG between IBD patients with lesions located in the colon and IBD patients with lesion locations in the ileum (colon vs. ileum). After the raw expression data were normalized by trimmed mean of M-values (TMM) using the edgeR package,[17] we performed differential expression analysis (colon vs. ileum). The screening threshold of DEG was set to a false discovery rate (FDR) <0.01 and $|\log FC| > 2$ (where FC is the fold change), and differential expression results were visualized as volcano plots.

Functional enrichment analysis

To explore the biological functions and biological pathways involved in genes associated with IBD patients with different lesion sites, we performed Gene Ontology (GO) functional enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses through the String database (https://string-db.org/) and ClueGo+CluePedia (http://www.cytoscape.org/download.php). A P-value <0.05 was considered as the cutoff criterion. Functional and pathway enrichment analysis results were visualized using R language (www.r-project.org).

Construction of the gene co-expression network

WGCNA analysis was performed using the WGCNA package.[18] The WGCNA algorithm is commonly used to construct weighted gene co-expression networks. The data pretreatment was different from our previous data preprocessing method. The obtained expression matrix was subjected to $\log_2(x+1)$ normalization treatment, and genes $\geq$75% of the median absolute deviation (MAD) and MAD ≥ 0.01 were screened. Missing values were

detected and processed. After pretreatment, the matrix was used for WGCNA analysis. The soft threshold is a very important part of the WGCNA algorithm. The soft threshold screening principle makes the constructed network more consistent with the scale-free network characteristics. Before performing soft threshold screening, we checked samples for outliers, and outliers were found in all samples. After the outliers were removed, the WGCNA network was constructed. Our analysis used a one-step approach to build the WGCNA network and the hierarchical clustering tree was used to perform each module. A module is defined as a group of genes with similar expression profiles. If some genes have similar expression pattern in a physiological process or in different tissues, these genes are thought to be functionally related, which can be defined as a module.

To select the appropriate model, the different lesion sites of IBD (colon or ileum) were considered external biological parameters, and Student's *t*-test was performed. The module most relevant to the external biological parameters (different lesion sites of IBD) was considered the key module.

Construction of protein–protein interaction network

Protein–protein interaction (PPI) suggests the processes by which two or more protein molecules form a protein complex through noncovalent bonds. PPI network analysis allows researchers to explore the molecular mechanisms of a disease or discover new drug targets. In the present study, we constructed a PPI network using the String database.[19] The minimum required interaction score was set to high confidence (0.7). The PPI network was visualized by using Cytoscape,[20] which was also used to calculate the connectivity of each node in the network. Nodes with a larger degree in the network have a greater role in the network.

Construction of key sub-module of PPI network

After construction of the PPI network, we used the MCODE plugin of Cytoscape to perform key sub-module analysis to obtain those sub-modules with biological significance in the PPI network. Node information of each node was calculated, including its neighbors, density, maximum k-value, the density of this k-core (core density), and the score value of the node. The score value of the node reflects the intensity of the node and its neighbors. Then, starting from the node with the largest score value, the getClusterCore() function was used to incorporate adjacent nodes that met the parameter conditions. Finally, the genes contained in the key sub-modules with MCODE score ≥ 5 were selected as hub genes for subsequent analysis.

Drug–gene interaction analysis

To explore whether a hub gene was therapeutically targeted or prioritized for drug development, the Drug–Gene Interaction database 2.0 (DGIdb2.0; http: //www.dgidb.org/) was used to predict the obtained hub genes.[21] The parameter settings were as follows: FDA approved and the drug database was limited to DrugBank. Finally, the drug–gene interaction network was visualized through Cytoscape.

Visualization of expression of hub genes in different lesion locations

Differential expression analysis of hub genes in different lesion sites (colon or ileum) of IBD was presented as boxplots using the ggpubr package. The Wilcoxon test was performed between different lesion sites.

Validation analysis

RNA-seq data (accession no: E-MTAB-5464) from the Array Express database were used to validate the identified hub genes of IBD. The datasets included 52 IBD samples (25 CD and 27 UC). We compared the expression differences of these hub genes in the colon and ileum of IBD and in CD and UC using the Wilcoxon test.

Results

Identification of DEG between the colon and ileum of IBD

Volcano plots show the DEG (FDR <0.01 and |logFC| >2) between the colon and ileum of IBD (colon or ileum) including 86 upregulated and 293 downregulated genes (Figure 1a). To explore the biological processes and pathways involved in these DEG, functional enrichment analysis was performed using the String database. The top eight terms of GO, including biological process (BP), cellular component (CC), and molecular function (MF), KEGG, and Reactome are shown separately in Figure 1b.

Construction of the gene co-expression network

First, we checked the sample for outliers, which were found and removed from all samples (Figure 2a). Then, the appropriate

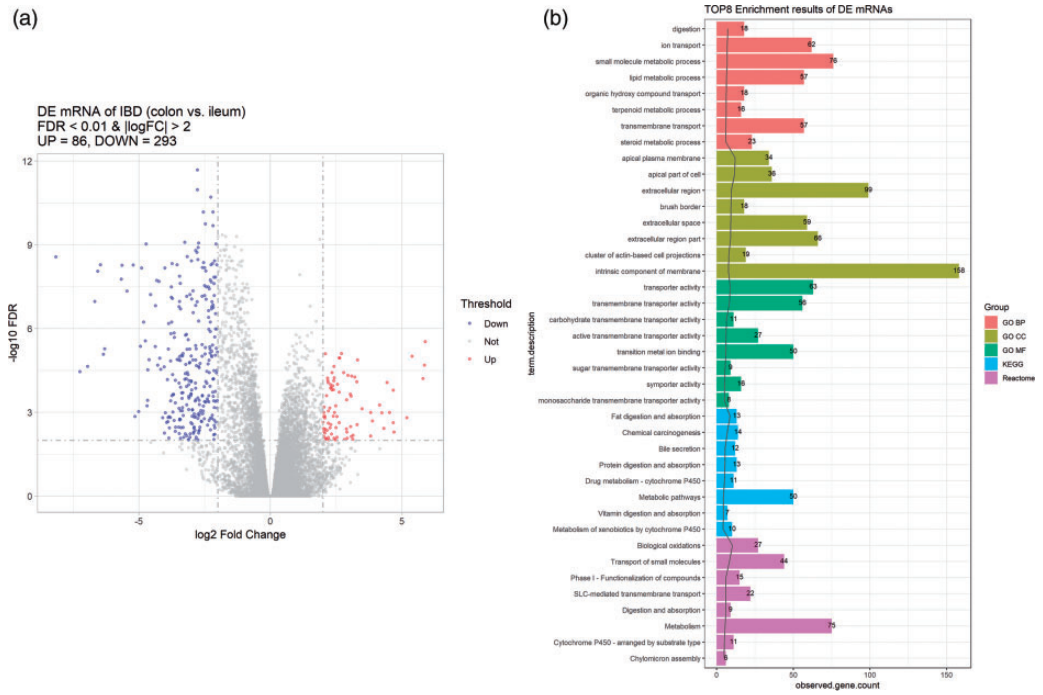

Figure 1. Differential expression analysis and functional enrichment analysis. (a) Differential expression analysis (colon vs. ileum). Red indicates upregulated genes in the colon compared with the ileum of IBD and blue indicates downregulated genes in the colon compared with the ileum of IBD. (b) Functional enrichment analysis showing the top eight enrichment terms of differentially expressed genes, including GO (BP, CC, and MF), KEGG, and Reactome. IBD, inflammatory bowel disease; GO, Gene Ontology; BB, biological process; CC, cellular component; MF, molecular function; KEGG, Kyoto Encyclopedia of Genes and Genomes; DE, differential expression; FDR, false discovery rate.

power value was screened. When the soft threshold power value was equal to 14, scale independence reached 0.8 and mean connectivity was higher than zero (Figure 2b). Therefore, the soft threshold power value was set to 14 for subsequent analysis. As shown in Figure 2c, 11 co-expression modules were identified according to the soft threshold power value, in which the gray module represented a gene that was not assigned to any module. Different modules are represented by different colors. The eigengene adjacency heatmap was used to identify correlations between different modules (Figure 2d). Modules that were divided into one branch may have similar functions. Furthermore, we explored module–trait relationships. As shown in Figure 2d, the green module had the highest correlation ($r = 0.7$, $P = 2e{-}07$) with different lesion sites of IBD (colon or ileum).

Identification of meaningful module with different inflammatory sites of IBD (colon or ileum)

To further identify meaningful modules with different lesion sites of IBD (colon or ileum), we verified modules in the gene co-expression network. As shown in Figure 3a, genes in the green module had the highest correlation with lymphovascular invasion. Moreover, the highest gene significance was found in the green module. Therefore, the green module was identified as meaningful for different inflammatory sites of IBD (colon or ileum). Membership in the green module was significantly related to gene

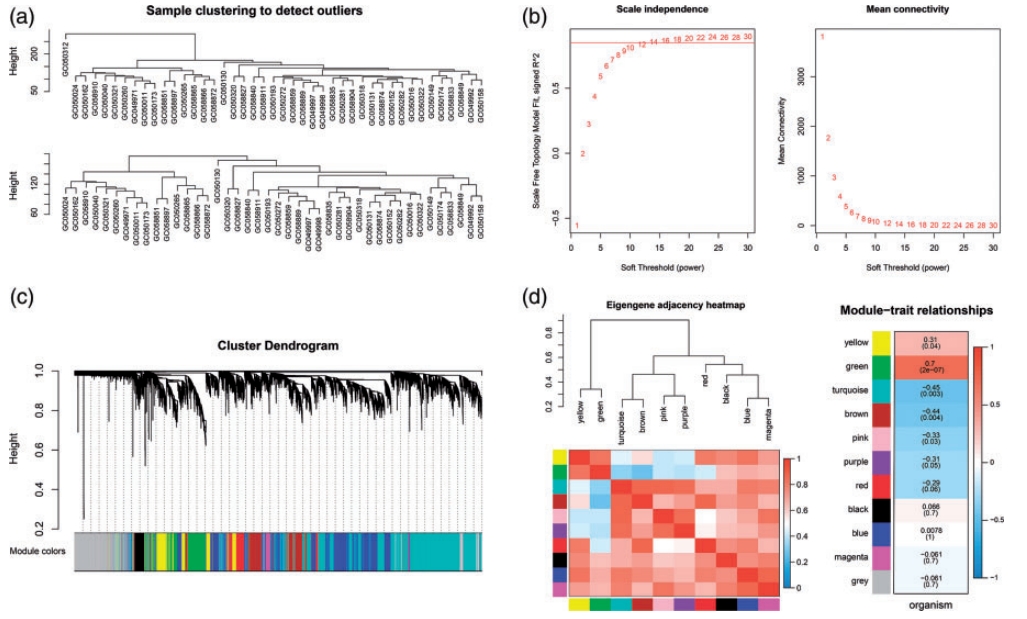

Figure 2. The gene co-expression network analysis. (a) Sample clustering to detect outliers. (b) Analysis of different soft-thresholding powers on the scale independence of co-expression modules. (c) Hierarchical clustering analysis showing each co-expression module; different modules are represented by different colors. (d) Eigengene adjacency heatmap (left) and module–trait relationships (right). The color change from blue (0) to red (0) in the heat map represents weak to strong connectivity of key genes in different modules. Module–trait relationships show the correlation between different modules and different lesion sites of IBD (colon or ileum). Each unit contains the corresponding correlation and *P*-value.

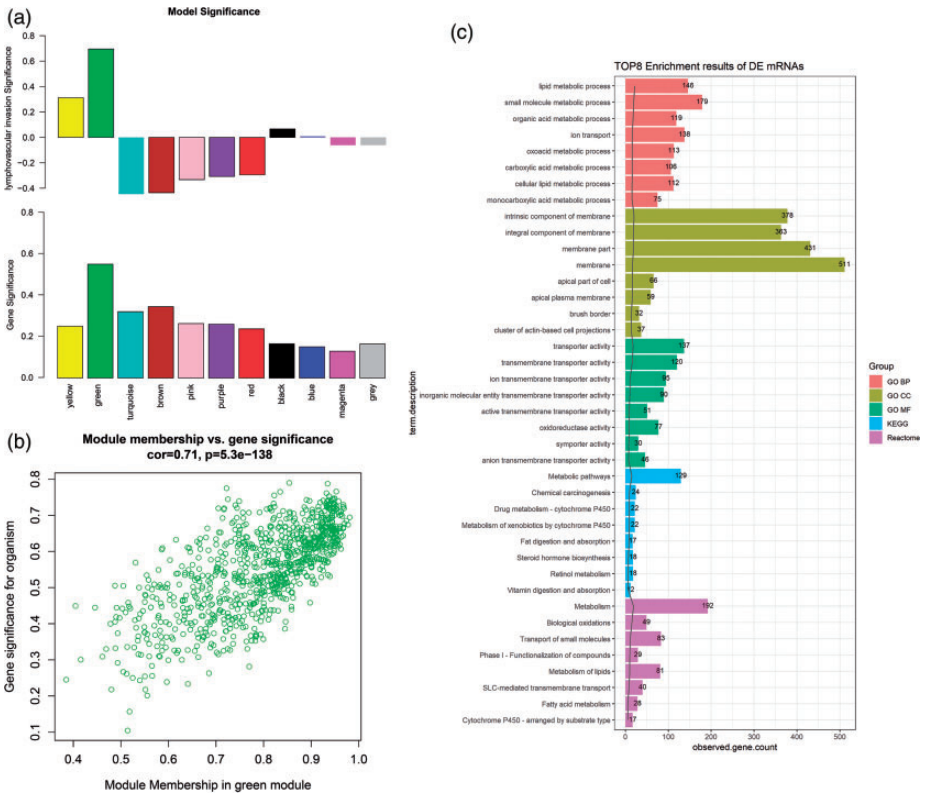

Figure 3. Identification of meaningful modules with different inflammatory lesion sites of IBD (colon or ileum) and functional enrichment analysis. (a) The correlation between different modules with lymphovascular invasion (top), and gene significance in different modules (bottom). (b) Module membership vs. gene significance in green module. (c) GO function enrichment and KEGG pathway enrichment analysis of genes in the green module. GO, Gene Ontology; BB, biological process; CC, cellular component; MF, molecular function; KEGG, Kyoto Encyclopedia of Genes and Genomes.

significance for different lesion sites of IBD in Figure 3b ($r = 0.71$, $P = 5.3e{-}138$). To further explore the biological functions and biological pathways of genes in the green module, the 879 genes were subjected to GO functional enrichment and KEGG pathway enrichment analysis. Figure 3c shows the top eight terms of GO, including BP, CC, and MF, KEGG, and Reactome.

Identification of hub genes in the green module

To obtain DEG between the colon and ileum of IBD, the 379 DEG were intersected with 879 genes in the green module (Figure 4a), and 200 genes were obtained for PPI using the String database. In the PPI network, there were 133 nodes and 237 relationship pairs (Figure 4b). The different sizes of the nodes represented the degree of the node in the PPI network. The larger the circle, the greater the degree of the node in the network, indicating that the node was more important. After obtaining key genes from the PPI network, the MCODE plug-in was used to perform further key sub-module analysis. A sub-module with a MCODE score ≥ 5 was considered a key sub-module, and three key

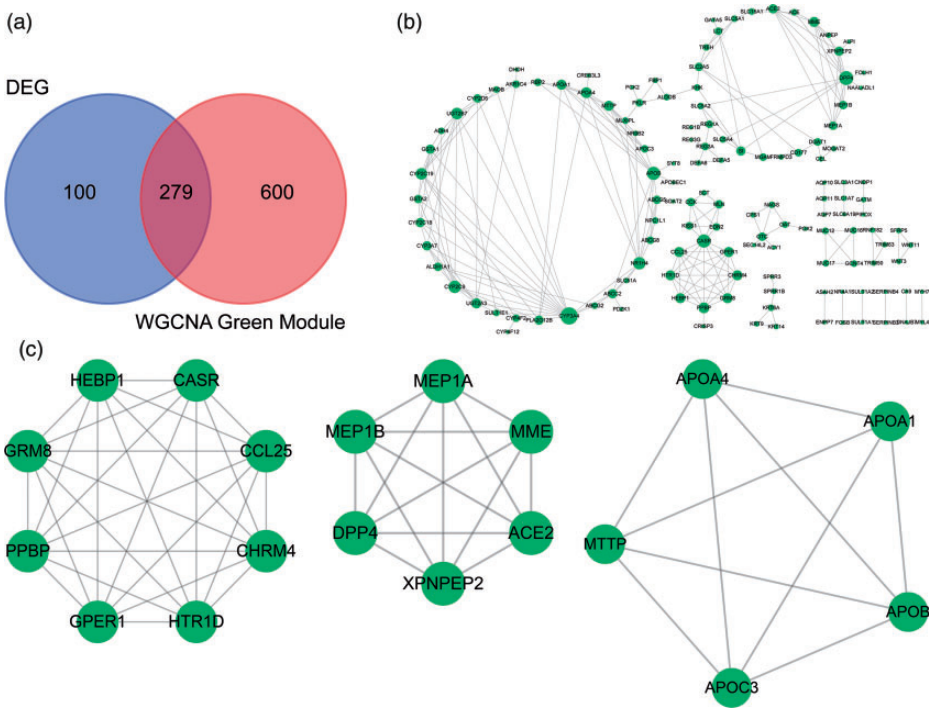

Figure 4. Identification of hub genes in the meaningful module. (a) Venn diagram showed the common genes between differentially expressed genes and genes in the green module. (b) PPI network constructed based on the common genes using the string database. (c) Analysis of key sub-modules in the PPI network using MCODE. DEG, differentially expressed genes; WGCNA, weighted gene co-expression network analysis; PPI, protein–protein interaction.

sub-modules were obtained. Sub-module 1 had a total of eight nodes, with an MCODE score of 8; sub-module 2 had a total of six nodes, with an MCODE score of 6; and sub-module 3 had a total of five nodes, with an MCODE score of 5 (Figure 4c). The 19 genes in the three key sub-modules were differentially expressed between the colon and ileum of IBD.

Functional enrichment analysis of hub genes and drug–gene network construction

To explore the biological functions and biological pathways involved in the 19 hub genes, functional enrichment analysis was performed. Figure 5a shows the top eight terms of GO, including BP, CC, and MF, KEGG, and Reactome using the String database. Pathway enrichment analysis of the 19 hub genes was performed using ClueGo+CluePedia. The 19 hub genes were mainly enriched in protein digestion and absorption, renin-angiotensin, vitamin digestion and absorption, cholesterol metabolism, and fat digestion and absorption (Figure 5b). To explore whether the 19 hub genes could be therapeutically targeted or prioritized for drug development, drug–gene relationships were predicted using the DGIdb2.0 database (Figure 5c). Among 19 hub genes, it was predicted that *CHRM4*, *HTR1D*, *DPP4*, *MTTP*, *ACE2*, *CASR*, and *MME* could be potential drug targets. Furthermore, there were 45 drug–gene

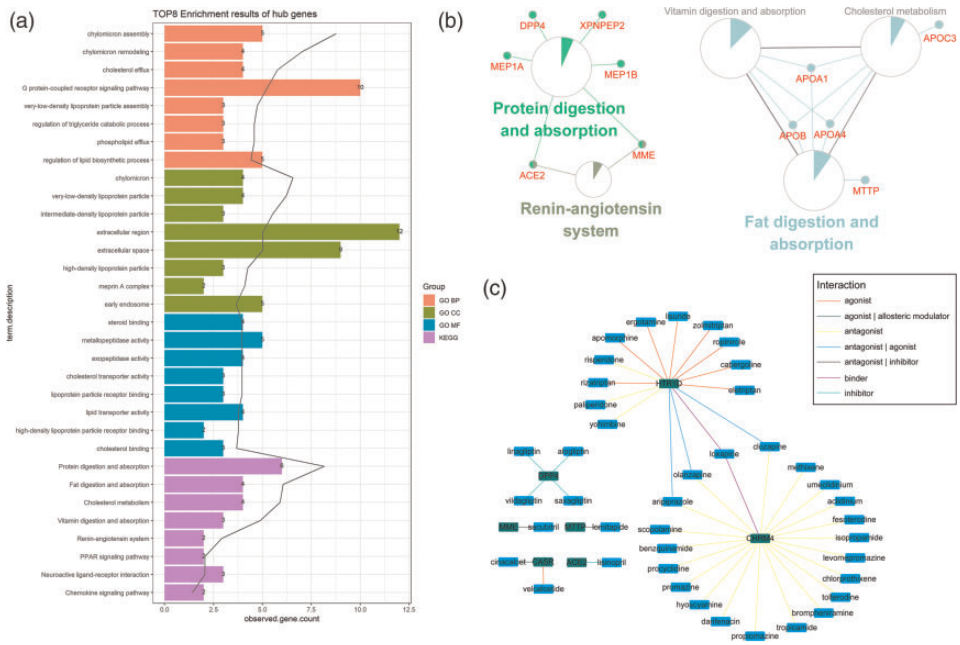

Figure 5. Function enrichment analysis of hub genes and drug-gene network construction. (a) The top eight terms of GO including BP, CC, and MF, KEGG, and Reactome using the String database. (b) Pathway enrichment analysis using ClueGo+CluePedia. (c) The construction of drug–gene network. GO, Gene Ontology; BB, biological process; CC, cellular component; MF, molecular function; KEGG, Kyoto Encyclopedia of Genes and Genomes.

relationships in the network. The *CHRM4* gene had the highest nodes, followed by *HTR1D*.

Expression of hub genes between the colon and ileum of IBD

We analyzed the expression of hub genes between the colon and ileum of IBD using the ggpubr package. We found that expression of *ACE2*, *APOA1*, *APOA4*, *APOB*, *APOC3*, *CASR*, *CCL25*, *CHRM4*, *DPP4*, *HEBP1*, *HTR1D*, *MEP1A*, *MEP1B*, *MME*, *MTTP*, and *XPNPEP2* were all significantly lower in the colon than in the ileum (Figure 6). Conversely, the expression of *GPER1*, *GRM8*, and *PPBP* were all higher in the colon than in the ileum. We further validated these hub genes in an independent dataset. We found that the expression patterns of these hub genes were similar to the above results (Figure 7). Moreover, we compared the expression differences of these hub genes in CD and UD. The results showed no statistically significant differences in expression of these hub genes between CD and UC (Figure 8).

Discussion

The pathogenesis of IBD is extremely complex, involving multiple pathways and cytokines.[22,23] The disease is unpredictable and incurable.[24] Therefore, understanding the mechanism underlying IBD would be helpful in finding the best treatment options as early as possible, which would improve long-term treatment results and quality of life in patients with IBD.[25,26]

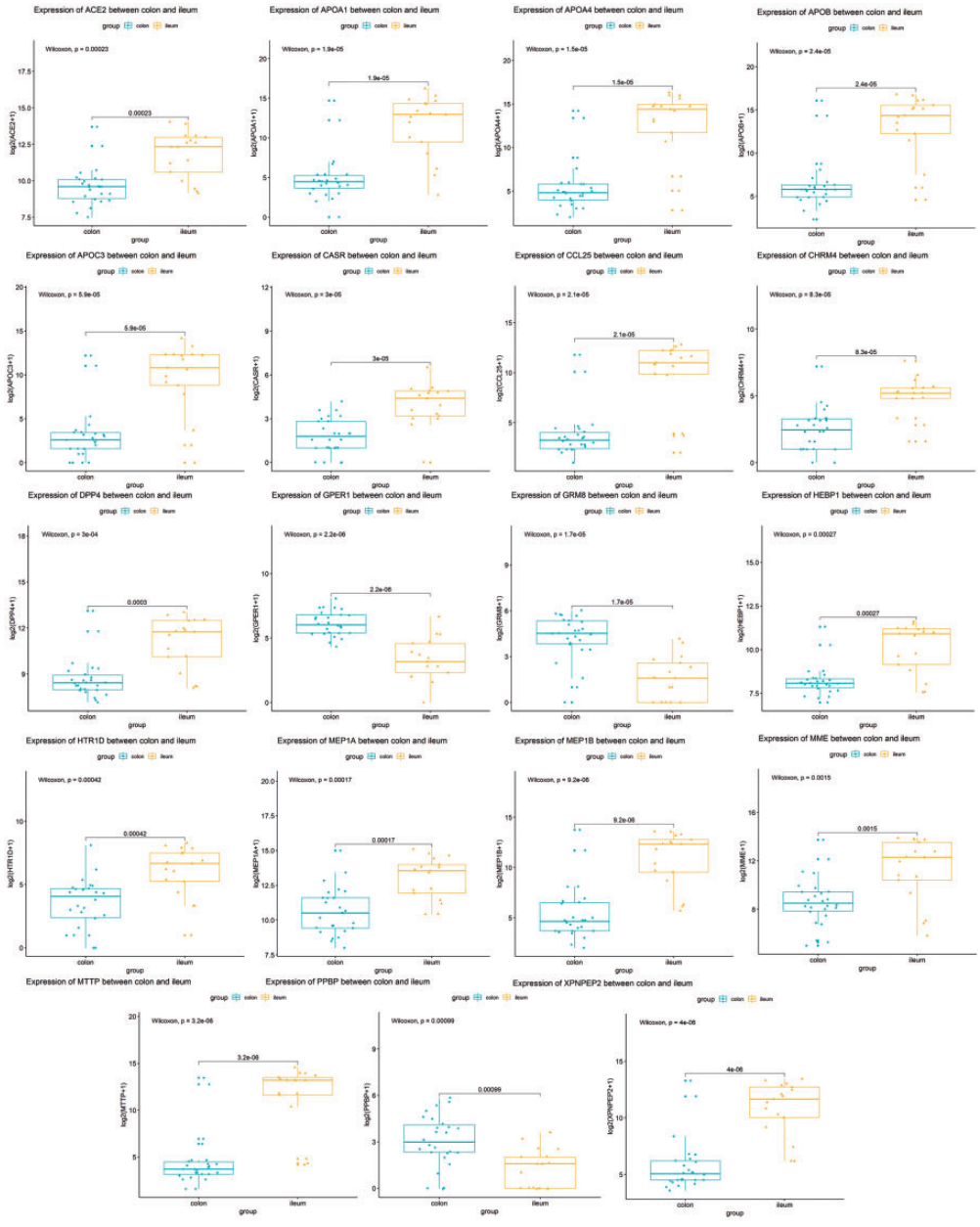

Figure 6. Boxplots showing the differential expression of hub genes between the colon and ileum of IBD. IBD, inflammatory bowel disease. The box shows the interquartile range, the line indicates the median, the whiskers indicate the minimum and maximum values, and symbols indicate the expression levels of genes.

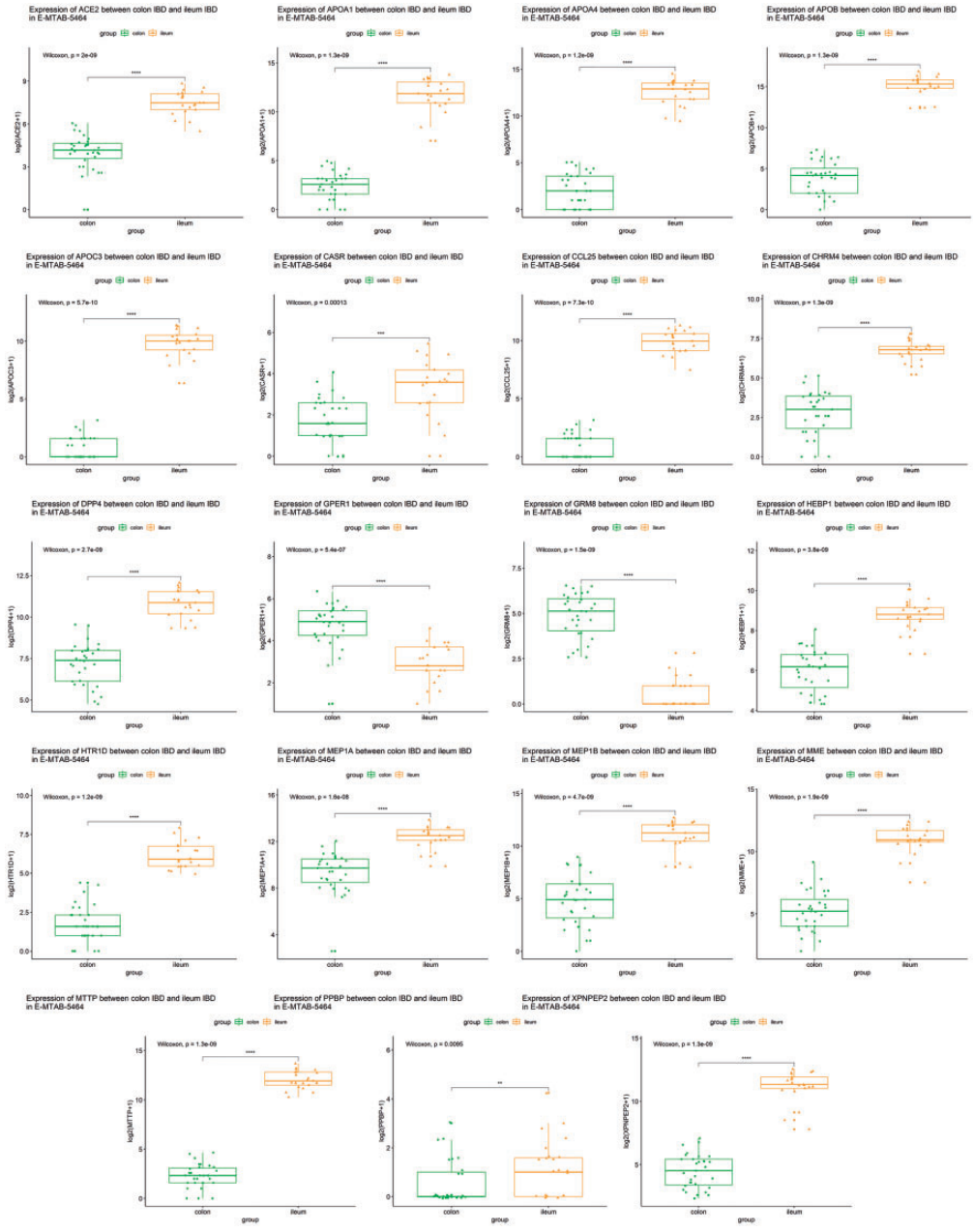

Figure 7. Validation of the differential expression of hub genes between the colon and ileum of IBD in an independent dataset. IBD, inflammatory bowel disease. The box shows the interquartile range, the line indicates the median, the whiskers indicate the minimum and maximum values, and symbols indicate the expression levels of genes.

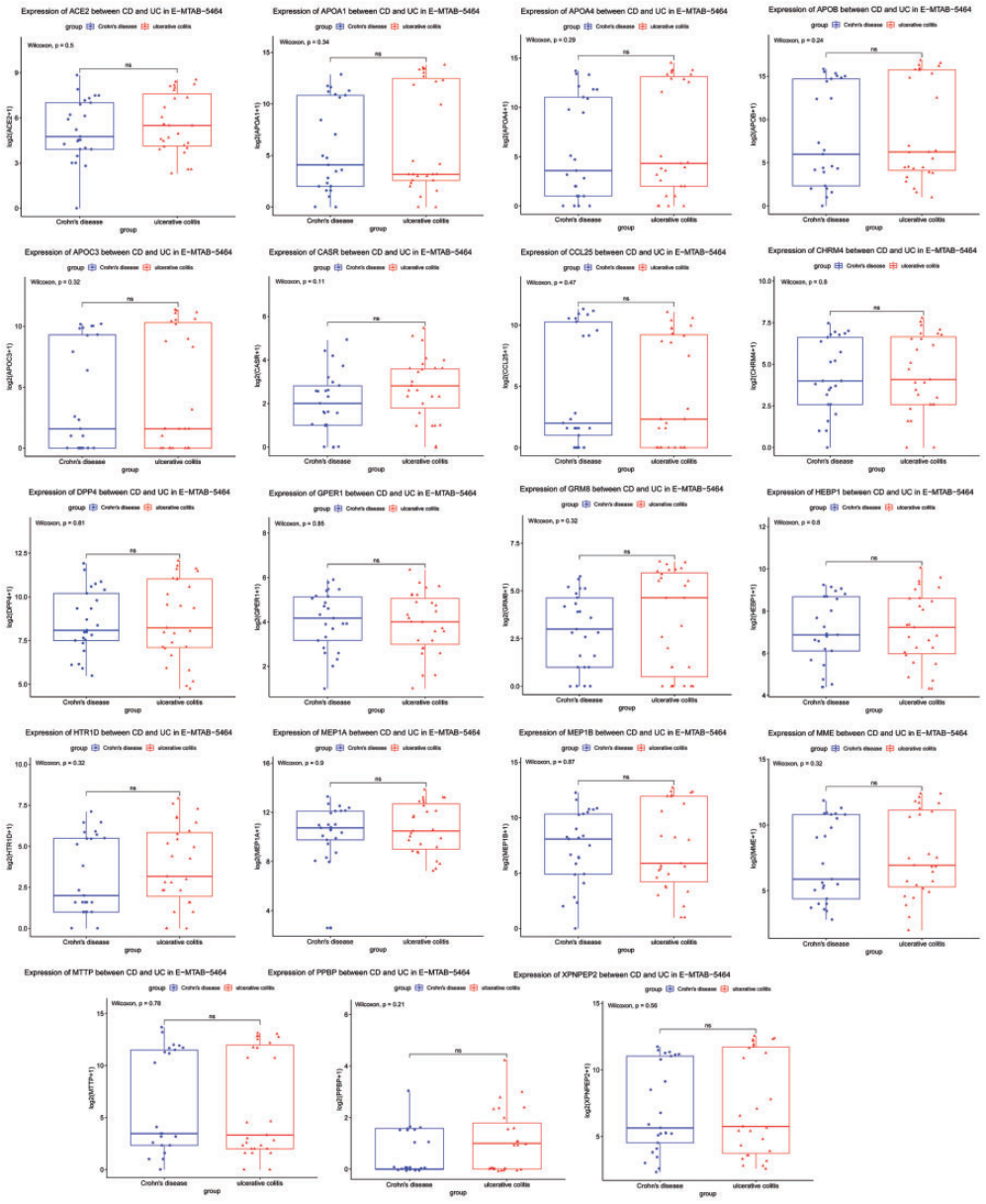

Figure 8. The differential expression of hub genes between samples from Crohn's disease and ulcerative colitis. CD, Crohn's disease; UC, ulcerative colitis. The box shows the interquartile range, the line indicates the median, the whiskers indicate the minimum and maximum values, and symbols indicate the expression levels of genes.

Gene expression patterns are altered in different inflammatory sites of IBD. Therefore, identification of DEG between the colon and ileum of IBD may help us understand the pathogenesis of IBD.

Traditional biological research is based on a single gene or protein, so it can only explain the local part of a biological system, and it is difficult to comprehensively explore the whole system. In contrast, WGCNA is an algorithm for mining module information from high-throughput data. In this method, a module is defined as a group of genes with similar expression profiles. If some genes have similar expression patterns within a physiological process or in different tissues, there is reason to believe that these genes are functionally related.[27] The clustering criteria of WGCNA have biological implications, such as the use of geometric distances between data.[28] Therefore, the results obtained using this method have higher credibility and are more conducive to identifying gene functions and internal associations from the whole biological function. In the present study, after constructing the co-expression network, we identified the green module as meaningful, because it had the highest correlation with different lesion sites of IBD. To identify DEG between the colon and ileum of IBD, we took the intersection of DEG and genes in the green module. Among these, we identified complex relationships according to the PPI network. After obtaining key genes from the PPI network, 19 hub genes were identified using the MCODE plug-in.

In this study, we focused on hub genes associated with lesion sites of IBD. Previous studies have found that mutations in *FUT3* (rs28362459 and rs3745635) are associated with lesion location in CD patients, suggesting that they may affect the lesion site of CD.[29] Additionally, *HEBP1* might participate in inflammation in IBD according to GWAS, which was consistent with our results.[30] The 19 hub genes were identified as sub-modules of PPI network, which were differentially expressed between the colon and ileum of IBD. The 19 hub genes were mainly involved in renin-angiotensin, vitamin digestion and absorption, and fat digestion and absorption, among others. These pathways are closely related to IBD. For example, the circulating component of the renin-angiotensin system is increased in patients with IBD.[31] In patients with IBD who are inhibited by angiotensin, hospitalization, surgery, and corticosteroid use are less frequent.[32] In the present study, of the 19 hub genes, *ACE2* is the main component of the renin-angiotensin system. We found that expression of *ACE2* was higher in the colon than in the ileum of patients with IBD. Drug–gene prediction results showed that *ACE2* could be a potential drug target. IBD is a group of idiopathic, chronic, and recurrent inflammatory conditions in the gastrointestinal tract. Vitamin D_3 supplementation in patients with IBD helps maintain the balance between inflammatory and inhibitory cytokines, which is important for disease management.[33,34] Available evidence suggests that IBD is a complex multifactorial disease in which immune disorders caused by genetic or environmental factors, or both, play an important role. Apolipoprotein plays an important role in cholesterol and lipid metabolism, which has been confirmed to alter innate and adaptive immune responses.[35] In the present study, we found that apolipoprotein A1 (*APOA1*), apolipoprotein A4 (*APOA4*), apolipoprotein B (*APOB*), and apolipoprotein C3 (*APOC3*) were associated with different lesion sites, including the colon and ileum, and were highly expressed in the colon. In the late stage of IBD, *APOA1* expression is elevated in serum. Antibiotic treatment can reduce the mean serum value of *APOA1*.[36] In patients with active IBD, infection and inflammation result in changes in lipids,

apolipoprotein, and lipoprotein profiles, and reduce cholesterol efflux, which may contribute to the development of cardiovascular diseases.[37] Functional enrichment analysis revealed that all of these genes participate in biological processes such as cholesterol metabolism and fat digestion and absorption. Previous studies have found a significant increase in chemokine CCL25 in patients with IBD; CCL25 promotes local intestinal inflammation and tissue damage in patients with IBD.[38,39] *MEP1A* is a susceptibility gene of UC, and is associated with intestinal inflammation.[40] Furthermore, the 19 hub genes could be predicted by the DGIdb2.0 database. Among the 19 hub genes, *CHRM4*, *HTR1D*, *DPP4*, *MTTP*, *ACE2*, *CASR*, and *MME* could be potential IBD drug targets, particularly *CHRM4* and *HTR1D*; these hub genes play a key role in the development of IBD.[41] We further validated these hub genes in an independent dataset. As expected, we found similar expression patterns of these hub genes between the colon and ileum of patients with IBD. However, one limitation of this study should be noted. Although this study was verified by an independent dataset, further experiments are required to validate our findings.

Conclusion

In the present study, we identified 19 DEG of IBD, which are regulated by different inflammatory sites. These genes participate in biological processes of IBD and could become potential drug targets. These 19 DEG might be important in the development of IBD.

Authors' contributions

Yong Xie conceived and designed the study; Yuting Zhang conducted most of the experiments and data analysis and wrote the manuscript; Bo Shen and Liya Zhuge participated in collecting data and helped to draft the manuscript. All authors reviewed and approved the manuscript.

Availability of data and material

The datasets analyzed during the current study are available from the corresponding author on reasonable request.

Declaration of conflicting interest

The authors declare that there is no conflict of interest.

Funding

This work was funded by National Key Research and Development Grogram of China (2016YFC1302201), the National Natural Science Foundation of China (81760105) and Special Funds of Graduate Student Innovation Project in Nanchang University (CX2016315).

ORCID iD

Yong Xie https://orcid.org/0000-0002-6880-4289

References

1. Aarestrup J, Jess T, Kobylecki CJ, et al. Cardiovascular risk profile among patients with inflammatory bowel disease: a population-based study of more than 100 000 individuals. *J Crohns Colitis* 2019; 13: 319–323.
2. Panes J, Jairath V and Levesque BG. Advances in use of endoscopy, radiology, and biomarkers to monitor inflammatory bowel diseases. *Gastroenterology* 2017; 152: 362–373.
3. Huang H, Fang M, Jostins L, et al. Fine-mapping inflammatory bowel disease loci to single-variant resolution. *Nature* 2017; 547: 173–178.
4. Teimoori-Toolabi L, Samadpoor S, Mehrtash A, et al. Among autophagy genes, ATG16L1 but not IRGM is associated with Crohn's disease in Iranians. *Gene* 2018; 675: 176–184.

5. Naganuma M, Aoyama N, Tada T, et al. Complete mucosal healing of distal lesions induced by twice-daily budesonide 2-mg foam promoted clinical remission of mild-to-moderate ulcerative colitis with distal active inflammation: double-blind, randomized study. *J Gastroenterol* 2018; 53: 494–506.
6. Ng SC, Shi HY, Hamidi N, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *Lancet* 2018; 390: 2769–2778.
7. Seyedabadi M, Rahimian R and Ghia JE. The role of alpha7 nicotinic acetylcholine receptors in inflammatory bowel disease: involvement of different cellular pathways. *Expert Opin Ther Targets* 2018; 22: 161–176.
8. de Lange KM, Moutsianas L, Lee JC, et al. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet* 2017; 49: 256–261.
9. Wojcik GL, Marie C, Abhyankar MM, et al. Genome-wide association study reveals genetic link between diarrhea-associated *Entamoeba histolytica* infection and inflammatory bowel disease. *MBio* 2018; 9: pii: e01668-18.
10. Hart AL, Lomer M, Verjee A, et al. What are the top 10 research questions in the treatment of inflammatory bowel disease? A priority setting partnership with the James Lind Alliance. *J Crohns Colitis* 2017; 11: 204–211.
11. Ghattamaneni NKR, Panchal SK and Brown L. An improved rat model for chronic inflammatory bowel disease. *Pharmacol Rep* 2019; 71: 149–155.
12. Marchioni Beery RM, Barnes EL, Nadkarni A, et al. Suicidal behavior among hospitalized adults with inflammatory bowel disease: a United States nationwide analysis. *Inflamm Bowel Dis* 2017; 24: 25–34.
13. Bogaert S, Laukens D, Peeters H, et al. Differential mucosal expression of Th17-related genes between the inflamed colon and ileum of patients with inflammatory bowel disease. *BMC Immunol* 2010; 11: 61.
14. Guo X, Xiao H, Guo S, et al. Identification of breast cancer mechanism based on weighted gene coexpression network analysis. *Cancer Gene Ther* 2017; 24: 333–341.
15. Brazma A, Parkinson H, Sarkans U, et al. ArrayExpress—a public repository for microarray gene expression data at the EBI. *Nucleic Acids Res* 2003; 31: 68–71.
16. Povey S, Lovering R, Bruford E, et al. The HUGO Gene Nomenclature Committee (HGNC). *Hum Genet* 2001; 109: 678.
17. Robinson MD, McCarthy DJ and Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010; 26: 139–140.
18. Langfelder P and Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008; 9: 559.
19. Mering CV, Huynen M, Jaeggi D, et al. STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res* 2003; 31: 258.
20. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003; 13: 2498–2504.
21. Wagner AH, Coffman AC, Ainscough BJ, et al. DGIdb 2.0: mining clinically relevant drug-gene interactions. *Nucleic Acids Res* 2016; 44: D1036–D1044.
22. Zheng X, Chen F, Zhang Q, et al. Salivary exosomal PSMA7: a promising biomarker of inflammatory bowel disease. *Protein Cell* 2017; 8: 686–695.
23. Click B, Anderson AM, Koutroubakis IE, et al. Peripheral eosinophilia in patients with inflammatory bowel disease defines an aggressive disease phenotype. *Am J Gastroenterol* 2017; 112: 1849–1858.
24. de Souza HSP. Etiopathogenesis of inflammatory bowel disease: today and tomorrow. *Curr Opin Gastroenterol* 2017; 33: 222–229.
25. Parlato M, Charbit-Henrion F, Pan J, et al. Human ALPI deficiency causes inflammatory bowel disease and highlights a key mechanism of gut homeostasis. *EMBO Mol Med* 2018; 10: pii: e8483.
26. Sandborn WJ, Hanauer S, Van Assche G, et al. Treating beyond symptoms with a view to improving patient outcomes in inflammatory bowel diseases. *J Crohns Colitis* 2014; 8: 927–935.

27. Li B, Pu K, Wu X, et al. Identifying novel biomarkers in hepatocellular carcinoma by weighted gene co-expression network analysis. *J Cell Biochem* 2019. DOI: 10.1002/jcb.28420.
28. Oros Klein K, Oualkacha K, Lafond MH, et al. Gene coexpression analyses differentiate networks associated with diverse cancers harboring TP53 missense or null mutations. *Front Genet* 2016; 7: 137.
29. Hu DY, Shao XX, Xu CL, et al. Associations of FUT2 and FUT3 gene polymorphisms with Crohn's disease in Chinese patients. *J Gastroenterol Hepatol* 2014; 29: 1778–1785.
30. Cagliani R, Pozzoli U, Forni D, et al. Crohn's disease loci are common targets of protozoa-driven selection. *Mol Biol Evol* 2013; 30: 1077–1087.
31. Garg M, Burrell LM, Velkoska E, et al. Upregulation of circulating components of the alternative renin-angiotensin system in inflammatory bowel disease: a pilot study. *J Renin Angiotensin Aldosterone Syst* 2015; 16: 559–569.
32. Jacobs JD, Wagner T, Gulotta G, et al. Impact of angiotensin II signaling blockade on clinical outcomes in patients with inflammatory bowel disease. *Dig Dis Sci* 2019; 64: 1938–1944.
33. O'Sullivan M. Vitamin D as a novel therapy in inflammatory bowel disease: new hope or false dawn? *Proc Nutr Soc* 2015; 74: 5–12.
34. Alhassan Mohammed H, Mirshafiey A, Vahedi H, et al. Immunoregulation of inflammatory and inhibitory cytokines by vitamin D3 in patients with inflammatory bowel diseases. *Scand J Immunol* 2017; 85: 386–394.
35. Al-Meghaiseeb ES, Al-Otaibi MM, Al-Robayan A, et al. Genetic association of apolipoprotein E polymorphisms with inflammatory bowel disease. *World J Gastroenterol* 2015; 21: 897–904.
36. Torrence AE, Brabb T, Viney JL, et al. Serum biomarkers in a mouse model of bacterial-induced inflammatory bowel disease. *Inflamm Bowel Dis* 2008; 14: 480–490.
37. Ripollés Piquer B, Nazih H, Bourreille A, et al. Altered lipid, apolipoprotein, and lipoprotein profiles in inflammatory bowel disease: consequences on the cholesterol efflux capacity of serum using Fu5AH cell system. *Metabolism* 2006; 55: 980–988.
38. Singh UP, Singh NP, Murphy EA, et al. Chemokine and cytokine levels in inflammatory bowel disease patients. *Cytokine* 2016; 77: 44–49.
39. Trivedi PJ, Bruns T, Ward S, et al. Intestinal CCL25 expression is increased in colitis and correlates with inflammatory activity. *J Autoimmun* 2016; 68: 98–104.
40. Banerjee S, Oneda B, Yap LM, et al. MEP1A allele for meprin A metalloprotease is a susceptibility gene for inflammatory bowel disease. *Mucosal Immunol* 2009; 2: 220–231.
41. Broedl UC, Schachinger V, Lingenhel A, et al. Apolipoprotein A-IV is an independent predictor of disease activity in patients with inflammatory bowel disease. *Inflamm Bowel Dis* 2007; 13: 391–397.