Research Article

Transcription Factors That Regulate the Pathogenesis of Ulcerative Colitis

Bing Zhang¹ and Tao Sun²

¹TCM Department, Hangzhou First People’s Hospital, China
²Hepatology Department, The Second Affiliated Hospital of Zhejiang Chinese Medical University, 318th Chaowang Road, Gongshu District, Zhejiang, 310005 Hangzhou, China

Correspondence should be addressed to Tao Sun; keweitaowy@163.com

Received 4 July 2020; Accepted 10 August 2020; Published 25 August 2020

Guest Editor: Tao Huang

Copyright © 2020 Bing Zhang and Tao Sun. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ulcerative colitis (UC) is one of the inflammatory bowel diseases (IBD) characterized by occurrence in the rectum and sigmoid colon of young adults. However, the functional roles of transcription factors (TFs) and their regulating target genes and pathways are not fully known in ulcerative colitis (UC). In this study, we collected gene expression data to identify differentially expressed TFs (DETFs). We found that differentially expressed genes (DEGs) were significantly enriched in the target genes of HOXA2, IKZF1, KLF2, XBP1, EGR2, ETV7, BACH2, CBFA2T3, HLF, and NFE2. TFs including BACH2, CBFA2T3, EGR2, ETV7, NFE2, and XBP1, and their target genes were significantly enriched in signaling by interleukins. BACH2 target genes were enriched in estrogen receptor- (ESR-) mediated signaling and nongenomic estrogen signaling. Furthermore, to clarify the functional roles of immune cells on the UC pathogenesis, we estimated the immune cell proportions in all the samples. The accumulated effector CD8 and reduced proportion of naïve CD4 might be responsible for the adaptive immune response in UC. The accumulation of plasma in UC might be associated with increased gut permeability. In summary, we present a systematic study of the TFs by analyzing the DETFs, their regulating target genes and pathways, and immune cells. These findings might improve our understanding of the TFs in the pathogenesis of UC.

1. Introduction

Ulcerative colitis (UC) is one of the inflammatory bowel diseases (IBD) characterized by occurrence in the rectum and sigmoid colon of young adults. However, the functional roles of transcription factors (TFs) and their regulating target genes and pathways are not fully known in ulcerative colitis (UC). In this study, we collected gene expression data to identify differentially expressed TFs (DETFs). We found that differentially expressed genes (DEGs) were significantly enriched in the target genes of HOXA2, IKZF1, KLF2, XBP1, EGR2, ETV7, BACH2, CBFA2T3, HLF, and NFE2. TFs including BACH2, CBFA2T3, EGR2, ETV7, NFE2, and XBP1, and their target genes were significantly enriched in signaling by interleukins. BACH2 target genes were enriched in estrogen receptor- (ESR-) mediated signaling and nongenomic estrogen signaling. Furthermore, to clarify the functional roles of immune cells on the UC pathogenesis, we estimated the immune cell proportions in all the samples. The accumulated effector CD8 and reduced proportion of naïve CD4 might be responsible for the adaptive immune response in UC. The accumulation of plasma in UC might be associated with increased gut permeability. In summary, we present a systematic study of the TFs by analyzing the DETFs, their regulating target genes and pathways, and immune cells. These findings might improve our understanding of the TFs in the pathogenesis of UC.
emerged as key regulators in several diseases [8, 9]. Heat shock transcription factor 2 could predict mucosal healing and promote mucosal repair by suppressing MAPK signaling and inhibit intestinal epithelial cell apoptosis in UC through the mitochondrial pathway [10, 11]. Moreover, RUNX3 is also associated with UC by regulating the immune-related target genes and pathways [12, 13]. However, there is a lack of systematic study analyzing the functional roles of TFs in the pathogenesis of UC. Therefore, we carried out the present study, aiming at identifying the critical TFs, their downstream target genes, and pathways involved in UC pathogenesis.

2. Materials and Methods

2.1. Datasets. The gene expression data were collected from the Gene Expression Omnibus (GEO) database with accession GSE128682, and the sample collection was described in an earlier study [14]. The counts for each sample were normalized by DESeq2 [15]. The pairs of transcription factor (TF) target genes were downloaded from three public databases including JASPAR [16], TRANSFAC [17], and CHEA [18].

2.2. Differential Expression Analysis. The count-based expression data were used for the differential expression analysis (DEA). R/Bioconductor DESeq2 [15] was employed to identify the differentially expressed genes (DEGs). The two-fold change and adjusted p value of 0.05 were used to determine the DEGs for each comparison.

2.3. Transcription Factor Target Genes and Pathway Enrichment Analysis. The Fisher’s exact test was used to identify the transcription factors (TFs) and pathways enriched by the DEGs. The DEGs with a significant correlation with their TFs were selected for this analysis and TFs with a large number of target genes (n > 2000) were excluded in the enrichment analysis. The enrichment analysis was implemented in the R clusterProfiler package with enricher function [19].

2.4. Immune Cell Proportion Estimation. The immune cell proportion was estimated by CIBERSORT, which used the gene expression profiles and immune cell-specific genes to characterize the cell composition of complex tissues [20]. The count-based expression data was normalized to Transcript Per Million (TPM) by R scater package (https://bioconductor.riken.jp/packages/3.4/bioc/html/scater.html), which was used for the CIBERSORT analysis.

2.5. Statistical Analysis. The two-sample comparison was tested by Wilcoxon rank-sum test or t test, and multiple-sample comparison was tested by analysis of variance (ANOVA). The Spearman correlation analysis was used to evaluate the correlation of two variables. Symbols of *, **, ***, and **** indicate the statistical significances of 0.05, 0.01, 0.001, and 0.0001, respectively.

3. Results

3.1. Identification of Differentially Expressed Transcription Factors. The mucosal biopsies had 14 ulcerative colitis (UC), 14 remission (R), and 16 healthy controls (N). With the three groups of mucosal biopsies, we compared one with the other two groups, respectively. UC had significantly different expression profiles as compared with R and N groups, with 3,202 and 2,517 differentially expressed genes (DEGs) in UC vs. N and UC vs. R (Supplementary Table S1), respectively. The comparison of R vs. N only identified 1,133 DEGs. Consistently, the comparisons of UC vs. N (n = 56) and UC vs. R (n = 46) had greater numbers of differentially expressed transcription factors (TFs) than that of R vs. N (n = 6) (Figure 1(a)). These results indicated that the transcriptomic profiles were significantly altered in UC samples as compared with samples of remission and healthy controls.

Totally, we identified 72 TFs significantly differentially expressed between the three groups (Supplementary Table S2). The hierarchical clustering analysis revealed that the UC samples could be clearly differentiated from the N and R samples by the TFs specifically upregulated in UC (Figure 1(b)). The TFs specifically upregulated in R and N samples also had the capability of classifying the two groups to some extent (Figure 1(b)). These results indicated that the TFs might be implicated in UC pathogenesis.

3.2. Expression Patterns of the Differentially Expressed Transcription Factors. To reveal the expression patterns of the differentially expressed transcription factors (DETFs), we conducted coexpression analysis of the 72 DETFs. Notably, four categories of DETFs (A, B, C, and D) could be identified by the coexpression analysis (Figure 2(a)). Further analysis of the expression patterns revealed that upregulated TFs in UC (N = R < UC) were highly enriched in groups A and C, upregulated TFs in R (N < R < UC) had higher proportion in group B, and downregulated TFs in UC (N = R > UC) were more frequently observed in group D (Figure 2(b)). These results indicated that three categories were observed in these DETFs.

3.3. Target Genes of the DETFs. As the TFs could promote or suppress the transcription of their target genes, we then investigated whether the target genes were also differentially expressed. Specifically, DEGs were significantly enriched in the target genes of HOXA2, IKZF1, KLF2, XBP1, EGR2, ETV7, BACH2, CBFA2T3, HLF, and NFE2 (Figure 3(a), Supplementary Table S3). Remarkably, BACH2, NFE2, IKZF1, EGR2, XBP1, CBFA2T3, and ETV7 were upregulated in UC (N = R < UC or N < R < UC), and HLF and HOXA2 were downregulated in UC (N = R > UC or N > R > UC) (Figure 3(b)). It should be noted that BACH2 and NFE2 were upregulated in UC (Figure 3(c)), and they had significantly more shared target genes (Figure 3(d)), suggesting that the two TFs might cooperate with each other to regulate their target genes.

3.4. Signaling Pathways That the DETFs and Target Genes May Participate in. To further identify the signaling
Figure 1: Differentially expressed transcription factors (DETFs). (a) The differential expression levels of the TFs. The X and Y axes represent the log2 fold change and -log10 (adjusted p value), respectively. (b) The gene expression profiles of the DETFs in UC, R, and N groups. The red and blue colors represent the high and low expression.
pathways regulated by the DETFs and target genes, we conducted a gene set enrichment analysis of the differentially expressed target genes of DETFs. We found that target genes of $BACH2$, $CBFA2T3$, $EGR2$, $ETV7$, $IKZF1$, $NFE2$, and $XBP1$ were significantly enriched in the pathways (Figure 4(a)). The virus infection pathways including human papillomavirus infection, Epstein-Barr virus infection, Hepatitis B, Kaposi sarcoma-associated herpesvirus infection, human...
immunodeficiency virus 1 infection, and immune-related pathways such as downstream signaling in naïve CD8+ T cells and signaling by interleukins were significantly enriched by these target genes (Figure 4(a)).

Particularly, TFs including BACH2, CBFA2T3, EGR2, ETV7, NFE2, and XBP1, and their target genes were significantly enriched in signaling by interleukins. The inflammatory factors such as IL6, IL18RAP, IL11, STAT5B, and CSF3...
were involved in the signaling by interleukins (Figure 4(b)). Furthermore, target genes of BACH2, including AKT3, GNGT2, MMP7, and MMP9, were involved in ESR-mediated signaling and non-genomic estrogen signaling. These results indicated that estrogen signaling and signaling by interleukins might be closely associated with the UC pathogenesis.

3.5. Immune Cells and Their Association with DETFs. As the inflammatory factors and pathways were potentially involved in UC pathogenesis, we investigated the relative abundance of immune cells in mucosal biopsies and their association with DETFs. The proportion of immune cells was estimated by CIBERSORT based on the gene expression profiles. Specifically, proportions of naive CD4, regulatory T cells (Tregs),
Figure 5: The association of transcription factors (TFs) with immune cell proportions. (a) The proportion of immune cells that were significantly accumulated or reduced in UC. (b) The Spearman correlation of TFs and immune cell proportions. The red and blue colors represent the positive and negative correlation.
and plasmacytoid dendritic cells (pDC) were decreased in UC, while effector CD8 and plasma were increased in UC compared with R and N groups (Figure 5(a)). Particularly, DC was found to be reduced in the R group (Figure 5(a)). The correlation analysis revealed that the nine DETFs with functional enrichment of pathways including BACH2, CBF1A2T3, EGR2, ETV7, NFE2, and XBP1 were highly correlated with effector CD8 and plasma (Figure 5(b)), indicating that these TFs might promote the infiltration of effector CD8 and plasma into the intestinal mucosal tissues.

4. Discussion

Transcription factors (TFs) are key proteins involved in regulating gene transcription in cells. However, the functional roles of TFs and their regulating target genes and pathways are still little known in ulcerative colitis (UC).

In the present study, we collected gene expression data of mucosal biopsies from 14 UC, 14 remission (R), and 16 healthy controls (N), and identified DEGs in the three groups, of which, 72 were identified as differentially expressed TFs (DETFs). Furthermore, the coexpression analysis of the DETFs revealed three categories of TFs, which were upregulated in UC (N = R < UC), upregulated in R (N < R < UC), and downregulated in UC (N = R > UC).

As the function of DETFs could result in dysregulation of their target genes, we found that DEGs were significantly enriched in the target genes of HOXA2, IKZF1, KLF2, XBP1, EGR2, ETV7, BACH2, CBFA2T3, HLF, and NFE2. As BACH2 and NFE2 proteins had similar protein structure [21], they had a greater number of shared target genes. BACH2 has interactions with NFE2L1 and NFE2L3 based on BIOGRID [22] protein–protein interaction (PPI), indicating that BACH2 might also have the potential to interact with NFE2. Both BACH2 and NFE2 were implicated in UC via regulating inflammation-related pathways [23, 24].

Among the TF target genes, inflammatory factors such as IL6, IL1B, IL17A, IL18RAP, IL11, STAT5B, and CSF3 were involved in the signaling by interleukins. The interleukin and receptors were frequently reported to promote the inflammatory phenotype in UC [25–28]. Notably, IL11 and IL18RAP were identified as susceptibility loci in UC [29, 30]. Furthermore, target genes of BACH2, including AKT3, NGN2T2, MMP7, and MMP9, were involved in ESR-mediated signaling and nongenomic estrogen signaling. As patients with UC have a higher risk for colorectal carcinoma (CRC) development [31] and the estrogen receptors (ER) alpha/beta balance has a relevant influence on colorectal carcinogenesis [32], we then speculated that the dysregulation of estrogen signaling might be associated with the risk of colorectal carcinogenesis.

Furthermore, to clarify the functional roles of immune cells on the UC pathogenesis, we estimated the immune cell proportions in all the samples. The accumulated effector CD8 and reduced proportion of naïve CD4 might be responsible for the adaptive immune response in UC, showing consistency with the previous study [33]. Notably, BACH2 and EGR2 could regulate CD8 cell differentiation, indicating that the high proportion of CD8 might be associated with the upregulation of BACH2 and EGR2 [34, 35]. The accumulation of plasma in UC might be associated with increased gut permeability [36].

In summary, we present a systematic study of the TFs by analyzing the DETFs, their regulating target genes and pathways, and immune cells. These findings might improve our understanding of the TFs in the pathogenesis of UC.

Data Availability

The gene expression data were collected from the Gene Expression Omnibus (GEO) database with accession GSE128682.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This manuscript is funded by National Natural Science Foundation of China (No. 81904152).

Supplementary Materials

Supplementary 1. Supplementary Table S1: UC had significantly different expression profiles as compared with R and N groups.

Supplementary 2. Supplementary Table S2: 72 TFs significantly differentially expressed between the three groups.

Supplementary 3. Supplementary Table S3: DEGs were significantly enriched in the target genes.

References

[1] M. Fumery, D. Duricova, C. Gower-Rousseau, V. Annese, L. Peyrin-Biroulet, and P. L. Lakatos, “Review article: the natural history of paediatric-onset ulcerative colitis in population-based studies,” *Alimentary Pharmacology & Therapeutics*, vol. 43, no. 3, pp. 346–355, 2016.

[2] Y. Yokoyama, T. Yamakawa, T. Hirano et al., “Current diagnostic and therapeutic approaches to cytomegalovirus infections in ulcerative colitis patients based on clinical and basic research data,” *International Journal of Molecular Sciences*, vol. 21, no. 7, p. 2438, 2020.

[3] R. M. Nanau and M. G. Neuman, “Metabolome and inflammation in inflammatory bowel disease,” *Translational Research*, vol. 160, no. 1, pp. 1–28, 2012.

[4] T. Tanaka, Y. Kobunai, Y. Yamamoto et al., “Increased copy number variation of mtDNA in an array-based digital PCR assay predicts ulcerative colitis-associated colorectal cancer,” *In Vivo*, vol. 31, no. 4, pp. 713–718, 2017.

[5] K. Kang, J. H. Bae, K. Han, E. S. Kim, Kim TO, and J. M. Yi, “A genome-wide methylation approach identifies a new hypermethylated gene panel in ulcerative colitis,” *International Journal of Molecular Sciences*, vol. 17, no. 8, p. 1291, 2016.

[6] V. Lorén, A. Garcia-Jaraquemada, J. E. Naves et al., “ANP32E, a protein involved in steroid-refractoriness in ulcerative colitis, identified by a systems biology approach,” *Journal of Crohn’s & Colitis*, vol. 13, no. 3, pp. 351–361, 2019.
[7] K. Dubois-Camacho, D. Diaz-Jimenez, M. de la Fuente et al., “Inhibition of miR-378a-3p by inflammation enhances IL-33 levels: a novel mechanism of alarmin modulation in ulcerative colitis,” *Frontiers in Immunology*, vol. 10, p. 2449, 2019.

[8] R. van der Lee, S. Correard, and W. W. Wasserman, “Deregulated regulators: disease-causing cis variants in transcription factor genes,” *Trends in Genetics*, vol. 36, no. 7, pp. 523–539, 2020.

[9] A. Chen and A. N. Koehler, “Transcription factor inhibition: lessons learned and emerging targets,” *Trends in Molecular Medicine*, vol. 26, no. 5, pp. 508–518, 2020.

[10] Y. Wen, J. Niu, F. Zhang et al., “Heat shock transcription factor 2 predicts mucosal healing and promotes mucosal repair of ulcerative colitis,” *Scandinavian Journal of Gastroenterology*, vol. 55, no. 6, pp. 677–686, 2020.

[11] W. Wang, F. Zhang, X. Li et al., “Heat shock transcription factor 2 inhibits intestinal epithelial cell apoptosis through the mitochondrial pathway in ulcerative colitis,” *Biochemical and Biophysical Research Communications*, vol. 527, no. 1, pp. 173–179, 2020.

[12] R. K. Weersma, L. Zhou, I. M. Nolte et al., “Runt-related transcription factor 3 is associated with ulcerative colitis and shows epistasis with solute carrier family 22, members 4 and 5,” *Inflammatory Bowel Diseases*, vol. 14, no. 12, pp. 1615–1622, 2008.

[13] C. Guo, F. Yao, K. Wu, L. Yang, X. Zhang, and J. Ding, “Chromatin immunoprecipitation and association study revealed a possible role of runt-related transcription factor 3 in the ulcerative colitis of Chinese population,” *Clinical Immunology*, vol. 135, no. 3, pp. 483–489, 2010.

[14] C. G. Fenton, H. Taman, J. Florholmen, S. W. Sorbye, and R. H. Paulsens, “Transcriptional signatures that define ulcerative colitis in remission,” *Inflammatory Bowel Diseases*, 2020.

[15] M. I. Love, W. Huber, and S. Anders, “Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2,” *Genome Biology*, vol. 15, no. 12, article 550, 2014.

[16] O. Fornes, J. A. Castro-Mondragon, A. Khan et al., “JASPAR 2020: update of the open-access database of transcription factor binding profiles,” *Nucleic Acids Research*, vol. 48, no. D1, pp. D87–D92, 2020.

[17] E. Wingender, “The TRANSFAC project as an example of framework technology that supports the analysis of genomic regulation,” *Briefings in Bioinformatics*, vol. 9, no. 4, pp. 326–332, 2008.

[18] A. Lachmann, H. Xu, J. Krishnan, S. I. Berger, A. R. Mazloom, and A. Ma’ayan, “ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments,” *Bioinformatics*, vol. 26, no. 19, pp. 2438–2444, 2010.

[19] G. Yu, L. G. Wang, Y. Han, and Q. Y. He, “ClusterProfiler: an R package for comparing biological themes among gene clusters,” *OMICs*, vol. 16, no. 5, pp. 284–287, 2012.

[20] A. M. Newman, C. L. Liu, M. R. Green et al., “Robust enumeration of cell subsets from tissue expression profiles,” *Nature Methods*, vol. 12, no. 5, pp. 453–457, 2015.

[21] F. Katsouka and M. Yamamoto, “Small Maf proteins (MafF, MafG, MafK): history, structure and function,” *Gene*, vol. 586, no. 2, pp. 197–205, 2016.

[22] R. Oughtred, C. Stark, B. J. Breitkreutz et al., “The BioGRID interaction database: 2019 update,” *Nucleic Acids Research*, vol. 47, no. D1, pp. D529–D541, 2019.

[23] K. Christodoulou, A. E. Wiskin, J. Gibson et al., “Next-generation exome sequencing of paediatric inflammatory bowel disease patients identifies rare and novel variants in candidate genes,” *Gut*, vol. 62, no. 7, pp. 977–984, 2013.

[24] Y. D. Jeon, J. H. Lee, Y. M. Lee, and D. K. Kim, “Puerarin inhibits inflammation and oxidative stress in dextran sulfate sodium-induced colitis mice model,” *Biomedicine & Pharmacotherapy*, vol. 124, article 109847, 2020.

[25] Y. Yan, N. Zhao, X. He, H. Guo, Z. Zhang, and T. Liu, “Mesenchymal stem cell expression of interleukin-35 protects against ulcerative colitis by suppressing mucosal immune responses,” *Cytotherapy*, vol. 20, no. 7, pp. 911–918, 2018.

[26] C. A. Parisinos, S. Sergiouli, M. Katsoulis et al., “Variation in interleukin 6 receptor gene associates with risk of Crohn’s disease and ulcerative colitis,” *Gastroenterology*, vol. 155, no. 2, pp. 303–306.e2, 2018, e302.

[27] M. Matusiewicz, K. Neubauer, I. Bednarz-Misa, S. Gorska, and M. Krzystek-Korpacka, “Systemic interleukin-9 in inflammatory bowel disease: association with mucosal healing in ulcerative colitis,” *World Journal of Gastroenterology*, vol. 23, no. 22, pp. 4039–4046, 2017.

[28] I. A. Sroufe, T. Gardner, K. A. Bresnahan, S. M. Quarnberg, and P. R. Wiedmeier, “Insights into the pathophysiology of ulcerative colitis: interleukin-13 modulates STAT6 and p38 MAPK activity in the colon epithelial sodium channel,” *The Journal of Physiology*, vol. 595, no. 2, pp. 421–422, 2017.

[29] A. Zhermakova, E. M. Festen, L. Franke et al., “Genetic analysis of innate immunity in Crohn’s disease and ulcerative colitis identifies two susceptibility loci harboring CARD9 and IL18RAP,” *American Journal of Human Genetics*, vol. 82, no. 5, pp. 1202–1210, 2008.

[30] W. Klein, A. Tromm, T. Griga et al., “A polymorphism in the IL11 gene is associated with ulcerative colitis,” *Genes and Immunity*, vol. 3, no. 8, pp. 494–496, 2002.

[31] M. Xue, L. Shi, W. Wang, S. Chen, and L. Wang, “An overview of molecular profiles in ulcerative colitis-related cancer,” *Inflammatory Bowel Diseases*, vol. 24, no. 9, pp. 1883–1894, 2018.

[32] M. Principi, M. Barone, M. Pricci et al., “Ulcerative colitis: from inflammation to cancer. Do estrogen receptors have a role?”, *World Journal of Gastroenterology*, vol. 20, no. 33, pp. 11496–11504, 2014.

[33] H. Rabe, M. Malmquist, C. Barkman et al., “Distinct patterns of naive, activated and memory T and B cells in blood of patients with ulcerative colitis or Crohn’s disease,” *Clinical and Experimental Immunology*, vol. 197, no. 1, pp. 111–129, 2019.

[34] R. Roychoudhuri, D. Clever, P. Li et al., “JASPAR CD8(+) T cell differentiation by controlling access of AP-1 factors to enhancers,” *Nature Immunology*, vol. 17, no. 7, pp. 851–860, 2016.

[35] N. Du, H. Kwon, P. Li et al., “EGR2 is critical for peripheral naïve T-cell differentiation and the T-cell response to influenza,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 46, pp. 16484–16489, 2014.

[36] P. P. Trivedi and G. B. Jena, “Role of α-lipoic acid in dextran sulfate sodium-induced ulcerative colitis in mice: studies on inflammation, oxidative stress, DNA damage and fibrosis,” *Food and Chemical Toxicology*, vol. 59, pp. 339–355, 2013.