Integrated Gene Expression Profiling Analysis Reveals Probable Molecular Mechanism and Candidate Biomarker in Anti-TNFα Non-Response IBD Patients

Yifan Liu
Yantao Duan
Yousheng Li
Department of General Surgery, Shanghai Ninth People’s Hospital, Shanghai JiaoTong University School of Medicine, Shanghai 200011, People’s Republic of China

Purpose: To explore the molecular mechanism and search for candidate biomarkers in the gene expression profile of IBD patients associated with the response to anti-TNFα agents.

Methods: Differentially expressed genes (DEGs) of response vs non-response IBD patients in datasets GSE12251, GSE16879, and GSE23597 were integrated using NetworkAnalyst. We conducted functional enrichment analysis of Gene Ontology and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and extracted hub genes from the protein–protein interaction network. The proportion of immune cell types was estimated via CIBERSORT. ROC curve analysis and binomial Lasso regression were applied to assess the expression level of hub genes in datasets GSE12251, GSE16879, and GSE23597, and another two datasets GSE107865 and GSE42296.

Results: A total of 287 DEGs were obtained from the integrated dataset. They were enriched in 14 Gene Ontology terms and 11 KEGG pathways. Polarization from M2 to M1 macrophages was relatively high in non-response individuals. We found nine hub genes (TLR4, TLR1, TLR8, CCR1, CD86, CCL4, HCK, and FCGR2A), mainly related to the interaction between Toll-like Receptor (TLR) pathway and FcγR signaling in non-response anti-TNFα individuals. FCGR2A, HCK, TLR1, TLR4, TLR8, and CCL4 show great value for prediction in intestinal tissue. Besides, FCGR2A, HCK, and TLR8 might be candidate blood biomarkers of anti-TNFα non-response IBD patients.

Conclusion: Over-activated interaction between FcγR-TLR axis in the innate immune cells of IBD patients might be used to identify non-response individuals and increased our understanding of resistance to anti-TNFα therapy.

Keywords: differentially expressed genes, inflammatory bowel disease, toll-like receptor pathway, FcγR signaling, anti-TNFα therapy

Introduction

Crohn’s disease (CD) and ulcerative colitis (UC), the two common forms of inflammatory bowel disease (IBD), are the chronic and relapsing inflammatory disorder of the intestine.1 The diseases can be of great importance to the affected persons, reducing the quality of life, as well as for the society, losing labor due to sickness, treatment, and medicine.1 Millions of people in North American and Europe are affected, as well as a rapid increase in newly industrialized countries.2

While classified as separate diseases, common driving disease mechanisms have been highlighted, such as the shared genetics risk factors, chronic inappropriate immune responses pathological features as well as joint involvement of various
biological pathways.\textsuperscript{1,3,4} Thus, several similar attempts have been made to treat this disease, monoclonal antibodies to anti-tumor necrosis factor α (anti-TNFα) have been one of the dominant strategies.\textsuperscript{5} However, patients must often be treated for an extended period to determine if the chosen drug is efficacious as a lack of accuracy in predicting the response of anti-TNFα agents. Meanwhile, approximately one-third of the patients benefit minimally or not at all from the treatment.\textsuperscript{5,6}

This arouses a question about the molecular events that facilitate the differential response of anti-TNFα therapy. Firstly, the Fab region of anti-TNFα agents neutralizes the biological activity of soluble TNFα and induct the apoptosis of the immune cell.\textsuperscript{7} Secondly, the Fc region of anti-TNFα agents induces Antibody-dependent cellular cytotoxicity (ADCC), complement-dependent cytotoxicity (CDC), and FcγR-mediated induction of wound healing macrophages.\textsuperscript{7,8} Thirdly, the synthesis of anti-drug antibodies (ADAs) and proteolytic degradation of matrix metalloproteinase (MMP) against anti-TNFα agents may decrease drug concentration.\textsuperscript{6,9} Hence, Anti-TNFα agents may have one or more known or unknown mechanisms of taking action as well as avoiding clearance. Interestingly, some of its underlying mechanisms are on a degree of overlapping or opposite to each other.\textsuperscript{8,10} Thus, a comprehensive comparison via genetics, gene expression data, and microbe provides an excellent to investigate the molecular mechanisms of resistance to anti-TNFα agents and predictive biomarkers.\textsuperscript{11–17}

So far, published studies using gene expression data from intestinal or blood samples of IBD patients collected before anti-TNFα therapy have identified several signature patterns of non-response patients.\textsuperscript{11–16} However, it is still difficult to predict the response of anti-TNFα therapy.\textsuperscript{16} We hypothesized that shared baseline molecular patterns may be associated with the clinical efficacy of anti-TNFα agents in IBD patients. The shared patterns in anti-TNFα non-response patients may enable its use as a predictive factor or even a novel target.

**Method**

**Microarray Data Extraction and Processing**

Six transcription datasets under accession GSE12251, GSE16879, GSE23597, GSE92415, GSE107865, and GSE42296 were downloaded from the Gene expression omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/). GSE12251, GSE16879, GSE23597, and GSE92415 discovered the gene expression profile of intestinal tissue samples between responses and non-response IBD patients before anti-TNFα therapy.\textsuperscript{11–13,15} GSE107865 and GSE42296 explored the blood gene expression profile of IBD patients.\textsuperscript{14,16} The raw expression data underwent background correction, normalization, and summarization using the robust multi-array average (RMA) algorithm in oligo.\textsuperscript{19,20}

**Integration of Microarray Datasets and DEGs Screening**

We integrated six cohorts of microarray datasets from three independent studies using NetworkAnalyst, a web interface for integrative statistical and visualizing tool.\textsuperscript{21} In the option of “multiple gene expression data” for the web interface, all processed expression data attaching with response status were uploaded. Following the integrity check of all uploaded data, the ComBat procedures implemented in the NetworkAnalyst tool were used to reduce potential study-specific batch effects.\textsuperscript{22} Principal component analysis (PCA) was applied to visualize and compare the sample clustering patterns before and after batch effect adjustment. Therefore, the efficiency of batch effect removal is estimated effective size (ES) is a well-established approach to integrate different datasets for increased statistical power and cross-study validation. The fixed model was selected in the current study to identify DEGs between response and non-response patients.\textsuperscript{23} In this analysis, combined P-value <0.05 and fold change >0.85 were used as the cutoff criteria.

**Functional Enrichment DEGs**

We performed the Gene Ontology (GO) analysis of BP, MF, and CC enrichment analysis of DEGs through integrated Discovery (DAVID).\textsuperscript{24,25} Then, the bubble chart showing P-value, Fold Enrichment, and gene counts were used to describe essential GO terms. Later, the Clue-go plug-in app of Cytoscape was utilized to visualize KEGG pathways through DAVID.\textsuperscript{26} We set the criteria as gene counts>5 and P-value<0.05.

**Verification and Supplement of DEGs Enrichment Result**

We used the Gene Set Enrichment Analysis (GSEA) to verify and supplement the result acquired by DEG enrichment analysis. GSEA is relative robustness to noise and outliers in the data. GSEA was performed using the
Composition of Immune Cells
We estimated the proportion of immune cell types of intestinal tissue from gene expression profile using CIBERSORT, an accurate and robust algorithm web tool for calculating the cell composition. We uploaded the gene expression data with response status to the CIBERSORT web (http://cibersort.stanford.edu/), choosing included LM22 signature and 500 permutations. Cases with a CIBERSORT output of p < 0.05, meaning that the inferred fractions of immune cell populations produced by CIBERSORT are accurate. For each sample, the calculated result of CIBERSORT output was normalized to 1 for further comparison.

PPI Network Analysis and Hub Gene Searching
We applied Search Tool for the Retrieval of Interacting Genes (STRING) to predict the functional protein interactions of DEGs. Cytoscape was used to compute the properties of the PPI network and for visualization. We extract the most highly connected cluster from the PPI network through MCODE analysis. The genes with high Maximal Clique Centrality (MCC) score calculated by CytoHubba in the extracted network were considered to be hub genes.

ROC Curve Analysis
The glmnet package (binomial Lasso) in R software was used to calculate and compression of linear models and preserve valuable variables. The expression level of hub genes and the status of responsiveness from eight cohorts (six cohorts form intestinal tissue and two cohorts from serum samples) were obtained. ROC analysis in PRISM 8.0 was used to count the area under the curve (AUC) and draw the ROC curves for each cohort separately. Thus, we investigated the ability of hub genes in prediction though the AUC value.

Result
Selection of Discovery and Verified Cohort
A total of six microarray datasets (GEO) were retrieved from Pubmed for anti-TNFα therapy IBD patients, followed by RMA normalized. GSE12251 discovered the gene expression profile of intestinal tissue samples between 12 responses to 10 non-response UC patients before the treatment of infliximab. GSE16879 detected the gene expression profile of 28 responses to 33 non-response IBD patients. GSE23597 analyzed the gene expression profile of 24 responses to 5 non-response UC patients. GSE92415 compare the gene expression profile of 32 responses to 27 non-response golimumab treated UC patient. GSE107865 and GSE42296 explored the blood gene expression profile 14 responses to 6 non-response or 17 responses to 5 non-response CD patient. Patient samples of GSE16879 were distinguished as three cohorts, including UC, CDi, CDC. GSE23597 is excluded for some of its patients is overlapped with GSE16879 in the process of data integration. Collectively, five cohorts from intestinal tissue (GSE12251, GSE16879UC, GSE16879CDi, GSE16879CDC, GSE92415) were selected as discovery cohort, which as eligible microarray datasets from integration. Six cohorts (GSE12251, GSE16879UC, GSE16879CDi, GSE16879CDC, GSE92415, GSE23597) from intestinal tissue and two cohorts from blood sample (GSE107865 and GSE42296) were chosen as verified cohorts. Detailed information for each dataset, including Cohort name, Sample source, IBD type, Definition time, Definition method, Anti-TNFα therapy, Purpose in the current study, and GEO accession number, is present in Table 1 (also in Table S1).

Screening of DEGs Between Response and Non-Response Patients
To identify the shared transcriptional signatures between response and non-response patients, we uploaded the data from the discovery cohort in succession to NetworkAnalyst. All uploaded data passed the integrity check. We performed “ComBat procedures” implemented in the NetworkAnalyst tool to reduce batch effects among different datasets. No apparent batch effects are found after this step, and the sample clustering patterns with or without batch effect adjustment were visualized by PCA plots (Figure 1A and B). All the five cohorts from the verified cohort were then integrated utilize ES statistical methods provided by NetworkAnalyst, which is the difference between two group means divided by standard deviation facilitate to reveal the DEGs between anti-TNFα agent response and non-response patients across different microarray datasets. Finally, a total of 288 DEGs between non-response and response IBD patients were obtained with the criteria of P-value <0.05 and log2(combined-fold change)>0.85. The 274 up-regulated genes and the 14 down-regulated genes were shown in the heat map as well as the volcano plot (Figures 1C and 2A, and Table S2).
Functional Annotation and Enrichment of
DEGs

Functional enrichment analyses were performed using DAVID to have further insight into the biological functions of DEGs. Considering the low count number of down-regulated DEGs, we focused on the up-regulated DEGs. For the GO analysis, "positive regulation of tumor necrosis factor production" was the most significant BP term with fold enrichment. We show the top 14 BP terms according to the P-value, gene counts, and fold enrichment (Figure 2B). These terms indicated that the gene expression differences of intestinal samples in non-response IBD patients were associated with an abundance of inflammatory cytokines and pathogen recognition. On the categories of Molecular function (MF), IgG binding, lipopolysaccharide-binding, and pattern recognition receptor activity were the prominent MF term. Intrinsic to membrane and integral to membrane were top 2 CC terms. We listed useful information of these GO terms in (Table 2). KEGG pathway enrichment analysis discovers 11 significantly enriched KEGG pathways with P-value <0.05 (Table 3). Among these pathways, the Toll-like receptor signaling pathway was the most important one according to the P-value, gene counts, FDR, and fold enrichment. The DEGs in the Toll-like receptor signaling pathway were concentrated in TLR (Figure S1). There were 11 DEGs genes that participated in this pathway (CD86, IL6, LY96, TLR1, TLR2, IL1B, TLR4, CXCL11, CCL4, TLR8, CD14, and SPP1). We visualize KEGG pathways by the ClueGO plug-in of Cytoscape v.3.4.0 (Figure 2C).

Verification and Supplement of DEGs

Enrichment Result
IL6 was significant differentially expressed between responders and non-responders (up-regulated in non-responders, adjusted P value<0.01), no significant difference was found in TNF and NOD2 (Figure 3A). IL6 was involved in the DEGs, while TNF and NOD2 were not included in the DEGs. GSEA was applied to verify the main result acquired by DEG enrichment analysis. GSEA result suggested upregulation of “inflammation response”, “Fc-gamma receptor

Table 1 Summary of Individual Studies Included

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Sample Source</th>
<th>IBD Type</th>
<th>Response Status</th>
<th>Definition Time</th>
<th>Definition Method</th>
<th>Anti-TNF Therapy</th>
<th>Purpose in the Current Study</th>
<th>Associated Publication</th>
<th>GEO Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE12251</td>
<td>Intestinal</td>
<td>UC</td>
<td>12</td>
<td>8 week</td>
<td>Endoscopic</td>
<td>Infliximab</td>
<td>Discovery and Validation</td>
<td>15</td>
<td>GSE12252</td>
</tr>
<tr>
<td>GSE16869Dc</td>
<td>Intestinal</td>
<td>CD</td>
<td>12</td>
<td>4–6 week</td>
<td>Endoscopic</td>
<td>Infliximab</td>
<td>Discovery and Validation</td>
<td>11</td>
<td>GSE16869</td>
</tr>
<tr>
<td>GSE16879Di</td>
<td>Intestinal</td>
<td>CD</td>
<td>12</td>
<td>4–6 week</td>
<td>Endoscopic</td>
<td>Infliximab</td>
<td>Discovery and Validation</td>
<td>11</td>
<td>GSE16870</td>
</tr>
<tr>
<td>GSE16879UC</td>
<td>Intestinal</td>
<td>UC</td>
<td>12</td>
<td>4–6 week</td>
<td>Endoscopic</td>
<td>Infliximab</td>
<td>Discovery and Validation</td>
<td>11</td>
<td>GSE16871</td>
</tr>
<tr>
<td>GSE92415</td>
<td>Intestinal</td>
<td>UC</td>
<td>32</td>
<td>8 week</td>
<td>Endoscopic</td>
<td>Golumimab</td>
<td>Discovery and Validation</td>
<td>13</td>
<td>GSE92415</td>
</tr>
<tr>
<td>GSE23597</td>
<td>Intestinal</td>
<td>UC</td>
<td>24</td>
<td>8 week</td>
<td>Endoscopic</td>
<td>Infliximab</td>
<td>Validation</td>
<td>12</td>
<td>GSE23597</td>
</tr>
<tr>
<td>GSE42296 (Blood)</td>
<td>PBMC</td>
<td>CD</td>
<td>14</td>
<td>6 week</td>
<td>CDAI</td>
<td>Infliximab</td>
<td>Validation</td>
<td>14</td>
<td>GSE42296</td>
</tr>
<tr>
<td>GSE107865 (Blood)</td>
<td>PBMC</td>
<td>CD</td>
<td>17</td>
<td>14 week</td>
<td>CRP, Calprotectin, Endoscopic</td>
<td>Infliximab</td>
<td>Validation</td>
<td>18</td>
<td>GSE107865</td>
</tr>
</tbody>
</table>

Notes: A total of six microarray datasets retrieved from Gene Expression Omnibus (GEO) were included in this study. Five cohorts from intestinal tissue (GSE12251, GSE16879CDc, GSE16879CDi, GSE16879UC, GSE92415) were selected as discovery cohorts. Eight cohorts (GSE12251, GSE16879CDc, GSE16879CDi, GSE16879UC, GSE92415, GSE23597, GSE107865, and GSE42296) were chosen as validation cohorts.

Abbreviations: UC, ulcerative colitis; CD, Crohn’s disease; CDi, Crohn’s ileitis; CDc, Crohn’s colon; PBMC, peripheral blood mononuclear cell; CRP, C reactive protein; Definition time, Time to definite response status; Definition method, Method to definite response status; CDAI, Crohn’s Disease Activity Index.
“signaling” and “Toll-like signing pathway” of on-response individuals in CD, UC, and IBD separately (p<0.01), which means the stable of DEGs enrichment analysis (Figure 3B–D). Also, whole down-regulated genes were applied to GSEA for search down-regulated function, oxidative phosphorylation was found.

Composition of Immune Cells Between Response and Non-Response Patients

Cases with a CIBERSORT output of p>0.05 means the estimated result is not accurate, which is not eligible for further analysis. Collectively, a total of 59 (100%), 24 (100%), 21 (87.5%) and 12 (63.2%) samples from GSE92415, GSE12251, GSE16869UC and GSE16869CDc were included, respectively. The dataset of GSE16869CDi was excluded, as more than 40% of the sample is not accurate (CIBERSORT output of p>0.05). The composition of immune cells in response and non-response from intestinal tissues was analyzed, and the fraction of 22 immune cells was shown in the violin plot (Figure 4A, also in Table S3). As shown in Figure 4B, the fractions M2 Macrophage is relatively low in the intestinal tissue from non-response patients than those from response patients. The proportions of M1 Macrophage in M1 and M2 Macrophage are also high in non-response individuals (FC 0.19–0.78).

Analyzing the PPI Network and Finding of Hub Genes

We constructed the PPI network involving 140 nodes (DEGs) and 975 edges through Cytoscape. The proteins expressed by DEGs are represented by the nodes and the physical interactions between two nodes means the edges,
which is obtained from string v11.0. The most highly connected cluster in the PPI network is extracted by MCODE plug-in in Cytoscape. The Maximal Clique Centrality is an important parameter, and the high score demonstrates the protein act as the center of the network and plays a critical role. The nodes with a higher score were shown as red or orange color, and the node size is related to the fold of change between non-response and response IBD patients (Figure 5A). The most prominent genes in this cluster were at the core of the whole PPI network, including TLR1, TLR2, TLR4, TLR8, FCGR2A, HCK, CD36, CCR1, CCL4, and they were all regulated as inflammation and chemotaxis (Table S4). Therefore, the night genes were considered as the hub gene. Subsequently, we analyzed Hub genes and their neighboring nodes for better visualization (Figure 5B).

Exploring Candidate Biomarkers by ROC Curves and Lasso Regression

Firstly, the Lasso regression model for GSE92415 (having the largest number of samples) was conducted to find an optimum linear combination of hub genes in predicting responsiveness (Figure 6A). Then, the ROC curve analysis of each hub gene and LASSO regression model (0.4795 and 0.2864 for CCL4 and FCGR2A) predicting the responsiveness of anti-TNFα agents in IBD patients from six cohorts (GSE12251, GSE16879, GSE23597, GSE92415, all form intestinal tissue).
Table 2 GO Terms Enrichment Results of DEGs

<table>
<thead>
<tr>
<th>GO Terms</th>
<th>Count</th>
<th>P value</th>
<th>Enriched Genes</th>
<th>Fold Enrichment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enriched from Up-Regulated Genes in BP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:0002760-positive regulation of tumor necrosis factor production</td>
<td>6</td>
<td>3.79E-06</td>
<td>TLR2, FCER1G, RIPK2, TLR4, CLEC7A, CD14</td>
<td>23.12478632</td>
</tr>
<tr>
<td>GO:0042116-macrophage activation</td>
<td>5</td>
<td>1.01E-04</td>
<td>SLC11A1, CD93, TLR1, TLR2, TLR4</td>
<td>19.27065527</td>
</tr>
<tr>
<td>GO:0002221-pattern recognition receptor signaling pathway</td>
<td>5</td>
<td>1.71E-04</td>
<td>IRAK3, TLR2, RIPK2, TLR4, CLEC7A</td>
<td>17.00351936</td>
</tr>
<tr>
<td>GO:0002680-regulation of tumor necrosis factor production</td>
<td>9</td>
<td>3.85E-08</td>
<td>IRAK3, TLR1, TLR2, FCER1G, RIPK2, TLR4, CLEC7A, IL10, CD14</td>
<td>16.78411911</td>
</tr>
<tr>
<td>GO:0009595-detection of biotic stimulus</td>
<td>5</td>
<td>2.17E-04</td>
<td>LY96, TLR4, CLEC7A, NLRP3, TLR8</td>
<td>16.05887939</td>
</tr>
<tr>
<td>GO:0002218-activation of innate immune response</td>
<td>5</td>
<td>2.71E-04</td>
<td>IRAK3, TLR2, RIPK2, TLR4, CLEC7A</td>
<td>15.21367521</td>
</tr>
<tr>
<td>GO:0032755-positive regulation of interleukin-6 production</td>
<td>5</td>
<td>2.71E-04</td>
<td>IL6, TLR2, IL1B, RIPK2, TLR4</td>
<td>15.21367521</td>
</tr>
<tr>
<td>GO:0002758-innate immune response-activating signal transduction</td>
<td>5</td>
<td>2.71E-04</td>
<td>IRAK3, TLR2, RIPK2, TLR4, CLEC7A</td>
<td>15.21367521</td>
</tr>
<tr>
<td>GO:0032675-regulation of interleukin-6 production</td>
<td>9</td>
<td>1.37E-07</td>
<td>IRAK3, IL6, CEPPB, TLR1, TLR2, IL1B, RIPK2, TLR4, IL10</td>
<td>14.45299145</td>
</tr>
<tr>
<td>GO:0007159-leukocyte adhesion</td>
<td>7</td>
<td>6.71E-06</td>
<td>ICAM1, SELP, OLR1, ITGAV, MSN, SELE, ITGAM</td>
<td>14.45299145</td>
</tr>
<tr>
<td>GO:0009620-response to fungus</td>
<td>5</td>
<td>6.95E-04</td>
<td>TLR2, TLR4, CLEC7A, PTX3, SI00A12</td>
<td>12.04415954</td>
</tr>
<tr>
<td>GO:0009237-response to molecule of bacterial origin</td>
<td>17</td>
<td>1.19E-12</td>
<td>SELP, IL6, PTGS2, SOCS3, LY96, TLR2, TLR4, IL10, IRAK3, SLC11A1, THBD, RIPK2, IL1B, IRG1, SELE, CD14, PAFR</td>
<td>11.26759206</td>
</tr>
<tr>
<td>GO:0050818-regulation of coagulation</td>
<td>8</td>
<td>5.61E-06</td>
<td>SELP, CAV1, PLEK, PDPN, SERPINE1, TLR4, ANX5A, PLAU</td>
<td>11.20838357</td>
</tr>
<tr>
<td>GO:0032496-response to lipopolysaccharide</td>
<td>15</td>
<td>4.39E-11</td>
<td>SELP, PTGS2, SOCS3, LY96, TLR4, IL10, IRAK3, SLC11A1, THBD, RIPK2, IL1B, IRG1, SELE, PAFR, CD14</td>
<td>11.26207126</td>
</tr>
<tr>
<td>Enriched from up-regulated genes in MF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:0019864-IgG binding</td>
<td>5</td>
<td>4.86E-06</td>
<td>FCGR2B, FCGR2C, FCGR1B, FCER1G, FCGR2A, FCGR3B</td>
<td>37.56655093</td>
</tr>
<tr>
<td>GO:0001530-lipopolysaccharide binding</td>
<td>6</td>
<td>8.57E-07</td>
<td>SELP, LY96, TLR2, TLR4, PAFR, CD14</td>
<td>30.03524074</td>
</tr>
<tr>
<td>GO:0003829-pattern recognition receptor activity</td>
<td>6</td>
<td>2.11E-06</td>
<td>LY96, TLR2, TLR4, CLEC7A, PAFR, CD14</td>
<td>25.75992063</td>
</tr>
<tr>
<td>GO:0019864-immunoglobulin binding</td>
<td>5</td>
<td>8.65E-05</td>
<td>FCGR2B, FCGR2C, FCGR1B, FCER1G, FCGR2A, FCGR3B</td>
<td>20.03549383</td>
</tr>
<tr>
<td>GO:004896-cytokine receptor activity</td>
<td>9</td>
<td>3.10E-06</td>
<td>IL1R1, IL1R1, OSMR, IL1RAP, CSF2RB, CSF3R, CXCR2, IL7R, IL13RA2</td>
<td>9.83560606</td>
</tr>
</tbody>
</table>

Notes: Significantly enriched GO terms with P-value <0.05 count ≥14 and were screened out. Fourteen BP terms were associated with an abundance of inflammatory cytokines and pathogen recognition.

Abbreviations: GO, Gene Ontology; DEGs, differentially expressed genes; BP, biological process; MF, molecular function.

is conducted successfully by PRISM 8.0 (Figure 6B); none of them were less than 0.500. The AUC of LASSO model is 0.7–1.0. Among these hub genes, FCGR2, TLR1, TLR2, and HCK was the distinguish one in intestine with the single gene AUC beyond 0.9 in at least one of study, which suggest it might have outstanding value for predicting the responsiveness in IBD patients. Besides, ROC curve analysis also applied in two cohorts (GSE107865 and GSE42296), FCGR2A (AUC: 0.669–0.706), HCK (AUC: 0.738–0.765), and TLR8 (AUC: 0.643–0.694) might be blood biomarkers (Figure 6C).

**Discussion**

Predicting the responsiveness of anti-TNFα treatment and understanding its molecular mechanism is becoming more and more apparent and urgent. Our finding through integrated analysis of several independently public available datasets of IBD patients before the treatment of anti-TNFα agents suggests that excessive interaction of FcγR signaling and Toll-Like Receptor (TLR) pathway may account for the different response to anti-TNFα agents in IBD patients. Beyond this, the lasso regression model containing FCGR2A and CCL4 can be used as predictive biomarkers.

Several publications have reported gene expression of IBD patients’ intestine biopsies with responsiveness to anti-TNFα agents.11–13,15 Ingrid Timmer et al found that unlike non-responders, the dysregulation of anti-microbial peptides (AMPs) in intestine mucosa almost normalized after Infliximab compared to their baseline samples in responders.11 Others have shown that Infliximab affects TH1, TH2, and TH17-cell differentiation pathways in
responders and that non-responders fail to regulate these three pathways. Nevertheless, no baseline gene panel to predict responsiveness was determined in the previous two studies. On the other hand, Ingrid Arijs et al also found that 5 top signature genes from a predictive panel of 109 probe sets in UC patients, which contains TNFRSF11B, STC1, PTGS2, IL11 and IL13α2. Then, the panel of 109 probe sets was refined in further study to a smaller set of 13 genes signature (CMTM2, C5AR1, FGF2, GK, HGF, L1RN, LILRA2, NAMPT, PAPPA, SNCA, SOD2, STEAP4, ZBED3). However, its accuracy of the 13 genes signature in an external validation was no better than chance, which limits its clinic use. The previous signature genes are not robust enough, which may arise from the relatively small sample size. We think it appropriate to reveal more robust and reproducible biological processes and hub genes through merging homogenizes studies. Therefore, 142 moderate-to-severity IBD patients refractory to conventional treatment were involved in our study for in-depth data re-mining.

The Gene Ontology (GO) terms of Biological Process (BP) revealed changes in the pathogenesis of non-response patients. IL-6 and TNF are the two most prominent BP terms, indicating that the abundance of inflammatory cytokines in non-response profile. The GSEA result of inflammation response confirms the involvement of IL-6 and TNF in UC, CD, and whole IBD patients. Noteworthy, because of the lack of random allocation, it has been challenging to determine if the excessive activation of inflammation response in non-response patients is independent of the severity of the disease.

However, the underlying biological process driving positive regulation of pro-inflammatory cytokines is more

<table>
<thead>
<tr>
<th>Table 3 KEGG Pathway Enrichment Results of DEGs</th>
</tr>
</thead>
<tbody>
<tr>
<td>KEGG Pathway</td>
</tr>
<tr>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>hsa04640: Hematopoietic cell lineage</td>
</tr>
<tr>
<td>hsa04620: Toll-like receptor signaling</td>
</tr>
<tr>
<td>hsa04610: Complement and coagulation cascades</td>
</tr>
<tr>
<td>hsa04060: Cytokine-cytokine receptor interaction</td>
</tr>
<tr>
<td>hsa04512: ECM-receptor interaction</td>
</tr>
<tr>
<td>hsa04630: Jak-STAT signaling pathway</td>
</tr>
<tr>
<td>hsa04210: Apoptosis</td>
</tr>
<tr>
<td>hsa04514: Cell adhesion molecules (CAMs)</td>
</tr>
<tr>
<td>hsa04510: Focal adhesion</td>
</tr>
<tr>
<td>hsa04062: Chemokine signaling pathway</td>
</tr>
<tr>
<td>hsa04010: MAPK signaling pathway</td>
</tr>
</tbody>
</table>

Notes: KEGG biological pathway enrichment analysis found that Toll-like receptor signaling pathway (P-value < 0.0001, count = 12, and fold enrichment = 4.61) was the most important one among the eleven significantly enriched pathways.

Abbreviation: KEGG, Kyoto Encyclopedia of Genes and Genomes.
critical. Four terms were directly relevant to pathogen recognition in the top six terms of BP, which is consistent with the previous finding that several genetic variations in pattern recognition receptors (PRRs) loci were associated with anti-TNFα response in autoimmune disease. Hence, we reasoned that the changes might contribute to a link between the innate immune cell and intestinal microbial. Host PRRs modulate microbial recognition and could thereby alert gut microbe community function and structure broadly. Modulated microbial, in turn, induces abnormal signal, cytokine secretion, and bacterial clearance in innate immune cells, thus highlighting the critical role of balance in the regulation of PRR-initiated outcomes in intestinal tissues.

Figure 3 Baseline biopsy gene expression and GSEA result between response and non-response in Discovery Cohorts. (A) Box plot indicates intestine tissue IL6, NOD2 and TNF expression. (B) GSEA result of Inflammation response, Fc-gamma receptor signaling and Toll-like signing pathway in CD (GSE16879CDi, GSE16879CDC) patients. (C) GSEA result of Inflammation response, Fc-gamma receptor signaling and Toll-like signing pathway in UC (GSE12251, GSE16879UC, GSE92415) patients. (D) GSEA result of Inflammation response, Fc-gamma receptor signaling and Toll-like signing pathway in IBD (GSE12251, GSE16879UC, GSE16879CDi, GSE16879CDC, GSE92415, GSE21597) patients.

Abbreviation: GSEA, Gene Set Enrichment Analysis.
interactions in non-response patients. On the one hand, the “Lipopolysaccharide receptor” acting as one of the significant PRR in detecting gram-negative microbe is the most significant GO terms of MF for up-regulated genes with the highest fold enrichment rate. Similarly, “Peptidoglycan” binding both gram-negative and gram-positive microbe is also enrichment. NOD2, the receptor of muramyl dipeptide (MDP), a fragment of bacterial peptidoglycan, is the first studied susceptibility factor to IBD. Although NOD2 mutations show the value of determining the risk of disease development, it is controversial whether NOD2 is related to the efficacy of anti-TNFα therapy. Supported by “IgG-mediated activation signaling” is the most significant GO term of MF. Within the colonic mucosa, monocytes, macrophages (MMP) as

Figure 4 Violin and Box plot of immune cell composition between response and non-response individual. (A) The violin plot indicates the composition of 22 immune cells between response and non-response individuals in 4 cohorts with CIBERSORT p < 0.05 for all eligible samples. The blue violin plot indicates non-response individual, and the red violin plot represents response patients. (B) The box plot indicates the composition of M1 macrophage, M2 macrophage, ration of M1/(M1+M2) macrophage, and its fold-change between response and non-response individuals.
the major FcγR-expressing cell types mediate effector functions of immunoglobulin G (IgG) antibodies.\textsuperscript{46, 47} Considering the polarization of M1 macrophages are relatively high in non-response individuals, theverse cycle is likely to occur between polarization and activation of MMP between gut microbe. Overproduction of anti-flagellin IgG, as well as voluminous IgG-opsonized microbes, are believed to lead to pathological IgG-FcγR-dependent immune response.\textsuperscript{46, 48} Also, within the persisting chronic pathological inflammatory states, the balance of MMP Activating and inhibitory FcγR (the A: I ratio) can be destroyed, which reduces the activation threshold of cells upon encounter of immune-complex or opsonized targets.\textsuperscript{44} As IgG widely emerged and the intrinsic A: I ratio skewed, over-Activated MMP, driven by FcγR signal, determines the magnitude of intestinal inflammation, which might be an essential cause of non-response to the anti-TNFα agents.\textsuperscript{49}

The KEGG analysis further shows changes in the signaling pathways in non-response patients. We prompt that “Toll-like receptor signaling pathway” as the hub KEGG pathway. Almost all DEGs involved in the Toll-like signaling pathway were up-regulated, which was consistent with the recent study found that polymorphisms of TLR pathway are valuable significant predictors for the response to anti-TNFα therapy among patients with autoimmune disease.\textsuperscript{50, 51}

Firstly, as motioned above, TLR pathway facility the production of IgG and thereby stimulate the FcγR signal. Secondly, FcγR stimulates phagocytosis.\textsuperscript{52} Consequently, more antigen is provided for the Toll-like pathway to recognize. Thirdly, the FcγR signaling could cross-link to the Toll-like pathway greatly up-regulates pro-inflammatory chemokines transcriptional, and Toll-like pathway agonist could partially rescue IgG/FcγR induction of inflammation in the setting of deficient chemokines expression.\textsuperscript{53, 54} Furthermore, several studies have highlighted a link between TLR and FcγR co-stimulation and the induction of a Th17 polarizing and pathogenic macrophage phenotype, which is constant to the characteristic of IBD.\textsuperscript{53–55} Taken together, FcγR stimulated by anti-commensal IgG signaling synergistic cooperating with toll-like pathway derives inappropriate inflammation in the gut, which might be crucial in the pathogenesis of non-response of anti-TNFα agents. These results suggest that the FcγR-TLR axis is responsible for the magnify reciprocal signaling, which ultimately impacts response status.

We constructed the PPI network with up-regulated DEGs lists the top-degree hub genes: TLR4, TLR1, TLR8, HCK, FCGR2A, CCL4, CCR1, and CD86. The KEGG analysis further shows changes in the signaling pathways in non-response patients. We prompt that “Toll-like receptor signaling pathway” as the hub KEGG pathway.
cross-validation, biological behavior has also been considered in the process of PPI construction.\textsuperscript{15} Thus, we detected nine diverse and non-overlapping hub genes. Importantly, all of the nine hub genes show distinct features that excessive interaction of the FcγR-TLR pathway axis in non-response individuals. FCGR2A, an active FcγR signal receptor, is a good example. One on hand, previous work has suggested that FCGR2A co-stimulate with Toll-Like pathway might massively induce pro-inflammatory cytokines and IgG.\textsuperscript{55} On the other hand, without the activation of the Toll-like pathway, FCGR2A also mediates beneficial functions via its ability to affinity to anti-TNFα monoclonal antibodies, which is one of the crucial ways for anti-TNFα agent taking effect.\textsuperscript{56,57} Therefore, given the versatility of FCGR2A and the critical role of the Fc region of anti-TNFα monoclonal antibodies, we speculate the dual role effect of FCGR2A in anti-TNFα treatment IBD patients. Additional studies are needed to confirm the dual role of FCGR2A. Besides, CCR1 is critical for the recruitment of effector immune cells to the site of inflammation. Meanwhile, it involved in IL-6 release of
CCL4, also known as macrophage inflammatory protein-1β (MIP-1β), plays a vital role in the migration of macrophage in the gut and can be active by TLR4. Our data differ from a previous study found that overexpression of CCR2-CCL7 is associated with non-response patients. In the future work, the chemokines and its receptor should be examined in terms of the specific type of cell recruited to the gut.

Interestingly, expression HCK, FCGR2A, and TLR8 also show predictive value in peripheral blood mononuclear cells (PBMCs). In particular, HCK might be an essential molecule that has not been thoroughly studied but is of great value to the prediction biomarker. Acting downstream of receptors that bind the Fc region of immunoglobulins, such as FCGR1A and FCGR2A, plays a vital role in the regulation of innate immune responses, phagocytosis, cell survival and proliferation, cell adhesion and migration. Numerous studies showed that active HCK exhibited increased susceptibility to TLR signaling inflammation and elevated innate immune response. Although the expression of HCK was proven to correlate with the development of inflammation positively, the recent investigation of HCK associated with autoimmune disease is that Behçet’s disease (BD) due to the broad regulation of LPS-Induced pathway, many results are contradictory.

From a clinical perspective, one implication of the current study finds biomarkers of anti-TNFα treatment. We have demonstrated a ROC analysis for predicting therapy response with a single hub gene or Lasso linear models. Thus, the model consists of FCGR2A and CCL4 shows great value of prediction in intestinal tissue.

Another clinical implication of the present study is that non-response to anti-TNFα treatment patients may elect an alternative therapy. Our findings suggest that no responders may represent a vicious circle of, including an up-regulated TLR pathway, skewed FcγR signaling, and distribute microbial, which hamper the resolution of inflammation via anti-TNFα agent treatments. Blocking the mechanism that restores the gut microbe community would be helpful, support by the finding that the functional links intestinal microbial ecosystem and therapeutic manipulation. Indeed, antibiotic showed that restoring the gut microbe prevent recurrence after surgery. Another way of blocking the pathogenic mechanism is to develop new agents with a higher affinity to FCGR2A.

In summary, our study has the merit of assessing transcriptional analysis from several independent cohorts though reanalyzed six publicly available microarray gene expression profiling (origin of samples varies from blood to intestinal) via different methods and gain different conclusions. Importantly, we found polarization from M2 to M1 macrophages in non-response individuals and nine hub genes. Thus, we propose our insight that the cross-talk between FcγR signaling and the TLR axis may be responsible for the different response to anti-TNFα agents in IBD patients. Besides, some hub genes have high predictive value.

**Conclusion**

The current study aims to explore the molecular mechanism and search for candidate biomarkers in the gene expression profile of IBD patients associated with the response to anti-TNFα agents. Totally 286 DEGs and nine hub genes were screened via bioinformatic approaches. Our prominent finding is that the interaction of the FcγR signaling and TLR axis may be responsible for the different responses to anti-TNFα agents in IBD patients, and the dual role effect of FCGR2A may be the underlying mechanism. However, there were some limitations. It should be noted that none of the cohorts involved in this study reported anti-TNFα agent trough levels of each patient, so it is difficult to figure out whether the non-response was caused by pharmacokinetic or pharmacodynamic. Further studies with therapeutic drug monitoring and dose optimization are needed. Besides, additional studies with a large sample size as well as direct evidence of microbes are useful to validate our results.

**Acknowledgments**

This project was supported from Natural Science Foundation of Shanghai (19ZR1429700); Research Project of Pudong New Area Health and Family Planning Commission (PW2018D-01); Clinical Research Project of Multi-Disciplinary Team, Shanghai Ninth People's Hospital, Shanghai JiaoTong University School of Medicine (201701008); Clinical Research Program of Shanghai Ninth People's Hospital, Shanghai JiaoTong University School of Medicine (JYLJ022). We also thank the encouragement and biostatistics support from Yeyang Zhu for this project.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**


21. Dovepress

Journal of Inflammation Research 2020:13

submit your manuscript | www.dovepress.com

DovePress

For personal use only.


