Clinical and Translational Research
Exploring the autophagy-related pathogenesis of active ulcerative colitis

Gong ZZ et al. Autophagy in active UC

Zhuo-Zhi Gong, Teng Li, He Yan, Min-Hao Xu, Yue Lian, Yi-Xuan Yang, Wei Wei, Tao Liu

Abstract
BACKGROUND
The pathogenesis of ulcerative colitis (UC) is complex, and recent therapeutic advances remain unable to fully alleviate the condition.

AIM
To inform the development of novel UC treatments, bioinformatics was used to explore the autophagy-related pathogenesis associated with the active phase of UC.

METHODS
The GEO database was searched for UC-related datasets that included healthy controls who met the screening criteria. Differential analysis was conducted to obtain differentially expressed genes (DEGs). Autophagy-related targets were collected and intersected with the DEGs to identify differentially expressed autophagy-related genes (DEARGs) associated with active UC. DEARGs were then subjected to KEGG, GO, and DisGeNET disease enrichment analyses using R software. Differential analysis of immune infiltrating cells was performed using the CiberSort algorithm. The least
absolute shrinkage and selection operator algorithm and protein-protein interaction network were used to narrow down the DEARGs, and the top five targets in the Dgree ranking were designated as core targets.

RESULTS
A total of 4822 DEGs were obtained, of which 58 were classified as DEARGs. SERPINA1, BAG3, HSPA5, CASP1, and CX3CL1 were identified as core targets. GO enrichment analysis revealed that DEARGs were primarily enriched in processes related to autophagy regulation and macroautophagy. KEGG enrichment analysis showed that DEARGs were predominantly associated with NOD-like receptor signaling and other signaling pathways. Disease enrichment analysis indicated that DEARGs were significantly linked to diseases such as malignant glioma and middle cerebral artery occlusion. Immune infiltration analysis demonstrated a higher presence of immune cells like activated memory CD4 T cells and follicular helper T cells in active UC patients than in healthy controls.

CONCLUSION
Autophagy is closely related to the active phase of UC and the potential targets obtained from the analysis in this study may provide new insight into the treatment of active UC patients.

**Key words:** Ulcerative colitis; Autophagy; Bioinformatic; Targets; Pathogenesis


**Core Tip:** This study used bioinformatics to explore the autophagy-related pathogenesis of ulcerative colitis (UC) during its active phase. A total of 58 differentially expressed autophagy-related genes (DEARGs) were found in gene expression datasets from UC
patients and healthy controls. Of these, SERPINA1, BAG3, HSPA5, CASPI, and CX3CL1 were identified as core targets. Enrichment analysis highlighted the involvement of DEARGs in autophagy regulation, and macroautophagy, in addition to NOD-like receptor signaling and other pathways. These DEARGs were also shown to be associated with diseases like malignant glioma and middle cerebral artery occlusion. Immune infiltration analysis revealed an increased presence of immune cells, including activated memory CD4 T cells and follicular helper T cells in active UC patients than in healthy controls. This study suggests that autophagy plays a significant role in the active phase of UC and identifies potential targets for novel UC treatments.

INTRODUCTION

Ulcerative colitis (UC) is a chronic, recurrent inflammatory disease in humans that profoundly affects normal functioning[1]. It has both active and remission phases that are classified according to disease severity. UC is characterized by symptoms such as weight loss, diarrhea, rectal bleeding, abdominal pain, and inflammation of the mucous membranes that extend from the rectum to the distal part of the colon[2]. Approximately 5 million people are affected by UC worldwide and recent studies indicate that the incidence is increasing[3]. UC is thought to result from a combination of genetic and environmental factors and is closely linked to compromised intestinal epithelial barriers, a dysregulated microbiome, and impaired immune responses[3].

Autophagy is a finely coordinated process that segregates misfolded proteins, damaged or aged organelles, and mutated proteins into double-membrane vesicles called autophagosomes. The autophagosomes later merge with lysosomes and degrade these components[4]. Three main forms of autophagy have been described to date: microautophagy, chaperone-mediated autophagy, and macroautophagy[5]. Autophagy is shown to be a key mediator in the pathophysiological processes of UC. From a physiological perspective, autophagy plays a critical role in maintaining intestinal balance, regulating interactions between gut microbiota and both the innate and adaptive immune systems, and protecting the host against intestinal pathogens[5]. Autophagy also
reduces endoplasmic reticulum stress associated with diverse inflammatory and immune disorders\cite{6}, helping to restore gut homeostasis. For example, estrogen-related receptor alpha (ESRRAa) contributes to maintaining intestinal balance by activating autophagy and regulating the gut microbiome, thereby protecting the host from inflammation and mitochondrial dysfunction\cite{7}. Meanwhile, autotaxin (ATX) inhibits autophagy through the mTOR pathway, causing significant damage to the intestinal epithelial barrier of colitis patients\cite{8}. During intestinal inflammation, specific bacterial species in the microbiome, including adherent invasive \textit{Escherichia coli}, can adhere to intestinal epithelial cells and evade autophagic elimination by phagocytic macrophages\cite{9}. The situational or excessive induction of autophagy can adversely impact cells by initiating autophagic cell death. The lack of erbin, a protein essential for epithelial cell polarity, markedly worsens the initiation of autophagic processes and autophagic cell death in mice with DSS-induced colitis\cite{10}. Damage to intestinal epithelial cells initiates inflammation and intensifies the severity of UC symptoms\cite{11}.

Given the importance of autophagy in preserving intestinal balance and the role of autophagy dysfunction in UC development, identifying autophagy-related disease predictors is essential for the design of new UC treatments. The current study uses bioinformatics to define gene expression patterns associated with the autophagy-related pathogenesis of active UC (Figure 1).

**MATERIALS AND METHODS**

\textit{Identification of active UC targets and difference analysis}

Datasets related to active UC which included normal control and active UC samples and had a sample size > 30 were obtained from the GEO database (https://www.ncbi.nlm.nih.gov/gds/). The selected data set was normalized and the “limma” package was downloaded using “Bioconductor.” R 4.3.1 software was then used to perform differential gene analysis on targets identified in the dataset. A $|\log FC| \geq 0.585$ and an adj. $P < 0.05$ were used to obtain differentially expressed genes (DEGs).
Acquisition of differentially expressed autophagy-related genes in patients with active UC

To obtain differentially expressed autophagy-related genes (DEARGs), autophagy-related genes were downloaded from the Human Autophagy Database (http://www.autophagy.lu/). Using the “Venn Diagram” package in R, autophagy-related targets were intersected with the DEGs, identifying DEARGs as the central targets for further analysis. DEARG heat maps were generated using the “limma” and “pheatmap” packages.

Analysis of immune cell infiltration

The immune microenvironment is typically composed of immune cells, inflammatory cells, fibroblasts, and mesenchymal stem cells, along with various cytokines and chemokines. Assessing immune cell infiltration is vital for predicting disease progression and treatment response. Several methods exist to analyze immune cell infiltration, including CiberSort, an inverse convolution algorithm developed by BinderG. This method calculates the cellular composition of complex tissues based on normalized gene expression data, allowing specific cell types to be quantified. The CiberSort deconvolution algorithm was used with 100 simulations and subsequent analyses were conducted with a significance threshold of \( P < 0.05 \) to determine the proportion of immune cells in different samples. The results were visualized using the “ggpubr” package in R.

Assessment of biological variables associated with the DEARGs

Gene ontology (GO) analysis categorizes genes into biological processes (BP), molecular functions (MF), and cellular components (CC), which help to inform their biological functions. The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a database that integrates genomic, chemical, and systemic information. It is often used for the functional annotation of genes to understand their associated activities and pathways of action. To further understand the target functions of autophagy in patients with active UC and the
associated signaling pathways, the “clusterProfiler” package was downloaded from Bioconductor, and GO and KEGG enrichment analysis of the DEARGs was conducted using R. The “clusterProfiler” package was downloaded from “Bioconductor” and the DEARGs were analyzed by GO and KEGG enrichment analysis using R with a threshold value of $P < 0.05$.

**Analysis of disease enrichment in DisGeNET**

To explore the role of autophagy in UC-related diseases, DEARGs were input into the Metascape platform (https://metascape.org/) using “*H. sapiens*” as the species setting for both “Input” and “Analysis.” The “Summary of enrichment analysis in DisGeNET” was then exported.

**Construction of the least absolute shrinkage and selection operator algorithm and protein-protein interaction network**

For more precise identification of the core targets, the least absolute shrinkage and selection operator (LASSO) algorithm was used along with the construction of a protein-protein interaction (PPI) network to refine DEARG selection and predict key biomarkers. The LASSO algorithm is more effective than ordinary least squares estimation at extracting essential variables and simplifying the model, particularly when using multiple variables. PPI analysis helps identify interactions among DEARGs and refine the selection.

The LASSO algorithm was used for DEARG validation and feature gene selection using the “glmnet” package in R. The identified genes were then uploaded to the String database (https://cn.string-db.org/) with a “minimum required interaction score” of 0.15 and the results were imported into Cytoscape 3.9.1. To further refine the selection, the “CytoNCA” plugin was used to rank the targets based on degree values, selecting the top five as core targets.

**RESULTS**
**Acquisition of DEGs**

GEO database screening identified two UC-related datasets: GSE87466 and GSE53306. GSE53306 includes data on differential gene expression between the active and quiescent stages of UC, providing insight into the disease characteristics. The dataset, which has information on 40 individuals, including 16 active UC cases and 12 normal controls, was published on December 13, 2014, and last updated on December 22, 2017. GSE87466 includes data on gene expression in mucosal biopsies from adult patients with moderate to severe active UC. The dataset has information on 87 UC active samples and 21 normal control samples and was first published on September 29, 2016, and last updated on March 2, 2019. The datasets were downloaded and exported, the data were de-duplicated and normalized, and differential analysis was conducted with R software using GSE87466 and GSE53306 as the base and supplemental datasets, respectively. GPL13158 and GPL14951 were used as the platform files. This analysis yielded 4,822 DEGs from the GSE87466 and GSE53306 datasets.

**DEARG acquisition**

A total of 232 autophagy-related genes were obtained from the Human Autophagy Database (http://www.autophagy.lu/). These autophagy-related genes were intersected with the DEGs, resulting in 58 DEARGs (Figure 2). R was then used to analyze the DEARGs and generate heat maps (Figure 3).

**Analysis of immune cell infiltration**

The Cibersort algorithm was used to evaluate immune cell infiltration in two distinct immune states. The following immune cell types were more abundant in active UC cases than in healthy controls: activated memory CD4 T cells, follicular helper T cells, γδ T cells, M0 macrophages, M1 macrophages, activated dendritic cells, activated mast cells, and neutrophils. The “ggpubr” package in R was used to visualize the differential analysis results of immune cell infiltration in each sample. A P < 0.05 was considered statistically significant.
The following immune cell types were significantly higher in the UC group than in the healthy control group: activated memory CD4 T cells ($P < 0.001$), follicular helper T cells ($P < 0.05$), gamma delta T cells ($P < 0.05$), M0 macrophages ($P < 0.001$), M1 macrophages ($P < 0.001$), activated dendritic cells ($P < 0.001$), activated mast cells ($P < 0.001$), and neutrophils ($P < 0.001$). Meanwhile, CD8 T cells ($P < 0.05$), resting memory CD4 T cells ($P < 0.05$), regulatory T cells (Tregs) ($P < 0.001$), activated NK cells ($P < 0.01$), monocytes ($P < 0.01$), M2 macrophages ($P < 0.001$), resting dendritic cells ($P < 0.05$), and resting mast cells ($P < 0.001$) were significantly higher in the healthy control group than in the active UC group. No significant differences were observed in naive B cells, memory B cells, plasma cells, naive CD4 T cells, resting NK cells, and eosinophils between active UC cases and healthy controls (Figure 4).

**Biological variables related to the DEARGs**

BP analysis revealed that the DEARGs were primarily associated with the regulation of autophagy, macroautophagy, autophagosome assembly, autophagosome organization, and vacuole organization. CC analysis showed that the DEARGs were primarily enriched in autophagosomes, phagophore assembly sites, and phagophore assembly site membranes. MF analysis found that DEARGs were mainly involved in chaperone binding, ubiquitin protein ligase binding, and heat shock protein binding (Figure 5A). KEGG enrichment analysis indicated that the DEARGs were predominantly enriched in autophagy-animal, autophagy-other, lipid and atherosclerosis, protein processing in the endoplasmic reticulum, and influenza A-related pathways (Figure 5B).

**DisGeNET disease enrichment analysis**

The Metascape “Summary of enrichment analysis in DisGeNET” revealed that the DEARGs were mainly enriched in malignant glioma, middle cerebral artery occlusion, infection, glomerulonephritis, and other diseases (Figure 6).

**Construction of the LASSO algorithm and PPI**
The LASSO algorithm narrowed the range of DEARGs and identified 13 targets: proliferation and apoptosis adaptor protein 15 (PEA15), heat shock 70-kDa protein 5 (HSPA5), caspase 1 (CASP1), serine protease inhibitor A1 (SERPINA1), C-X3-C chemokine ligand 1 (CX3CL1), Bcl2-associated athanogene 3 (BAG3), tumor protein p53 inducible nuclear protein 2 (TP53INP2), and peroxisomal biogenesis factor 14 (PEX14) (Figure 7). Their relationships were further established using the String database. The Cytoscape 3.9.1 software “CytoNCA” plug-in was used to sort the 13 targets according to their degree values, and the top five were selected as the core targets. The Fold Change (logFC) of these targets was obtained from the difference analysis results. All five were up-regulated and had the following parameters: SERPINA1 (logFC = 1.051), BAG3 (logFC = 0.661), HSPA5 (logFC = 0.790), CASP1 (logFC = 1.231), and CX3CL1 (logFC = 0.837) (Figure 8).

**DISCUSSION**

The current study identified HSPA5, CASP1, SERPINA1, CX3CL1, and BAG3 as core autophagy-related targets in active UC, all of which were upregulated during the disease. Key signaling pathways linked to these targets included autophagy in animals, other autophagy pathways, and lipid and atherosclerosis pathways. DisGeNET enrichment analysis found that middle cerebral artery occlusion, glomerulonephritis, and active UC were interrelated risk factors associated with autophagy. Active UC patients were found to have significantly higher counts of activated memory CD4 T cells, follicular helper T cells, gamma delta T cells, M0 macrophages, M1 macrophages, activated dendritic cells, mast cells, and neutrophils than healthy controls.

The results, including those predicted using core targeting and immune infiltration analysis, are supported by existing literature. R-HSPA5 is a specific form of HSPA5 that is localized in the endoplasmic reticulum (ER) and shown to play a critical role in autophagy-mediated lysosomal protein hydrolysis. Significant overexpression of HSPA5 mRNA and protein is found in UC patient tissues. CASP1 expression in macrophages impairs autophagy, triggering inflammatory vesicle activation, a key factor
associated with diseases such as active UC\textsuperscript{[15]}. Activated CASP1 is critical for DSS-induced colitis\textsuperscript{[16-18]}. Soendergaard \textit{et al}\textsuperscript{[19]} identified SERPINA1 as a potential biomarker of mild to moderate UC activity. Elevated CX3CL1 levels interact with CX3CR1, inhibiting autophagy in Kupffer cells\textsuperscript{[20]}. CX3CL1 induces the infiltration and activation of CX3CR1-expressing cells, stimulating iNOS expression, a key mediator in DSS-induced colitis. These findings suggest that CX3CL1 may have potential for use in UC treatments\textsuperscript{[21,22]}. Effector memory CD4 T cells induce IL-7 expression, leading to inflammation in the lamina propria of the intestinal mucosa\textsuperscript{[23]}. During UC pathogenesis, Th1 cell functional abnormalities and imbalances disrupt the immune barrier of the intestinal mucosa, triggering immune disorders and the development of UC\textsuperscript{[24]}. Inagaki-Ohara \textit{et al}\textsuperscript{[25]} found that UC naturally occurs in γδ T-cell-deficient mice. Mast cells (MC), commonly found in capillaries of the intestinal mucosa, are increased in inflammatory bowel disease (IBD) patients. They contain basophilic granules that release inflammatory factors during stress\textsuperscript{[26,27]}. The IL-33/ST2 pathway, together with IgE signaling, mediates MC degranulation, inducing the release of inflammatory factors and initiating a cascade response\textsuperscript{[28]}. Reduced NK cell levels affect mucosal flora responses, resulting in immune abnormalities and inflammatory changes in the colon\textsuperscript{[29]}. Cherfane \textit{et al}\textsuperscript{[30]} identified an association between the number of peripheral blood mononuclear cells in UC patients and disease activity, suggesting that monocyte count could serve as a potential UC biomarker. The Th1/Th2 cell imbalance, along with the overexpression and activation of co-stimulatory molecules on dendritic cells, can trigger monocyte migration to the intestine, causing inflammation and potential damage to the intestinal mucosa\textsuperscript{[31,32]}. The disease enrichment analysis results discussed here are confirmed by prior studies. Different stages of IBD are linked to the development of thrombosis, with IBD episodes or activity serving as a primary risk factor\textsuperscript{[33]}. The predicted core genes are critical to the pathogenesis+ADs- hSPA5, for example, offering neuroprotection in ischemic strokes\textsuperscript{[34]}. BAG3 overexpression is shown to improve neurological outcomes associated with middle cerebral artery embolism in mice, reducing infarct volume and enhancing cell survival by activating autophagy and inhibiting apoptosis\textsuperscript{[35]}. Ischemia-induced neuronal
autophagy is shown to exacerbate microglial inflammation post-stroke, possibly due to the reduced CX3CL1 expression in autophagic neurons\textsuperscript{36}. Growing evidence suggests that autophagy plays a role in renal disease pathogenesis\textsuperscript{37,38}. Glomerulonephritis emergence or exacerbation often coincides with IBD and subsides following effective IBD treatment\textsuperscript{30}. CX3CL1 and CXCL10, induced by the core targets in this study, initiate activated leukocyte infiltration into glomerular cells\textsuperscript{40-42}. The NLRP3-ASC-caspase-1 inflammasome mitigates glomerular dysfunction by producing IL-1\textsuperscript{43,44}.

**CONCLUSION**

In summary, HSPA5, CASP1, SERPINA1, CX3CL1, and BAG3 were identified as core autophagy-related targets that are upregulated in active UC patients. These targets are associated with key signaling pathways, including autophagy in animals, other autophagy pathways, and lipid and atherosclerosis pathways. DisGeNET enrichment analysis also revealed a significant connection between middle cerebral artery occlusion, glomerulonephritis, and the autophagy-related pathogenesis of active UC. In addition, active UC patients exhibited significantly higher counts of various immune cells than healthy controls, indicating the occurrence of immune dysregulation. These findings provide valuable insight into the role of autophagy in UC pathogenesis and have potential implications for the development of novel targeted therapies.

**ARTICLE HIGHLIGHTS**

*Research background*

The etiology of ulcerative colitis (UC), a chronic inflammatory bowel disease (IBD), remains poorly understood. The pathogenesis of UC is complex and is influenced by genetic, environmental, and immune-related factors. While some recent progress has been made in the development of effective UC treatments, few patients experience complete relief of their symptoms. Thus, finding new therapeutic avenues to improve UC patient quality of life remains an urgent need. Autophagy is a cellular self-degradation and repair process that can help remove harmful proteins and organelles from cells and
maintain intracellular homeostasis. Recent studies suggest that autophagy may play a key role in the pathogenesis and progression of IBD.

*Research motivation*

The motivation of this study was to provide an in-depth investigation of the autophagy-related pathogenesis of active phase UC. Bioinformatics analysis was used to better understand whether autophagy plays a key role in active UC and which autophagy-related genes may contribute to the disease process.

*Research objectives*

This study sought to provide new ideas and potential therapeutic targets for the treatment of active UC to better understand the pathogenesis of the disease and improve clinical symptoms.

*Research methods*

A bioinformatics approach was used to compare gene expression data between patients with active UC and healthy controls to identify core genes associated with autophagy and to obtain more information about the role of autophagy in this disease.

*Research results*

HSPA5, CASP1, SERPINA1, CX3CL1, and BAG3 were identified as core targets associated with autophagy-related pathogenesis in active UC, all of which were upregulated. Key signaling pathways linked to these targets include autophagy in animals, other autophagy pathways, and lipids and atherosclerosis pathways. DisGeNET enrichment analysis showed that middle cerebral artery occlusion, glomerulonephritis, and active UC were interrelated risk factors associated with autophagy. Active UC patients had significantly higher counts of activated memory CD4 T cells, follicular helper T cells, gamma delta T cells, M0 macrophages, M1 macrophages, activated dendritic cells, mast cells, and neutrophils than healthy controls.
Research conclusions

HSPA5, CASP1, SERPINA1, CX3CL1, and BAG3 were identified as core autophagy-related targets in active UC patients, all of which were upregulated. These targets are associated with key signaling pathways, including autophagy in animals, other autophagy pathways, and lipid and atherosclerosis pathways. DisGeNET enrichment analysis revealed a significant connection between middle cerebral artery occlusion, glomerulonephritis, and the autophagy-related pathogenesis of active UC. In addition, active UC patients had significantly elevated counts of various immune cells, indicating that immune function is dysregulated. These findings provide valuable insight into the role of autophagy in UC pathogenesis and could be used to inform the development of targeted therapeutic interventions.

Research perspectives

Future research in this field should focus on better understanding the molecular mechanisms by which HSPA5, CASP1, SERPINA1, CX3CL1, and BAG3 contribute to autophagy in patients with active UC. Investigating the specific roles of these core targets in UC pathogenesis and their interactions with the identified key signaling molecules should be a priority. Interventions that target the core autophagy-related genes and pathways could offer promising treatment options for active UC patients. It is also important to further explore the immune dysregulation observed in UC patients, particularly the elevated immune cell counts, to understand better the inflammatory processes involved and inform the development of immunomodulatory strategies to manage UC.

10
ACKNOWLEDGEMENTS

We thank the GEO database for providing their platforms and contributors for their valuable datasets.
REFERENCES


<table>
<thead>
<tr>
<th>1</th>
<th><a href="http://www.frontiersin.org">www.frontiersin.org</a></th>
<th>49 words — 1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>aging-us.com</td>
<td>34 words — 1%</td>
</tr>
<tr>
<td>3</td>
<td><a href="http://www.dovepress.com">www.dovepress.com</a></td>
<td>34 words — 1%</td>
</tr>
<tr>
<td>4</td>
<td><a href="http://www.tandfonline.com">www.tandfonline.com</a></td>
<td>33 words — 1%</td>
</tr>
<tr>
<td>5</td>
<td><a href="http://www.mdpi.com">www.mdpi.com</a></td>
<td>30 words — 1%</td>
</tr>
<tr>
<td>6</td>
<td><a href="http://www.researchsquare.com">www.researchsquare.com</a></td>
<td>28 words — 1%</td>
</tr>
<tr>
<td>7</td>
<td>f6publishing.blob.core.windows.net</td>
<td>21 words — 1%</td>
</tr>
<tr>
<td>8</td>
<td>assets.researchsquare.com</td>
<td>20 words — 1%</td>
</tr>
<tr>
<td>9</td>
<td>Xuefeng Zhang, Wanchun Yang. &quot;Identification of a Novel Prognostic Model for Patients with Head and Neck Squamous Cell Carcinoma Based on Autophagy-related Genes&quot;, Research Square Platform LLC, 2022</td>
<td>19 words — 1%</td>
</tr>
<tr>
<td>#</td>
<td>Title</td>
<td>Source</td>
</tr>
<tr>
<td>----</td>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>pubmed.ncbi.nlm.nih.gov</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td><a href="http://www.ncbi.nlm.nih.gov">www.ncbi.nlm.nih.gov</a></td>
</tr>
</tbody>
</table>