91644_Auto_Edited.docx

WORD COUNT 3362
TIME SUBMITTED 06-FEB-2024 04:35PM
PAPER ID 106710387

6
Name of Journal: *World Journal of Clinical Cases*
Manuscript NO: 91644
Manuscript Type: ORIGINAL ARTICLE

Clinical and Translational Research

Identification of marker genes associated with m6A and autophagy in ulcerative colitis

Xiao-Yan Liu, Dan Qiao, Ya-Li Zhang, Zi-xuan Liu, You-Lan Chen, Ren-Ye Que, Hong-Yan Cao, Yan-Cheng Dai

Abstract

BACKGROUND

Both N6-methyladenosine (m6A) methylation and autophagy are considered relevant to the pathogenesis of ulcerative colitis (UC). However, a systematic exploration of the role of the com-bination of m6A methylation and autophagy in UC remains to be performed.

AIM

To elucidate the autophagy-related genes of m6A with a diagnostic value for UC.

METHODS

The correlation between m6A-related genes and autophagy-related genes (ARGs) was analyzed. Finally, gene set enrichment analysis (GSEA) was performed on the characteristic genes. Additionally, the expression levels of four characteristic genes were verified in DSS-induced colitis in mice.

RESULTS

GSEA indicated that BAG3, P4HB and TP53INP2 were involved in the inflammatory response and TNF-α signalling *via* NF- κB. Furthermore, PCR results showed significantly higher mRNA levels of BAG3 and P4HB and lower mRNA levels of FMR1 and TP53INP2 in the DSS group compared to the control group.

CONCLUSION

This study identified four m6A-ARGs that predict the occurrence of UC, thus providing a scientific reference for further studies on the pathogenesis of UC.

INTRODUCTION

Ulcerative colitis (UC) is a complex, chronic, immune-mediated, colitis disease. The UC lesions are mostly located in the sigmoid colon and rectum even the whole colon [1]. The clinical manifestations of UC can take many forms, with bloody diarrhoea as the most obvious early symptom. Other symptoms include abdominal pain, bloody stool, weight loss, tenesmus and vomiting. UC is characterised by the alternation of the active phase and remission phase; however, its clinical process is unclear [2]. The incidence rate of UC is on the rise worldwide, with treatment proving to be difficult. Approximately 15% of patients with UC experience an aggressive course, and some will even develop colorectal cancer. Current non-surgical treatment options include 5-aminosalicylic acid (5-ASA), glucocorticoids, immunosuppressants, biological agents and probiotics, which are limited by high recurrence rates and varying side effects. Current research on the pathogenesis of UC mainly focuses on microbiota, genetics, immunity and intestinal mucosal barrier. However, the exact pathogenesis of UC remains unclear. Therefore, exploring the aetiology of UC is of great significance for the diagnosis and treatment of UC [3].

N6-methyladenosine (m6A) methylation is one of the most common RNA modifications and plays a key role in the development and progression of various diseases. It has been reported that changes related to m6A are associated with intestinal

microbiota changes and gastrointestinal cancer development [4]. Currently, studies have investigated m6A methylation in UC, which revealed the role of m6A methylation in the pathogenesis of UC [5, 6].

Autophagy is a conservative degradation process, which is critical for regulating major cellular functions and biological metabolic processes. Thus, impaired autophagy could lead to many diseases, including cancer, cardiomyopathy, neurodegenerative diseases and ageing. Moreover, autophagy disorder can also lead to inflammation, intestinal barrier destruction and intestinal homeostasis imbalance, thus increasing the risk of colon disease[7]. Recently, autophagy-related gene polymorphisms have been reported to be strongly associated with an increased risk of UC. Additionally, the therapeutic effect of certain UC drugs is indicated to be mediated by regulating the autophagy pathway[8]. Recent studies have shown that the methylation of m6A RNA can regulate autophagy gene expression and affect autophagy function. Both m6A modification and autophagy play a key role in the occurrence and development of human diseases; however, the combined role of m6A and autophagy in UC remains unexplored[9].

In this study, we used publicly available data related to UC and comprehensive bioinformatics methods to elucidate the autophagy-related genes of m6A with a diagnostic value for UC. Additionally, animal models were also used to validate and explore the potential regulatory mechanism of m6A-autophagy in UC, thereby contributing to the development of treatment options for patients with UC.

MATERIALS AND METHODS

Data Extraction

UC-related datasets (GSE8747 and GSE75214) were downloaded from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/). The training set was the GSE87473 dataset, which comprised 21 normal and 106 UC samples. The GSE75214 dataset, comprising 11 normal and 97 UC samples, was used as the external validation set. In total, 222 ARGs were acquired from the Human Autophagy database

(http://www.autophagy.lu/index.html) (**Table S1**). Furthermore, 23 m6A-related genes were obtained from the published literature (**Table S1**)[10].

Identification and Functional Analysis of Differentially Expressed Genes (DEGs) Between the UC and Normal Groups

The mRNA expression levels between the UC and normal groups in the GSE87473 dataset were compared using the 'limma' package (version 3.52.4) with adj $P < 0.05$ and $|\log_2 FC| \geq 0.5$ [11]. Subsequently, the 'clusterProfiler' R package (version 4.4.4) was used to perform the biological functional enrichment analysis of DEGs between the UC and normal groups with Gene Ontology (GO) and Kyoto Encyclopaedia of Genes and Genomes (KEGG)($P < 0.05$)[12] .

Identification and Establishment of a Protein-Protein Interaction (PPI) Network of m6A-Autophagy-Related DEGs (m6A-AR DEGs)

Correlations between the m^6A-related genes and ARGs were calculated using the 'rcorr' function of the 'Hmisc' R package (version 4.7-1) ($P < 0.05$). The intersection of the m6A-autophagy-related genes (m^6A-ARGs) and DEGs was performed using the Venn tool to acquire m^6A-AR DEGs. STRING was adopted to explore the PPIs of m^6A-AR DEGs. Furthermore, Cytoscape (version 3.8.0) was utilised to visualise the co-expression network of m^6A-ARGs and the PPI network of m^6A-AR DEGs.

Screening of Characteristic Genes

Univariate logistic regression was used to initially screen variables in the identified m^6A-AR DEGs. Additionally, the least absolute shrinkage and selection operator (LASSO) and support vector machines (SVM) were further applied for screening characteristic genes. Moreover, the diagnostic value for UC of the characteristic gene was assessed using receiver operating characteristic (ROC) curves and the area under the curve (AUC).

Immune Infiltration Analysis

The CIBERSORT algorithm (version 1.03) was utilised to assess the infiltrating abundance of 22 immune cells between the UC and normal groups in the training set[13]. Differences between UC and normal groups were analysed using a Wilcox test, and a boxplot was plotted using the 'ggplot2' R package (version 3.3.6). Additionally, the correlation between characteristic genes and differential immune cells was analysed using the cor function of R language.

Gene Set Enrichment Analysis (GSEA) of Characteristic Genes

Based on the median value of the expression of the characteristic genes, the samples in GSE87473 were grouped into high and low-expression groups, and differential analysis was performed. Hallmark gene sets were downloaded using the 'msigdbr' R package (version 7.5.1) as a reference. Sorted DEGs were subjected to enrichment analysis (adj. $P < 0.05$).

Expression of Characteristic Genes in External Validation Datasets

To further demonstrate the validity of our results, the expression levels of the characteristic genes were compared between the UC and normal groups in the GSE87473 and GSE75214 datasets for external validation.

Animal

C57BL/6 mice (6-8 wk old, Quality Certification of Laboratory Animals: SCXK (Shanghai) 2017-0012) were housed in the Animal Experiment Centre of Shanghai University of Traditional Chinese Medicine, at a temperature of 22℃±2℃, alternating between light and dark (SPF class), 50±10% relative humidity. All animals had unlimited access to standard diet and water and general health status was checked daily by veterinarians. All animals (n = 20) were divided randomly into two groups (five animals per cage): control group (n = 10) and dextran sulfate sodium (DSS) group (n = 10). The mice in the DSS group were given DSS solution (3.5%) for seven days. Mice in the control group drank water normally. All animal experiments conformed to

the internationally accepted principles for the care and use of laboratory animals (No. PZSHUTCM210611001, the Animal Ethics Committee of the Shanghai University of Traditional Chinese Medicine).

Evaluation of disease activity index

Colitis was evaluated by disease activity index (DAI) score, including body weight, stool consistency, and fecal occult blood or gross bleeding[14]. DAI was calculated by grading on a scale of 0 to 4 using the following parameters: loss of body weight (0: normal; 1: 0–5%; 2: 5–10%; 3: 10–15%; 4: >15%); stool consistency (0: normal; 2: loose stools; 4: watery diarrhea); and occult blood (0: negative; 2: positive; 4: gross bleeding). The final result was expressed as the average of the three.

Sample collection

At the end of the experiments, all animals were euthanized by intraperitoneal injection of 80 mg/kg 1% sodium pentobarbital. After euthanasia, the whole colon was collected and the feces and surrounding connective tissue were removed. Freshly isolated organs were kept frozen at –80 °C for RNA isolation and analyses of gene expression.

Quantitative real-time polymerase chain reaction (qPCR)

After the experiment, colonic tissues were collected for quantitative real-time polymerase chain reaction (Table 1). Total RNA was extracted with TRIzol by homogenizing the tissue. Total RNA quality was assessed by measuring the absorbance at 260 and 280 nm using a NanoDrop-2000c spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and the 260/280 ratio ranged between 1.8 and 2.0. Then samples were used to reverse transcription and synthesize cDNA by the Evo M-MLV RT Premix. The samples were diluted using the SYBR® Green Premix Pro Taq HS qPCR Kit (High Rox Plus) (Accurate Biology, China). q-PCR and melting-curve analyses were performed using StepOnePlus Real-Time PCR System (Applied Biosystems, Foster City, CA, USA).The PCR reactions were set for 40 cycles. Each cycle was fixed at 95 °C for 30

s, 95 °C for 3 s, and 60 °C for 30 s. These primers have been validated for specificity and efficiency using conventional RT-PCR. Relative gene expression was calculated based on $2^{-\Delta\Delta Ct}$ method.

RESULTS

Acquisition and Enrichment Analysis of DEGs between the UC and Normal Groups

In total, 3512 DEGs were identified between the UC and normal groups in GSE87473 (Figure 1A-B). Enrichment analysis revealed a total of 1247 GO BP, 143 GO MF and 97 GO CC that were related to DEGs, such as small molecule catabolic process, oxidoreductase activity, catabolism of organic acids process, acting on CH-OH group of donors, mitochondrial matrix, peroxisomal matrix and microbody lumen (Figure 1C). Moreover, there were 67 KEGG pathways such as retinol metabolism, co-factors biosynthesis and cytokine-cytokine receptor interaction, that were related to DEGs (Figure 1D).

Construction of the PPI Network of m6A -AR DEGs

All 23 m6A-related genes and 222 ARGs were observed to be related to each other. Figure 2A presents the co-expression network of m6A-autophagy-related genes that contains 14 m6A-related genes and 64 ARGs ($|r| > 0.8$ and $P < 0.05$) (Table S2). A total of 43 m6A-AR DEGs were obtained by the intersection of m6A-autophagy-related genes and DEGs between the UC and normal groups (Figure 2B). Moreover, using the PPI network of m6A-AR DEGs, a strong reciprocal relationship was observed between heat shock protein family A (Hsp70) member 5(HSPA5) and dnaJ heat shock protein family member B9 (DNAJB9), HSPA5 and tumour protein P53 inducible nuclear protein 2 (TP53INP2), death-associated protein kinase 2 (DAPK2) and mitogen-activated protein kinase 3 (MAPK3) (Figure 2C). However, four proteins did not interact with other proteins, namely Cathepsin L (CTSL), eukaryotic elongation factor 2 kinase (EEF2K), a regulator of G protein signalling 19 (RGS19) and tumour suppressor candidate 1 (TUSC1) (Figure 2C).

Acquisition of Characteristic Genes

Univariate logistic regression revealed that all 43 m6A-AR DEGs were associated with the occurrence of UC (Figure 3A). Nine characteristic genes were further identified using LASSO, namely BLC2-associated athanogene 3 (BAG3), CC Chemokine Ligand 2 (CCL2), prolyl 4-hydroxylase subunit beta (P4HB), proliferation and apoptosis adaptor protein 15 (PEA15), serpin family a member 1(SERPINA1), fragile X mental retardation 1 (FMR1), MAPK3, TP53INP2 and TUSC1 (Figure 3B). From the SVM algorithm, we obtained nine eigengenes, namely DNAJB9, BAG3, TP53IMP2, recombinant human caspase-1 (CASP1), P4HB, breast tumour kinase (Brk)/protein-tyrosine kinase 6 (PTK6), DAPK2, FMR1, leucine-rich pentatricopeptide repeat containing (LRPPRC) (Figure 3C). Finally, four characteristic genes (FMR1, BAG3, P4HB and TP53IMP2) were selected through the cross-talk between LASSO and SVM (Figure 3D). The diagnostic accuracy of the four characteristic genes was evaluated using ROC curve analysis in GSE87466 and GSE75214. The AUC values of the four genes were greater than 0.7 in both datasets, indicating their high predictive accuracy for the occurrence of UC (Figure 3E-F).

Immuno-infiltration Analysis in the UC and Normal Groups

The infiltrating abundance of the 22 immune cells between the UC and normal groups was demonstrated using a bar chart, which showed that the content of B cells and T cells was higher in the UC group (Figure 4A). The correlation of infiltration levels of the 22 immune cells with each other was demonstrated using a heat map (Figure 4B). Moreover, 14 immune cells significantly differed between the UC and normal groups (Figure 4C), such as M0 macrophages, neutrophils and activated NK cells. Finally, correlation analysis between characteristic genes and differential immune cells showed that TP53INP2 had the strongest positive correlation with M2 macrophages, while FMR1 had the strongest negative correlation with naive B cells (Figure 4D).

GSEA of Characteristic Genes

To further understand the impact of the characteristic genes on the development of UC, we performed GSEA of the characteristic genes. BAG3, P4HB and TP53INP2 were found to be involved in the inflammatory response and tumour necrosis factor-α (TNF-α) signalling through nuclear factor kappa-B (NF-κB) (Figure 5A-C). Notably, BAG3 and P4HB were positively associated with the two signalling pathways, whereas TP53INP2 was negatively correlated. FMR1 was mainly enriched in adipogenesis, MYC targets V1, E2F targets, oxidative phosphorylation and fatty acid metabolism (Figure 5D).

The mRNA Levels of Characteristic Genes

The visualised data exhibited the expressions of BAG3, FMR1, P4HB and TP53IMP2. The expression of four characteristic genes were shown in the GSE87473 and GSE75214 datasets（Figure 6A-B）. The expression of the four characteristic genes between the normal and UC groups was significantly different. Moreover, the expression trends of the four genes were consistent in both datasets, with BAG3 and P4HB expressions elevated in the UC group whereas TP53INP2 and FMR1 expressions were lowered. The results suggested that these four genes had great diagnostic value in predicting the occurrence of UC.

Validation of the expression levels of four characteristic genes in DSS-induced colitis in mice

The control group mice had a good mental state, sensitive reaction, shiny hair, gradually increased body mass and formed faeces; The mice in the DSS group had a poor mental state, slow movement, matted hair, significantly reduced body weight and mucus purulent stool. Compared with the control group, the DAI of the DSS group mice increased significantly on the 7^{th} day of administration (Figure 7A). HE staining showed that the colon tissue structure in the control group mice was complete and orderly arranged, the goblet cells and crypt structure were normal, and there was no congestion, oedema or ulcer. Meanwhile, in the DSS group, the colonic tissue was

destroyed, goblet cells and crypts disappeared, a large number of inflammatory cells infiltrated the tissue and large ulcerative lesions were observed (Figure 7B). PCR also revealed significantly higher mRNA levels of BAG3 and P4HB and lower mRNA levels of FMR1 and TP53INP2 in the DSS group compared to the control group (Figure 7C). The flowchart systematically describes our study (Figure 8).

DISCUSSION

The role of m^{6}A modification, especially concerning autophagy regulation, in human diseases such as obesity, heart disease, azoospermia or oligozoospermia, intervertebral disc degeneration and cancer has been extensively studied by scholars. Such studies contribute to the optimisation of treatment strategies for various diseases[9, 15, 16]. However, research on the mechanism of m^{6}A autophagy interaction in colitis is still lacking.

In our study, four characteristic genes (FMR1, BAG3, P4HB and TP53INP2) were acquired using the machine learning algorithms. FMR1 is related to m^{6}A, while BAG3, P4HB and TP53INP2 are related to autophagy. FMR1, the gene responsible for fragile X syndrome, encodes the fragile X mental retardation protein (FMRP)[17]. BAG proteins compete with Hip for binding to the Hsc70/Hsp70 ATPase domain and promote substrate release. Additionally, diseases associated with BAG3 include myopathy, myofibrillar and cardiomyopathy[18]. P4HB is an important endoplasmic reticulum (ER) molecular chaperone. Traditionally, it regulates the post-translational modification of proteins in the ER, which in turn is crucial for cell proliferation, apoptosis and autophagy regulation [19]. The protein encoded by TP53INP2 promotes autophagy and is essential for proper autophagosome formation and processing[20]. Recent studies report that the relationship between these four genes and intestinal diseases is mainly studied in colorectal cancer, with very little attention being paid to colitis.

In this study, through machine learning and animal model verification, the expression of Fmr1 and TP53INP2 was observed to be reduced in colitis. Researchers report that the relative abundance of Proteus, Deironobacteria and Bacteroides in Fmr1 knockout

mice was higher than that in the wild-type (WT) mice, whereas that of Firmicutes and Tenericutes was lower in Fmr1 knockout mice than in the WT mice[21]. Therefore, we speculate that Fmr1 could affect the occurrence and development of colitis through the intestinal flora and its metabolites. Studies also report that the downregulation of TP53INP2 inhibits epithelial-to-mesenchymal transition (EMT) *via* the GSK-3β /β-catenin/Snail1 pathway in bladder cancer[22]. We hypothesised that TP53INP2 could inhibit the occurrence of EMT by regulating EMT-related transcription factors, thus blocking the key link required for UC-associated colorectal cancer (CAC) transformation. In the current study, the expression of BAG3 and P4HB was elevated in colitis. Studies have revealed that BAG3 overexpression promoted HCT-116 cell growth, migration and invasion in vitro. Contrastingly, BAG3 knockout inhibited HCT-116 cell growth, migration and invasion[23]. In view of the similar biological characteristics of HCT-116 and HT-29/ intestinal epithelial cells, we speculate that BAG3 is involved in the signalling pathway related to cell proliferation, migration, invasion and chemical resistance control in colitis. Moreover, P4HB has been reported to be highly expressed in patients with colorectal cancer and closely related to the degree of cancer differentiation[24]. Thus, we hypothesise that P4HB has potential as a molecular marker in the diagnosis and treatment of human CAC.

GSEA revealed that BAG3, P4HB and TP53INP2 were involved in the NF-κB/TNF-α signalling pathway. The NF-κB signalling pathway plays an important role in the development of UC. The loss of the NF-κB signalling pathway and its regulatory factors lead to pathological changes in the UC intestinal tract, which, acting as positive feedback, further activates NF-κB and aggravates inflammation. Cytokines such as TNF-α, interleukin (IL)-6 and IL-1β are influenced by NF-κB and regulate immunity and inflammation in different ways[25]. Therefore, BAG3, P4HB and TP53INP2-related NF- κB/TNF-α signalling pathways have curative potential in clinical UC treatment[26].

Immuno-infiltration analysis showed that TP53INP2 had the highest positive correlation with M2 macrophages, while FMR1 had the highest negative correlation with naive B cells. Intestinal macrophages are involved in intestinal immune

homeostasis and intestinal inflammation. The imbalance of the classical activated pro-inflammatory phenotype (M1)/alternative activated anti-inflammatory phenotype (M2) macrophage polarization can lead to intestinal inflammation. In UC patients, intestinal inflammation is closely related to the imbalance of intestinal M1/M2 macrophages polarization.[27]. Therefore, targeted therapy that promotes macrophage polarisation by regulating TP53INP2 can reconstruct the homeostasis of intestinal immune microenvironment and restore post-inflammatory tissue homeostasis, which is a new focus of UC therapy. Through single-cell RNA sequencing of B cells from three cohorts of patients with UC, researchers have drawn the composition, transcription and clonal map of the intestinal mucosa and circulating B cells and found major perturbations within the mucosal B cell compartment, including an expansion of naive B cells and IgG+ plasma cells with curtailed diversity and maturation. These findings suggest that B cells play an important role in the pathogenesis of UC[28]. Nevertheless, the mechanism of how *FMR1* regulates naive B cells requires further study.

CONCLUSION

In conclusion, our study analysed the diagnostic value of m^{6}A and ARGs in predicting the occurrence of UC. A total of four key genes (FMR1, BAG3, P4HB and TP53INP2) were identified and verified using animal models, providing a foundation for the clinical diagnosis and treatment of UC. However, m6A and ARGs how to participate in the occurrence and development of UC, as well as the genes identification are as possible markers for assessing UC severity and developing innovative UC targeted therapeutic approaches. The specific regulatory mechanisms of these genes need further experimental research and clinical application research.

91644_Auto_Edited.docx

ORIGINALITY REPORT

11%
SIMILARITY INDEX

PRIMARY SOURCES

1	Chen Chen, Bo Lan, Guanghong Xie, Zhaoyang Liu. "Analysis and identification of ferroptosis-related genes in ulcerative colitis", Scandinavian Journal of Gastroenterology, 2023 Crossref	83 words — 2%
2	www.ncbi.nlm.nih.gov Internet	64 words — 2%
3	www.researchgate.net Internet	63 words — 2%
4	www.cancerindex.org Internet	39 words — 1%
5	pubmed.ncbi.nlm.nih.gov Internet	29 words — 1%
6	f6publishing.blob.core.windows.net Internet	23 words — 1%
7	Qi Lv, Yao Xing, Jian Liu, Dong Dong, Yue Liu, Hongzhi Qiao, Yinan Zhang, Lihong Hu. "Lonicerin targets EZH2 to alleviate ulcerative colitis by autophagy-mediated NLRP3 inflammasome inactivation", Acta Pharmaceutica Sinica B, 2021 Crossref	16 words — < 1%

8	dokumen.pub Internet	16 words — < 1%
9	www.science.org Internet	15 words — < 1%
10	www.frontiersin.org Internet	14 words — < 1%
11	www.spandidos-publications.com Internet	13 words — < 1%
12	translational-medicine.biomedcentral.com Internet	12 words — < 1%
13	www.rgd.mcw.edu Internet	12 words — < 1%

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	OFF
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	< 12 WORDS