# scientific reports

OPEN

# Identification of ubiquitination-related key biomarkers and immune infiltration in Crohn's disease by bioinformatics analysis and machine learning

Wei Chen[1], Zeyan Xu[2], Haitao Sun[3], Wen Feng[4]✉ & Zhenhua Huang[3]✉

Crohn's disease (CD) is a chronic inflammatory bowel disease with an unknown etiology. Ubiquitination plays a significant role in the pathogenesis of CD. This study aimed to explore the functional roles of ubiquitination-related genes in CD. Differentially expressed ubiquitination-related genes were identified by intersecting differentially expressed genes (DEGs) from the GSE95095 dataset in the Gene Expression Omnibus (GEO) database with a set of ubiquitination-related genes. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed. Key genes were selected by combining hub genes from the protein-protein interaction (PPI) network with feature genes identified by Lasso and Random Forest (RF) algorithms. Additionally, the correlation between key genes and immune infiltration was assessed, and Gene Set Enrichment Analysis (GSEA) of key genes was conducted. The efficacy of key genes was validated using ROC curves in an external dataset, and their expression was confirmed in LPS-induced Caco-2 cells through RT-qPCR. A total of 32 ubiquitination-related DEGs were identified, and two key genes (UBE2R2, NEDD4L) were selected. The infiltration of M2 macrophages was reduced in CD patients, with UBE2R2 expression negatively correlated and NEDD4L expression positively correlated with M2 macrophage infiltration. GSEA indicated that UBE2R2 was enriched in terpenoid backbone biosynthesis, regulation of autophagy, and limonene and pinene degradation, while NEDD4L was enriched in lysosome, Wnt signaling, and calcium signaling pathways. ROC curves demonstrated superior efficacy for NEDD4L. In LPS-induced Caco-2 cells, UBE2R2 expression increased, while NEDD4L expression decreased. A comprehensive analysis of the functional relationship between ubiquitination-related genes and CD can enhance understanding of CD pathogenesis and suggest potential therapeutic targets.

**Keywords**  Biomarker, Bioinformatics, Machine learning, Ubiquitination, Crohn's disease

Crohn's disease (CD) is a chronic inflammatory bowel disease characterized by persistent inflammation of the gastrointestinal tract, most commonly affecting the small intestine and colon[1,2]. CD often leads to ulcers, strictures, and fistulas, resulting in symptoms such as abdominal pain, diarrhea, and weight loss[3]. Despite advancements in understanding CD, its etiology remains incompletely understood, involving complex interactions among genetic predisposition, environmental factors, and dysregulated immune responses[4,5]. Current treatments aim to manage inflammation and alleviate symptoms, but challenges such as treatment resistance, recurrent flares, and significant impacts on patients' quality of life remain.

Ubiquitination is a highly regulated process that plays an important role in maintaining protein homeostasis in cells[6]. Adding ubiquitin to a protein can alter its stability, localization, and activity, and can also serve as a signal to be degraded by the proteasome[7]. This process is crucial for controlling the expression of immune-related proteins, including cytokines, chemokines, and enzymes involved in the initiation and resolution

[1]Department of Gastroenterology, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 201620, China. [2]Department of Gastroenterology, Zhongda Hospital, Southeast University, Nanjing 210009, China. [3]General Surgery Department, Jiangsu University Affiliated People's Hospital, Zhenjiang 212000, China. [4]Department of Gastroenterology, Songjiang Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 201600, China. ✉email: 734001525@shsmu.edu.cn; zhenhuahuangzj@163.com

of inflammatory responses[8]. Chronic inflammation is a complex pathological condition resulting from dysregulated signaling pathways and immune responses. Ubiquitination is a key regulator of these processes because it controls the expression of pro-inflammatory cytokines and chemokines through proteasome-mediated degradation of transcription factors and their regulatory factors[9]. This regulation is essential for the relief of inflammation because it prevents the continued activation of inflammatory pathways that can lead to tissue damage and disease.

With the advancement of bioinformatics technologies, a comprehensive framework integrating genetics, epigenetics, proteomics, transcriptomics, and metabolic processes has emerged[10]. This framework has been increasingly applied to explore the diverse mechanisms driving the onset and progression of CD[11]. Recent studies leveraging bioinformatics tools have highlighted the potential role of ferroptosis in CD pathogenesis and identified possible therapeutic targets associated with this process[12]. This study aimed to investigate the regulatory role of ubiquitination in CD by bioinformatics analysis, and we sought to elucidate the mechanisms by which ubiquitination regulates key signaling pathways in CD to provide a basis for new therapeutic approaches.

## Materials and methods
### Microarray data
The GSE95095 dataset from a GEO database (https://www.ncbi.nlm.nih.gov/geo/), which consists of 24 patients with CD and 12 cases of normal control of intestinal tissue microarray data. Gene expression profiles in this dataset were constructed on the GPL14951 illumina HumanHT-12 WG-DASL V4.0 R2 platform. Additionally, the GSE83448 dataset contains the gene microarray data of 19 CD patients and 14 normal controls to verify the reliability of the findings from the GSE95095 dataset. The gene expression profiles in this dataset were constructed based on the GPL18134 CodeLink Human Whole Genome Array platform.

### Differential expression analysis
The process of identifying differential expression genes (DEGs) involved comparing the expression patterns of CD patients with those of normal healthy individuals. A stringent filter was applied by selecting genes with $|log2(FC)| > 1$ and an adjusted P-value $< 0.05$. In addition, we sought to identify ubiquitination-related genes that may be involved in the disease's pathology. To achieve this, we obtained a gene set from the Genecards database (https://www.genecards.org/) that was associated with ubiquitination. We narrowed down this list to 373 genes, which were selected based on a Relevance score higher than 10. Finally, we identified the intersection between these ubiquitination-related genes and the DEGs we previously identified. This allowed us to narrow down a subset of DEGs that were also related to the ubiquitin system.

### Machine learning
The Lasso algorithm is a powerful tool for analyzing and understanding high-dimensional data, especially for screening the key marker features of CD[13]. One way to choose an appropriate penalty parameter λ is to use ten-fold cross-validation, thus determining the best value of λ that leads to the lowest cross-validation error. The "glmnet" package is a valuable tool to implement this method, facilitating the analysis of high-dimensional data[14]. Random Forest builds a prediction model by sampling the variables, and the feature importance ranking can be obtained, and the analysis is run using the "randomForest" package[15].

### Protein-protein interaction
Using the STRING database, we searched for and created protein-protein interaction (PPI) networks for DEGs related to ubiquitination. We considered any interaction score greater than 0.4 to be statistically significant and omitted any single target protein nodes. The resulting data from the STRING database was then visualized into a PPI network via Cytoscape (version 3.10.0)[16]. Cytoscape is a powerful tool that enabled us to create a high-quality, interactive diagram of the PPI network, facilitating a more comprehensive understanding of the complex interrelationships between the identified proteins. To further analyze the PPI network, we utilized the cytoHubba plugin in Cytoscape[17].

### Immune infiltration analysis
The study used CIBERSORT and quanTIseq algorithm, to calculate the level of immune infiltration in various tissue samples[18]. In addition, we further analyzed the relationship between immune cells and key genes using the GraphPad Prism software (version 8.0.2) and calculated the proportion of immune cells in CD and normal control samples.

### Functional enrichment analysis
The Gene Ontology (GO) enrichment analysis is widely applied in functional studies of large-scale transcriptome or genome data, defining concepts for describing gene function, including biological processes (BP), cellular components (CC), and molecular functions (MF). To characterize the function and pathways of identified feature genes, we used biological tools for GO and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis to determine the function and metabolic pathways[19–21]. The clusterProfiler package (R Foundation) was used to explore the biological significance and pathways behind many genes. Single-gene GSEA analysis is developed from the basis of GSEA and used to analyze the correlation and regulatory mechanisms between a single gene and biological processes or disease states.

### Cell culture
The Caco-2 cell line was obtained from the Institute of Biochemistry and Cell Biology (Shanghai, China), and was cultured in DMEM medium containing 10%FBS, 100 U/mL penicillin, and 100 mg/mL streptomycin and

incubated with 5% $CO_2$ at 37 ℃. The cells were incubated with 10 ng/mL lipopolysaccharide (LPS) for 24 h as an in vitro CD model for subsequent gene expression verification experiments.

### Quantitative reverse transcription polymerase chain reaction (RT-qPCR)

Total RNA was extracted from cells using TRIzol (Takara, Tokyo, Japan), while total RNA from intestinal tissues was isolated using the same method, followed by cDNA synthesis using the RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). The expression of the target genes was detected using SYBR® Green Real-time PCR Master Mix (Toyobo, Osaka, Japan). Specific gene expression was detected using the following primers:

| Gene | primer(5'-3') | |
| --- | --- | --- |
| | F: | R: |
| human-UBE2R2 | ACGAGTTACCGCCTTCTTCA | TGAAGGCTGCTGGATCTCAA |
| human-NEDD4L | ATTAGGACGGGAGCTTGGAG | TACGGGGTTGAGAATGCAGT |
| human-GAPDH | CCACAGTCCATGCCATCAC | CGTTCAGCTCAGGGATGAC |
| mouse-UBE2R2 | ATGGAGGGACAGCAAAGGAA | GCTGTTGTCATTGGAAGGCA |
| mouse-NEDD4L | CTTTGCTGTATCTGCGTCCC | CAGTTGGAGAAGATCGCTGC |
| mouse-GAPDH | CAACTCCCACTCTTCCACCT | GAGTTGGGATAGGGCCTCTC |

### Construction of CD mice model

Male C57BL/6 mice (6–8 weeks old) were purchased from the Experimental Animal Center of the Shanghai Institutes for Biological Sciences. During chronic TNBS administration, mice were presensitized by applying 1% TNBS (Sigma-Aldrich) to their backs for one week. After a 24-hour fasting period, 100 µl of TNBS in 50% ethanol was administered rectally for six consecutive weeks. The concentrations of TNBS were 0.5%, 1.5%, 2.0%, 2.5%, 2.5%, and 2.5%, respectively. Seven days after the last administration, mice were sacrificed by intraperitoneal injection of 100 mg/kg sodium pentobarbital, and the distal colon was collected for analysis. All methods were carried out in accordance with relevant guidelines and regulations. All methods are reported in accordance with ARRIVE guidelines. All mouse experiments were approved by the Animal Protection and Utilization Committee of the Southeast University (Approval number: 20230301020). For histological analysis, mouse intestinal tissues were fixed in 4% paraformaldehyde (P0099, Beyotime, China) and embedded in paraffin. After sectioning the tissues, hematoxylin and eosin (HE) staining (G1120, Solarbio, China) was used to assess histological changes.

### Immunohistochemistry

For immunohistochemistry (IHC), paraffin-embedded tissue sections were subjected to deparaffinization, rehydration, and antigen retrieval using EDTA solution. To block nonspecific binding, the sections were incubated with normal goat serum. Primary antibodies against UBE2R2 (Invitrogen, PA5-31003, 1:1000) and NEDD4L (Abcam, ab46521, 1:1000) were applied and incubated overnight at 4 °C. The following day, after reheating the sections, they were treated with 3% hydrogen peroxide for 8 min to quench endogenous peroxidase activity, followed by washing with PBS. The sections were then incubated with the corresponding secondary antibody at room temperature for 1 h. Staining was performed using diaminobenzidine (DAB, Sigma, USA), and the slides were counterstained with hematoxylin. Images of the stained slides were captured and analyzed.

### Statistical analysis

GraphPad Prism software (version 8.0.2) and R software (version 4.0.3) were utilized for data analysis and visualization. Student's t-test (two-tailed) was employed to compare differences between two groups for normally distributed data. For non-normally distributed data, the Mann-Whitney U test was used. For correlations, Pearson's correlation coefficient was used to assess the strength and direction of linear relationships between two variables. Receiver operating characteristic (ROC) curve analysis was conducted to evaluate the diagnostic or predictive power of specific genes, and the area under the curve (AUC) was calculated to quantify performance. All statistical tests were performed with a significance threshold set at $P < 0.05$.

## Results

### Identification of ubiquitination-related DEGs

In order to identify DEGs between CD and normal controls base on the GSE95095 dataset, the limma package in the R software was utilized. A cutoff standard was set with $|log2(FC)| > 1$ and adjusted P-value < 0.05 (Fig. 1A). The volcano plot was used to display both upregulated and downregulated genes. As shown in Fig. 1B and 32 ubiquitination-related DEGs (Ub-DEGs) were obtained by intersecting 2824 DEGs with 373 ubiquitination-related genes. A heatmap was used to present the distribution of these 32 Ub-DEGs in each sample (Fig. 1C). In addition, we analyzed the correlation between each Ub-DEG (Fig. 1D).

### Biological function analysis

To further analyze the biological function of Ub-DEGs, GO and KEGG pathway analyses were performed on the up-regulated and down-regulated Ub-DEGs, respectively. The GO analysis results of up-regulated Ub-DEGs were mainly enriched in axonal transport of mitochondrion, inclusion body, and ubiquitin protein ligase binding (Fig. 2A). The KEGG pathway analysis results of up-regulated Ub-DEGs were mainly enriched in Pathways of
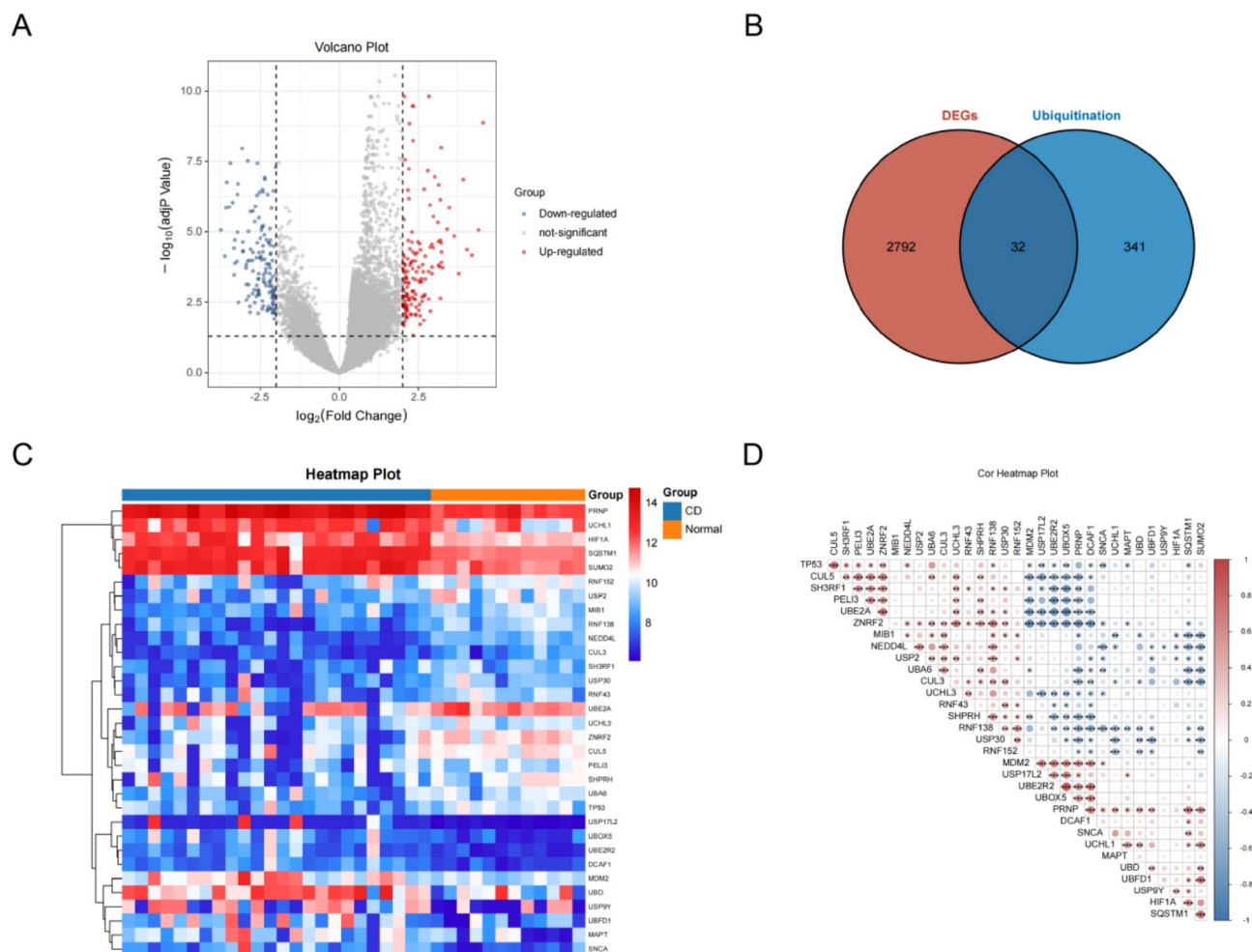
**Fig. 1**. Ub-DEGs identification. (**A**) Volcano plot of DEGs between CD and normal controls, where green dots represent down-regulated genes and red dots represent up-regulated genes. (**B**) Venn diagram between DEGs and ubiquitination related genes. (**C**) Heat map of 32 Ub-DEGs expression distribution. (**D**) Heat map of correlation between 32 Ub-DEGs.

neurodegeneration-multiple diseases, Ubiquitin mediated proteolysis, and Mitophagy-animal (Fig. 2B). The GO analysis results of down-regulated Ub-DEGs were mainly enriched in protein polyubiquitination, site of DNA damage, and ubiquitin-protein transferase activity (Fig. 2C). The KEGG pathway analysis results of down-regulated Ub-DEGs were mainly enriched in Ubiquitin mediated proteolysis, Mitophagy-animal, Wnt signaling pathway, and Ferroptosis (Fig. 2D).

## Identification of key genes

The PPI of 32 Ub-DEGs was analyzed using the String database and visualized by Cytoscape (Fig. 3A). To identify the most important genes for further exploration, we utilized both the BottleNeck algorithm and the EcCentricity algorithm through the cytoHubba plugin in the Cytoscape tool. Specifically, we analyzed 32 Ub-DEGs by these two algorithms, and each analysis yielded a list of 10 top-ranked candidate genes (Fig. 3B, C). Additionally, among Ub-DEGs, 13 promising genes were identified using the Lasso algorithm, and 16 important genes were identified using the RF algorithm. (Fig. 3D, E). Finally, key genes (UBE2R2, NEDD4L) were identified by taking the intersection of the gene sets obtained by the three algorithms (Fig. 3F).

## Infiltrating landscape of immune cells

To analyze the infiltration of immune cells, we used two algorithms, CIBERSORT and quanTIseq, for analysis. Firstly, the quanTIseq algorithm was used to obtain the proportions of immune cell infiltration in each sample (Fig. 4A). Secondly, the differences in immune cell scores between the CD and Normal groups were compared. In the CD group, the infiltration of Macrophages M2 was significantly lower than in the Normal group, while the infiltration of Monocytes and Tregs was significantly higher than in the Normal group (Fig. 4B). In addition, we explored the correlation between key genes and various immune cells. UBE2R2 was significantly positively correlated with Neutrophils and Tregs, but significantly negatively correlated with NK cells and Dendritic cells (Fig. 4C). NEDD4L was significantly positively correlated with Macrophages M2, but significantly negatively
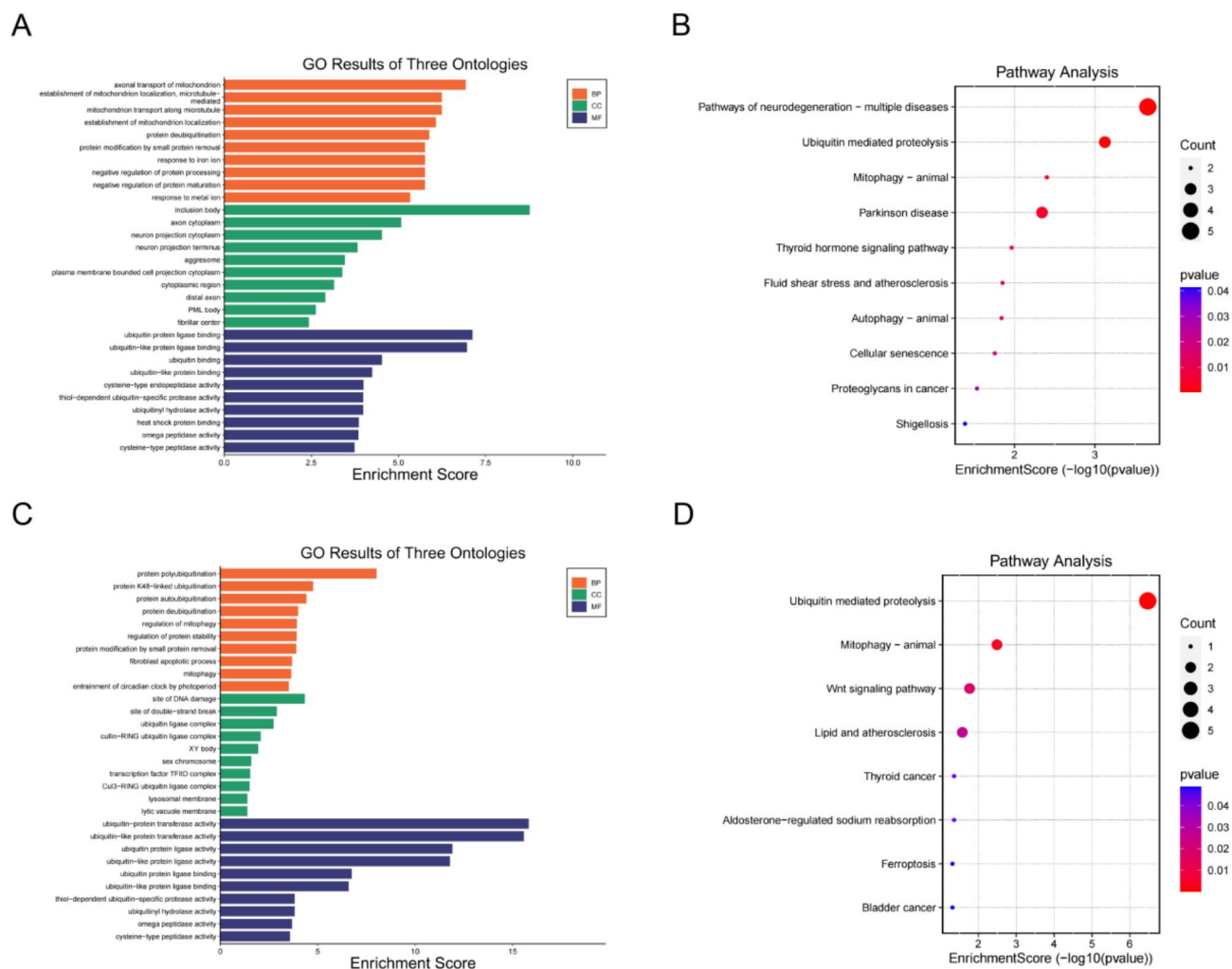
**Fig. 2.** GO and KEGG analysis of Ub-DEGs. (**A**) GO analysis of up-regulated Ub-DEGs. The longer the length, the higher the enrichment. (**B**) KEGG analysis of up-regulated Ub-DEGs. The larger the bubble, the greater the number of genes involved in the pathway. The redder the color, the more enriched it is. (**C**) GO analysis of down-regulated Ub-DEGs. (**D**) KEGG analysis of down-regulated Ub-DEGs.

correlated with B cells (Fig. 4D). Finally, we used the CIBERSORT algorithm to analyze the correlation between key genes and various immune cells again. UBE2R2 was positively correlated with T cells gamma delta, NK cells resting, and Macrophages M0, but negatively correlated with Plasma cells and Macrophages M2 (Fig. 4E). NEDD4L was positively correlated with Plasma cells, T cells CD4 memory resting, T cells regulatory (Tregs), and Macrophages M2, but negatively correlated with B cells memory, T cells CD8, and Dendritic cells activated (Fig. 4F).

### GSEA of key genes
The functional pathway enrichment analysis of UBE2R2 and NEDD4L was further performed by GSEA. GSEA analysis OF UBE2R2 showed that UBE2R2 was enriched to terpenoid backbone biosynthesis, regulation of autophagy, and limonene and pinene degradation (Fig. 5A). GSEA analysis of NEDD4L showed that NEDD4L was enriched to lysosome, Wnt signaling pathway, and calcium signaling pathway (Fig. 5B).

### Validation of key genes
In the GSE95095 dataset, the expression of UBE2R2 in the CD group was significantly higher than that in the Normal group, while the expression of NEDD4L in the CD group was significantly lower than that in the Normal group (Fig. 6A). ROC curve analysis was conducted on the accuracy of UBE2R2 and NEDD4L in the GSE95095 dataset, with AUC values of 94.0972% and 90.9722%, respectively (Fig. 6B). In addition, the ROC curve analysis results in the GSE83448 dataset showed AUC values of 69.9248% and 95.8847% for UBE2R2 and NEDD4L, respectively (Fig. 6C). Therefore, NEDD4L may have better predictive performance. Previous studies have used LPS-stimulated Caco-2 cells as an in vitro model of CD and TNBS-induced enteritis in rats as an in vivo model[22,23]. By RT-qPCR, we found that LPS promoted UBE2R2 expression but inhibited NEDD4L expression in Caco-2 cells (Fig. 6D). In addition, UBE2R2 expression was upregulated and NEDD4L expression
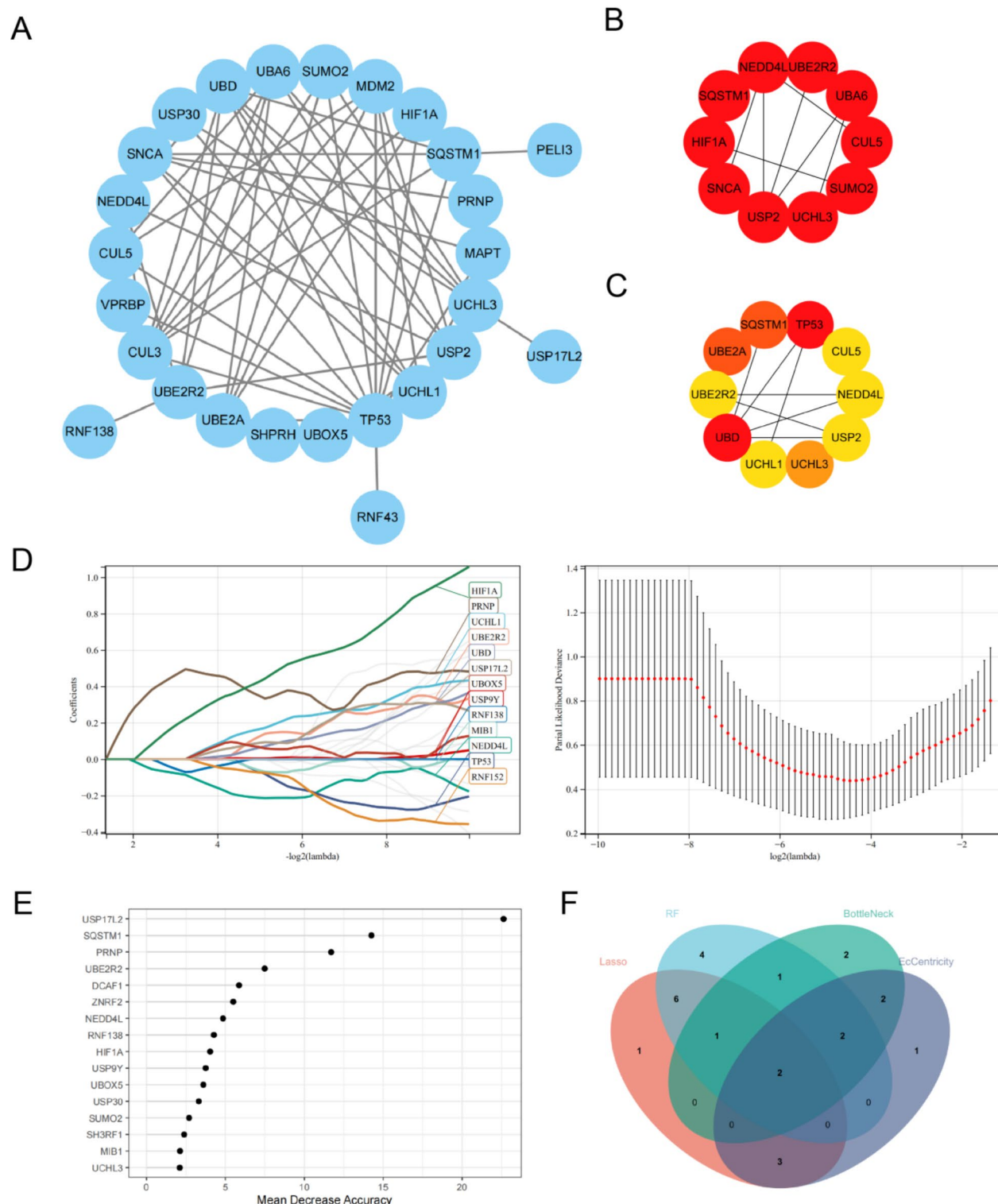
**Fig. 3.** Screening of key genes. (**A**) PPI network of Ub-DEGs. (**B**) Top ten genes of BottleNeck algorithm of cytoHubba plug-in in Cytoscape software. (**C**) Top ten genes of EcCentricity algorithm of cytoHubba plug-in in Cytoscape software. (**D**) Distribution of coefficients in the Lasso model. Plot of parameters in the Lasso model. (**E**) The ranking is based on the mean decrease accuracy value in the RF model. (**F**) Venn diagram between genes obtained by the three algorithms.
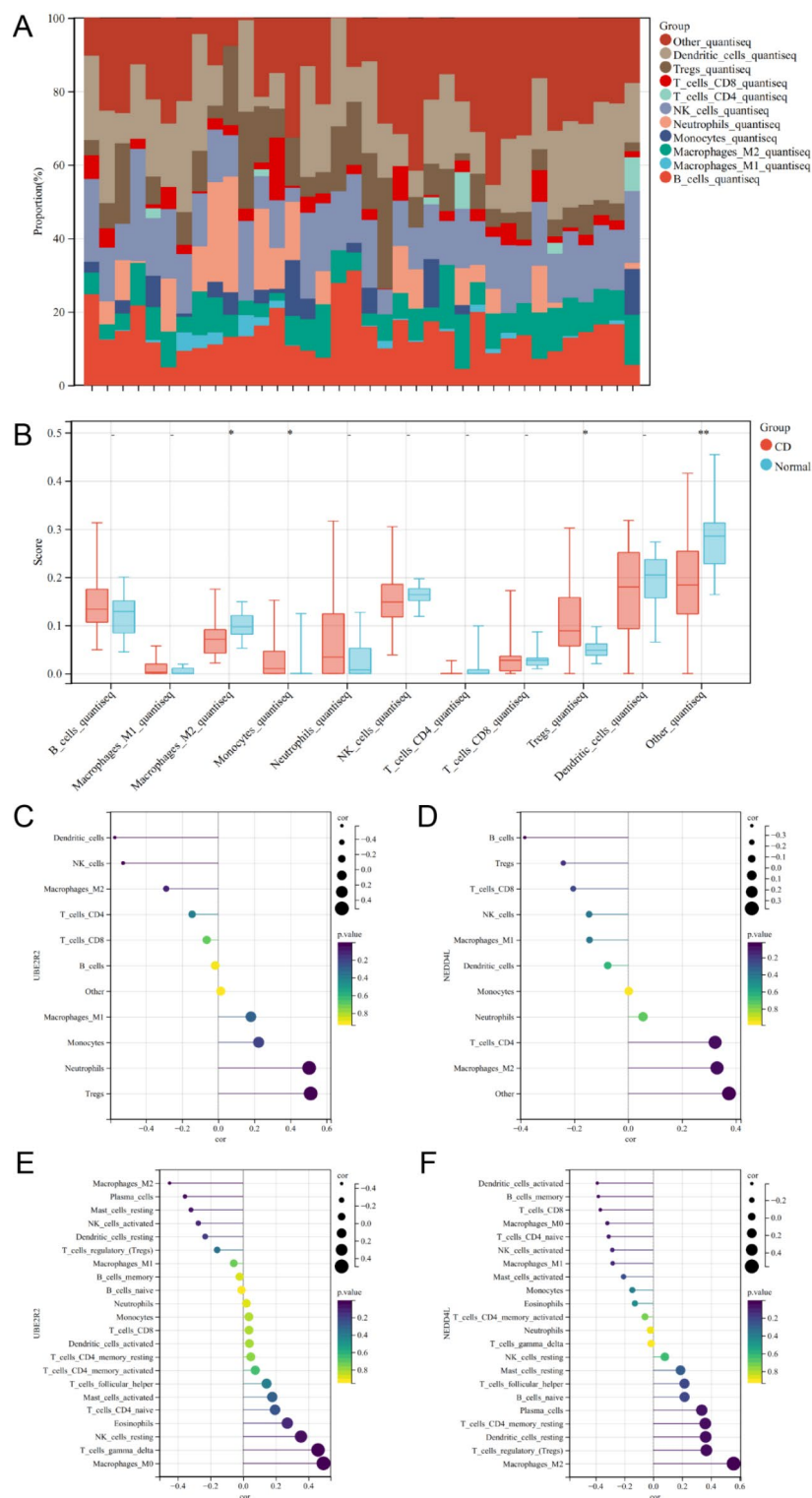
**Fig. 4**. Analysis of immune infiltration. (**A**) The proportion of each immune cell in the sample. (**B**) Differences in immune cell infiltration between CD and Normal groups. (**C**) The correlation between UBE2R2 and immune cells was based on the quanTIseq algorithm. (**D**) The correlation between NEDD4L and immune cells was based on the quanTIseq algorithm. (**E**) The correlation between UBE2R2 and immune cells was based on the CIBERSORT algorithm. (**F**) The correlation between NEDD4L and immune cells was based on the CIBERSORT algorithm. $*p < 0.05$, $**p < 0.01$.
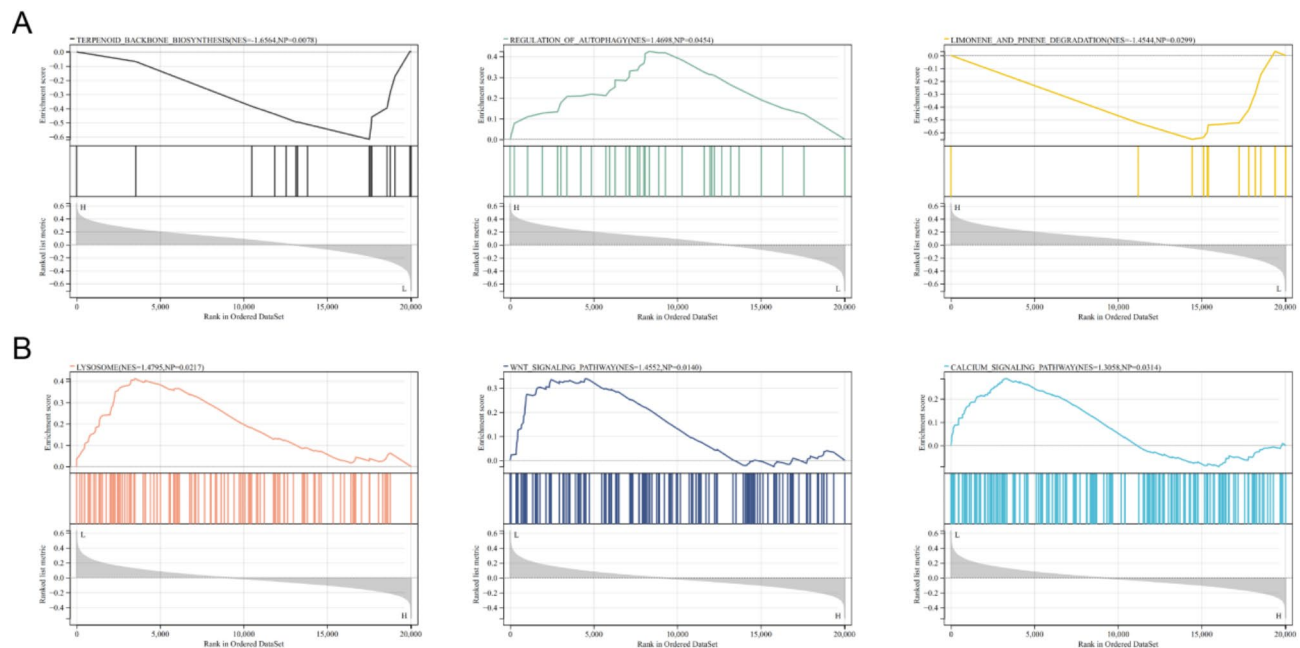
**Fig. 5**. KEGG pathway enriched by key genes were analyzed by GSEA. (**A**) UBE2R2 was enriched in terpenoid backbone biosynthesis, regulation of autophagy, and limonene and pinene degradation. (**B**) NEDD4L was enriched in lysosome, Wnt signaling pathway, and calcium signaling pathway.

was downregulated in TNBS-treated mice intestinal tissues by RT-qPCR and IHC (Fig. 6E–G). This is consistent with our previous study.

## Discussion

Ubiquitination is a post-translational modification process in which a small protein called ubiquitin is added to target proteins. This process plays a crucial role in regulating protein stability, degradation, and function, thereby affecting various cellular processes[24,25]. Alterations in the ubiquitin-proteasome system have been observed in patients with CD, which is responsible for regulating protein degradation in cells. Dysregulation of this system can lead to the accumulation of misfolded or abnormal proteins, triggering inflammatory responses and promoting the development of CD. Additionally, studies have shown that dysregulation of specific ubiquitin ligases and deubiquitinases can affect the immune response and intestinal inflammation in CD[26]. Abnormal activation or inhibition of these enzymes can disrupt the balance of immune cells in the intestine, leading to chronic inflammation and tissue damage. Apart from its role in immune regulation, the ubiquitination process is also associated with intestinal barrier function and tissue repair, both of which are crucial in the pathogenesis of CD[27]. Impairment of ubiquitination can disrupt the integrity of the intestinal barrier, allowing harmful substances to penetrate the intestinal wall and trigger an immune response. Furthermore, defects in ubiquitin-mediated tissue repair mechanisms may hinder the healing of inflamed intestinal tissues, perpetuating the cycle of inflammation and contributing to disease progression[8]. Therefore, based on the GEO database, we explored the mechanisms underlying the interaction between ubiquitination and CD using a set of ubiquitination-related genes and identified key biomarkers.

In this study, GO and KEGG analysis results showed that Ub-DEGs were enriched in protein hydrolysis, mitochondrial autophagy, DNA damage, and ferroptosis, which is consistent with previous studies. Mitochondria are commonly referred to as the powerhouse of the cell and play a crucial role in various aspects of intestinal epithelial cell function[28]. They are responsible for generating ATP energy through oxidative phosphorylation (OXPHOS) and also play a critical role in regulating cell metabolism, immune responses, stress reactions, and programmed cell death[29]. Mitochondrial redox imbalance refers to the imbalance in the generation and clearance of reactive oxygen species (ROS) and reactive nitrogen species (RON) within the mitochondria[30]. These imbalances have significant effects on the gut microbiota and intestinal environment[31]. Mitochondrial OXPHOS is the process by which mitochondria generate cellular energy currency ATP. During this process, ROS is produced as a byproduct, which is usually regulated by the cellular ROS clearance system. However, ROS overload can cause mitochondrial oxidative stress damage, leading to the activation of the PARKIN-PINK1 pathway and inducing mitochondrial autophagy[32]. Inhibition of mitochondrial autophagy can also lead to the accumulation of mtROS, activating NLRP3 and promoting the production of inflammatory factor IL-1, resulting in intestinal inflammation[33]. In addition, ROS overload induces NLRP3 de-ubiquitination, promoting inflammasome formation and aggravating intestinal inflammation[34]. Ferroptosis is a unique form of cell death characterized by the accumulation of intracellular iron ions and the production of ROS. Increased ROS levels simultaneously initiate the ferroptosis program in intestinal epithelial cells. GPX4 can reduce lipid peroxidation and inhibit ferroptosis of intestinal epithelial cells, thereby protecting the intestinal mucosa[35]. Therefore,
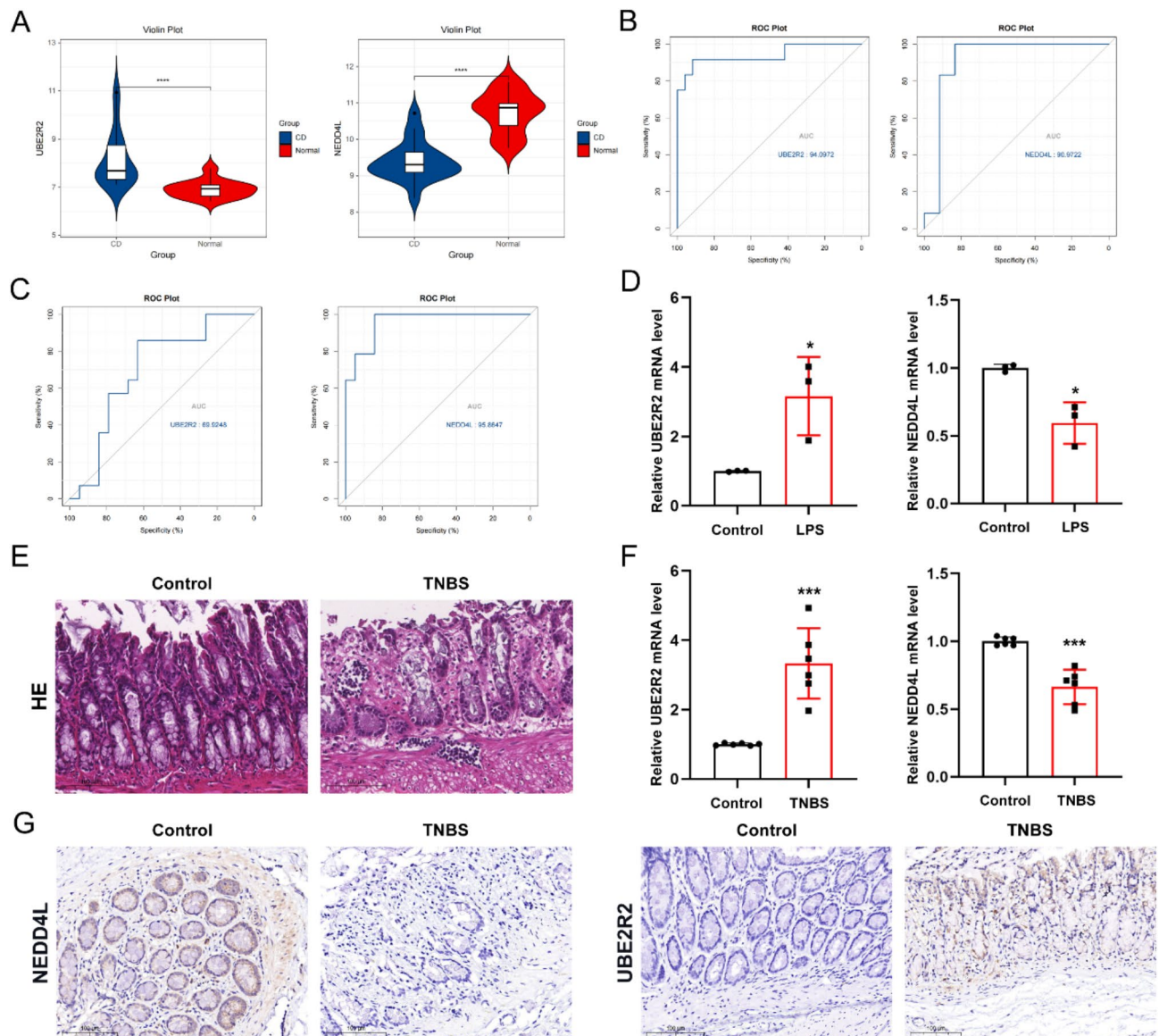
**Fig. 6**. Validation of key genes. (**A**) Differential expression of UBE2R2 and NEDD4L was identified in the GSE95095 dataset. (**B**) The ROC curves of UBE2R2 and NEDD4L were in the GSE95095 dataset. (**C**) The ROC curves of UBE2R2 and NEDD4L were in the GSE83448 dataset. (**D**) Expression levels of UBE2R2 and NEDD4L mRNA in Caco-2 cells induced by LPS. (**E**) The HE staining in Control and CD mice. (**F**) Expression levels of UBE2R2 and NEDD4L mRNA in Control and CD mice. (**G**) Immunohistochemical staining of UBE2R2 and NEDD4L in Control and CD mice. $*p < 0.05$, $***p < 0.001$, $****p < 0.0001$.

Mitochondrial autophagy may be involved in disease progression by regulating the release of inflammatory factors, while DNA damage and ferroptosis may promote the development of CD through cell injury and destruction. However, Ubiquitination may play an important role in regulating such a mechanism.

We combined the BottleNeck algorithm and EcCentricity algorithm from Cytoscape software, as well as the Lasso algorithm, to filter the Ub-DEGs. Through this, we obtained two key genes (UBE2R2 and NEDD4L). Both genes have good predictive efficiency and are validated in external datasets. We also found increased UBE2R2 expression and decreased NEDD4L expression in LPS-treated Caco-2 cells. Similarly, UBE2R2 expression was increased while NEDD4L expression was decreased in the intestinal tissues of TNBS-treated mice. To further explore the roles of UBE2R2 and NEDD4L, we noted that the expression of UBE2R2 was negatively correlated with the infiltration of M2 macrophages, while the expression of NEDD4L was positively correlated with the infiltration of M2 macrophages in both immune infiltration analysis results. UBE2R2 and NEDD4L may be potential diagnostic biomarkers and immunotherapy targets for CD. Furthermore, GSEA analysis revealed that UBE2R2 is involved in the regulation of autophagy, while NEDD4L is involved in the Wnt signaling pathway. The covalent connection between ubiquitin and specific target proteins is mainly achieved through a stepwise enzymatic cascade reaction. Ubiquitin is attached to the substrate through the synergistic action of ubiquitin-

activating enzyme (E1), ubiquitin-conjugating enzyme (E2), and ubiquitin ligase (E3). Among them, the E2 enzyme is the crucial transfer point for ubiquitination, interacting with both E1 and E3 and plays a decisive role in determining whether the marked protein will be degraded or involved in non-protein hydrolysis processes[36]. Research has shown that UBE2R2 can increase the ubiquitination and degradation of p53[37]. NEDD4L promotes the ubiquitination of MEKK2 and inhibits the protein stability of MEKK2, thereby inhibiting the production of pro-inflammatory cytokines and chemokines induced by IL-17, IL-1α, and TNF-α[38]. In addition, in inflammatory bowel diseases, ABIN3 negatively regulates programmed necrosis-induced intestinal inflammation by limiting the ubiquitination of RIPK3[26]. The protein level of RNF183 is overexpressed in the inflamed colon tissues of mice stimulated with TNBS and IBD patients. It interacts with IκBα and induces ubiquitin-mediated degradation of IκBα, thereby promoting the activation of NF-κB signaling and the release of inflammatory factors[39]. Therefore, exploring the functional mechanisms of ubiquitination in CD provides a perspective on the disease's pathogenesis and development.

This study has certain limitations. It is a bioinformatics research based on public databases, and more clinical sample data is needed to validate the results. In addition, the specific mechanism of action of key genes requires further in vivo and in vitro experiments for in-depth study.

## Conclusion

Based on the identification of Ub-DEGs in CD using the GEO database, we screened out two key genes (UBE2R2 and NEDD4L) through PPI network and Lasso algorithm and observed the immune cell infiltration landscape in CD. More importantly, we emphasized the correlation between UBE2R2 and NEDD4L and immune cells, as well as the possible functional mechanisms they may be involved in. Understanding the complex relationship between CD and ubiquitination may provide us with a new understanding of the potential mechanisms driving the disease and thus develop more effective treatment methods.

## Data availability

The datasets analysed during the current study are available in the Gene Expression Omnibus repository (https ://www.ncbi.nlm.nih.gov/geo/).

## References

1. Magalhaes, F. C. B. et al. Crohn's disease: Review and standardization of nomenclature. *Radiol. Bras.* **56** (2), 95–101 (2023).
2. Seyedian, S. S., Nokhostin, F. & Malamir, M. D. A review of the diagnosis, prevention, and treatment methods of inflammatory bowel disease. *J. Med. Life.* **12** (2), 113–122 (2019).
3. Choudhury, A. et al. Differentiating gastrointestinal tuberculosis and Crohn's disease: A comprehensive review. *BMC Gastroenterol.* **23** (1), 246 (2023).
4. Singh, N. & Bernstein, C. N. Environmental risk factors for inflammatory bowel disease. *United Eur. Gastroenterol. J.* **10** (10), 1047–1053 (2022).
5. Park, J. & Cheon, J. H. Incidence and prevalence of inflammatory bowel disease across Asia. *Yonsei Med. J.* **62** (2), 99–108 (2021).
6. Martinez-Ferriz, A., Ferrando, A., Fathinajafabadi, A. & Farras, R. Ubiquitin-mediated mechanisms of translational control. *Semin Cell. Dev. Biol.* **132**, 146–154 (2022).
7. Cockram, P. E. et al. Ubiquitination in the regulation of inflammatory cell death and cancer. *Cell. Death Differ.* **28** (2), 591–605 (2021).
8. Ruan, J., Schluter, D., Naumann, M., Waisman, A. & Wang, X. Ubiquitin-modifying enzymes as regulators of colitis. *Trends Mol. Med.* **28** (4), 304–318 (2022).
9. Madiraju, C., Novack, J. P., Reed, J. C. & Matsuzawa, S. I. K63 ubiquitination in immune signaling. *Trends Immunol.* **43** (2), 148–162 (2022).
10. Rodriguez, R. & Krishnan, Y. The chemistry of next-generation sequencing. *Nat. Biotechnol.* **41** (12), 1709–1715 (2023).
11. Gao, X. et al. Integrative multi-omics deciphers the spatial characteristics of host-gut microbiota interactions in Crohn's disease. *Cell. Rep. Med.* **4** (6), 101083 (2023).
12. Huang, R. et al. Identifying immune cell infiltration and effective diagnostic biomarkers in Crohn's disease by bioinformatics analysis. *Front. Immunol.* **14**, 1162473 (2023).
13. Liu, C. et al. Identification of crucial genes for predicting the risk of atherosclerosis with system lupus erythematosus based on comprehensive bioinformatics analysis and machine learning. *Comput. Biol. Med.* **152**, 106388 (2023).
14. Jiang, H., Chen, H., Wang, Y. & Qian, Y. Novel molecular Subtyping Scheme based on in Silico Analysis of Cuproptosis Regulator gene patterns optimizes Survival Prediction and Treatment of Hepatocellular Carcinoma. *J. Clin. Med.* ;**12**(18). (2023).
15. Wang, Y. et al. Unveiling the key genes, environmental toxins, and drug exposures in modulating the severity of ulcerative colitis: A comprehensive analysis. *Front. Immunol.* **14**, 1162458 (2023).
16. Majeed, A. & Mukhtar, S. Protein-protein Interaction Network Exploration using Cytoscape. *Methods Mol. Biol.* **2690**, 419–427 (2023).
17. Liang, Z. et al. Identification of shared gene signatures and molecular mechanisms between chronic kidney disease and ulcerative colitis. *Front. Immunol.* **14**, 1078310 (2023).
18. Wang, X., Jing, H. & Li, H. A novel cuproptosis-related lncRNA signature to predict prognosis and immune landscape of lung adenocarcinoma. *Transl. Lung Cancer Res.* **12** (2), 230–246 (2023).
19. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28** (1), 27–30 (2000).
20. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28** (11), 1947–1951 (2019).
21. Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* **51** (D1), D587–D592 (2023).
22. Shen, Z. et al. Roseburia intestinalis stimulates TLR5-dependent intestinal immunity against Crohn's disease. *EBioMedicine* **85**, 104285 (2022).
23. Liu, J. et al. Diosmetin Maintains Barrier Integrity by reducing the expression of ABCG2 in Colonic epithelial cells. *J. Agric. Food Chem.* **71** (23), 8931–8940 (2023).
24. Mark, K. G. & Rape, M. Ubiquitin-dependent regulation of transcription in development and disease. *EMBO Rep.* **22** (4), e51078 (2021).

25. Celebi, G., Kesim, H., Ozer, E. & Kutlu, O. The effect of dysfunctional ubiquitin enzymes in the Pathogenesis of Most Common diseases. *Int. J. Mol. Sci.* ;**21**(17). (2020).
26. Zhou, M. et al. ABIN3 negatively regulates necroptosis-induced intestinal inflammation through recruiting A20 and restricting the ubiquitination of RIPK3 in inflammatory bowel disease. *J. Crohns Colitis.* **15** (1), 99–114 (2021).
27. Zhang, H. et al. RNF186 regulates EFNB1 (ephrin B1)-EPHB2-induced autophagy in the colonic epithelial cells for the maintenance of intestinal homeostasis. *Autophagy* **17** (10), 3030–3047 (2021).
28. Zhang, Y., Zhang, J. & Duan, L. The role of microbiota-mitochondria crosstalk in pathogenesis and therapy of intestinal diseases. *Pharmacol. Res.* **186**, 106530 (2022).
29. Castellanos, E. & Lanning, N. J. Phosphorylation of OXPHOS Machinery subunits: Functional implications in Cell Biology and Disease. *Yale J. Biol. Med.* **92** (3), 523–531 (2019).
30. Willems, P. H., Rossignol, R., Dieteren, C. E., Murphy, M. P. & Koopman, W. J. Redox Homeostasis and mitochondrial dynamics. *Cell. Metab.* **22** (2), 207–218 (2015).
31. Mossad, O. et al. Gut microbiota drives age-related oxidative stress and mitochondrial damage in microglia via the metabolite N(6)-carboxymethyllysine. *Nat. Neurosci.* **25** (3), 295–305 (2022).
32. Tanaka, K. The PINK1-Parkin axis: An overview. *Neurosci. Res.* **159**, 9–15 (2020).
33. Xia, B. et al. Lactobacillus johnsonii L531 ameliorates enteritis via elimination of damaged mitochondria and suppression of SQSTM1-dependent mitophagy in a Salmonella infantis model of piglet diarrhea. *FASEB J.* **34** (2), 2821–2839 (2020).
34. Liu, L. et al. Giardia Duodenalis and its secreted PPIB trigger inflammasome activation and Pyroptosis in macrophages through TLR4-Induced ROS Signaling and A20-Mediated NLRP3 deubiquitination. *Cells* **10**(12). (2021).
35. Wang, X. et al. Resveratrol reduces ROS-induced ferroptosis by activating SIRT3 and compensating the GSH/GPX4 pathway. *Mol. Med.* **29** (1), 137 (2023).
36. Zhang, X. et al. Mechanism and Disease Association with a Ubiquitin conjugating E2 enzyme: UBE2L3. *Front. Immunol.* **13**, 793610 (2022).
37. Zhang, L. et al. SMYD3 promotes epithelial ovarian cancer metastasis by downregulating p53 protein stability and promoting p53 ubiquitination. *Carcinogenesis* **40** (12), 1492–1503 (2019).
38. Li, H. et al. E3 ubiquitin ligase NEDD4L negatively regulates inflammation by promoting ubiquitination of MEKK2. *EMBO Rep.* **23** (11), e54603 (2022).
39. Yu, Q. et al. E3 ubiquitin ligase RNF183 is a Novel Regulator in Inflammatory Bowel Disease. *J. Crohns Colitis.* **10** (6), 713–725 (2016).

## Acknowledgements

## Author contributions

WC designed the topic and conducted the bioinformatics analysis and the writing of the manuscript; ZYX performed the experiments; HTS carried out the visualization; WF supervised the project and provided the funding; ZHH revised the manuscript and provided the funding.

## Funding

## Declarations

## Competing interests

The authors declare no competing interests.

## Ethics declarations

Not applicable.

## Additional information

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-88148-4.

**Correspondence** and requests for materials should be addressed to W.F. or Z.H.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.