BMC Pharmacology and Toxicology

RESEARCH

Open Access

Identifying Lactylation-related biomarkers and therapeutic drugs in ulcerative colitis: insights from machine learning and molecular docking

Yao Yang[1,2†], Xu Sun[2†], Bin Liu[3†], Yunshu Zhang[2], Tong Xie[1,2], Junchen Li[1,2*], Jifeng Liu[1*] and Qingkai Zhang[1*]

Abstract

Background Ulcerative colitis (UC), a chronic relapsing-remitting inflammatory bowel disease. Recent studies have shown that lactylation modifications may be involved in metabolic-immune interactions in intestinal inflammation through epigenetic regulation, but their specific mechanisms in UC still require in-depth validation.

Methods We conducted comparative analyses of transcriptomic profiles, immune landscapes, and functional pathways between UC and normal cohorts. Lactylation-related differentially expressed genes were subjected to enrichment analysis to delineate their mechanistic roles in UC. Through machine learning algorithms, the diagnostic model was established. Further elucidating the mechanisms and regulatory network of the model gene in UC were GSVA, immunological correlation analysis, transcription factor prediction, immunofluorescence, and single-cell analysis. Lastly, the CMap database and molecular docking technology were used to investigate possible treatment drugs for UC.

Results Twenty-two lactylation-related differentially expressed genes were identified, predominantly enriched in actin cytoskeleton organization and JAK-STAT signaling. By utilizing machine learning methods, 3 model genes (S100A11, IFI16, and HSDL2) were identified. ROC curves from the train and test cohorts illustrate the superior diagnostic value of our model. Further comprehensive bioinformatics analyses revealed that these three core genes may be involved in the development of UC by regulating the metabolic and immune microenvironment. Finally, regorafenib and R-428 were considered as possible agents for the treatment of UC.

†Yao Yang, Xu Sun and Bin Liu have contributed equally to this work and share first authorship.

*Correspondence:
Junchen Li
lijc03@dmu.edu.cn
Jifeng Liu
jifeng0213@163.com
Qingkai Zhang
dlkaiyu@163.com

Full list of author information is available at the end of the article

Conclusion This study offers a novel strategy to early UC diagnosis and treatment by thoroughly characterizing lactylation modifications in UC.

Keywords Ulcerative colitis, Lactylation, Machine learning, Biomarkers, Molecular docking

Introduction

Ulcerative colitis (UC), a chronic relapsing bowel illness, is typified by fibrosis, mucosal damage, and intestinal inflammation [1]. Abdominal discomfort, weight loss, and bloody diarrhea are the most typical signs of UC. In recent years, UC has borne a substantial global illness burden and a growing incidence across the globe [2]. Although the precise origin of UC is uncertain, recent research indicates that immunological responses, genetic predisposition, and microbial imbalances may all be involved [3, 4]. The immune system is crucial to the development and incidence of UC, which mostly shows up as a dysregulation of the immunological response to intestinal microbes, resulting in the gastrointestinal tract's increasing destruction and poor repair [5]. Current therapy approaches are still insufficient, and many patients do not have lasting remission despite advancements in our knowledge of the pathophysiology of UC. The need for new biomarkers and treatment targets is highlighted by the heterogeneity of UC and its intricate molecular and immunological underpinnings.

The post-translational modification known as lactylation, which adds lactate-derived lactyl groups to lysine residues, is a crucial regulatory mechanism in cellular metabolism and gene expression [6]. Protein function can be altered by lactylation, which also affects a number of biological functions, such as inflammation and immunological responses [7, 8]. Recent studies have demonstrated the significance of lactylation in controlling the metabolism and activity of immune cells, pointing to a possible involvement in the pathophysiology of UC. Furthermore, a mouse model of DSS colitis has demonstrated that intestinal inflammation can be reduced by lactic acid ingestion through gut bacteria [9], whereas TNBS-induced colitis is prevented by topical lactic acid treatment [10]. Thus, increased intestinal lactate might be linked to the development of UC, but its precise role needs to be further investigated. Additionally, there have been no reports on the expression and functional roles of lactylation-related genes in UC.

Recent developments in machine learning and bioinformatics have improved our ability to precisely evaluate massive data sets. Key regulatory genes may be prioritized, hidden patterns in gene expression data can be found, and prediction models for illness diagnosis and treatment response can be built using machine learning methods. In this study, model genes were screened by machine analysis of multiple datasets to develop a robust diagnostic model that offered fresh perspectives on the immunological and metabolic processes that connect lactylation and UC, as well as offered possible biomarkers and treatment targets for UC intervention in the future.

Methods

Data collection from GEO databases

The three UC datasets in this study were all from the GEO database. GSE87466 [11] (87 UC and 21 normal samples) and GSE38713 [12] (30 UC and 13 normal samples) were combined into the training set, and GSE75214 [13] (97 UC and 11 normal samples) was the validation set. Then, we acquired 332 lactylation-related genes from the previous studies [14, 15]. The "ComBat" function from the "sva" package was then used to adjust for batch effects [16]. We used principal component analysis (PCA) to compare the data quality before and after batch removal in order to evaluate the effectiveness of this modification [17].

Identifying differentially expressed genes (DEGs) and lactylation-related genes of UC

The "limma" software was used to identify DEGs between the normal and UC groups, with an adjusted p value of less than 0.05 and a threshold of $|\log2FC| > 0.585$ [18]. Gene Set Enrichment Analysis (GSEA) is utilized to compare the biological process activity or pathway enrichment of UC and normal samples [19]. Finally, by using a Venn diagram to intersect lactylation-related genes and DEGs, lactylation-related DEGs of UC were discovered.

Functional enrichment analysis

Protein interactions were examined in the STRING database (https://cn.string-db.org/), with a composite score of more than 0.15 indicating the scientific validity of the interactions [20]. A database called GeneMANIA (https://genemania.org/) has ranked genes according to their functional tests [21].Functional enrichment studies were conducted using the Metascape database (https://metascape.org/), constructed to provide a comprehensive resource for annotating and analyzing gene lists to evaluate the biological roles and routes implicated in particular genes [22].

Identification of UC core lactylation-related genes by machine learning algorithms

Three machine learning methods were applied to identify hub genes using a unified input matrix of 151 samples × 22 DEGs, with binary classification labels ("Control" vs. "Treat"). LASSO ("glmnet" R package) employed L1

regularization and cross-validation (nfolds = 10) to optimize λ.min [23, 24]. RF ("randomForest" R package) trained 500 trees (ntree = 500), adopted out-of-bag error for optimal tree selection, and filtered genes with importance > 4 [25]. SVM-RFE ("e1071" R package) utilized recursive feature elimination with 10-fold validation, selecting the minimal-error feature subset [26]. The core genes for the subsequent study were then chosen from the intersections of the three subsets. The "pROC" tool was used to calculate the ROC curves and evaluate the model's predictability [27].

Cell culture and quantitative real-time PCR analysis

The Caco-2 cell line was cultivated in MEM with 1% penicillin-streptomycin and 20% fetal bovine serum. The HT-29 cell line was cultivated in McCoy's 5 A with 1% penicillin–streptomycin and 10% fetal bovine serum. Caco-2 and HT-29 cells were exposed to lipopolysaccharide (LPS) (1 μg/mL) for 24 h to induce a colitis model, according to relevant literature [28, 29].

Using the Trizol reagent, total RNA was isolated from cells and a reverse transcription kit was used to create cDNA. A fluorescent dye-based assay based on SYBR Green I was then used to evaluate the target genes' expression levels. The ΔΔCt method was used to assess and quantify the levels of RNA expression, and the t-test was used to compare the target gene expression levels between the two groups. IL-6 primer (forward: GGTGT TGCCTGCTGCCTTCC; reverse: GTTCTGAAGAGGT GAGTGGCTGTC), IL-1β primer (forward: GACCTGG ACCTCTGCCCTCTG; reverse: GCCTGCCTGAAGC CCTTGC), TNF-α primer (forward: TGGCGTGGAGC TGAGAGATAACC; reverse: GACGGCGATGCGGCT GATG), S100A11 primer (forward: CTCGCTCAGCTC CAACATG; reverse: AGGAACTCTGTCTTGGAGAG), IFI16 primer (forward: ACTGAGTACAACAAAGCCA TTTGA; reverse: TTGTGACATTGTCCTGTCCCCAC), HSDL2 primer (forward: AAGCCACTCAAGCAATCT ATCTG; reverse: GCTCTCCATATCCGACATTCCC), and β-actin primer (forward: CCTGGGCATGGAGTCC TGTG; reverse: TCTTCATTGTGCTGGGTGCC) were the primers used.

Gene set variation analysis (GSVA) and Immunoinfiltration analysis

Using the "GSVA" software, enrichment analysis of GO and KEGG pathways was carried out for model genes [30]. We examined the immunological microenvironmental differences between UC patients and healthy individuals using the CIBERSORT method, which determines the percentages of infiltration of 22 distinct immune cell types. Furthermore, the relationship between model genes and immune cells was investigated further, drawing on the findings of CIBERSORT.

Transcription factor (TF) prediction, cellular localization, single-cell mapping analysis of hub genes

The hub genes for TFs are predicted using TFTF online [31] which integrates the five main TF-Target online tools: hTFtarget [32], ENCODE [33], JASPAR [34], GTRD [35], and ChIP_Atlas [36]. Meanwhile, the Human Protein Atlas (HPA: https://www.proteinatlas.org/) was utilized to examine the hub genes' immunofluorescence and single-cell type atlases.

Finding possible small molecule medicines for UC

The CMap database (https://clue.io/) links diseases, genes, and drugs based on similar or opposing gene expression profiles [37]. The DEGs of UC were added to the CMap database in order to identify potential small molecule medications for the treatment of UC. The AutoDock tool was then used to determine the hydrogenation and charge of the proteins after the protein structures of the feature genes were obtained from the PDB database. To find the molecular structure of the active substance, use the PubChem database. Rotatable bonds and the charge balance of small molecules are examined using the AutoDock device. To generate docking energy, AutoDock Vina runs docking simulations. Lastly, the PyMol tool was used to analyze the docking complex.

Results

Analysis of DEGs and functional roles in UC

The study was illustrated in the flowchart (Fig. 1). First, we used the "sva" software tool to integrate two UC datasets, and the differences between batches are essentially removed (Supplementary Fig. 1A-D). Then, UC and normal samples were subjected to differential expression analysis in order to uncover potential pathogenic genes. Supplementary Fig. 1E displays the DEGs of normal people and UC patients. The CIBERSORT algorithm was used to assess the proportion of various immune cell subpopulations infiltrated in each sample. The results showed that UC patients exhibited higher levels of B cell memory infiltration, while normal individuals had higher levels of Tregs (Supplementary Fig. 1F). GSEA enrichment analysis showed that in UC patients, viral protein interaction with cytokine and cytokine receptor, TNF signaling pathway, and inflammatory bowel disease pathway were enriched (Fig. 2A-C), while valine, leucine and isoleucine degradation, pyruvate metabolism, and butanoate metabolism were down-regulated (Fig. 2D-F).

Selection and functional enrichment of lactylation-related DEGs

To identify potential lactylation-related DEGs, we analyzed their intersection using a Venn diagram. This approach revealed 22 lactylation-related DEGs (Fig. 3A). The STRING database screened the DEGs PPI network

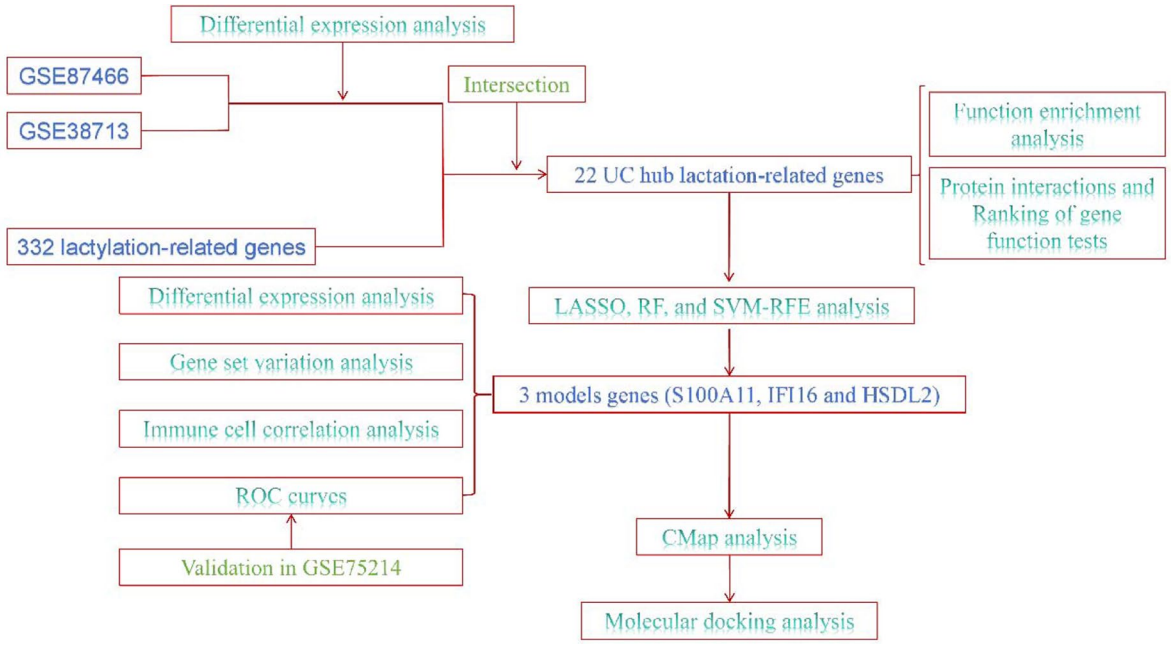

Fig. 1 The process of data analyzing in this study

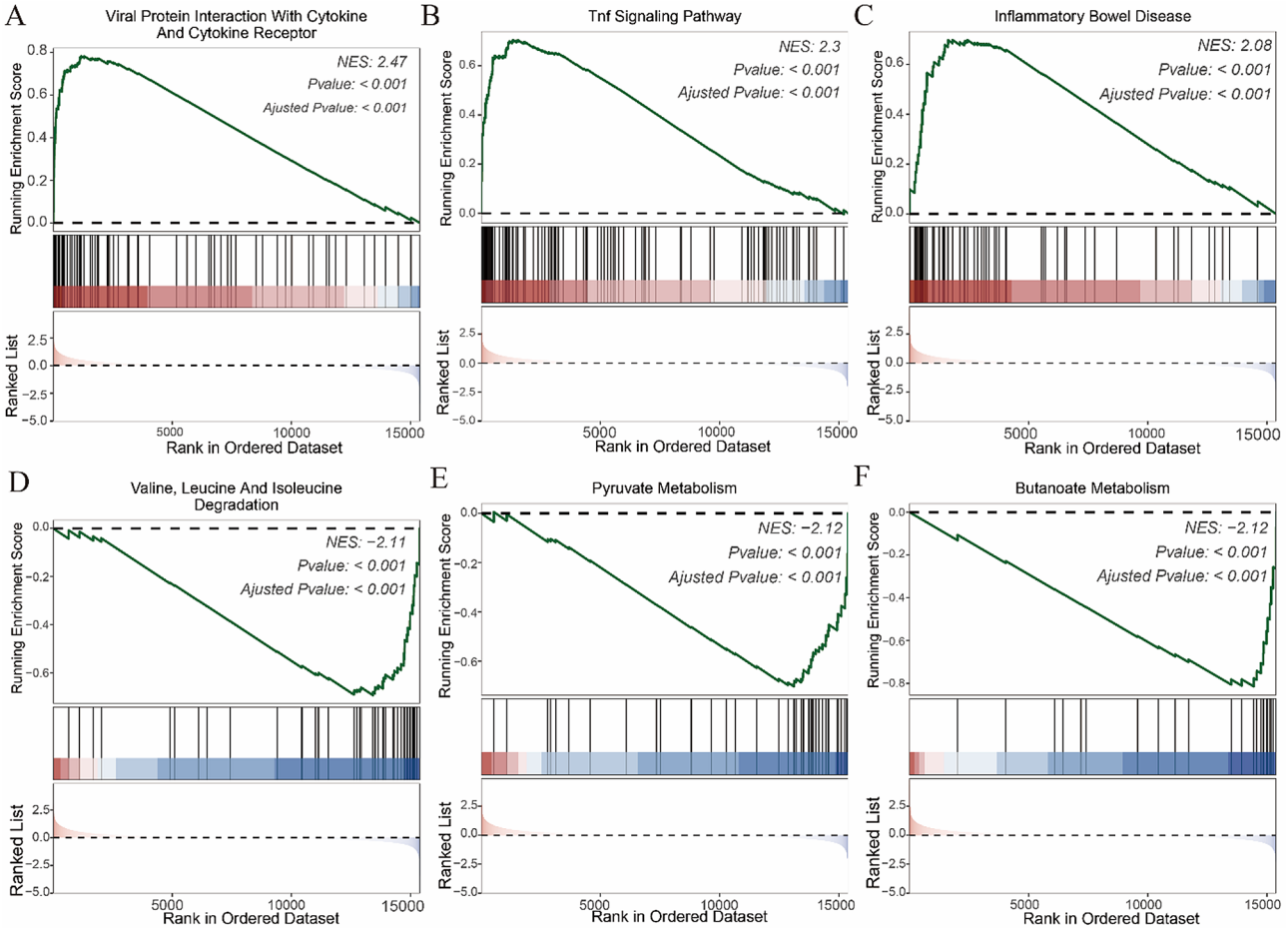

Fig. 2 GSEA for UC samples. (**A-C**) Pathways enriched in UC samples; (**D-F**) Pathways downregulated in UC samples

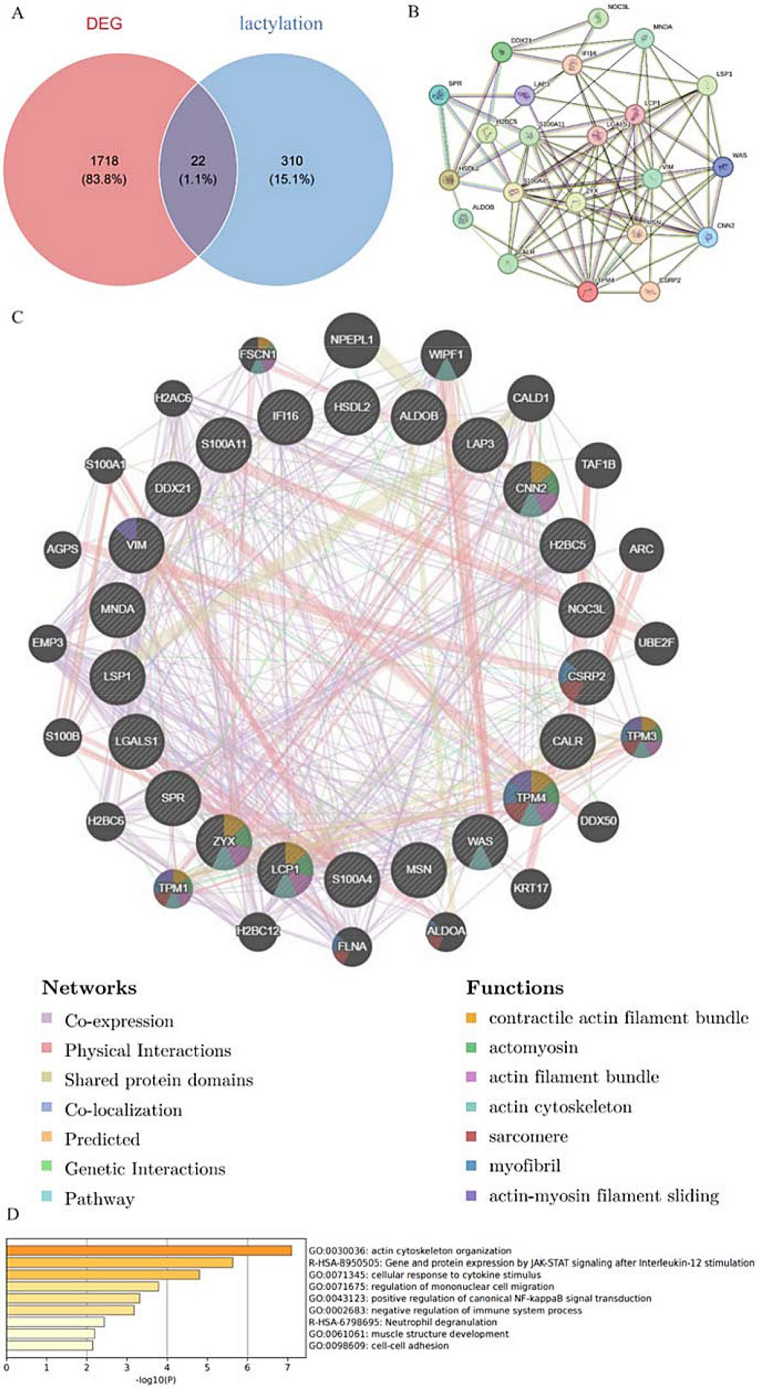

Fig. 3 Selection of hub genes. (**A**) Venn diagram of the lactylation-related genes and the DEGs of UC; (**B**) PPI of the intersection genes; (**C**) GeneMANIA analysis of the intersection genes; (**D**) Functional enrichment analysis of the intersection genes based on Metascape database

and displayed possible interactions between proteins represented by similar DEGs (Fig. 3B). Next, in order to anticipate connections between colocalization, pathways, shared protein domains, co-expression, and prediction, we analyzed 22 intersection genes using GeneMANIA. As shown in Fig. 3C, the network shows that they were mostly involved in the positive regulation of the actin cytoskeleton, contractile actin filament bundle, and actin filament bundle. They were most heavily enriched in JAK-STAT signaling and actin cytoskeleton organization, according to functional enrichment analysis based on the the Metascape database (Fig. 3D).

Construction of the diagnostic model for UC

Three machine learning models were created using the 22 lactylation-related DEGs to find possible biomarkers for the diagnosis of UC. Using LASSO regression, nine genes were identified as possible diagnostic markers (Fig. 4A-B). The SVM algorithm further refined this list, extracting seventeen candidate biomarkers (Fig. 4C-D). In the RF algorithm, four genes with importance

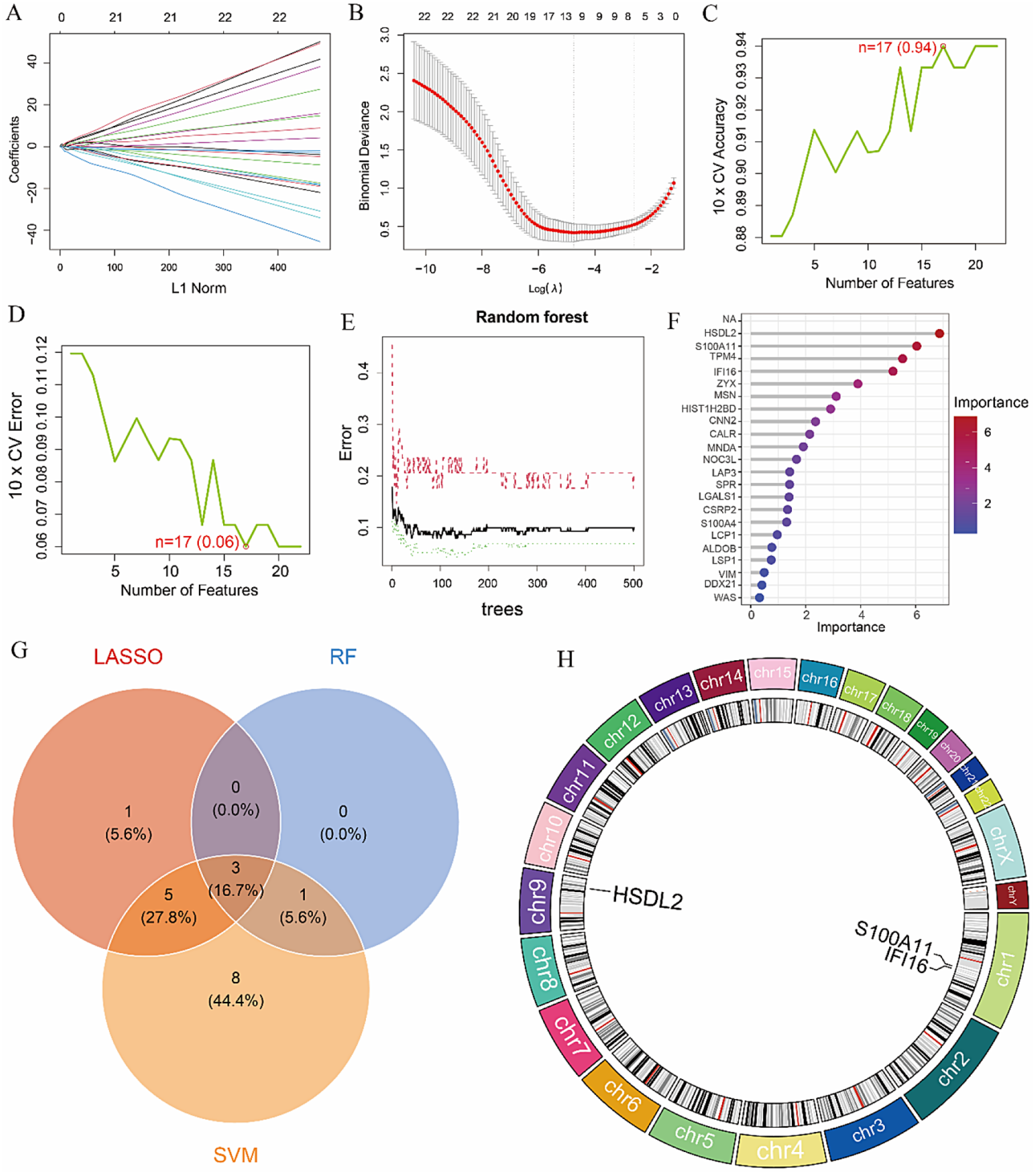

Fig. 4 Using machine learning techniques to determine the model genes for UC. (**A-B**) The LASSO regression selected 9 genes based on minimum lambda values; (**C-D**) 17 genes were identified using SVM; (**E-F**) RF assessed the importance of all genes to obtain four genes with scores for importance greater than 4; (**G**) The Venn diagram showing the intersection of three machine learning models; (**H**) Chromosomal locations of 3 model genes

scores greater than 4 were selected for further analysis (Fig. 4E-F). Finally, the Venn diagram was used to intersect these results, and three key genes—S100A11, IFI16, and HSDL2—were highlighted as robust diagnostic biomarkers (Fig. 4G). Their location on the chromosome was shown in Fig. 4H.

Evaluation of the diagnostic model

The correlations among the three key genes are illustrated in Fig. 5A. Specifically, S100A11 exhibited a positive correlation with IFI16, whereas HSDL2 was negatively correlated with both S100A11 and IFI16. Expression analysis revealed that S100A11 and IFI16

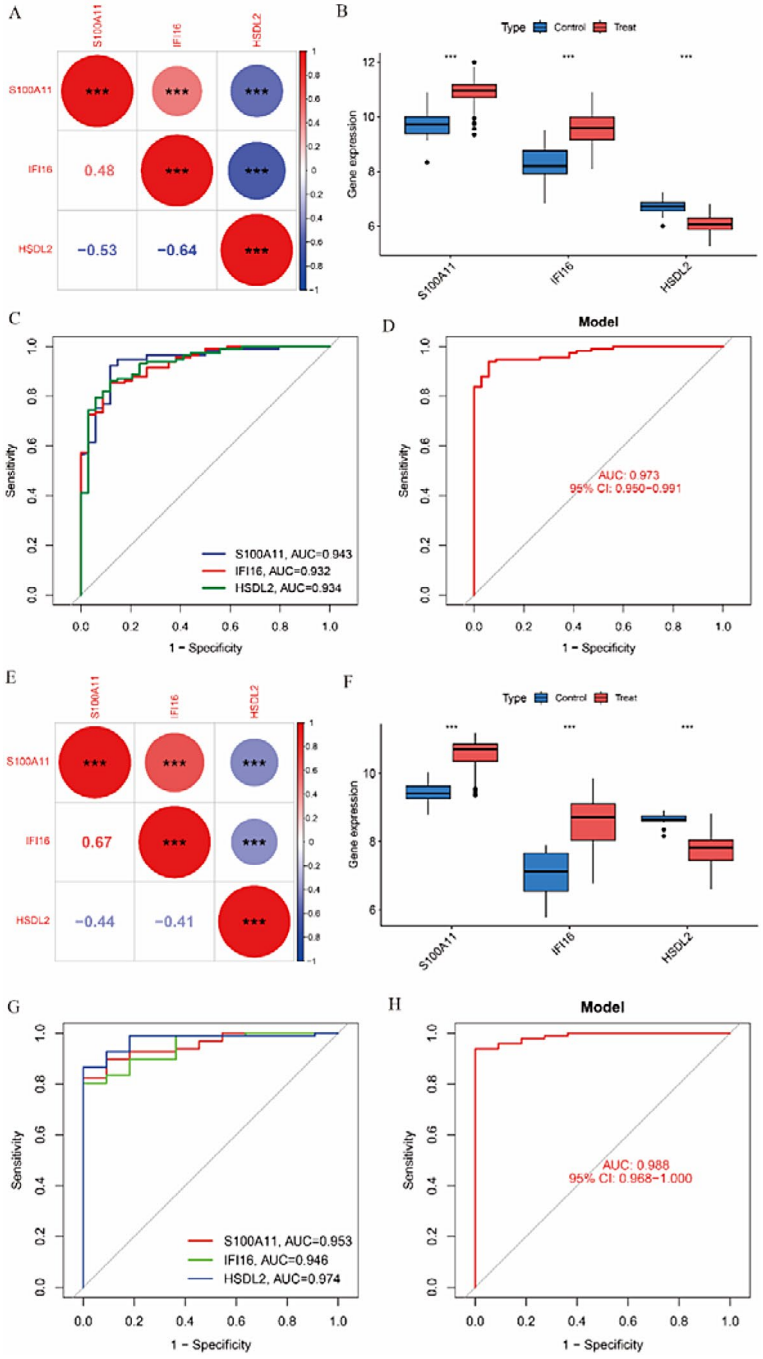

Fig. 5 Diagnostic effect of the model on UC. (**A**) Correlation analysis between the three pattern genes in the training set; (**B**) Box plots illustrating the expression difference in S100A11, IFI16, and HSDL2 between UC and normal samples of the training set; (**C-D**) ROC curves for the three genes and model in the training set; (**E**) Correlation analysis between the three genes in the validation set; (**F**) Box plots illustrating the expression difference in S100A11, IFI16, and HSDL2 between UC and normal samples of the validation set; (**G-H**) ROC curves for the three genes and model in the validation set

were significantly upregulated, while HSDL2 was downregulated in UC patients (Fig. 5B). ROC curve analysis showed that these three genes had good predictive ability (Fig. 5C). Notably, the three-gene prediction model achieved outstanding diagnostic performance, with an AUC of 0.973 (Fig. 5D). To further substantiate the model's accuracy, validation set analysis was conducted, reaffirming the robustness and diagnostic efficacy of our three-gene model (Fig. 5E-H).

To further validate the expression levels of the three model genes in UC, we established cellular models of UC by stimulating the cell lines Caco-2 and HT-29 with LPS. As shown in Supplementary Fig. 2A-B, the expression levels of IL-6, IL-1β and TNF-α in the model group were significantly higher than those in the control group, indicating that the UC model had been successfully established. Subsequently, we assessed the expression levels of the three model genes in the cell models. The results showed that the expression of S100A11 and IFI16 was up-regulated in the UC group while the opposite was true for HSDL2, which was generally consistent with the results of bioinformatics analysis (Supplementary Fig. 2C-D).

GSVA and immune correlation analysis of model genes

Each model gene pathway enrichment difference was then analyzed using GSVA in order to further investigate the possible processes of the three model genes in UC. The high S100A11 and IFI16 expression group was mainly enriched in some metabolism-related processes and pathways, while the opposite was true for HSDL2 (Supplementary Fig. 3A-F). The connection between the three model genes and immunological chemicals was then investigated in further detail. According to the findings, S100A11 primarily had a negative correlation with T cells CD8 and a positive correlation with neutrophils and macrophages M1. (Fig. 6A-D). IFI16 was positively correlated with neutrophils and macrophages M0 and negatively correlated with macrophages M2 (Fig. 6E-H). In contrast, HSDL2 was negatively correlated with macrophages M0 and neutrophils. There was a positive correlation with macrophages M2 (Fig. 6I-L). Figure 6M shows the relationship between model genes and immune cells and the strength of the correlation between immune cells. These results suggest that these three model genes may work together to play a role in UC by regulating these metabolic processes and immune cells.

The TFs, Immunofluorescence, and single-cell maps analysis of hub genes

TFs play a key role in controlling gene expression and determining the phenotypes of cells and organisms. Thus, in order to characterize TFs in S100A11, IFI16, and HSDL2, we integrated five databases (Fig. 7A-C). Subsequently, the intersection of them was taken to obtain two common TFs, CTCF and STAT1, which may be involved in regulating these three model genes (Fig. 7D). The cellular location of three model genes was then investigated. The immunofluorescence analysis revealed distinct subcellular localizations of the target proteins: S100A11 was detected in the nucleoplasm, cytosol, and nuclear speckles; IFI16 was observed in the nucleoplasm, nucleoli, and cytosol; while HSDL2 showed exclusive mitochondrial localization. In the representative immunofluorescence images, the target proteins are shown in green, microtubules are labeled in red, and nuclei are counterstained in blue (Fig. 7E-G). Additionally, we used the HPA database to assess their expression in the colon at the single-cell data level. S100A11 was mainly distributed in Paneth cells and distal enterocytes (Fig. 7H). IFI16 was mainly distributed in T cells (Fig. 7I). HSDL2 was mainly distributed in undifferentiated cells and enteroendocrine cells (Fig. 7J).

Finding possible small-molecule drugs to treat UC

Additionally, we entered the DEGs in UC into the CMap database for analysis in order to look for possible medications to treat UC and prevent the disease from developing. Supplementary Table 1 lists the top 10 smaller-molecule drugs that were anticipated to be promising medications for the treatment of UC. AutoDock was used for molecular docking of 10 most sensitive drugs and 3 model genes. The molecular docking results of ten drugs and model genes were shown in Supplementary Table 2. It can be found that regorafenib and R-428 have the best docking effect with the 3 model genes, indicating that they may be potential therapeutic drugs for patients with UC. The docking fraction of regorafenib with S100A11 and R-428 with IFI16 and HSDL2 was less than − 9.0 kcal/mol, indicating that these two drugs have high binding affinity with target genes and can interact well with model genes and were most likely to be potential therapeutic drugs and action targets for UC. The binding postures and loci of the regorafenib and R-428 to the model gene with the strongest affinity were shown in Fig. 8A-C, where red represents the compound and yellow dashed lines represent hydrogen bond interactions.

Discussion

UC is a common inflammatory condition affecting the gastrointestinal tract [38]. While the pathogenesis of epithelial damage caused by an abnormal inflammatory reaction is unclear, it has been hypothesized that its etiology may be related to factors such as diet, genetics, and intestinal flora [39, 40]. The incidence and prevalence of UC have increased globally in recent years due to changes in the dietary choices and style of living of people in emerging nations [41]. However, there are certain

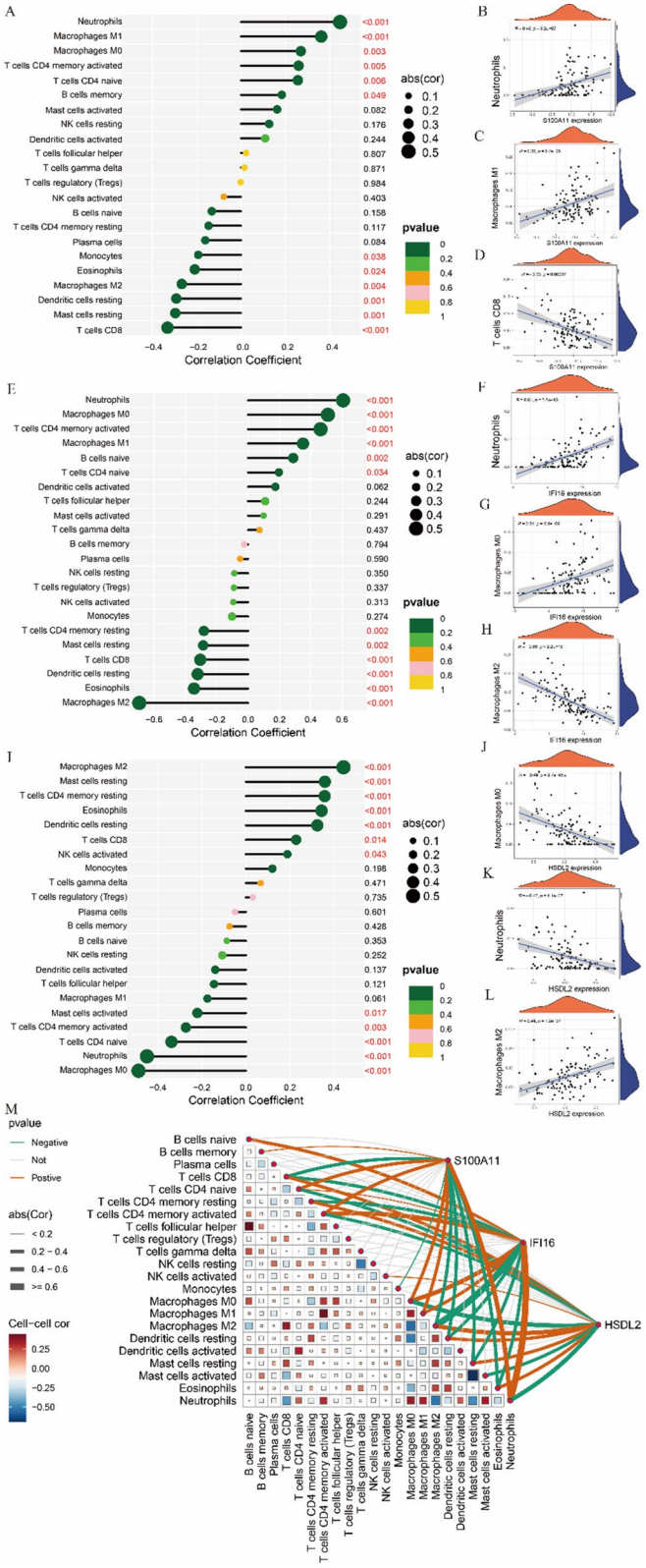

Fig. 6 Immune cell correlation analysis of model genes. (**A**) Correlations between various immune cell types and S100A11; Correlation between S100A11 and (**B**) neutrophils, (**C**) macrophages M1, (**D**) CD8 T cells; (**E**) Correlations between various immune cell types and IFI16; Correlation between IFI16 and (**F**) neutrophils, (**G**) macrophages M0, (**H**) macrophages M2; (**I**) Correlations between various immune cell types and HSDL2; Correlation between HSDL2 and (**J**) macrophages M0, (**K**) neutrophils, (**L**) macrophages M2; (**M**) Analysis of correlation between model genes and immune cells

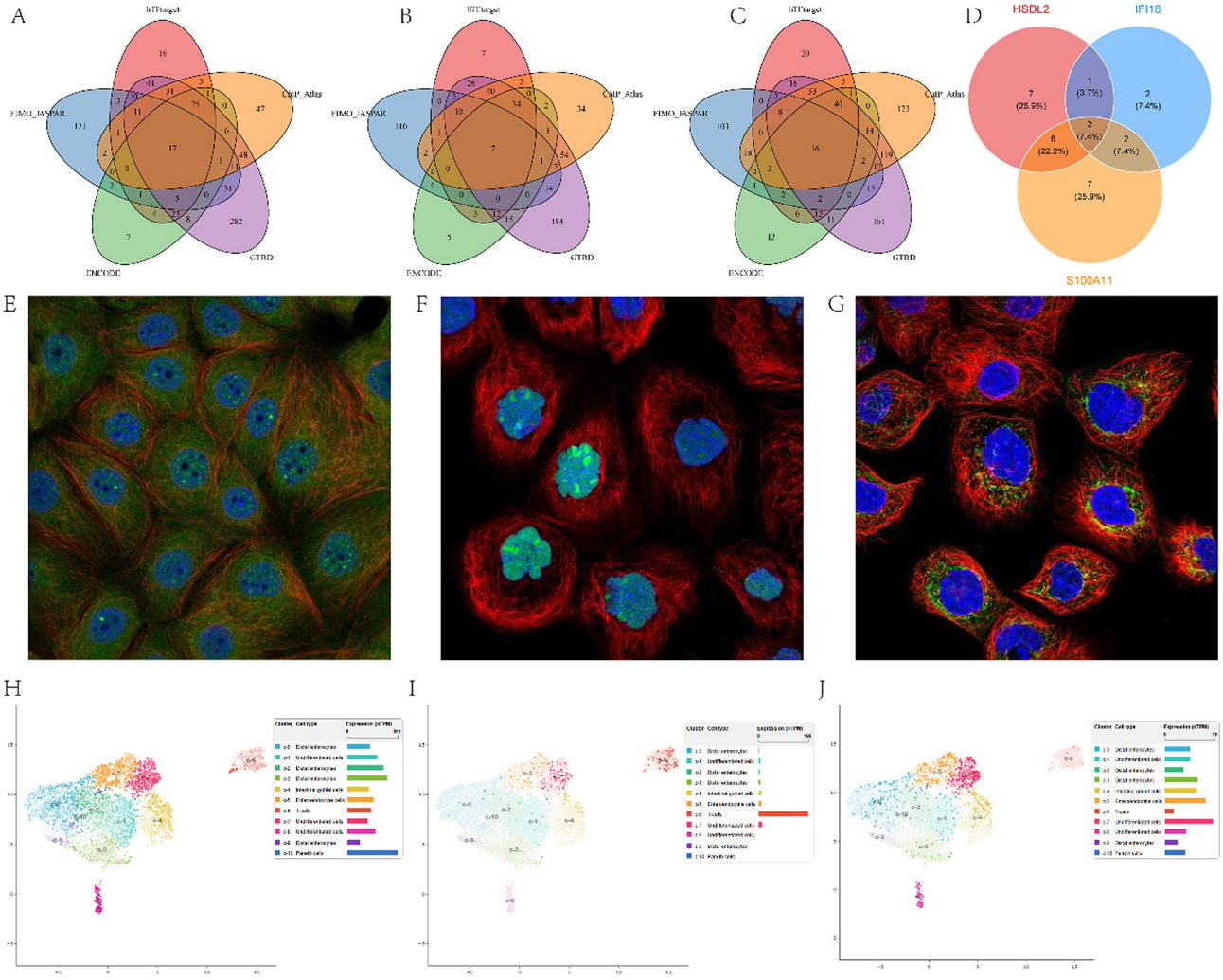

Fig. 7 TFs, immunofluorescence, and single-cell mapping analysis of model genes. (**A-C**) Prediction of TFs for S100A11, IFI16, and HSDL2 based on different databases; (**D**) The TFs of the three model genes are taken as intersections; (**E-G**) The immunofluorescence of S100A11, IFI16, and HSDL2 based on the HPA database; (**H-J**) The single-cell mapping of S100A11, IFI16, and HSDL2 based on the HPA database

challenges in diagnosing and treating UC because there aren't enough reliable diagnostic markers or medication therapeutic targets. It has been demonstrated that lactylation, a post-translational alteration resulting from lactate, controls gene expression and protein function, hence impacting a number of cellular functions [42]. With the growing use of machine learning and integrated bioinformatics analysis in recent years, we are gaining important knowledge about complex diseases by identifying new genes, potential biomarkers, mechanistic insights, and therapeutic targets from massive data sets [43, 44]. In order to better understand the molecular pathogenesis and treatment strategy of UC, we employed a number of thorough bioinformatics techniques to thoroughly analyze the gene sets related to lactylation. This will also help us come up with ideas for early clinical screening and intervention of UC, which will improve the prognosis of this patient population.

Initially, 22 shared genes were identified through a series of analyses integrating UC and lactylation-related genes. Functional enrichment analysis revealed that these genes are primarily enriched in actin cytoskeleton organization and gene and JAK-STAT signaling pathway. These findings suggest that these genes may play a crucial role in lactylation modification in UC, and the identified enrichment pathways could represent potential pathogenic mechanisms underlying lactylation in UC. Machine learning was then used to identify three pattern genes, among which S100A11 and IFI16 were highly expressed, while HSDL2 was low in UC patients. This suggests that S100A11 and IFI16 may be key genes in promoting the development of UC, while HSDL2 may inhibit the progression of UC. Additionally, the three-gene prediction model showed good diagnostic performance, and the AUC value was 0.973. The results of the validation set further confirm our findings.

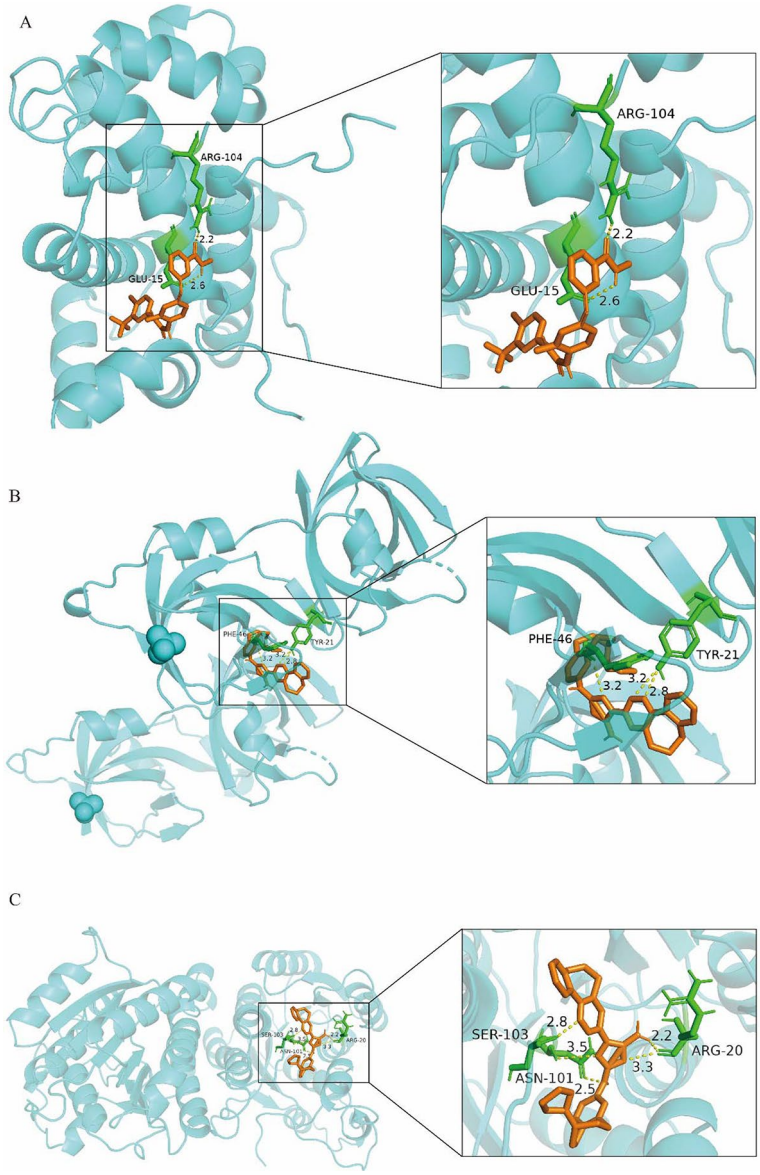

Fig. 8 Molecular docking experiment. (**A**) Molecular docking model of regorafenib and S100A11; (**B**) Molecular docking model of R-428 and IFI16; (**C**) Molecular docking model of R-428 and HSDL2

A member of the S100 protein family, S100A11 affects cell proliferation, differentiation, and cytokine production and is linked to a number of inflammatory and metabolic disorders [45–48]. Previous research has demonstrated tryptophan metabolism, specifically via the kynurenine pathway, is essential for controlling and triggering immunological responses in a number of inflammatory conditions, including UC [49, 50]. Notably, S100A11 is closely linked to dysregulation of the tryptophan metabolism pathway [51]. Therefore, by modulating tryptophan metabolism, S100A11 may influence the progression of UC, potentially serving as a key regulatory factor in disease pathogenesis. IFI6 is an interferon-stimulated protein that plays a pivotal role in immune regulation. Studies have shown that IFI6 activates the JAK/STAT signaling pathway, thereby mediating the immune response to type I interferons and contributing to the progression of UC [52]. The SDR family member HSDL2 catalyzes the reduction and oxidation of vitamins, carbohydrates, steroids, retinoids, and fatty acids [53]. Through the alteration of signaling and metabolic pathways, HSDL2 plays crucial roles in the pathophysiology of a number of disorders, including obesity, cancer, and Alzheimer's disease [54–56]. Through its ability to catalyze the metabolism of different amino acids in UC patients, HSDL2 may prevent the progression of UC.

To further explore the potential functions and molecular mechanisms of S100A11, IFI16, and HSDL2 in UC, we

performed GSVA. According to our research, S100A11 is predominantly increased in pathways linked to the breakdown of branched-chain amino acids (BCAAs), such as isoleucine, leucine, and valine. Meanwhile, IFI16 was found to be enriched in metabolism-related pathways such as pyruvate. Conversely, HSDL2 is predominantly downregulated in processes related to amino acid betaine biosynthesis, histidine metabolism, and the degradation of valine, leucine, and isoleucine. Previous studies have established a strong link between UC and altered amino acid metabolism. UC patients exhibit significantly lower concentrations of leucine and other BCAAs compared to healthy individuals [57–60]. Furthermore, UC has been associated with elevated pyruvate levels and decreased blood histidine concentrations [60]. Notably, reduced histidine levels have been identified as a predictor of acute disease recurrence within six months and one year [59, 61]. These findings suggest that S100A11, IFI16, and HSDL2 may contribute to UC pathogenesis by modulating the metabolism of BCAAs, histidine, and pyruvate.

Immune infiltration analysis further revealed that S100A11 and IFI16 were positively correlated with neutrophils and macrophages M0 and M1 and negatively correlated with macrophage M2, while HSDL2 was the opposite. One significant characteristic of UC, an autoimmune illness, is the recurrence of colonic inflammation, which is caused by the recruitment and local polarization of macrophages to the M1 type, which in turn causes further inflammation [62]. Macrophages are essential for both acute and chronic inflammatory processes. Specifically, overactivated M1-type cells that release enough proinflammatory chemicals to trigger inflammatory storms can result in an imbalanced M1/M2 macrophage ratio. In contrast, M2 macrophages function to resolve inflammation and restore immune homeostasis [63, 64]. Immune homeostasis is regulated by intestinal macrophages, which are crucial for both innate and classical antimicrobial immunity [65]. In UC, a compromised intestinal barrier results in increased antigen penetration into the lamina propria, disrupting immune tolerance and leading to extensive infiltration of proinflammatory macrophages and neutrophils [66, 67]. Therefore, we speculated that S100A11 and IFI16 have an antagonistic relationship with HSDL2 to regulate the disease progression of UC by regulating the function of neutrophils and macrophages.

Finally, drug screening was conducted in order to find drugs that can effectively treat UC. The results showed that regorafenib and R-428 had the lowest free binding energy with the three model genes. Patients with hepatocellular carcinoma [68], gastrointestinal stromal tumors [69], and colorectal cancer [70] are treated with the tyrosine kinase inhibitor regorafenib. Numerous targets are impacted by regorafenib, such as vascular endothelial growth factor receptors 1–3, fibroblast growth factor receptors 1 and 2, and tyrosine kinase 2 with immunoglobulin-like and epidermal growth factor-like domains [71]. Furthermore, R-428 is presently being studied for the treatment of a number of pathologic disorders, such as mesothelioma, acute myeloid leukemia, melanoma, and myelodysplastic syndrome. Tyrosine-protein kinase receptor inhibition is how it works [72]. Although no previous studies had mentioned regorafenib and R-428 as drugs for the treatment of UC, their mechanism of action suggested that regorafenib and R-428 may have a positive impact on the treatment of UC by regulating immune response and inhibiting angiogenesis.

While our study pioneers the exploration of lactylation-related genes in UC through integrated bioinformatics and machine learning approaches, we acknowledge several inherent limitations. Although qPCR validation partially supported our computational findings, the absence of functional experiments represents a critical constraint. Future work will focus on elucidating the molecular mechanisms of model genes in UC through comprehensive in vitro/in vivo studies. Our retrospective design and limited sample size may affect the generalizability of conclusions, necessitating future validation through large-scale prospective cohorts to confirm clinical relevance. Nevertheless, these computational insights provide a valuable foundation for prioritizing targets in subsequent experimental investigations of lactylation-mediated pathophysiology in UC.

Conclusion

Our study identified three lactylation-related genes as promising biomarkers for the diagnosis of UC. Additionally, our findings suggest that regorafenib and R-428 may serve as potential therapeutic agents for UC. These findings underscore the critical role lactylation plays in the development of UC and offer a strong basis for the creation of targeted diagnostic and treatment approaches that use these biomarkers to manage UC.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40360-025-00939-7.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

None.

Author contributions

YY: Conceptualization, Investigation, Writing – original draft. XS: Investigation, Formal analysis, Writing – original draft. BL: Visualization, Writing – original draft. YZ: Data curation, Writing – original draft. TX: Writing – original draft. JL: Data curation, Writing – review & editing. JL: Methodology, Data curation, Writing – review & editing. QZ: Conceptualization, Writing – review & editing.

Funding
None.

Data availability
The datasets generated and/or analysed during the current study are available in the [GEO] repository. https://www.ncbi.nlm.nih.gov/geo/.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of General Surgery, The First Affiliated Hospital of Dalian Medical University, Dalian, Liaoning, China
[2]Institute of Integrative Medicine, Dalian Medical University, Dalian, Liaoning, China
[3]Health Team, The 92914th Military Hospital of PLA, Lingao, Hainan, China

Received: 1 March 2025 / Accepted: 29 April 2025
Published online: 13 May 2025

References

1. Ordás I, Eckmann L, Talamini M, Baumgart DC, Sandborn WJ. Ulcerative colitis. Lancet (London England). 2012;380(9853):1606–19.
2. Lu PD, Zhao YH. Targeting NF-κB pathway for treating ulcerative colitis: comprehensive regulatory characteristics of Chinese medicines. Chin Med. 2020;15:15.
3. de Souza HS, Fiocchi C. Immunopathogenesis of IBD: current state of the Art. Nat Reviews Gastroenterol Hepatol. 2016;13(1):13–27.
4. El-Far YM, Elsherbiny NM, El-Shafey M, Said E. The interplay of the inhibitory effect of Nifuroxazide on NF-κB/STAT3 signaling attenuates acetic acid-induced ulcerative colitis in rats. Environ Toxicol Pharmacol. 2020;79:103433.
5. Ungaro R, Mehandru S, Allen PB, Peyrin-Biroulet L, Colombel JF. Ulcerative colitis. Lancet (London England). 2017;389(10080):1756–70.
6. Gan PR, Wu H, Zhu YL, Shu Y, Wei Y. Glycolysis, a driving force of rheumatoid arthritis. Int Immunopharmacol. 2024;132:111913.
7. Wang YH, Gao P, Wang YQ, Xu LZ, Zeng KW, Tu PF. Small-molecule targeting PKM2 provides a molecular basis of lactylation-dependent fibroblast-like synoviocytes proliferation Inhibition against rheumatoid arthritis. Eur J Pharmacol. 2024;972:176551.
8. Zhang Y, Gao Y, Wang Y, Jiang Y, Xiang Y, Wang X, Wang Z, Ding Y, Chen H, Rui B, et al. RBM25 is required to restrain inflammation via ACLY RNA splicing-dependent metabolism rewiring. Cell Mol Immunol. 2024;21(11):1231–50.
9. Chen L, Li R, Wang Z, Zhang Z, Wang J, Qiao Y, Huang Y, Liu W. Lactate-utilizing bacteria ameliorates DSS-induced colitis in mice. Life Sci. 2022;288:120179.
10. Iraporda C, Romanin DE, Bengoa AA, Errea AJ, Cayet D, Foligné B, Sirard JC, Garrote GL, Abraham AG, Rumbo M. Local treatment with lactate prevents intestinal inflammation in the TNBS-Induced colitis model. Front Immunol. 2016;7:651.
11. Li K, Strauss R, Ouahed J, Chan D, Telesco SE, Shouval DS, Canavan JB, Brodmerkel C, Snapper SB, Friedman JR. Molecular comparison of adult and pediatric ulcerative colitis indicates broad similarity of molecular pathways in disease tissue. J Pediatr Gastroenterol Nutr. 2018;67(1):45–52.
12. Planell N, Lozano JJ, Mora-Buch R, Masamunt MC, Jimeno M, Ordás I, Esteller M, Ricart E, Piqué JM, Panés J, et al. Transcriptional analysis of the intestinal mucosa of patients with ulcerative colitis in remission reveals lasting epithelial cell alterations. Gut. 2013;62(7):967–76.
13. Vancamelbeke M, Vanuytsel T, Farré R, Verstockt S, Ferrante M, Van Assche G, Rutgeerts P, Schuit F, Vermeire S, Arijs I, et al. Genetic and transcriptomic bases of intestinal epithelial barrier dysfunction in inflammatory bowel disease. Inflamm Bowel Dis. 2017;23(10):1718–29.
14. Huang H, Chen K, Zhu Y, Hu Z, Wang Y, Chen J, Li Y, Li D, Wei P. A multi-dimensional approach to unravel the intricacies of lactylation related signature for prognostic and therapeutic insight in colorectal cancer. J Translational Med. 2024;22(1):211.
15. Cheng Z, Huang H, Li M, Liang X, Tan Y, Chen Y. Lactylation-related gene signature effectively predicts prognosis and treatment responsiveness in hepatocellular carcinoma. Pharmaceuticals (Basel). 2023:16(5).
16. Liu J, Zhong L, Zhang Y, Ma J, Xie T, Chen X, Zhang B, Shang D. Identification of novel biomarkers based on lipid metabolism-related molecular subtypes for moderately severe and severe acute pancreatitis. Lipids Health Dis. 2024;23(1):1.
17. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The Sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012;28(6):882–3.
18. Pan X, Liu J, Zhong L, Zhang Y, Liu C, Gao J, Pang M. Identification of lipid metabolism-related biomarkers for diagnosis and molecular classification of atherosclerosis. Lipids Health Dis. 2023;22(1):96.
19. Wang YX, Li XL, Zhang LH, Li HN, Liu XM, Song W, Pang XF. Machine learning algorithms assist early evaluation of enteral nutrition in ICU patients. Front Nutr. 2023;10:1060398.
20. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023;51(D1):D638–46.
21. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The genemania prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res. 2010;3(suppl_2):W214–20.
22. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019;10(1):1523.
23. Bai J, Huang JH, Price CPE, Schauer JM, Suh LA, Harmon R, Conley DB, Welch KC, Kern RC, Shintani-Smith S, et al. Prognostic factors for polyp recurrence in chronic rhinosinusitis with nasal polyps. J Allergy Clin Immunol. 2022;150(2):352–e361357.
24. Blanco JL, Porto-Pazos AB, Pazos A, Fernandez-Lozano C. Prediction of high anti-angiogenic activity peptides in Silico using a generalized linear model and feature selection. Sci Rep. 2018;8(1):15688.
25. Rigatti SJ. Random forest. J Insurance Med (New York NY). 2017;47(1):31–9.
26. Yang L, Pan X, Zhang Y, Zhao D, Wang L, Yuan G, Zhou C, Li T, Li W. Bioinformatics analysis to screen for genes related to myocardial infarction. Front Genet. 2022;13:990888.
27. Qin Q, Gu Z, Li F, Pan Y, Zhang T, Fang Y, Zhang L. A diagnostic model for Alzheimer's disease based on blood levels of Autophagy-Related genes. Front Aging Neurosci. 2022;14:881890.
28. Wang Y, Xu T, Wang W. C9orf72 alleviates DSS–Induced ulcerative colitis via the cGAS-STING pathway. Immun Inflamm Dis. 2025;13(1):e70139.
29. Li J, Chen Y, Yu Q, Li S, Zhang X, Cheng Y, Fu X, Li J, Zhu L. Estrogen receptor B alleviates colitis in intestinal epithelial cells and activates HIF-1a and ATG-9a-mediated autophagy. Exp Cell Res. 2025;447(2):114520.
30. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013;14:7.
31. Wang J. TFTF: an R-Based integrative tool for decoding human transcription factor-target interactions. Biomolecules 2024:14(7).
32. Zhang Q, Liu W, Zhang HM, Xie GY, Miao YR, Xia M, Guo AY. hTFtarget: A comprehensive database for regulations of human transcription factors and their targets. Genomics Proteom Bioinf. 2020;18(2):120–8.
33. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57–74.
34. Rauluseviciute I, Riudavets-Puig R, Blanc-Mathieu R, Castro-Mondragon JA, Ferenc K, Kumar V, Lemma RB, Lucas J, Cheneby J, Baranasic D, et al. JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles. Nucleic Acids Res. 2024;52(D1):D174–82.
35. Kolmykov S, Yevshin I, Kulyashov M, Sharipov R, Kondrakhin Y, Makeev VJ, Kulakovskiy IV, Kel A, Kolpakov F. GTRD: an integrated view of transcription regulation. Nucleic Acids Res. 2021;49(D1):D104–11.
36. Zou Z, Ohta T, Miura F, Oki S. ChIP-Atlas 2021 update: a data-mining suite for exploring epigenomic landscapes by fully integrating ChIP-seq, ATAC-seq and Bisulfite-seq data. Nucleic Acids Res. 2022;50(W1):W175–82.
37. Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK, et al. A next generation

connectivity map: L1000 platform and the first 1,000,000 profiles. Cell. 2017;171(6):1437–e14521417.
38. Kobayashi T, Siegmund B, Le Berre C, Wei SC, Ferrante M, Shen B, Bernstein CN, Danese S, Peyrin-Biroulet L, Hibi T. Ulcerative colitis. Nat Reviews Disease Primers. 2020;6(1):74.
39. Park JH, Peyrin-Biroulet L, Eisenhut M, Shin JI. IBD Immunopathogenesis: A comprehensive review of inflammatory molecules. Autoimmun Rev. 2017;16(4):416–26.
40. Rosenzwajg M, Lorenzon R, Cacoub P, Pham HP, Pitoiset F, El Soufi K, C RI, Bernard C, Aractingi S, Banneville B, et al. Immunological and clinical effects of low-dose interleukin-2 across 11 autoimmune diseases in a single, open clinical trial. Ann Rheum Dis. 2019;78(2):209–17.
41. Park SC, Jeen YT. Anti-integrin therapy for inflammatory bowel disease. World J Gastroenterol. 2018;24(17):1868–80.
42. Chen L, Huang L, Gu Y, Cang W, Sun P, Xiang Y. Lactate-Lactylation Hands between Metabolic Reprogramming and Immunosuppression. Int J Mol Sci. 2022:23(19).
43. Yin X, Xing W, Yi N, Zhou Y, Chen Y, Jiang Z, Ma C, Xia C. Comprehensive analysis of lactylation-related gene sets and mitochondrial functions in gastric adenocarcinoma: implications for prognosis and therapeutic strategies. Front Immunol. 2024;15:1451725.
44. Fu W, Wang T, Lu Y, Shi T, Yang Q. The role of lactylation in plasma cells and its impact on rheumatoid arthritis pathogenesis: insights from single-cell RNA sequencing and machine learning. Front Immunol. 2024;15:1453587.
45. Sobolewski C, Abegg D, Berthou F, Dolicka D, Calo N, Sempoux C, Fournier M, Maeder C, Ay AS, Clavien PA, et al. S100A11/ANXA2 belongs to a tumour suppressor/oncogene network deregulated early with steatosis and involved in inflammation and hepatocellular carcinoma development. Gut. 2020;69(10):1841–54.
46. Zeng X, Guo H, Liu Z, Qin Z, Cong Y, Ren N, Zhang Y, Zhang N. S100A11 activates the Pentose phosphate pathway to induce malignant biological behaviour of pancreatic ductal adenocarcinoma. Cell Death Dis. 2022;13(6):568.
47. Zhang L, Zhu T, Miao H, Liang B. The calcium binding protein S100A11 and its roles in diseases. Front Cell Dev Biology. 2021;9:693262.
48. Andrés Cerezo L, Šumová B, Prajzlerová K, Veigl D, Damgaard D, Nielsen CH, Pavelka K, Vencovský J, Šenolt L. Calgizzarin (S100A11): a novel inflammatory mediator associated with disease activity of rheumatoid arthritis. Arthritis Res Therapy. 2017;19(1):79.
49. Nikolaus S, Schulte B, Al-Massad N, Thieme F, Schulte DM, Bethge J, Rehman A, Tran F, Aden K, Häsler R, et al. Increased Tryptophan metabolism is associated with activity of inflammatory bowel diseases. Gastroenterology. 2017;153(6):1504–1516. e1502.
50. Shi Y, Luo S, Zhai J, Chen Y. A novel causative role of imbalanced kynurenine pathway in ulcerative colitis: upregulation of KMO and KYNU promotes intestinal inflammation. Biochim Et Biophys Acta Mol Basis Disease. 2024;1870(2):166929.
51. Chen G, Qi H, Jiang L, Sun S, Zhang J, Yu J, Liu F, Zhang Y, Du S. Integrating single-cell RNA-Seq and machine learning to dissect Tryptophan metabolism in ulcerative colitis. J Translational Med. 2024;22(1):1121.
52. Sajid M, Ullah H, Yan K, He M, Feng J, Shereen MA, Hao R, Li Q, Guo D, Chen Y, et al. The functional and antiviral activity of interferon Alpha-Inducible IFI6 against hepatitis B virus replication and gene expression. Front Immunol. 2021;12:634937.
53. Oppermann UC, Salim S, Tjernberg LO, Terenius L, Jörnvall H. Binding of amyloid beta-peptide to mitochondrial hydroxyacyl-CoA dehydrogenase (ERAB): regulation of an SDR enzyme activity with implications for apoptosis in Alzheimer's disease. FEBS Lett. 1999;451(3):238–42.
54. Di Marco F, Vecchio GM, Di Stefano B, Pesce A, Puleo S. Metastatic hepatocellular carcinoma to the occipito-cervical junction: a unique case and literature review. ANZ J Surg. 2019;89(9):E414–6.
55. Juan C, Hua Q, Ruping Z, Tingting W. miRNA-489 as a biomarker in diagnosis and treatment of cervical cancer. Bratisl Lek Listy. 2018;119(5):278–83.
56. Xu Y, Yu J, Liu T, Meng F, Kong D, Lou G. Loss of FBXW7 is related to the susceptibility and poor prognosis of cervical squamous carcinoma. Biomarkers: Biochem Indic Exposure Response Susceptibility Chemicals. 2016;21(4):379–85.
57. Bjerrum JT, Steenholdt C, Ainsworth M, Nielsen OH, Reed MA, Atkins K, Günther UL, Hao F, Wang Y. Metabonomics uncovers a reversible proatherogenic lipid profile during Infliximab therapy of inflammatory bowel disease. BMC Med. 2017;15(1):184.
58. Diab J, Hansen T, Goll R, Stenlund H, Jensen E, Moritz T, Florholmen J, Forsdahl G. Mucosal metabolomic profiling and pathway analysis reveal the metabolic signature of ulcerative colitis. Metabolites. 2019:9(12).
59. Probert F, Walsh A, Jagielowicz M, Yeo T, Claridge TDW, Simmons A, Travis S, Anthony DC. Plasma nuclear magnetic resonance metabolomics discriminates between high and low endoscopic activity and predicts progression in a prospective cohort of patients with ulcerative colitis. J Crohn's Colitis. 2018;12(11):1326–37.
60. Tews HC, Schmelter F, Kandulski A, Büchler C, Schmid S, Schlosser S, Elger T, Loibl J, Sommersberger S, Fererberger T, et al. Unique metabolomic and lipidomic profile in serum from patients with Crohn's disease and ulcerative colitis compared with healthy control individuals. Inflamm Bowel Dis. 2024;30(12):2405–17.
61. Hisamatsu T, Ono N, Imaizumi A, Mori M, Suzuki H, Uo M, Hashimoto M, Naganuma M, Matsuoka K, Mizuno S, et al. Decreased plasma histidine level predicts risk of relapse in patients with ulcerative colitis in remission. PLoS ONE. 2015;10(10):e0140716.
62. Neurath MF. Cytokines in inflammatory bowel disease. Nat Rev Immunol. 2014;14(5):329–42.
63. Shapouri-Moghaddam A, Mohammadian S, Vazini H, Taghadosi M, Esmaeili SA, Mardani F, Seifi B, Mohammadi A, Afshari JT, Sahebkar A. Macrophage plasticity, polarization, and function in health and disease. J Cell Physiol. 2018;233(9):6425–40.
64. Atri C, Guerfali FZ, Laouini D. Role of human macrophage polarization in inflammation during infectious diseases. Int J Mol Sci. 2018:19(6).
65. Smith PD, Smythies LE, Shen R, Greenwell-Wild T, Gliozzi M, Wahl SM. Intestinal macrophages and response to microbial encroachment. Mucosal Immunol. 2011;4(1):31–42.
66. Kaser A, Adolph TE, Blumberg RS. The unfolded protein response and Gastrointestinal disease. Semin Immunopathol. 2013;35(3):307–19.
67. Martini E, Krug SM, Siegmund B, Neurath MF, Becker C. Mend your fences: the epithelial barrier and its relationship with mucosal immunity in inflammatory bowel disease. Cell Mol Gastroenterol Hepatol. 2017;4(1):33–46.
68. Arai H, Battaglin F, Wang J, Lo JH, Soni S, Zhang W, Lenz HJ. Molecular insight of regorafenib treatment for colorectal cancer. Cancer Treat Rev. 2019;81:101912.
69. Blay JY, Kang YK, Nishida T, von Mehren M. Gastrointestinal stromal tumours. Nat Reviews Disease Primers. 2021;7(1):22.
70. Llovet JM, Montal R, Sia D, Finn RS. Molecular therapies and precision medicine for hepatocellular carcinoma. Nat Reviews Clin Oncol. 2018;15(10):599–616.
71. Grothey A, Blay JY, Pavlakis N, Yoshino T, Bruix J. Evolving role of regorafenib for the treatment of advanced cancers. Cancer Treat Rev. 2020;86:101993.
72. Urra G, Valdés-Muñoz E, Suardiaz R, Hernández-Rodríguez EW, Palma JM, Ríos-Rozas SE, Flores-Morales CA, Alegría-Arcos M, Yáñez O, Morales-Quintana L et al. From proteome to potential drugs: integration of subtractive proteomics and ensemble docking for drug repurposing against Pseudomonas aeruginosa RND superfamily proteins. Int J Mol Sci. 2024:25(15).

Publisher's note