


# Identification of Anoikis-Related Genes in Driving Immune-Inflammatory Responses in Ulcerative Colitis Based on Bioinformatics Analysis and Machine Learning

Wenxiu Diao<sup>1,2,\*</sup>, Xu Huang<sup>1,\*</sup>, Wensha Huang<sup>1,\*</sup>, Jing Jiang<sup>2</sup>, Wentao Li<sup>2</sup>, He Liu<sup>2</sup>, Bo Yan<sup>1</sup> , Lei Shen<sup>1</sup>

<sup>1</sup>Department of Gastroenterology, Renmin Hospital of Wuhan University, Wuhan, People's Republic of China; <sup>2</sup>Department of Cardiology, No. 971 hospital of The People's Liberation Army Navy, Qingdao, People's Republic of China

\*These authors contributed equally to this work

Correspondence: Lei Shen; Bo Yan, Department of Gastroenterology, Renmin Hospital of Wuhan University, Wuhan, People's Republic of China, Email leishenwuhan@126.com; byanwuhan@126.com

**Background:** Ulcerative colitis (UC) is a challenging chronic intestinal inflammation. Anoikis, a type of programmed cell death triggered by detachment from the extracellular matrix, is crucial in various physiological and pathological contexts. This study aims to explore the biological and clinical implications of anoikis-related genes (ARGs) in UC.

**Methods:** Gene expression microarrays from normal and UC mucosal tissues focused on ARGs. Differentially expressed genes (DEGs) related to anoikis in UC were identified. Weighted gene co-expression network analysis (WGCNA) screened UC-related module genes. GO, KEGG, GSEA, and GSVA analyses were used to uncover mechanisms. Machine learning identified hub ARG-DEGs highly correlated with UC, and diagnostic nomograms assessed their diagnostic potential. The CIBERSORT algorithm analyzed changes in the UC immune microenvironment related to hub UC-ARGs. Potential drugs, miRNAs, and transcription factors (TFs) interacting with these hub UC-ARGs were investigated, and animal experiments verified their expression.

**Results:** 49 ARG-DEGs were identified, mainly linked to the PI3K-AKT signaling pathway, inflammatory signal regulation, and extracellular matrix (ECM)-receptor interactions. Notably, CDH3 and SERPINA1 showed significant diagnostic potential for UC, confirmed by the Wilcoxon rank-sum test, independent validation sets, Western blot, and immunohistochemical staining. Significant variations in immune cell infiltration and activation within UC samples correlated with hub UC-ARGs were observed using the CIBERSORT algorithm.

**Conclusion:** Anoikis may drive UC progression by initiating an immune inflammatory response. CDH3 and SERPINA1 are promising biomarkers and therapeutic targets for UC.

**Keywords:** ulcerative colitis, anoikis, biomarkers, machine learning

## Introduction

Ulcerative colitis (UC) is primarily characterized by an increase in bowel movements and the presence of blood in the stool. Its course is often prolonged and recurrent, significantly impacting patients' quality of life. In 2023, the prevalence of UC was estimated to be 5 million cases around the world, and the incidence is increasing worldwide.<sup>1,2</sup> Persistent autoimmune abnormalities and intestinal mucosal damage are foundational to the development of this disease, yet the complete etiological mechanisms underlying UC remain incompletely elucidated.<sup>2</sup> Consequently, ongoing exploration of the molecular mechanisms involved in the onset and progression of UC is crucial for discovering novel diagnostic biomarkers and effective therapeutic approaches.

Anoikis represents a unique mode of programmed cell death, distinct from apoptosis, autophagy, and pyroptosis.<sup>3,4</sup> This process is initiated by the disruption of cell adhesion to the extracellular matrix (ECM),<sup>5</sup> leading to inappropriate cellular death that serves to preserve tissue homeostasis. Activation of anoikis encompasses two mechanisms: the mitochondrial apoptotic pathway and the Fas/FasL pathway.<sup>4</sup> When cells are stimulated by factors such as inflammation and oxidative stress, these stimuli lead to a reduction in mitochondrial membrane potential and the subsequent release of oxygen radicals. This, in turn, triggers the activation of intramitochondrial kinases, ultimately culminating in the activation of downstream effector molecules associated with anoikis.<sup>6</sup> Alternatively, the binding of receptors and ligands can also mediate the initiation of anoikis, resulting in the activation of downstream effector molecules.<sup>7</sup>

Recently, anoikis has been shown to play an important role in pathological processes associated with diseases such as cancer, diabetes and cardiovascular disease.<sup>3,8,9</sup> Recent research has highlighted the role of specific molecular pathways, including the AKT, MAPK and NF- $\kappa$ B signaling cascades, in initiating the process of anoikis.<sup>9–11</sup> Notably, these same pathways have also been associated with the progression of UC, highlighting a potential role for anoikis in the disease's progression.<sup>2</sup>

The involvement of genes associated with anoikis in UC remains understudied, which has hindered a comprehensive understanding of their overall significance. This study endeavors to bridge this research gap by investigating biomarkers and molecular mechanisms pertaining to anoikis-related genes (ARGs) in UC patients. To achieve this, we will utilize differential gene expression (DEG) analysis, weighted gene co-expression network analysis (WGCNA), and machine learning algorithms. Furthermore, we aim to establish a nomogram model based on hub ARG-DEGs to distinguish UC cases while simultaneously exploring variations within the immune microenvironment. Animal models will be employed to validate the expression patterns of these hub ARG-DEGs.

## Materials and Methods

### Data Acquisition and Processing

From the GEO database, we retrieved gene expression profile microarrays corresponding to both normal individuals and patients with UC. We employed the R package limma for normalization of each dataset. To mitigate batch effects, we integrated two datasets (GSE59071, GSE92415) using the R package sva.<sup>12</sup> For the ensuing analyses, the consolidated dataset was designated as the training set, while the GSE87466 and GSE75214 datasets served as the validation set. A total of 509 ARGs were retrieved from the Harmonizome web portal and the GeneCards database. Figure 1 illustrates the analysis process of this study.

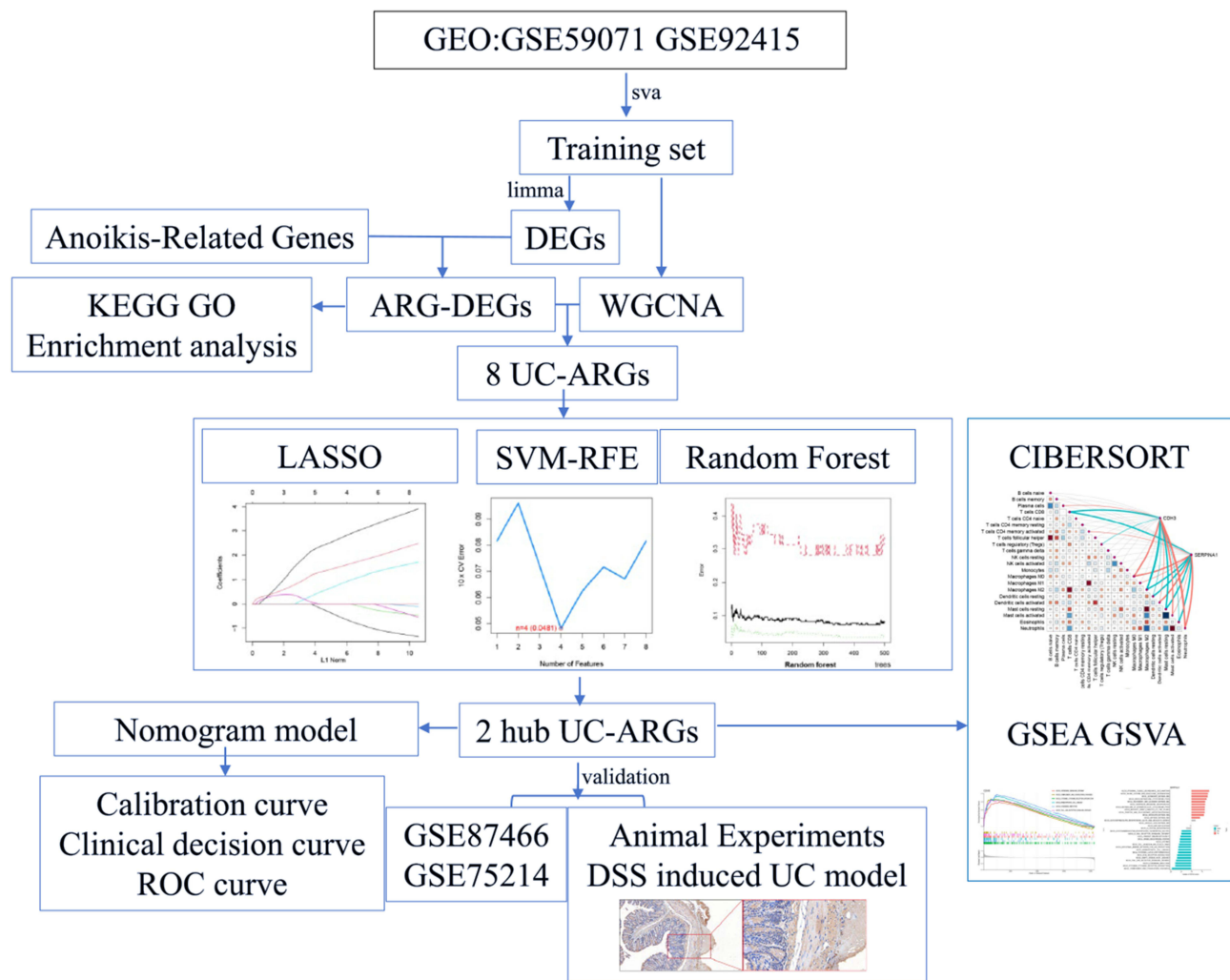
### Differential Gene Expression and Gene Enrichment Analyses

The R package “limma” was employed to identify differentially expressed genes (DEGs) between UC and healthy controls using the Wilcoxon's rank-sum test. Visualize the results of DEG analysis by using the “pheatmap” R package. By intersecting the DEGs with ARGs, the anoikis-related DEGs (ARG-DEGs) were discerned using a Venn diagram for clarity.

Subsequently, to explore the biological mechanisms of these ARG-DEGs, GO and KEGG enrichment analyses were conducted using the R packages “clusterProfiler” and “org.Hs.eg.db”,<sup>13</sup> enabling a comprehensive exploration of the functional implications of the identified ARG-DEGs.

### Weighted Gene Co-Expression Network Analysis (WGCNA)

Utilizing the Sanger Box WGCNA Platform,<sup>14</sup> we constructed a WGCNA based on the training dataset. To ensure data integrity, we employed the goodSamplesGenes function to identify and address any missing values. Subsequently, the function of “pickSoftThreshold” was utilized to rigorously filter and validate the optimal soft threshold. The adjacency matrix was generated using the Pearson correlation coefficient. Within these modules, genes were selected that demonstrated the strongest association with UC traits, using a stringent criterion of MM > 0.7 and GS > 0.4 as the threshold for identifying hub module genes.



**Figure 1** The analysis flow chart of this study.

## Identification of Key Features Using Machine Learning (ML) Models

To improve the accuracy of hub ARG-DEGs diagnostic UC, three machine learning algorithms were used to identify hub UC-ARGs. LASSO regression, a widely used data mining technique,<sup>15</sup> was implemented using the R package “glmnet”. By integrating the ARG-DEGs into the diagnostic model and conducting ten cross-validations, we ultimately identified significant ARG-DEGs using LASSO. SVM-RFE, a prevalent machine learning approach based on embedded methods,<sup>16</sup> was applied using the R package “e1071”. Through this method, we pinpointed the most significant variables and eliminated feature vectors generated by the SVM. The Random Forest algorithm was configured with a decision tree count of 500.<sup>17</sup> The intersection of the UC-ARGs filtered by these three machine learning methods was defined as the hub UC-ARGs. To validate our findings, the GSE87466 dataset was utilized to perform receiver operating characteristic (ROC) curve analysis.

## Construction of a Diagnostic Model

To further improve its clinical applicability, we utilized the R package rms to create a nomogram that incorporates the hub UC-ARGs.<sup>18</sup> The reliability and precision of our nomogram model were rigorously assessed through calibration curve analysis, decision curve analysis (DCA), and evaluation of the ROC curve.

## Gene Set Enrichment Analysis (GSEA) and Gene Set Variation Analysis (GSVA)

To identify the most significant KEGG pathways that differentiate the UC group from the control group, we performed single-gene GSEA on the hub UC-ARGs.<sup>19</sup> A gene set was considered significantly enriched if the p-value was less than 0.05. Subsequently, we conducted single-gene GSVA using the “GSVA” R package to evaluate the enrichment of these hub UC-ARGs in specific functions and pathways.<sup>20</sup> For this analysis, we utilized the gene sets “c2.cp.kegg.v7.4.symbols” and “c5.go.bp.v7.5.1.symbols” sourced from the Molecular Signatures Database (MSigDB). Functional pathways exhibiting a GSVA score  $|t|$  exceeding 0.5 were considered as significantly enriched.

## Assessment of Immune Cell Infiltration in UC

To analyze the infiltrating immune cells, we calculated their gene expression profiles in UC data using the CIBERSORT algorithm along with a reference set (LM22) comprising 22 distinct immune cells, with 1000 permutations conducted to estimate the relative abundance of these immune cells.<sup>21</sup> Subsequently, we assessed and visually represented the discrepancies and correlations of infiltrating immune cells between UC and controls using the R packages “corrplot” and “boxplot”. To explore the connection between the hub UC-ARGs and the immune cells, we performed Spearman correlation analysis using R software. We subsequently visualized the results using the R packages “ggplot2” and “linkET”.

## Drug Prediction

To explore the hub UC-ARGs as potential therapeutic targets for existing medications or novel druggable gene candidates, we leveraged the DGIdb (<https://www.dgldb.org/>) to investigate protein (or gene)-drug interactions.<sup>22</sup> The DGIdb platform offers a comprehensive suite of advanced functionalities, enabling users to browse, search, and refine drug-gene interaction data sourced from over 30 reputable repositories. Through searching this database, potential drugs targeting hub UC-ARGs were acquired.

## miRNA-TF-mRNA Regulatory Network of Hub UC-ARGs

To identify the microRNAs (miRNAs) targeting the hub UC-ARGs, we utilized the miRTarBase<sup>23</sup> and Targetscan databases.<sup>24</sup> Additionally, we employed a p-value threshold of 0.05 as a screening criterion to predict the transcription factors (TFs) associated with these hub UC-ARGs, utilizing the Enrichr database (<http://amp.pharm.mssm.edu/Enrichr/>) for these analyses.<sup>25</sup> Subsequently, we constructed regulatory networks depicting the interactions among miRNAs, TFs, and mRNAs, and visualized these networks using Cytoscape version 3.9.2.

## Animal Experiments

This study adhered to the ARRIVE guidelines. Animal experimentation and care were strictly conducted in accordance with the regulations outlined by the National Institutes of Health (NIH) concerning the ethical treatment and use of animals in research. Approval was granted by the Ethics Committee of Renmin Hospital of Wuhan University (Approved number: WDRY20230402A).

Male C57BL/6J mice, aged 8 weeks, were procured from Hunan SJA Laboratory Animal Co., Ltd. All mice were housed in a laboratory at  $21^{\circ}\text{C} \pm 2^{\circ}\text{C}$  and  $50\% \pm 5\%$  humidity. Twelve mice were randomly divided into Control group and DSS group. Mice in the DSS group were given 3% DSS solution (from MP Biomedicals, USA) in their drinking water for seven days, while control mice were given sterile water for seven days.

## Hematoxylin–Eosin (HE) Staining

Colon tissue sections from fixed mice were cut to a thickness of 5  $\mu\text{m}$  and subsequently dewaxed and dehydrated. The sections then underwent staining with hematoxylin, followed by differentiation in a hydrochloric acid-alcohol solution. Afterward, they were returned to a blue color in a weakly alkaline aqueous solution and further processed with alcoholic eosin. Upon completion of these staining procedures, the sections were dehydrated, rendered transparent, and sealed with



neutral gum. Lastly, microscopic examination was conducted to observe the pathological changes within the colon tissues.

## Immunohistochemistry (IHC) Assay

Colon sections from mice were initially deparaffinized and then immersed in an antigen repair solution. They were subjected to microwave repair for a duration of 20 minutes, followed by washing with PBS. To block endogenous peroxidase activity, freshly prepared 3% hydrogen peroxide was applied. Subsequently, drops of bovine serum albumin (BSA) were added, and the sections were incubated at room temperature for 20 minutes to achieve closure. The primary antibody was then applied, and the sections were washed thoroughly with PBS. Color development was achieved using DBA, followed by restaining with hematoxylin. After dehydrating the sections, they were sealed. Finally, microscopic images of the colon sections were captured (Olympus Corporation, Japan).

## Western Blotting Assay

As previously described,<sup>26</sup> Western blot (WB) analysis was performed on colon tissue samples obtained from each group. Protein extraction was achieved using RIPA lysate containing both protease and phosphatase inhibitors. The protein samples then underwent SDS-PAGE, membrane transfer, blocking, and washing steps. This was followed by an overnight incubation with the primary antibody and subsequent washing. The next day, the secondary antibody was added for incubation, washed off, and developed. Lastly, the gray values of the resultant bands were quantified using Image J software. The antibodies utilized were CDH3 (Rabbit, 1:2000, Proteintech, China), SERPINA1 (Rabbit, 1:2000, Proteintech, China), and GAPDH (Rabbit, 1:5000, Proteintech, China).

## Statistical Analysis

Statistical analyses were carried out utilizing R software (version 4.2.2). For group comparisons, a *t*-test was employed when comparing two groups, while one-way ANOVA was used for comparisons involving three or more groups. To assess correlations, Spearman's rank correlation analysis was conducted with the “ggpubr” and “stats” R packages. The threshold for statistical significance was set at a *p*-value of less than 0.05.

## Results

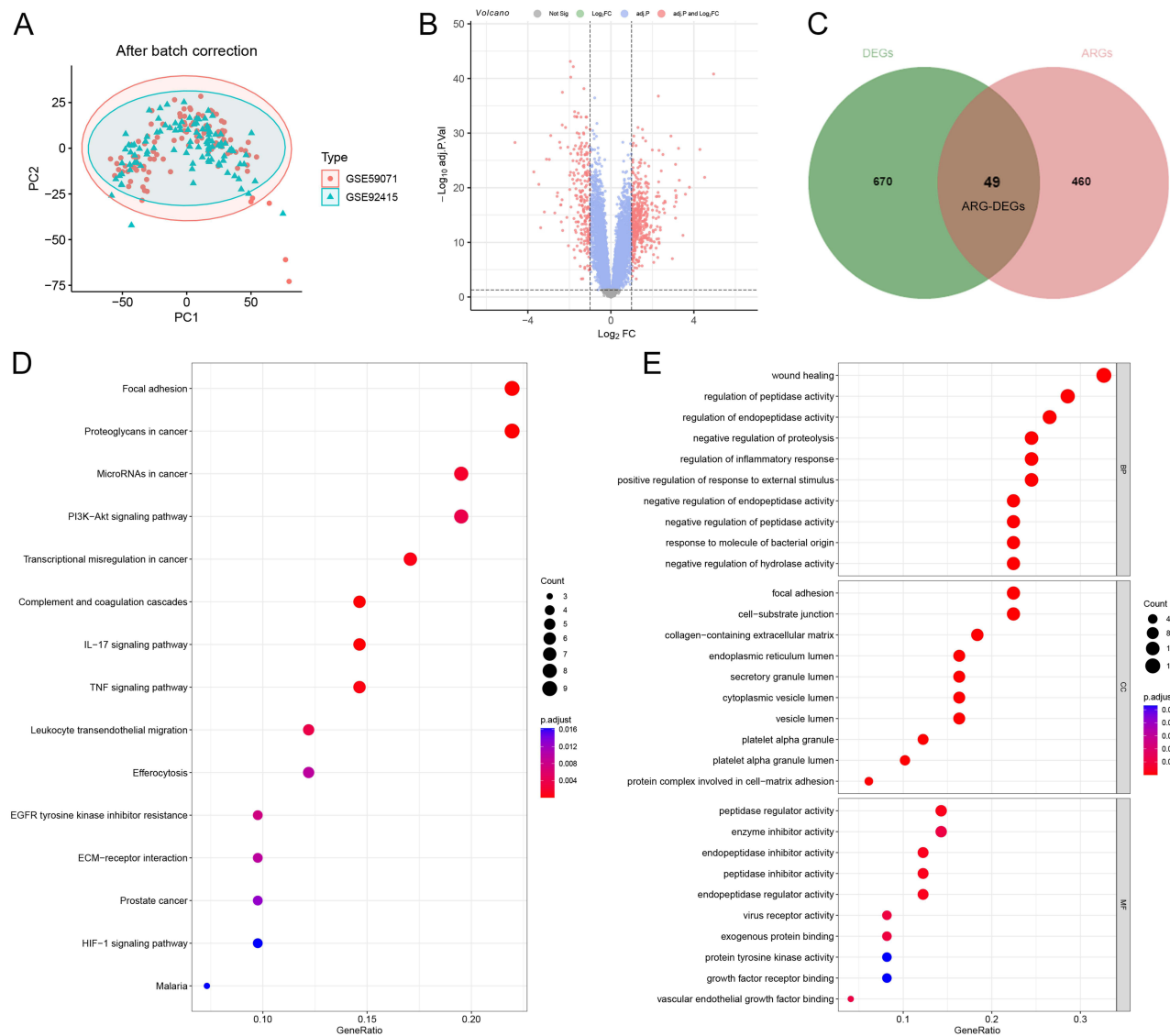
### Data Preprocessing and Identification of Anoikis-Related DEGs

In this study, we combined two datasets, GSE59071 and GSE92415, which together comprised 80 UC and 129 healthy controls. After removing batch effects, we presented the PCA results for each sample in [Figure 2A](#). Using differential expression analysis, we identified 719 DEGs ([Figure 2B](#),  $|\log_2FC| > 1$ ,  $FDR < 0.05$ ). By intersecting this set of DEGs with a predefined list of 509 ARGs, we pinpointed 49 DEGs that were specifically linked to anoikis, termed anoikis-related DEGs (ARG-DEGs) ([Figure 2C](#)).

Further functional analyses indicated that these ARG-DEGs were significantly enriched in multiple biological pathways, including focal adhesion, the PI3K-AKT signaling pathway, inflammatory signal regulation, ECM-receptor interaction, wound healing, and peptidase activity regulation, among others ([Figure 2D](#) and [E](#)). This enrichment suggests the potential biological significance and functional roles of these genes in the context of anoikis.

### Identification of Key Module Genes Associated With UC

To identify hub modules associated with UC, we employed WGCNA. The optimal soft threshold was determined to be 7, ensuring a robust network construction. As the mean connectivity approached zero, the scale-free fit index along the ordinate axis augmented, and the signed  $R^2$  value converged towards the threshold of 0.85, indicative of a suitable model fit ([Figure 3A](#)). Subsequently, utilizing the dynamic tree-cut algorithm in conjunction with similar merging strategies, we identified 13 distinct modules ([Figure 3B](#)). The midnight blue module, characterized by a strong correlation ( $|\text{cor}| > 0.69$ ,  $P < 0.05$ ), emerged as significantly associated with UC ([Figure 3C](#)). To further refine our analysis, we focused on the



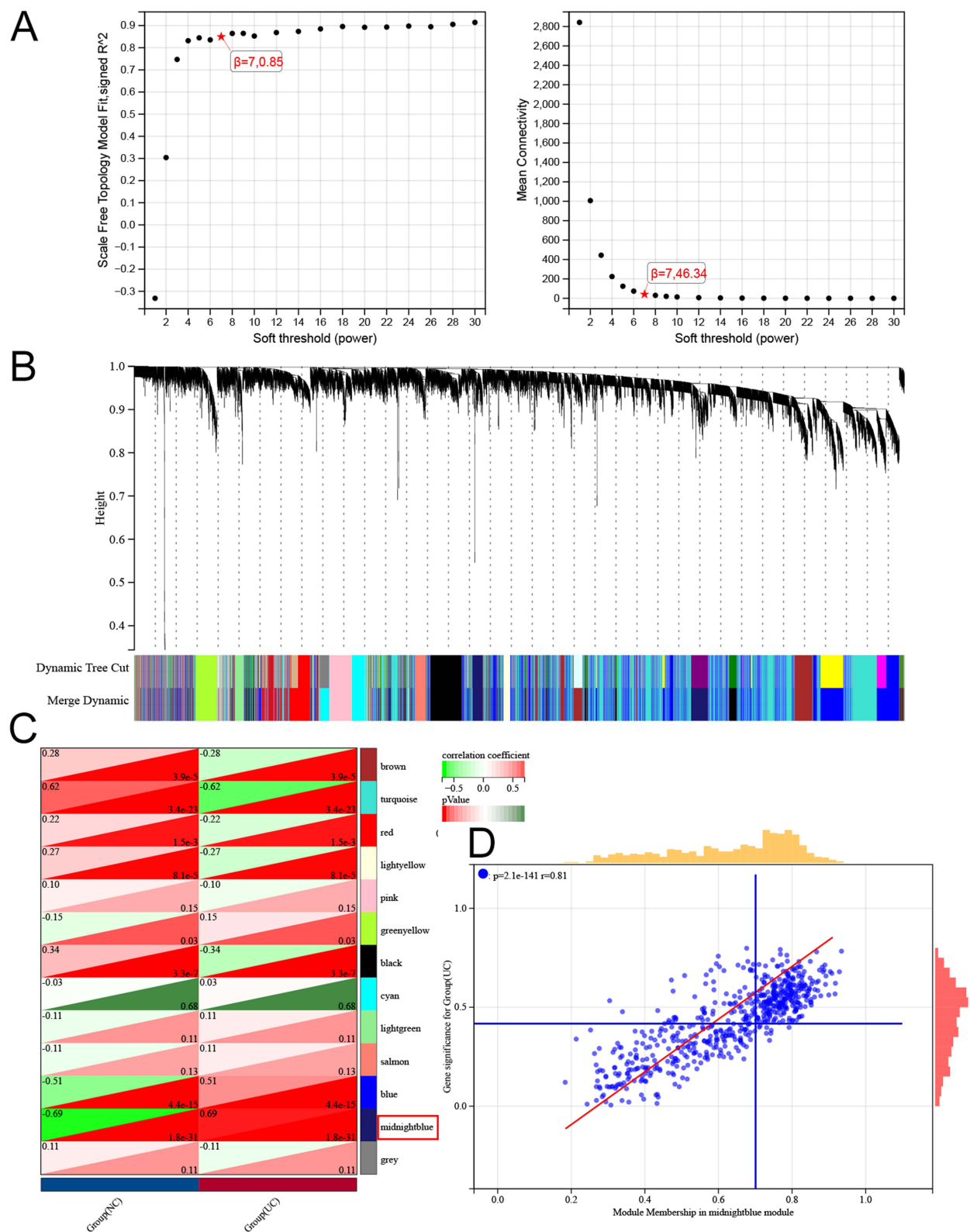
**Figure 2** Identification of ARG-DEGs. **(A)** PCA of two datasets after batch effect correction. **(B)** Volcano map showing differential gene expression patterns between UC and HC. **(C)** Venn diagram illustrating the overlapping genes between differentially expressed genes (DEGs) and anoikis-related genes (ARGs). **(D)** bubble plots displaying the top 15 significant KEGG enrichment pathway. **(E)** GO enrichment analysis of ARG-DEGs.

269 hub genes within this module, which were deemed relevant to UC based on stringent criteria of GS>0.4 and MM>0.7 (Figure 3D).

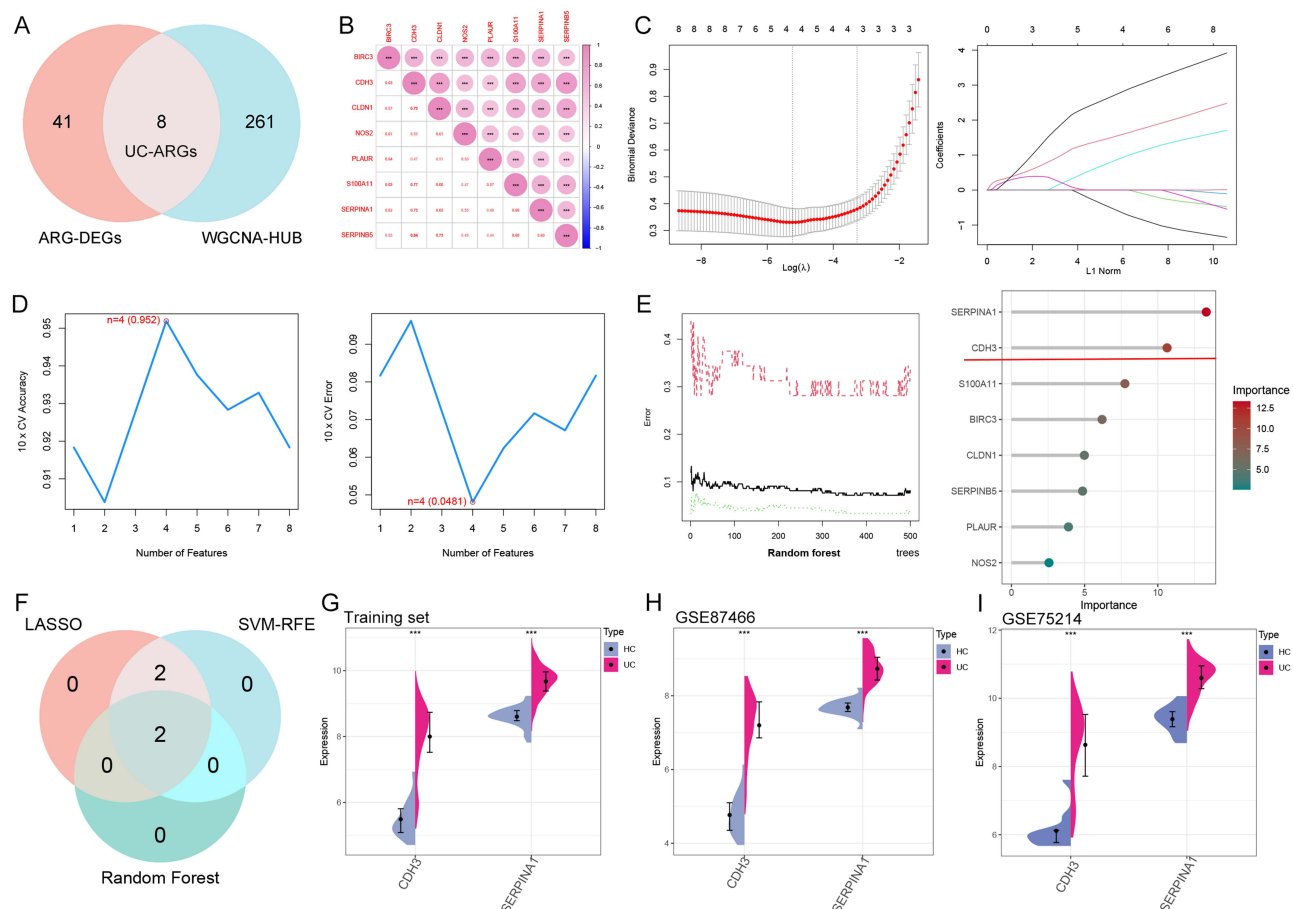
### Identification and Validation of Hub UC-ARGs

By intersecting the set of ARG-DEGs with the hub module genes, we identified eight potential hub UC-ARGs, including BIRC3, CDH3, CLDN1, NOS2, PLAUR, S100A11, SERPINA1, and SERPINA5 (Figure 4A). Subsequently, we conducted a gene-gene correlation analysis for these genes, revealing significant intercorrelations among them. Notably, the correlation coefficient (r) between CDH3 and SERPINA1 stood out as r=0.73, and p<0.001, indicating a strong relationship between these two genes (Figure 4B).

To enhance the diagnostic accuracy of ARG-DEGs in UC, we conducted three distinct machine learning algorithms for the purpose of further identifying hub UC-ARGs. Specifically, the LASSO analysis identified four genes: BIRC3, CDH3, PLAUR, and SERPINA1 (Figure 4C). Similarly, the SVM-RFE method also identified CDH3, SERPINA1, BIRC3, and PLAUR as significant (Figure 4D). Furthermore, using a gene importance score threshold of greater than 10



**Figure 3** WGCNA. (A) Determine the best soft threshold. (B) Gene dendrogram obtained by hierarchical clustering. (C) Heat map of module-trait relationships. (D) Scatterplot of correlations between GS and MM in midnight blue module.



**Figure 4** Identification of hub UC-ARGs by machine learning. **(A)** Venn diagram illustrating the overlapping genes (UC-ARGs) between ARG-DEGs and WGCNA hub module genes. **(B)** The heatmap for correlation analysis of 8 UC-ARGs. **(C)** LASSO coefficient analysis. Vertical dashed lines are depicted at the optimal lambda and ten cross-validations for the selection of adjustment parameters in the LASSO model. Each curve corresponds to one gene. **(D)** Maximum accuracy and minimum error plots of the SVM-RFE algorithm for screening hub UC-ARGs. **(E)** Analysis of the correlation between the quantity of 500 random forest trees and the error rate, and evaluation of the relative importance ranking of UC-ARGs. **(F)** The Venn diagram illustrating the overlapping genes (2 hub UC-ARGs) among LASSO, SVM-RFE, and Random Forest. Violin plot illustrating differential expression of hub UC-ARGs between healthy controls and UC patients in the training set **(G)** and validation set GSE87466 **(H)**, GSE75214 **(I)**. **(H)** ROC curve analysis of hub UC-ARGs in the training set. **(I)** ROC curve analysis of hub UC-ARGs in the GSE87466 validation set. \*\*\* $p < 0.001$ .

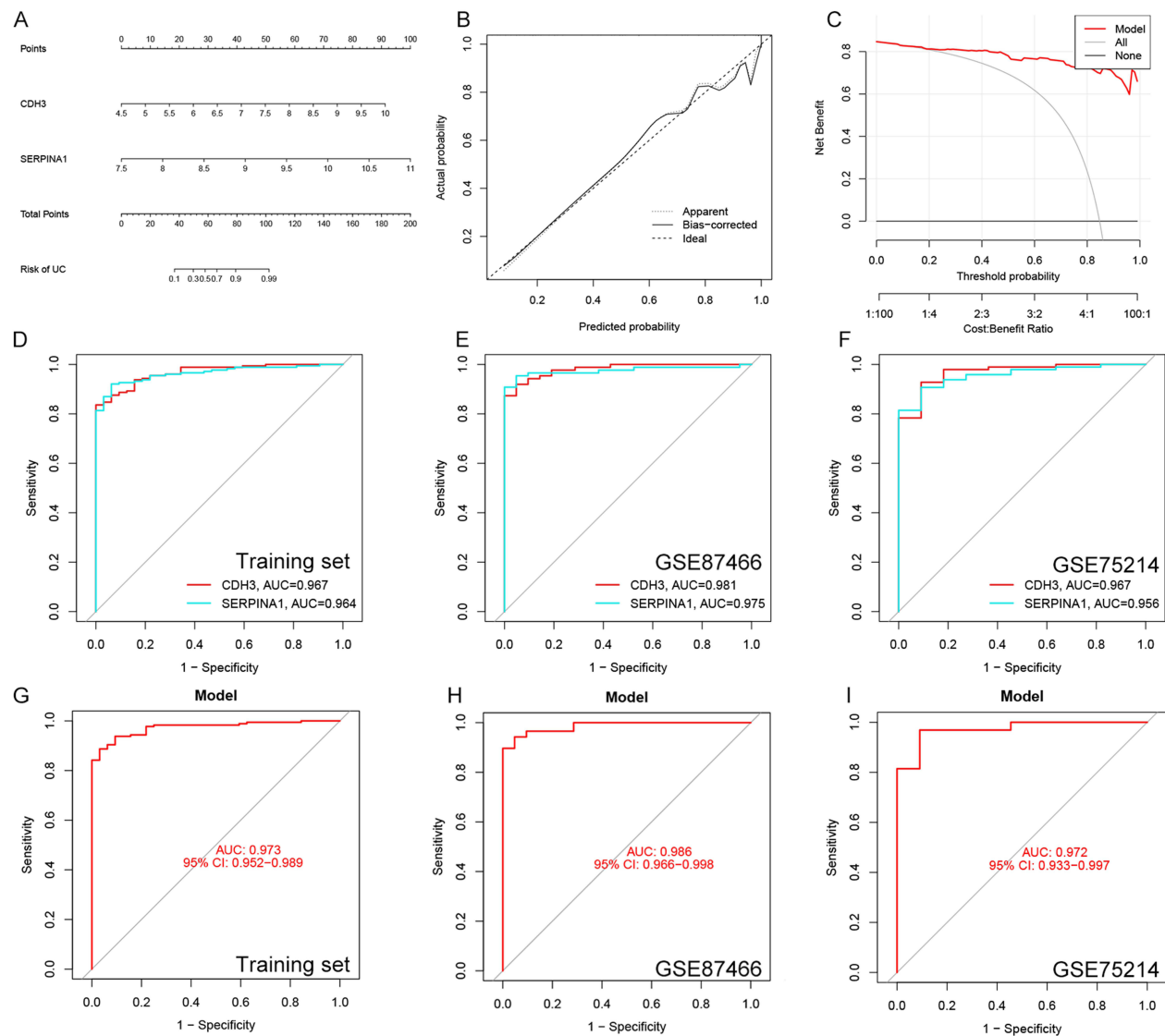
in the Random Forest algorithm, CDH3 and SERPINA1 were highlighted as key genes (Figure 4E). Upon integrating the outcomes from these three algorithms, we pinpointed two hub UC-ARGs: CDH3 and SERPINA1 (Figure 4F).

Analyzing the expression patterns of hub UC-ARGs within the training dataset (Figure 4G) and two external validation datasets (Figure 4H and I), we observed a notable upregulation of these genes in UC samples.

## Construction of Hub UC-ARGs Diagnostic Model

We devised a diagnostic nomogram based on the expression patterns of two hub UC-ARGs (Figure 5A). The predictive accuracy of the nomogram model for the training dataset was validated using calibration curves (Figure 5B). Additionally, the decision curve analysis (DCA) demonstrated that decisions made based on this model were beneficial for UC patients (Figure 5C).

To assess their diagnostic utility, we conducted a receiver operating characteristic (ROC) curve analysis. The results indicated that two hub UC-ARGs, specifically CDH3 and SERPINA1, exhibited high diagnostic accuracy for UC within the training set, with respective AUC values of 0.967 and 0.964 (Figure 5D). Furthermore, the ROC analysis outcomes for the external validation datasets, GSE87466 and GSE75214, confirmed the robustness of these findings. Specifically, CDH3 and SERPINA1 achieved AUC values of 0.981 and 0.975, respectively, for GSE87466 (Figure 5E), and AUC values of 0.967 and 0.956, respectively, for GSE75214 (Figure 5F). When subjected to rigorous 10-fold cross-validation,



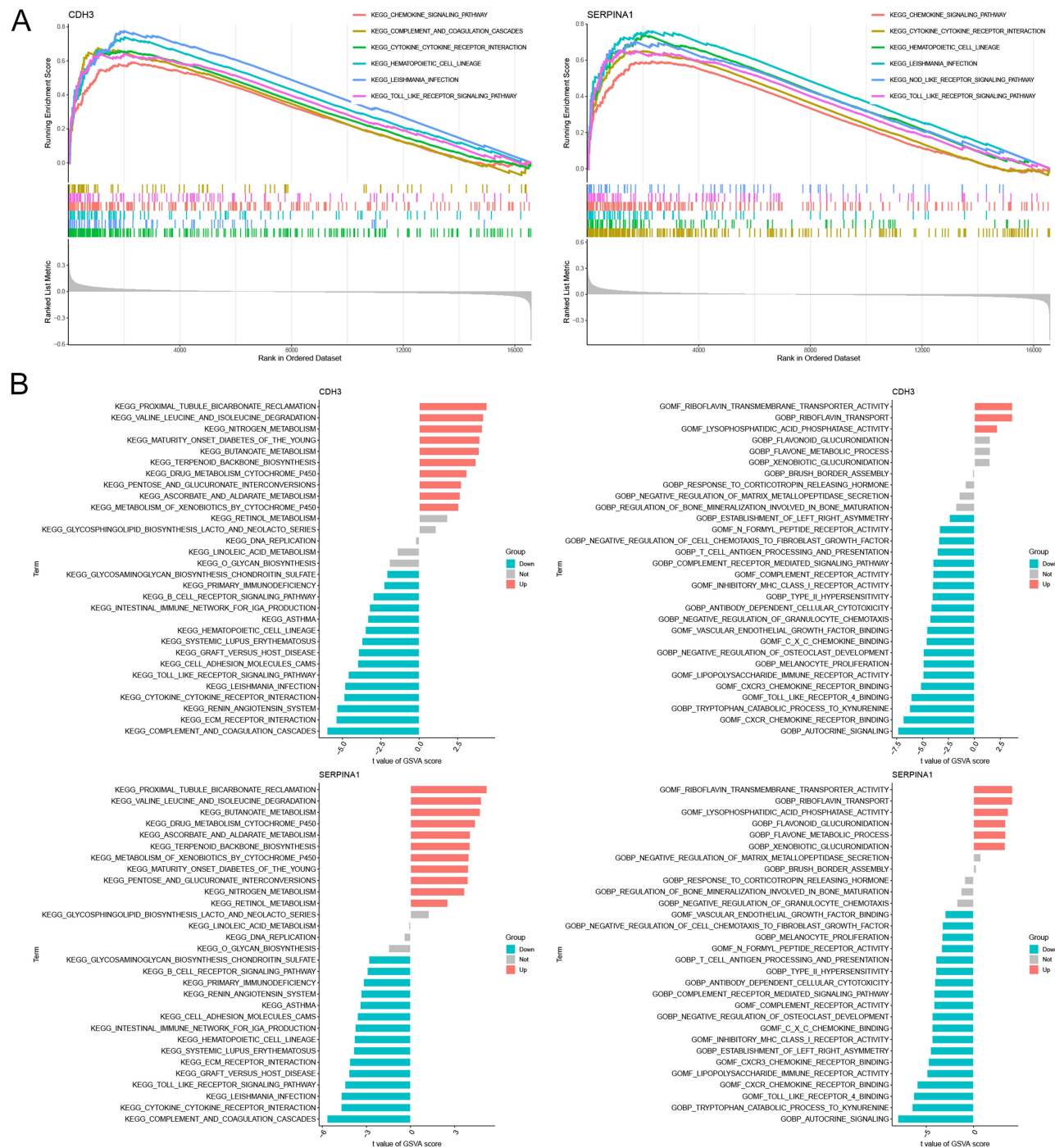
**Figure 5** Hub UC-ARGs risk prediction model. (A) Nomogram of hub UC-ARGs in the diagnosis of UC. Within this nomogram, “Points” function as a quantitative indicator reflecting the scores assigned to individual candidate genes, while “Total Points” represent the cumulative scores derived from the aggregation of all listed genes. Evaluation of the predictive efficiency and clinical utility of the nomogram in the training set: (B) Calibration Curves and (C) Decision Curve Analysis. ROC curve analysis of hub UC-ARGs in the training set (D) and validation set GSE87466 (E), GSE75214 (F). ROC curve analysis of nomogram in the training set (G) and validation set GSE87466 (H), GSE75214 (I).

the nomogram exhibited an impressive AUC of 0.973 for the training set (Figure 5G), and AUC values of 0.986 (Figure 5H) and 0.972 (Figure 5I) for the validation sets.

## Enrichment Analysis of Hub UC-ARGs

To delve deeper into the potential functions of hub UC-ARGs in UC, we conducted a single-gene GSEA and GSVA (Figure 6). The results retrieved from the KEGG database suggested that CDH3 and SERPINA1 participate in several notable pathways: the chemokine signaling pathway, hematopoietic cell lineage, leishmaniasis infection, and the toll-like receptor pathway (Figure 6A). Furthermore, the GSVA results pointed to an enhancement of various biological processes and pathways in the context of UC (Figure 6B). These up-regulated processes specifically include proximal tubule bicarbonate reclamation, the degradation of valine, leucine, and isoleucine, as well as riboflavin transmembrane transporter activity. Additionally, a negative regulation was observed in complement and coagulation cascades, among other pathways.

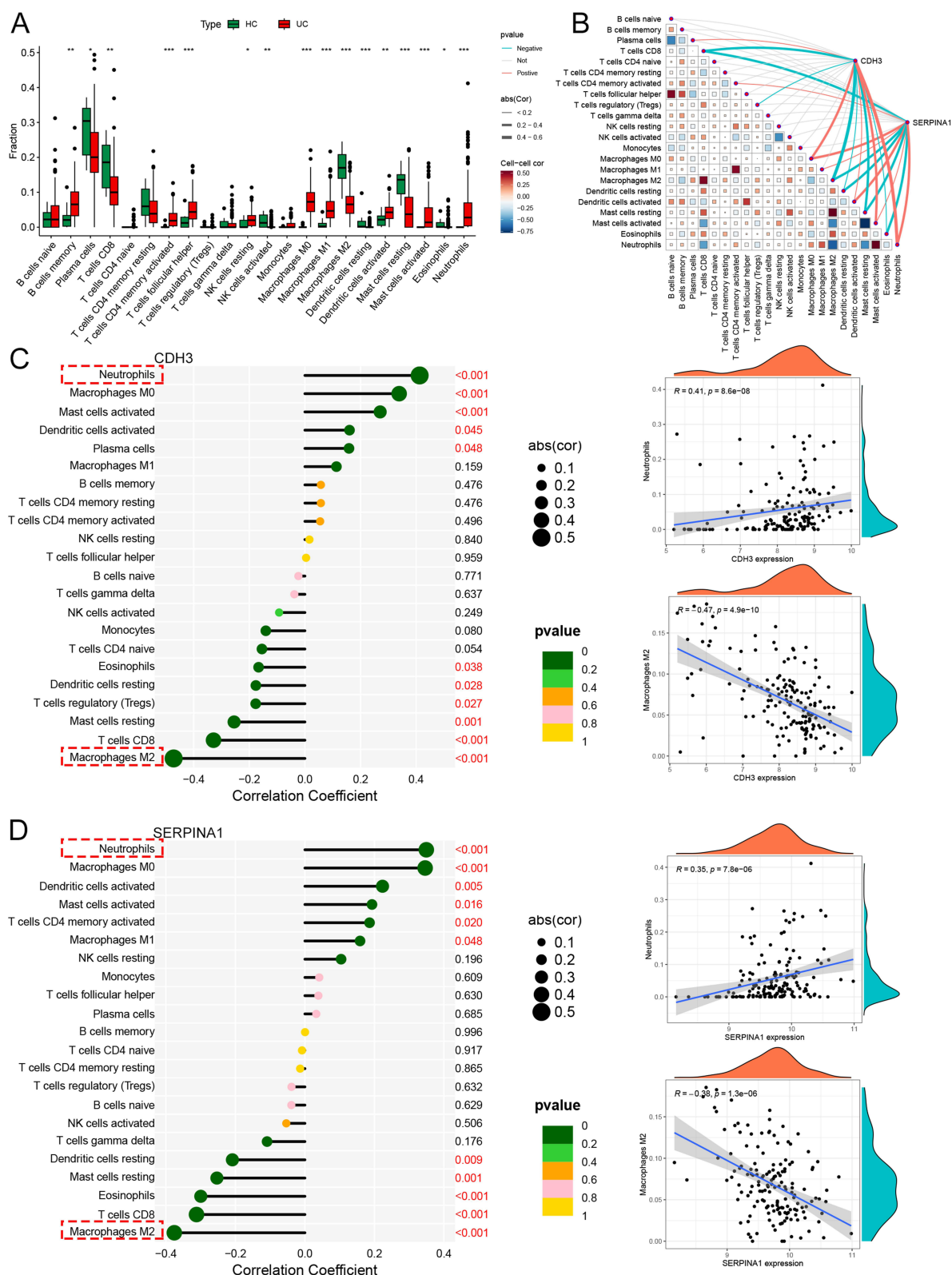




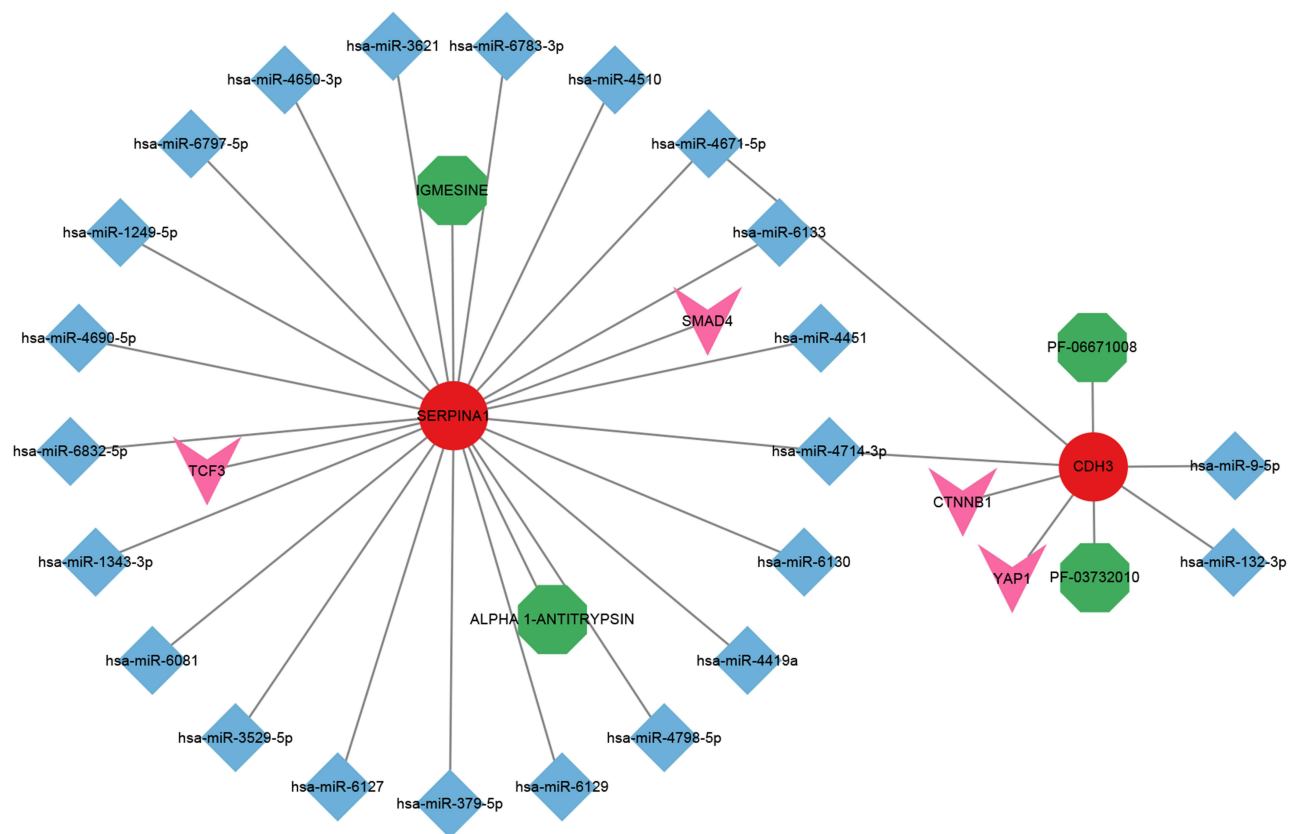
**Figure 6** GSEA and GSVA Enrichment Analysis of Hub UC-ARGs. (A) GSEA and (B)GSVA Enrichment Analysis of Hub UC-ARGs.

# The Role of CDH3 and SERPINA1 in UC Immune Microenvironment

Given the interplay between the pathophysiology of UC patients and their immune microenvironment, an in-depth investigation into this immune milieu was conducted. It is worth highlighting that 16 immune cell populations exhibited significant variations in abundance within UC samples (Figure 7A). Subsequently, Correlation analysis was performed to elucidate the relationships between hub UC-ARGs and the distinct immune cell subsets (Figure 7B). Our findings indicated that these genes positively correlated with neutrophils while exhibiting a negative correlation with M2 macrophages (Figure 7C and D).



**Figure 7** CIBERSORT immune infiltration. (A) Box diagram of differences in the infiltration of 22 immune cells in healthy control and UC samples. (B) Correlation heat map of hub UC-ARGs with 22 kinds of immune cells. (C) Correlation Analysis between CDH3 and 22 Types of Immune Cells. (D) Correlation Analysis between SERPINA1 and 22 Types of Immune Cells. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .



**Figure 8** Drug and miRNA-TF-mRNA regulatory network. Visualization of Target Gene mRNAs (Red Circles), miRNAs (Blue Squares), TFs (Pink V-Shapes), and Predicted Drugs (Green Octagons).

## Drug Prediction and miRNA-TF-mRNA Regulatory Network

Through database searches, we identified four potential drugs targeting hub UC-ARGs (Figure 8, green graphic representation). By predicting the miRNAs and TFs associated with hub UC-ARGs, we utilized Cytoscape to visualize the network, which encompassed 23 miRNAs, 4 TFs, and 2 mRNAs. Notably, miRNA-4671-5p and miRNA-4714-3p were identified as common regulatory factors within this network (Figure 8).

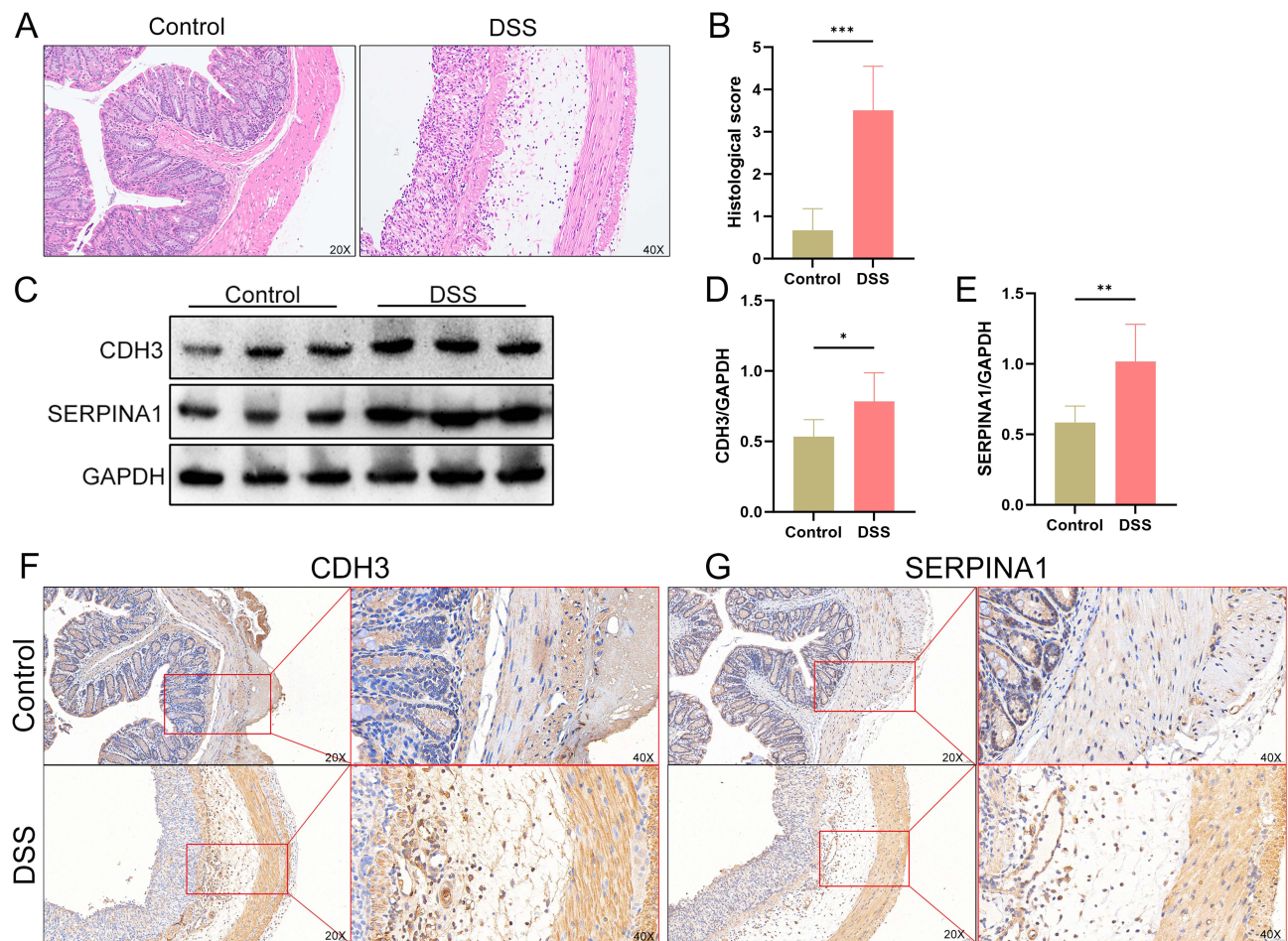
## Verification of Hub Gene Expression in Colon Tissue Induced by DSS in Mice

A colitis animal model was successfully established through the administration of 3% DSS, which was corroborated by HE staining and histopathological scoring (Figure 9A and B). To further verify the expression levels of CDH3 and SERPINA1. Western blot analysis was conducted in the colonic tissues of mice with DSS-induced colitis, in comparison with control mice. Consistent with our previous findings, the expression of these genes was significantly upregulated in the colonic tissues of the DSS group (Figure 9C–E). Immunohistochemical staining also corroborated this observation (Figure 9F and G).

## Discussion

Ulcerative colitis (UC) is a prevalent chronic IBD that significantly reduces the quality of life of patients due to its recurrent and intractable nature.<sup>27</sup> Despite the progress made in therapeutic approaches for UC, further lowering recurrence rates and achieving disease eradication remain a formidable global challenge. Anoikis, a distinct form of programmed cell death induced by the absence or defective adhesion of cells to the ECM, holds a pivotal position in numerous biological phenomena, encompassing developmental processes, tissue maintenance, disease initiation, and tumor progression.<sup>28</sup> Studies have shown that glucomannan derived from Aloe vera gel maintains the integrity of the





**Figure 9** Verification of colonic tissue gene expression induced by DSS in mice with colitis. **(A)** Representative H&E staining of colonic tissues from each group. **(B)** Statistical results for histological damage scores in each group ( $n = 6$ ). **(C)** Western blot analysis of CDH3 and SERPINA1 levels in colon tissues. Quantitative Western blot analysis results for **(D)** CDH3 and **(E)** SERPINA1. mean  $\pm$  SD. ( $n=6$ ) \* $P < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ . Representative immunohistochemical staining images of **(F)** CDH3 and **(G)** SERPINA1 expression in colon tissue.

intestinal barrier by alleviating anoikis through modulation of the Nrf2-mitochondrial axis.<sup>29</sup> However, there is still a scarcity of reports clarifying the complex mechanisms underlying anoikis in UC. Our investigation delves into the significance of anoikis in UC by pinpointing hub ARG-DEGs and immune cell populations implicated, offering a fresh insight into the underlying mechanisms and therapeutic approaches for this disease.

In this study, we conducted a comprehensive analysis to identify and characterize hub ARG-DEGs in UC. By integrating two large datasets and employing rigorous bioinformatics methods, we have made several novel findings that contribute to the understanding of UC at the molecular level.

Firstly, our differential expression analysis identified 719 DEGs between UC samples and healthy controls. By intersecting this set with a predefined list of anoikis-related genes, we pinpointed 49 ARG-DEGs. These genes were significantly enriched in multiple biological pathways, including focal adhesion, the PI3K-AKT signaling pathway, inflammatory signal regulation, and ECM-receptor interaction. These pathways are recognized for their vital functions in cellular survival, proliferation, and inflammatory responses, and disturbances within them have been associated with the development of UC.<sup>30–32</sup> Hence, our results imply that these ARG-DEGs may play a role in the initiation and advancement of UC by participating in these signaling pathways.

Secondly, we employed WGCNA to identify hub modules associated with UC. The midnight blue module emerged as significantly associated with UC, and we further refined our analysis to focus on 269 hub genes within this module. By intersecting these hub genes with the ARG-DEGs, we identified eight potential hub UC-ARGs. Notably, CDH3 and

SERPINA1 stood out as the most significant genes based on their strong correlation and consistent identification across multiple machine learning algorithms. The results indicate that CDH3 and SERPINA1 may have crucial roles in the development of UC and could potentially be utilized as diagnostic biomarkers or therapeutic interventions.

The diagnostic accuracy of CDH3 and SERPINA1 was further validated through ROC curve analysis, which demonstrated high AUC values for both the training and external validation datasets. This robust performance highlights the potential clinical utility of these genes in diagnosing UC. Furthermore, we constructed a diagnostic nomogram based on the expression patterns of CDH3 and SERPINA1, which exhibited impressive predictive precision and was advantageous for UC patients in decision-making. These findings underscore the potential of our identified hub UC-ARGs as valuable tools in the clinical management of UC.

CDH3 (Cadherin-3), a member of the cadherin family, is a cell adhesion molecule that plays pivotal roles in various cellular processes, including cell differentiation, migration, and tumorigenesis.<sup>33</sup> Studies have indicated that the expression of CDH3 in colorectal cancer tissues is closely associated with inflammatory responses, suggesting its potential as a biomarker for the inflammatory status in colorectal cancer.<sup>34</sup> Notably, colitis is intimately linked to the development of tumors. SERPINA1 is a plasma protein primarily functioning to inhibit serine proteases, such as neutrophil elastase. It participates in a multitude of physiological processes, encompassing coagulation, inflammation, immune responses, and cellular migration.<sup>35</sup> It belongs to the same proinflammatory cytokine cluster as interleukin-6 (IL-6) and C-reactive protein (CRP).<sup>36</sup> The levels of SERPINA1 are associated with inflammatory conditions. Notably, the expression level of SERPINA1 in tissues, particularly in colorectal cancer tissues, has been demonstrated to have a close correlation with inflammatory responses.<sup>37</sup> This finding implies its potential application as a biomarker for inflammatory states in cancer and other diseases.

To gain further insights into the potential roles of CDH3 and SERPINA1 in UC, we conducted enrichment analysis using the KEGG database and GSVA. The results suggested that these genes participate in several notable pathways, including chemokine signaling, hematopoietic cell lineage, and toll-like receptor signaling. These pathways are known to be involved in immune regulation and inflammation, which are central to the pathophysiology of UC. Additionally, the GSVA results pointed to an enhancement of various biological processes and pathways in UC, further supporting the potential roles of CDH3 and SERPINA1 in this disease.

Thirdly, we investigated the correlations between hub UC-ARGs and immune cell subsets in the UC microenvironment. Our findings indicated that CDH3 and SERPINA1 positively correlated with neutrophils while exhibiting a negative correlation with M2 macrophages. This suggests that these genes may play a role in modulating the immune response in UC, which could have implications for disease progression and treatment.

Finally, we successfully established an animal model of colitis in mice through pharmacological induction. Further validation and confirmation of the expression of CDH3 and SERPINA1 in colitis were conducted through HE staining, Western blot analysis, and immunohistochemical staining. These findings were consistent with our bioinformatics analysis results, further affirming the robustness of our analytical outcomes.

However, the study is subject to certain limitations. Firstly, the analysis relies on publicly available databases, which inherently lack primary sequencing data and may consequently introduce selection bias. Secondly, the limited sample size presents challenges in comprehensively accounting for potential confounding factors, such as gender, age, comorbidities, and ethnicity. Additionally, the experiment only validated the expression patterns without delving into the underlying mechanisms. Furthermore, the conclusion that CDH3 and SERPINA1 contribute to UC progression by modulating immune-inflammatory responses is primarily based on correlations and enrichment analyses. Considering these limitations, there is a pressing need for extensive, multicenter, prospective studies with larger sample sizes to rigorously explore the predictive potential of these signature genes in ulcerative colitis.

## Conclusion

In summary, our study has identified and characterized key anoikis-related differentially expressed genes in ulcerative colitis, offering novel insights into the underlying molecular mechanisms of this disease. The identified genes, notably CDH3 and SERPINA1, demonstrate high diagnostic accuracy. Future research is required to further validate these findings and explore the functional roles of these genes in the context of UC.



## Data Sharing Statement

The data from this study are publicly accessible in the Gene Expression Omnibus (GEO) database.

## Ethics and Dissemination

In accordance with the “Measures for Ethical Review of Life Science and Medical Research Involving Human Subjects” issued by China on February 18, 2023, research activities that meet specific conditions are exempt from ethical review, and this study meets the exemption criteria. Specifically, as per Item 1, Article 32, the research does not involve direct interaction with human subjects and poses no potential risk to their physical and mental health. Only anonymized data obtained legally are used, without identifying individual subjects. According to Item 2, Article 32, the research methods and procedures are based on established scientific principles and do not involve any new or untested techniques that may harm human subjects. All data in this study are sourced from the publicly - accessible Gene Expression Omnibus (GEO) database, which meets relevant ethical and legal requirements. Thus, this research project doesn’t need review and approval from the relevant ethical review committee.

## Acknowledgment

We acknowledge GEO database and contributors for sharing datasets.

## Funding

There is no funding to report.

## Disclosure

The authors report no conflicts of interest in this work.

## References

1. Cohen RD, Yu AP, Wu EQ, Xie J, Mulani PM, Chao J. Systematic review: the costs of ulcerative colitis in Western countries. *Aliment Pharmacol Ther.* 2010;31(7):693–707. doi:10.1111/j.1365-2036.2010.04234.x
2. Le Berre C, Honap S, Peyrin-Biroulet L. Ulcerative colitis. *Lancet.* 2023;402(10401):571–584. doi:10.1016/S0140-6736(23)00966-2
3. Wang Y, Cheng S, Fleishman JS, et al. Targeting anoikis resistance as a strategy for cancer therapy. *Drug Resist Updat.* 2024;75:101099. doi:10.1016/j.drug.2024.101099
4. Han YH, Wang Y, Lee SJ, Jin MH, Sun HN, Kwon T. Regulation of anoikis by extrinsic death receptor pathways. *Cell Commun Signal.* 2023;21(1):227. doi:10.1186/s12964-023-01247-5
5. Mei J, Jiang XY, Tian HX, et al. Anoikis in cell fate, physiopathology, and therapeutic interventions. *Med Commun.* 2024;5(10):e718. doi:10.1002/mco.2718
6. Zou Y, Xu L, Wang W, et al. Muscone restores anoikis sensitivity in TMZ-resistant glioblastoma cells by suppressing TOP2A via the EGFR/Integrin beta1/FAK signaling pathway. *Phytomedicine.* 2024;129:155714. doi:10.1016/j.phymed.2024.155714
7. Satyavarapu EM, Nath S, Mandal C. Desialylation of Atg5 by sialidase (Neu2) enhances autophagosome formation to induce Anchorage-dependent cell death in ovarian cancer cells. *Cell Death Discov.* 2021;7(1):26. doi:10.1038/s41420-020-00391-y
8. Tan Y, Wang L, Chen G, et al. Hyaluronate supports hESC-cardiomyocyte cell therapy for cardiac regeneration after acute myocardial infarction. *Cell Prolif.* 2020;53(12):e12942. doi:10.1111/cpr.12942
9. Taddei ML, Giannoni E, Fiaschi T, Chiarugi P. Anoikis: an emerging hallmark in health and diseases. *J Pathol.* 2012;226(2):380–393. doi:10.1002/path.3000
10. Zhi Z, Ouyang Z, Ren Y, et al. Non-canonical phosphorylation of Bmf by p38 MAPK promotes its apoptotic activity in anoikis. *Cell Death Differ.* 2022;29(2):323–336. doi:10.1038/s41418-021-00855-3
11. Sun Q, Yang Z, Li P, et al. A novel miRNA identified in GRSF1 complex drives the metastasis via the PIK3R3/AKT/NF-kappaB and TIMP3/MMP9 pathways in cervical cancer cells. *Cell Death Dis.* 2019;10(9):636.
12. 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):882–883. doi:10.1093/bioinformatics/bts034
13. Yu G, Wang LG, Han Y, He QY. ClusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012;16(5):284–287. doi:10.1089/omi.2011.0118
14. Shen W, Song Z, Zhong X, et al. Sangerbox: a comprehensive, interaction-friendly clinical bioinformatics analysis platform. *Imeta.* 2022;1(3):e36. doi:10.1002/imt.236
15. Frost HR, Amos CI. Gene set selection via LASSO penalized regression (SLPR). *Nucleic Acids Res.* 2017;45(12):e114. doi:10.1093/nar/gkx291
16. Nedaie A, Najafi AA. Support vector machine with Dirichlet feature mapping. *Neural Netw.* 2018;98:87–101. doi:10.1016/j.neunet.2017.11.006
17. Zhou J, Huang J, Li Z, et al. Identification of aging-related biomarkers and immune infiltration characteristics in osteoarthritis based on bioinformatics analysis and machine learning. *Front Immunol.* 2023;14:1168780. doi:10.3389/fimmu.2023.1168780

18. Deng B, Chen Y, He P, et al. Identification of Mitophagy-Associated Genes for the Prediction of Metabolic Dysfunction-Associated Steatohepatitis Based on Interpretable Machine Learning Models. *J Inflamm Res.* **2024**;17:2711–2730. doi:10.2147/JIR.S450471
19. Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP. GSEA-P: a desktop application for Gene Set Enrichment Analysis. *Bioinformatics.* **2007**;23(23):3251–3253. doi:10.1093/bioinformatics/btm369
20. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf.* **2013**;14(1):7. doi:10.1186/1471-2105-14-7
21. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* **2015**;12(5):453–457. doi:10.1038/nmeth.3337
22. Cannon M, Stevenson J, Stahl K, et al. DGIdb 5.0: rebuilding the drug-gene interaction database for precision medicine and drug discovery platforms. *Nucleic Acids Res.* **2024**;52(D1):D1227–D1235. doi:10.1093/nar/gkad1040
23. Huang HY, Lin YC, Cui S, et al. miRTarBase update 2022: an informative resource for experimentally validated miRNA-target interactions. *Nucleic Acids Res.* **2022**;50(D1):D222–D230. doi:10.1093/nar/gkab1079
24. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. *Elife.* **2015**;4. doi:10.7554/eLife.05005
25. Reis-de-Oliveira G, Carregari VC, Sousa G, Martins-de-Souza D. OmicScope unravels systems-level insights from quantitative proteomics data. *Nat Commun.* **2024**;15(1):6510. doi:10.1038/s41467-024-50875-z
26. Yan B, Li X, Zhou L, et al. Inhibition of IRAK 1/4 alleviates colitis by inhibiting TLR4/ NF-kappaB pathway and protecting the intestinal barrier. *Bosn J Basic Med Sci.* **2022**;22(6):872–881. doi:10.17305/bjbm.2022.7348
27. Cosnes J, Gower-Rousseau C, Seksik P, Cortot A. Epidemiology and natural history of inflammatory bowel diseases. *Gastroenterology.* **2011**;140(6):1785–1794. doi:10.1053/j.gastro.2011.01.055
28. Dai Y, Zhang X, Ou Y, et al. Anoikis resistance--protagonists of breast cancer cells survive and metastasize after ECM detachment. *Cell Commun Signal.* **2023**;21(1):190. doi:10.1186/s12964-023-01183-4
29. Zhang D, Zhou X, Zhang K, Yu Y, Cui SW, Nie S. Glucomannan from Aloe vera gel maintains intestinal barrier integrity via mitigating anoikis mediated by Nrf2-mitochondria axis. *Int J Biol Macromol.* **2023**;235:123803. doi:10.1016/j.ijbiomac.2023.123803
30. Yao T, Wu Y, Fu L, Lv J, Lv L, Li L. Christensenellaceae minuta modulates epithelial healing via PI3K-AKT pathway and macrophage differentiation in the colitis. *Microbiol Res.* **2024**;289:127927. doi:10.1016/j.micres.2024.127927
31. Peng Y, Zhu J, Li Y, Yue X, Peng Y. Almond polysaccharides inhibit DSS-induced inflammatory response in ulcerative colitis mice through NF-kappaB pathway. *Int J Biol Macromol.* **2024**;281(Pt 1):136206. doi:10.1016/j.ijbiomac.2024.136206
32. Pang X, Song H, Li X, et al. Transcriptomic analyses of treatment-naïve pediatric ulcerative colitis patients and exploration of underlying disease pathogenesis. *J Transl Med.* **2023**;21(1):30. doi:10.1186/s12967-023-03881-6
33. Faraj Tabrizi P, Peters I, Schimansky I, et al. Alteration of Cadherin 3 Expression and DNA Methylation in Association with Aggressive Renal Cell Carcinoma. *Int J mol Sci.* **2023**;24(22):16476. doi:10.3390/ijms242216476
34. Song J, Jin Y, Fan S, et al. CDH3 Is an Effective Serum Biomarker of Colorectal Cancer Distant Metastasis Patients. *J Cancer.* **2024**;15(16):5218–5229. doi:10.7150/jca.98337
35. Mazzuca C, Vitiello L, Travaglini S, et al. Immunological and homeostatic pathways of alpha -1 antitrypsin: a new therapeutic potential. *Front Immunol.* **2024**;15:1443297. doi:10.3389/fimmu.2024.1443297
36. McElvaney OJ, O'Connor E, McEvoy NL, et al. Alpha-1 antitrypsin for cystic fibrosis complicated by severe cytokinemic COVID-19. *J Cyst Fibros.* **2021**;20(1):31–35. doi:10.1016/j.jcf.2020.11.012
37. Ma Y, Chen Y, Zhan L, et al. CEBPB-mediated upregulation of SERPINA1 promotes colorectal cancer progression by enhancing STAT3 signaling. *Cell Death Discov.* **2024**;10(1):219. doi:10.1038/s41420-024-01990-9

## Journal of Inflammation Research

### Publish your work in this journal

The Journal of Inflammation Research is an international, peer-reviewed open-access journal that welcomes laboratory and clinical findings on the molecular basis, cell biology and pharmacology of inflammation including original research, reviews, symposium reports, hypothesis formation and commentaries on: acute/chronic inflammation; mediators of inflammation; cellular processes; molecular mechanisms; pharmacology and novel anti-inflammatory drugs; clinical conditions involving inflammation. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/journal-of-inflammation-research-journal>

**Dovepress**  
Taylor & Francis Group