

Immune-metabolic Crosstalk Driven by Pyroptosis in Colorectal Cancer: Identification of an Immunoregulatory Prognostic Signature and Mechanistic Pathways

Daoli Liu, Bohuang Zeng, Yuming Gong, Jian Wang, Donghua Gao, Yangfen He, Yi Zhou, Jiayi Wu, Rong Zhang, and Huajing Long

The Affiliated Traditional Chinese Medicine Hospital, Guangzhou Medical University, Guangzhou, Guangdong, China

Received: 19 May 2026; Received in revised form: 13 July 2026; Accepted: 19 July 2026

ABSTRACT

Pyroptosis is an inflammatory form of programmed cell death involved in antitumor immunity and tumor microenvironment regulation. However, its relationships with immune signaling and metabolic adaptation in colorectal cancer (CRC) remain incompletely understood.

An integrated bioinformatics analysis was performed using Gene Expression Omnibus datasets GSE20916 and GSE73360 and The Cancer Genome Atlas colorectal adenocarcinoma cohort. Pyroptosis- and glucose metabolism-related genes were retrieved from public databases and the literature. Differential expression, functional enrichment, Cox regression, least absolute shrinkage and selection operator modeling, gene set enrichment and variation analyses, and protein-protein interaction analyses were conducted.

Fifty-six pyroptosis- and metabolism-related differentially expressed genes were identified. Enrichment analyses highlighted immune and inflammatory pathways, including IL-17 signaling, cytokine-mediated responses, neutrophil extracellular trap formation, chemokine receptor interactions, and IL-10 signaling. A seven-gene prognostic signature comprising ELANE, BRCA1, FKBP10, H1-5, CXCL8, GATA1, and FNDC5 was established, with receiver operating characteristic areas under the curve exceeding 0.90. Risk-stratified analyses revealed enrichment of Wnt/ β -catenin and Notch signaling, suggesting associations between immune regulation and metabolic adaptation. Protein-protein interaction analysis identified CXCL8, ELANE, BRCA1, GATA1, and FNDC5 as hub genes potentially involved in tumor-associated immune remodeling.

These findings suggest interconnected pyroptosis, immune, and metabolic processes in CRC and identify a prognostic signature with potential biomarker value. Inflammatory cytokine signaling and Wnt/Notch-associated pathways warrant further investigation to clarify their mechanistic roles and relevance to immunotherapeutic stratification.

Keywords: Colorectal cancer; Cytokine Signaling; Immune-metabolic reprogramming; Prognostic signature; Pyroptosis; Tumor immunology

Corresponding Author: Huajing Long, MM;
The Affiliated Traditional Chinese Medicine Hospital, Guangzhou
Medical University, Guangzhou, Guangdong, China. Tel: (+86
159) 1850 0203, Email: huajing6668@163.com

INTRODUCTION

Colorectal carcinoma (CRC) is the third most frequently diagnosed malignancy and the second leading

cause of cancer-related death worldwide, with approximately 1.9 million new cases and 935 000 deaths annually and a sharp drop in survival once the disease metastasizes.¹ Current surgery, chemotherapy, radiotherapy, and targeted agents are limited by high recurrence (30% to 50% within 5 years) and treatment-related toxicity, and therapy resistance remains a major obstacle.^{2,3}

Beyond these clinical challenges, a key mechanistic question is how programmed cell death and cellular metabolism jointly shape CRC. Pyroptosis—an inflammatory form of programmed cell death—and glycolytic reprogramming are both recognized drivers of tumor progression, yet the molecular crosstalk linking pyroptosis to glucose metabolism in CRC is poorly understood.^{4,5} Dissecting this interplay is logically demanding because the two processes are deeply entangled: metabolic rewiring fuels the inflammatory cascade that pyroptosis triggers, while pyroptotic cell death in turn remodels the metabolic landscape of the tumor microenvironment.

To address this gap, we integrated multi-omics transcriptomic data to characterize pyroptosis- and glucose metabolism-related alterations in CRC and to build a prognostic signature that captures their crosstalk. Using TCGA and GEO cohorts, we performed differential-expression, enrichment, Cox/LASSO modeling, and network analyses to identify candidate regulators and mechanistic pathways with potential clinical relevance.⁶

This study examined transcriptional changes in pyroptotic regulators and glycolytic pathway components in CRC. Using advanced bioinformatics methodologies, we conducted a comprehensive analysis of publicly available datasets to identify key genes and pathways involved in these processes. Through systematic data normalization and comparative expression profiling, we performed detailed pathway mapping. The methodological framework integrated multivariate prognostic modeling via Cox regression, enhanced by least absolute shrinkage and selection operator (LASSO)-based variable selection for dimensionality reduction.⁶ This integrated approach was employed to clarify the pathophysiological pathways underlying CRC while identifying potential molecular targets for therapy.

This investigation sought to define the functional interplay between pyroptotic regulators and glycolytic metabolism in CRC by developing an integrated

prognostic signature. Through construction of this molecular model, we aimed to improve prognostic accuracy and establish a mechanistic framework for patient stratification, facilitating personalized therapeutic approaches in oncology.⁷ The findings are expected to advance precision medicine by revealing key relationships between pyroptosis and metabolic reprogramming, potentially identifying novel therapeutic targets and supporting biomarker-guided management strategies to improve patient outcomes.⁸

MATERIALS AND METHODS

Data Download

Transcriptomic data acquisition and processing were conducted using TCGAAbiolinks⁹ (v2.30.0) to retrieve The Cancer Genome Atlas colorectal adenocarcinoma dataset (TCGA-COADREAD) from the Genomic Data Commons portal. Following quality control and exclusion of samples lacking complete clinical annotations, the final cohort included 601 CRC cases and 48 nonmalignant controls, with raw sequencing counts converted to FPKM values. Corresponding clinical metadata were obtained through the UCSC Xena platform,¹⁰ with detailed cohort characteristics summarized in Table 1.

Transcriptomic datasets GSE20916 and GSE73360 were downloaded from the Gene Expression Omnibus (GEO) repository¹¹ using GEOquery (v2.70.0),¹² with GSE73360 used as an independent validation cohort. Both datasets originated from *Homo sapiens* colorectal tissue: CRC samples from tumor tissue and control samples from colonic mucosa. The GSE20916 dataset (GPL570 platform) comprised 36 CRC samples from macroscopic sources and 10 normal rectal mucosa samples from microscopic sources, whereas GSE73360 (GPL17586 platform) contained 55 CRC samples from macroscopic sources and 31 normal rectal mucosa samples. All CRC samples and controls were included in subsequent analyses.

Gene annotation data were retrieved from GeneCards¹³ using strict filtering criteria to identify pyroptosis-related genes (PRGs) and glucose metabolism-associated genes (GMRGs). Searching “Pyroptosis” with protein-coding status and a relevance score >1 identified 366 (PRGs), with parallel analysis for GMRGs yielding 4376 candidates. Additional PubMed mining using “Pyroptosis” and “Glucose Metabolism” as search terms retrieved 35 PRGs¹⁴ and 290 GMRGs,¹⁵

Pyroptotic-metabolic Signatures in Colorectal Cancer

respectively. After merging and deduplication, 371 PRGs and 4543 GMRGs were compiled. Intersection analysis identified 223 overlapping genes involved in pyroptosis and glucose metabolism, i.e., pyroptosis and glucose metabolism-related genes (P&GMRGs). GeneCards and PubMed were selected because they provide complementary, curated coverage of both experimentally validated and literature-supported gene sets, whereas TCGA and GEO supplied cohort-specific expression evidence; overlapping genes emerging from multiple sources were prioritized as the most robust candidates, and analogous curated repositories would be expected to yield a largely concordant P&GMRG set (the stepwise filtering strategy is summarized in Figure 1). Because the two source gene sets (PRGs and

GMRGs) share substantial functional overlap (notably inflammation-related genes), we applied an additional functional-deduplication step in which identifiers from GeneCards and PubMed were unified to HGNC-approved symbols and only unique genes were carried forward, thereby reducing redundancy before intersection.

The limma package in R¹⁶ (v3.58.1) was used to process GSE20916 and GSE73360 independently. Probe reannotation, standardization, and normalization were performed to correct batch effects and generate adjusted expression matrices. The adjusted GSE20916 dataset contained 36 CRC and 10 control cases; adjusted GSE73360 included 55 CRC and 31 control cases. Each dataset was validated separately.

Table 1. Baseline characteristics of the overall cohort.

Characteristics	Overall
Age, n (%)	601
≥60	430 (72%)
<60	171 (28%)
Gender, n (%)	601
Female	279 (46%)
Male	322 (54%)
Stage, n (%)	601
Stage I	109 (18%)
Stage II	226 (38%)
Stage III	179 (30%)
Stage IV	87 (14%)

Differentially Expressed P&GMRGs in CRC

Based on the TCGA-COADREAD cohort, samples were divided into CRC cases and nonmalignant controls. Differential gene expression analysis was conducted using DESeq2¹⁷ (v1.42.0) to compare transcriptional profiles between malignant and normal tissues. Thresholds of $|\log FC| > 1$ and adjusted $p < 0.05$ defined differentially expressed genes (DEGs): $\log FC > 1$ indicated upregulation and $\log FC < -1$ indicated downregulation. p values were adjusted using the Benjamini–Hochberg (BH) method. Expression patterns were visualized using volcano plots generated via the R package ggplot2 (v3.4.4), illustrating significance versus fold-change relationships.

To identify differentially expressed P&GMRGs (P&GMRDEGs), results from TCGA-COADREAD variance analysis ($|\log FC| > 1$ and adjusted $p < 0.05$) were intersected with predefined P&GMRGs, with results

visualized via Venn diagrams. The resulting P&GMRDEGs were further analyzed through hierarchical clustering using pheatmap (v1.0.12), highlighting expression profiles for the top 20 most dysregulated genes.

Gene Ontology and Kyoto Encyclopedia of Genes and Genomes Pathway Enrichment Analyses

Functional annotation of P&GMRDEGs was performed through Gene Ontology (GO)¹⁸ analysis, encompassing biological process (BP), cellular component (CC), and molecular function (MF) GO terms. Pathway enrichment was evaluated using the Kyoto Encyclopedia of Genes and Genomes (KEGG)¹⁹ database. Both analyses were conducted via clusterProfiler (v4.10.0),²⁰ applying stringent significance thresholds (adjusted $p < 0.05$ and $q < 0.05$) with BH correction for multiple testing. Reassessment

under this stricter threshold confirmed that the principal immune and inflammatory pathways (IL-17 signaling, IL-10 signaling, neutrophil extracellular trap formation, chemokine receptor interactions) remained highly significant; pathways such as systemic lupus erythematosus and alcoholism, although still meeting the nominal threshold, are interpreted cautiously as reflecting shared inflammatory gene modules rather than CRC-specific biology.

Risk Model Construction for CRC Prognosis

Prognostic modeling for the TCGA-COADREAD cohort was performed using sequential univariate and multivariate Cox regression analyses via the survival package²¹ (v3.5-7). This framework assessed the prognostic importance of P&GMRDEGs while incorporating relevant clinical variables. Initially, univariate Cox regression identified variables with $p < 0.10$, which were then included in LASSO regression for refined feature selection.

Feature selection was executed using LASSO regression²² (glmnet v4.1-8, family="cox") with 10-fold cross-validation on P&GMRDEGs identified through univariate Cox analysis. This regularization method enhances generalizability by penalizing model complexity ($\lambda \times |\text{slope coefficients}|$) to improve predictive performance. The final multivariate Cox regression model incorporated genes selected via LASSO optimization. The risk score was calculated using the following formula:

$$\text{Risk score} = \sum_i \text{coefficient}(\text{gene}_i) \times \text{mRNA expression}(\text{gene}_i).$$

Cox regression results were visualized through forest plots showing survival associations for the included P&GMRDEGs. LASSO outputs were illustrated with coefficient trajectory plots and risk score distribution diagrams, demonstrating variable selection and prognostic stratification.

Validation of Prognostic Risk Models for CRC

To assess predictive accuracy, risk scores were computed for each TCGA-COADREAD sample using gene expression profiles and model coefficients. The cohort was stratified into high- and low-risk subgroups based on the median risk score.

Comparative survival analysis between risk subgroups was performed using Kaplan–Meier

analysis²³ (survival package v3.5-7), with overall survival (OS) differences shown via risk-stratified survival curves. Using ggplot2, risk factors, risk scores, risk groups, OS, and model gene expression distributions were visualized.

To verify the independent prognostic value of risk stratification, univariate and multivariate Cox regression analyses incorporated risk scores with clinical variables (age, sex, and tumor stage). Results were displayed as forest plots, illustrating independent prognostic effects. Differential expression of model genes between risk-stratified subgroups was examined through comparative expression profiling. Calibration curves were used to evaluate the prognostic model's predictive accuracy and discriminative ability by comparing predicted versus observed survival outcomes across risk strata.

Gene Set Enrichment Analysis

Differential expression analysis of risk-stratified subgroups was performed using DESeq2, identifying significantly dysregulated genes ($|\log\text{FC}| > 1$ and adjusted $p < 0.05$; $\log\text{FC} > 1$ upregulated and $\log\text{FC} < -1$ downregulated). Results were visualized through volcano plots generated via ggplot2. Gene set enrichment analysis (GSEA)²⁴ was employed to assess the collective behavior of predefined gene sets within phenotype-correlated rankings to reveal biological significance. Transcriptomic data from TCGA-COADREAD stratified by risk status were ranked by logFC values. GSEA was performed using clusterProfiler²⁰ with the following parameters: random seed=2020, minimum gene set size=10, and maximum gene set size=500. Additionally, c2 curated gene sets were retrieved from MSigDB (v2023.2.Hs.symbols), applying significance thresholds of adjusted $p < 0.05$ and $q < 0.05$ with BH correction for multiple testing.

Gene Set Variation Analysis

Gene set variation analysis (GSVA)²⁵ was applied as a nonparametric, unsupervised method to transform individual gene expression matrices into sample-level enrichment profiles. Additionally, c2 curated pathway gene sets were obtained from MSigDB²⁶ (v2023.2.Hs.symbols). Using GSVA (v1.50.0), pathway enrichment differences between high- and low-risk TCGA-COADREAD cohort subgroups were assessed, with significance set at adjusted $p < 0.05$ following BH correction.

Protein–Protein Interaction Network

Protein–protein interaction (PPI) networks represent complex systems of interconnected proteins that regulate cellular processes through molecular interactions. These networks are central to signal transduction, transcriptional regulation, metabolic balance, and cell cycle control, shaping physiological and pathological states. Systematic analysis of PPIs clarifies not only the functional relationships among proteins but also the organization of biological systems and mechanisms underlying disease and metabolism. The STRING database (<https://cn.string-db.org/>)²⁷ provides experimentally verified and predicted PPIs. Using STRING, we constructed a PPI network for model gene–encoded proteins with an interaction confidence threshold > 0.400 . Network topology analysis identified densely interconnected regions corresponding to potential molecular complexes with specific biological roles. Hub genes were determined based on connectivity and incorporated into downstream analyses. Network visualization and analysis were achieved via Cytoscape.²⁸ Hub genes were defined as model-gene–encoded proteins with the highest node degree (connectivity) in the resulting network; among the seven model genes, the five with the greatest degree—*BRCAl*, *CXCL8*, *ELANE*, *GATA1*, and *FNDC5*—were selected as hubs. The two remaining model genes (*FKBP10* and *H1-5*) showed lower connectivity and were therefore not classified as hubs, although they remain in the prognostic signature. Raising the confidence threshold to 0.400 preserved the same set of hub genes.

The GeneMANIA database²⁹ was used for functional gene annotation and prioritization, employing integrated genomic and proteomic data to identify functionally related genes. For hub gene analysis, GeneMANIA generated interaction networks using weighted genomic datasets and similarity metrics, predicting gene functions from shared interaction partners and biological characteristics, thereby supporting construction of comprehensive PPI networks.

Hub Gene Differential Expression Validation and Receiver Operating Characteristic Curve Analysis

Differential expression of hub genes was evaluated across the TCGA-COADREAD, GSE20916, and GSE73360 datasets by comparing CRC and normal tissue samples. Diagnostic potential was assessed using receiver operating characteristic (ROC) curve analysis via pROC³⁰ (v1.18.5). Area under the curve (AUC)

values were calculated to quantify discriminative performance for CRC detection. AUC values range from 0.5 to 1.0, where values close to 1.0 indicate strong diagnostic power, 0.5–0.7 reflects low accuracy, 0.7–0.9 indicates moderate accuracy, and >0.9 shows high accuracy.

Statistical Analysis

All analyses were performed in R (v4.2.2). For two-group comparisons of continuous variables, Student's t-test was used for parametric data and the Wilcoxon rank-sum test was applied for nonparametric data. Multigroup comparisons were performed using the Kruskal–Wallis test. Spearman's correlation was applied to quantify associations between molecular variables. Unless otherwise specified, all tests were two-tailed, with $p < 0.05$ considered statistically significant. In particular, the differential-expression validation of hub genes between CRC ($n = 601$) and normal ($n = 48$) tissues, which exhibits a large and unbalanced group-size disparity, was analyzed exclusively with the nonparametric Wilcoxon rank-sum test rather than the t-test. Parametric t-tests were reserved for balanced comparisons with verified normality; all other two-group comparisons used the Wilcoxon rank-sum test.

RESULTS

Analysis Workflow

Figure 1 illustrates the analytical workflow for P&GMRDEGs.

Correction of CRC Datasets

Using limma, GSE20916 and GSE73360 datasets were independently annotated, standardized, and normalized to generate adjusted GEO datasets. Distribution boxplots (Figure 2A and 2B) were generated to visualize expression value differences in GSE20916 before and after correction. Equivalent distribution boxplots were also generated for GSE73360 (Figure 2C and 2D). Boxplot analysis confirmed uniform expression distributions across normalized CRC datasets (GSE20916 and GSE73360), indicating successful data correction and standardization.

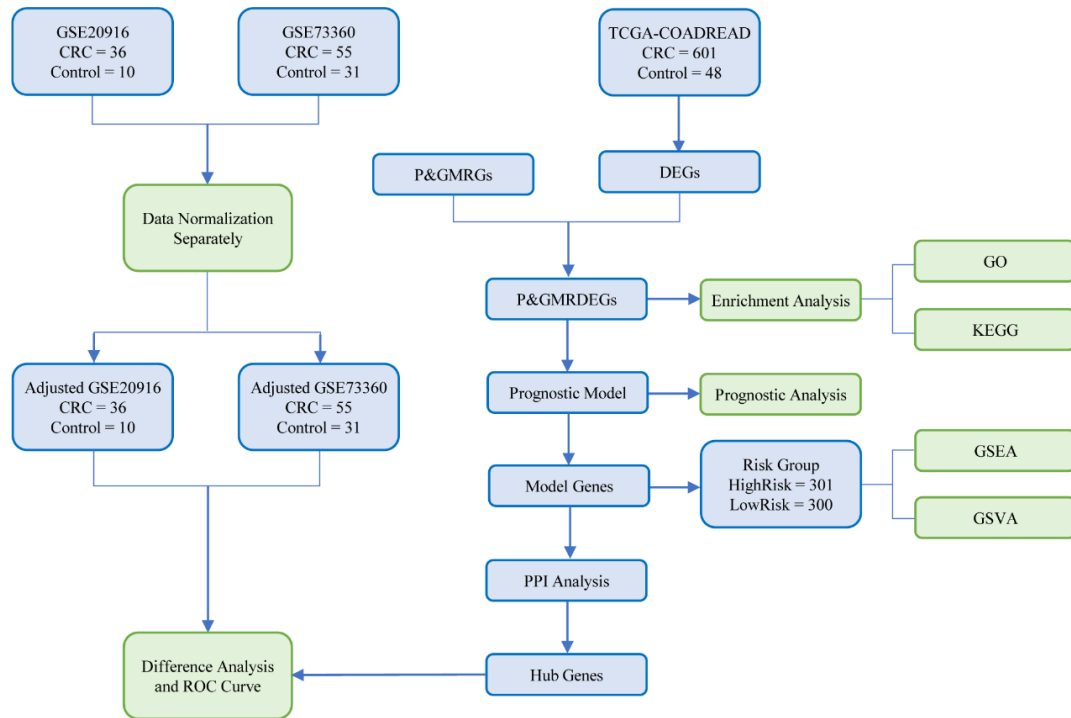


Figure 1. Workflow Chart of P&GMRDEG Analysis. CRC, colorectal cancer; TCGA; The Cancer Genome Atlas; COADREAD: colon adenocarcinoma, rectal adenocarcinoma; P&GMRG: pyroptosis and glucose metabolism-related gene; DEG, differentially expressed gene; P&GMRDEG, pyroptosis and glucose metabolism-related differentially expressed gene; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; GSEA, gene set enrichment analysis; GSVA, gene set variation analysis; PPI, protein-protein interaction; ROC, receiver operating characteristic.

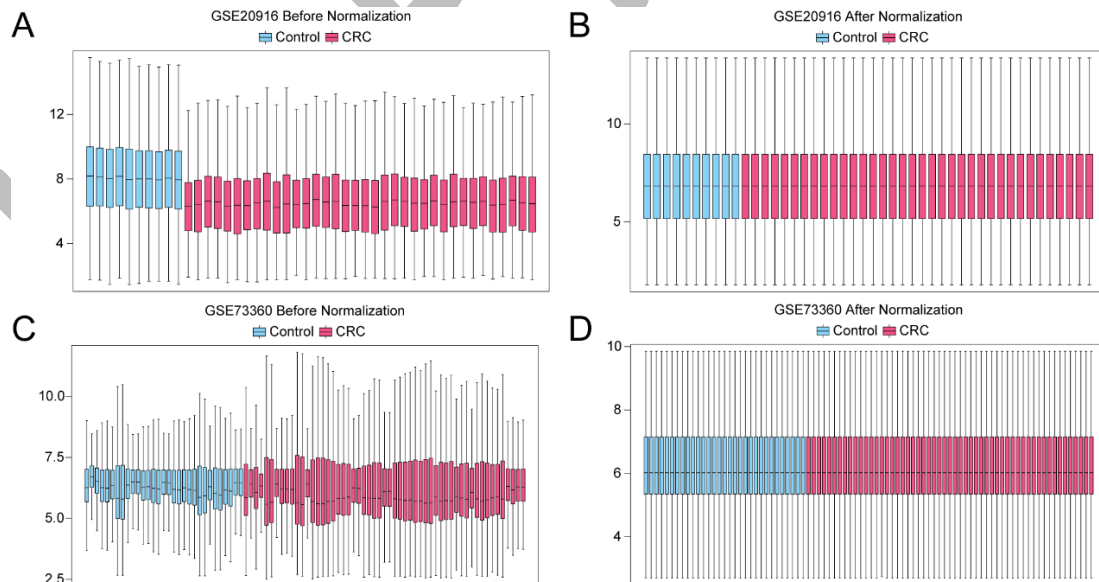


Figure 2. Normalization of Datasets GSE20916 and GSE73360. A. Boxplot of GSE20916 distribution before correction. B. Boxplot of corrected GSE20916 distribution. C. Boxplot of GSE73360 distribution before correction. D. Boxplot of corrected GSE73360 distribution. CRC, colorectal cancer. Pink represents CRC samples; blue represents control samples.

P&GMRDEGs in CRC

Transcriptomic profiling of the TCGA-COADREAD cohort revealed 10 271 DEGs meeting the criteria $|\log FC| > 1$ and adjusted $p < 0.05$, including 6790 upregulated and 3481 downregulated transcripts. Volcano plot visualization (Figure 3A) highlighted distinct expression patterns between CRC and normal tissue samples.

To identify P&GMRDEGs, genes meeting $|\log FC| > 1$ and adjusted $p < 0.05$ were intersected with P&GMRGs, as illustrated in a Venn diagram (Figure 3B). The analysis yielded 56 P&GMRDEGs, including

GZMB, GZMA, ELANE, CASP7, H2AX, CTSG, VEGFA, TREM1, FOXP3, FLNA, MMP9, BRCA1, CD55, GSTP1, GZMA, SERPINH1, HP, LCN2, RASGRF1, S100A9, FKBP10, PVALB, H1-5, PCSK9, IL17A, VDR, PINK1, CEBPB, SIGLEC15, P2RX7, FPR2, GABARAP, FASN, CXCL8, MST1, CITED2, TLR3, HSP90AB1, EZH2, DNMT3B, H4C1, H4C3, H3C13, H4C5, H4C2, H4C4, H2BC15, H2BC17, DRD2, OSM, GATA1, CAMP, FNDC5, IL1A, IL1B, and NOD2. Hierarchical clustering of the top 20 most dysregulated genes by $|\log FC|$ was visualized through heatmap analysis (Figure 3C).

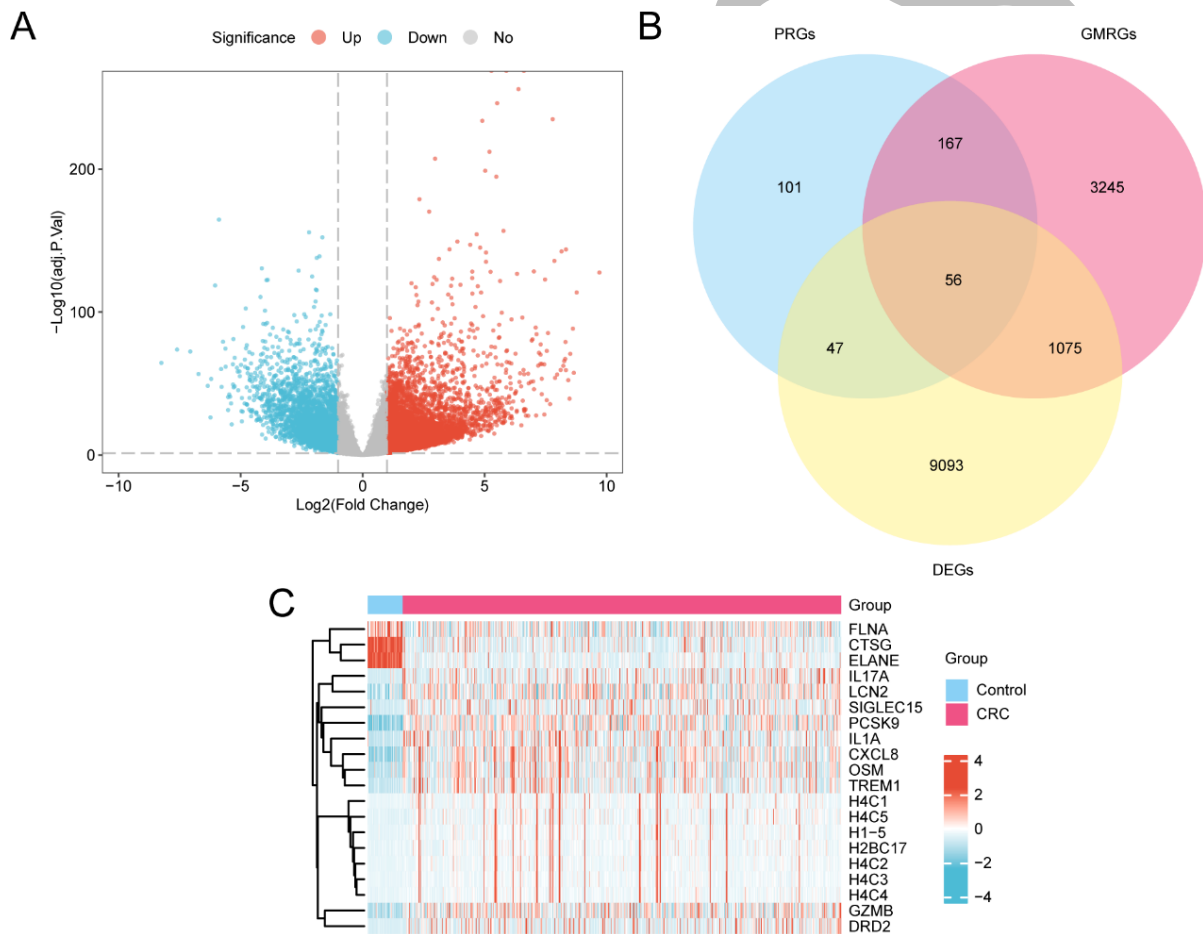


Figure 3. Differential Gene Expression Analysis. A. Volcano plot of DEGs between CRC and control samples in TCGA-COADREAD. B. Venn diagram of DEGs and P&GMRGs in TCGA-COADREAD. C. Heatmap of P&GMRDEG expression in TCGA-COADREAD. CRC, colorectal cancer; DEG, differentially expressed gene; PRG, pyroptosis-related gene; GMRG, glucose metabolism-related gene; P&GMRDEG, pyroptosis and glucose metabolism-related differentially expressed gene; TCGA, The Cancer Genome Atlas; COADREAD, colon adenocarcinoma, rectal adenocarcinoma. Pink represents CRC; blue represents controls. Heatmap: red indicates high expression; blue indicates low expression.

GO and KEGG Pathway Enrichment Analyses

Functional enrichment of the 56 P&GMRDEGs via GO and KEGG pathway analyses revealed significant enrichment in BP, CC, and MF GO terms as well as signaling pathways associated with CRC. As shown in Table 2, these genes were predominantly enriched in lipopolysaccharide response and related bacterial defense pathways. Enriched BP terms included response to molecule of bacterial origin, defense response to bacterium, myeloid cell differentiation, and nucleosome assembly; enriched CC terms included nucleosome, CENP-A-containing nucleosome, CENP-A-containing chromatin, chromosome, centromeric core domain, and secretory granule lumen; enriched MF terms included structural constituent of chromatin, protein heterodimerization activity, serine-type endopeptidase activity, serine-type peptidase activity, and serine hydrolase activity. Additional KEGG pathway enrichment was observed for neutrophil extracellular trap formation, systemic lupus erythematosus, IL-17 signaling, alcoholism, and inflammatory bowel disease. These results were visualized in a bubble plot (Figure 4A). Network diagrams were generated to illustrate relationships among enriched BP, CC, and MF GO terms as well as KEGG pathways (Figure 4B–E). We note that the keyword-based retrieval of pyroptosis-related genes inevitably captured inflammation- and angiogenesis-associated genes (eg, *VEGFA*, *MMP9*) that are not core pyroptosis executors such as *GSDMD* or *CASP1*. This breadth reflects the intrinsically inflammatory nature of pyroptosis and was intentionally retained so that the enrichment analysis would capture the full pyroptosis-associated transcriptional program rather than only executor molecules; consequently, functional terms were enriched for inflammatory responses, consistent with pyroptosis biology, and we prioritized pyroptosis executors when defining hub genes. Collectively, the enriched programs form an integrated immune-metabolic network: IL-17 and IL-10 cytokine signaling establish the inflammatory tone, chemokine signaling (eg, *CXCL8*) recruits myeloid cells, and Wnt/ β -catenin and Notch signaling govern stemness and immune evasion; together they convert pyroptosis-associated inflammation into sustained metabolic remodeling and tumor progression, providing the biological framework that links our hub genes to CRC outcome.

At the mechanistic level, glucose metabolism and pyroptotic signaling are reciprocally coupled. Enhanced glycolysis supplies ATP and biosynthetic precursors that sustain inflammasome assembly and NF- κ B/STAT3 activation, thereby amplifying gasdermin D cleavage and the release of IL-1 β and IL-18 that characterize pyroptosis. Conversely, pyroptosis-derived IL-1 β /IL-18 and DAMPs reshape the colorectal tumor microenvironment by recruiting myeloid cells and reinforcing a glycolytic, inflammatory stromal state. Within CRC, this feed-forward loop links metabolic reprogramming to inflammatory cell death and helps explain why genes bridging both processes (eg, *CXCL8*, *ELANE*) are enriched in high-risk tumors.

Beyond pathway activity, the signature genes are positioned to shape immune-cell composition and the tumor microenvironment. *CXCL8* and *ELANE* are potent drivers of neutrophil and myeloid recruitment and sustain a chemokine-rich inflammatory milieu, whereas the cytokine programs highlighted above (IL-17, IL-10) modulate T-cell and myeloid function. Because these inflammatory circuits intersect with immune-checkpoint regulation, the signature may indirectly reflect checkpoint-expression states; future studies should explicitly correlate the risk groups with infiltrating immune subsets and PD-L1/PD-1 status to clarify the immunobiological basis of the prognostic signal.

CRC Prognosis Risk Model

To construct a CRC prognostic model, univariate Cox regression was applied to P&GMRDEGs in the TCGA-COADREAD dataset. Genes showing $p < 0.10$ were included in LASSO regression for feature selection. Subsequent multivariate Cox regression incorporated selected genes to predict survival outcomes. Visualization of univariate and multivariate Cox regression results is shown in forest plots (Figure 5A and 5B). LASSO coefficient trajectories were determined (Figure 5C) and the corresponding regression model was visualized (Figure 5D). The final multivariate Cox regression model included seven genes: *ELANE*, *BRCA1*, *FKBP10*, *H1-5*, *CXCL8*, *GATA1*, and *FNDC5*.

Pyroptotic-metabolic Signatures in Colorectal Cancer

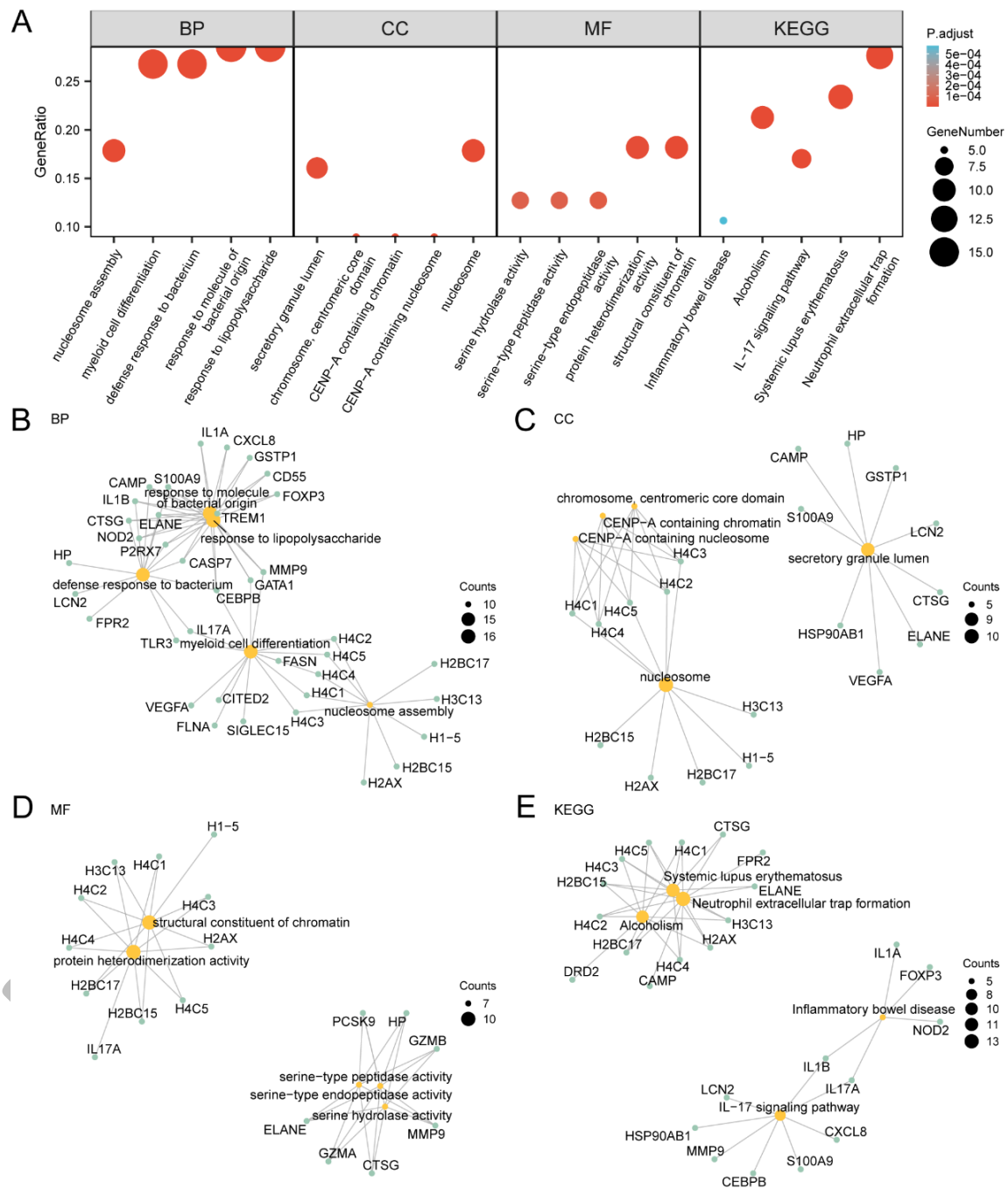


Figure 4. GO and KEGG Pathway Enrichment Analysis of P&GMRDEGs. A. Bubble chart of GO and KEGG pathway enrichment results. Abscissa displays GO and KEGG terms. B–E. Network diagrams showing GO and KEGG pathway enrichment results for P&GMRDEGs: BP, CC, MF, and KEGG pathways, respectively. Orange nodes represent terms; green nodes represent genes; connecting lines indicate gene–term relationships. P&GMRDEG, pyroptosis and glucose metabolism–related differentially expressed gene; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; BP, biological process; CC, cellular component; MF, molecular function. In the bubble chart, bubble size indicates the number of genes, and bubble color corresponds to the adjusted *p* value (darker red denotes smaller values, whereas blue denotes larger values). GO and KEGG pathway enrichment analyses were filtered using adjusted $p < 0.05$ and $q < 0.05$, with *p* values corrected using the Benjamini–Hochberg method.

Table 2. Results of GO and KEGG pathway enrichment analyses for P&GMRDEGs.

Ontology	ID	Description	Gene ratio	Background ratio	<i>p</i>	Adjusted <i>p</i>	<i>q</i>
BP	GO:0032496	Response to lipopolysaccharide	16/56	348/18870	2.71×10^{-15}	6.79×10^{-12}	4.03×10^{-12}
BP	GO:0002237	Response to molecule of bacterial origin	16/56	369/18870	6.77×10^{-15}	8.48×10^{-12}	5.03×10^{-12}
BP	GO:0042742	Defense response to bacterium	15/56	330/18870	2.71×10^{-14}	2.27×10^{-11}	1.34×10^{-11}
BP	GO:0030099	Myeloid cell differentiation	15/56	430/18870	1.26×10^{-12}	7.90×10^{-10}	4.69×10^{-10}
BP	GO:0006334	Nucleosome assembly	10/56	118/18870	1.73×10^{-12}	8.69×10^{-10}	5.16×10^{-10}
CC	GO:0000786	Nucleosome	10/56	134/19886	3.76×10^{-12}	6.31×10^{-10}	4.59×10^{-10}
CC	GO:0043505	CENP-A-containing nucleosome	5/56	18/19886	1.23×10^{-9}	6.88×10^{-8}	5.00×10^{-8}
CC	GO:0061638	CENP-A-containing chromatin	5/56	18/19886	1.23×10^{-9}	6.88×10^{-8}	5.00×10^{-8}
CC	GO:0034506	Chromosome, centromeric core domain	5/56	19/19886	1.66×10^{-9}	6.99×10^{-8}	5.08×10^{-8}
CC	GO:0034774	Secretory granule lumen	9/56	322/19886	2.66×10^{-7}	7.07×10^{-6}	5.14×10^{-6}
MF	GO:0030527	Structural constituent of chromatin	10/55	97/18496	2.35×10^{-13}	5.27×10^{-11}	3.92×10^{-11}
MF	GO:0046982	Protein heterodimerization activity	10/55	344/18496	6.05×10^{-8}	6.77×10^{-6}	5.03×10^{-6}
MF	GO:0004252	Serine-type endopeptidase activity	7/55	174/18496	8.02×10^{-7}	5.99×10^{-5}	4.44×10^{-5}
MF	GO:0008236	Serine-type peptidase activity	7/55	192/18496	1.55×10^{-6}	7.97×10^{-5}	5.92×10^{-5}
MF	GO:0017171	Serine hydrolase activity	7/55	196/18496	1.78×10^{-6}	7.97×10^{-5}	5.92×10^{-5}
KEGG	hsa04613	Neutrophil extracellular trap formation	13/47	192/8841	1.17×10^{-11}	1.53×10^{-9}	1.17×10^{-9}
KEGG	hsa05322	Systemic lupus erythematosus	11/47	139/8841	1.05×10^{-10}	6.80×10^{-9}	5.23×10^{-9}
KEGG	hsa04657	IL-17 signaling pathway	8/47	95/8841	2.94×10^{-8}	1.27×10^{-6}	9.73×10^{-7}
KEGG	hsa05034	Alcoholism	10/47	188/8841	3.89×10^{-8}	1.27×10^{-6}	9.73×10^{-7}
KEGG	hsa05321	Inflammatory bowel disease	5/47	65/8841	2.22×10^{-5}	5.78×10^{-4}	4.44×10^{-4}

BP: biological process; CC: cellular component; CENP-A: centromere protein A; GO: Gene Ontology; IL: interleukin; KEGG: Kyoto Encyclopedia of Genes and Genomes; MF: molecular function; P&GMRDEG: pyroptosis and glucose metabolism-related differentially expressed gene.

Pyroptotic-metabolic Signatures in Colorectal Cancer

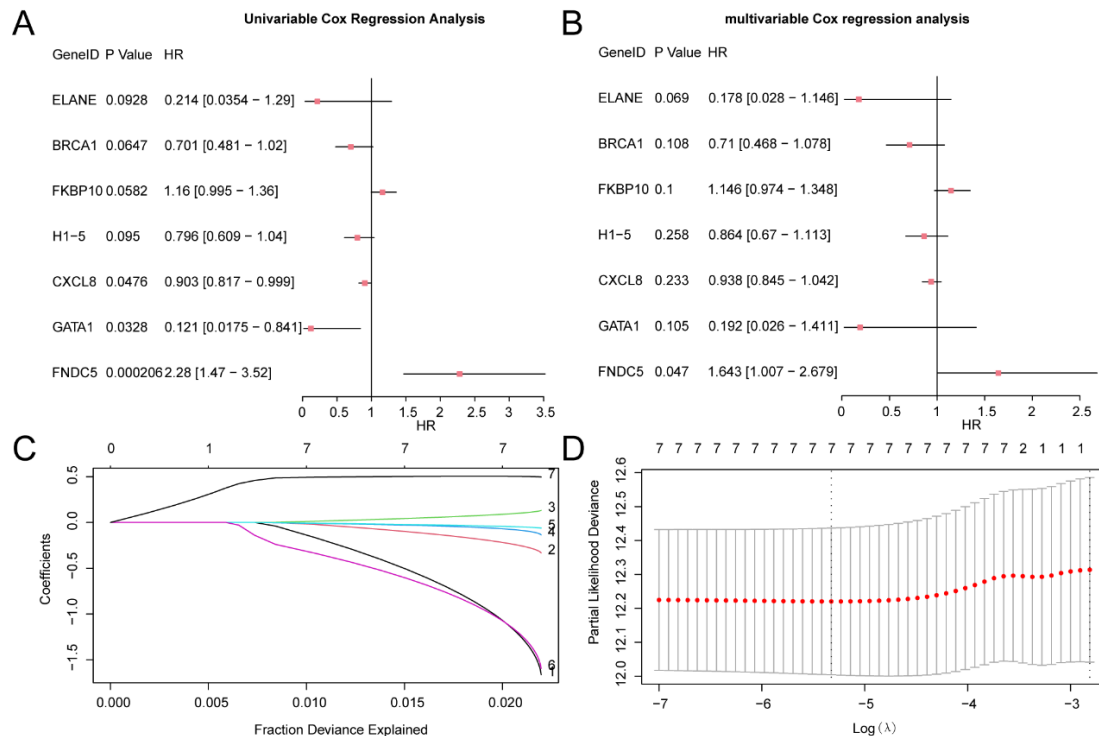


Figure 5. LASSO and Cox Regression Analyses. A. Forest plot of univariate Cox regression analysis. B. Forest plot of multivariate Cox regression analysis. C. LASSO regression variable trajectory plot. D. Prognostic risk model plot. LASSO, least absolute shrinkage and selection operator.

Validation of the CRC Prognostic Risk Model

To validate model accuracy, risk scores for each TCGA-COADREAD sample were calculated as follows:

$$\text{Risk score} = 1.7259 \times \text{ELANE} + 0.3427 \times \text{BRCA1} + 0.1362 \times \text{FKBP10} + 0.1464 \times \text{H1-5} + 0.0637 \times \text{CXCL8} + 0.4963 \times \text{GATA1} + 1.6521 \times \text{FNDC5}.$$

Based on the median risk score, patients were divided into high- and low-risk groups, and Kaplan–Meier survival analysis (Figure 6A) revealed significant differences in OS between these groups ($p < 0.001$). Risk score distribution, survival status, and model gene expression were visualized using ggplot2 (Figure 6B). To assess independent prognostic value, univariate and multivariate Cox regression analyses were conducted, integrating clinical parameters from the TCGA-COADREAD dataset. Variables with $p < 0.10$ in univariate analysis were included in multivariate modeling. Results are shown in forest plots (Figure 6C and 6D) and summarized in Table 3. Comparative expression analysis of the seven model genes (*ELANE*, *BRCA1*, *FKBP10*, *H1-5*, *CXCL8*, *GATA1*, and *FNDC5*) between the high- and low-risk subgroups revealed

significant differences ($p < 0.001$; Figure 6E). While this comparison is largely confirmatory of the model-gene distribution across risk groups, it validates that the signature genes are differentially expressed along the risk axis.

Model calibration at 1-year (Figure 6F), 3-year (Figure 6G), and 5-year (Figure 6H) intervals compared predicted and observed survival probabilities. Alignment with the ideal reference line indicated superior predictive accuracy, with optimal clinical performance observed at the 1-year timepoint.

GSEA

Risk-stratified differential expression analysis of TCGA-COADREAD samples identified 1715 significantly dysregulated genes ($|\log FC| > 1$ and adjusted $p < 0.05$), comprising 30 upregulated and 1685 downregulated transcripts. Volcano plot visualization (Figure 7A) highlighted distinct expression patterns between the high- and low-risk subgroups.

To assess the impact of gene expression on CRC phenotypes, GSEA was used to compare all genes between the high- and low-risk subgroups. Functional

enrichment analysis revealed significant activation in BP, CC, and MF, visualized via bubble plots (Figure 7B), with detailed results provided in Table 4. Key enriched pathways included PID Wnt signaling (Figure

7C), Wnt-mediated signaling (Figure 7D and 7E), and Notch signaling (Figure 7F and 7G), highlighting critical molecular mechanisms in CRC pathogenesis.

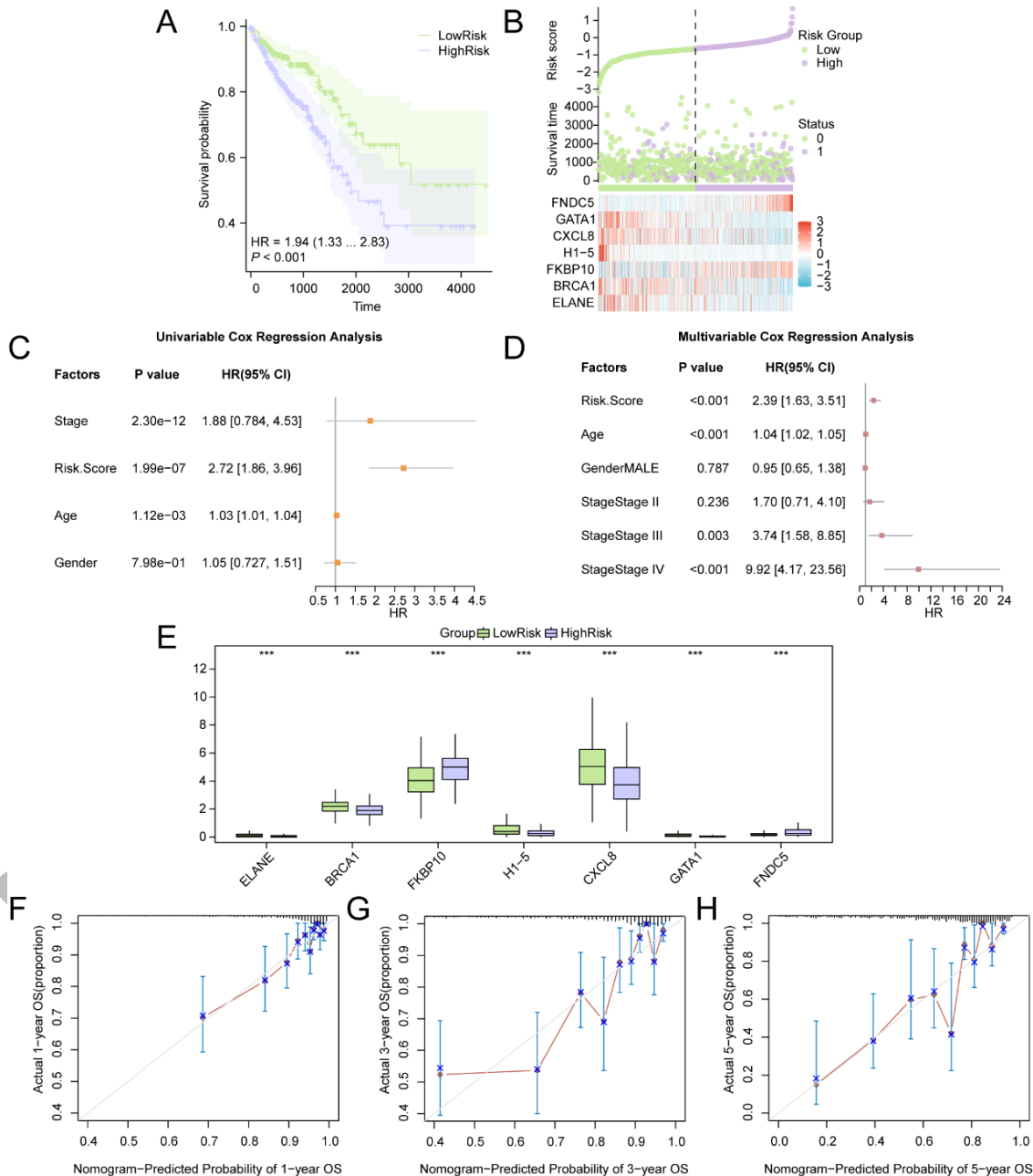


Figure 6. Validation of the Prognostic Model. A. Kaplan–Meier survival curves comparing high- and low-risk groups in TCGA-COADREAD. B. Risk factor plot of CRC samples from TCGA-COADREAD. C. Forest plot of risk scores and clinical variables in univariate Cox regression analysis. D. Forest plot of risk scores and clinical variables in multivariate Cox regression analysis. E. Comparison of model gene expression between high- and low-risk groups. F–H. Calibration curves for 1-year, 3-year, and 5-year survival predictions, respectively. CRC, colorectal cancer; TCGA, The Cancer Genome Atlas; COADREAD, colon adenocarcinoma, rectal adenocarcinoma; OS, overall survival. *** $p < 0.001$ (highly statistically significant).

Pyroptotic-metabolic Signatures in Colorectal Cancer

Table 3. Cox regression analysis results.

Characteristics	Total (N)	Univariate analysis		Multivariate analysis	
		<i>p</i>	Hazard ratio (95% CI)	<i>p</i>	Hazard ratio (95% CI)
Stage	601	2.30×10 ⁻¹²	1.88 [0.784, 4.53]	NA	NA
Stage II	226	NA	NA	0.236	1.70 [0.71, 4.10]
Stage III	179	NA	NA	0.003	3.74 [1.58, 8.85]
Stage IV	87	NA	NA	<0.001	9.92 [4.17, 23.56]
Risk score	601	1.99×10 ⁻⁰⁷	2.72 [1.86, 3.96]	<0.001	2.39 [1.63, 3.51]
Age	601	1.12×10 ⁻⁰³	1.03 [1.01, 1.04]	<0.001	1.04 [1.02, 1.05]
Gender	601	7.98×10 ⁻⁰¹	1.05 [0.727, 1.51]	NA	NA
Gender (male)	322	NA	NA	0.787	0.95 [0.65, 1.38]

CI: confidence interval; NA: not applicable.

Table 4. GSEA results for risk groups.

ID	Set size	Enrichment score	NES	<i>p</i>	Adjusted <i>p</i>	<i>q</i>
WP_VITAMIN_K_METABOLISM_AND_ACTIVATION_OF_DEPENDING_PROTEINS	12	0.854624	2.552657	2.08×10 ⁻⁰⁶	1.61×10 ⁻⁰⁴	1.57×10 ⁻⁰⁴
PAPASPYRIDONOS_UNSTABLE_ATEROSCLEROTIC_PLAQUE_DN	43	0.607170	2.534724	1.39×10 ⁻⁰⁷	1.13×10 ⁻⁰⁵	1.11×10 ⁻⁰⁵
NIKOLSKY_BREAST_CANCER_7P15_AMPLICON	11	0.866043	2.530258	6.39×10 ⁻⁰⁶	4.81×10 ⁻⁰⁴	4.70×10 ⁻⁰⁴
REACTOME_GAMMA_CARBOXYLATION_TRANSPORT_AND_AMINO_TERMINAL_CLEAVAGE_OF_PROTEINS	10	0.869203	2.378238	1.22×10 ⁻⁰⁵	8.69×10 ⁻⁰⁴	8.49×10 ⁻⁰⁴
ABE_INNER_EAR	41	0.585384	2.354010	1.91×10 ⁻⁰⁶	1.50×10 ⁻⁰⁴	1.47×10 ⁻⁰⁴
EBAUER_MYOGENIC_TARGETS_OF_PAX3_FOXO1_FUSION	44	0.546159	2.309547	2.39×10 ⁻⁰⁶	1.83×10 ⁻⁰⁴	1.78×10 ⁻⁰⁴
BIOCARTA_EXTRINSIC_PATHWAY	12	0.765233	2.285655	2.79×10 ⁻⁰⁴	1.51×10 ⁻⁰²	1.47×10 ⁻⁰²
REACTOME_DEFECTS_OF_CONTACT_ACTIVATION_SYSTEM_CAS_AND_KALLIKREIN_KININ_SYSTEM_KKS	14	0.693999	2.218728	8.21×10 ⁻⁰⁴	3.64×10 ⁻⁰²	3.56×10 ⁻⁰²
NIKOLSKY_BREAST_CANCER_20Q11_AMPLICON	31	0.622012	2.206805	1.22×10 ⁻⁰⁵	8.69×10 ⁻⁰⁴	8.49×10 ⁻⁰⁴
PID_WNT_SIGNALING_PATHWAY	28	0.599853	2.201400	9.35×10 ⁻⁰⁵	5.58×10 ⁻⁰³	5.45×10 ⁻⁰³

Table 4. Continued...

ID	Set size	Enrichment score	NES	<i>p</i>	Adjusted <i>p</i>	<i>q</i>
REACTOME_SIGNALING_BY_WNT	309	0.515397	1.283511	4.99×10^{-04}	2.42×10^{-02}	2.37×10^{-02}
REACTOME_TCF_DEPENDENT_SIGNALING_IN_RESPONSE_TO_WNT	212	0.610967	1.506275	4.67×10^{-08}	3.98×10^{-06}	3.88×10^{-06}
REACTOME_SIGNALING_BY_NOTCH	219	0.620642	1.532778	4.66×10^{-09}	4.40×10^{-07}	4.30×10^{-07}
REACTOME_PRE_NOTCH_EXPRESSION_AND_PROCESSING	94	0.788543	1.902855	1.00×10^{-10}	1.08×10^{-08}	1.06×10^{-08}
REACTOME_CHROMATIN_MODIFICATIONS_DURING_THE_MATERNAL_TO_ZYGOTIC_TRANSITION_MZT	54	0.852359	2.005814	1.00×10^{-10}	1.08×10^{-08}	1.06×10^{-08}
REACTOME_HDACS_DEACETYLATE_HISTONES	78	0.836053	2.006402	1.00×10^{-10}	1.08×10^{-08}	1.06×10^{-08}
REACTOME_ACTIVATED_PKN1_STIMULATES_TRANSCRIPTION_OF_AR_ANDROGEN_RECEPTOR_REGULATED_GENES_KLK2_AND_KLK3	51	0.856980	2.011313	1.00×10^{-10}	1.08×10^{-08}	1.06×10^{-08}
REACTOME_DNA_METHYLATION	48	0.863054	2.018193	1.00×10^{-10}	1.08×10^{-08}	1.06×10^{-08}
REACTOME_SIRT1_NEGATIVELY_REGULATES_RRNA_EXPRESSION	52	0.862070	2.025982	1.00×10^{-10}	1.08×10^{-08}	1.06×10^{-08}
REACTOME_ERCC6_CSB_AND_EHMT2_G9A_POSITIVELY_REGULATE_RRNA_EXPRESSION	60	0.858425	2.031650	1.00×10^{-10}	1.08×10^{-08}	1.06×10^{-08}

GSEA: gene set enrichment analysis; NES: normalized enrichment score.

Pyroptotic-metabolic Signatures in Colorectal Cancer

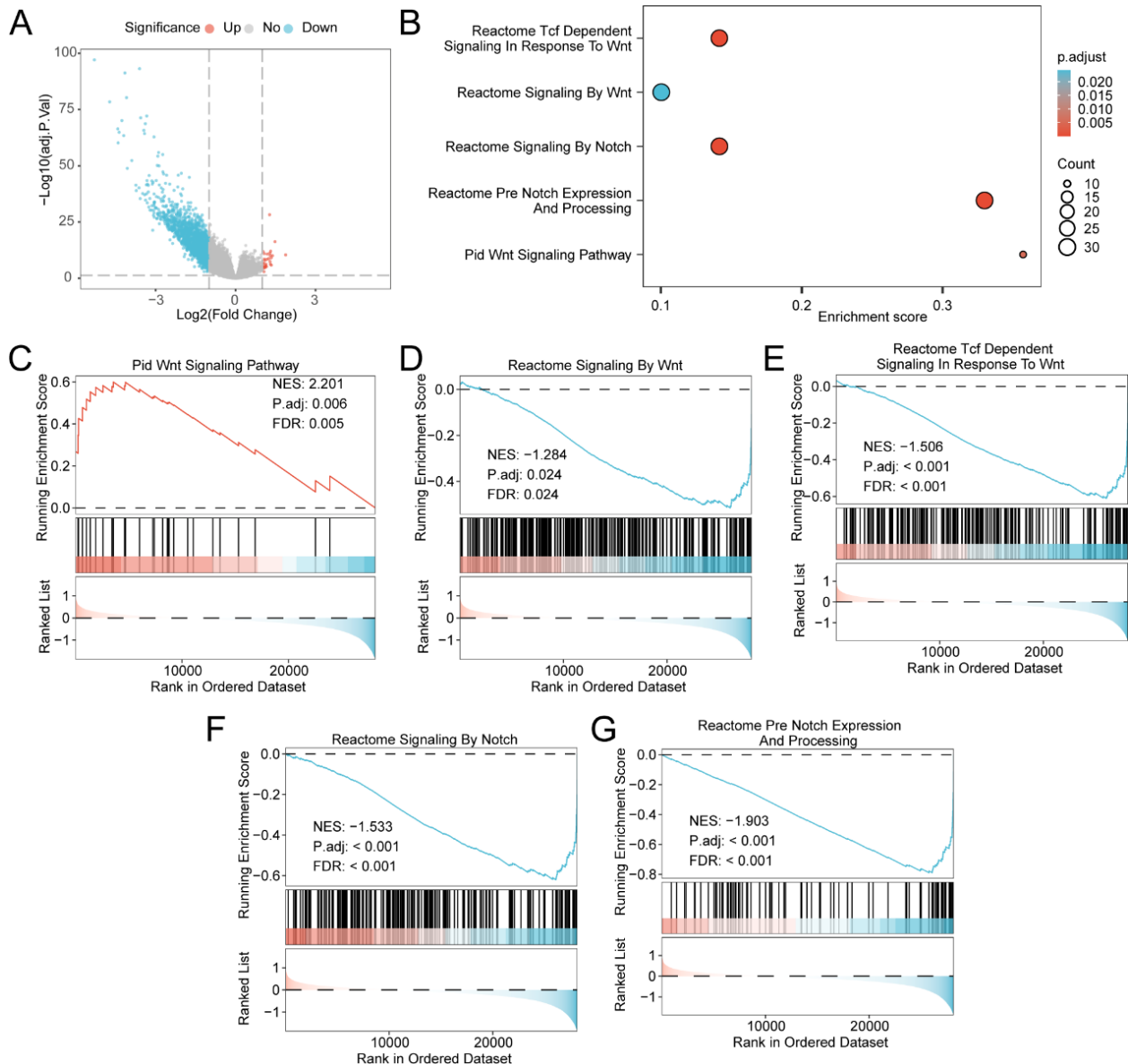


Figure 7. Differential Gene Expression Analysis and GSEA for Risk Groups. A. Volcano plot of DEGs between high- and low-risk groups in TCGA-COADREAD. B. Bubble chart of the top five enriched gene sets from GSEA. C–G. Enrichment analysis plots for the PID Wnt signaling pathway, Reactome signaling by Wnt, Reactome TCF-dependent signaling in response to Wnt, Reactome signaling by Notch, and Reactome pre-Notch expression and processing, respectively. CRC, colorectal cancer; TCGA, The Cancer Genome Atlas; COADREAD, colon adenocarcinoma, rectal adenocarcinoma; GSEA, gene set enrichment analysis. In the bubble plot, bubble size represents gene set size, and color represents adjusted *p* value size (red indicates smaller values; blue indicates larger values). Screening criteria for GSEA were adjusted *p*<0.05 and *q*<0.05, with *p* values corrected using the Benjamini–Hochberg method.

GSVA

GSVA was performed on TCGA-COADREAD samples using the c2.cp.v2023.2.Hs.symbols gene set to identify pathway enrichment differences between high- and low-risk subgroups (Table 5). The top 20 most dysregulated pathways (adjusted $p < 0.05$; ranked by $|\log FC|$) were visualized in a heatmap (Figure 8A), illustrating distinct pathway activation patterns between risk subgroups. Mann–Whitney U tests confirmed significant enrichment differences ($p < 0.05$) between these groups (Figure 8B). Key differentially activated pathways included cytokine-mediated inflammatory responses, IL-10 signaling, chemokine receptor interactions, and immune regulation, emphasizing distinct molecular profiles in high- versus low-risk patients. Notably, although GSEA identified PID_WNT_SIGNALING (adjusted $p = 5.58 \times 10^{-3}$) as enriched in the high-risk group, WNT/ β -catenin did not appear among the Top 20 pathways in the GSVA heatmap (Figure 8A). This discrepancy arises because GSEA evaluates rank-based enrichment across the full gene list, whereas GSVA estimates sample-level pathway activity; the WNT program is therefore captured robustly by GSEA but shows a more moderate, less extreme sample-level signal in GSVA. Both observations are consistent with a real but moderate WNT/ β -catenin involvement in high-risk CRC, and the adjusted p value is now stated explicitly.

PPI Network

PPI network analysis using STRING identified five hub genes (*BRCA1*, *CXCL8*, *ELANE*, *GATA1*, and *FNDC5*) among the seven model genes (Figure 9A). Functional gene interaction networks were further explored using GeneMANIA, which predicted interactions with 20 functionally similar proteins, illustrating shared expression patterns, protein domains, and biological functions (Figure 9B).

Differentially Expressed Hub Gene Validation and ROC Curve Analysis

Expression of the five hub genes in CRC (*BRCA1*, *CXCL8*, *ELANE*, *GATA1*, and *FNDC5*) was examined across the TCGA-COADREAD, GSE20916, and GSE73360 datasets. In TCGA-COADREAD, all five hub genes were significantly dysregulated ($p < 0.001$) between CRC and normal tissues (Figure 10A). In

GSE20916, two hub genes, *BRCA1* and *CXCL8*, showed significant differential expression ($p < 0.001$; Figure 10 B). In GSE73360, two hub genes, *BRCA1* and *ELANE*, were also significant ($p < 0.001$; Figure 10 C).

ROC curve analysis using pROC revealed that *BRCA1* and *ELANE* had superior diagnostic accuracy ($AUC > 0.9$) for CRC classification (Figure 10D and 10E). *CXCL8*, *GATA1*, and *FNDC5* exhibited moderate discriminative performance ($0.7 < AUC < 0.9$). In GSE20916, ROC curve analysis showed that *BRCA1* and *CXCL8* expression achieved $AUC > 0.9$ (Figure 10G), whereas in GSE73360, *BRCA1* showed $AUC > 0.9$ while *ELANE* and *GATA1* exhibited $0.7 < AUC < 0.9$ (Figure 10H and 10I). All group comparisons in Figure 10 were performed with the Wilcoxon rank-sum test given the unbalanced sample sizes.

Pyroptotic-metabolic Signatures in Colorectal Cancer

Table 5. GSVA results for risk groups.

ID	logFC	AveExpr	t	p	Adjusted p	B
BIOCARTA IL5 PATHWAY	0.471160	0.005417	7.372825	5.41×10^{-13}	5.38×10^{-10}	18.966492
KEGG MEDICUS REFERENCE RECRUITMENT AND FORMATION OF THE MCC	0.457864	0.000399	7.009018	6.32×10^{-12}	3.77×10^{-9}	16.615461
BIOCARTA CYTOKINE PATHWAY	0.434055	0.005650	7.984530	6.94×10^{-15}	2.07×10^{-11}	23.140702
WP CYTOKINES AND INFLAMMATORY RESPONSE	0.415935	0.006308	7.629019	9.02×10^{-14}	1.35×10^{-10}	20.681483
BIOCARTA STEM PATHWAY	0.399794	0.004775	6.816805	2.22×10^{-11}	1.11×10^{-8}	15.414056
WP IDO METABOLIC PATHWAY	0.392416	0.017537	7.079340	3.96×10^{-12}	2.96×10^{-9}	17.062088
WP COVID 19 ADVERSE OUTCOME PATHWAY	0.386274	0.009864	6.765478	3.10×10^{-11}	1.17×10^{-8}	15.098078
WP REGULATION OF SISTER CHROMATID SEPARATION AT THE METAPHASE ANAPHASE TRANSITION	0.378667	0.021599	6.470469	1.99×10^{-10}	4.25×10^{-8}	13.321892
BIOCARTA IL17 PATHWAY	0.369160	0.014703	6.051410	2.49×10^{-9}	3.24×10^{-7}	10.917816
WP GENES ASSOCIATED WITH THE DEVELOPMENT OF RHEUMATOID ARTHRITIS	0.361348	0.010819	6.476672	1.92×10^{-10}	4.25×10^{-8}	13.358533
WP IL 10 ANTI INFLAMMATORY SIGNALING PATHWAY	0.359135	0.017968	6.228362	8.73×10^{-10}	1.30×10^{-7}	11.915736
REACTOME INTERLEUKIN 10 SIGNALING	0.359084	0.002369	6.762942	3.15×10^{-11}	1.17×10^{-8}	15.082517
BIOCARTA INFLAM PATHWAY	0.355780	0.005691	6.427315	2.60×10^{-10}	5.18×10^{-8}	13.067825
BIOCARTA ASBCELL PATHWAY	0.353511	0.004581	5.080310	5.00×10^{-7}	1.34×10^{-5}	5.902064
WP CONTROL OF IMMUNE TOLERANCE BY VASOACTIVE INTESTINAL PEPTIDE	0.352602	0.006438	5.765020	1.29×10^{-8}	1.06×10^{-6}	9.356850
BIOCARTA GRANULOCYTES PATHWAY	0.342868	0.000237	5.658767	2.34×10^{-8}	1.55×10^{-6}	8.794926
REACTOME CHEMOKINE RECEPTORS BIND CHEMOKINES	0.334319	0.012784	6.697800	4.78×10^{-11}	1.58×10^{-8}	14.684565
BIOCARTA IL10 PATHWAY	0.331008	0.012352	5.814628	9.76×10^{-9}	8.83×10^{-7}	9.622408
KEGG MEDICUS REFERENCE ORGANIZATION OF THE INNER KINETOCHORE	0.327289	0.003800	5.575438	3.70×10^{-8}	2.15×10^{-6}	8.360809
BIOCARTA DC PATHWAY	0.326325	0.009856	5.416566	8.72×10^{-8}	3.94×10^{-6}	7.549260

AveExpr: average expression; B: log odds; COVID-19: coronavirus disease 2019; DC: dendritic cell; GSVA: gene set variation analysis; IDO: indoleamine 2,3-dioxygenase; IL: interleukin; KEGG: Kyoto Encyclopedia of Genes and Genomes; logFC: log2 fold change; MCC: mitotic checkpoint complex; WP: WikiPathways.

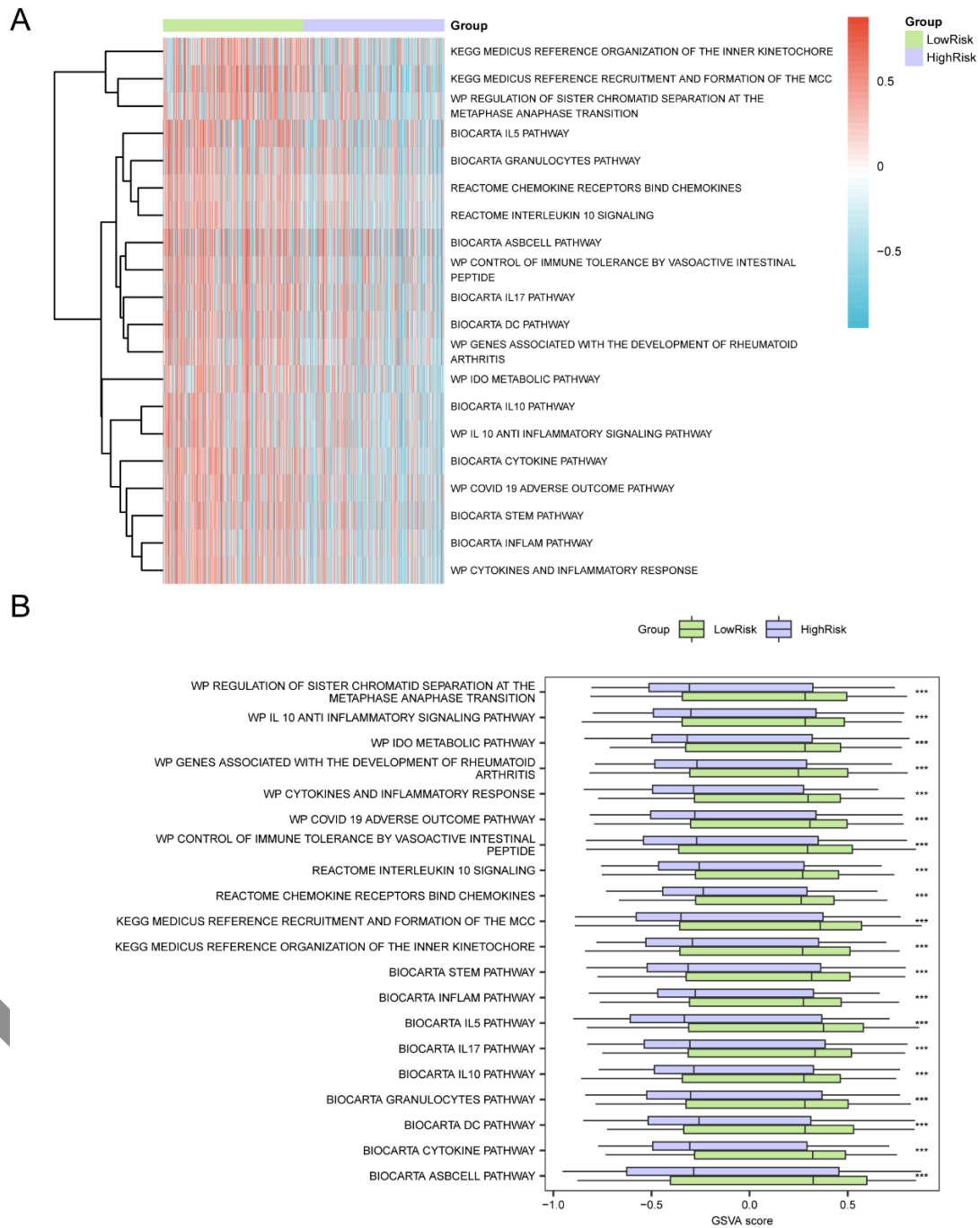


Figure 8. GSVA Analysis for Risk Groups. A. Heat map of GSVA results. B. Group comparison plot of GSVA results between high- and low-risk CRC groups in TCGA-COADREAD. CRC, colorectal cancer; TCGA, The Cancer Genome Atlas; COADREAD, colon adenocarcinoma, rectal adenocarcinoma; GSVA, gene set variation analysis. * $p < 0.001$ (highly statistically significant). Purple represents the high-risk group; green represents the low-risk group. GSVA set selection criteria were adjusted $p < 0.05$, with p value correction via the Benjamini–Hochberg method. In the heat map, blue represents low enrichment, and red represents high enrichment.**

Pyroptotic-metabolic Signatures in Colorectal Cancer

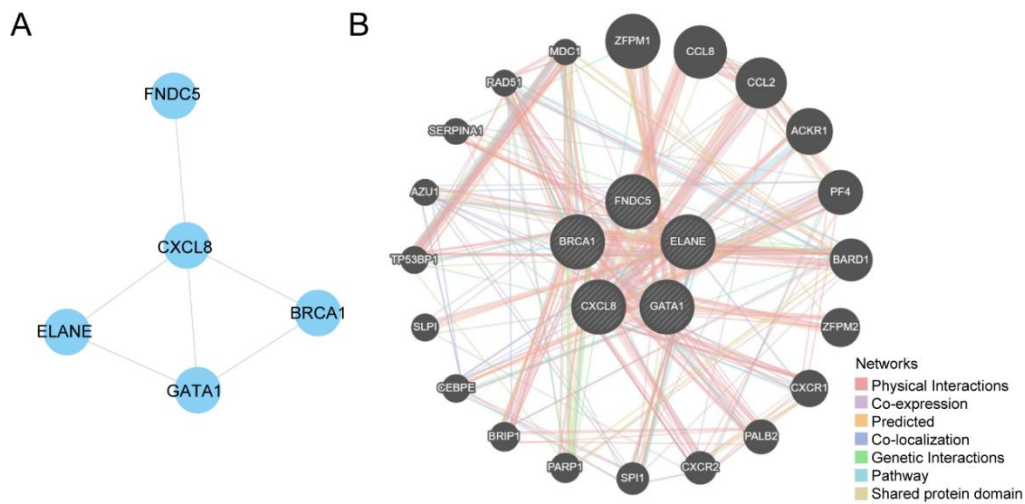
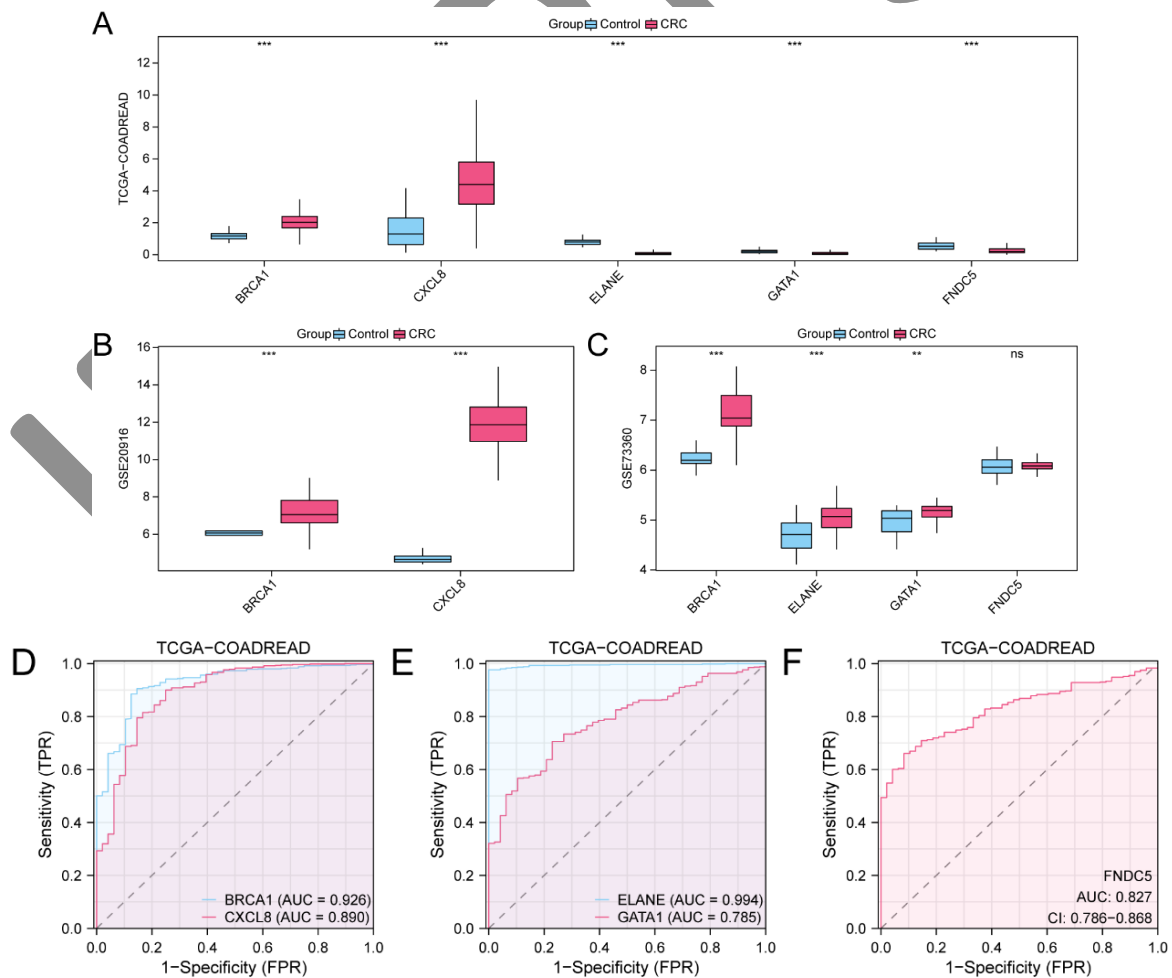


Figure 9. PPI network analysis. (A) PPI network of model genes constructed using the STRING database. (B) GeneMANIA prediction of the functional relationships among hub genes and functionally similar genes. Node colors correspond to functional categories. PPI, protein-protein interaction.



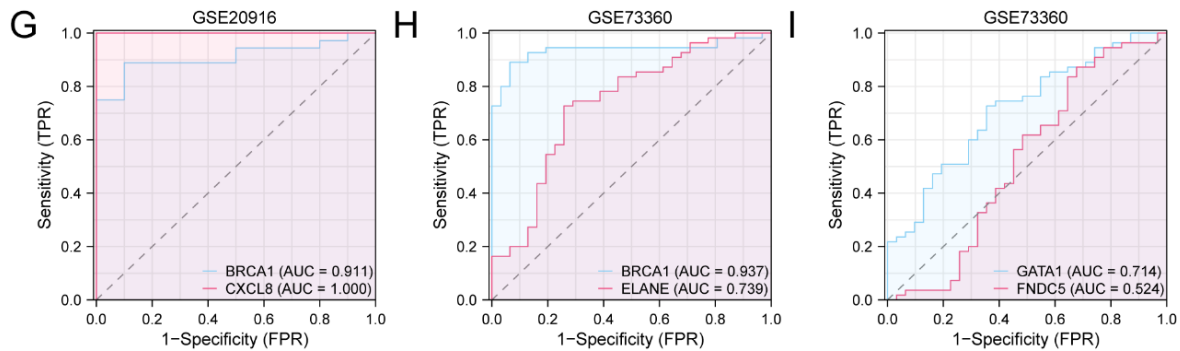


Figure 10. Differential Expression Validation and ROC Curve Analysis. A–C. Comparison of hub genes between CRC and control samples in TCGA-COADREAD, GSE20916, and GSE73360, respectively. D–F. ROC curves for *BRCA1* and *CXCL8*, *ELANE* and *GATA1*, and *FNDC5* in TCGA-COADREAD, respectively. G. ROC curves for *BRCA1* and *CXCL8* in GSE20916. H. ROC curves for *BRCA1* and *ELANE* in GSE73360. I. ROC curves for *GATA1* and *FNDC5* in GSE73360. NS, $p \geq 0.05$ (not statistically significant); ** $p < 0.01$ (statistically significant); *** $p < 0.001$ (highly statistically significant). AUC 0.5–0.7: low predictive accuracy; AUC 0.7–0.9: moderate accuracy; AUC > 0.9: high accuracy. CRC, colorectal cancer; TCGA, The Cancer Genome Atlas; COADREAD, colon adenocarcinoma, rectal adenocarcinoma; ROC, receiver operating characteristic; AUC, area under the curve; TPR, true positive rate; FPR, false positive rate. Blue represents controls, and pink represents CRC. For print and grayscale clarity, this figure can be split into differential-expression panels A–C and ROC panels D–I; each subfigure is mapped explicitly above.

DISCUSSION

CRC is recognized as one of the most prevalent malignancies globally, imposing substantial health risks and economic burdens on patients and society. Its high incidence and mortality rates necessitate a deeper understanding of underlying mechanisms and therapeutic targets. Current treatments, including surgery, chemotherapy, radiotherapy, and targeted therapies, are limited by high recurrence rates and major side effects, underscoring the need for novel approaches to improve patient outcomes.^{31,32}

Despite persistent challenges posed by molecular heterogeneity, emerging evidence highlights pyroptosis and metabolic reprogramming as critical drivers of CRC progression and therapeutic response. Although Li et al.³³ established prognostic relevance for PRGs, and Zhang et al.³⁴ demonstrated glucose metabolism's role in chemoresistance, the interplay between these processes remains inadequately characterized. No validated gene signature currently integrates both mechanisms for clinical prognostication. To address this gap, we developed a multi-omics-derived prognostic model capturing pyroptosis–glucose metabolism crosstalk, aiming to enhance risk stratification and support personalized treatment strategies in CRC.

Using an integrative computational framework, we systematically analyzed transcriptional alterations in pyroptotic regulators and metabolic components. Integration of GEO and TCGA datasets identified 56 pathway-associated genes with significant dysregulation. These markers exhibit dual functionality, contributing to both disease pathogenesis and clinical outcome prediction in CRC. The resulting risk model offers new avenues for CRC diagnosis and treatment.^{35,36}

Differential expression profiling of P&GMRGs in CRC revealed 56 candidate P&GMRDEGs, elucidating crosstalk between metabolic reprogramming and programmed cell death.⁴ These molecular targets show therapeutic potential, supported by prior studies documenting analogous pathway involvement in diverse malignancies.⁴ In particular, the genes *ELANE*, *BRCA1*, and *CXCL8* were identified as central regulators within the network. *ELANE*, encoding neutrophil elastase, mediates extracellular matrix degradation and inflammatory activation in the tumor microenvironment. *BRCA1*, beyond its canonical role in DNA repair, may participate in metabolic adaptation and influence glycolytic flux in CRC cells. *CXCL8*, encoding a proinflammatory chemokine, promotes angiogenesis and immune evasion, correlating with advanced disease stages. Less-characterized genes, including *FKBP10* and

Pyroptotic-metabolic Signatures in Colorectal Cancer

HI-5, were also dysregulated. *FKBP10*, encoding an endoplasmic reticulum resident chaperone, may support protein folding under metabolic stress. *HI-5*, encoding a histone H1 family member, may modulate chromatin organization and gene expression during pyroptosis. The prognostic signature derived from these genes aligns with evidence linking inflammatory cell death to energy metabolism in cancer.³⁷ Notably, incorporating multi-omics data enhances the model's biological plausibility and reproducibility. Although some observations align with previous studies, such as the established role of PRGs in CRC prognosis, the inclusion of glucose metabolism components introduces a novel perspective. To our knowledge, this study is among the first to connect pyroptotic machinery with metabolic reprogramming in CRC through a comprehensive bioinformatics approach, revealing previously unrecognized mechanistic interactions and providing a unified biomarker signature for risk stratification. Among the less-characterized members, *FNDC5* (encoding the irisin precursor) carried the largest positive coefficient (1.65) in the model, indicating that higher *FNDC5* expression is associated with an elevated risk score and worse prognosis in this cohort despite prior reports of irisin's metabolic protective effects; this context-dependent association may reflect irisin's pleiotropic roles in glucose handling and inflammation within the CRC microenvironment. *GATA1*, an erythroid/hematopoietic transcription factor, also contributed positively (coefficient 0.4963) and may modulate immune and stromal programs in CRC. For all seven genes the regression coefficients are positive, meaning that higher expression of each gene increments the risk score; the magnitude of each coefficient reflects that gene's independent contribution after LASSO-based adjustment for collinearity among the inflammatory genes. Regarding *ELANE*, although high expression conferred increased risk in our model (hazard ratio 1.72), we note that neutrophil elastase can also exert antitumor effects by cleaving *GSDMD* to induce pyroptosis and recruit antitumor immunity; its net association with adverse outcome therefore reflects a context-dependent balance between protumoral inflammation/ECM remodeling and pyroptosis-mediated tumor suppression. Regarding *CXCL8*, although its multivariate Cox coefficient (0.0637) is small—reflecting that its strong baseline expression is partly captured by correlated inflammatory genes selected through LASSO—it was retained because it

remained an independent predictor ($p < 0.05$ in multivariate analysis) and carries established biological relevance to CRC inflammation and angiogenesis; its modest weight therefore reflects shared signal rather than lack of importance.

GO and KEGG pathway analyses revealed significant enrichment of DEGs in immune and inflammatory response pathways, underscoring their critical roles in CRC biology. For instance, the IL-17 signaling pathway, implicated in tumor microenvironment remodeling, emerged as a key mediator, promoting proinflammatory cytokine and chemokine release, which facilitates neutrophil recruitment and activation, thereby fostering a chronic inflammatory state conducive to tumor progression and metastasis.⁶ Furthermore, synergistic pyroptosis–metabolic reprogramming interplay appears to be partly mediated through NF- κ B and STAT3 signaling, which not only amplify inflammatory responses but also regulate genes involved in cell survival, proliferation, and immune evasion. This duality highlights pyroptosis' paradoxical role: simultaneously inducing antitumor cell death and protumoral inflammation. It also emphasizes the complexity of targeting this pathway therapeutically. Such network-level insights align with established bioinformatic approaches that underscore the importance of pathway crosstalk in cancer.⁶ Specifically, the enrichment of glycolytic and amino acid metabolic pathways alongside immune activation suggests that metabolic rewiring supplies energy and biosynthetic precursors to sustain inflammation and tumor growth. This functional duality warrants further mechanistic investigation focusing on the contextual determinants of pyroptosis-related immune modulation, which will support CRC therapy development.

Integration of *ELANE*, *BRCA1*, *FKBP10*, *HI-5*, *CXCL8*, *GATA1*, and *FNDC5* into a prognostic stratification framework offers substantial potential for optimizing therapeutic decision-making. Empirical validation confirmed strong predictive performance for OS, highlighting its translational value in personalizing CRC management. This genetic signature–based prediction paradigm aligns with contemporary precision oncology initiatives, enabling treatment customization via genomic profiling.³⁸ Future research directions should focus on multiparameter model refinement and external validation across diverse demographic cohorts to establish broader clinical applicability. Clinically, this signature could complement conventional TNM staging

as a molecular risk layer to refine patient stratification and could guide enrollment in adjuvant or immunotherapy trials; however, its integration into routine workflows would first require prospective validation in independent, treatment-annotated cohorts and standardization of the expression-profiling platform.

In translational terms, the identified hub genes and the seven-gene signature hold promise for several applications. *ELANE*, *CXCL8*, and *FNDC5* could serve as candidate biomarkers for patient stratification and as targets for adjuvant or immunotherapy strategies, while the signature may help select patients likely to benefit from immune checkpoint blockade or metabolic interventions. Prospective studies linking the risk groups to treatment response will be essential to translate these findings into clinical practice.

Topological analysis of the PPI network provided insights into the cooperative mechanisms underlying CRC. Identification of key network hubs and associated signaling cascades offers a rational foundation for targeted therapeutic development. Pharmacological intervention at critical network nodes may enable synergistic treatment strategies capable of overcoming conventional drug resistance in CRC.³⁹

This investigation systematically delineated the functional landscape of pyroptotic regulators and metabolic pathway components in CRC through integrative bioinformatics.⁴ The findings demonstrate their dual involvement in disease progression and outcome prediction, corroborating pancancer studies highlighting the transcriptional interplay between programmed cell death and metabolic reprogramming.³⁷ These findings yield mechanistic insights and translational potential for CRC management; however, experimental validation is required to fully characterize the underlying molecular networks.

Several methodological constraints must be acknowledged. First, exclusive reliance on computational approaches, although informative, may not fully capture the biological complexity of in vivo systems. Second, sample size limitations could affect statistical validity and broader generalizability. Third, technical heterogeneity across datasets, including potential batch effects, may have introduced confounding factors in expression profiling. Finally, although the proposed prognostic framework shows promise, its clinical application will require rigorous external validation in larger patient cohorts to confirm predictive robustness and therapeutic relevance. In

addition, because feature selection (LASSO) and model construction were both performed in the same TCGA discovery cohort, a degree of overfitting cannot be excluded; we mitigated this by employing 10-fold cross-validation during LASSO regularization and by validating the signature in independent GEO cohorts (GSE20916, GSE73360), but further bootstrap validation and external prospective cohorts are required to fully establish robustness. Specific constraints also include heterogeneity across the TCGA and GEO datasets, the retrospective and de-identified nature of the public data with missing treatment and outcome details, potential batch effects despite limma correction, limited ethnic diversity of the included cohorts, and the absence of prospective clinical validation. These factors should be acknowledged when generalizing the findings.

In conclusion, this study integrated computational biology methodologies to delineate the functional interplay between programmed cell death regulators and metabolic pathway components in CRC. The proposed prognostic framework enhances our mechanistic understanding of CRC progression while offering clear translation potential. Future research should prioritize experimental validation and prospective multicenter trials to substantiate these findings and facilitate their integration into standardized oncology practice.

Acknowledgments

The authors would like to thank the public databases and research communities for providing the data and tools that made this study possible, including The Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO), and the developers of R/Bioconductor packages used in this analysis. We also acknowledge the support from our institutions and colleagues for their insightful discussions and technical assistance.

STATEMENT OF ETHICS

Not applicable. This study used only publicly available, de-identified data.

FUNDING

Not applicable.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

ACKNOWLEDGMENTS

Not applicable.

DATA AVAILABILITY

The datasets analyzed during this study are publicly available from the following sources:

The Cancer Genome Atlas Colorectal Adenocarcinoma (TCGA-COADREAD) data were accessed via the Genomic Data Commons (GDC) portal using TCGAAbiolinks.

Gene Expression Omnibus (GEO) datasets GSE20916 and GSE73360 were downloaded from the NCBI GEO repository. All processed data and code used in this study are available from the corresponding author upon reasonable request.

AI ASSISTANCE DISCLOSURE

Not applicable.

REFERENCES

- Morgan E, Arnold M, Gini A, Lorenzoni V, Cabaasag CJ, Laversanne M, et al, Global burden of colorectal cancer in 2020 and 2040: incidence and mortality estimates from GLOBOCAN. *Gut*. 2023;72(2):338-44.
- Biller LH, Schrag D. Diagnosis and treatment of metastatic colorectal cancer: a review. *JAMA*. 2021;325(7):669-85.
- Ciardiello F, Ciardiello D, Martini G, Napolitano S, Tabernero J, Cervantes A. Clinical management of metastatic colorectal cancer in the era of precision medicine. *CA Cancer J Clin*. 2022;72(4):372-401.
- Kumar P, Paul RK, Roy HS, Yeasin M, Ajit, Paul AK. Big data analysis in computational biology and bioinformatics. *Methods Mol Biol*. 2024;2719:181-97.
- Zhou P, Zhang S, Wang M, Zhou J. The induction mechanism of ferroptosis, necroptosis, and pyroptosis in inflammatory bowel disease, colorectal cancer, and intestinal injury. *Biomolecules*. 2023;13(5):820.
- Kusonmano K. Gene expression analysis through network biology: bioinformatics approaches. *Adv Biochem Eng Biot*. 2017;160(15-32).
- Tripepi G, Heinze G, Jager KJ, Stel VS, Dekker FW, Zoccali C. Risk prediction models. *Nephrol Dial Transplant*. 2013;28(8):1975-80.
- Nikolić J, Loncar-Turukalo T, Sladojević S, Marinković M, Janjić Z. Melanoma risk prediction models. *Vojnosanit Pregl*. 2014;71(8):757-66.
- Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, et al, TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res*. 2016;44(8):e71.
- Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, et al, Visualizing and interpreting cancer genomics data via the Xena platform. *Nat Biotechnol*. 2020;38(6):675-8.
- Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al, NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res*. 2013;41(Database issue):D991-5.
- Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and Bioconductor. *Bioinformatics*. 2007;23(14):1846-7.
- Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al, The GeneCards suite: from gene data mining to disease genome sequence analyses. *Curr Protoc Bioinformatics*. 2016;54:1-30.
- Deng M, Sun S, Zhao R, Guan R, Zhang Z, Li S, et al, The pyroptosis-related gene signature predicts prognosis and indicates immune activity in hepatocellular carcinoma. *Mol Med*. 2022;28(1):16.
- Zhang C, Wang M, Ji F, Peng Y, Wang B, Zhao J, et al, A novel glucose metabolism-related gene signature for overall survival prediction in patients with glioblastoma. *Biomed Res Int*. 2021;2021:8872977.
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al, Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
- Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res*. 2019;47(D1):D419-26.
- Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*. 2000;28(1):27-30.
- Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284-7.
- Deng X, Thompson JA. An R package for survival-based

- gene set enrichment analysis. *Research Square*. 2023;3-3367968.
22. Engebretsen S, Bohlin J. Statistical predictions with glmnet. *Clin Epigenetics*. 2019;11(1):123.
23. Rich JT, Neely JG, Paniello RC, Voelker CCJ, Nussenbaum B, Wang EW. A practical guide to understanding Kaplan–Meier curves. *Otolaryngol Head Neck Surg*. 2010;143(3):331-6.
24. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *P Natl Acad Sci Usa*. 2005;102(43):15545-50.
25. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013;14:7.
26. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. 2011;27(12):1739-40.
27. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al, STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019;47(D1):D607-13.
28. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al, Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498-504.
29. Franz M, Rodriguez H, Lopes C, Zuberi K, Montojo J, Bader GD, et al, GeneMANIA update 2018. *Nucleic Acids Res*. 2018;46(W1):W60-4.
30. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al, pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12:77.
31. Shen Z, Ye Y, Bin L, Yin M, Yang X, Jiang K, et al, Metabolic syndrome is an important factor for the evolution of prognosis of colorectal cancer: survival, recurrence, and liver metastasis. *Am J Surg*. 2010;200(1):59-63.
32. Farvid MS, Sidahmed E, Spence ND, Mante Angua K, Rosner BA, Barnett JB. Consumption of red meat and processed meat and cancer incidence: a systematic review and meta-analysis of prospective studies. *Eur J Epidemiol*. 2021;36(9):937-51.
33. Li R, Zhang S, Liu G. Identification and validation of a pyroptosis-related prognostic model for colorectal cancer. *Funct Integr Genomic*. 2022;23(1):21.
34. Zhang J, Zou S, Fang L. Metabolic reprogramming in colorectal cancer: regulatory networks and therapy. *Cell Biosci*. 2023;13(1):25.
35. Xu JZ, Xia QD, Sun JX, Liu CQ, Lu JL, Xu MY, et al, Establishment of a novel indicator of pyroptosis regulated gene transcription level and its application in pan-cancer. *Sci Rep*. 2023;13(1):17911.
36. Mima K, Cao Y, Chan AT, Qian ZR, Nowak JA, Masugi Y, et al, *Fusobacterium nucleatum* in colorectal carcinoma tissue according to tumor location. *Clin Transl Gastroen*. 2016;7(11):e200.
37. Wen W, Zhao Z, Zheng Z, Zhao S, Zhao H, Cheng X, et al, Rare variant association analyses reveal the significant contribution of carbohydrate metabolic disturbance in severe adolescent idiopathic scoliosis. *J Med Genet*. 2024;61(7):666-76.
38. Yao Y, Wang D, Zheng L, Zhao J, Tan M. Advances in prognostic models for osteosarcoma risk. *Heliyon*. 2024;10(7):e28493.
39. Malyarchuk BA. Genetic features of lipid and carbohydrate metabolism in Arctic peoples. *Biochemistry. Biokhimiia*. 2024;89(7):1192-201.