Int J Clin Exp Pathol 2022;15(4):168-182
www.ijcep.com /ISSN:1936-2625/IJCEP0141845

Original Article

A pyroptosis-related prognosis model to predict survival in colorectal cancer patients

Meiyan Chen[1*], Jingyi Zhang[1*], Xiaoyan Lin[2], Xiaosan Zhu[1#], Tao Xie[3#]

Departments of [1]Gastroenterology, [2]Dermatology, [3]Obstetrics and Gynecology, Chenggong Hospital, Xiamen University, Xiamen 361003, Fujian, China. [*]Co-first authors. [#]Equal contributors.

Received January 19, 2022; Accepted March 15, 2022; Epub April 15, 2022; Published April 30, 2022

Abstract: Pyroptosis is a recently-identified pathway of host cell death that is stimulated by a range of microbial infections. Emerging evidence indicates pyroptosis plays crucial roles in tumor growth, disease progression, and migration of different cancer cells. However, the clinical significance of pyroptosis in tumor behavior prognosis, as well as the underlying mechanism in different cancers remains elusive. Here, by evaluating the expression level of pyroptosis genes in colorectal cancer (CRC) patients from the TCGA cohort and GEO cohort (GSE39582), we identified pyroptosis-related DEGs and then built a 13-gene risk model by applying the LASSO Cox regression algorithm. Furthermore, functional analysis using GSEA and GSEV revealed that our prognostic model may function through regulating immune responses and tumor biogenesis pathways. Significant infiltration of activated immune cells (e.g. cytotoxic T cells) was observed in the low risk score group. The selected gene set was further validated in the GEO cohort. Time-dependent ROC curves confirmed that our risk score model is robust in predicting 1, 3 and 5-year overall survival in CRC patients. Overall, we have identified a pyroptosis-related gene signature that consists of 13 genes, which serves as a potent indicator of CRC prognosis. Thus, our model provides insights in how to make better clinical decision in the future.

Keywords: Pyroptosis, colorectal cancer, gene signature, prognosis

Introduction

Colorectal cancer (CRC) is ranked as the fifth most common cancer in both genders, representing 7.9% of new cancer incidence yearly in the US [1]. Colon adenocarcinoma (COAD) is the most dominant form of CRC, accounting for up to 53% of cases. COAD has the highest mortality rate of gastrointestinal cancers and the incidence of COAD has risen steadily, especially in developed countries [2, 3]. It was reported that ~295,000 new colorectal cancer cases are diagnosed yearly and ~134,000 deaths are recorded around the globe. Although the five-year survival has greatly improved recently, CRC remains the third leading cause of mortality in the United States. Localized CRC can be treated with surgery, radiation, and local ablation. Despite the advance of curative treatment, more than 30% patients eventually develop metastases, which requires more complicated therapy with high mortality [4, 5]. The key to successful treatment of CRC is to identify the disease at early stage. However, early clinical manifestations of CRC are diverse and may give rise to a range of non-specific and often misattributed symptoms [6-8]. Hence, there is an immediate need to develop better tools for risk assessment of CRC, to facilitate the management.

There are various proposed hallmarks of tumor cells that lead to growth and metastasis. Inhibited cell death is the most well recognized one [9]. Evidence suggests that malignant cells take advantage of distinct strategies to circumvent cell death pathways that are critical to normal development and physiological status [10]. Different types of cell death function as significant mechanisms for anti-cancer treatment [11, 12]. Recently, a new form of programmed cell death, pyroptosis, was identified, which is different from necrosis, apoptosis, necroptosis, or autophagy. It is believed that pyroptosis is

triggered by proinflammatory signals and occurs frequently upon infection by a range of intracellular pathogens [13].

Numerous studies have focused on investigating the molecular mechanism of pyroptosis regulation and induction; however, the relationship between pyroptosis and cancer remains largely elusive [14]. It was reported that pyroptosis is related to gastric cancer (GC). The expression of GSDMD is decreased in GC cells compared to adjacent non-cancerous cells. GSDMD is a major player in pyroptosis, which can be cleaved by inflammatory caspases upon external stimulation. The resulting N-terminal fragment of GSDMD forms membrane pores, activating the pyroptosis process [15]. Nevertheless, our understanding of pyroptosis in cancer is uncertain, especially whether pyroptosis affects the clinical outcome of cancer patients.

In the present work, we first identified the differential expression genes (DEGs) related to pyroptosis, then constructed and validated a prognostic model upon the expression of selected DEGs to successfully predict the survival of COAD patients. Then, we performed multiple functional analyses to identify the potential molecular mechanism of pyroptosis in COAD. We also explored the difference of immune responses in patients in both low and high-risk groups based on our model and examined the pyroptosis-related immunity and frontline treatment of COAD. The study should provide about the function of pyroptosis in COAD and help discover novel metrics of risk assessment for COAD patients.

Material and methods

Data collection

The University of California Santa Cruz (UCSC) Xena portal (https://xenab rowser.net/) was used to download the clinical records and RNA-seq data of the TCGA COAD patients. The RNA-seq data of the GEO cohort (GSE39582) were acquired from GEO portal (https://www.ncbi.nlm.nih.gov/geo/). For the TCGA cohort, the gene expression profiles from 41 normal tissues and 458 tumor tissues were included for analysis. The FPKM (Fragments Per Kilobase Million) from the downloaded gene expression files was converted to TPM (Transcripts Per Kilobase Million) for downstream analysis. For GSE39582, the microarray data from 19 normal tissues and 566 colon cancer tissues were included for model validation.

Kaplan-Meier analysis

Survival analysis was conducted on TCGA COAD patients with complete clinical records. The patients were categorized based on consensus analysis or risk scores and subjected to Kaplan-Meier estimator analysis using Survival package in R. Log rank test was utilized to define statistical significance. The survival curves were generated using Survminer in R.

DEG analysis

Identification of differential expression genes was performed using limma R package. The cut-off used to identify significant DEGs was fold change > |±1.5| and FDR (adjusted *p*-values) < 0.05. Volcano plot was depicted for DEG visualization.

GSVA and GSEA

Gene set variation and gene set enrichment analysis were performed respectively to identify the significant pathways and signatures related to pyroptosis. MSigDB gene sets were used for both methods [16]. GSEA analysis was carried out using Cluster Profiler in R. Significant pathways were defined if |NES| (normalized enrichment score) > 1, NOM P < 0.05 and FDR < 0.05 with 1000 permutations. GSVA analysis was carried out using GSVA in R. A heatmap was created using pheatmap in R.

Construction and validation of the risk score model

The TCGA cohort and GSE39582 dataset were used for model construction and validation, respectively. A LASSO regression analysis was performed using glmnet package in R. The final model was determined based on the optimal lambda generated from 1000-fold cross-validation. Univariate Cox regression analysis was conducted to identify gene features with significant prognostic value. Feature selection was achieved by recursive feature elimination (RFE) with random forest as classifier and 10-fold cross-validation to avoid model overfitting. The risk score was computed as the linear combination of the gene expression of each feature multiplied by the LASSO coefficient:

Risk score = -0.0375*MMP3-0.1895*TNIP3+0.1378*SLC2A3-0.1375*MMP10+0.0295*SERPINE1+0.2242*NECAB3-0.3972*RAB3B+0.0458*CD720.0716*FAS+0.1111*ADAM8+0.0341*LZTS3+0.2799*ANGPTL4-0.1612*RN7SL3.

To validate the risk score model, each patient from GSE39582 was classified into low or high-risk groups based on the risk score calculated from the prediction model. The median risk score was set as the cut-off threshold. Kaplan-Meier curves were generated to estimate overall survival. Using pROC package in R, receiver operating characteristic (ROC) curves were created and the area under the curve (AUC) was computed to predict the survival rate up to five years. The correlation between clinical factors and the risk score was evaluated to assess performance of the risk score model.

Immune score analysis

The CIBERSORT, MCPcounter, single-sample gene set enrichment analysis (ssGSEA), and TIMER algorithms were used to determine the immune cell components or immune responses between low and high-risk groups based on the pyroptosis-related gene signature [17-20]. The results of immune response generated from the four algorithms were visualized by heatmap. The correlation between the clinical factors and the computed immune scores was shown in the same heatmap.

Statistical analysis

Unpaired *t* test was applied to determine significant differences between 2 clusters after passing the normality Shapiro-Wilk test. Wilcoxon test was used to compare high and low risk groups due to the lack of a normal distribution. All the other tests were two-sided. All analysis was conducted using R3.6.1. *P* values < 0.05 were considered significant.

Results

Classification of COAD based on pyroptosis-related genes

Pyroptosis is an inflammatory-associated cell death, that has been widely implicated in various cancer types. For example, pyroptosis could promote inflammatory cell death, while impeding the undesired proliferation and invasion of malignant cells [21]. To understand the connection between pyroptosis and COAD, we collected a total of 37 pyroptosis-related genes from the current literature and performed a consensus clustering analysis with the expression profile of these genes from all 440 COAD patients in the TCGA cohort. When tuning the clustering variable (k) (ranging from 2 to 10), we uncovered that the optimal cluster number is two, which has the highest cophenetic correlation coefficient (**Figure 1A** and **1B**). Further PCA analysis confirmed that the COAD patients could be suitably divided into two clusters based on pyroptosis-related genes (**Figure 1C**). We then started to evaluate whether the expression of pyroptosis genes are correlated with any clinicopathologic factors besides the consensus clusters. The expression level of 25 pyroptosis genes were significantly differentiated between cluster 1 and cluster 2 (**Figure 1D**). In contrast, there was no strong correlation observed between these genes and seven other clinical characteristics (age, gender, and various metrics in cancer TNM staging). Additionally, Kaplan-Meier survival analysis showed that COAD patients in cluster 1 and cluster 2 hadve significantly different survival probabilities (**Figure 1E**). Altogether, our results suggest that pyroptosis was strongly correlated with COAD and served as an independent indicator of COAD prognosis.

Identification of pyroptosis-related DEGs

To study whether pyroptosis is related to COAD prognosis, we first identified the differentially expressed genes (DEGs) between the two pyroptosis clusters in the TCGA cohort by performing DEG analysis. DEGs were considered significant if adjusted *P* was less than 0.05 and fold change of expression level was larger than 1.5. A total of 1217 DEGs, with 1012 up-regulated and 105 down-regulated genes were identified between cluster 1 and cluster 2 (**Figure 2A**). Next, we explored the biologic processes and pathways associated with the pyroptosis-related DEGs by performing GO enrichment and KEGG pathway analyses (**Figure 2B**). More than 15 pathways were found to be significantly enriched in cluster 2 of COAD. The top gene sets are involved in regulation of immune and inflammatory responses, such as cytokine receptor binding, T cell activation, and rheumatoid arthritis.

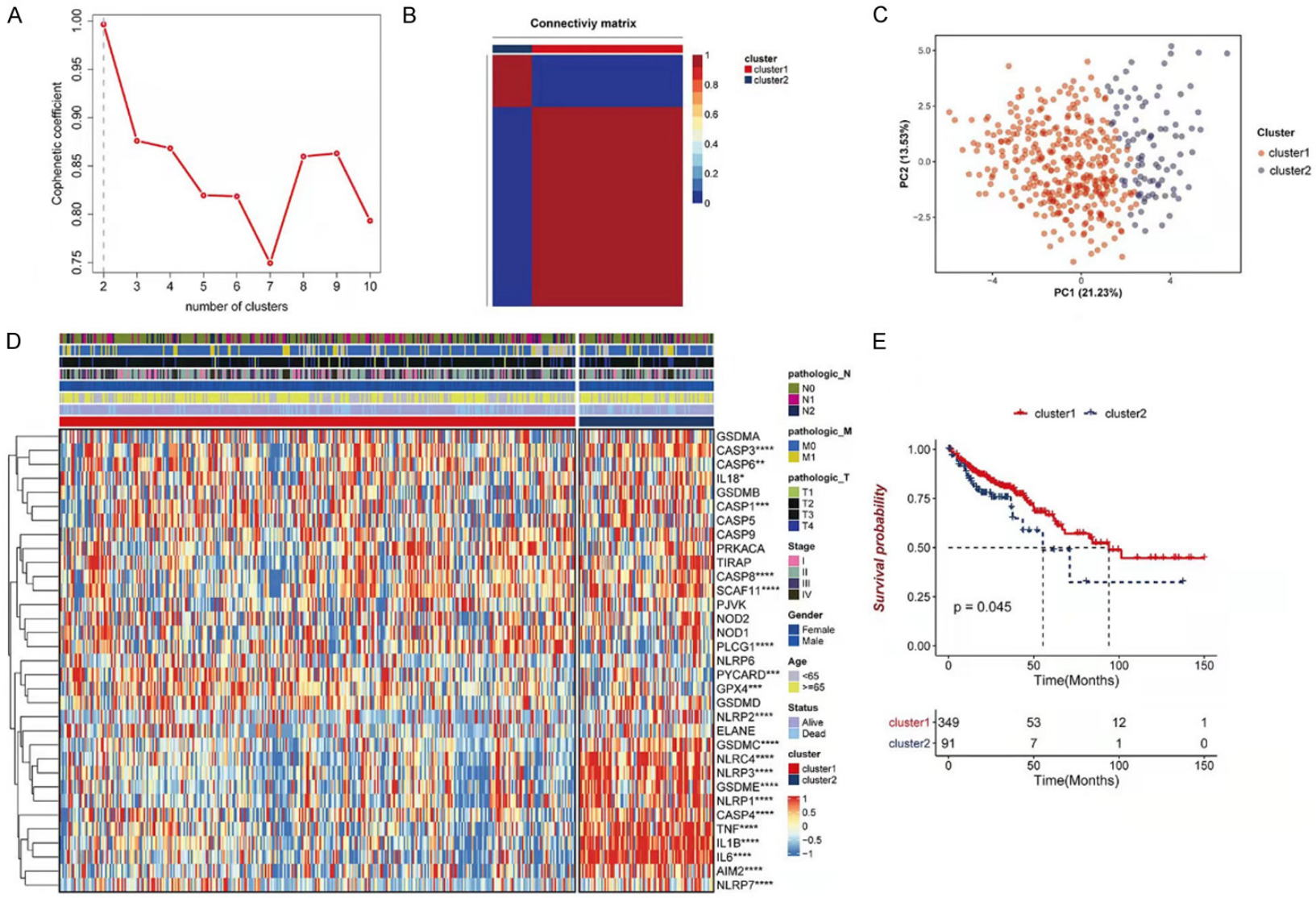


Figure 1. Colorectal cancer (CRC) patients were clustered into two subgroups based on the expression level of pyroptosis-related genes. Nonnegative matrix factorization (NMF) consensus analysis of the TCGA cohort revealed a good consensus for k = 2. A. Comparison of cophenetic coefficients among k clusters. B. Consensus matrix for the expression of all pyroptosis-related genes across all TCGA COAD samples at k = 2. C. PCA analysis showing the separation of cluster 1 (red) and cluster 2 (blue) based on the first two principal components. D. Heatmap showing the correlation between clinicopathologic factors and the pyroptosis-related subtypes. E. Kaplan-Meier analysis of overall survival of COAD patients in Cluster 1 (red) and Cluster 2 (blue) stratified by the expression of pyroptosis-related genes.

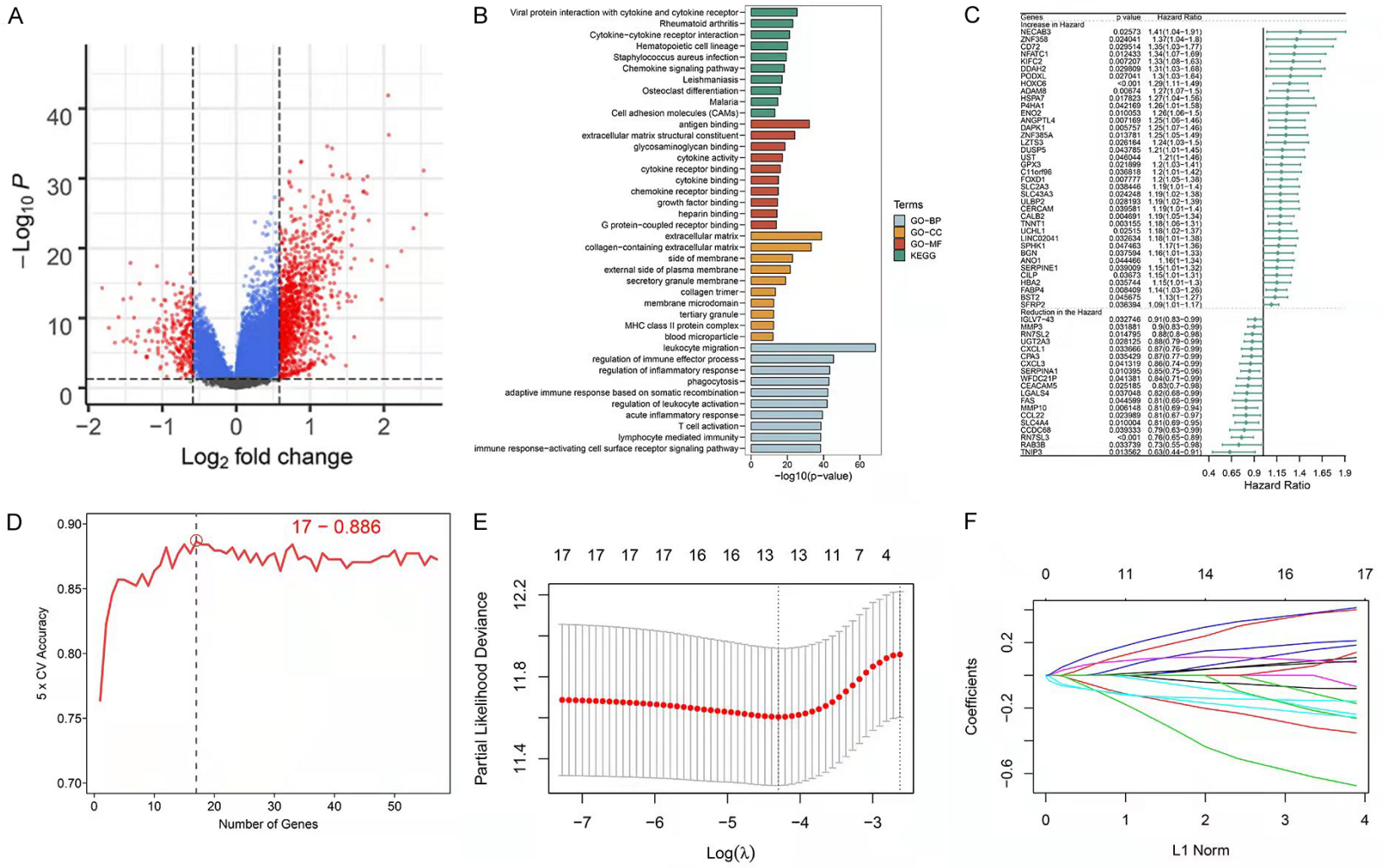


Figure 2. Construction of the prognostic risk model based on pyroptosis-related gene signatures. A. Volcano plot showing the DEGs between cluster 1 and cluster 2 in TCGA cohort. Black and blue dots represent the genes without significant changes, while red dots outside the left and right dashed lines donate the genes with decreased or increased expression, respectively. B. Bar plots showing the significant enriched pathways based on all DEGs. GO: Gene Ontology; BP: Biological Processes; CC: Cellular Component; MF: Molecular Function. C. Forest plot showing the 57 prognostic gene variables selected from univariate Cox analysis. D. Random forest was applied to select the optimal number of variables for the predictive model. Seventeen genes were chosen with the corresponding highest accuracy (0.886) after 5-fold cross-validation. E. Additional variable selection using LASSO and the lambda parameter was tuned with 1000-fold cross validation. The final gene signature was chosen with the lowest partial likelihood deviation computed against the log (lambda). F. Coefficient profiles of the 17 gene features from the univariate Cox regression model.

Construction and validation of risk score model

To explore the predictive power of these DEGs in overall survival, we performed univariate Cox regression analysis on the TCGA COAD cohort. Fifty-seven DEGs were significantly correlated with the overall survival of patients (**Figure 2C**). Particularly, 38 gene features displayed a rising risk of poor COAD prognosis (hazard ratio > 1). On the other hand, we found that the gene expression levels of the remaining 19 DEGs were associated with improved risk of COAD prognosis. Random forest was applied to select the optimal number of variables for the predictive model. Seventeen genes were chosen with the corresponding highest accuracy (0.886) after 5-fold cross-validation (**Figure 2D**). To reduce the risk of overfitting, we applied and conducted LASSO regression to select the most representative gene features in our prognostic model. After feature selection, 13 out of the 57 DEGs were retained for prognosis prediction, which showed non-zero coefficients (**Figure 2E** and **2F**). Hence, the risk score was calculated as the linear combination of the gene expression levels of 13 variables weighted by the relative coefficient. Among the 13 gene variables, six were low-risk indicators, while the rest five had positive coefficients, indicating unfavorable outcome.

The risk scores of each COAD patient were derived from the above model, and the 440 COAD patients were then classified into two groups (low-risk or high-risk) based on the median risk score (**Figure 3A**). Notably, patients in the high-risk group had more deaths and a shorter recurrence-free survival time than those in the low-risk group. Kaplan-Meier survival analysis further revealed that the overall survival time was significantly shorter in the patient group with high risk score, compared to the low-risk score group (**Figure 3B**). To assess the robustness of the prognosis model, we conducted a time-dependent ROC curve analysis. The result indicates that the 13-gene classifier was a powerful prognostic factor to predict COAD patient survival at one, three- and five-year time windows, the AUCs of which were calculated as 0.746, 0.729 and 0.802, respectively (**Figure 3C**). We further evaluated the prognostic power of our risk score model on an independent dataset (GSE39582) by carrying out the Kaplan-Meier analysis (**Figure 3D**). We found that the lower the risk score was, the better the patient survival outcome, which confirmed that risk score is valuable for COAD prognosis.

To further characterize the 13 predictive DEGs in our model, we depicted them along the chromosome by RCircos (**Figure 3E**). They are not evenly distributed along the genome; for instance, none are located in chr2 and chr3, although they are among the largest chromosomes. To examine the correlation of risk score and the clinicopathologic factors, the gene expression profile of the 13 predictive DEGs and the clinical features were presented in the heatmap. We found that all the factors were diversely distributed across the whole spectrum of risk score ranges and a similar result was observed on the original 33 DEGs identified before feature selection (**Figure 3F**).

Infiltrating immune cells are an integral component of the tumor microenvironment and are reported to play key roles in invasion and metastasis of cancer and related drug resistance. To determine cell-mediated immunity in COAD, we applied four different algorithms to quantify cell fractions from gene expression profiles based on our prognosis model. The heatmap of immune responses based on CIBERSORT, MCP counter, ssGSEA, and TIMER algorithms is shown in **Figure 4**. Based on the enrichment scores of the more than 16 types of immune cells, we found that both adaptive and innate immune responses were significantly different between COAD patients from the two different risk groups.

Confirmation of independent prognostic value of the risk model

We studied the correlation between risk score and other potential confounding factors (clinicopathologic features, such as age, gender, and various pathologic staging metrics) to evaluate the prognostic power of our model. The univariate Cox regression analysis proved that the risk score could independently predict the survival of COAD patients (HR = 3.89, 95% CI: 2.76-5.5), like other tumor staging factors. Additional multivariate Cox analysis confirmed that the risk score was an independent prognostic factor affecting survival in the TCGA cohort of COAD patients (HR = 3.23, 95% CI: 2.17-4.8) (**Figure 5A**). Altogether, our data revealed that prognosis was significantly relat-

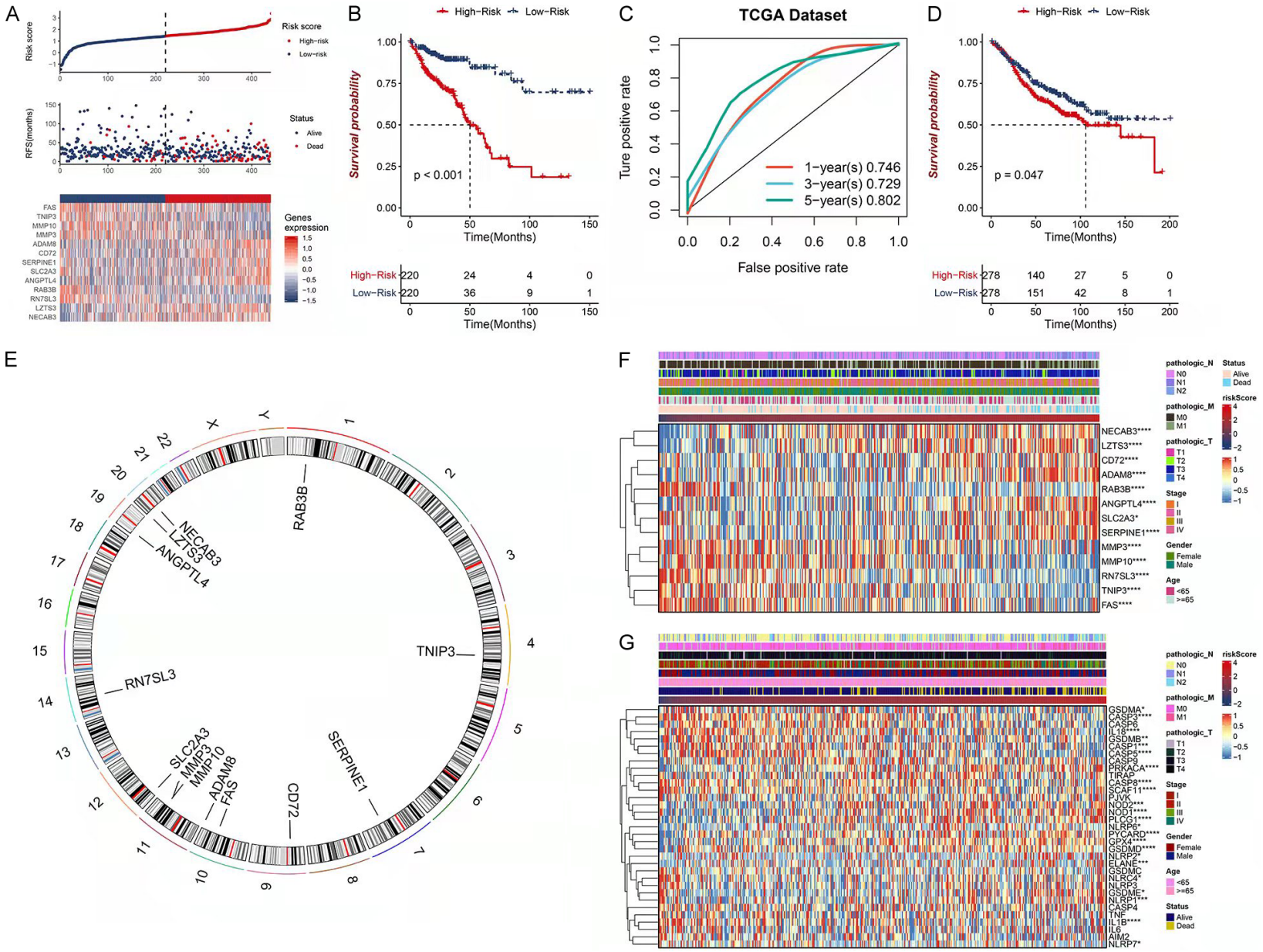

A
Risk score
High-risk
Low-risk
RFS(months)
Status
Alive
Dead
FAS
TNIP3
MMP10
MMP3
ADAM8
CD72
SERPINE1
SLC2A3
ANGPTL4
RAB3B
RN7SL3
LZTS3
NECAB3
Genes expression
B
High−Risk
Low−Risk
Survival probability
p < 0.001
Time(Months)
High−Risk 220 24 4 0
Low−Risk 220 36 9 1
C
TCGA Dataset
Ture positive rate
False positive rate
1−year(s) 0.746
3−year(s) 0.729
5−year(s) 0.802
D
High−Risk
Low−Risk
Survival probability
p = 0.047
Time(Months)
High−Risk 278 140 27 5 0
Low−Risk 278 151 42 8 1
E
RAB3B
NECAB3
LZTS3
ANGPTL4
TNIP3
RN7SL3
SLC2A3
MMP3
MMP10
ADAM8
FAS
CD72
SERPINE1
F
NECAB3****
LZTS3****
CD72****
ADAM8****
RAB3B****
ANGPTL4****
SLC2A3*
SERPINE1****
MMP3****
MMP10****
RN7SL3****
TNIP3****
FAS****
pathologic_N
N0
N1
N2
pathologic_M
M0
M1
pathologic_T
T1
T2
T3
T4
Stage
I
II
III
IV
Gender
Female
Male
Age
<65
>=65
Status
Alive
Dead
riskScore
G
GSDMA*
CASP3****
CASP6
IL18****
GSDMB**
CASP1***
CASP5****
CASP9
PRKACA****
TIRAP
CASP8****
SCAF11****
PJVK
NOD2***
NOD1****
PLCG1****
NLRP6*
PYCARD****
GPX4****
GSDMD****
NLRP2*
ELANE***
GSDMC
NLRC4*
NLRP3
GSDME*
NLRP1***
CASP4
TNF
IL1B****
IL6
AIM2
NLRP7*

Figure 3. Validation of the risk score model. (A) Scatter plot showing the distribution of the risk scores (top) and recurrence free survival time (bottom) of patients in the TCGA cohort. The cut-off value for low and high-risk group classification was indicated by the dash lines. (B, D) Kaplan-Meier analysis of overall survival of patients in the TCGA (B) and GSE39582 (D) cohort stratified by high or low risk scores. (C) Time-ROC curve analysis of the 13-gene signature in the TCGA dataset in 1, 3, 5-year. (E) Circular visualization of chromosomal positions of the 13 predictive DEGs. (F, G) Heatmap showing the expression of the 13 gene features or 33 pyroptosis genes associated with the clinicopathologic factors ranked by their risk score in the TCGA cohort.

ed to cancer stage, including pathologic M ($P < 0.001$), pathologic N ($P < 0.001$), and risk score ($P < 0.001$) derived from our prognostic model.

Somatic mutations intensely contribute to cancer development and generation of novel tumor epitopes. Given that fact that mutational burden could be an explanation for clinical outcome, such as overall survival, we performed somatic mutation analysis on the COAD patients stratified by risk score computed by our prognosis model. Distinct mutation patterns were observed in the samples from both groups (**Figure 5B**). For the high-risk group, the top 5 mutated genes were: APC, TP53, TTN, KRAS and SYNE1. However, the most frequently altered genome regions for low risk group were APC, TTN, TP52, KRAS and MUC16. The total tumor mutation burden was significantly different between the two groups (**Figure 5C**).

Functional analysis based on the risk model

To further explore the molecular functions of the selected gene signature in disease progression, GSEA and GoTerm analysis was conducted to identify the signaling pathways associated with the survival rate. Among all the gene sets we tested in MSigDB database, the most significant enriched pathways ($|NES| > 1$, NOM P-value < 0.05 and FDR < 0.05) were associated with various types of immune response (**Figure 6**). The top hits were toll like receptor, Rig I like receptor, NK cell mediated cytotoxicity, T cell mediated immunity, and regulation of immune effector response. Most of these pathways are associated with various types of immune response. In addition, pathways involved in colorectal cancer were also significantly enriched in this analysis. Taken together, our discovery suggests that apyroptosis-related prognostic signature is involved in regulating immune and tumor biogenesis pathways.

Prediction of chemotherapy response based on the risk model

Nowadays chemotherapy is still the frontline treatment for various cancer types. To explore the response of COAD patients with different risk scores, we predicted the IC50 of nine most frequently used chemo drugs in clinical practice for every included patient from the TCGA cohort. As a result, a total of 6 drugs demonstrated significantly lower IC50 in the low-risk group, indicating these patients were more sensitive to these chemotherapies and could respond better compared to the high-risk group (**Figure 7**).

Discussion

Pyroptosis is a newly-appreciated type of programmed cell death, morphologically and mechanistically distinct from apoptosis or necrosis. It features the involvement of Caspas-1 and the downstream response of rapid plasma-membrane rupture and release of proinflammatory intracellular contents [22, 23]. Although initially thought to be associated with inflammatory diseases, emerging evidence have revealed a complex relationship between pyroptosis and cancer. Pyroptosis induces inflammation-mediated cell death in malignant cells, resulting in inhibitory of cancer growth and migration. Scientists reported the abnormal expression of some pyroptotic inflammasomes found in tumor cells. For instance, Gao *et al* recently revealed that downregulation of GSDMD attenuates tumor proliferation through impaired EGFR/Akt signaling and is a promising estimator of prognosis in non-small cell lung cancer [24]. Moreover, pyroptosis provides a novel direction for alternative cancer therapy as it participates in the pathogenesis of tumor [25]. Despite the large number of studies on the molecular regulators and underlying mechanism of pyroptosis in cancer development, the clinical significance of pyroptosis for disease prognosis remained unclear.

In this study, we focused on the prognostic value of pyroptosis in colorectal carcinoma, which is the third most prevalent cancer around the globe. In contrast to breast cancer, which is frequently diagnosed at an early stage and is

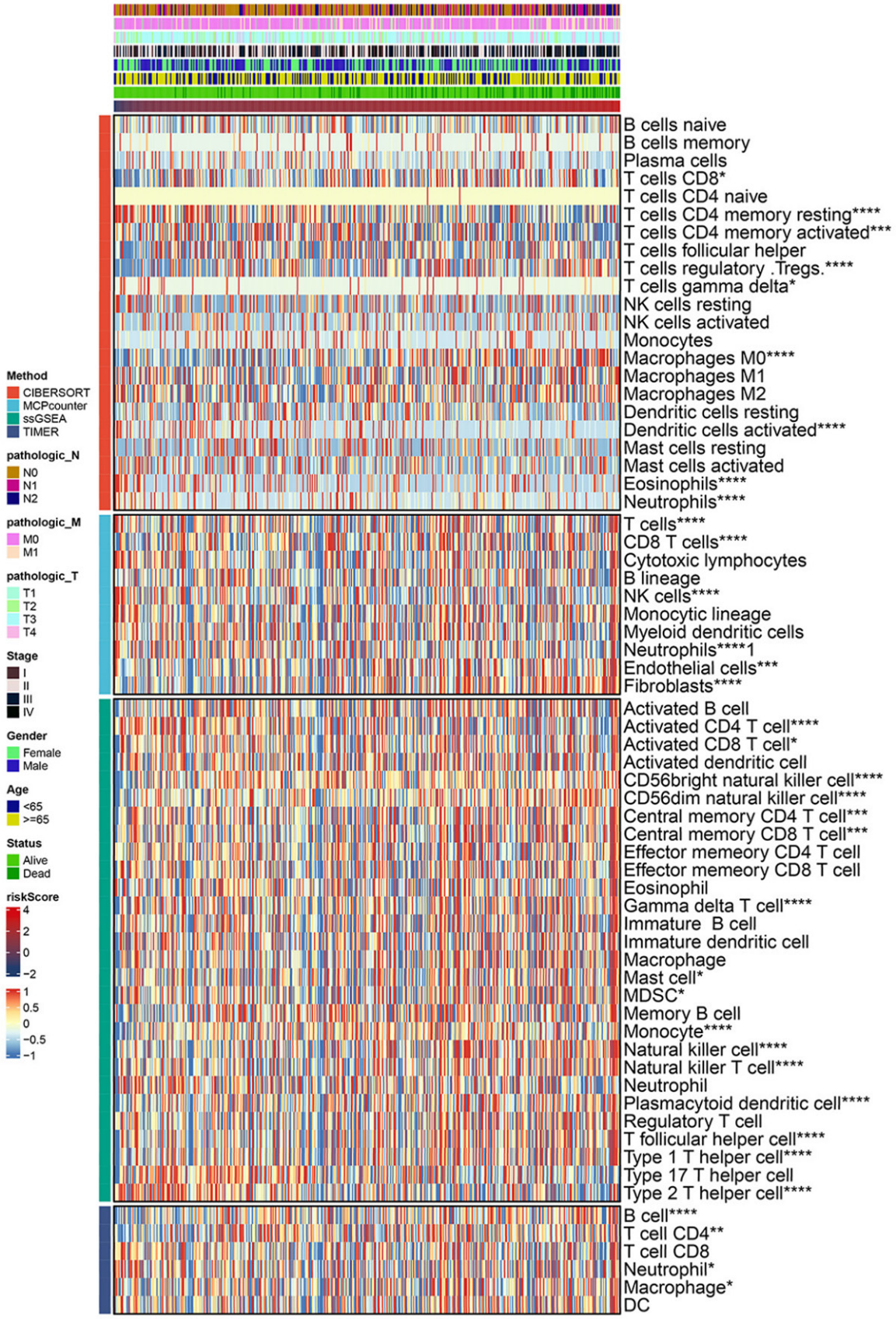


Figure 4. Correlation between immune responses and risk score. Heatmap showing the comparison of immune responses based on CIBERSORT, MCPcounter, ssGSEA, and TIMER algorithms among high and low risk groups, represented by expression-based interrogation of the tumor immune infiltrates. *P < 0.05; **P < 0.01; ***P < 0.001.

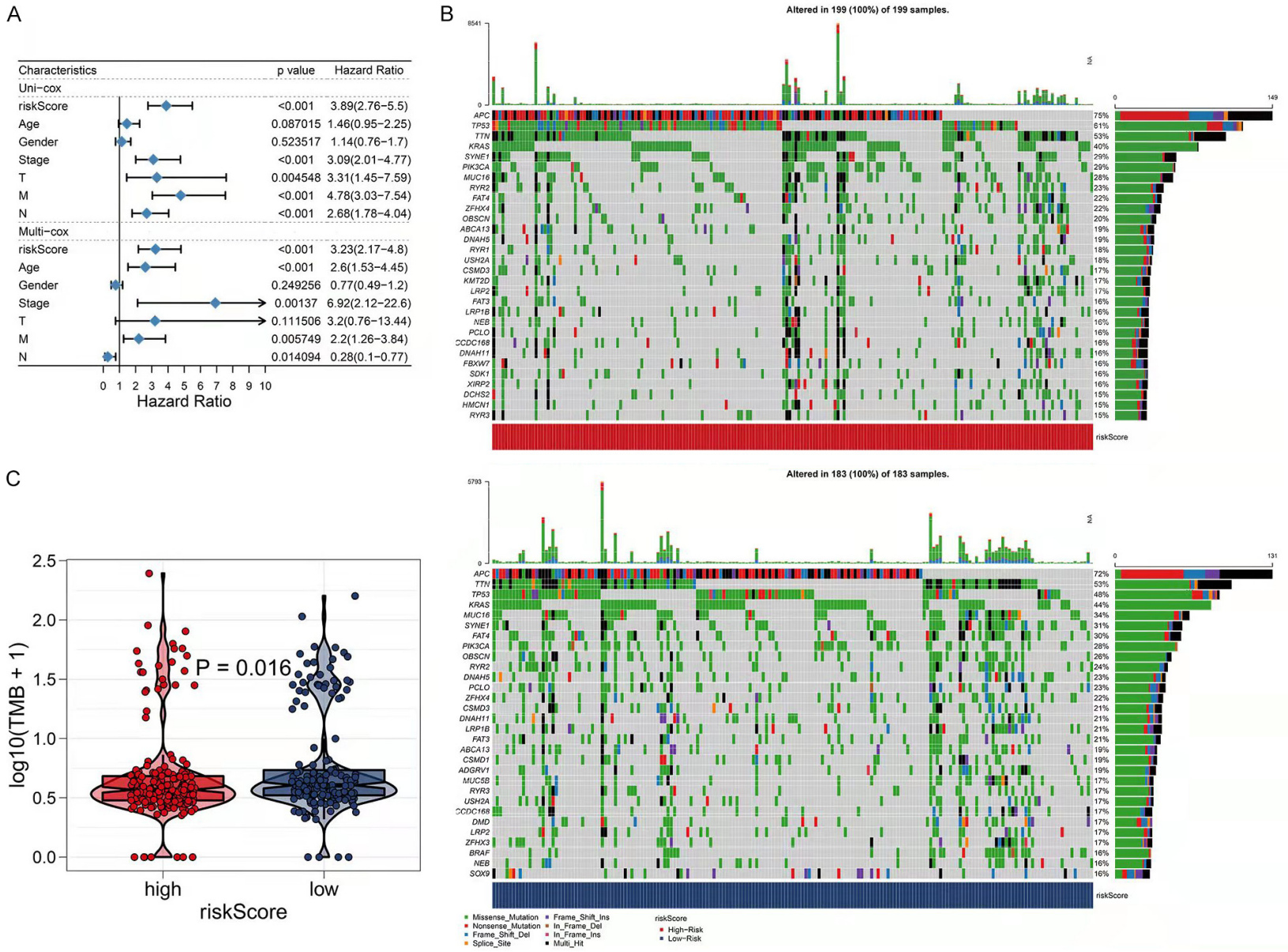
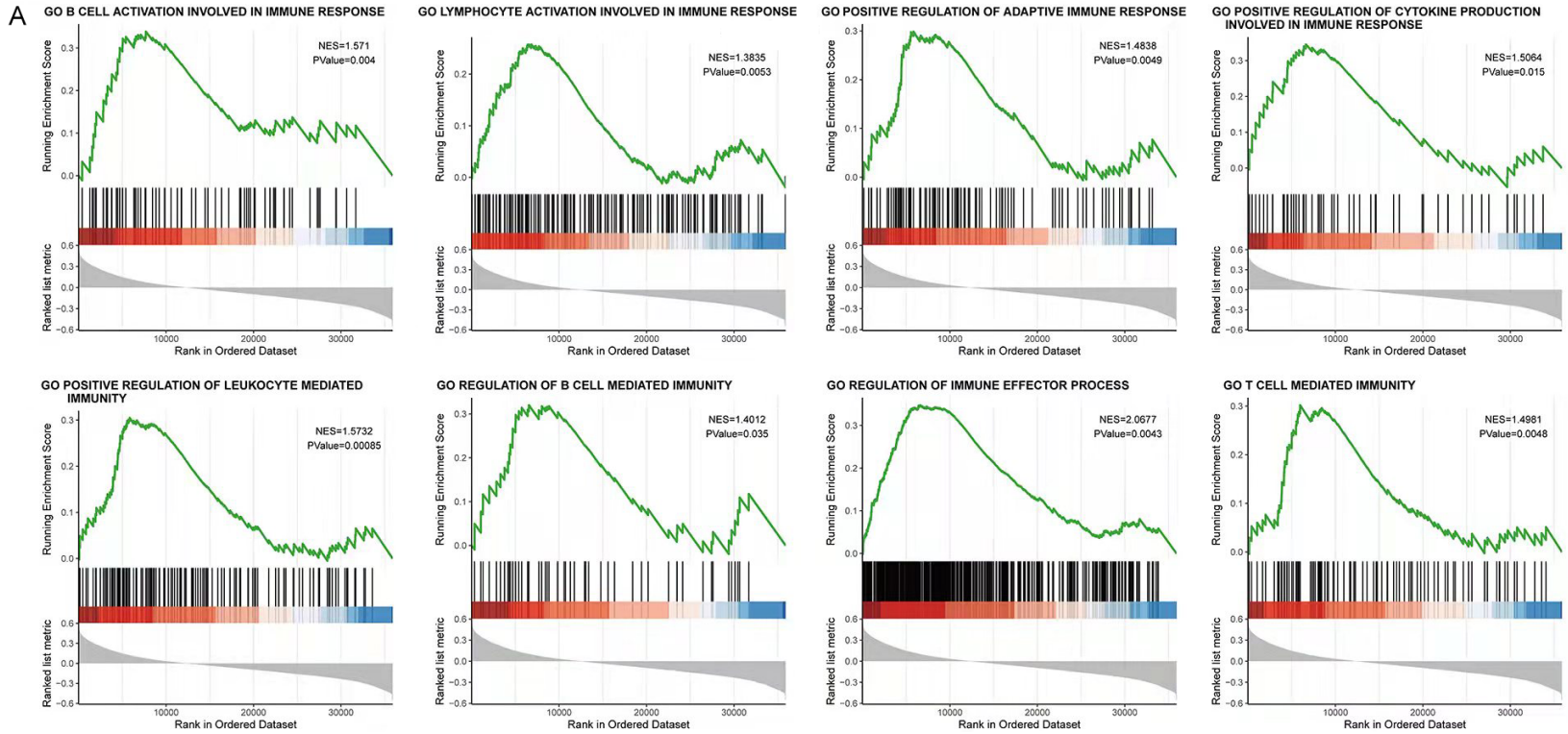

A
Characteristics
p value
Hazard Ratio
Uni−cox
riskScore
<0.001
3.89(2.76−5.5)
Age
0.087015
1.46(0.95−2.25)
Gender
0.523517
1.14(0.76−1.7)
Stage
<0.001
3.09(2.01−4.77)
T
0.004548
3.31(1.45−7.59)
M
<0.001
4.78(3.03−7.54)
N
<0.001
2.68(1.78−4.04)
Multi−cox
riskScore
<0.001
3.23(2.17−4.8)
Age
<0.001
2.6(1.53−4.45)
Gender
0.249256
0.77(0.49−1.2)
Stage
0.00137
6.92(2.12−22.6)
T
0.111506
3.2(0.76−13.44)
M
0.005749
2.2(1.26−3.84)
N
0.014094
0.28(0.1−0.77)
0 1 2 3 4 5 6 7 8 9 10
Hazard Ratio
B
Altered in 199 (100%) of 199 samples.
APC
TP53
TTN
KRAS
SYNE1
PIK3CA
MUC16
RYR2
FAT4
ZFHX4
OBSCN
ABCA13
DNAH5
RYR1
USH2A
CSMD3
KMT2D
LRP2
FAT3
LRP1B
NEB
PCLO
CCDC168
DNAH11
FBXW7
SDK1
XIRP2
DCHS2
HMCN1
RYR3
riskScore
C
P = 0.016
log10(TMB + 1)
high
low
riskScore
Altered in 183 (100%) of 183 samples.
APC
TTN
TP53
KRAS
MUC16
SYNE1
FAT4
PIK3CA
OBSCN
RYR2
DNAH5
PCLO
ZFHX4
CSMD3
DNAH11
LRP1B
FAT3
ABCA13
CSMD1
ADGRV1
MUC5B
RYR3
USH2A
CCDC168
DMD
LRP2
ZFHX3
BRAF
NEB
SOX9
riskScore
Missense_Mutation
Nonsense_Mutation
Frame_Shift_Del
Splice_Site
Frame_Shift_Ins
In_Frame_Del
In_Frame_Ins
Multi_Hit
riskScore
High−Risk
Low−Risk

Figure 5. Correlation between the risk score, clinical chrematistics, and mutation burden. A. Uni-and multivariate Cox regression analysis for the prediction of overall survival in COAD patients. B. Oncoplots showing the top 30 genes with most frequent mutations in the high (top) and low (bottom) risk groups in the TCGA cohort. Each column represents a patient and its frequency of mutations (the top barplot). Each row represents a gene and the corresponding frequency of mutations (right barplot). C. Violin plot showing the comparison of tumor mutation burdens between high and low risk groups.



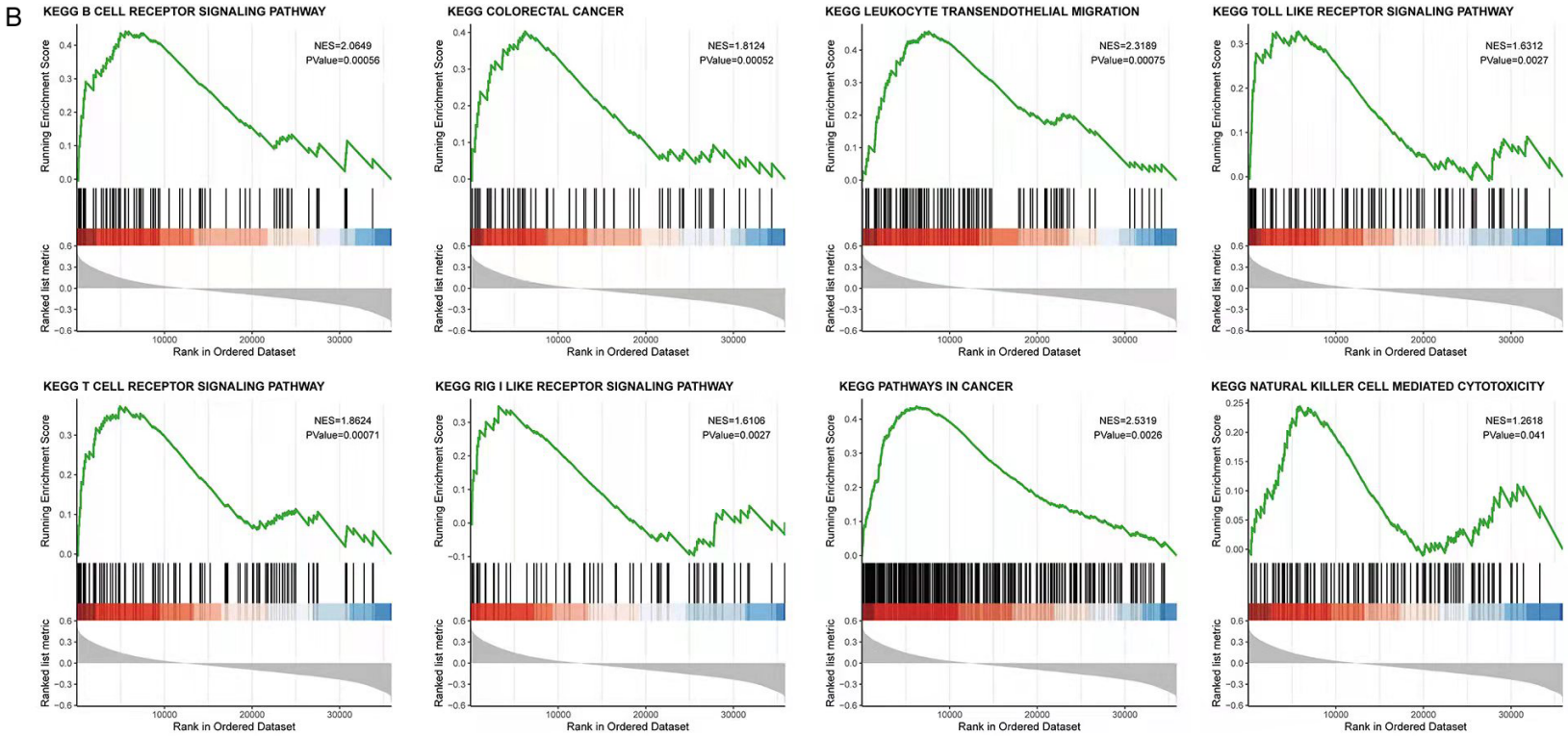


Figure 6. Enriched pathways in high risk COAD patients defined from the risk score model. A. Top enriched GO pathways. B. Top enriched KEGG pathways revealed by GO analysis.

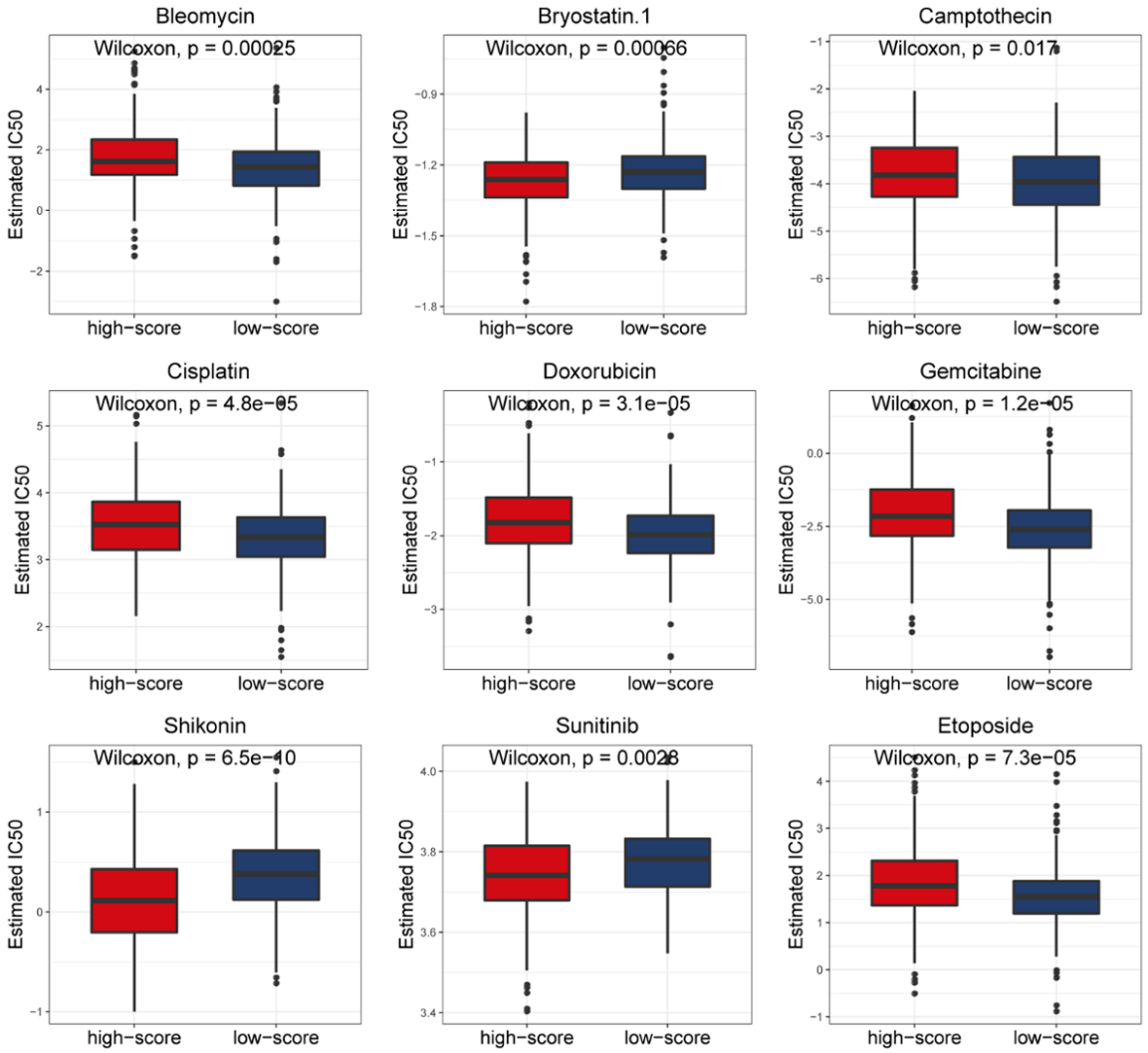


Figure 7. Significant differences between estimated IC50 for chemo drugs between high risk and low risk patients. Box plots showing the estimated IC50 from 9 first line chemo therapeutics drugs in COAD patients stratified by low and high risk scores.

widely treated by conservative surgery, COAD has a high incidence of metastasis. Here, a LASSO Cox regression model was used to select the relevant DEGs identified from TCGA patients for COAD prognosis. The resultant risk score model contained 13 pyroptosis-related DEGs. We utilized the previously reported 37 pyroptosis genes and identified DEGs based on pyroptosis status. Subsequently, the LASSO method was applied for feature selection and to establish the prognostic model. The effectiveness and robustness of the model was validated by an additional dataset GSE39582, where the prognostic power of the model was up to 0.804 for 5-year survival. On the other hand, to determine the correlation between the risk model and the clinical outcome, we conducted multivariate Cox analysis and found that there were significant differences in N stage, T stage, and grade between the risk groups. Thus, our risk score was significantly associated with prognosis.

Tumor mutational burden has been implicated as a predictive marker for cancer prognosis and drug resistance. By conducting the somatic mutation analysis, we showed that an elevated risk score was linked with a greater mutation rate of oncogenes, including TP53 and KRAS, consistent with the previous reports on both colon cancer and other cancer types [26, 27].

Additionally, to gain functional insights of the pyroptosis-related gene signatures, GSEA analysis was performed and signaling pathways involved in immune responses and colorectal cancer were enriched by multiple categories, such as NK cell mediated cytotoxicity, and T cell mediated immunity. Actually, a significant alteration in immune response has been discovered in patients with colorectal cancer [28]. CRC patients maintained decreased level of total Th1 CD4+ cells, accompanied by reduced cytokine production. In contrast, the Th2 lymphocyte population was not affected, and in some patients the cells numbers even slighted elevated [29, 30]. It is known that this kind of immune suppression starts gradually at both a molecular and cellular level [31, 32]. As the disease progresses, the tumor immune tolerance eventually spreads to the whole organism. Likewise, studies have shown that natural killer (NK) cells provide immune surveillance of cancer [33]. One study has shown that EGFR inhibitors can enhance the susceptibility to NK cell-mediated lysis of colon cancer cells through suppressing of the protein kinase C pathway [34].

Taken together, we have established a risk model for COAD prognosis, based on 13 pyroptosis-related genes. Our model has proven accurate and robust in predicting the overall survival of COAD patients, independently from the TNM staging system. Moreover, functional analysis identified potential molecular targets and plausible drug responses in COAD treatment. Hence, our study provides insights into prognosis and disease management of COAD patients.

Acknowledgements

This work was supported by the Xiamen Key Laboratory of Biliary Tract Disease.

Disclosure of conflict of interest

None.

Address correspondence to: Xiaosan Zhu, Department of Gastroenterology, Chenggong Hospital, Xiamen University, Xiamen 361003, Fujian, China. E-mail: zhuxs223@xmu.edu.cn; Tao Xie, Department of Obstetrics and Gynecology, Chenggong Hospital, Xiamen University, Xiamen 361003, Fujian, China. E-mail: dair016@163.com

References

[1] Siegel RL, Miller KD, Fuchs HE and Jemal A. Cancer statistics, 2021. CA Cancer J Clin 2021; 71: 359.

[2] Gu MJ, Huang QC, Bao CZ, Li YJ, Li XQ, Ye D, Ye ZH, Chen K and Wang JB. Attributable causes of colorectal cancer in China. BMC Cancer 2018; 18: 38.

[3] Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Soneson C, Marisa L, Roepman P, Nyamundanda G, Angelino P, Bot BM, Morris JS, Simon IM, Gerster S, Fessler E, De Sousa E Melo F, Missiaglia E, Ramay H, Barras D, Homicsko K, Maru D, Manyam GC, Broom B, Boige V, Perez-Villamil B, Laderas T, Salazar R, Gray JW, Hanahan D, Tabernero J, Bernards R, Friend SH, Laurent-Puig P, Medema JP, Sadanandam A, Wessels L, Delorenzi M, Kopetz S, Vermeulen L and Tejpar S. The consensus molecular subtypes of colorectal cancer. Nat Med 2015; 21: 1350-1356.

[4] McGuire S. World Cancer Report 2014. Geneva, Switzerland: World Health Organization, international agency for research on cancer, WHO Press, 2015. Adv Nutr 2016; 7: 418-419.

[5] Chen ZX, Chen MY, Xue ZY and Zhu XS. Comprehensive analysis of gene expression profiles identifies a P4HA1-related gene panel as a prognostic model in colorectal cancer patients. Cancer Biother Radiopharm 2021; 36: 693-704.

[6] Walter FM, Emery JD, Mendonca S, Hall N, Morris HC, Mills K, Dobson C, Bankhead C, Johnson M, Abel GA, Rutter MD, Hamilton W and Rubin GP. Symptoms and patient factors associated with longer time to diagnosis for colorectal cancer: results from a prospective cohort study. Br J Cancer 2016; 115: 533-541.

[7] Lin JS, Perdue LA, Henrikson NB, Bean SI and Blasi PR. Screening for colorectal cancer: updated evidence report and systematic review for the US preventive services task force. JAMA 2021; 326: 279.

[8] Biller LH and Schrag D. Diagnosis and treatment of metastatic colorectal cancer: a review. JAMA 2021; 325: 669-685.

[9] Fouad YA and Aanei C. Revisiting the hallmarks of cancer. Am J Cancer Res 2017; 7: 1016-1036.

[10] Smets LA. Programmed cell death (apoptosis) and response to anti-cancer drugs. Anticancer Drugs 1994; 5: 3-9.

[11] Blagosklonny MV. Cell death beyond apoptosis. Leukemia 2000; 14: 1502-1508.

[12] Kanduc D, Mittelman A, Serpico R, Sinigaglia E, Sinha AA, Natale C, Santacroce R, Di Corcia MG, Lucchese A, Dini L, Pani P, Santacroce S, Simone S, Bucci R and Farber E. Cell death:

apoptosis versus necrosis (review). Int J Oncol 2002; 21: 165-170.
[13] Bergsbaken T, Fink SL and Cookson BT. Pyroptosis: host cell death and inflammation. Nat Rev Microbiol 2009; 7: 99-109.
[14] Fang Y, Tian S, Pan Y, Li W, Wang Q, Tang Y, Yu T, Wu X, Shi Y, Ma P and Shu Y. Pyroptosis: a new frontier in cancer. Biomed Pharmacother 2020; 121: 109595.
[15] Wang Y, Yin B, Li D, Wang G, Han X and Sun X. GSDME mediates caspase-3-dependent pyroptosis in gastric cancer. Biochem Biophys Res Commun 2018; 495: 1418-1425.
[16] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES and Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005; 102: 15545-15550.
[17] Chen B, Khodadoust MS, Liu CL, Newman AM and Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. Methods Mol Biol 2018; 1711: 243-259.
[18] Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautès-Fridman C, Fridman WH and de Reyniès A. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol 2016; 17: 249.
[19] Foroutan M, Bhuva DD, Lyu R, Horan K, Cursons J and Davis MJ. Single sample scoring of molecular phenotypes. BMC Bioinformatics 2018; 19: 404.
[20] Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, Li B and Liu XS. TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. Cancer Res 2017; 77: e108-e110.
[21] Minton K. Pyroptosis heats tumour immunity. Nat Rev Immunol 2020; 20: 274-275.
[22] Fink SL and Cookson BT. Apoptosis, pyroptosis, and necrosis: mechanistic description of dead and dying eukaryotic cells. Infect Immun 2005; 73: 1907-1916.
[23] Frantz S, Ducharme A, Sawyer D, Rohde LE, Kobzik L, Fukazawa R, Tracey D, Allen H, Lee RT and Kelly RA. Targeted deletion of caspase-1 reduces early mortality and left ventricular dilatation following myocardial infarction. J Mol Cell Cardiol 2003; 35: 685-694.
[24] Gao J, Qiu X, Xi G, Liu H, Zhang F, Lv T and Song Y. Downregulation of GSDMD attenuates tumor proliferation via the intrinsic mitochondrial apoptotic pathway and inhibition of EGFR/Akt signaling and predicts a good prognosis in non-small cell lung cancer. Oncol Rep 2018; 40: 1971-1984.
[25] Wang YY, Liu XL and Zhao R. Induction of pyroptosis and its implications in cancer management. Front Oncol 2019; 9: 971.
[26] Dolatkhah R, Somi MH, Shabanloei R, Farassati F, Fakhari A and Dastgiri S. Main risk factors association with proto-oncogene mutations in colorectal cancer. Asian Pac J Cancer Prev 2018; 19: 2183-2190.
[27] Ades S, Kumar S, Alam M, Goodwin A, Weckstein D, Dugan M, Ashikaga T, Evans M, Verschraegen C and Holmes CE. Tumor oncogene (KRAS) status and risk of venous thrombosis in patients with metastatic colorectal cancer. J Thromb Haemost 2015; 13: 998-1003.
[28] Evans C, Dalgleish AG and Kumar D. Review article: immune suppression and colorectal cancer. Aliment Pharmacol Ther 2006; 24: 1163-1177.
[29] Nakayama H, Kitayama J, Muto T and Nagawa H. Characterization of intracellular cytokine profile of CD4(+) T cells in peripheral blood and tumor-draining lymph nodes of patients with gastrointestinal cancer. Jpn J Clin Oncol 2000; 30: 301-305.
[30] Heriot AG, Marriott JB, Cookson S, Kumar D and Dalgleish AG. Reduction in cytokine production in colorectal cancer patients: association with stage and reversal by resection. Br J Cancer 2000; 82: 1009-1012.
[31] Whiteside TL. Immune suppression in cancer: effects on immune cells, mechanisms and future therapeutic intervention. Semin Cancer Biol 2006; 16: 3-15.
[32] Krysko DV, Garg AD, Kaczmarek A, Krysko O, Agostinis P and Vandenabeele P. Immunogenic cell death and DAMPs in cancer therapy. Nat Rev Cancer 2012; 12: 860-875.
[33] Schmitt C, Ghazi B and Bensussan A. NK cells and surveillance in humans. Reprod Biomed Online 2008; 16: 192-201.
[34] Lee YK, Lee SY, Park JR, Kim RJ, Kim SR, Roh KJ and Nam JS. Dysadherin expression promotes the motility and survival of human breast cancer cells by AKT activation. Cancer Sci 2012; 103: 1280-1289.